Structural and functional characterization of SARS-CoV-2 nucleocapsid protein mutations identified in Turkey by using *in silico* approaches

Betul Akcesme[1], Burcin Erkal[2], Zehra Yaren Donmez[1]

[1]Department of Basic Medical Sciences, Division of Medical Biology, Faculty of Medicine, University of Health Sciences, Istanbul, Turkey; [2]Department of Molecular Biology and Genetics, Faculty of Arts and Sciences, Yıldız Technical University, Istanbul, Turkey

Summary. – Missense mutations in the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus may cause changes in the structure of proteins. The nucleocapsid (N) protein is an important target for drugs and vaccines. The main purpose of this study is to detect missense mutations in the SARS-CoV-2 N protein and to reveal the effects of these mutations on protein structure by using *in silico* approaches. 161 missense mutations of the N protein were determined in 2286 SARS-CoV-2 genomes derived from the GISAID EpiCoV database in the Turkish population. Identified 161 missense mutations were analyzed by using sequence and structure-based methods to predict effects of mutation on function and structure of SARS-CoV-2 N protein. These analyzes revealed that some mutations showed deleterious effects and change of stability and flexibility of nucleocapsid protein. D3L, S194L, S235F, and P13L (Omicron variant) mutations were further analyzed in our study due to their importance in the literature and in our results. Even though, our findings are essential for research of SARS-CoV-2 virus, *in vitro* and *in vivo* validations are necessary.

Keywords: nucleocapsid protein; SARS-CoV-2; missense mutations; protein stability; protein flexibility

Introduction

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which affects the whole world rapidly since the end of 2019 and causes an epidemic of viral pneumonia, first appeared in Wuhan in China. The World Health Organization (WHO) declared coronavirus disease 2019 (COVID-19) as a pandemic on 11th March (WHO, 2022a).

Up to date, the number of cases seen all over the world were estimated at around 450 million and the number of deaths at 6 million. However, in Turkey the number of

cases recorded, was around 14 million and the number of deaths around 95 thousand (WHO, 2022b).

Coronaviruses are enveloped and positive-strand RNA viruses belonging to the family *Coronaviridae*. Their genome is ~30 kb in size. The genome consists of specific genes, spike protein (S), envelope protein (E), membrane protein (M), nucleocapsid protein (N) and open reading frames (ORF3a, ORF6, ORF7a, ORF7b, ORF8, ORF9 and ORF10) forming protease and replicase (1a–1b) (Naqvi *et al.*, 2020). Among these proteins, spike protein has a special role. S protein facilitates entry into the host cell via binding of angiotensin-converting enzyme 2 (ACE2) cell surface receptor. SARS-CoV-2 is transported through the respiratory route and affects lungs (Cascarina and Ross, 2020). Among the symptoms seen in patients infected with the SARS-CoV-2 are high fever, cough, shortness of breath, fatigue, muscle and headache, loss of smell and taste (Zeng *et al.*, 2020).

Among other proteins, SARS-CoV-2 nucleocapsid protein has an undeniable importance. The N protein is

---

a multifunctional RNA-binding protein that is required for the viral RNA transcription and replication processes. The primary function of the N protein is to bind to the viral RNA genome, form the RNP complex and drive viral mRNA transcription and replication. It plays many important roles in the regulation of host cell metabolism, such as cytoskeletal organization and immune regulation (Savastano *et al*., 2020; Khan *et al*., 2021). Furthermore, the double-stranded RNA-binding activity of the N protein acts to combat antiviral responses mediated by host RNA silencing. RNA silencing acts as a viral repressor. High amount of double-stranded RNA is expressed during infection and has been reported to induce both humoral and cellular immune responses after infection (Peng *et al*., 2020). Some studies demonstrate that the N protein regulates host-pathogen interactions and has an impact on actin organization, host cell cycle and apoptosis pathways (Ding *et al*., 2016; Pizzato *et al*., 2022). Further, it has been shown that interaction of SARS-CoV-2 N with envelope protein as well as with the membrane protein are essential to further trigger the other molecular mechanisms during the viral infection (Tseng *et al*., 2014; He *et al*., 2004). Structural E protein has a vital role for viral assembly and infectivity, whereas structural M protein promotes membrane fusion of virus and host cell through protein-protein interaction (Gorkhali *et al*., 2021).

N protein has two structural domains, N-terminal domain (NTD) and C-terminal domain (CTD). These two domains are linked by an intrinsically disordered region (IDR) called the binding region. It is known that NTD is responsible for RNA binding and CTD has both oligomerization and RNA binding function. Also, NTD and CTD have two IDRs which are placed at the beginning and ending of the nucleocapsid protein, called the N-arm and C-tail. The IDR has important role in modulating the RNA-binding activity of NTD and CTD and oligomerization. The linker region also has a Serine-Arginine (SR) region containing phosphorylation sites in which primary phosphorylation plays role in further functioning of the protein. Each of these is a potential drug target of antiviral inhibitors (Peng *et al*., 2020; Khan *et al*., 2021).

Amino acid sequences of proteins determine their structure and function. A single amino acid mutation can change molecular function and 3D structure of proteins and causes diseases (Stone and Sidow, 2005). Therefore, it is important to understand the missense mutations and their impact on gene expression and protein functions.

Missense mutations can disrupt not only protein stability, but also interactions of other proteins, nucleic acids, polysaccharides, ligands, and metal ions (Dehghanpoor *et al*., 2018). Missense mutations became even more important with the SARS-CoV-2 virus (Teng *et al*., 2021). Especially, how potential mutations that may occur in the virus will affect the viral function and how it will change its interaction with the host cell is a matter of curiosity for the whole world (Mohammadi *et al*., 2021). Mutagenesis studies in physical proteins can provide insight into the effects of amino acid substitutions. However, these studies require time and financial resources (Jacob *et al*., 2021). In recent studies, bioinformatic analysis are preferred to predict the effects of a mutation on a protein structure.

In this study, our aim is to determine SARS-CoV-2 nucleocapsid protein related mutations identified in Turkey and to analyze defined mutations to understand their possible effects on protein function and structure which can change their flexibility and stability features.

Materials and Methods

Identification of viral sequences. SARS-CoV-2 genome sequences were obtained from GISAID EpiCoV database (GISAID, 2021) between 01 March 2020 and 15 March 2021. These sequences were eliminated using following filters. Location was defined "Europe/Turkey" to evaluate only Turkey related sequences and 2557 genome sequences were included. In addition, full genome length (sequence size 29 kilobases or greater), high coverage (< 0.05% specific mutation, no insertion or deletion specified, < 0.1% sequence) were included and low coverages were excluded (low coverage excluded) (>5% Ns). 2286 viral genome sequences were found and downloaded in FASTA format.

Detection of nucleocapsid protein related mutations. Missense mutations related to nucleocapsid protein were found in genome sequences using the Nextclade v1.7.1 (Aksamentov *et al*., 2021). Nextclade is a network tool that facilitates the analyses of mutations in the SARS-CoV-2 genome comparing to reference sequence (MN908947.3). FASTA format files of 2286 genome sequences was uploaded to the Nextclade. The results were saved in CSV and TSV formats.

Sequence-based analysis of nucleocapsid protein. The SARS-CoV-2 nucleocapsid protein sequence (P0DTC9) was downloaded from the NCBI database. PROVEAN (Choi and Chan, 2015) and PredictSNP (Bendl *et al*., 2014) network servers were used for sequence-based analysis to predict the mutation effect on protein function. The file in TXT format containing the P0DTC9 encoded SARS-CoV-2 nucleocapsid protein sequence and defined missense mutations were uploaded and analyzed.

Obtaining the 3D structure of the nucleocapsid protein. The three-dimensional N protein structure was downloaded from the I-Tasser (Yang *et al*., 2015). The I-Tasser server uses PDB coded 3D protein structures such as PDB ID: 6M3M and PDB ID: 6YUN to reveal the desired structure. The code (QHD43423.pdb) for the nucleocapsid protein was downloaded from the I-Tasser. The template modelling (TM) score of this structure was 0.97.

Structure-based analysis of nucleocapsid protein. Structure-based analysis was performed with the DynaMut web server

Table 1. Missense mutations located on different domains of N protein

Position of missense mutations	Number of missense mutations
N-Arm	39
NTD	29
Linker	44
CTD	23
C-Tail	26

Protein regions of N protein

to predict the impact of mutation on protein stability and flexibility (Rodrigues *et al*., 2018). A file in txt format containing the N protein structure encoded [QHD43423.pdb] and defined missense mutations were uploaded and analyzed.

Results

Identification of missense mutations

161 different missense mutations of N protein were found in 2286 sample. Number of mutations found in different regions of nucleocapsid protein are shown in Table 1.

Mutation effects on nucleocapsid protein

Eighteen deleterious and 143 neutral missense mutations were detected by performing analysis with PROVEAN network server. In addition, resulting from analysis by PredictSNP network server, 65 missense mutations were identified as deleterious whereas 96 missense mutations were neutral. Moreover, 70 missense mutations were found to be deleterious and 91 missense mutations were neutral in SNAP database. 91 missense mutations were found as deleterious and 70 missense mutations were found as neutral in SIFT database. Finally, 87 missense mutations were found as deleterious and 74 missense mutations were found as neutral using MAPP tool (Supplementary Table 2). 49 mutations identified as deleterious in at least 4 servers among PROVEAN, PredictSNP, PhD-SNP, PolyPhen 1 and 2, SIFT, SNAP are shown in the Table 2.

Stabilization and flexibility of nucleocapsid protein analysis

ΔΔG DynaMut values represent change of stabilization while ΔΔSVib ENCoM values represent change of flexibi-

lity. If the values are higher than 0 kcal/mol, the stability and flexibility of protein increases. If the values are less than 0 kcal/mol, the stability and flexibility of protein decreases. Both ΔΔG DynaMut and ΔΔSVib ENCoM values of 49 mutations identified in the Table 2 are also shown in the Table 3.

103 mutations increased stability of the N protein 3D structure based on stabilization analysis performed with the DynaMut web server. On the other hand, 58 mutations decreased stability of the protein structure. Missense mutations that increased the flexibility of N protein 3D structure are shown in Table 2 (Supplementary Table 3).

According to protein entropy analysis performed by the DynaMut network server; 77 missense mutations increased flexibility of the 3D structure (Supplementary Table 4) while 84 missense mutations decreased the flexibility. Number of mutations in different domains of N protein that either increased or decreased stability and flexibility are shown in the Table 4.

Selection of key mutations on protein structure

D3L, S194L, S235F mutations among 161 missense mutations were selected for our study, as key mutations for further analysis due to their high frequency and highlighted importance in the literature. Besides high frequency mutations, we selected P13L mutation as key mutation. Although P13L mutation frequency is very low (n = 1) in our SARS-CoV-2 data, it is one of the mutations found in Omicron variant. This variant is in the recent list of Variant of Concern (VOC) in the WHO and its importance is increasing every day.

Interatomic interactions of key mutations

Interatomic interaction analysis of D3L, S194L, S235F and P13L mutations were performed with DynaMut net-

Table 2. Mutations identified as deleterious in at least 4 servers among PROVEAN, PredictSNP, PhD-SNP, PolyPhen 1 and 2, SIFT, SNAP

Mutation	Gene location	Frequency	Mutation	Gene location	Frequency
D3L	A28281T,T28282A	382	S194L	C28854T	164
D3Y	G28280T	33	T198I	C28866T	42
A35V	C28377T	17	G204L	G28883C,G28884T	23
G25C	G28346T	6	N196Y	A28859T	13
Q9H	G28300T	5	S188P	T28835C	8
G34W	G28373T	4	R195I	G28857T	7
P13L	C28311T	3	R191L	G28845T	3
R40C	C28391T	3	P207H	C28893A	2
G25V	G28347T	2	G204P	G28883C, G28884C	1
G44C	G28403T	2	Q241L	A28995T	1
S33I	G28371T	2	R185C	C28826T	1
D22Y	G28337T	1	S180C	A28811T	1
Q9R	A28299G	1	S201C	A28874T	1
R14C	C28313T	1	S202C	A28877T	1
R32H	G28368A	1	S206F	C28890T	1
S21P	T28334C	1	T334I	C29274T	9
S2F	C28278T	1	A308S	G29195T	2
T16M	C28320T	1	T271I	C29085T	2
D128Y	G28655T	6	V350G	T29322G	2
D144H	G28703C	2	P364Q	C29364A	1
D63Y	G28460T	1	S327L	C29253T	1
P151A	C28724G	1	D377Y	G29402T	22
P151S	C28724T	1	D401Y	G29474T	21
V72L	G28487C	1	S416L	C29520T	2
Q418H	G29527T	1			

work server (Fig. 1). For S194L mutation, it was observed that the carbonyl bond and ionic interaction in the wild-type structure were disrupted, and new hydrogen bond formation was formed. For S235F mutation, it was determined that the hydrogen bond in the wild-type structure was broken, and weak hydrogen bond formation occurred. New bonds are formed and clearly observed because of P13L mutation. Finally, no serious change was observed between D3L wild-type and mutant.

Entropy energy change and atomic instability analysis of key mutations

Δ vibrational entropy energy between wild-type and mutant of D3L, S194L, S235F and P13L was also identified (Fig. 2). Amino acids are colored according to the vibrational entropy change due to mutation. Blue and red colors represent rigidification and gaining flexibility of the structure, respectively. S194L and D3L mutations increase the flexibility of N protein whereas S235F and P13L mutations reduce the flexibility of the N protein according to the entropy energy change analysis performed with the DynaMut network server.

Discussion

SARS-CoV-2, which caused the COVID-19 pandemic, deeply affect the whole world with its contagiousness. The virus is transmitted to humans and causes serious symptoms. It is the focus of attention of researchers to solve the structure of the virus, to diagnose and to find the treatment of the disease (Peng *et al.*, 2020).

