Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Computers in Biology and Medicine 132 (2021) 104315

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: http://www.elsevier.com/locate/compbiomed

A hijack mechanism of Indian SARS-CoV-2 isolates for relapsing contemporary antiviral therapeutics

R. Prathiviraj, S. Saranya, M. Bharathi, P. Chellapandi *

Industrial Systems Biology Lab, Department of Bioinformatics, School of Life Sciences, Bharathidasan University, Tiruchirappalli, 620024, Tamil Nadu, India

ARTICLE INFO

Keywords:
COVID-19
SARS-CoV-2
Antivirals
Genome variation
Fold-rate
Phylogeography

ABSTRACT

Coronavirus disease (COVID-19) rapidly expands to a global pandemic and its impact on public health varies from country to country. It is caused by a new virus severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). It is imperative for relapsing current antiviral therapeutics owing to randomized genetic drift in global SARS-CoV-2 isolates. A molecular mechanism behind the emerging genomic variants is not yet understood for the prioritization of selective antivirals. The present computational study was aimed to repurpose existing antivirals for Indian SARS-CoV-2 isolates by uncovering a hijack mechanism based on structural and functional characteristics of protein variants. Forty-one protein mutations were identified in 12 Indian SARS-CoV-2 isolates by analysis of genome variations across 460 genome sequences obtained from 30 geographic sites in India. Two unique mutations such as W6152R and N5928H found in exonuclease of Surat (GBRC275b) and Gandhinagar (GBRC239) isolates. We report for the first time the impact of folding rate on stabilizing/retaining a sequence-structure-function-virulence link of emerging protein variants leading to accommodate hijack ability from current antivirals. Binding affinity analysis revealed the effect of point mutations on virus infectivity and the drug-escaping efficiency of Indian isolates. Emodin and artinemol suggested herein as repurposable antivirals for the treatment of COVID-19 patients infected with Indian isolates. Our study concludes that a protein folding rate is a key structural and evolutionary determinant to enhance the receptor-binding specificity and ensure hijack ability from the prevalent antiviral therapeutics.

1. Introduction

Coronavirus disease (COVID-19) is a global epidemic caused by a new virus severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The World Health Organization (WHO) declared this outbreak a global health emergency on March 11, 2020 [1]). COVID-19 is affecting 213 countries and territories around the world and 2 international conveyances. As of February 10, 2021, 107 M confirmed, and 2.33 M death cases have been reported worldwide. India is the 2nd most-affected 20 countries with COVID-19 with 8.41 M confirmed cases of which 125 K people have succumbed to the infection (https://www.worldometers.info/coronavirus/). The SARS-CoV-2 outbreak spread rapidly worldwide by human-to-human contact [2]. No drugs or vaccines have been approved by the U.S. Food and Drug Administration (FDA) for the prevention or treatment of COVID-19. The FDA authorized Remdesivir as an emergency antiviral for the treatment of severe COVID-19 patients on May 1, 2020 [3]. The WHO has suggested randomized standard care or active treatment (remdesivir, chloroquine/hydroxychloroquine plus azithromycin, lopinavir/ritonavir, or lopinavir/ritonavir plus interferon-β1a) to confirmed cases of COVD-19 [4]. Several systematic investigations on target-based antiviral discovery have been in progress for the development of effective therapies for COVID-19 infections [5].

SARS-CoV-2 is an enveloped single positive-stranded RNA virus [1]. Have completed the first genome sequence (accession: NC_045512.2; Wuhan-Hu-1) of SARS-CoV-2 isolated from Wuhan, Hubei province, China. This reference genome consists of 29.9 Kb organized into specific genes encoding structural (spike, envelope, membrane, and nucleocapsid) and non-structural proteins (Nsp1-16). Spike glycoprotein plays an important role in viral infectivity and disease transmission by interacting with the human ACE-2 receptor [6]). Papain-like protease (PLPro), 3-chymotrypsin-like protease (3CLPro), RNA-dependent RNA polymerase (RdRp), helicase, exoribonuclease, and 2′-O-methyltransferase are promising therapeutic targets for antivirals development [5]. Besides, ORF3a, ORF6, ORF7a ORF8, and ORF10 genes are accessory proteins essential for virus assembly and infection [7]).

