VirusDis. (July–September 2022) 33(3):244–250
https://doi.org/10.1007/s13337-022-00781-z

ORIGINAL ARTICLE

Prediction of infectivity of SARS-CoV-2 virus based on Spike-hACE-2 interaction

Dwaipayan Chaudhuri[1] · Joyeeta Datta[1] · Satyabrata Majumder[1] · Kalyan Giri[1]

Received: 13 December 2021 / Accepted: 21 June 2022 / Published online: 5 August 2022

Abstract The COVID-19 pandemic caused by SARS-CoV-2 results almost 3 M death worldwide and till continuing in spite of having several vaccine against the virus. One of the main reasons is the mutations occur in the virus to cope with the environment. Detail study of genomics and proteomics level of each components may help to combat the situation. Spike (S) protein that covers the surface of the virus helps in entry by encountering the host receptor Human Angiotensin-Converting Enzyme-2 (hACE-2) with other different roles. In this study, we accomplish our work with the mutations in receptor binding domain (RBD) of Spike (S) protein considering different aspects like the hACE-2 variants in human populations to get an idea about the varying infectivity of different strains for different population. Several other parameters affecting the viral infectivity and in different diseased condition were also studied which may guide to a better insight in developing future therapeutics.

Keywords Spike · hACE-2 · Population variants · Cancer variants · Network analysis · Binding affinity

Introduction

Coronavirus disease 2019 i.e., COVID-19 is caused by a new coronavirus which has spread worldwide and eventually turned into a global pandemic [1]. The COVID-19 was discovered to be caused by a member of beta-coronavirus family, later named Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) [1]. It is the seventh known coronavirus to infect humans; four of these coronaviruses namely 229E, NL63, OC43, and HKU1 only cause slight symptoms of the common cold, while the other three, SARS-CoV, MERS-CoV, and SARS-CoV-2, are able to cause severe symptoms and even death, with fatality rates of 10%, 37%, and 5%, respectively also showing that in comparison to its two predecessors this virus has a lower mortality rate.

The virus has a single stranded RNA genome (29,811 nts) encoding 29 proteins [2]. Glycosylated Spike (S) proteins cover the surface of SARS-CoV-2 virus and bind to the host cell receptor Angiotensin-Converting Enzyme 2 (hACE-2), assisting in viral cell entry [3]. Following cell entry, the viral RNA gets released and translated into polyproteins and this replication-transcription process of the viral RNA genome occurs via protein cleavage and assembly of the replicase–transcriptase complex. The Viral RNA replication process ends up with the synthesis of the structural proteins followed by their assembling and packaging in the host cell, after which viral particles get released [4]. The SARS-CoV-2 Spike protein is highly conserved among all human coronaviruses and is involved in receptor recognition, viral attachment, and entry into host cells which makes the protein one of the most important factors in virus infectivity and host pathogenesis coupled with host selectivity as well [5].

The Spike protein (180–200 kDa) consists of an extracellular N-terminus, a transmembrane (TM) domain to anchor with the viral membrane, and a short intracellular C-terminal segment [6]. The Spike protein is 1273 aa long and consists of a signal peptide comprising of amino acids 1–13 located at the N-terminus, the S1 subunit (14–685 residues) follows the signal peptide, and S2 subunit (686–1273 residues) comprises the C-terminal domain. The S1 and S2 domains are responsible for receptor binding and membrane

✉ Kalyan Giri
kalyan.dbs@presiuniv.ac.in

[1] Department of Life Sciences, Presidency University, 86/1 College Street, Kolkata 700073, India

fusion respectively. The S1 subunit, itself has an N-terminal domain (14–305 residues) and a receptor binding domain (RBD) which extends from 319–541 residues which is followed by the fusion peptide in the S2 segment which extends from 788–806 residues thus mediating membrane fusion, a heptapeptide repeat sequence 1 extending from 912–984 residues and a HR2 from 1163–1213 residues, transmembrane domain from 1213–1237 residues and finally a cytoplasm domain extending from 1237–1273 residues [7]. The protein exists in a metastable, pre-fusion conformation which rearranges in post-fusion conformation on spot when the virus interacts with the host cell, in turn allows the virus to fuse with the host cell membrane. The polysaccharide molecules involved in protein glycosylation, function to camouflage the proteins, thus helping to evade from the surveillance of the host immune system during entry [8].