In the *Coronaviridae* virus family, the N protein is a multifunctional RNA-binding protein required for viral RNA transcription and replication (Mohammadi *et al.*, 2021). Some studies demonstrated that the interaction between the N and E protein has role in the release of
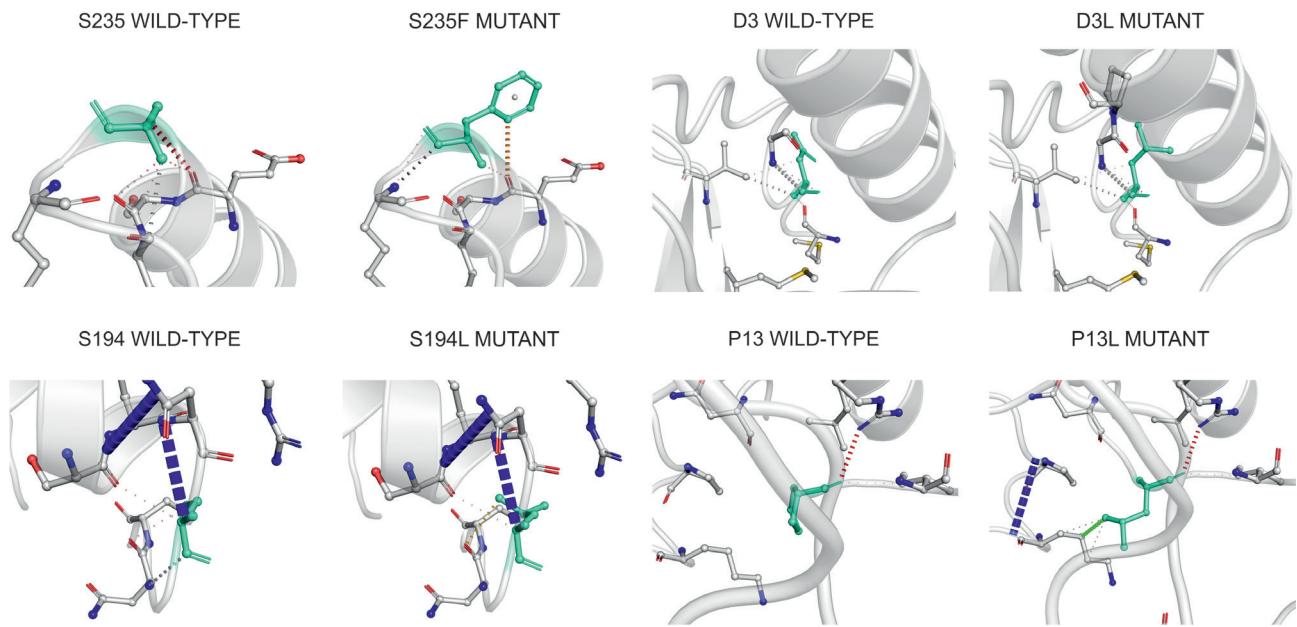
Fig. 1

Prediction of interatomic interaction changes upon targeted D3L, S194L, S235F, and P13L mutations

Red color shows hydrogen bonds, orange color weak hydrogen bonds, blue color halogen bonds, yellow color ionic interactions, light blue color aromatic interaction, green hydrophobic contact, pink carbonyl interactions, gray color Van der Waals interactions. Wild-type and mutant side chains are colored light green. Interactions and bonds between them are shown with dashed lines.

Table 3. DynaMut, structure-based analysis of mutations identified as deleterious in at least 4 servers among PROVEAN, PredictSNP, PhD-SNP, PolyPhen 1 and 2, SIFT, SNAP

Mutation	ΔΔG DynaMut (kcal/mol)	ΔΔG mCSM (kcal/mol)	ΔΔG SDM (kcal/mol)	ΔΔG DUET (kcal/mol)	ΔΔS ENCoM (kcal^{-1}/mol^{-1})	ΔΔSVibENCoM (kcal^{-1}/mol^{-1})
D3L	-0,548	0,194	-0,190	0,380	0,044	0,044
D3Y	-0,382	-0,065	-0,430	-0,153	-0,030	-0,030
A35V	1,13	-0,270	1,150	0,296	-0,174	-0,174
G25C	-0,092	-0,966	-0,420	-0,842	-0,022	-0,022
Q9H	0,062	-0,500	0,740	-0,199	-0,365	-0,365
G34W	-1,112	-1,143	-2,140	-1,531	0,073	0,073
P13L	1,523	-0,306	-0,040	0,008	-0,693	-0,693
R40C	0,137	0,043	-0,400	-0,005	0,883	0,883
G25V	0,063	-0,494	0,150	-0,146	0,010	0,010
G44C	-0,843	-1,079	-0,320	-0,938	0,186	0,186
S33I	0,216	-0,226	1,370	0,356	-0,055	-0,055
D22Y	1,464	0,094	0,540	0,267	-1,928	-1,928
Q9R	0,411	-0,045	0,54	0,303	-0,397	-0,397
R14C	-0,223	0,023	-0,240	0,003	0,852	0,852
R32H	0,233	-0,624	0,120	-0,529	-0,055	-0,055
S21P	-0,140	-0,187	-0,050	0,024	-0,358	-0,358
S2F	1,005	-0,920	0,360	-0,740	-0,712	-0,712
T16M	-0,316	-0,187	0,620	0,045	-0,034	-0,034
D128Y	-0,036	0,237	0,120	0,196	0,244	0,244

Table 3. (continued)

Mutation	ΔΔG DynaMut (kcal/mol)	ΔΔG mCSM (kcal/mol)	ΔΔG SDM (kcal/mol)	ΔΔG DUET (kcal/mol)	ΔΔS ENCoM (kcal⁻¹/mol⁻¹)	ΔΔSVibENCoM (kcal⁻¹/mol⁻¹)
D144H	0,210	-0,377	-0,040	-0,385	-0,322	-0,322
D63Y	2,299	1,688	0,330	1,484	-0,43	-0,943
P151A	-0,129	-0,525	0,510	-0,198	0,297	0,297
P151S	-0,323	-0,999	-1,050	-1,016	-0,055	-0,055
V72L	0,348	-0,362	-1,210	-0,325	-0,626	-0,626
S194L	0,647	-0,446	2,270	0,313	0,044	0,044
T198I	0,398	-0,275	1,370	0,169	-0,070	-0,070
G204L	0,013	-0,441	-1,860	-0,565	-0,901	-0,901
N196Y	-1,169	-0,641	1,109	-0,542	0,334	0,334
S188P	0,335	0,492	-1,620	0,375	0,141	0,141
R195I	0,481	0,002	0,280	0,199	0,349	0,349
R191L	-0,358	-0,854	0,710	-0,380	-0,021	-0,021
P207H	0,053	-0,015	0,910	0,198	-0,051	-0,051
G204P	0,684	-0,327	-3,610	-0,844	-0,572	-0,572
Q241L	0,346	0,166	1,170	0,693	0,089	0,089
R185C	-1,292	-2,110	0,330	-1,721	1,042	1,042
S180C	0,265	-0,076	0,830	0,239	-0,059	-0,059
S201C	0,321	-0,212	1,180	0,174	0,176	0,176
S202C	-0,377	-0,206	0,620	0,048	0,315	0,315
S206F	0,390	-0,864	0,940	-0,621	0,016	0,016
T334I	-0,065	-0,267	0,640	0,200	0,267	0,267
A308S	0,159	-0,891	-1,140	-0,774	-0,104	-0,104
T271I	0,569	-0,219	0,890	0,336	-0,398	-0,398
V350G	-0,544	-1,871	-0,960	-1,855	0,296	0,296
P364Q	-0,247	-0,400	0,640	0,072	-0,024	-0,024
S327L	0,753	-0,317	1,030	0,098	-0,122	-0,122
D377Y	0,953	-0,389	0,770	-0,046	-0,545	-0,545
D401Y	1,253	-0417	0,780	-0,068	-0,487	-0,487
S416L	-0,308	-0,347	0,980	0,054	0,086	0,086
Q418H	0,350	-0,212	0,910	-0,001	-0,222	-0,222

Table 4. Number of mutations in the domains of N protein that either increase or decrease stability and flexibility

Stabilization results				Flexibility results			
Increase		Decrease		Increase		Decrease	
Position	No. of mutations	Position	No. of mutations	Position	No. of mutations	Position	No. of mutations
N-Arm	22	N-Arm	17	N-Arm	17	N-Arm	22
NTD	17	NTD	12	NTD	11	NTD	18
Linker	36	Linker	8	Linker	21	Linker	23
CTD	10	CTD	13	CTD	12	CTD	11
C-Tail	18	C-Tail	8	C-Tail	16	C-Tail	10

the SARS-CoV from the cell whereas the N–M protein complex is involved in the formation of the coronavirus (Lopez *et al.*, 1994; He *et al.*, 2004; Boscarino *et al.*, 2008; Tseng *et al.*, 2014). Therefore, the interaction between each SARS-CoV-2 mutant N protein and the E protein may increase virus production, on the other hand decreased binding affinity of the S194L N-M protein complex may weaken virus formation (Wu *et al.*, 2021). In this study, ΔΔG DynaMut values (0.647 kcal/mol) revealed that S194L mutation increases the stabilization of the 3D structure. ΔΔSvib ENCoM values (0,044 kcal^{-1}/mol^{-1}) revealed that S194L mutation decreases the flexibility of the N protein 3D structure. In the sequence-based mutation pathogenicity analysis, the S194L mutation is found to be deleterious. It was determined that carbonyl bond and ionic interaction were disrupted, and a new hydrogen bond formation is seen with S194L mutation. It is thought that the S194L mutation affects the binding affinity of the N-M protein complex and the interaction between the N and E proteins.

In recent studies, it is demonstrated that the SR-rich linker region plays an important role in the intracellular

signaling pathways via phosphorylation in serine side chains (Wootton *et al.*, 2002; Mcbride *et al.*, 2014; Azad, 2021). In our study, 44 missense mutations were detected in SR-rich linker region. S194L, G204L, N196Y, S188P, R195I, R191L, R185C, S180C, S201C, S202C, and S206F mutations are detected as deleterious in the sequence-based analyzes in our study. It was determined that the ΔΔG DUET values of G204P (-0,844 kcal/mol), G204Q (-0,623 kcal/mol) mutations destabilize the 3D structure of the N protein and the flexibility of the structure is decreased (ΔΔSVib ENCoM (-0,572 kcal^{-1}/mol^{-1}) (-0,938 kcal^{-1}/mol^{-1}). It was also observed that the R185C (-1,292 kcal/mol) mutation destabilizes the 3D structure of the N protein. Therefore, it is possible that these missense mutations in the serine side chains may affect phosphorylation-dependent signaling.

NTD and CTD regions of N protein are known to be involved in RNA binding (Zeng *et al.*, 2020). In a recent study, by using the PROVEAN network server, 13 and 12 deleterious mutations are detected in the CTD and NTD domain of the N protein, respectively (Das and Roy, 2021). The deleterious D128Y and P151S mutations in NTD and T334I and T271I mutations in CTD are consistent with

Flexibility: D3L

Flexibility: P13L

Flexibility: S194L
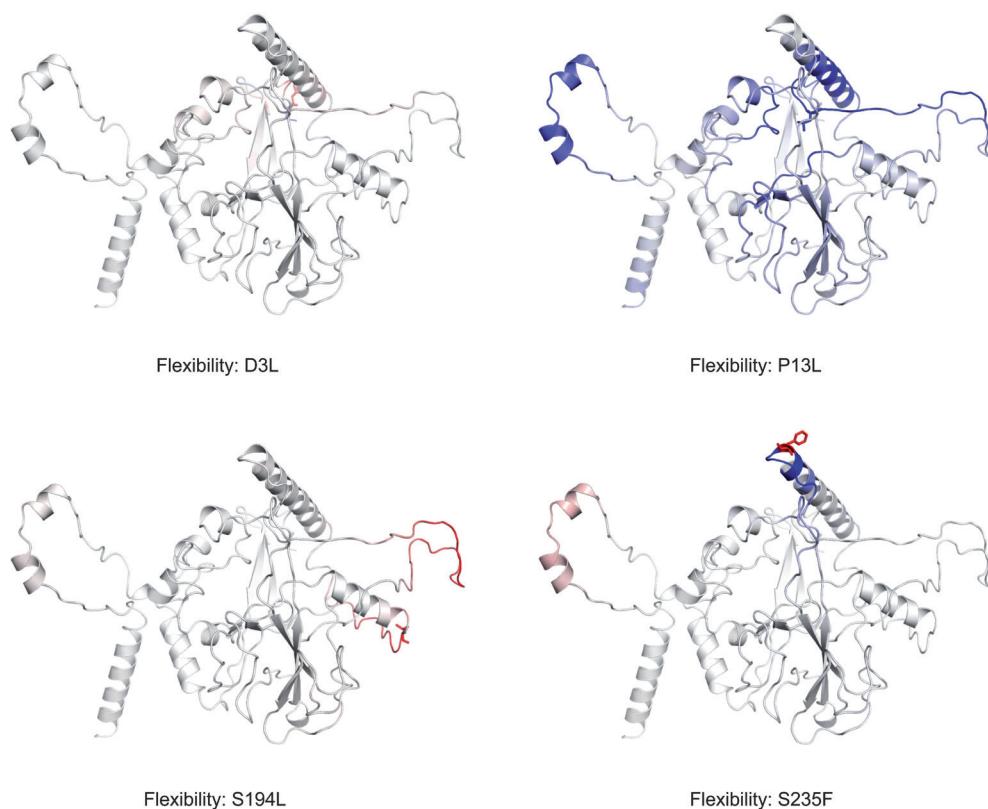
Flexibility: S235F

Fig. 2
**Vibrational entropy changes upon targeted D3L, S194L, S235F, and P13L mutations**
Blue color indicates rigidification of the structure and red color indicates an increase in flexibility.

our study. In the analysis of PredictSNP server, 6 and 5 deleterious mutations were detected in the NTD and CTD domains of N protein, respectively. How these mutations affect binding ability of the protein to the RNA should be the subject of further *in vitro* studies.

Azad *et al.* (2021) revealed that one of the B and T cell epitopes is located between the 305-340 side chains of the N protein. Mutations in the 305–340 side chains may change the epitope characteristics and as a result immunological response of the host may be affected. In this study, 11 mutations in the CTD (305–340 amino acids) are detected. T334I is found as deleterious by using PROVEAN and PredictSNP tool. Furthermore, T334A, M234I and P326S mutations are found to destabilize the 3D structure of the N protein by using the DynaMut network server. Therefore, it reinforces the notion that mutations occurring in various populations and regions should be considered in the development of vaccines and drugs targeting the N protein of SARS-CoV-2.

Irregular regions do not have a well resolved tertiary structure; however, the IDRs of N proteins have been found to play an important role in binding with viral genomic RNA (Chang *et al.*, 2008). 109 mutations in the IDR regions were detected in our study. The analysis with PROVEAN and PredictSNP network servers revealed that S194L, G204L, N196Y, S188P, R195I, R191L, R185C, S180C, S201C, S202C and S206F mutations are deleterious. As a result of the ΔΔG DUET values, it was determined that the G204L (-0,565 kcal/mol) mutation decreases the stability of the N protein structure, and the R195I (0,199 kcal/mol) mutation increases the stability of the N protein structure. In addition, it is determined that the R195I (0,349 kcal$^{-1/}$mol^{-1}) mutation increases the flexibility of the 3D structure of the N protein according to ΔΔSVib ENCoM values. These mutations may have a great potential to alter the N protein binding stability of the viral genome.

Drugs that can interfere with N protein have great pharmacological interest. Drugs inhibiting the CTD region are promising candidates for preventing virus formation. Researchers have identified several promising drugs including Conivaptan, Ergotamine, Venetoclax, Rifapentine against SARS-CoV-2 (Kadioglu *et al.*, 2021). T334I and T271 mutations in the CTD of the N protein were found to be deleterious in the analysis performed with PROVEAN and PredictSNP network servers in our study. Mutations detected in this region give great opportunity for studies related to drug-protein interactions. We think that mutations detected in this region of nucleocapsid will contribute greatly to drug-protein interaction studies.