* Corresponding author.
E-mail address: pchellapandi@gmail.com (P. Chellapandi).

https://doi.org/10.1016/j.compbiomed.2021.104315
Received 10 February 2021; Accepted 2 March 2021
Available online 6 March 2021

Enormous genome sequences of SARS-CoV-2 are available to the research community. Genomic structure and variations of SARS-CoV-2 are pro-vital for studying virulence, infectivity, and COVID-19 outbreak. SARS-CoV-2 isolates are distributed non-geographically with the prevalence of single nucleotide mutations as the major genomic variations. Genetic variance analyses revealed some subtypes, deleterious, and frequent mutation events in SARS-CoV-2 isolates [8–11]. Therefore, the emergence of new mutations in SARS-CoV-2 may force the subsequent establishment of a new escape or hijack mechanism for antivirals.

The viral hijack mechanism is an emphatic one to escape the host immune systems and antiviral therapies. Human immunodeficiency virus 1 (HIV-1) hijacks its genomic and structural perspective for survival advantage in a host [12], cell-to-cell transmission [13], and membrane traffic to promote infection [14]. A host cellular pathway is hijacked by SARS-CoV for supplying replication membranes for viral protein synthesis [15,16]. HIV-1 virion infectivity factor is found as a molecular determinant to escape from antiviral APOBEC3 (A3) proteins [17]. There is no report on the hijack ability of SARS-CoV-2 from antivirals.

The sequence-structure-function-conformational dynamic space-folding rate-virulence link revealed as a molecular determinant of some virulence proteins [18–27]. Accordingly, we attempted to know how SARS-CoV-2 protein variants hijack or escape from current antivirals and to increase host-specific infectivity in phylogenetic populations. A protein thermodynamic stability accounts for a large fraction of observed mutational effects in RNA viruses. Destabilizing mutations inflict a fitness penalty proportional to the reduction in folded protein [28,29]. A biophysical model developed from this hypothesis was used to study RNA viruses escaping antibody stress [30].

The folding kinetics of a protein depends on the reduced hydrophobic density and enhanced polar-polar repulsive interactions [31]. The folding rates of spike glycoprotein and membrane protein are essential in the formation of the viral envelope of coronaviruses. At the specific folding rates, both proteins are accumulated at the site of virus assembly (pre-Golgi membranes) [32]. Thus, viruses follow different blueprints of interference with protein folding to ensure their intracellular survival like bacteria [33]. Therefore, our study was intended to propose a hijack mechanism of Indian SARS-CoV-2 isolates for relapsing current antivirals by computing structural and functional determinants of identified protein variants. This study reports significant protein mutations from these isolates and stratifies them structurally. Our findings on selective variants would be useful for systematic repurposing of antivirals on a regional basis to mitigate the spreading of COVID-19.

2. Materials and methods

2.1. Dataset

In this study, a complete SARS-CoV-2 nucleotide sequence (accession: NC_045512.2; Wuhan-Hu-1 isolate) was downloaded as a reference genome from the National Centre for Biotechnology Information (NCBI)-SARS-CoV-2 resources (https://www.ncbi.nlm.nih.gov/sars-cov-2/). This resource consists of 59, 482 Sequence Read Archive runs, and 24,380 nucleotide records by 24, September 2020. The Sequence Read Archive runs were neglected in the phylogenomic study as poorly annotated coding regions from raw sequence data, the impact of recombination, and incomplete lineage sorting [34]. A total of 16, 557 SARS-CoV-2 nucleotide sequences (sequence length ranged from 29600 to 31000 bp) isolated from different countries were used in this study. Among six geographic origins, 11, 578 sequences isolated from North and South America, 1424 from Oceania, 1571 from Asia, 460 from Europe, and 189 from Africa. The 460 genomes were sequenced in India of which 361 genomes sequenced from 277 confirmed COVID-19 cases (132 female and 229 male) by Gujarat Biotechnology Research Centre, Gandhi Nagar, India. The two genomes were deposited from Kerala State, three genomes from Una, Himachal Pradesh (Una), and one genome from Hyderabad, Telangana State. The number of ambiguous nucleotides was <1% in the trimmed sequences. The dataset (accession, isolate name, country, region, collection date) used for the tree building is available in **Supplementary file 1**. The computational process for proposing a hijack mechanism of Indian SARS-CoV-2 isolates for relapsing current antiviral therapeutics is represented in Fig. 1.

