



SUSCEPTIBILITY OF SPIKE GLYCOPROTEIN AND RNA-DEPENDENT RNA POLYMERASE OF SARS-CoV-2 TO MUTATION: IN SILICO STRUCTURAL DYNAMICS STUDY

Toluwase Hezekiah Fatoki¹, Jude Akinyelu¹, Oluwafijimi Yomi Adetuyi¹, Temitope Olawale Jeje¹, Uchechukwu Nebo², Jesupemi Mercy Fatoki³, Tolulope Mercy Kupolati⁴

Abstract

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a microorganism that causes coronavirus disease 2019 (COVID-19). Mutations affect evolutionary conservation of microorganisms. The fast pace evolutionary changes are currently affecting pathogenicity of SARS-CoV-2. In this study, the structural fluctuations of the amino acid residues in the spike glycoprotein and RNA-dependent RNA polymerase (nsp12) of SARS-CoV-2 were investigated by in silico approach using structural flexibility dynamics to decipher susceptibility to mutation. The result of this study implicated key amino acid residues (with rmsf) which could be very susceptible to mutation, which include residues 50 (3.79 Å), 119 (4.56 Å), 120 (3.53 Å), 220 (3.84 Å), 265 (4.31 Å) of RNA-dependent RNA polymerase (nsp12), as well as residues 477 (4.21 Å), 478 (4.82 Å), 479 (5.40 Å), 481 (5.94 Å), 560 (4.63 Å), 704 (4.02 Å), 848 (4.58 Å), 1144 (4.56 Å) and 1147 (4.61 Å) of spike glycoprotein. The SARS-CoV-2 mutations destabilized the overall protein structure in multiples of amino acid residues which could interfere with active site leading to insensitivity or resistance to the inhibitors. Mutation T478K of Spike glycoprotein showed the highest deviation in the structure. Overall, spike glycoprotein has the highest number of mutations, and these variants could increase the risk to human health if not mitigated in the population.

Running title: Susceptibility of SARS-CoV-2 proteins to mutations

Keywords: SARS-CoV-2, omicron variant, spike glycoprotein, RNA-dependent RNA polymerase, mutation, structural analysis

¹Department of Biochemistry, Federal University Oye-Ekiti, Oye-Ekiti, Nigeria

²Department of Microbiology, Federal University Oye-Ekiti, Oye-Ekiti, Nigeria

³Department of Microbiology, Federal University of Technology, Akure, Nigeria

⁴Department of Zoology, Ekiti State University, Ado-Ekiti, Nigeria

*Correspondence: toluwase.fatoki@fuoye.edu.ng

Full list of author information is available at the end of article

Introduction

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a pathogen that causes coronavirus disease 2019 (COVID-19) which was first recorded at Wuhan city, China in December 2019, and it has caused 262,866,050 confirmed cases which has resulted to over 5,224,519 deaths worldwide as of 3 December 2021 [1]. Mutations affect evolutionary conservation of microorganisms. The emergence of new variants is an expected occurrence as it has been the case of SARS-CoV-2. A variation of SARS-CoV-2 due to evolution and adaptation processes has been detected globally. While most emerging mutations will not have significant impact on the spread of the virus, some mutations or combinations of mutations may provide the virus with a selective advantage, such as increased transmissibility or the ability to evade the host immune response. In that instances, the variants could increase the risk to human health and are considered to be variants of concern (VOC) [2].

Since the outbreak of COVID-19 pandemic, new SARS-CoV-2 variants have been reported from the United Kingdom, France, South Africa, Denmark, and Nigeria [2,3]. The fast pace evolutionary changes are currently affecting pathogenicity of SARS-CoV-2 as the third-wave slams the global communities, and this called for urgent *in silico*, *in vitro*, and *in vivo* studies [4,5]. According to Huang et al. [6], SARS-CoV-2 protein mutations have two imperative features: (i) a higher ability to infect and enhanced pathogenicity than the bat-like SARS-CoV, and (ii) a lower pathogenicity than SARS-CoV. It has been reported that the existing approved vaccines which include two mRNA-based anti-SARS-CoV-2 vaccines; BNT162b2 (Pfizer-BioNTech) and mRNA-1273 (Moderna), as well as three inactivated virus-based vaccines; BBIBP-CorV vaccine (Sinopharm), BBV152 (Bharat Biotech International Limited), and CoronaVac (Sinovac Biotech), were significantly less protective against Beta variant of coronavirus [7].