Hu *et al.* (2021) have determined the interaction sites on the N protein with Ceftriaxone, Arbidol and Lopinavir drugs. It was found that Arbidol could bind to the NTD region while Lopinavir could bind to the CTD region of the N protein (Hu *et al.*, 2021). T334 (T334I: -0,065 kcal/mol, T334A:-0,653 kcal/mol) mutations are detected in the side chains where the lopinavir drug interacts. This mutation destabilized the N protein. Moreover, according to the ΔΔSVib ENCoM T334I: 0,149 kcal$^{-1/}$mol^{-1}, T334A: 0,267 kcal$^{-1/}$mol^{-1}) value that increased the flexibility of the 3D structure of the N protein. In addition, mutations were detected in the T166 (-0,226 kcal$^{-1/}$mol^{-1}) and L167 (-0,543 kcal$^{-1/}$mol^{-1}) side chain where the arbidol drug interacts, and as a result of the ΔΔSVib ENCoM value, these mutations were found to reduce the flexibility of the N protein structure. It is important to determine how mutations in these specific structural areas and how the interactions of these small molecules within these areas will change.

Vilar and Isom emphasized that the possible effects of D3L, S235F, and S194L mutations should be investigated in further studies (Vilar and Isom, 2021). D3L, S235F, and S194L mutations, also detected in our study, that have been detected more frequently in the SARS-CoV-2 genome are mostly deleterious. Therefore, D3L, S194L, S235F, and P13L mutations were selected as key mutations for the further analysis in our study. For chemical bond and interaction analysis, atomic instability analysis of N protein, and analysis of deformation energies of selected mutations were calculated using DynaMut web server.

In the chemical bond and interaction analysis with the DynaMut web server, no change was detected with D3L mutation. As a result of the ΔΔSVib ENCoM value of the S235F (-0,135 kcal$^{-1/}$mol^{-1}) mutation, it was determined that the flexibility of the 3D structure of the N protein decreased. Otherwise, it was determined that the hydrogen bond in the wild-type structure was disrupted and weak hydrogen formation occurred in S235F mutation. The last and the most important mutation P13L seen in Omicron variant causes formation of new bonds in the mutant form of the protein. Further, rigidification is also observed due to mutation P13L. ΔΔS$_{Vib}$ ENCoM value is -0.693 kcal.mol^{-1}.K^{-1} and it decreases molecule flexibility

Overall, our results revealed that mutations in the SARS-CoV-2 N protein can alter its binding affinity and interatomic interactions. The integrative bioinformatics methods are significantly time and cost saving for research communities to gain more insight about structural and molecular mechanism of COVID-19. However, our findings need to be validated with further *in vitro* and *in vivo* studies to contribute to development of drug and vaccine research against SARS-CoV-2. Furthermore, our study shows that more *in silico* and laboratory research is required to investigate the precise role of mutations in the N protein in replication and pathogenesis of SARS-CoV-2 and to contribute to development of novel therapeutics against COVID-19.

Supplementary information is available in the online version of the paper.

References

Aksamentov I, Roemer C, Hodcroft EB, Neher RA (2021): Nextclade: clade assignment, mutation calling and quality control for viral genomes. J. Open Source Softw. 6(67), 3773. https://doi.org/10.21105/joss.03773

Azad GK (2021): Identification and molecular characterization of mutations in nucleocapsid phosphoprotein of SARS-CoV-2. PeerJ 9, e10666. https://doi.org/10.7717/peerj.10666

Bendl J, Stourac J, Salanda O, Pavelka A, Wieben ED, Zendulka J, Brezovsky J, Damborsky J (2014): PredictSNP: robust and accurate consensus classifier for prediction of disease-related mutations. PLOS Comput. Biol. 10, e1003440. https://doi.org/10.1371/journal.pcbi.1003440

Boscarino JA, Logan HL, Lacny JJ, Gallagher TM (2008): Envelope protein palmitoylations are crucial for murine coronavirus assembly. J. Virol. 82(6), 2989–2999. https://doi.org/10.1128/JVI.01906-07

Cascarina SM, Ross ED (2020): A proposed role for the SARS-CoV-2 nucleocapsid protein in the formation and regulation of biomolecular condensates. The FASEB J. 34(8), 9832–9842. https://doi.org/10.1096/fj.202001351

Chang CK, Hsu YL, Chang YH, Chao FA, Wu MC, Huang YS, Hu CK, Huang TH (2008): Multiple nucleic acid binding sites and intrinsic disorder of severe acute respiratory syndrome coronavirus nucleocapsid protein: Implications for Ribonucleocapsid Protein Packaging. J. Virol. 83(5), 2255–2264. https://doi.org/10.1128/JVI.02001-08

Choi Y, Chan AP (2015): PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics 31(16), 2745–2747. https://doi.org/10.1093/bioinformatics/btv195

Das JK, Roy S (2021): A study on non-synonymous mutational patterns in structural proteins of SARS-CoV-2. Genome 64(7), 665–678. https://doi.org/10.20944/preprints202008.0621.v2

Dehghanpoor R, Ricks E, Hursh K, Gunderson S, Farhoodi R, Haspel N, Hutchinson B, Jagodzinski F (2018): Predicting the effect of single and multiple mutations on protein structural stability. Molecules 23(2), E251. https://doi.org/10.3390/molecules23020251

Ding B, Qin Y, Chen M (2016): Nucleocapsid proteins: roles beyond viral RNA packaging. Wiley Interdiscip. Rev. RNA 7(2), 213–226. https://doi.org/10.1002/wrna.1326

GISAID (2021) [Online] Website https://www.gisaid.org// [accessed March 15, 2021]

Gorkhali R, Koirala P, Rijal S, Mainali A, Baral A, Bhattarai HK (2021): Structure and function of major SARS-CoV-2 and SARS-CoV proteins. Bioinform. Biol. Insights 22;15:11779322211025876. https://doi.org/10.1177/11779322211025876

He R, Dobie F, Ballantine M, Leeson A, Li Y, Bastien N, Cutts T, Andonov A, Cao J, Booth TF, Plummer FA, Tyler S, Baker L, Li X (2004): Analysis of multimerization of the SARS coronavirus nucleocapsid protein. Biochem. Biophys. Res. Commun. [online] 316(2), 476–483. https://doi.org/10.1016/j.bbrc.2004.02.074

Hu X, Zhou Z, Li F, Xiao Y, Wang Z, Xu J, Dong F, Zheng H, Yu, R (2021): The study of antiviral drugs targeting SARS-CoV-2 nucleocapsid and spike proteins through large-scale compound repurposing. Heliyon 7(3), e06387. https://doi.org/10.1016/j.heliyon.2021.e06387

Jacob JJ, Vasudevan K, Pragasam AK, Gunasekaran K, Veeraraghavan B, Mutreja A (2021): Evolutionary tracking of SARS-CoV-2 genetic variants highlights an intricate balance of stabilizing and destabilizing mutations. mBio 12(4), e0118821. https://doi.org/10.1128/mBio.01188-21

Kadioglu O, Saeed M, Greten HJ, Efferth T (2021): Identification of novel compounds against three targets of SARS CoV-2 coronavirus by combined virtual screening and supervised machine learning. Comput. Biol. Med. 133, 104359. https://doi.org/10.1016/j.compbiomed.2021.104359

Khan MT, Irfan M, Ahsan H, Ahmed A, Kaushik AC, Khan AS, Chinnasamy S, Ali A, Wei DQ (2021): Structures of SARS-CoV-2 RNA-binding proteins and therapeutic targets. Intervirology 64(2), 55–68. https://doi.org/10.1159/000513686

Lopez S, Yao JS, Kuhn RJ, Strauss EG, Strauss JH (1994): Nucleocapsid-glycoprotein interactions required for assembly of alphaviruses. J. Virol. 68(3), 1316–1323. https://doi.org/10.1128/jvi.68.3.1316-1323.1994

McBride R, van Zyl M, Fielding B (2014): The coronavirus nucleocapsid is a multifunctional protein. Viruses 6(8), 2991–3018. https://doi.org/10.3390/v6082991

Mohammadi E, Shafiee F, Shahzamani K, Ranjbar MM, Alibakhshi A, Ahangarzadeh S, Javanmard SH (2021): Novel and emerging mutations of SARS-CoV-2: Biomedical implications. Biomed. Pharmacother. 139, 111599. https://doi.org/10.1016/j.biopha.2021.111599

Naqvi AAT, Fatima K, Mohammad T, Fatima U, Singh IK, Singh A, Atif SM, Hariprasad G, Hasan GM, Hassan MdI (2020): Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach. Biochim. Biophys. Acta. Molecular Basis of Disease 1866(10), 165878. https://doi.org/10.1016/j.bbadis.2020.165878

Peng Y, Du N, Lei Y, Dorje S, Qi J, Luo T, Gao GF, Song H (2020): Structures of the SARS -CoV-2 nucleocapsid and their perspectives for drug design. EMBO J., 39(20), e105938. https://doi.org/10.15252/embj.2020105938

Pizzato M, Baraldi C, Boscato Sopetto G, Finozzi D, Gentile C, Gentile MD, Volpini, L (2022): SARS-CoV-2 and the host cell: a tale of interactions. Front. Virol., 46. https://doi.org/10.3389/fviro.2021.815388

Rodrigues Carlos HM, Pires DE, Ascher DB (2018): DynaMut: predicting the impact of mutations on protein con-

formation, flexibility and stability. Nucleic acids res. 46, W350-W355. https://doi.org/10.1093/nar/gky300

Savastano A, Ibáñez de Opakua A, Rankovic M, Zweckstetter M (2020): Nucleocapsid protein of SARS-CoV-2 phase separates into RNA-rich polymerase-containing condensates. Nat. Commun. 27;11(1), 6041. https://doi.org/10.1038/s41467-020-19843-1

Stone EA, Sidow A (2005): Physicochemical constraint violation by missense substitutions mediates impairment of protein function and disease severity. Genome Res. 15(7), 978–986. https://doi.org/10.1101/gr.3804205

Teng S, Sobitan A, Rhoades R, Liu D, Tang Q (2021): Systemic effects of missense mutations on SARS-CoV-2 spike glycoprotein stability and receptor-binding affinity. Brief. Bioinform. 22(2), 1239–1253. https://doi.org/10.1093/bib/bbaa233

Tseng YT, Wang SM, Huang KJ, Wang CT (2014): SARS-CoV envelope protein palmitoylation or nucleocapsid association is not required for promoting virus-like particle production. J. Biomed. Sci. 21(1), 34. https://doi.org/10.1186/1423-0127-21-34

Vilar S, Isom DG (2021): One Year of SARS-CoV-2: How much has the virus changed? Biology 10(2), 91. https://doi.org/10.3390/biology10020091

Wootton SK, Rowland RRR, Yoo D (2002): Phosphorylation of the porcine reproductive and respiratory syndrome virus nucleocapsid protein. J. Virol. 76(20), 10569–10576 https://doi.org/10.1128/JVI.76.20.10569-10576.2002

Wu S, Tian C, Liu P, Guo D, Zheng W, Huang X, Zhang Y, Liu L (2021): Effects of SARS-CoV-2 mutations on protein structures and intraviral protein-protein interactions. J. Med. Virol. 93(4), 2132–2140. https://doi.org/10.1002/jmv.26597

Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y (2015): The I-TASSER Suite: Protein structure and function prediction. Nat. Methods, 12, 7–8. https://doi.org/10.1038/nmeth.3213

Zeng W, Liu G, Ma H, Zhao D, Yang Y, Liu M, Mohammed A, Zhao C, Yang Y, Xie J, Ding C, Ma X, Weng J, Gao Y, He H, Jin T (2020): Biochemical characterization of SARS-CoV-2 nucleocapsid protein. Biochem. Biophys. Res. Commun. 527(3), 618–623. https://doi.org/10.1016/j.bbrc.2020.04.136

WHO (2022a): [Online] Website https://www.who.int/emergencies/diseases/novel-coronavirus-2019 [accessed March 29, 2022]

WHO (2022b): [Online] Website https://covid19.who.int/region/euro/country/tr [accessed March 15, 2022]

SUPPLEMENTARY INFORMATION

Structural and functional characterization of SARS-CoV-2 nucleocapsid protein mutations identified in Turkey by using in silico approaches

Betul Akcesme[1], Burcin Erkal[2], Zehra Yaren Donmez[1]

[1]Department of Basic Medical Sciences, Division of Medical Biology, Faculty of Medicine, University of Health Sciences, Istanbul, Turkey; [2]Department of Molecular Biology and Genetics, Faculty of Arts and Sciences, Yıldız Technical University, Istanbul, Turkey

Supplementary Table 1. All missense mutations detected in Turkish population of nucleocapsid protein

MUTATION	GENE LOCATION	FREQUENCY	MUTATION	GENE LOCATION	FREQUENCY
R203K	G28881A, G28882A	1158	P383S	C29420T	14
G204R	G28883C	1117	N196Y	A28859T	13
S235F	C28977T	387	D128N	G28655A	11
D3L	A28281T, T28282A	382	T334I	C29274T	9
S194L	C28854T	164	P67L	C28473T	8
T205I	C28887T	118	S188P	T28835C	8
T198I	C28866T	42	S193T	G28851C	8
T362I	C29358T	41	Q7K	C28292A	7
D3Y	G28280T	33	R195I	G28857T	7
M234I	G28975T	28	M322I	G29239A	7
P13S	C28310T	26	G25C	G28346T	6
G204L	G28883C, G28884T	23	D128Y	G28655T	6
S202N	G28878A	22	Q9H	G28300T	5
D377Y	G29402T	22	P199L	C28869T	5
A220V	C28932T	21	A376T	G29399A	5
D401Y	G29474T	21	P383L	C29421T	5
A211V	C28905T	19	T379I	C29409T	5
A35V	C28377T	17	D22N	G28337A	4
E378Q	G29405C	17	G34W	G28373T	4
P199S	C28868T	16	S2Y	C28278A	4
G238V	G28986T	15	L139F	G28690T	4

E-mail: betul.akcesme@sbu.edu.tr; burcinerkall@gmail.com, zehrayaren@outlook.com phone: +905343064763.