2.2. Phylogeographic tree construction and analysis

We performed multiple sequence alignment of all SARS-CoV-2 sequences to a reference genome using MAFFT v7.0 (Multiple Alignment using Fast Fourier Transform) with truncating the bases in the 5′-UTR and 3′-UTR without losing key sites [35]. After alignment, the tree was generated for entire isolates based nucleotide variations with the NCBI-SARS-CoV-2 resources and Virus Pathogen Database and Analysis Resource (ViPR) [36]). The phylogenomic tree was constructed using the Neighbor-Joining algorithm [37] with the Jukes-Cantor evolutionary model [38] and 1000 bootstraps. The identical nodes were collapsed and then the tree was visualized by Interactive Tree Of Life (iTOL) v5.0 [39]). In this study, 95 SARS-CoV-2 sequences from different coronavirus hotspot countries and 177 sequences from Indian origin were selected and then aligned to a reference genome using MAFFT. The sampling trees were generated with the Bayesian Evolutionary Analysis by Sampling Tree (BEAST) v2.5 [40] with 1000 bootstrapping. Each tree was weighted proportional to its posterior probability. Evolutionary distances were computed between pairs of sequences based on mutational substitutions. Finally, phylogenomic tree nodes were annotated based on geographic regions using the iTOL tool.

2.3. Analysis of genomic variations

We computed the extent of variation at each position of nucleotide sequences uploaded to the ViPR genome variation analysis tool and visualized aligned nucleotide sequences using the ViPR interactive alignment viewer. Aligned sequences generated from the ViPR tool were further viewed in the BioEDIT v7.2 software [41] for the detection of mutation patterns compared to a reference genome. All nucleotide variants identified in the coding regions converted to encoded amino acid residues. The sequence similarity comparison was carried to find out protein variants from a reference genome by searching identical residues against a reference genome using the NCBI BLAST tool [42]. The optimal global protein sequence alignment was conducted between a reference genome and selected Indian SARS-CoV-2 isolates by the EMBOSS Needle tool (https://www.ebi.ac.uk/Tools/psa/emboss_needle/). To identify cross conservation and to understand the nature of amino acid changes, we used the MicroGMT software, a mutation tracker for SARS-CoV-2 [43].

2.4. Structural and functional characterization of protein variants

The structural and functional characteristics of protein variants identified from selected Indian SARS-CoV-2 isolates were predicted by various bioinformatics tools. All predicted characteristics were compared with proteins from a reference genome (Wuhan-Hu-1). The structural flexibility deviation of protein variants from reference proteins was analyzed based on a library composed of structural prototypes using the PredyFlexy server [44]. The folding rate of identified protein variants was computed from the sequences with the FOLD-RATE based on a set of 49 diverse amino acid properties [45]. The average amino acid property (P_{ave}) was calculated using the following equation.

