ORIGINAL RESEARCH

A survey of ORF8 sequence and immunoinformatics features during alpha, delta, and wild type peaks of the SARS-CoV-2 pandemic in Iran

Alireza Tabibzadeh[1], Mohammad Hadi Karbalaie Niya[1,2], Hossein Keyvani[1], Sajad Karampoor[2], Parastoo Yousefi[1], Mohammad Hossein Razizadeh[1], Leila Mousavizadeh[1], Maryam Esghaei[1]*

1 Department of Virology, School of Medicine, Iran University of Medical Sciences, Tehran, Iran.
2 Gastrointestinal and Liver Diseases Research Center, Iran University of Medical Sciences, Tehran, Iran.

***Corresponding Authors: Maryam Esghaei**; E-mail: Maryam.esghaei@gmail.com

Abstract

Background

The Coronavirus disease 2019 (COVID-19) pandemic influences all around the world. The SARS-CoV-2 ORF8 accessory gene represents multiple functions in virus-host interaction. The current study aimed to compare the ORF8 substitutions and epitope features of these substitutions in the various SARS-CoV-2 outbreaks including delta, alpha, and wild type variants in Iran from 2020 to 2022. In addition, we evaluate B cell, HLA I and II epitopes, by in-silico approach to ORF8 binding site prediction.

Methods

The samples were collected from patients diagnosed with SARS-CoV-2 infection via a real-time PCR assay. Then, a conventional PCR was carried out for ORF8 mutations analysis and further Sanger sequencing. Possible important alterations in epitope features of the ORF8 were evaluated by epitope mapping. B cell, HLA class I and II epitopes, evaluated by online databases ABCpred, NetMHCpan-4.1, and NetMHCIIpan-3.2, respectively.

Results

The current study results could not represent novel variations in seven full-length ORF8 sequences or major ORF8 deletions in 80 evaluated samples. In addition, we could not find any ORF8 Δ382 during each outbreak of variants. Epitope mapping represents differences between the Alpha and other variants, especially in B cell potential epitopes and HLA I.

Conclusion

The immunoinformatic evaluation of ORF8 suggested epitopes represent major differences for the Alpha variant in comparison with other variants. In addition, having mild pathogenesis of the Omicron variant does not seem to be associated with ORF8 alteration by phylogenetic evaluation. Future in-vitro studies for a clear conclusion about the epitope features of ORF8 are required.

Keywords: SARS-CoV-2, COVID-19, ORF8, alpha variant, Epitope Mapping

Introduction

The genome of SARS-CoV-2 is about 30 kb in length and encodes four structural, sixteen non-structural, and six accessory genes including ORF3a, ORF6, ORF7a, ORF7b, ORF8, and ORF10[1-3]. SARS-CoV-2 ORF8 is different from other members of the Coronaviridae family in which its properties are related to the pathogenicity of the virus[4,5]. This protein is composed of 121 amino acids (aa) which have a signal peptide in its N-terminal hydrophobic domain (1–15 aa) and an ORF8 chain (16–121 aa)[6,7]. A biologically functional section known as function motif (VLVVL) of SARS-CoV-1 ORF8b induces cell stress pathways and activation of macrophages but it is absent in ORF8 of SARS-CoV-2[8]. It was found that in the later stages of the SARS-CoV-1 pandemic, a 29 nucleotide deletion in the ORF8 protein caused it to split into ORF8a (39 aa) and ORF8b (84aa) sections, which seems to be important for virus pathogenesis attenuation[9]. While such deletion has not been observed in SARS-CoV-2, a 382-nucleotide deletion variant (Δ382) was identified that resulted in the deletion of the whole ORF8 protein. Patients infected with this variant showed less severe symptoms such as low cytokine activity[10]. In addition, since the ORF8 plays a role in the transmission and replication of the virus, this variant showed less efficient replication in human cells compared with wild types[11].

Along with the spike gene, the ORF8 is among the most divergent genes in SARS-CoV-2 compared with bat and pangolin coronaviruses (12). Since the ORF8 contributes to intriguing functions of the SARS-CoV-2, better comprehension of its function and changes can help us better understand the infectivity of the virus in the future. The current study aimed to investigate the ORF8 variations or possible deletions in the various SARS-CoV-2 outbreaks including delta, alpha, and wild-type variants in Iran. Also, we run an in-silico analysis for B cell, HLA class I and II epitopes, for the binding site of ORF8 in different variants (Wild type, Alpha, Delta, and Omicron).