Supplementary Table 1. (continued)

MUTATION	GENE LOCATION	FREQUENCY	MUTATION	GENE LOCATION	FREQUENCY
G212V	G28908T	4	A12G	G28308G	1
G238C	G28985T	4	A12S	G28307T	1
P365S	C29366T	4	D22Y	G28337T	1
N8T	A28296C	3	D3N	G28280A	1
P13L	C28311T	3	D3Q	G28280C, T28282A	1
R40C	C28391T	3	E31G	A28365G	1
A119S	G28628T	3	G19A	G28329C	1
A90S	G28541T	3	N29S	A28359G	1
T135I	C28677T	3	N8Y	A28295T	1
T166I	C28770T	3	Q28K	C82A	1
G214C	G289113T	3	Q7L	A28293T	1
R191L	G28845T	3	Q9R	A28299G	1
I320F	A29231T	3	R10L	G28302T	1
A414S	G29513T	3	R14C	C28313T	1
E378A	A29406C	3	R32H	G28368A	1
G25V	G28347T	2	S21P	T28334C	1
G44C	G28403T	2	S2F	C28278T	1
S23L	C28341T	2	S33G	A28370G	1
S33I	G28371T	2	S37P	T28382C	1
D144H	G28703C	2	S37T	T28382A	1
E136Q	G28679C	2	T16M	C28320T	1
H145Y	C28706T	2	T24A	A28343G	1
L167F	G28774T	2	A152S	G28727T	1
N126S	C28473T	2	A173V	C28791T	1
P80R	C28512G	2	D128E	C28657A	1
M210I	G28903T	2	D63Y	G28460T	1
P207H	C28893A	2	E118Q	G28625C	1
P207S	C28892T	2	M101I	G28576T	1
G215S	G28916A	2	N140S	A28692G	1
A308S	G29195T	2	P142S	C28697T	1
H300Y	C29171T	2	P151A	C28724G	1
M317T	T29223C	2	P151S	C28724T	1
P364S	A29147C	2	P67S	C28472T	1
T271I	C29085T	2	Q163K	C28760A	1
T334A	A1000G	2	T141I	C28695T	1
V324I	G29243A	2	T49I	C28419T	1
V350G	T29322G	2	V72L	G28487C	1
D377A	A29403C	2	A208S	G28895T	1
Q380L	C29411T, C29412T	2	A211G	28906	1
Q389H	G1167C	2	A218V	C28926T	1
S413T	G29511C	2	G204P	G28883C, G28884C	1
S416L	C29520T	2	G204Q	G28883C, G28884A	1
T417I	C29523T, T29524C	2	M210N	28906	1

Supplementary Table 1. (continued)

MUTATION	GENE LOCATION	FREQUENCY	MUTATION	GENE LOCATION	FREQUENCY
N213S	A28911G	1	P364Q	C29364A	1
Q229R	A28959G	1	Q289H	G29140C	1
Q241L	A28995T	1	S327L	C29253T	1
R185C	C28826T	1	T296I	C29160T	1
R185G	C28826G	1	V350A	T29322C	1
R209S	A28900C	1	V350L	G29321C	1
S180C	A28811T	1	K375E	A29396G	1
S201C	A28874T	1	P365L	C29367T	1
S202C	A28877T	1	P365Q	C29367A	1
S206F	C28890T	1	Q389L	A29439T	1
S232N	G28968A	1	Q390H	A29443C	1
V246A	T29010C	1	Q418H	G29527T	1
I292L	A29147C	1	S413I	G29511T	1
I337V	A29282G	1	T366I	C29370T	1
K299N	A29170C	1	T379A	A29408G	1
L339F	G29290T	1	T391I	C29445T	1
P326S	C29249T	1			

Supplementary Table 2. Sequence based analysis of all mutations with PROVEAN and PredictSNP tools

Wild residue	Position	Target residue	PROVEAN	PredictSNP prediction	PredictSNP expected accuracy	MAPP prediction	MAPP expected accuracy	PhD-SNP prediction	PhD-SNP expected accuracy	PolyPhen-1 prediction	PolyPhen-1 expected accuracy	PolyPhen-2 prediction	PolyPhen-2 expected accuracy	SIFT prediction	SIFT expected accuracy	SNAP prediction	SNAP expected accuracy	nsSNPAnalyzer prediction	nsSNPAnalyzer expected accuracy	PANTHER prediction	PANTHER expected accuracy	Annotations
S	2	F	NEUTRAL	DELETERIOUS	0.64351714	DELETERIOUS	0.42653673	NEUTRAL	0.83210379	DELETERIOUS	0.59445019	DELETERIOUS	0.47429306	DELETERIOUS	0.79280784	DELETERIOUS	0.72038776	UNKNOWN	0.0	UNKNOWN	0.0	
S	2	Y	NEUTRAL	DELETERIOUS	0.60548272	NEUTRAL	0.70484581	NEUTRAL	0.71871412	DELETERIOUS	0.74491225	DELETERIOUS	0.47429306	DELETERIOUS	0.79280784	DELETERIOUS	0.72038776	UNKNOWN	0.0	UNKNOWN	0.0	
D	3	Q	NEUTRAL	DELETERIOUS	0.60548272	NEUTRAL	0.64229075	NEUTRAL	0.660879	DELETERIOUS	0.74491225	DELETERIOUS	0.81142888	DELETERIOUS	0.79280784	DELETERIOUS	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	
D	3	N	NEUTRAL	DELETERIOUS	0.54946365	NEUTRAL	0.72246696	NEUTRAL	0.83210379	DELETERIOUS	0.59445019	DELETERIOUS	0.67524116	DELETERIOUS	0.79280784	DELETERIOUS	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	
D	3	Y	NEUTRAL	DELETERIOUS	0.7556615	DELETERIOUS	0.62143928	NEUTRAL	0.660879	DELETERIOUS	0.74491225	DELETERIOUS	0.81142888	DELETERIOUS	0.79280784	DELETERIOUS	0.72038776	UNKNOWN	0.0	UNKNOWN	0.0	
D	3	L	NEUTRAL	DELETERIOUS	0.65494636	DELETERIOUS	0.55997001	NEUTRAL	0.83210379	DELETERIOUS	0.74491225	DELETERIOUS	0.81142888	DELETERIOUS	0.79280784	DELETERIOUS	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	Natural variant in strain: B.1.1.7 (mapped from position 3 in UniProt P0DTC9)
Q	7	K	NEUTRAL	NEUTRAL	0.74796037	DELETERIOUS	0.81934033	NEUTRAL	0.7828765	NEUTRAL	0.66884082	NEUTRAL	0.72272272	NEUTRAL	0.52894034	NEUTRAL	0.500114676	UNKNOWN	0.0	UNKNOWN	0.0	
Q	7	L	NEUTRAL	NEUTRAL	0.6025641	NEUTRAL	0.5089955	NEUTRAL	0.83210379	NEUTRAL	0.66884082	NEUTRAL	0.67635271	DELETERIOUS	0.52778524	DELETERIOUS	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	
N	8	Y	NEUTRAL	DELETERIOUS	0.60548272	DELETERIOUS	0.63268366	NEUTRAL	0.660879	DELETERIOUS	0.59445019	NEUTRAL	0.60960961	DELETERIOUS	0.52778524	DELETERIOUS	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	
N	8	T	NEUTRAL	NEUTRAL	0.73834499	NEUTRAL	0.74273128	NEUTRAL	0.7828765	NEUTRAL	0.66884082	NEUTRAL	0.78928929	DELETERIOUS	0.45233975	NEUTRAL	0.500114676	UNKNOWN	0.0	UNKNOWN	0.0	
Q	9	H	NEUTRAL	DELETERIOUS	0.63151762	DELETERIOUS	0.76086957	NEUTRAL	0.83210379	DELETERIOUS	0.66884082	DELETERIOUS	0.40745501	DELETERIOUS	0.45233975	DELETERIOUS	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	
Q	9	R	NEUTRAL	NEUTRAL	0.6025641	DELETERIOUS	0.81934033	NEUTRAL	0.7828765	NEUTRAL	0.66884082	NEUTRAL	0.69369369	DELETERIOUS	0.42969871	NEUTRAL	0.58359402	UNKNOWN	0.0	UNKNOWN	0.0	
R	10	L	NEUTRAL	NEUTRAL	0.82622462	NEUTRAL	0.65667166	NEUTRAL	0.83210379	NEUTRAL	0.66884082	NEUTRAL	0.64364364	DELETERIOUS	0.52778524	DELETERIOUS	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	
A	12	S	NEUTRAL	NEUTRAL	0.68365861	NEUTRAL	0.74449339	NEUTRAL	0.71871412	NEUTRAL	0.66884082	NEUTRAL	0.67635271	NEUTRAL	0.76224399	NEUTRAL	0.70878378	UNKNOWN	0.0	UNKNOWN	0.0	
A	12	G	NEUTRAL	NEUTRAL	0.60548272	NEUTRAL	0.64229075	NEUTRAL	0.71871412	DELETERIOUS	0.66884082	DELETERIOUS	0.50257069	DELETERIOUS	0.45949821	NEUTRAL	0.76823399	UNKNOWN	0.0	UNKNOWN	0.0	
P	13	S	NEUTRAL	DELETERIOUS	0.71871275	DELETERIOUS	0.73127753	NEUTRAL	0.71871412	DELETERIOUS	0.74491225	DELETERIOUS	0.81142888	DELETERIOUS	0.79280784	DELETERIOUS	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	
P	13	L	NEUTRAL	DELETERIOUS	0.7556615	DELETERIOUS	0.76086957	NEUTRAL	0.68183996	DELETERIOUS	0.74491225	DELETERIOUS	0.81142888	DELETERIOUS	0.79280784	DELETERIOUS	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	
R	14	C	NEUTRAL	DELETERIOUS	0.71871275	DELETERIOUS	0.84182909	NEUTRAL	0.71871412	DELETERIOUS	0.74491225	DELETERIOUS	0.67524116	DELETERIOUS	0.79280784	DELETERIOUS	0.80510276	UNKNOWN	0.0	UNKNOWN	0.0	
T	16	M	NEUTRAL	DELETERIOUS	0.65307311	DELETERIOUS	0.5089955	NEUTRAL	0.83210379	DELETERIOUS	0.74491225	DELETERIOUS	0.67524116	DELETERIOUS	0.79280784	DELETERIOUS	0.80510276	UNKNOWN	0.0	UNKNOWN	0.0	
G	19	A	NEUTRAL	NEUTRAL	0.65494636	DELETERIOUS	0.57314329	NEUTRAL	0.71871412	NEUTRAL	0.66884082	NEUTRAL	0.70240481	DELETERIOUS	0.42969871	NEUTRAL	0.58359402	UNKNOWN	0.0	UNKNOWN	0.0	
S	21	P	NEUTRAL	DELETERIOUS	0.54946365	DELETERIOUS	0.77494374	NEUTRAL	0.7828765	DELETERIOUS	0.59445019	DELETERIOUS	0.46330848	DELETERIOUS	0.52778524	DELETERIOUS	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	
D	22	N	NEUTRAL	DELETERIOUS	0.7556615	NEUTRAL	0.7596831	NEUTRAL	0.660879	DELETERIOUS	0.59445019	DELETERIOUS	0.67524116	DELETERIOUS	0.52778524	DELETERIOUS	0.80510276	UNKNOWN	0.0	UNKNOWN	0.0	
D	22	Y	NEUTRAL	DELETERIOUS	0.6025641	DELETERIOUS	0.63268366	NEUTRAL	0.7828765	DELETERIOUS	0.74491225	DELETERIOUS	0.81142888	DELETERIOUS	0.79280784	DELETERIOUS	0.84845361	UNKNOWN	0.0	UNKNOWN	0.0	
S	23	L	NEUTRAL	NEUTRAL	0.6025641	DELETERIOUS	0.9197901	NEUTRAL	0.7828765	DELETERIOUS	0.66884082	DELETERIOUS	0.67635271	DELETERIOUS	0.52778524	DELETERIOUS	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	
T	24	A	NEUTRAL	NEUTRAL	0.60548272	DELETERIOUS	0.71814093	NEUTRAL	0.7828765	NEUTRAL	0.66884082	NEUTRAL	0.8730899	DELETERIOUS	0.42969871	NEUTRAL	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	
G	25	V	NEUTRAL	DELETERIOUS	0.65494636	DELETERIOUS	0.71814093	NEUTRAL	0.7828765	NEUTRAL	0.59445019	NEUTRAL	0.60960961	DELETERIOUS	0.79280784	DELETERIOUS	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	
G	25	C	NEUTRAL	DELETERIOUS	0.73834499	DELETERIOUS	0.57121439	NEUTRAL	0.7828765	DELETERIOUS	0.74491225	DELETERIOUS	0.54177378	DELETERIOUS	0.79280784	DELETERIOUS	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	
Q	28	K	NEUTRAL	NEUTRAL	0.73834499	NEUTRAL	0.78854626	NEUTRAL	0.83210379	DELETERIOUS	0.66884082	DELETERIOUS	0.39845758	NEUTRAL	0.52894034	NEUTRAL	0.500114676	UNKNOWN	0.0	UNKNOWN	0.0	
N	29	S	NEUTRAL	NEUTRAL	0.73834499	NEUTRAL	0.76123348	NEUTRAL	0.7828765	NEUTRAL	0.66884082	NEUTRAL	0.78928929	NEUTRAL	0.6841637	DELETERIOUS	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	
E	31	G	NEUTRAL	NEUTRAL	0.74825175	DELETERIOUS	0.42653673	NEUTRAL	0.83210379	NEUTRAL	0.66884082	NEUTRAL	0.8730899	NEUTRAL	0.89670526	NEUTRAL	0.58359402	UNKNOWN	0.0	UNKNOWN	0.0	
R	32	H	NEUTRAL	DELETERIOUS	0.71871275	DELETERIOUS	0.76536732	NEUTRAL	0.7828765	DELETERIOUS	0.74491225	DELETERIOUS	0.67524116	DELETERIOUS	0.79280784	DELETERIOUS	0.84845361	UNKNOWN	0.0	UNKNOWN	0.0	
S	33	I	NEUTRAL	DELETERIOUS	0.54946365	DELETERIOUS	0.63268366	NEUTRAL	0.83210379	DELETERIOUS	0.59445019	NEUTRAL	0.60960961	DELETERIOUS	0.45233975	DELETERIOUS	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	
S	33	G	NEUTRAL	NEUTRAL	0.82622462	NEUTRAL	0.85110132	NEUTRAL	0.83210379	NEUTRAL	0.66884082	NEUTRAL	0.74974975	NEUTRAL	0.60819234	NEUTRAL	0.55439739	UNKNOWN	0.0	UNKNOWN	0.0	
G	34	W	NEUTRAL	DELETERIOUS	0.7556615	DELETERIOUS	0.57314329	NEUTRAL	0.660879	DELETERIOUS	0.74491225	DELETERIOUS	0.81142888	DELETERIOUS	0.79280784	DELETERIOUS	0.72038776	UNKNOWN	0.0	UNKNOWN	0.0	

Supplementary Table 2. (continued)