The World Health Organization (WHO) has provided new nomenclature for VOC on 31 May 2021 using Greek letters: Alpha for B.1.1.7; Beta for B.1.351; Gamma for P.1; Delta for B.1.617.2, and others. Beta variant (B.1.351 VOC-20DEC-02 (VOC-202012/02), GH/501Y.V2) of coronavirus is characterized by 17 mutations, which are: eight spike protein mutations (D80A, D215G, 241/243 deletion, K417N, E484K, N501Y, D614G, and A701V), four open reading frame 1a (orf1a) mutations, one orf1b mutation, one nucleocapsid (N) protein mutation and three others [8]. RNA-dependent RNA polymerase (RdRp) is one of the non-structural proteins (NSPs) derived from the polyprotein encoded ORF1 of SARS-CoV-2 [9]. Non-structural protein (nsp) of SARS-CoV-2 which include 3-Chymotrypsin-like protease (3CL^{pro}), Papain-like protease (PL^{pro}), helicase (nsp13), RdRp

(nsp12), 2'-O-ribose methyltransferase (nsp16) has been investigated computationally, as potential targets for therapeutic discovery [10]. V843F and A889V are the two mutations reported for PL^{pro}, with V843F mutation being the most prevalent mutation in the clinical samples [11]. It has been reported that PL^{pro} (PDB ID: 5Y3E) is less deformed and could be less susceptible during viral mutations [10]. *In silico* approach has been used to compare annotated functional domains of SARS-CoV and SARS-CoV-2 sequences. An *in silico* molecular dynamics study on the protein structure was used to predict amino acid residues T478K mutation in Spike protein and that might significantly alter the electrostatic surface of the protein and their binding with human receptor angiotensin-converting enzyme 2 (ACE2) as well as prevention and curative measures such as vaccines, antiviral drugs, and monoclonal antibodies [12]. In this study, the structural fluctuations of the amino acid residues in the spike glycoprotein and nsp12 of SARS-CoV-2 were investigated by *in silico* approach to decipher susceptibility to mutation.

Materials and methods

Structure preparation

Briefly, From the crystal structure of Spike glycoprotein and RNA-dependent RNA polymerase (nsp12) of SARS-CoV-2 were obtained from the protein databank (www.rcsb.org/pdb) in pdb format and they were prepared for simulation by removing all water molecules, multichain, and heteroatoms using PyMol v 2.0.7.

Molecular structure dynamics simulation of normal SARS-CoV-2 proteins

The prepared crystal structures of Spike Glycoprotein (PDB ID: 7CAB, chain A) and RNA-dependent RNA polymerase (PDB ID: 7BW4, chain A) of SARS-CoV-2 were subjected to fast simulation of structural flexibility using CAB-flex 2.0 server [13], all parameters were at default settings. The contact map and root-mean square fluctuations (RMSF) of atoms in the server analyzed protein was obtained.

Molecular structure dynamics simulation of mutated SARS-CoV-2 proteins

The prepared crystal structures of Spike Glycoprotein (PDB ID: 7CAB, chain A) of SARS-CoV-2 was edited with the following mutations: **T478K, E484K, N501Y, D614G, A701V** and RNA-dependent RNA polymerase (PDB ID: 7BW4, chain A) of SARS-CoV-2 was edited with the following mutations: **G228C, P323L, V330E, D824Y** [3,8,9,14]. These were subjected to fast simulation of structural flexibility using CAB-flex 2.0 server [13], all parameters were at default settings. The contact map and root-mean square fluctuation (RMSF) of atoms in the server analyzed protein was obtained.