Materials and methods

Sample collection

The study was confirmed by the Iran University of Medical Sciences ethical committee (ethical code: IR.IUMS.FMD.

https://dx.doi.org/10.4314/mmj.v35i2.5

REC.1400.187). All patients with the confirmed SARS-CoV-2 infection by real-time PCR and Cq values less than 25 were included. All the nasopharyngeal swabs (NP) were collected in a viral transport medium (VTM) and transferred to the virology laboratory by considering the optimum cold chain. In this study, we evaluated 80 COVID-19 patients for ORF8 mutations during delta, alpha, and wild-type of the SARS-CoV-2 in Iran from 2020 to 2022. The ORF8 major deletions were screened in 80 patients.

RNA extraction and cDNA synthesis

The RNA was extracted by using the FavorPrep™ Viral Nucleic Acid Extraction Kit (Favorgen, Taiwan) based on the manufacturer's protocols. The extracted RNA quality was evaluated by the Nanodrop spectrophotometry NanoDrop™ One/OneC (Termofisher, USA). The extracted RNA was used for the cDNA synthesis by the PCR-specific primers and cDNA synthesis kit (Yekta Tajhiz azma, Iran) based on the manufacturer's protocol.

ORF8 mutation evaluation

The semi-nested PCR was used for the mutation evaluation. The primers get from a previous study by Young et al.[10]. The mentioned primers amplify an 880 base pair section of the partial ORF7a and a complete ORF7b and ORF8 in wild-type viruses. In addition, primers can detect a 380 base pair deletion of the ORF8 based on the specific band size in electrophoresis. The PCR was performed by 2X super MasterMix (Yekta Tajhiz azma, Iran) and the thermal program and mixture based on the Young et al.[25].

The PCR products used for Sanger sequencing by ABI 3730 XL sequencer for bidirectional sequencing. The Sanger sequencing reads were evaluated by the NCBI BLAST algorithm. The multiple sequence alignment was performed in the CLC workbench version 20 and the MEGA X was used for the phylogenetic evaluation. The phylogenetic evaluation was performed by the neighbor-joining. Bootstrapping was used for the statistical evaluation of the phylogenetic tree and the results with less than 70 replicates were removed from the tree. All of the mutations and sequences for the variants are obtained from the Next Strain based on the Global Initiative on Sharing All Influenza Data (GISAID) data.

HLA epitope mapping

The HLA I epitope mapping was performed by the online database NetMHCpan-4.1 and based on the previous study by Porto et al.[13,14]. The most prevalent HLA I alleles include HLA-A*02, HLA-B*35, HLA-C*12 (studied from the Iranian population) HLA-A*02:01, HLA-B*51:01, HLA-C*06:02 (studied from Saudi Arabia population) were used for HLA I epitope mapping[15,16]. Default parameters and threshold with 9 amino acid lengths were used.

HLA II epitope mapping was performed by the online database NetMHCIIpan-3.2[17]. Default parameters and threshold with 15 amino acid lengths were used. HLA-DRB1*11, DRB1*07:01, HLA-DQA1*01, -DQB1*03, DQB1*02:01, HLA-DPB1*04:01, DPB1*02:01, DPA1*01:03 DPB1*04:01, DPB1*02:01 were used for the HLA II mapping of the ORF8 amino acid sequence[15,16].

B cell epitopes

B cell epitopes of ORF8 were evaluated by the artificial neural network-based B-cell epitope prediction server named ABCpred. Value 0.5 is considered as threshold and window length of 16 amino acids. Based on the manual of the method these parameters can predict the epitopes with 65% accuracy and equal sensitivity and specificity[13,18].

Results

ORF8 mutations

From all of the included patients, only 7 high-quality full reads for the ORF8 were obtained. Based on the band size on the electrophoresis of the 80 evaluated patients there was no major mutation (382 deletions of the ORF8) during all of the evaluated outbreaks of SARS-CoV-2 variants. All of the 80 samples screened for major deletions represent a normal 880 base pair band (880 base pair represents no major deletion in ORF8) in gel electrophoresis (Supplement figure 1). The mutation analysis revealed all of the expected variants associated (Alpha and delta) specific mutations in ORF8 Figure 1.