Wild residue	Position	Target residue	PROVEAN	PredictSNP prediction	PredictSNP expected accuracy	MAPP prediction	MAPP expected accuracye	PhD-SNP prediction	PhD-SNP expected accuracy	PolyPhen-1 prediction	PolyPhen-1 expected accuracy	PolyPhen-2 prediction	Poly-Phen-2 expected accuracy	SIFT prediction	SIFT expected accuracy	SNAP prediction	SNAP expected accuracy	nsSNPAnalyzer prediction	nsSNPAnalyzer expected accuracy	PANTHER prediction	PANTHER expected accuracy	Annotations
A	35	V	NEUTRAL	DELETERIOUS	0.63587604	DELETERIOUS	0.46176912	NEUTRAL	0.83210379	DELETERIOUS	0.59445019	DELETERIOUS	0.60008997	DELETERIOUS	0.52778524	DELETERIOUS	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	
S	37	P	NEUTRAL	NEUTRAL	0.82622462	NEUTRAL	0.85110132	NEUTRAL	0.89245476	NEUTRAL	0.66884082	NEUTRAL	0.8730899	NEUTRAL	0.89670526	NEUTRAL	0.83058824	UNKNOWN	0.0	UNKNOWN	0.0	
S	37	T	NEUTRAL	NEUTRAL	0.73688811	DELETERIOUS	0.58770615	NEUTRAL	0.89245476	NEUTRAL	0.66884082	NEUTRAL	0.74074074	NEUTRAL	0.64501779	NEUTRAL	0.55439739	UNKNOWN	0.0	UNKNOWN	0.0	
R	40	C	NEUTRAL	DELETERIOUS	0.65494636	DELETERIOUS	0.40929535	NEUTRAL	0.7828765	DELETERIOUS	0.74491225	DELETERIOUS	0.67524116	DELETERIOUS	0.79280784	DELETERIOUS	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	
G	44	C	NEUTRAL	NEUTRAL	0.6025641	DELETERIOUS	0.57121439	NEUTRAL	0.7828765	DELETERIOUS	0.74491225	NEUTRAL	0.60960961	DELETERIOUS	0.52778524	NEUTRAL	0.50014676	UNKNOWN	0.0	UNKNOWN	0.0	
T	49	I	NEUTRAL	NEUTRAL	0.6025641	DELETERIOUS	0.65667166	NEUTRAL	0.71871412	NEUTRAL	0.66884082	NEUTRAL	0.76252505	DELETERIOUS	0.45232975	DELETERIOUS	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	
D	63	Y	NEUTRAL	DELETERIOUS	0.71871275	DELETERIOUS	0.63268366	NEUTRAL	0.71871412	DELETERIOUS	0.59445019	DELETERIOUS	0.59318766	DELETERIOUS	0.79280784	DELETERIOUS	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	
P	67	S	NEUTRAL	NEUTRAL	0.73834499	NEUTRAL	0.70484581	NEUTRAL	0.7828765	NEUTRAL	0.66884082	DELETERIOUS	0.50257069	NEUTRAL	0.71326803	NEUTRAL	0.61167883	UNKNOWN	0.0	UNKNOWN	0.0	
P	67	L	NEUTRAL	DELETERIOUS	0.50595948	DELETERIOUS	0.57314329	NEUTRAL	0.7828765	NEUTRAL	0.66884082	DELETERIOUS	0.60008997	DELETERIOUS	0.79280784	NEUTRAL	0.50014676	UNKNOWN	0.0	UNKNOWN	0.0	
V	72	L	NEUTRAL	DELETERIOUS	0.71871275	DELETERIOUS	0.76086957	DELETERIOUS	0.67620995	NEUTRAL	0.66884082	DELETERIOUS	0.81142888	DELETERIOUS	0.79280784	DELETERIOUS	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	
P	80	R	NEUTRAL	NEUTRAL	0.63151762	DELETERIOUS	0.65667166	NEUTRAL	0.71871412	NEUTRAL	0.66884082	NEUTRAL	0.7247472	NEUTRAL	0.66429207	DELETERIOUS	0.72038776	UNKNOWN	0.0	UNKNOWN	0.0	
A	90	S	NEUTRAL	NEUTRAL	0.82622462	NEUTRAL	0.65022026	NEUTRAL	0.660879	NEUTRAL	0.66884082	NEUTRAL	0.62762763	NEUTRAL	0.75690116	NEUTRAL	0.50014676	UNKNOWN	0.0	UNKNOWN	0.0	
M	101	I	NEUTRAL	NEUTRAL	0.82622462	NEUTRAL	0.72246696	NEUTRAL	0.58230958	NEUTRAL	0.66884082	NEUTRAL	0.71171171	NEUTRAL	0.66429207	NEUTRAL	0.50014676	UNKNOWN	0.0	UNKNOWN	0.0	
E	118	Q	NEUTRAL	NEUTRAL	0.65307311	NEUTRAL	0.76123348	DELETERIOUS	0.58885542	NEUTRAL	0.66884082	DELETERIOUS	0.54177378	NEUTRAL	0.76224399	NEUTRAL	0.55439739	UNKNOWN	0.0	UNKNOWN	0.0	
A	119	S	NEUTRAL	DELETERIOUS	0.54946365	DELETERIOUS	0.76611694	DELETERIOUS	0.60798122	NEUTRAL	0.66884082	DELETERIOUS	0.40745501	DELETERIOUS	0.42966871	NEUTRAL	0.55439739	UNKNOWN	0.0	UNKNOWN	0.0	
N	126	S	NEUTRAL	NEUTRAL	0.82622462	NEUTRAL	0.74977974	NEUTRAL	0.83210379	NEUTRAL	0.66884082	NEUTRAL	0.7077071	NEUTRAL	0.83526269	NEUTRAL	0.70878378	UNKNOWN	0.0	UNKNOWN	0.0	
D	128	E	NEUTRAL	NEUTRAL	0.82622462	NEUTRAL	0.76123348	NEUTRAL	0.660879	NEUTRAL	0.66884082	NEUTRAL	0.68268268	NEUTRAL	0.89670526	NEUTRAL	0.66524217	UNKNOWN	0.0	UNKNOWN	0.0	
D	128	N	NEUTRAL	NEUTRAL	0.73834499	NEUTRAL	0.70484581	NEUTRAL	0.44670846	NEUTRAL	0.66884082	DELETERIOUS	0.39845758	NEUTRAL	0.66429207	NEUTRAL	0.55439739	UNKNOWN	0.0	UNKNOWN	0.0	
D	128	Y	DELETERIOUS	DELETERIOUS	0.71871275	DELETERIOUS	0.5089955	DELETERIOUS	0.81731169	NEUTRAL	0.66884082	DELETERIOUS	0.54177378	DELETERIOUS	0.79280784	DELETERIOUS	0.80510276	UNKNOWN	0.0	UNKNOWN	0.0	
T	135	I	NEUTRAL	NEUTRAL	0.61421911	DELETERIOUS	0.62143928	NEUTRAL	0.83210379	NEUTRAL	0.66884082	DELETERIOUS	0.39845758	DELETERIOUS	0.45949821	NEUTRAL	0.70878378	UNKNOWN	0.0	UNKNOWN	0.0	
E	136	Q	NEUTRAL	NEUTRAL	0.82622462	NEUTRAL	0.74273128	NEUTRAL	0.71871412	NEUTRAL	0.66884082	NEUTRAL	0.64364364	NEUTRAL	0.69991095	NEUTRAL	0.70878378	UNKNOWN	0.0	UNKNOWN	0.0	
L	139	F	NEUTRAL	NEUTRAL	0.63151762	NEUTRAL	0.64229075	NEUTRAL	0.7828765	DELETERIOUS	0.59445019	DELETERIOUS	0.56233933	NEUTRAL	0.64501779	NEUTRAL	0.61167883	UNKNOWN	0.0	UNKNOWN	0.0	
N	140	S	NEUTRAL	NEUTRAL	0.82622462	NEUTRAL	0.64229075	NEUTRAL	0.68183996	NEUTRAL	0.66884082	NEUTRAL	0.69369369	DELETERIOUS	0.76135352	NEUTRAL	0.58359402	UNKNOWN	0.0	UNKNOWN	0.0	
T	141	I	NEUTRAL	NEUTRAL	0.6025641	DELETERIOUS	0.55997001	NEUTRAL	0.7828765	DELETERIOUS	0.66884082	DELETERIOUS	0.50257069	NEUTRAL	0.52778524	NEUTRAL	0.55439739	UNKNOWN	0.0	UNKNOWN	0.0	
P	142	S	NEUTRAL	NEUTRAL	0.73834499	NEUTRAL	0.72246696	NEUTRAL	0.83210379	NEUTRAL	0.66884082	NEUTRAL	0.64364364	DELETERIOUS	0.42966871	NEUTRAL	0.58359402	UNKNOWN	0.0	UNKNOWN	0.0	
D	144	H	NEUTRAL	DELETERIOUS	0.65494636	DELETERIOUS	0.77494374	NEUTRAL	0.83210379	DELETERIOUS	0.74491225	DELETERIOUS	0.67524116	DELETERIOUS	0.45949821	DELETERIOUS	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	
H	145	Y	NEUTRAL	NEUTRAL	0.73688811	DELETERIOUS	0.70484581	NEUTRAL	0.68183996	NEUTRAL	0.66884082	NEUTRAL	0.62762763	DELETERIOUS	0.42966871	NEUTRAL	0.50014676	UNKNOWN	0.0	UNKNOWN	0.0	
P	151	S	DELETERIOUS	DELETERIOUS	0.71871275	DELETERIOUS	0.85832084	NEUTRAL	0.660879	DELETERIOUS	0.59445019	DELETERIOUS	0.64974293	DELETERIOUS	0.79280784	DELETERIOUS	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	
P	151	A	DELETERIOUS	DELETERIOUS	0.60548272	DELETERIOUS	0.77494374	NEUTRAL	0.660879	NEUTRAL	0.66884082	DELETERIOUS	0.39845758	DELETERIOUS	0.79280784	DELETERIOUS	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	
A	152	S	NEUTRAL	NEUTRAL	0.82622462	NEUTRAL	0.7596631	NEUTRAL	0.89245476	NEUTRAL	0.66884082	NEUTRAL	0.64364364	NEUTRAL	0.89670526	NEUTRAL	0.70878378	UNKNOWN	0.0	UNKNOWN	0.0	
Q	163	K	NEUTRAL	NEUTRAL	0.73834499	DELETERIOUS	0.46176912	NEUTRAL	0.83210379	NEUTRAL	0.66884082	NEUTRAL	0.64364364	NEUTRAL	0.60819234	NEUTRAL	0.58359402	UNKNOWN	0.0	UNKNOWN	0.0	
T	166	I	NEUTRAL	NEUTRAL	0.82622462	NEUTRAL	0.70484581	NEUTRAL	0.83210379	NEUTRAL	0.66884082	NEUTRAL	0.67635271	NEUTRAL	0.79073909	NEUTRAL	0.66524217	UNKNOWN	0.0	UNKNOWN	0.0	

Supplementary Table 2. (continued)