In this study, we have performed systematic mutations in the receptor binding domain (RBD) of Spike glycoprotein to identify some specific mutations which lead to increase in binding affinity that plays crucial role in SARS-CoV-2 infectivity. The effect of mutations on the structural stability, vibrational entropy and network connectivity, used to determine the importance of the various residues were also studied and the existing mutations were mapped. The hACE-2 receptor variant positions were also mapped based on their linkage with disease phenotypes as well as based on population distribution of the variants. This study aimed to link the effect of the mutations on the virus infectivity as well as try to give an insight on the effect of these mutant and wildtype virus strains to varied diseased conditions caused due to receptor variants and also effect of the various strains at the population specific level. This study may serve as the gateway for therapeutic research with an eye on the virus protein mutations as also the population specificity.

Materials and methods

Mutational effect on binding affinity

The Spike glycoprotein of the SARS-CoV-2 binds the hACE-2 receptor on the cell surface in 3 stages. First the trimer of the Spike protein interacts with a single receptor which leads to change in conformation thus leading to interaction with the second receptor molecule and finally the third receptor molecule. All the 3 states were taken for this study and their PDB IDs are 7A94, 7A97 and 7A98. Also the structure (PDB ID 6LZG) of only Receptor binding domain to the receptor was taken into consideration (Fig. 1). The receptor binding domain which is the major region for this study extends from 333 to 527. Mutations those are effective on binding affinity was calculated using the BeatMusic server [9] and the results which increase the binding affinity