The mutation analysis revealed no novel mutation in ORF8 evaluated sequences. The phylogenetic evaluation of the ORF8 segment illustrated the similarity of the evaluated sequence with previously reported samples Figure 2. The phylogenetic study represents a significant similarity between wild type and the Omicron variant ORF8 amino acid sequences.

The epitopes of the HLA I

The evaluated sequence of the ORF8 demonstrates a wide range of potential sequences for the HLA-I binding. There were three alterations in the alpha variant in important

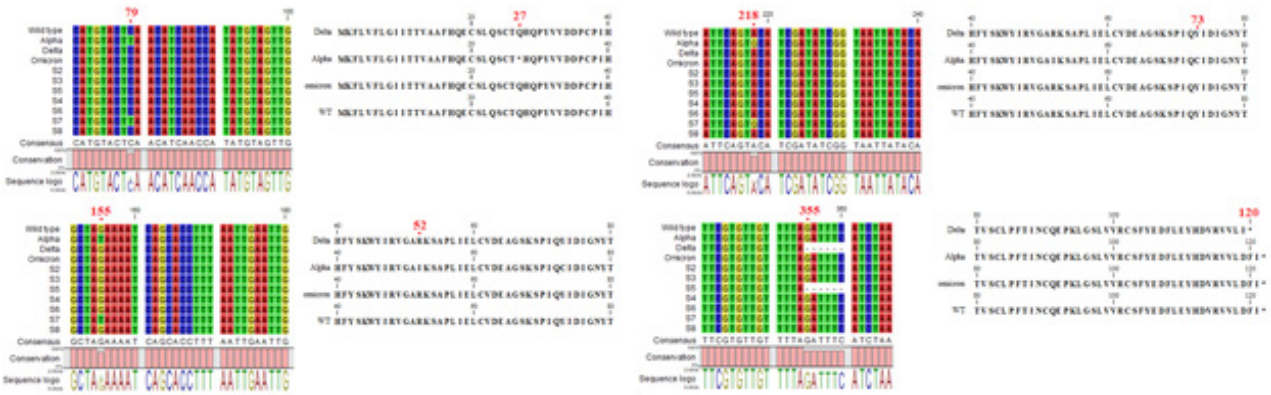

Figure 1. The ORF8 mutations in all evaluated samples (n=7) and important variants of concern. The nucleotides are in color and amino acids are in white. The substations in nucleotides and reflected changes in amino acid sequences are marked with red numbers.

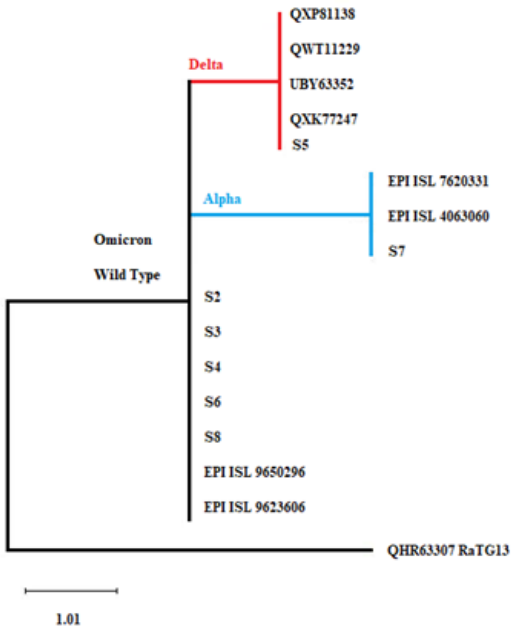

Figure 2. The ORF8 phylogenetic evaluation was performed by the neighbor joining. Bootstrapping was used for the statistical evaluation of the phylogenetic tree and the results with less than 70 replicates were removed from the tree. RaTG13 is used as an outgroup, S2-7 represents the sequenced samples and other evaluated nucleotide sequences are obtained from GSAID by the EPI numbers. Some delta variant amino acid sequences are provided from the NCBI and refereed by the accession numbers. The Alpha variants are routed as Blue and the delta variants are routed as red. The name of the variants in each branch is write

epitope regions for HLA I binding (Figure 3). These mutations in epitope regions for HLA I are the alpha variants specific. The first potential epitope with alteration in Alpha variant is "RKSAPLIEL" at positions 52-60 with strong binding to HLA-C06:02. The amino acids at positions 67-76 represent "SPIQYIDI" as a potential epitope for HLA-B51:01 and HLA-B35:01 with an alteration in the Alpha variant. Furthermore, "YIDIGNYTV" at positions 72-81 for HLA-A02:01 and HLA-A02:02 represents a potential epitope with an alteration in the Alpha variant. Raw data of epitopes in different HLA I and binding potential are provided in Supplement 1.