Assessment of mutational effect on the SARS-CoV-2 protein

The structural differences between the normal and mutated proteins were estimated by model-by-model basis on the Dali Server (<http://ekhidna2.biocenter.helsinki.fi/dali/#tabs-3>, [15]). The root-mean square deviations (RMSD) were obtained.

Results

The result of this study showed that RNA-dependent RNA polymerase (nsp12) with amino acid residues (rmsf) 40 (3.00 Å), 50 (3.79 Å), 119 (4.56 Å), 120 (3.53 Å), 220 (3.84 Å), 263 (3.33 Å), 265 (4.31 Å), 270 (3.30 Å), and 499 (3.00 Å) could be very susceptible to mutation as shown in **figure 1**. The result of this study showed that spike glycoprotein with amino acid residues (rmsf) 282 (3.65 Å), 477 (4.21 Å), 478 (4.82 Å), 479 (5.40 Å), 481 (5.94 Å), 560 (4.63 Å), 703 (3.94 Å), 704 (4.02 Å), 848 (4.58 Å), 1041 (3.94 Å), 1144 (4.56 Å) and 1147 (4.61 Å) could be very susceptible to mutation as shown in **figure 2**.

The SAR-CoV-2 mutations destabilized the overall protein structure in multiples of amino acid residues which could interfere with active site leading to insensitivity or resistance to the inhibitors as seen in Figure 1C and 2C. Mutation T478K of Spike glycoprotein showed the highest deviation in the structure while other mutations seem to limitedly affected the structure.

Discussion

The results of comparison of protein sequences SARS-CoV-2 and SARS-CoV by Blastp, has shown that orf8 and orf10 proteins were not present in SARS-CoV, while surface glycoprotein, nsp2, nsp3, orf3a, and orf6 proteins, have percent identity of less than 80% of variation between SARS-CoV-2 and SARS-CoV [16]. According to the alignment analysis, the closest similarity was between the two COVID-19 genomes (99%) compared with the two bat CoV [4]. The orf1a polyprotein of SARS-CoV-2 is similar to that SARS-CoV by 80.58% identity [16]. Spike glycoprotein Receptor-Binding Domain (RBD; aa 319-541), has vital functional and antigenic properties, and any mutation in this region has potential for reduced susceptibility to neutralizing antibodies elicited by vaccination or prior infection [17].

D614G mutation in spike glycoprotein and P323L in RNA-dependent RNA polymerase (nsp12) are the globally dominant variants with a high frequency [14]. Variant of Nextstrain cluster 20A.EU2 has been reported with mutation of S477N in spike protein of the new SARS-CoV-2, which was rapidly increased in France at the start of the second wave, probably due to initiation effects, and prevalent also in Belgium, Czech, Denmark, Hungary, Netherlands, and Switzerland [18]. The P681H mutation is also characteristic of the new SARS-CoV-2 variants from the United Kingdom and Nigeria [3].