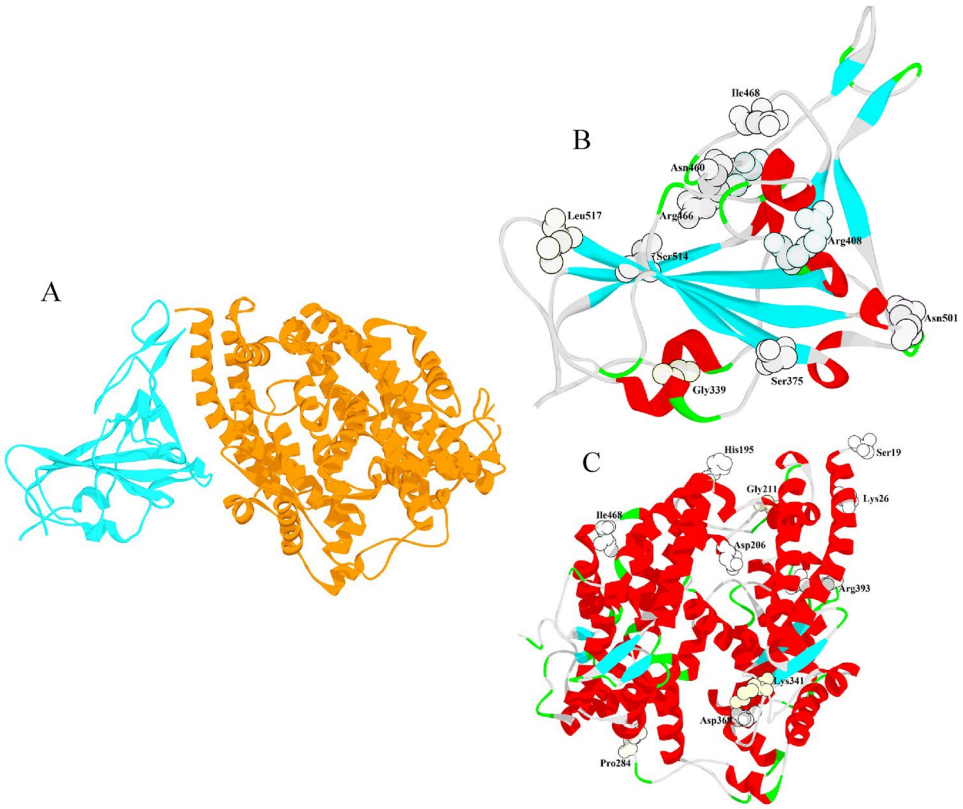

Fig. 1 (**A**) Complex structure of Spike-hACE-2 (6lzg.pdb), (**B**) mutations of Spike proteins and (**C**) mutations of hACE-2. Mutant residues are represented in Van der Waals representations using probe size 1.04 Å

for all the aforementioned 4 structures were taken into consideration for further analysis.

Effect on structural stability

The 165 mutations selected in the previous step were analyzed to study the effect of the mutations on the structural stability of the RBD domain with receptor using the DynaMut server [10]. This data lead to determination of whether the mutation stabilizes the structure or causes a destabilization which would in turn contribute to evolution.

Effect on vibrational entropy

DynaMut server [10] was further used to calculate the change in vibrational entropy that indicates the effect of the mutation on the flexibility of the structure. The increase of flexibility point at change in exposure to solvent, degrees of freedom and hydrophobicity, it can also indicate a loss of rigidity in the structure and push the structure towards lower thermal stability [11, 12].

Mutant structure preparation

The mutant structures were prepared by incorporating the mutation using the DUET web server. The steric clashes were minimized by refining the mutant complexes using GalaxyRefineComplex [13, 14] web server after which the binding affinity values were calculated using the Prodigy web server [15, 16] to ascertain the mutation effects on the binding affinity which serve as the stepping stone to determine the virus infectivity.

Network properties change owing to mutation in Spike protein

Several properties were analyzed for those mutant protein structures to get information about the mutational effects on the protein structures using NAPS web server [17]. Various properties that have been taken into consideration in this study are as follows: (i) degree of a node to know the connectivity with the neighboring node, (ii) closeness centrality measures the extent of flow of information between one node and its neighbor, (iii) Betweenness indicates the shortest path between two nodes, (iv) clustering co-efficient determines the tendency of a node to cluster together, (v) eigenvector centrality measures the influence of one node upon the rest and (vi) eccentricity denotes the easiness of a node to be functionally active in protein network.

Database analysis

The nCoV database [18] was used to obtain the Spike protein mutation data to ascertain the real existence of Spike mutants and have also led to an increase in the binding affinity in the complex. The HuVarBase [19] was used to obtain the various variants available in the 19–614 region that is found in the 6LZG.pdb complex. This was done to get the comparative infectivity data in case of the hACE-2 variants when the patient encounters the ancestral or the mutant Spike protein virus strains. Population variation of hACE-2 was done using the sequences recorded in GNOM AD database [20]. Population specific variations were considered for further analysis.

Results

Effects of Spike mutation on binding affinity of host-virus protein complex

A total of 3705 potential systematic mutations in the region were performed. Based on the change in binding affinity, a final selection of 165 was done which constituted of only the mutations that led to increase in the binding affinity in the complex. In all these 165 cases, the binding affinity between the Spike protein and receptor increased for all the 4 conformations considered in this study. This increase of affinity can be directly linked to infectivity of the virus as the RBD mutant strain with N501Y mutation which has been seen to show a greater infectivity is included in this list of increased binding affinity mutations. Also another study showed in vitro effect of mutations on infectivity and all the mutations in this ancestral strain which led to decrease in affinity were excluded from the list [21]. Again this points to the fact that the decrease of infectivity could be linked with lowering of the binding affinity in the complex. In this entire study only the ancestral structure was taken and not D614G as this mutation is not only part of RBD but also affects the infectivity by modifying the endocytosis rate of the viral particle in the cell and not its interaction with the host receptor.