The epitopes of the HLA II

The evaluated sequences of the ORF8 and HLA II are shown in (Figure 4). A wide range of weak binding positions is introduced in the ORF8 sequence. There is one epitope region with strong binding to HLA II with specific mutations in the Alpha variant (Figure 4). This epitope is "HFYSKWYIRVGARKSA" at positions 40-60 with strong binding to HLA DRB1:11:01. Furthermore, raw data of epitopes and binding potential are provided in Supplement 2.

The epitopes of B cell

Evaluation of the B cell epitopes exhibits a great diversity between all suspected VOCs and Alpha variants. A deletion mutation in the amino acid residue 27 changes a predicted B cell epitopes in the entire sequence of the Alpha variant. In addition, substitutions in amino acid residues number 52 and 73 lead to alteration in predicted B cell epitopes. Of ten suggested B cell epitopes in all VOCs and the Alpha variant only three at positions 83 to 120 were common (Figure 5). More information about the Score and the amino acid sequence for epitopes are provided in Supplement 3.

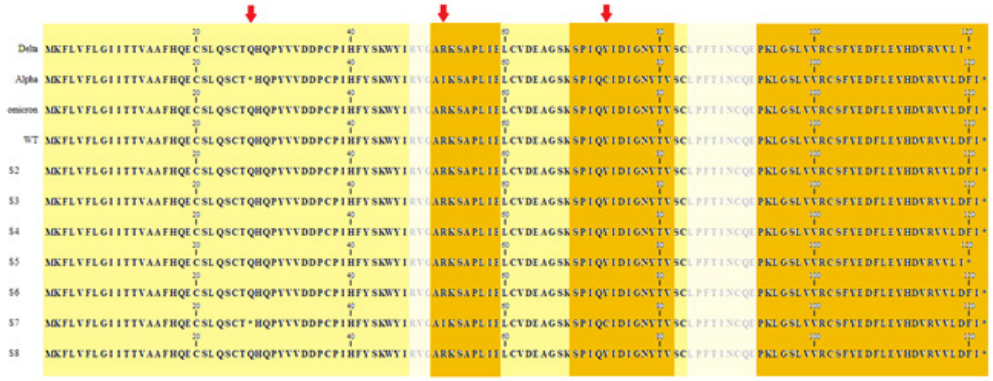

Figure 3, full-lengths 121 amino acid sequence of the wild type, Omicron, Delta (119 aa), and alpha (119 aa) variants alignment. The highlighted sequence with the darker yellow color represents high potential HLA I epitopes with weak binding of the peptide to HLA I. The orange highlighted sequences are displays potential strong binding sequences with HLA I. The red arrows are amino acid substitution positions in strong binding sequences with HLA I which all are associated with the alpha variant. The amino acid composition for strong binding sequences with HLA I is not variable between Omicron, wild type, and delta variants.

Discussion

Current studies indicate multiple roles for ORF8 in SARS-CoV-2 pathogenesis[7,8]. ORF8 also plays a role in modulating the host immune system by down-regulating MHC class I molecules and thereby, evading the infected cell from the cytotoxic T lymphocytes. It also inhibits the HLA type 1 interferon signaling pathway, a crucial part of the host antiviral response[19,20]. This protein also enhances the survivability of the infected cell by regulating unfolded protein response (UPR) induced due to the endoplasmic reticulum stress by triggering the ATF-6 activation[21,22]. Previous data