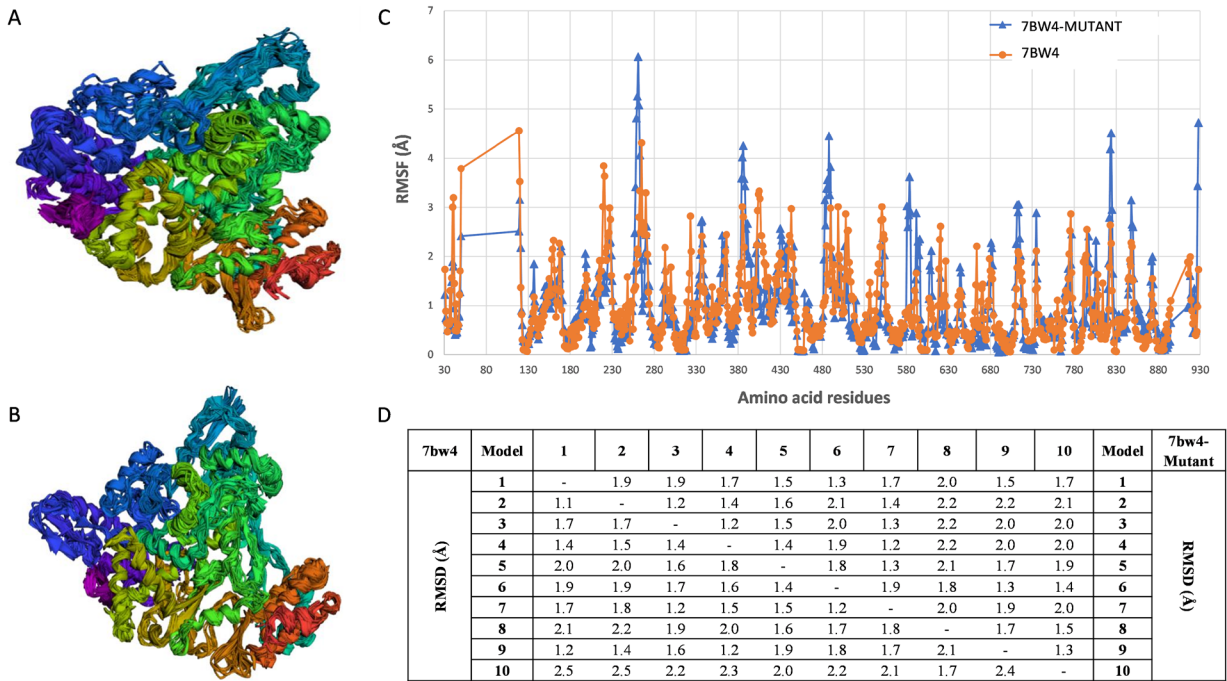


FIGURE 1 Structural dynamics of RNA-dependent RNA polymerase (PDB ID: 7BW4, chain A) and its mutant consisting of G228C, P323L, V330E, and D824Y. (A) Superimposition of 10 model structures of 7BW4 from CAB-flex 2.0 server. (B) Superimposition of 10 model structures of 7BW4 mutant from CAB-flex 2.0 server. (C) RMSF of amino acids residues in 7BW4 and 7BW4 mutant. (D) Comparison of model protein structures from Dali Server