Mutational effect on structural stability of host-virus protein complex

Of the 165 mutations selected and considered for further analysis, 68% of the mutations were shown to be having a stabilizing influence on the protein complex structure. This stabilization would help the virus to maintain the mutation and serve to create a new line of evolution for the virus, while the rest destabilizing ones would aid in further evolution of the virus since the destabilizing nature helps the protein cross its evolutionary fitness barrier quite easily [22].

Mutational effect on vibrational entropy of host-virus protein complex

In this change in vibrational entropy analysis we found that, 29.6% of the mutations led to increase in the structural flexibility. The other 70.4% led to decrease of flexibility thus giving a proper shape to the binding site of the receptor molecule. This rigidity could also thus play a role in increase in the binding affinity as this rigidity helps the molecule maintain the active site in proper conformation and thus not much change in conformation is needed for the interaction and thus the conformation search for the proper conformation during the binding process would be much easier than if the binding site is highly flexible in nature.

Refinement of mutant Spike-hACE-2 complex

Only 10 (out of 165 mutations) were selected for further analysis because of real existence in nature. GalaxyRefine-Complex server was used All the 10 mutant virus protein-host (mutant Spike-hACE-2) structures were examined in detail to refine and remove the various clashes in the mutant structures where only the sidechains of the residue was changed to generate mutant structures. To solve this and get an optimized structure, structural refinement was performed and the model with least outliers and least rotational outliers were taken into consideration for further analysis in the study.

Structural stability analysis of refined mutant Spike-hACE-2 complex

The ΔG and dissociation constant (K_D) values of wildtype and all the 10 refined protein–protein complexes (mutant Spike-hACE-2) were calculated using the PRODIGY server tabulated in Table 1. The wildtype host-virus protein complex possess ΔG value of -12.4 kcal/mol and k_D value of 1.90E-09 M. ΔG values of all those 10 refined structures have shown to become more negative, lead to increase in binding affinity and greater stability of each structure. Decrease in dissociation constant (K_D) for majority of the cases (compared to the wildtype Spike-hACE-2 complex) indicates strong binding of mutant Spike-hACE-2 complex.

Network analysis to study effect of mutations on the protein structures and its topological properties

The effect of various mutations on protein structure and it's connectivity with the neighboring residues have been described by the mentioned properties mentioned in methodology. All these properties give a comprehensive information about the effect of the mutations in the protein structure. Each analysis has been examined with respect to particular node in that network. Results of each mutation and their effects are tabulated in Table 2. Mutational effect on the degree of a node do not influence for most of the cases. Rest of the properties have shown to follow a mixed trends with their respective effects.

Effect of Spike mutations upon hACE-2 variants

This study has been done in two consecutive ways with respect to the binding affinity (calculating ΔG) of mutant Spike and hACE-2 variants. Firstly, the effect of Spike mutant upon pathogenic hACE-2 variants were examined and secondly, the mutational effect were checked upon hACE-2 variants of different geographical area. HuVarBase was used to obtain the various pathogenic hACE-2 variants affected in 19–614 region found in the 6LZG.pdb complex (Fig. 1). Total 6 pathogenic variants of hACE-2 are recorded and the effect of the various SARS-CoV-2 mutants upon these pathogenic (cancer related) variants could be anticipated by checking the binding affinity (calculating ΔG) values of the variants-mutant Spike

Table 1 Binding affinity on the basis of calculated ΔG (kcal/mol) and K_D (in M) values of mutant Spike-wildtype hACE-2. All the mutant possess higher infectivity than the wildtype host-virus protein complex (ΔG: − 12.4 in Kcal/mol, K_D: 1.90E–09 M)