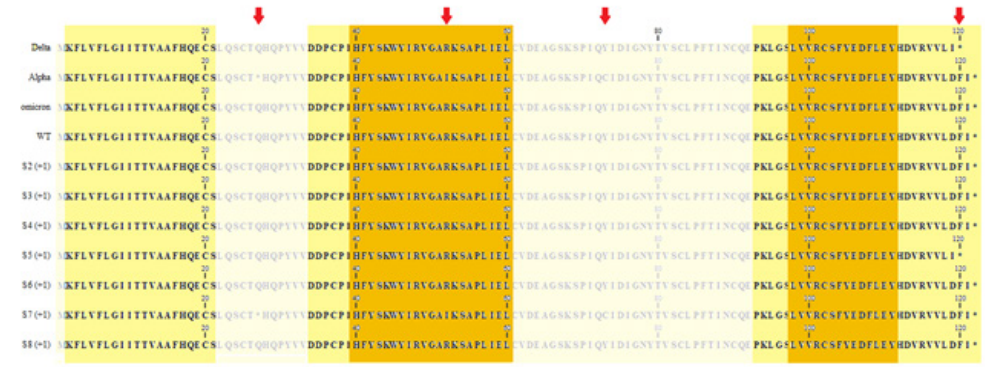

Figure 4, full length 121 amino acid sequence of the wild type, Omicron, Delta (119 aa), and alpha (119a) variants alignment. The highlighted sequence with the darker yellow color shows high potential HLA II epitopes with weak binding of the peptide to HLA II. The orange highlighted sequences show potential strong binding sequences with HLA II. The red arrows are amino acid substitution in different variants. The amino acid sections 40 to 60 show strong binding sequences with HLA II and are variable alpha variants.

https://dx.doi.org/10.4314/mmj.v35i2.5

demonstrated the importance of ORF8 in influencing the severity of COVID-19 as well as influencing T cell immune response[5]. Based on total information related to ORF8 and its role in SARS-CoV-2 pathogenesis, the current study aimed to investigate the ORF8 mutations during delta, alpha, and wild-type outbreaks of the SARS-CoV-2 pandemic in Iran. The current study did not report any particular ORF8 major deletions among 80 evaluated Iranian COVID-19 patients. Furthermore, there were no novel nucleotide mutations in 7 full-length ORF8 evaluated sequences. Sequences were related to Alpha, delta, and wild-type variants. Meanwhile, there were no specific mutations in the ORF8 of the Omicron variant (21L clade). However, epitope mapping indicates differences between the Alpha variant and other variants, especially in B cell epitopes and HLA I. B cell potential epitopes in the N terminal site of the Alpha variant are completely different from other variants. In addition, three and one of the potential epitope regions in HLA I and II, respectively, reveal specific mutations in the Alpha variant. These mutations and potential epitope regions in ORF8 of different variants highlight the importance of further in-vitro studies for potential differences in ORF8 of the variants. According to our data, a close relationship between Omicron variant and wild-type virus is represented (Figure 2). In this regard, it could be concluded that the ORF8 sequence (as a hypothesis for ORF8 dividing in ORF8a and 8b in SARS-CoV-1) may not lead to affect SARS-CoV2 pathogenicity by itself and it cannot be the only cause influencing immune interaction or virus attenuation at least in Omicron variant. Although, there are some limitations in our study including the low number of evaluated samples. This preliminary data about ORF8 major deletions or single nucleotide mutations should not be considered for the total population of Iranian COVID-19 patients and further investigation with larger sample sizes is required. Another was in the limited considered HLA types based on the previous studies for Iranian and Arab populations[15,16]. The epitopes evaluation was a prototype data and this data needs to be confirmed by in-vitro or other more complete evaluation with more in-slico databases approaches. Another study has mentioned the conserved structure of some non-structural proteins of the SARS-CoV-2 such as ORF8 and used them as anti-SARS-CoV-2 vaccine targets by massive and great in-silico pipeline[23]. As a glimpse, it seems to be our current study data is against the Safavi et al. study[23]. However, it should be considered that the Safavi et al. was published in 2020, there were no clues about the variants, and it seems they only used available the wild-type stain amino acid sequences. Furthermore, another limitation of the current study was the preliminary in-silico approaches. The current study aimed to screen any differences between variants of SARS-CoV-2 antigen diversity. The results of the current study need to be confirmed by other epitope mapping data bases for MHC and B cell epitopes in further studies.