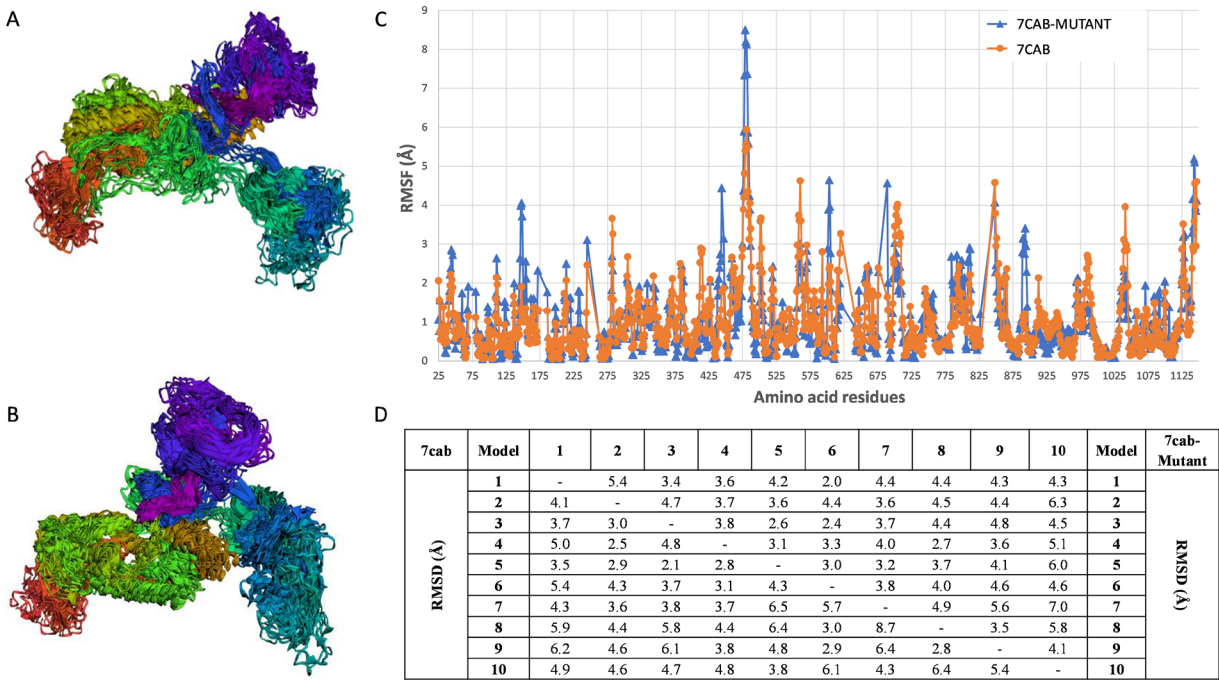


FIGURE 2 Structural dynamics of Spike glycoprotein (PDB ID: 7CAB, chain A). and its mutant consisting of T478K, E484K, N501Y, D614G, and A701V. (A) Superimposition of 10 model structures of 7CAB from CAB-flex 2.0 server. (B) Superimposition of 10 model structures of 7CAB mutant from CAB-flex 2.0 server. (C) RMSF of amino acids residues in 7CAB and 7CAB mutant. (D) Comparison of model protein structures from Dali Server

A recent report on the distribution of the Spike mutation S:T478K and its recent growth in prevalence in the SARS-CoV-2 population of Mexico, United States, and several European countries, showed that S:T478K is frequently co-occurring with three other Spike mutations; D614G, P681H and T732A, with 99.83%, 93.8% and 88.7% co-occurrence respectively as well as two amino acid residues mutation in Nucleocapsid (N:RG203KR), and mutations in nsp12 [9]. Moreover, bioinformatics study has predicted G476S, and T478I mutation in spike glycoprotein and G25Y, G228C, V330E, L810H and D824Y in nsp12 [14]. A study has reported that mutations in 2891, 3036, 14408, 23403 and 28881 positions of SARS-CoV-2 genomic sequence, are predominantly observed in Europe, whereas those located at positions 17746, 17857 and 18060 are exclusively present in North America [19]. These mutations were belonged to the sequence of ORF1ab (1397 nsp2, 2891 nsp3, 14408 nsp12 (RdRp), 17746 and 17857 nsp143, 18060 nsp14), S (23403, spike protein) and ORF9a (28881, nucleocapsid phosphoprotein), respectively [19]. Interestingly, the docking site for inhibitors is not within the catalytic domain of the RdRp [10] but the docking site for inhibitors of the spike glycoprotein is on the Receptor-Binding Domain in the N-terminal region of S1 which is responsible for viral attachment to target cells via the ACE2 receptor [17,20].

The VOC B.1.1.529 (Omicron) belongs to Pango lineage B.1.1.529, Nextstrain clade 21K, is the most divergent variant, and characterized by 30 amino acid

substitutions, 3 deletions and one insertion in the spike protein compared to the original virus (A67V, del69-70, T95I, del142-144, Y145D, del211, L212I, ins214EPE, G339D, S371L, S373P, S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493R, G496S, Q498R, N501Y, Y505H, T547K, D614G, H655Y, N679K, P681H, N764K, D796Y, N856K, Q954H, N969K, L981F) [21]. Most of the substitutions in the sequence coding the spike protein have been found associated with almost complete escape sera from neutralization by convalescent and vaccinee sera [22].

Conclusions

This study has shown key amino acid residues of SARS-CoV-2 spike glycoprotein and nsp12 that could be susceptibility to mutation. Overall, spike glycoprotein has highest number of mutations and these variants could increase the risk to human health if not mitigated in the population. Consequently, it is imperative to investigate and characterize SARS-CoV-2 spike glycoprotein and nsp12 (RdRp) mutation in order to assess possible drug-resistance viral phenotypes. However, based on the existing classification, these variants may not be of interest or high consequence, in that currently no coherent proof that efficacy of prevention and curative measures such as vaccines, antiviral drugs, and monoclonal antibodies, are significantly decreased in comparison to earlier perpetrating variants [21,23]. As of 1 December 2021, a total of 7,885,790,810 vaccine doses have been administered globally [1].