Mutants (reported to have real existence)	Number of reported sequence	ΔG in kcal/mol of mutant Spike-hACE-2 complex	Dissociation constant K_D (in M) of mutant Spike-hACE-2 complex
G339S	11	− 13.5	2.90E–10
S375P	1	− 12.7	1.10E–09
R408K	67	− 13.5	2.90E–10
N460T	10	− 13.5	3.10E–10
R466I	6	− 13.7	2.20E–10
I468T	11	− 14.2	1.00E–10
N501Y	29,415	− 13.2	4.70E–10
S514F	21	− 13.1	5.90E–10
L517F	26	− 13.4	3.50E–10
L517I		− 13.3	4.50E–10

Table 2 Spike protein mutations and their effects on protein network properties

Residue	Closeness	Betweenness	Clustering co-eff	Eigenvector centrality	Eccentricity	Remarks
wt339	0.078	0.0004	0.73	0.005	22	Easy transmission and clustering, influence of the node on the entire network and mutant residue become unreachable to other network component
G339S	0.083↑	0.0004	0.8↑	0.011↑	19↓	
wt514	0.087	0.033	0.44	0.22	21	Easy of flow of information through the node to make more influential in the network
S514F	0.09↑	0.033	0.44	0.52↑	21	
wt375	0.092	0.01	0.67	0.1	20	Very small difference in information transmission but high loss in influence of the residue in the network. Occurrence of residues in more shortest paths and increase in accessibility of the residue
S375P	0.09↓	0.02↓	0.67	0.02↓	20	
wt466	0.089	0.004	0.43	0.09	21	Loss of a single connectivity closeness, less influence and functional accessibility of the node in the network
R466I	0.1↑	0.007↑	0.53↑	0.006	18↓	
wt408	0.1	0.071	0.6	0.06	18	Leads tough for the functional accessibility of the residue in the network
R408K	0.11↑	0.11↑	0.6	0.006↓	16↓	
wt501	0.123	0.17	0.42	0.14	16	Gain of one new connectivity, loss of Influence of the residue in the network, more functional accessibility
N501Y	0.134↑	0.36↑	0.4↓	0.05↓	14↓	
wt460	0.083	0.003	0.5	0.03	22	Minimal effect on the spike protein network, but possess 2 times more influence in spite of being tougher to access functionally
N460T	0.084↑	0.003	0.5	0.06↑	21↓	
wt468	0.091	0.022	0.53	0.06	21	Decreasing the tendency of the node to form clusters, the influence of the node (by 10 times) and functional accessibility
I468T	0.098↑	0.054↑	0.4↓	0.006↓	19↓	
wt517	0.08	0.014	0.53	0.06	22	Loss of a single connectivity, decrease the accessibility (largely based on structure and minute loss of based on function). Node become more influential in the protein network
L517I	0.08	0.002↓	0.7↑	0.15↑	22	
L517F	0.08	0.0017↓	0.7↑	0.14↑	21↓	

structures. Analyzing the results of binding affinity on the basis of calculated ΔG values of mutant spike and population variants of hACE-2 related to cancer (Supplementary Table 1), it can be said that the mutant strains of Spike proteins are more infective compared to the wildtype strain for the various variants of hACE-2 linked to cancer. Thus it can be said that certain cancer patients are much more susceptible to certain mutant strains than others. Geographical prevalence of the hACE-2 variants are recorded in GNOM AD database. Analysis of binding affinity on the basis of calculated ΔG (in Kcal/mol) values of mutant spike and population variants of hACE-2 (Supplementary Table 2) shows that all the mutant Spike proteins are most susceptible to the European population supporting the spread history of the virus throughout the continent (Fig. 2).

Discussion

The Spike glycoprotein of SARS-CoV-2 is one of the major determinants of viral infectivity. The protein has 2 subunits, S1 which is responsible for host receptor interaction and S2 which leads to membrane fusion so that the virus genome can enter the host cell to proliferate. Receptor binding domain specially the receptor binding motif of S1 subunit is a crucial region responsible for the host binding. The virus first identified in Wuhan, China in December 2019 has undergone strain diversification and certain strains have been tested to have greater infectivity.