The immunogenic features of the ORF8 introduced this protein as a considerable candidate for immunoassay platforms[24]. The ORF8 is a highly immunogenic and rapid serological response to ORF8 by IgM, IgG, and IgA which is demonstrated[25]. In this regard, amino acid residues 41 to 71 are introduced as the external domain of the ORF8[25]. Based on the mutation evaluation, there are two important mutations in the external domain of the ORF8 in the Alpha variant, which could be important in different antigenic features (R52I, Y73C) of the protein. Our current in-silico evaluation represents a strong HLA binding for ORF8 in this particular domain either. Conducted in-vitro study for the B cell epitopes in ORF8 highlighted the importance of immune dominant epitopes "MKFLVFLGIITTVAAFHQE", " Y V V D D P C P I H F Y S K W Y I R V G " and "GSKSPIQYIDIGNYTVSCLP"[26]. The last epitope represents an amino acid substitution in the alpha variant. The comparison between in-vitro and in-silico results for B cell epitopes seems to be acceptable based on the suggested epitopes. These alterations in ORF8 epitope content by immunoinformatic analysis on the B 1.1.7 linage (alpha variant) are comprehensively reported by Hussain et al.[27]. These findings are supported by current study results and suggest a possible important epitope alteration in the Alpha variant but not the other variants (including the Omicron variant) in comparison with the wild type.

Conclusion

The immunoinformatic evaluation of ORF8 suggested epitopes represent major differences for the Alpha variant in comparison with other variants, especially in HLA I and B cell epitopes. These epitope alterations could be one of the possible reasons for changes in the Alpha variant pathogenesis. The study could not find any ORF8 major deletions in 80 evaluated samples from Iran from 2020 to 2022. In addition, having mild pathogenesis of the Omicron variant does not seem to be associated with ORF8 alteration by phylogenetic evaluation of the amino acid sequences. However, more studies by a greater population need to be done to comprehend the results. Future in-vitro studies for a clear conclusion about the epitope features of ORF8 are required. In addition, future studies about possible reasons for the milder pathogenesis of the Omicron variant are suggested.

Conflict of interest

No conflict of interests.

Authors contribution

AT, MHK and PY experiments, HK, SK, MHR and LM manuscript preparation, ME director and manuscript procreation.

Acknowledgements

Acknowledgements from all the collaborative authors.

References

1.Shang J, Han N, Chen Z, Peng Y, Li L, Zhou H, et al. Compositional diversity and evolutionary pattern of coronavirus accessory proteins. Briefings in bioinformatics. 2021;22(2):1267-78.

2.Li J-Y, Liao C-H, Wang Q, Tan Y-J, Luo R, Qiu Y, et al. The ORF6, ORF8 and nucleocapsid proteins of SARS-CoV-2 inhibit type I interferon signaling pathway. Virus research. 2020;286:198074.

3.Kim D, Lee J-Y, Yang J-S, Kim JW, Kim VN, Chang H. The architecture of SARS-CoV-2 transcriptome. Cell. 2020;181(4):914-21. e10.

4.Mohammad S, Bouchama A, Mohammad Alharbi B, Rashid M, Saleem Khatlani T, Gaber NS, et al. SARS-CoV-2 ORF8 and SARS-CoV ORF8ab: genomic divergence and functional convergence. Pathogens. 2020;9(9):677.

5.Zhang Y, Chen Y, Li Y, Huang F, Luo B, Yuan Y, et al. The ORF8 protein of SARS-CoV-2 mediates immune evasion through down-regulating MHC-I. Proceedings of the National Academy of Sciences. 2021;118(23).

6.Kamau A, Kulmanov M, Arold ST, Pain A, Gojobori T, Duarte CM. Functional pangenome analysis provides insights into the origin, function and pathways to therapy of SARS-CoV-2 coronavirus. bioRxiv. 2020.

7.Alam I, Kamau AA, Kulmanov M, Jaremko Ł, Arold ST, Pain A, et al. Functional pangenome analysis shows key features of e protein are preserved in sars and sars-cov-2. Frontiers in cellular and infection microbiology. 2020;10:405.

8.Wu J, Yuan X, Wang B, Gu R, Li W, Xiang X, et al. Severe acute respiratory syndrome coronavirus 2: from gene structure to pathogenic mechanisms and potential therapy. Frontiers in microbiology. 2020;11:1576.

9.Schaecher SR, Pekosz A. SARS coronavirus accessory gene expression and function. Molecular biology of the SARS-coronavirus. 2010:153-66.

10.Young BE, Fong S-W, Chan Y-H, Mak T-M, Ang LW, Anderson DE, et al. Effects of a major deletion in the SARS-CoV-2 genome on the severity of infection and the inflammatory response: an observational cohort study. The Lancet. 2020;396(10251):603-11.