Wild residue	Position	Target residue	PROVEAN	PredictSNP prediction	PredictSNP expected accuracy	MAPP prediction	MAPP expected accuracy	PhD-SNP prediction	PhD-SNP expected accuracy	PolyPhen-1 prediction	PolyPhen-1 expected accuracy	PolyPhen-2 prediction	PolyPhen-2 expected accuracy	SIFT prediction	SIFT expected accuracy	SNAP prediction	SNAP expected accuracy	nsSNPAnalyzer prediction	nsSNPAnalyzer expected accuracy	PANTHER prediction	PANTHER expected accuracy	Annotations
L	167	F	NEUTRAL	DELETERIOUS	0.54946365	DELETERIOUS	0.63268366	NEUTRAL	0.83210379	DELETERIOUS	0.74491225	DELETERIOUS	0.59318766	DELETERIOUS	0.42969871	NEUTRAL	0.61167883	UNKNOWN	0.0	UNKNOWN	0.0	
A	173	V	NEUTRAL	NEUTRAL	0.82622462	NEUTRAL	0.70484581	NEUTRAL	0.83210379	NEUTRAL	0.66894082	NEUTRAL	0.63363363	NEUTRAL	0.89670526	NEUTRAL	0.76823399	UNKNOWN	0.0	UNKNOWN	0.0	
S	180	C	DELETERIOUS	DELETERIOUS	0.71871275	DELETERIOUS	0.76536732	NEUTRAL	0.71871412	DELETERIOUS	0.74491225	DELETERIOUS	0.64974293	DELETERIOUS	0.52778524	DELETERIOUS	0.80510276	UNKNOWN	0.0	UNKNOWN	0.0	
R	185	G	DELETERIOUS	NEUTRAL	0.63151762	NEUTRAL	0.73480176	NEUTRAL	0.71871412	NEUTRAL	0.66894082	DELETERIOUS	0.54177378	NEUTRAL	0.67141585	DELETERIOUS	0.80510276	UNKNOWN	0.0	UNKNOWN	0.0	
R	185	C	DELETERIOUS	DELETERIOUS	0.71871275	DELETERIOUS	0.55997001	NEUTRAL	0.660879	DELETERIOUS	0.74491225	DELETERIOUS	0.81142888	DELETERIOUS	0.46949821	DELETERIOUS	0.88519637	UNKNOWN	0.0	UNKNOWN	0.0	
S	188	P	DELETERIOUS	DELETERIOUS	0.64351714	DELETERIOUS	0.5089955	NEUTRAL	0.89245476	DELETERIOUS	0.59445019	DELETERIOUS	0.64974293	DELETERIOUS	0.52778524	DELETERIOUS	0.80510276	UNKNOWN	0.0	UNKNOWN	0.0	
R	191	L	DELETERIOUS	DELETERIOUS	0.65494636	DELETERIOUS	0.84182909	NEUTRAL	0.83210379	DELETERIOUS	0.59445019	DELETERIOUS	0.6417224	DELETERIOUS	0.79280784	DELETERIOUS	0.80510276	UNKNOWN	0.0	UNKNOWN	0.0	
S	193	T	NEUTRAL	NEUTRAL	0.73834499	NEUTRAL	0.64229075	NEUTRAL	0.83210379	NEUTRAL	0.66894082	NEUTRAL	0.7027027	DELETERIOUS	0.42969871	NEUTRAL	0.61167883	UNKNOWN	0.0	UNKNOWN	0.0	
S	194	L	DELETERIOUS	DELETERIOUS	0.65494636	DELETERIOUS	0.7711928	NEUTRAL	0.7828765	DELETERIOUS	0.59445019	DELETERIOUS	0.60089974	DELETERIOUS	0.79280784	DELETERIOUS	0.72038776	UNKNOWN	0.0	UNKNOWN	0.0	
R	195	I	DELETERIOUS	DELETERIOUS	0.60548272	DELETERIOUS	0.81934033	NEUTRAL	0.68183996	DELETERIOUS	0.74491225	NEUTRAL	0.60960961	DELETERIOUS	0.52778524	DELETERIOUS	0.80510276	UNKNOWN	0.0	UNKNOWN	0.0	
N	196	Y	DELETERIOUS	DELETERIOUS	0.65494636	DELETERIOUS	0.65667166	NEUTRAL	0.71871412	DELETERIOUS	0.74491225	DELETERIOUS	0.39845758	DELETERIOUS	0.52778524	DELETERIOUS	0.72038776	UNKNOWN	0.0	UNKNOWN	0.0	
T	198	I	NEUTRAL	DELETERIOUS	0.71871275	DELETERIOUS	0.76086957	NEUTRAL	0.71871412	DELETERIOUS	0.59445019	DELETERIOUS	0.64974293	DELETERIOUS	0.79280784	DELETERIOUS	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	
P	199	S	NEUTRAL	NEUTRAL	0.65307311	NEUTRAL	0.73127753	NEUTRAL	0.68183996	DELETERIOUS	0.59445019	DELETERIOUS	0.56233933	NEUTRAL	0.76135352	NEUTRAL	0.70878378	UNKNOWN	0.0	UNKNOWN	0.0	
P	199	L	NEUTRAL	NEUTRAL	0.63151762	NEUTRAL	0.6778169	NEUTRAL	0.7828765	DELETERIOUS	0.59445019	DELETERIOUS	0.56233933	NEUTRAL	0.68210151	NEUTRAL	0.55439739	UNKNOWN	0.0	UNKNOWN	0.0	
S	201	C	DELETERIOUS	DELETERIOUS	0.65494636	DELETERIOUS	0.42653673	NEUTRAL	0.7828765	DELETERIOUS	0.74491225	DELETERIOUS	0.81142888	DELETERIOUS	0.52778524	DELETERIOUS	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	
S	202	C	DELETERIOUS	DELETERIOUS	0.71871275	DELETERIOUS	0.57121439	NEUTRAL	0.68183996	DELETERIOUS	0.74491225	DELETERIOUS	0.81142888	DELETERIOUS	0.79280784	DELETERIOUS	0.80510276	UNKNOWN	0.0	UNKNOWN	0.0	
S	202	N	NEUTRAL	DELETERIOUS	0.54946365	DELETERIOUS	0.55997001	NEUTRAL	0.83210379	DELETERIOUS	0.59445019	DELETERIOUS	0.6417224	NEUTRAL	0.64501779	DELETERIOUS	0.80510276	UNKNOWN	0.0	UNKNOWN	0.0	
R	203	K	NEUTRAL	NEUTRAL	0.63151762	NEUTRAL	0.78854626	NEUTRAL	0.7828765	DELETERIOUS	0.66894082	NEUTRAL	0.60089974	NEUTRAL	0.52894034	DELETERIOUS	0.72038776	UNKNOWN	0.0	UNKNOWN	0.0	
G	204	R	NEUTRAL	DELETERIOUS	0.60548272	NEUTRAL	0.73480176	NEUTRAL	0.83210379	DELETERIOUS	0.74491225	DELETERIOUS	0.81142888	DELETERIOUS	0.79280784	DELETERIOUS	0.72038776	UNKNOWN	0.0	UNKNOWN	0.0	
G	204	Q	NEUTRAL	NEUTRAL	0.62831343	NEUTRAL	0.74823944	NEUTRAL	0.71871412	DELETERIOUS	0.74491225	DELETERIOUS	0.81142888	NEUTRAL	0.7366548	NEUTRAL	0.55439739	UNKNOWN	0.0	UNKNOWN	0.0	
G	204	P	NEUTRAL	DELETERIOUS	0.71871275	DELETERIOUS	0.65667166	NEUTRAL	0.71871412	DELETERIOUS	0.74491225	DELETERIOUS	0.81142888	DELETERIOUS	0.79280784	DELETERIOUS	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	
G	204	L	DELETERIOUS	DELETERIOUS	0.65494636	DELETERIOUS	0.48350825	NEUTRAL	0.7828765	DELETERIOUS	0.74491225	DELETERIOUS	0.81142888	DELETERIOUS	0.79280784	DELETERIOUS	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	
T	205	I	NEUTRAL	NEUTRAL	0.73688811	DELETERIOUS	0.57121439	NEUTRAL	0.83210379	NEUTRAL	0.66894082	NEUTRAL	0.8730899	NEUTRAL	0.67141585	NEUTRAL	0.50014676	UNKNOWN	0.0	UNKNOWN	0.0	
S	206	F	DELETERIOUS	DELETERIOUS	0.71871275	DELETERIOUS	0.65667166	NEUTRAL	0.660879	DELETERIOUS	0.74491225	DELETERIOUS	0.67524116	DELETERIOUS	0.79280784	DELETERIOUS	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	
P	207	S	NEUTRAL	DELETERIOUS	0.54946365	NEUTRAL	0.74977974	NEUTRAL	0.7828765	NEUTRAL	0.66894082	DELETERIOUS	0.81142888	DELETERIOUS	0.42969871	DELETERIOUS	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	
P	207	H	NEUTRAL	DELETERIOUS	0.65494636	DELETERIOUS	0.76086957	NEUTRAL	0.89245476	DELETERIOUS	0.74491225	DELETERIOUS	0.81142888	DELETERIOUS	0.52778524	DELETERIOUS	0.72038776	UNKNOWN	0.0	UNKNOWN	0.0	
A	208	S	NEUTRAL	NEUTRAL	0.68365861	DELETERIOUS	0.42653673	NEUTRAL	0.68183996	NEUTRAL	0.66894082	DELETERIOUS	0.55077121	NEUTRAL	0.75800712	NEUTRAL	0.70878378	UNKNOWN	0.0	UNKNOWN	0.0	
R	209	S	NEUTRAL	NEUTRAL	0.63151762	DELETERIOUS	0.63348018	NEUTRAL	0.68183996	DELETERIOUS	0.59445019	DELETERIOUS	0.6417224	NEUTRAL	0.76936676	NEUTRAL	0.50014676	UNKNOWN	0.0	UNKNOWN	0.0	
M	210	I	NEUTRAL	NEUTRAL	0.82622462	NEUTRAL	0.64229075	NEUTRAL	0.83210379	NEUTRAL	0.66894082	NEUTRAL	0.78928929	NEUTRAL	0.75800712	NEUTRAL	0.66524217	UNKNOWN	0.0	UNKNOWN	0.0	

Supplementary Table 2. (continued)

Wild residue	Position	Target residue	PROVEAN	PredictSNP prediction	PredictSNP expected accuracy	MAPP prediction	MAPP expected accuracye	PhD-SNP prediction	PhD-SNP expected accuracy	PolyPhen-1 prediction	PolyPhen-1 expected accuracy	PolyPhen-2 prediction	PolyPhen-2 expected accuracy	SIFT prediction	SIFT expected accuracy	SNAP prediction	SNAP expected accuracy	nsSNPAnalyzer prediction	nsSNPAnalyzer expected accuracy	PANTHER prediction	PANTHER expected accuracy	Annotations
M	210	N	NEUTRAL	NEUTRAL	0.82622462	NEUTRAL	0.77004405	NEUTRAL	0.83210379	NEUTRAL	0.66884082	NEUTRAL	0.8730899	NEUTRAL	0.7747106	NEUTRAL	0.55439739	UNKNOWN	0.0	UNKNOWN	0.0	
A	211	V	NEUTRAL	NEUTRAL	0.73834499	NEUTRAL	0.64229075	NEUTRAL	0.7828765	DELETERIOUS	0.66884082	DELETERIOUS	0.47429306	NEUTRAL	0.75690116	NEUTRAL	0.66524217	UNKNOWN	0.0	UNKNOWN	0.0	
A	211	G	NEUTRAL	NEUTRAL	0.73688811	NEUTRAL	0.72246696	NEUTRAL	0.68183996	NEUTRAL	0.66884082	DELETERIOUS	0.60089974	NEUTRAL	0.72662511	NEUTRAL	0.76823399	UNKNOWN	0.0	UNKNOWN	0.0	
G	212	V	NEUTRAL	NEUTRAL	0.75291375	DELETERIOUS	0.85832084	NEUTRAL	0.7828765	NEUTRAL	0.66884082	NEUTRAL	0.78928929	NEUTRAL	0.75690116	NEUTRAL	0.70878378	UNKNOWN	0.0	UNKNOWN	0.0	
N	213	S	NEUTRAL	NEUTRAL	0.82622462	NEUTRAL	0.85110132	NEUTRAL	0.83210379	NEUTRAL	0.66884082	NEUTRAL	0.78928929	NEUTRAL	0.82295374	NEUTRAL	0.66524217	UNKNOWN	0.0	UNKNOWN	0.0	
G	214	C	NEUTRAL	DELETERIOUS	0.54946365	DELETERIOUS	0.5089955	NEUTRAL	0.7828765	DELETERIOUS	0.74491225	DELETERIOUS	0.55077121	DELETERIOUS	0.45232975	NEUTRAL	0.50014676	UNKNOWN	0.0	UNKNOWN	0.0	
G	215	S	NEUTRAL	NEUTRAL	0.82622462	NEUTRAL	0.85110132	NEUTRAL	0.83210379	NEUTRAL	0.66884082	NEUTRAL	0.60960961	NEUTRAL	0.76135352	NEUTRAL	0.55439739	UNKNOWN	0.0	UNKNOWN	0.0	
A	218	V	NEUTRAL	NEUTRAL	0.6530731	DELETERIOUS	0.46176912	NEUTRAL	0.83210379	NEUTRAL	0.66884082	DELETERIOUS	0.56233933	NEUTRAL	0.72662511	NEUTRAL	0.66524217	UNKNOWN	0.0	UNKNOWN	0.0	
A	220	V	NEUTRAL	NEUTRAL	0.6025641	DELETERIOUS	0.85832084	NEUTRAL	0.83210379	NEUTRAL	0.66884082	DELETERIOUS	0.54177378	DELETERIOUS	0.45949821	NEUTRAL	0.58359402	UNKNOWN	0.0	UNKNOWN	0.0	
Q	229	R	NEUTRAL	NEUTRAL	0.75291375	DELETERIOUS	0.8065967	NEUTRAL	0.83210379	NEUTRAL	0.66884082	NEUTRAL	0.8730899	NEUTRAL	0.75690116	NEUTRAL	0.70878378	UNKNOWN	0.0	UNKNOWN	0.0	
S	232	N	NEUTRAL	DELETERIOUS	0.54946365	DELETERIOUS	0.63268366	NEUTRAL	0.89245476	DELETERIOUS	0.59445019	DELETERIOUS	0.60089974	DELETERIOUS	0.45232975	NEUTRAL	0.50014676	UNKNOWN	0.0	UNKNOWN	0.0	
M	234	I	NEUTRAL	NEUTRAL	0.82622462	NEUTRAL	0.76563877	NEUTRAL	0.89245476	NEUTRAL	0.66884082	NEUTRAL	0.8730899	NEUTRAL	0.76936776	NEUTRAL	0.50014676	UNKNOWN	0.0	UNKNOWN	0.0	
S	235	F	NEUTRAL	DELETERIOUS	0.54946365	NEUTRAL	0.74823944	NEUTRAL	0.71871412	DELETERIOUS	0.74491225	DELETERIOUS	0.64974293	DELETERIOUS	0.52778524	DELETERIOUS	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	Natural variant: in strain: B.1.1.7 (mapped from position 235 in UniProt P0DTC9)
G	238	V	NEUTRAL	NEUTRAL	0.62831343			NEUTRAL	0.71871412	DELETERIOUS	0.59445019	NEUTRAL	0.67635271	DELETERIOUS	0.45949821	NEUTRAL	0.58359402	UNKNOWN	0.0	UNKNOWN	0.0	
G	238	C	NEUTRAL	DELETERIOUS	0.62127533	DELETERIOUS	0.5089955	NEUTRAL	0.71871412	DELETERIOUS	0.74491225	DELETERIOUS	0.45308483	DELETERIOUS	0.52778524	DELETERIOUS	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	
Q	241	L	NEUTRAL	DELETERIOUS	0.50595948	DELETERIOUS	0.85682159	NEUTRAL	0.89245476	NEUTRAL	0.66884082	DELETERIOUS	0.47429306	DELETERIOUS	0.52778524	DELETERIOUS	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	
V	246	A	NEUTRAL	DELETERIOUS	0.50595948	DELETERIOUS	0.57121439	NEUTRAL	0.83210379	NEUTRAL	0.66884082	DELETERIOUS	0.54177378	DELETERIOUS	0.79280784	NEUTRAL	0.55439739	UNKNOWN	0.0	UNKNOWN	0.0	
T	271	I	DELETERIOUS	DELETERIOUS	0.52145411	DELETERIOUS	0.76086957	NEUTRAL	0.68183996	DELETERIOUS	0.74491225	DELETERIOUS	0.59318766	DELETERIOUS	0.45232975	DELETERIOUS	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	
Q	289	H	NEUTRAL	DELETERIOUS	0.50595948	DELETERIOUS	0.79559471	NEUTRAL	0.83210379	DELETERIOUS	0.74491225	NEUTRAL	0.60089974	NEUTRAL	0.70614426	NEUTRAL	0.50014676	UNKNOWN	0.0	UNKNOWN	0.0	
I	292	L	NEUTRAL	NEUTRAL	0.82622462	NEUTRAL	0.76611694	NEUTRAL	0.83210379	NEUTRAL	0.66884082	NEUTRAL	0.8730899	NEUTRAL	0.89670526	NEUTRAL	0.76823399	UNKNOWN	0.0	UNKNOWN	0.0	
T	296	I	NEUTRAL	NEUTRAL	0.75291375	DELETERIOUS	0.72246696	NEUTRAL	0.58230958	NEUTRAL	0.66884082	DELETERIOUS	0.78928929	NEUTRAL	0.66429207	NEUTRAL	0.66524217	UNKNOWN	0.0	UNKNOWN	0.0	
K	299	N	NEUTRAL	NEUTRAL	0.6025641	NEUTRAL	0.63268366	NEUTRAL	0.660879	NEUTRAL	0.66884082	DELETERIOUS	0.60089974	DELETERIOUS	0.45949821	DELETERIOUS	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	
H	300	Y	NEUTRAL	NEUTRAL	0.6025641	DELETERIOUS	0.81934033	NEUTRAL	0.55202703	NEUTRAL	0.66884082	NEUTRAL	0.8730899	DELETERIOUS	0.42969871	DELETERIOUS	0.72038776	UNKNOWN	0.0	UNKNOWN	0.0	
A	308	S	NEUTRAL	DELETERIOUS	0.71871275	DELETERIOUS	0.63268366	NEUTRAL	0.660879	DELETERIOUS	0.59445019	DELETERIOUS	0.54177378	DELETERIOUS	0.79280784	DELETERIOUS	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	
M	317	T	NEUTRAL	NEUTRAL	0.73688811	NEUTRAL	0.81934033	NEUTRAL	0.68183996	NEUTRAL	0.66884082	NEUTRAL	0.63363363	NEUTRAL	0.75690116	NEUTRAL	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	
I	320	F	NEUTRAL	NEUTRAL	0.73834499	NEUTRAL	0.65022026	NEUTRAL	0.71871412	DELETERIOUS	0.66884082	DELETERIOUS	0.39845758	NEUTRAL	0.76936776	NEUTRAL	0.70878378	UNKNOWN	0.0	UNKNOWN	0.0	
M	322	I	NEUTRAL	NEUTRAL	0.82622462	DELETERIOUS	0.69903084	NEUTRAL	0.7828765	NEUTRAL	0.66884082	NEUTRAL	0.67635271	NEUTRAL	0.67141585	NEUTRAL	0.50014676	UNKNOWN	0.0	UNKNOWN	0.0	
V	324	I	NEUTRAL	NEUTRAL	0.78834499	NEUTRAL	0.42653673	NEUTRAL	0.71871412	NEUTRAL	0.66884082	NEUTRAL	0.63363363	NEUTRAL	0.75800712	NEUTRAL	0.61167883	UNKNOWN	0.0	UNKNOWN	0.0	
P	326	S	NEUTRAL	NEUTRAL	0.82622462	DELETERIOUS	0.73480176	NEUTRAL	0.7828765	NEUTRAL	0.66884082	NEUTRAL	0.8730899	NEUTRAL	0.69991095	NEUTRAL	0.66524217	UNKNOWN	0.0	UNKNOWN	0.0	
S	327	L	DELETERIOUS	NEUTRAL	0.6025641	DELETERIOUS	0.76536732	NEUTRAL	0.71871412	NEUTRAL	0.66884082	NEUTRAL	0.78928929	DELETERIOUS	0.45232975	DELETERIOUS	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	
T	334	I	DELETERIOUS	DELETERIOUS	0.54946365	DELETERIOUS	0.57121439	NEUTRAL	0.58230958	DELETERIOUS	0.59445019	NEUTRAL	0.60960961	DELETERIOUS	0.42969871	DELETERIOUS	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	
T	334	A	NEUTRAL	NEUTRAL	0.63151762	DELETERIOUS	0.58889722	NEUTRAL	0.7828765	NEUTRAL	0.66884082	NEUTRAL	0.68268268	NEUTRAL	0.70614426	DELETERIOUS	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	