The extent of infectivity of the virus is also highly linked to the expression level of its specific receptor in the cell which in this case is the hACE-2 receptor. Lung has been seen to be one of the major organs with highly expressed hACE-2 according to several studies. hACE-2 expression was significantly elevated in lung adenocarcioma (LUAD) and lung squamous cell carcinoma (LUSC) compared to normal tissues; thus making the patients with these types of cancer more susceptible to SARS-CoV-2 infection [23]. The expression level of hACE-2 has also been seen to be aberrantly expressed in many tumors [24].

The hACE-2 protein level was seen to be high in renal cancer and colorectal cancer according to protein atlas which points to the direct relation between SARS-CoV-2 infectivity in these cancers due to the high expression of the receptors as has been seen for lung cancer patients too. The protein is also seen to be highly expressed in GI tract, liver, kidney and reproductive tract with average expression in lungs. This

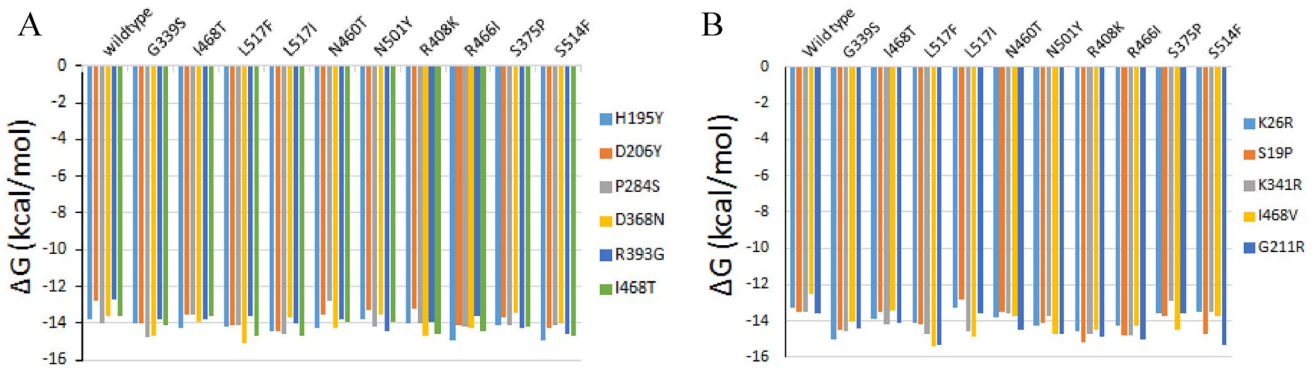

Fig. 2 Effect of Spike protein mutations upon hACE-2 (**A**) disease variants and (**B**) population variants on the basis of binding affinity (ΔG kcal/mol) of host-virus protein–protein complex

also makes since as COVID infection has been associated with kidney dysfunction as also affecting the GI tract and lungs causing symptoms in digestive and respiratory systems. The data is also obtained from protein atlas. In the colon the protein is maximally expressed in C6 enterocytes [25].

Cancer patients are reported three times more susceptible to SARS-CoV-2 infection with possible poor prognosis than individuals without cancer because of their systemic immunosuppressive state caused by the malignancy and anticancer treatments, such as chemotherapy or surgery. The mortality rate of cancer patients who contracted SARS-CoV-2 virus was reported to be 6% in comparison to 1% for healthy people in China. According to the World Health Organization, the case fatality rate for COVID-19 patients with cancer as a comorbid condition was 7.6% vs. a case fatality rate of 3.8% in the entire COVID-19 population [26]. Several studies show that cancer subjects are at a higher risk of critical events in (48–54% of cases vs. 16% in the general population) and death (5.6–29% vs. 3.4% in the general population) [27]. Several studies have shown that patients with hematological, lung or breast cancer are more vulnerable than comorbid conditions involving other cancers. In the Veneto study, breast and hematological cancers were associated with a higher risk of both hospitalization and death. Lung cancer was also associated with a fourfold higher risk of death owing to SARS-CoV-2 infection [28].