11.Parlikar A, Kalia K, Sinha S, Patnaik S, Sharma N, Vemuri SG, et al. Understanding genomic diversity, pan-genome, and evolution of SARS-CoV-2. PeerJ. 2020;8:e9576.

12.Boni MF, Lemey P, Jiang X, Lam TT-Y, Perry BW, Castoe TA, et al. Evolutionary origins of the SARS-CoV-2 sarbecovirus lineage responsible for the COVID-19 pandemic. Nature microbiology. 2020;5(11):1408-17.

13.Porto PS, Anjos D, Dábilla N, da Fonseca SG, Souza M. Immunoinformatic construction of an adenovirus-based modular vaccine platform and its application in the design of a SARS-CoV-2 vaccine. Infection, Genetics and Evolution. 2020;85:104489.

14.Jurtz V, Paul S, Andreatta M, Marcatili P, Peters B, Nielsen M. NetMHCpan-4.0: improved peptide–MHC class I interaction predictions integrating eluted ligand and peptide binding affinity data. The Journal of Immunology. 2017;199(9):3360-8.

15.Jawdat D, Uyar FA, Alaskar A, Müller CR, Hajeer A. HLA-A, -B, -C, -DRB1, -DQB1, and -DPB1 Allele and Haplotype Frequencies of 28,927 Saudi Stem Cell Donors Typed by Next-Generation Sequencing. Frontiers in Immunology. 2020;11.

16.Abedini F, Rahmanian N, Heidari Z, Feizi A, Sherkat R, Rezaei M. Diversity of HLA class I and class II alleles in Iran populations: Systematic review and Meta-Analaysis. Transplant Immunology. 2021;69:101472.

17.Jensen KK, Andreatta M, Marcatili P, Buus S, Greenbaum JA, Yan Z, et al. Improved methods for predicting peptide binding affinity to MHC class II molecules. Immunology. 2018;154(3):394-406.

18.Saha S, Raghava GPS. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins: Structure, Function, and Bioinformatics. 2006;65(1):40-8.

19.Park MD. Immune evasion via SARS-CoV-2 ORF8 protein? Nature Reviews Immunology. 2020;20(7):408-.

20.Kumar S, Nyodu R, Maurya VK, Saxena SK. Host immune response and immunobiology of human SARS-CoV-2 infection. Coronavirus disease 2019 (COVID-19): Springer; 2020. p. 43-53.

21.Sung S-C, Chao C-Y, Jeng K-S, Yang J-Y, Lai MM. The 8ab protein of SARS-CoV is a luminal ER membrane-associated protein and induces the activation of ATF6. Virology. 2009;387(2):402-13.

22.Hassan SS, Aljabali AA, Panda PK, Ghosh S, Attrish D, Choudhury PP, et al. A unique view of SARS-CoV-2 through the lens of ORF8 protein. Computers in biology and medicine. 2021;133:104380.

23.Safavi A, Kefayat A, Mahdevar E, Abiri A, Ghahremani F. Exploring the out of sight antigens of SARS-CoV-2 to design a candidate multi-epitope vaccine by utilizing immunoinformatics approaches. Vaccine. 2020;38(48):7612-28.

24.Hachim A, Kavian N, Cohen CA, Chin AW, Chu DK, Mok CK, et al. ORF8 and ORF3b antibodies are accurate serological markers of early and late SARS-CoV-2 infection. Nature immunology. 2020;21(10):1293-301.

25.Wang X, Lam J-Y, Wong W-M, Yuen C-K, Cai J-P, Au SW-N, et al. Accurate diagnosis of COVID-19 by a novel immunogenic secreted SARS-CoV-2 orf8 protein. MBio. 2020;11(5):e02431-20.

26.Wang X, Lam J-Y, Chen L, Au SW-N, To KK, Yuen K-Y, et al. Mining of linear B cell epitopes of SARS-CoV-2 ORF8 protein from COVID-19 patients. Emerging microbes & infections. 2021;10(1):1016-23.

27.Hussain M, Shabbir S, Amanullah A, Raza F, Imdad MJ, Zahid S. Immunoinformatic analysis of structural and epitope variations in the spike and Orf8 proteins of SARS-CoV-2/B. 1.1. 7. Journal of medical virology. 2021;93(7):4461-8.