Thus, there should be concerted efforts toward administration and development of holistic vaccines that will directly target the existing and emergent variants of SARS-CoV-2.

Ethical approval

The conducted research is not related to either human or animal use.

Acknowledgement

Not applicable.

Corresponding author

Toluwase Hezekiah Fatoki, Department of Biochemistry, Federal University Oye-Ekiti, PMB 373 Oye-Ekiti, Ekiti State, Nigeria, e-mail: hezekiahfatoki@gmail.com, toluwase.fatoki@fuoye.edu.ng.

Conflict of interest

The authors declare they have no conflict of interest.

References

- World Health Organization. Coronavirus Disease (COVID-19) Dashboard [Internet]. Geneva: WHO; 2020 [updated 2021 Dec 02; cited 2021 Dec 03]. Available from: <https://covid19.who.int>.
- European Centre for Disease Prevention and Control. Risk related to spread of new SARS-CoV-2 variants of concern in the EU/EEA. Stockholm: ECDC; 2020. 26 p.
- Centers for Disease Control and Prevention. Science brief: Emerging SARS-CoV-2 variants [Internet]. USA: CDC; 2020 [updated 2021 Jan 28; cited 2021 Jul 30]. Available from: <https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/scientific-brief-emerging-variants.html>.
- Dawood AA. Mutated COVID-19 may foretell a great risk for mankind in the future. *NMNI* 2020;35:100673; DOI:10.1016/j.nmni.2020.100673.
- Maison DP, Ching LL, Shikuma CM, Nerurkar VR. Genetic characteristics and phylogeny of 969-bp S gene sequence of SARS-CoV-2 from Hawai'i Reveals the worldwide emerging P681H mutation. *Hawaii J Health Soc Welf*. 2021;80(3):52-61.
- Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X, Cheng Z, Yu T, Xia J, Wei Y, Wu W, Xie X, Yin W, Li H, Liu M, Xiao Y, Gao H, Guo L, Xie J, Wang G, Jiang R, Gao Z, Jin Q, Wang J, Cao B. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*. 2020;395(10223):497-506. DOI: 10.1016/S0140-6736(20)30183-5.
- Khateeb J, Li Y, Zhang H. Emerging SARS-CoV-2 variants of concern and potential intervention approaches. *Crit Care*. 2021;25:244; DOI:10.1186/s13054-021-03662-x.
- Latif AA, Mullen JL, Alkuzweny M, Tsueng G, Cano M, Haag E, Zhou J, Zeller M, Hufbauer E, Matteson N, Wu C, Andersen KG, Su AI, Gangavarapu K, Hughes LD. B.1.351 Lineage Report. [Internet]. USA: outbreak.info; 2021 [updated 2021 Nov 30; cited 2021 Dec 03]. Available from: <https://outbreak.info/situation-reports?pango=B.1.351>.
- Di Giacomo S, Mercatelli D, Rakhimov A, Giorgi FM. Preliminary report on severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) Spike mutation T478K. *J. Med Virol*. 2021;93:5638-5643; DOI:10.1002/jmv.27062.
- Fatoki TH, Ibraheem O, Ogunyemi IO, Akinmoladun AC, Ugbo HU, Adeseko CJ, Awofisayo OA, Olusegun SJ, Enibukun JM. Network analysis, sequence and structure dynamics of key proteins of coronavirus and human host, and molecular docking of selected phytochemicals of nine medicinal plants. *J Biomol Struct Dyn*. 2021;39(16):6195-6217; DOI:10.1080/07391102.2020.1794971.
- Hossain MU, Bhattacharjee A, Emon MTH, Chowdhury ZM, Ahammad I, Mosaib MG, Moniruzzaman M, Rahman MH, Islam MN, Ahmed I, Amin MR, Rashed A, Das KC, Keya CA, Salimullah M. Novel mutations in NSP-1 and PLPro of SARS-CoV-2 NIB-1 genome mount for effective therapeutics. *J Genet Eng Biotechnol*. 2021;19(1):52; DOI:10.1186/s43141-021-00152-z.