Protein network analysis was performed to see the effect of the mutation on the role of the residue with respect to the entire structure. Most mutations led to an increase in the closeness of the node while almost all the mutations led to an increase of the betweenness. Changes were seen in the clustering co-efficient, eigenvector centrality and eccentricity as well as a result of the mutations. The mutations thus in most cases led to an increase in availability of the residue in the various shortest paths connecting 2 residues in the protein. All mutations apart from those at a separate position led to change in the information flow through the residue and also led to changing of the tendency to be included in various clusters. The mutations also led to change in the functional influence and functional activeness of the residues. Thus all these point to the effect of the mutations not only at the structural level but also at the functional level showing the overall influence of the same on the protein.

In this study, the effect of mutations of the viral Spike protein upon hACE-2 pathogenic variants and geographical population variants are examined on the basis of binding affinity (calculating ΔG) of virus-host (Spike-hACE-2) protein complex. Many studies have already proved that strong binding between virus-host proteins is the pin greater infectivity of the virus. We have identified the RBD mutations which can possibly increase the binding affinity of the Spike-receptor complex leads to increased propensity of the Spike protein to interact with the host receptor thus in the process increasing the probability of a successful infection. It has been shown that for both pathogenic variants and geographical population variants, the mutant Spike proteins possess high binding affinity (i.e. more negative ΔG) with the host proteins. So, it may be concluded that these mutations which increase the binding affinity between the proteins actually lead to a greater infectivity when the other parameters are kept constant and only the receptor binding ability is taken into consideration.

This study manifest the effect of the mutations of Spike protein upon various variants of the host receptor would serve some positive guidance for specific targeted therapy which would be highly specific based on which mutant strain is causing the infection as also on the factor that which population the virus is infecting.

Conclusion

Though several vaccines have been invented, SARS-CoV-2 infection continues throughout the world. Future development of medications against the virus requires detail study of each important part of the virus. This study focus on the mutational effect of Spike protein in viral infectivity taking other parameters into consideration. Variants of h-ACE2 protein linked with different types of cancer, have shown some extent of additional fatality as well as this study manifest the effect of mutation variants depending on different geographical populations. Contemplating all the results of this study, it can be suggested that due to high mutation in both virus and host receptor (RBD) it would require different medications specially targeted for particular variants.

Author contributions Protocol designed and conceptualized by D.C., manuscript preparation and data analysis done by J.D and S.M. Project was done under the supervision of K.G. The manuscript was reviewed and approved by all authors.

Declarations

Conflicts of interest The authors declare that there are no conflict of interests.