Supplementary Table 2. (continued)

Wild residue	Position	Target residue	PROVEAN	PredictSNP prediction	PredictSNP expected accuracy	MAPP prediction	MAPP expected accuracy	PhD-SNP prediction	PhD-SNP expected accuracy	PolyPhen-1 prediction	PolyPhen-1 expected accuracy	PolyPhen-2 prediction	Poly-Phen-2 expected accuracy	SIFT prediction	SIFT expected accuracy	SNAP prediction	SNAP expected accuracy	nsSNPAnalyzer prediction	nsSNPAnalyzer expected accuracy	PANTHER prediction	PANTHER expected accuracy	Annotations
I	337	V	NEUTRAL	NEUTRAL	0.73688811	DELETERIOUS	0.55997001	NEUTRAL	0.83210379	NEUTRAL	0.66884082	NEUTRAL	0.63363363	NEUTRAL	0.70614426	NEUTRAL	0.55439739	UNKNOWN	0.0	UNKNOWN	0.0	
L	339	F	NEUTRAL	NEUTRAL	0.73834499	NEUTRAL	0.71742958	NEUTRAL	0.89245476	NEUTRAL	0.66884082	DELETERIOUS	0.50257069	NEUTRAL	0.81300089	NEUTRAL	0.61167883	UNKNOWN	0.0	UNKNOWN	0.0	
V	350	G	NEUTRAL	DELETERIOUS	0.86908365	DELETERIOUS	0.5089955	DELETERIOUS	0.60798122	DELETERIOUS	0.59445019	DELETERIOUS	0.43123393	DELETERIOUS	0.79280784	DELETERIOUS	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	
V	350	A	NEUTRAL	NEUTRAL	0.65307311	DELETERIOUS	0.5089955	NEUTRAL	0.68183996	NEUTRAL	0.66884082	NEUTRAL	0.70240481	NEUTRAL	0.71326803	DELETERIOUS	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	
V	350	L	NEUTRAL	NEUTRAL	0.82622462	NEUTRAL	0.64229075	NEUTRAL	0.71871412	NEUTRAL	0.66884082	NEUTRAL	0.8730899	NEUTRAL	0.72662511	NEUTRAL	0.70878378	UNKNOWN	0.0	UNKNOWN	0.0	
T	362	I	NEUTRAL	NEUTRAL	0.68365861	NEUTRAL	0.77004405	NEUTRAL	0.660879	NEUTRAL	0.66884082	NEUTRAL	0.789288929	DELETERIOUS	0.45232975	DELETERIOUS	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	
P	364	S	NEUTRAL	DELETERIOUS	0.60548272	DELETERIOUS	0.63268366	NEUTRAL	0.71871412	DELETERIOUS	0.59445019	DELETERIOUS	0.649974293	NEUTRAL	0.52894034	DELETERIOUS	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	
P	364	Q	NEUTRAL	DELETERIOUS	0.7556615	DELETERIOUS	0.766611694	NEUTRAL	0.660879	DELETERIOUS	0.74491225	DELETERIOUS	0.81142888	DELETERIOUS	0.79280784	DELETERIOUS	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	
P	365	S	NEUTRAL	NEUTRAL	0.63151762	NEUTRAL	0.6778169	NEUTRAL	0.83210379	DELETERIOUS	0.59445019	DELETERIOUS	0.67524116	NEUTRAL	0.70614426	NEUTRAL	0.66524217	UNKNOWN	0.0	UNKNOWN	0.0	
P	365	Q	NEUTRAL	DELETERIOUS	0.62831343	NEUTRAL	0.71742958	NEUTRAL	0.7828765	DELETERIOUS	0.74491225	DELETERIOUS	0.81142888	NEUTRAL	0.71326803	NEUTRAL	0.66524217	UNKNOWN	0.0	UNKNOWN	0.0	
P	365	L	NEUTRAL	DELETERIOUS	0.60548272	DELETERIOUS	0.63268366	NEUTRAL	0.7828765	DELETERIOUS	0.74491225	DELETERIOUS	0.81142888	DELETERIOUS	0.52778524	NEUTRAL	0.61167883	UNKNOWN	0.0	UNKNOWN	0.0	
T	366	I	NEUTRAL	DELETERIOUS	0.54946365	DELETERIOUS	0.58770615	NEUTRAL	0.7828765	DELETERIOUS	0.59445019	DELETERIOUS	0.67524116	DELETERIOUS	0.52778524	NEUTRAL	0.61167883	UNKNOWN	0.0	UNKNOWN	0.0	
K	375	E	NEUTRAL	NEUTRAL	0.75203963	NEUTRAL	0.65022026	NEUTRAL	0.83210379	NEUTRAL	0.66884082	DELETERIOUS	0.60089974	NEUTRAL	0.64501779	NEUTRAL	0.500114676	UNKNOWN	0.0	UNKNOWN	0.0	
A	376	T	NEUTRAL	NEUTRAL	0.82622462	NEUTRAL	0.65286344	NEUTRAL	0.89245476	NEUTRAL	0.66884082	NEUTRAL	0.789288929	NEUTRAL	0.89670526	NEUTRAL	0.61167883	UNKNOWN	0.0	UNKNOWN	0.0	
D	377	A	NEUTRAL	DELETERIOUS	0.50595948	NEUTRAL	0.71742958	NEUTRAL	0.97902098	DELETERIOUS	0.59445019	DELETERIOUS	0.67524116	DELETERIOUS	0.52778524	DELETERIOUS	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	
D	377	Y	NEUTRAL	DELETERIOUS	0.65494636	DELETERIOUS	0.40929535	NEUTRAL	0.89245476	DELETERIOUS	0.74491225	DELETERIOUS	0.81142888	DELETERIOUS	0.79280784	DELETERIOUS	0.72038776	UNKNOWN	0.0	UNKNOWN	0.0	
E	378	A	NEUTRAL	NEUTRAL	0.65307311	NEUTRAL	0.7596831	NEUTRAL	0.89245476	NEUTRAL	0.66884082	DELETERIOUS	0.60089974	NEUTRAL	0.42969871	NEUTRAL	0.61167883	UNKNOWN	0.0	UNKNOWN	0.0	
E	378	Q	NEUTRAL	NEUTRAL	0.6025641	NEUTRAL	0.765663877	NEUTRAL	0.83210379	DELETERIOUS	0.59445019	DELETERIOUS	0.64717224	NEUTRAL	0.42969871	NEUTRAL	0.58359402	UNKNOWN	0.0	UNKNOWN	0.0	
T	379	I	NEUTRAL	NEUTRAL	0.82622462	NEUTRAL	0.63348018	NEUTRAL	0.83210379	NEUTRAL	0.66884082	NEUTRAL	0.8730899	NEUTRAL	0.75690116	NEUTRAL	0.58359402	UNKNOWN	0.0	UNKNOWN	0.0	
T	379	A	NEUTRAL	NEUTRAL	0.82622462	NEUTRAL	0.9137931	NEUTRAL	0.89245476	NEUTRAL	0.66884082	NEUTRAL	0.8730899	NEUTRAL	0.89670526	NEUTRAL	0.76823399	UNKNOWN	0.0	UNKNOWN	0.0	
Q	380	L	NEUTRAL	NEUTRAL	0.62813343	DELETERIOUS	0.9137931	NEUTRAL	0.89245476	DELETERIOUS	0.66884082	DELETERIOUS	0.55077121	NEUTRAL	0.71326803	DELETERIOUS	0.58359402	UNKNOWN	0.0	UNKNOWN	0.0	
P	383	S	NEUTRAL	NEUTRAL	0.6025641	NEUTRAL	0.64229075	NEUTRAL	0.83210379	NEUTRAL	0.59445019	DELETERIOUS	0.40745501	DELETERIOUS	0.79280784	NEUTRAL	0.66524217	UNKNOWN	0.0	UNKNOWN	0.0	
P	383	L	NEUTRAL	NEUTRAL	0.68365861	DELETERIOUS	0.48350825	NEUTRAL	0.7828765	NEUTRAL	0.66884082	NEUTRAL	0.74974975	DELETERIOUS	0.42969871	NEUTRAL	0.61167883	UNKNOWN	0.0	UNKNOWN	0.0	
Q	389	H	NEUTRAL	DELETERIOUS	0.50059548	NEUTRAL	0.64229075	NEUTRAL	0.83210379	DELETERIOUS	0.74491225	DELETERIOUS	0.63431877	DELETERIOUS	0.79280784	DELETERIOUS	0.500114676	UNKNOWN	0.0	UNKNOWN	0.0	
Q	389	L	NEUTRAL	NEUTRAL	0.6025641	DELETERIOUS	0.7711928	NEUTRAL	0.89245476	DELETERIOUS	0.66884082	DELETERIOUS	0.55077121	DELETERIOUS	0.79280784	NEUTRAL	0.61167883	UNKNOWN	0.0	UNKNOWN	0.0	
Q	390	H	NEUTRAL	NEUTRAL	0.75291375	DELETERIOUS	0.760861957	NEUTRAL	0.83210379	NEUTRAL	0.66884082	NEUTRAL	0.682682868	NEUTRAL	0.67141585	NEUTRAL	0.500114676	UNKNOWN	0.0	UNKNOWN	0.0	
T	391	I	NEUTRAL	NEUTRAL	0.6025641	NEUTRAL	0.78325991	NEUTRAL	0.83210379	DELETERIOUS	0.59445019	DELETERIOUS	0.67524116	NEUTRAL	0.45949821	NEUTRAL	0.55439739	UNKNOWN	0.0	UNKNOWN	0.0	
D	401	Y	NEUTRAL	DELETERIOUS	0.65494636	DELETERIOUS	0.57121639	NEUTRAL	0.71871412	DELETERIOUS	0.74491225	DELETERIOUS	0.54177378	DELETERIOUS	0.79280784	DELETERIOUS	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	
S	413	I	NEUTRAL	DELETERIOUS	0.521145411				0.71871412	DELETERIOUS	0.59445019	DELETERIOUS	0.649974293	DELETERIOUS	0.79280784	NEUTRAL	0.58359402	UNKNOWN	0.0	UNKNOWN	0.0	
S	413	T	NEUTRAL	NEUTRAL	0.61421911				0.83210379	NEUTRAL	0.66884082	NEUTRAL	0.59318766	NEUTRAL	0.79280784	NEUTRAL	0.66524217	UNKNOWN	0.0	UNKNOWN	0.0	
A	414	S	NEUTRAL	NEUTRAL	0.65307311	DELETERIOUS	0.57121639	NEUTRAL	0.83210379	NEUTRAL	0.66884082	NEUTRAL	0.62762763	NEUTRAL	0.45949821	NEUTRAL	0.58359402	UNKNOWN	0.0	UNKNOWN	0.0	
S	416	L	NEUTRAL	DELETERIOUS	0.65494636	DELETERIOUS	0.760861957	NEUTRAL	0.7828765	DELETERIOUS	0.59445019	DELETERIOUS	0.60089974	DELETERIOUS	0.79280784	DELETERIOUS	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	
T	417	I	NEUTRAL	NEUTRAL	0.60548272	NEUTRAL	0.65286344	NEUTRAL	0.71871412	DELETERIOUS	0.59445019	DELETERIOUS	0.67524116	DELETERIOUS	0.79280784	DELETERIOUS	0.55551884	UNKNOWN	0.0	UNKNOWN	0.0	
Q	418	H	NEUTRAL	DELETERIOUS	0.65494636	DELETERIOUS	0.55997001	NEUTRAL	0.89245476	DELETERIOUS	0.74491225	DELETERIOUS	0.63431877	DELETERIOUS	0.79280784	DELETERIOUS	0.62208185	UNKNOWN	0.0	UNKNOWN	0.0	

Supplementary Table 3. Structure based analysis (DynaMut, mCSM, SDM and DUET results (continued) of all mutations of N protein abtained with DynaMut server to predict stability