- Rezaei S, Sefidbakht Y, Uskoković V. Comparative molecular dynamics study of the receptor-binding domains in SARS-CoV-2 and SARS-CoV and the effects of mutations on the binding affinity. *J Biomol Struct Dyn*. 2020;1-20; DOI:10.1080/07391102.2020.1860829.
- Kuriata A, Gierut AM, Oleniecki T, Ciemny MP, Kolinski A, Kurcinski M, Kmiecik S. CABS-flex 2.0: a web server for fast simulations of flexibility of protein structures. *Nucleic Acids Res*. 2018;46(W1):W338-W343; DOI:10.1093/nar/gky356.
- Ikram A, Naz A, Awan FM, Rauff B, Obaid A, Hakim MS, Malik A, The impact of mutations on the structural and functional properties of SARS-CoV-2 proteins: A comprehensive bioinformatics analysis. *bioRxiv* [Preprint] 2021; DOI:10.1101/2021.03.01.433340.
- Holm L. Using Dali for Protein Structure Comparison. *Methods Mol Biol*. 2020;2112:29-42; DOI:10.1007/978-1-0716-0270-6_3.
- Xu J, Zhao S, Teng T, Abdalla AE, Zhu W, Xie L, Wang Y and Guo X, Systematic comparison of two animal-to-human transmitted human coronaviruses: SARS-CoV-2 and SARS-CoV. *Viruses*. 2020;12:244; DOI:10.3390/v12020244.
- Tchesnokova V, Kulasekara H, Larson L, Bowers V, Rechkina E, Kisiela D, Sledneva Y, Choudhury D, Maslova I, Deng K, Kutumbaka K, Geng H, Fowler C, Greene D, Ralston J, Samadpour M, Sokurenko E. Acquisition of the L452R mutation in the ACE2-binding interface of spike protein triggers recent massive expansion of SARS-CoV-2 variants. *J Clin Microbiol*. 2021;59(11):e0092121; DOI:10.1128/JCM.00921-21.
- Hodcroft EB, Zuber M, Nadeau S, Vaughan TG, Crawford KHD, Althaus CL, Reichmuth ML, Bowen JE, Walls AC, Corti D, Bloom JD, Veersler D, Matteo D, Hernando A, Comas I, González-Candelas F; SeqCOVID-SPAIN consortium, Stadler T, Neher RA. Spread of a SARS-CoV-2 variant through Europe in the summer of 2020. *Nature*. 2021;595(7869):707-712; DOI:10.1038/s41586-021-03677-y.
- Pachetti M, Marini B, Benedetti F, Giudici F, Mauro E, Storici P, Masciovecchio C, Angeletti S, Ciccozzi M, Gallo RC, Zella D, Ippodrino R. Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant. *J Transl Med*. 2020;18(1):179; DOI:10.1186/s12967-020-02344-6.
- Adeoye AO, Oso BJ, Olaoye IF, Tijjani H, Adebayo AI. Repurposing of chloroquine and some clinically approved antiviral drugs as effective therapeutics to prevent cellular entry and replication of coronavirus. *J Biomol Struct Dyn*. 2021;39(10):3469-3479; DOI:10.1080/07391102.2020.1765876.
- Centers for Disease and Control Prevention. SARS-CoV-2 Variant Classifications and Definitions [Internet]. USA: CDC; 2020 [updated 2021 Dec 01; cited 2021 Dec 03]. Available from: <https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-info.html>.
- Schmidt F, Weisblum Y, Rutkowska M, Poston D, Da Silva J, Zhang F, Bednarski E, Cho A, Schaefer-Babajew DJ, Gaebler C, Caskey M, Nussenzweig MC, Hatzioannou T, Bieniasz PD. High genetic barrier to SARS-CoV-2 polyclonal neutralizing antibody escape. *Nature*. 2021;1-9; DOI:10.1038/s41586-021-04005-0.
- Janik E, Niemcewicz M, Podogrocki M, Majsterek I, Bijak M. The emerging concern and interest SARS-CoV-2 variants. *Pathogens*. 2021;10:633; DOI:10.3390/pathogens10060633.