References

1. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med. 2020;382(8):727–33.
2. Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet. 2020;395(10224):565–74.
3. Letko M, Marzi A, Munster V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. Nat Microbiol. 2020;5:562–9.
4. Fehr AR, Perlman S. Coronaviruses: an overview of their replication and pathogenesis. Methods Mol Biol. 2015;1282:1–23.
5. Huang Y, Yang C, Xu Xf, et al. Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. Acta Pharmacol Sin. 2020;41:1141–9.
6. Bosch BJ, van der Zee R, de Haan CA, Rottier PJ. The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. J Virol. 2003;77(1):8801–11.
7. Xia S, Zhu Y, Liu M, Lan Q, Xu W, Wu Y, et al. Fusion mechanism of 2019-nCoV and fusion inhibitors targeting HR1 domain in spike protein. Cell Mol Immunol. 2020;17:765–7.
8. Watanabe Y, Allen JD, Wrapp D, McLellan JS, Crispin M. Site-specific glycan analysis of the SARS-CoV-2 spike. Science. 2020;369(6501):330–3.
9. Dehouck Y, Kwasigroch JM, Rooman M, Gilis D. BeAtMuSiC: Prediction of changes in protein-protein binding affinity on mutations. Nucleic Acids Res. 2013;41:W333-339.
10. Rodrigues CHM, Pires DEV, Ascher DB. DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability. Nucleic Acids Res. 2018;46(W1):W350–5.
11. Frappier V, Najmanovich R. Vibrational entropy differences between mesophile and thermophile proteins and their use in protein engineering. Protein Sci. 2015;24(4):474–83.
12. Miao Z, Cao Y. Quantifying side-chain conformational variations in protein structure. Sci Rep. 2016;6:37024.
13. Pires DE, Ascher DB, Blundell TL. DUET: a server for predicting effects of mutations on protein stability using an integrated computational approach. Nucleic Acids Res. 2014;42:W314–9.
14. Shin WH, Lee GR, Heo L, Lee H, Seok C. Prediction of Protein Structure and Interaction by GALAXY protein modeling programs. Bio Design. 2014;2(1):1–11.
15. Ko J, Park H, Heo L, Seok C. GalaxyWEB server for protein structure prediction and refinement. Nucleic Acids Res. 2012;40(W1):W294–7.
16. Vangone A, Bonvin AM. Contacts-based prediction of binding affinity in protein-protein complexes. Elife. 2015;4:e07454.
17. Xue LC, Rodrigues JP, Kastritis PL, Bonvin AM, Vangone A. PRODIGY: a web server for predicting the binding affinity of protein-protein complexes. Bioinformatics. 2016;32(23):3676–8.
18. Chakrabarty B, Parekh N. NAPS: network analysis of protein structures. Nucleic Acids Res. 2016;44(W1):W375–82.
19. Ganesan K, Kulandaisamy A, Binny Priya S, Gromiha MM. HuVarBase: a human variant database with comprehensive information at gene and protein levels. PLoS ONE. 2019;14(1):e0210475.
20. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141, 456 humans. Nature. 2020;581:434–43.
21. Li Q, Wu J, Nie J, Li Z, Hao H, Liu S, Zhao C, Zhang Q, Liu H, Nie L, Qin H, Wang M, Lu Q, Li X, Sun Q, Liu J, Zhang L, Li X, Huang W, Wang Y. The impact of mutations in SARS-CoV-2 spike on viral infectivity and antigenicity. Cell. 2020;182(5):1284–94.
22. Storz JF. Compensatory mutations and epistasis for protein function. Curr Opin Struct Biol. 2018;50:18–25.
23. Kong Q, Xiang Z, Wu Y, Gu Y, Guo J, Geng F. Analysis of the susceptibility of lung cancer patients to SARS-CoV-2 infection. Mol Cancer. 2020;19(1):80.
24. Zhang H, Quek K, Chen R, Chen J, Chen B. Expression of the SARS-CoV-2 receptor ACE2 reveals the susceptibility of COVID-19 in non-small cell lung cancer. J Cancer. 2020;11(18):5289–92.
25. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A, Olsson I, Edlund K, Lundberg E, Navani S, Szigyarto CA, Odeberg J, Djureinovic D, Takanen JO, Hober S, Alm T, Edqvist PH, Berling H, Tegel H, Mulder J, Rockberg J, Nilsson P, Schwenk JM, Hamsten M, von Feilitzen K, Forsberg M, Persson L, Johansson F, Zwahlen M, von Heijne G, Nielsen J, Pontén F. Proteomics. Tissue-based map of the human proteome. Science. 2015;347(6220):1260419.
26. Jyotsana N, King MR. The Impact of COVID-19 on cancer risk and treatment. Cell Mol Bioeng. 2020;13(4):1–7.
27. Allegra A, Pioggia G, Tonacci A, Musolino C, Gangemi S. Cancer and SARS-CoV-2 infection: diagnostic and therapeutic challenges. Cancers (Basel). 2020;12(6):1581.
28. Rugge M, Zorzi M, Guzzinati S. SARS-CoV-2 infection in the Italian Veneto region: adverse outcomes in patients with cancer. Nat Cancer. 2020;1:784–8.