MUTATION	ΔΔG DynaMut (kcal/mol)	ΔΔG mCSM (kcal/mol)	ΔΔG SDM (kcal/mol)	ΔΔG DUET (kcal/mol)
D3L	-0,548	0,194	-0,19	0,38
D3Y	-0,382	-0,065	-0,43	-0,153
P13S	-0,272	-0,745	-0,84	-0,562
A35V	1,13	-0,27	1,15	0,296
Q7K	-0,411	-0,316	0,04	-0,083
G25C	-0,092	-0,966	-0,42	-0,842
Q9H	0,062	-0,5	0,74	-0,199
D22N	0,605	0,564	0,39	0,826
G34W	-1,112	-1,143	-2,14	-1,531
S2Y	1,318	-0,682	0,16	-0,56
N8T	-0,344	-0,207	-0,4	0,032
P13L	1,523	-0,306	-0,04	0,008
R40C	0,137	0,043	-0,4	-0,005
G25V	0,063	-0,494	0,15	-0,146
G44C	-0,843	-1,079	-0,32	-0,938
S23L	-0,459	-0,426	0,94	0,033
S33I	0,216	-0,226	1,37	0,356
A12G	-0,88	-1,353	-0,39	-1,199
A12S	0,413	-1,411	-1,62	-1,419
D22Y	1,464	0,094	0,54	0,267
D3N	-0,217	0,321	-0,07	0,532
D3Q	-0,463	0,311	-0,74	0,465
E31G	-0,094	-0,635	-0,94	-0,683
G19A	0,441	-0,454	0,16	-0,101
N29S	-0,043	0,023	-0,53	0,258
N8Y	0,219	-0,28	-0,18	-0,337
Q7L	0,477	0,017	1,28	0,436
Q9R	0,411	-0,045	0,54	0,303
R10L	0,025	-0,188	0,45	0,052
R14C	-0,223	0,023	-0,24	0,003
R32H	0,233	-0,624	0,12	-0,529
S21P	-0,14	-0,187	-0,05	0,024
S2F	1,005	-0,92	0,36	-0,74
S33G	1,158	-0,142	1,45	0,455
S37P	0,448	-0,162	-0,23	0,024
S37T	0,333	-0,275	0,57	0,189
T16M	-0,316	-0,187	0,62	0,045
T24A	0,195	-0,265	0,61	0,114
Q28K	0,413	-0,053	0,52	0,413
D128N	-0,402	0,004	-0,2	0,082
P67L	0,244	-0,385	-0,04	-0,201
D128Y	-0,036	0,237	0,12	0,196

Supplementary Table 3. (continued)

MUTATION	ΔΔG DynaMut (kcal/mol)	ΔΔG mCSM (kcal/mol)	ΔΔG SDM (kcal/mol)	ΔΔG DUET (kcal/mol)
L139F	-0,137	-0,938	-0,41	-0,942
A119S	0,64	-0,995	-0,85	-0,971
A90S	0,03	-1,159	1,04	0,938
T135I	1,482	-0,39	1,47	0,255
T166I	0,88	-0,174	1,07	0,316
D144H	0,21	-0,377	-0,04	-0,385
E136Q	0,034	0,347	-0,87	0,367
H145Y	0,809	1,127	-0,4	1,035
L167F	1,096	-0,989	-0,46	-1,006
N126S	-0,08	-0,026	-0,44	0,209
P80R	1,025	0,419	1,85	0,783
A152S	-0,09	-0,202	-0,85	-0,131
A173V	0,848	-0,546	1,74	0,26
D128E	-0,003	-0,357	0,25	-0,08
D63Y	2,299	1,688	0,33	1,484
E118Q	-0,743	-0,109	-0,36	0,001
M101I	0,093	-0,321	0,68	0,396
N140S	-0,279	-0,244	-1,15	-0,199
P142S	-0,546	-0,548	-0,84	-0,394
P151A	-0,129	-0,525	0,51	-0,198
P151S	-0,323	-0,999	-1,05	-1,016
P67S	0,078	-1,091	-0,84	-1,055
Q163K	-0,309	-0,459	-0,43	-0,329
T141I	1,026	-0,036	1,28	0,521
T49I	0,45	0,416	0,65	0,798
V72L	0,348	-0,362	-1,21	-0,325
R203K	0,749	-0,179	0,1	0,007
G204R	1,064	-0,139	-1,91	-0,155
S235F	0,312	-0,968	0,58	-0,82
S194L	0,647	-0,446	2,27	0,313
T205I	0,867	-0,141	1,07	0,371
T198I	0,398	-0,275	1,37	0,169
M234I	0,237	0,053	0,44	0,692
G204L	0,013	-0,441	-1,86	-0,565
S202N	0,681	0,013	0,78	0,348
A220V	0,104	-0,278	-1,03	-0,124
A211V	-0,165	-0,134	0,01	0,13
P199S	0,263	-0,454	-0,58	-0,342
G238V	0,108	-0,408	0,09	-0,069
N196Y	-1,169	-0,641	1,109	-0,542
S188P	0,335	0,492	-1,62	0,375
S193T	0,052	-0,336	0,69	0,07
R195I	0,481	0,002	0,28	0,199

Supplementary Table 3. (continued)

MUTATION	ΔΔG DynaMut (kcal/mol)	ΔΔG mCSM (kcal/mol)	ΔΔG SDM (kcal/mol)	ΔΔG DUET (kcal/mol)
P199L	0,287	-0,336	0,7	0,043
G212V	0,014	-0,274	0,12	0,106
G238C	0,088	-0,96	-0,23	-0,792
G214C	1,523	-0,765	0,53	-0,343
R191L	-0,358	-0,854	0,71	-0,38
G215S	0,187	-0,43	-0,64	-0,136
M210I	0,051	0,198	0,55	0,782
P207H	0,053	-0,015	0,91	0,198
P207S	-0,708	-0,336	-0,26	-0,157
A208S	-0,221	-0,499	-0,85	-0,442
A211G	0,271	-0,051	-0,15	0,18
A218V	0,267	-0,631	-1,46	-0,617
G204P	0,684	-0,327	-3,61	-0,844
G204Q	0,324	-0,444	-2,78	-0,623
M210N	0,814	0,8	0,42	1,168
N213S	0,18	0,148	-0,48	0,463
Q229R	1,086	-0,009	0,34	0,386
Q241L	0,346	0,166	1,17	0,693
R185C	-1,292	-2,11	0,33	-1,721
R185G	0,109	-1,437	0,31	-1,059
R209S	0,307	-0,144	-1,48	0,475
S180C	0,265	-0,076	0,83	0,239
S201C	0,321	-0,212	1,18	0,174
S202C	-0,377	-0,206	0,62	0,048
S206F	0,39	-0,864	0,94	-0,621
S232N	0,153	-0,125	0,59	0,367
V246A	-0,099	-0,963	-0,47	-0,781
T362I	0,994	-0,146	1,95	0,606
T334I	-0,065	-0,267	0,64	0,2
M322I	-0,546	-0,498	0,51	0,177
I320F	0,092	-0,872	-0,59	-0,849
A308S	0,159	-0,891	-1,14	-0,774
H300Y	0,502	0,845	-0,39	0,862
M317T	0,316	0,444	-1,39	0,66
P364S	-0,537	-0,588	-0,84	-0,392
T271I	0,569	-0,219	0,89	0,336
T334A	-0,653	-0,311	-0,45	-0,152
V324I	-0,26	-0,684	-0,28	-0,356
V350G	-0,544	-1,871	-0,96	-1,855
I292L	0,144	-1,208	0,2	-0,814
I337V	-0,318	-0,249	-0,12	0,049
K299N	-0,446	-0,288	0,28	-0,112
L339F	-0,134	-1,047	0,12	-0,874

Supplementary Table 3. (continued)

MUTATION	ΔΔG DynaMut (kcal/mol)	ΔΔG mCSM (kcal/mol)	ΔΔG SDM (kcal/mol)	ΔΔG DUET (kcal/mol)
P326S	-0,487	-0,582	-0,26	-0,413
P364Q	-0,247	-0,4	0,64	0,072
Q289H	-0,157	-1,039	0,06	-0,842
S327L	0,753	-0,317	1,03	0,098
T296I	1,195	-0,405	1,93	0,355
V350A	-0,25	-1,43	0,98	-0,843
V350L	0,74	-0,27	0,61	0,28
D377Y	0,953	-0,389	0,77	-0,046
D401Y	1,253	-417	0,78	-0,068
E378Q	-0,247	0,027	-1,19	0,017
P383S	0,194	-0,887	0,01	-0,478
A376T	0,088	-0,757	-1,33	-0,623
P383L	0,612	-0,503	1,57	0,248
T379I	-0,278	-0,141	0,9	0,332
P365S	0,211	-0,416	-0,84	-0,214
A414S	-0,115	-0,159	-2,04	-0,148
E378A	-0,256	-0,511	-0,68	-0,558
D377A	-0,062	-0,175	1,85	0,555
Q380L	0,154	0,188	1,4	0,094
Q389H	0,584	-0,41	0,7	-0,239
S413T	0,333	-0,003	0,27	0,477
S416L	-0,308	-0,347	0,98	0,054
T417I	0,043	-0,268	0,9	0,116
K375E	0,351	0,012	0,99	0,621
P365L	0,055	-0,359	-0,04	-0,032
P365Q	0,177	-0,386	-0,81	-0,177
Q389L	0,705	0,227	1,11	0,602
Q390H	-0,322	-0,055	0,85	0,235
Q418H	0,35	-0,212	0,91	-0,001
S413I	0,175	-0,417	1,02	0,168
T366I	0,34	-0,031	0,46	0,314
T379A	-0,33	-0,348	0,68	0,069
T391I	0,245	-0,17	1,03	0,311

Supplementary Table 4. Structure based analysis (DynaMut, mCSM, SDM and DUET results) of all mutations of N protein obtained with DynaMut server to predict stability

MUTATION	ΔΔS ENCoM (kcal^{-1}/mol^{-1})	ΔΔSVibENCoM (kcal^{-1}/mol^{-1})	MUTATION	ΔΔS ENCoM (kcal^{-1}/mol^{-1})	ΔΔSVibENCoM (kcal^{-1}/mol^{-1})
D3L	0,044	0,044	D128Y	0,244	0,244
D3Y	-0,03	-0,03	L139F	-0,197	-0,197
P13S	0,472	0,472	A119S	-0,222	-0,222
A35V	-0,174	-0,174	A90S	0,008	0,008
Q7K	0,509	0,509	T135I	-0,422	-0,422
G25C	-0,022	-0,022	T166I	-0,226	-0,226
Q9H	-0,365	-0,365	D144H	-0,322	-0,322
D22N	0,162	0,162	E136Q	-0,015	-0,015
G34W	0,073	0,073	H145Y	0,144	0,144
S2Y	-0,692	-0,692	L167F	-0,543	-0,543
N8T	-0,051	-0,051	N126S	-0,024	-0,024
P13L	-0,693	-0,693	P80R	-0,54	-0,54
R40C	0,883	0,883	A152S	-0,137	-0,137
G25V	0,01	0,01	A173V	-0,511	-0,511
G44C	0,186	0,186	D128E	0,17	0,17
S23L	-0,034	-0,034	D63Y	-0,43	-0,943
S33I	-0,055	-0,055	E118Q	0,39	0,39
A12G	0,42	0,42	M101I	0,419	0,419
A12S	-0,029	-0,029	N140S	0,03	0,03
D22Y	-1,928	-1,928	P142S	-0,06	-0,06
D3N	0,059	0,059	P151A	0,297	0,297
D3Q	0,042	0,042	P151S	-0,055	-0,055
E31G	0,616	0,616	P67S	-0,009	-0,009
G19A	-0,015	-0,015	Q163K	0,296	0,296
N29S	0,454	0,454	T141I	-0,315	-0,315
N8Y	-0,191	-0,191	T49I	0,152	0,152
Q7L	0,843	0,843	R203K	-0,107	-0,107
Q9R	-0,397	-0,397	G204R	-2,522	-2,522
R10L	0,838	0,838	S235F	-0,135	-0,135
R14C	0,852	0,852	S194L	0,044	0,044
R32H	-0,055	-0,055	T205I	-0,289	-0,289
S21P	-0,358	-0,358	T198I	-0,07	-0,07
S2F	-0,712	-0,712	M234I	0,057	0,057
S33G	-0,134	-0,134	G204L	-0,901	-0,901
S37P	-0,565	-0,565	S202N	-0,215	-0,215
S37T	-0,242	-0,242	A220V	-0,018	-0,018
T16M	-0,034	-0,034	A211V	-0,074	-0,074
T24A	0,04	0,04	P199S	-0,44	-0,44
Q28K	-0,15	-0,15	G238V	-0,041	-0,041
V72L	-0,626	-0,626	N196Y	0,334	0,334
D128N	0,209	0,209	S188P	0,141	0,141
P67L	-0,169	-0,169	S193T	0,01	0,01

Supplementary Table 4. (continued)

MUTATION	ΔΔS ENCoM (kcal^{-1}/mol^{-1})	ΔΔSVibENCoM (kcal^{-1}/mol^{-1})	MUTATION	ΔΔS ENCoM (kcal^{-1}/mol^{-1})	ΔΔSVibENCoM (kcal^{-1}/mol^{-1})
R195I	0,349	0,349	V350G	0,296	0,296
P199L	-0,021	-0,021	I292L	0,079	0,079
G212V	0,003	0,003	I337V	0,464	0,464
G238C	-0,004	-0,004	K299N	0,383	0,383
G214C	-1,357	-1,357	L339F	0,13	0,13
G215S	-0,259	-0,259	P326S	0,012	0,012
R191L	0,134	0,134	P364Q	-0,024	-0,024
M210I	0,469	0,469	Q289H	-0,05	-0,05
P207H	-0,051	-0,051	S327L	-0,122	-0,122
P207S	-0,01	-0,01	T296I	-0,422	-0,422
A208S	0,023	0,023	V350A	0,265	0,265
A211G	0,088	0,088	V350L	-0,181	-0,181
A218V	-0,388	0,388	D377Y	-0,545	-0,545
G204P	-0,572	-0,572	D401Y	-0,487	-0,487
G204Q	-0,938	-0,938	E378Q	0,072	0,072
M210N	0,81	0,81	P383S	-0,006	-0,006
N213S	-0,162	-0,162	A376T	-0,041	-0,041
Q229R	-0,36	-0,36	P383L	-0,058	-0,058
Q241L	0,089	0,089	T379I	-0,043	-0,043
R185C	1,042	1,042	P365S	0,115	0,115
R185G	1,753	1,753	A414S	-0,031	-0,031
R209S	1,731	1,731	E378A	0,217	0,217
S180C	-0,059	-0,059	D377A	0,249	0,249
S201C	0,176	0,176	Q380L	0,094	0,094
S202C	0,315	0,315	Q389H	-0,119	-0,119
S206F	0,016	0,016	S413T	0,143	0,143
S232N	0,046	0,046	S416L	0,086	0,086
V246A	0,311	0,311	T417I	0,073	0,073
T334A	0,149	0,149	K375E	0,127	0,127
T362I	-0,23	-0,23	P365L	0,131	0,131
T334I	0,267	0,267	P365Q	0,141	0,141
M322I	0,175	0,175	Q389L	0,01	0,01
I320F	-0,68	-0,68	Q390H	0,019	0,019
A308S	-0,104	-0,104	Q418H	-0,222	-0,222
H300Y	-0,175	-0,175	S413I	0,077	0,077
M317T	0,342	0,342	T366I	-0,238	-0,238
P364S	0,003	0,003	T379A	0,143	0,143
T271I	-0,398	-0,398	T391I	0,052	0,052
V324I	-0,114	-0,114			

ΔΔS ENCoM and ΔΔSVib ENCoM analysis of all mutations of N protein obtained with DynaMut server to predict flexibility