$$P(ave)(i) = \sum_{j=1}^{N} P(j) \Big/ N$$

where $P(j)$ is the property value of jth residue and the summation is over

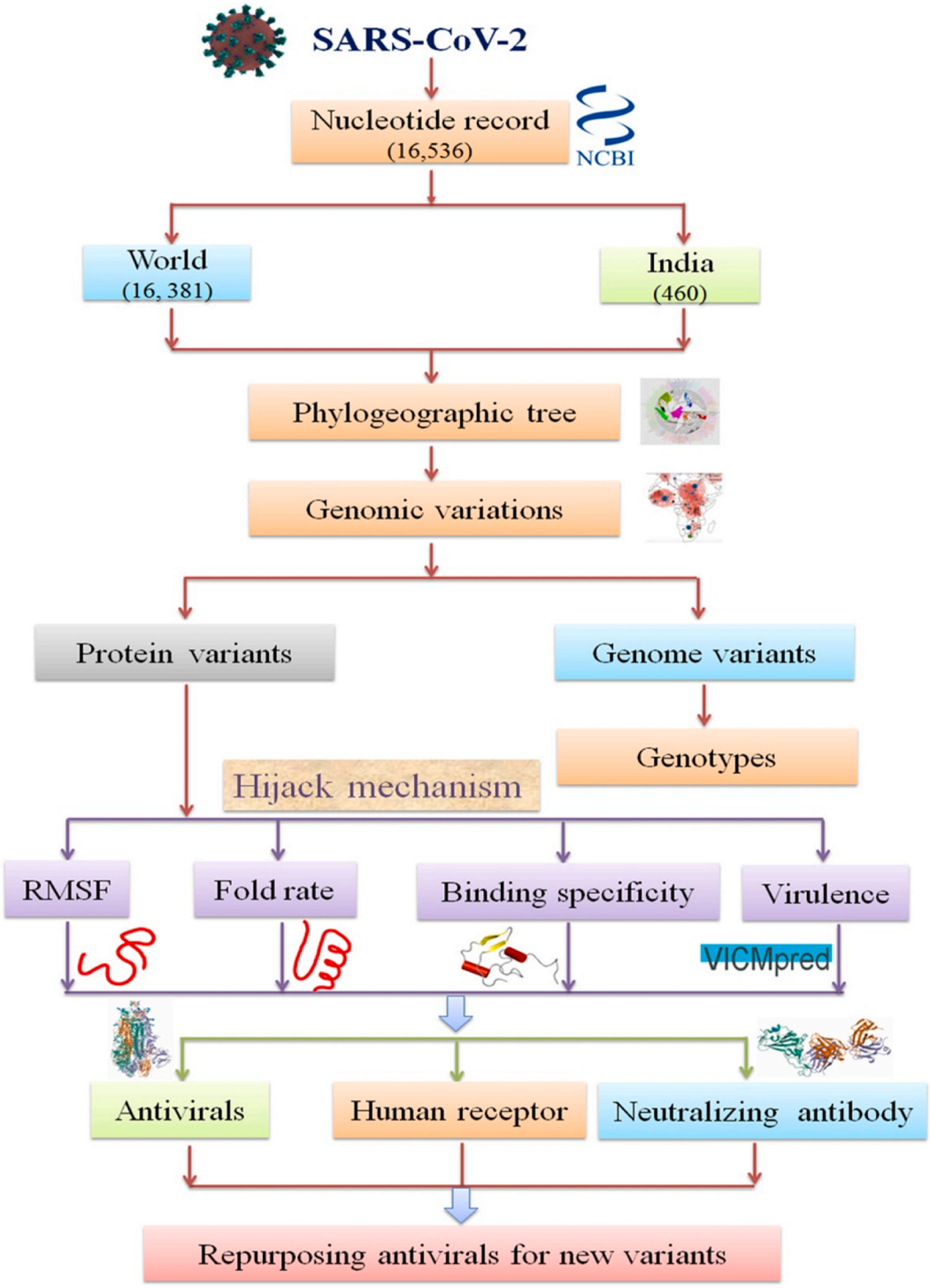


Fig. 1. Flow chart for proposing a hijack mechanism of Indian SARS-CoV-2 isolates for relapsing current antiviral therapeutics.

the total number of residues in a protein (*N*). Multiple regression techniques were used to correlate predicted folding rate related to experimental folding rate *lnkf* (i). The protein folding kinetic order and folding rate of the variants were calculated from their 3D structures using K-FOLD based on a support vector machine [46]. The protein variants retaining any virulence property, information molecule, and involvement in the cellular metabolic process were predicted from amino acid and dipeptide composition. It was conducted with the VICMpred server using Matthew's correlation coefficient [47].

2.5. Characterization of the binding specificity of protein variants

The FDA-approved antiviral drugs were retrieved from the DrugBANK v5.0 [4,48]. A cryo-electron microscopic structure of SARS-CoV-2 binding human receptor ACE-2-B^{0}AT1 (PDB ID: 6M17) was downloaded from the Protein Data Bank (PDB). The crystal structures of human neutralizing antibodies (anti-SARS-CoV-2 antibodies) CC12.1, CR3022 (PDB ID: 6XC3), m396 (PDB ID: 2G75), and 80R (PDB ID: 2GHW) were obtained from the PDB [49]. The 3-D structures of the proteins from both Wuhan-Hu-1 and selected Indian isolates were built by the

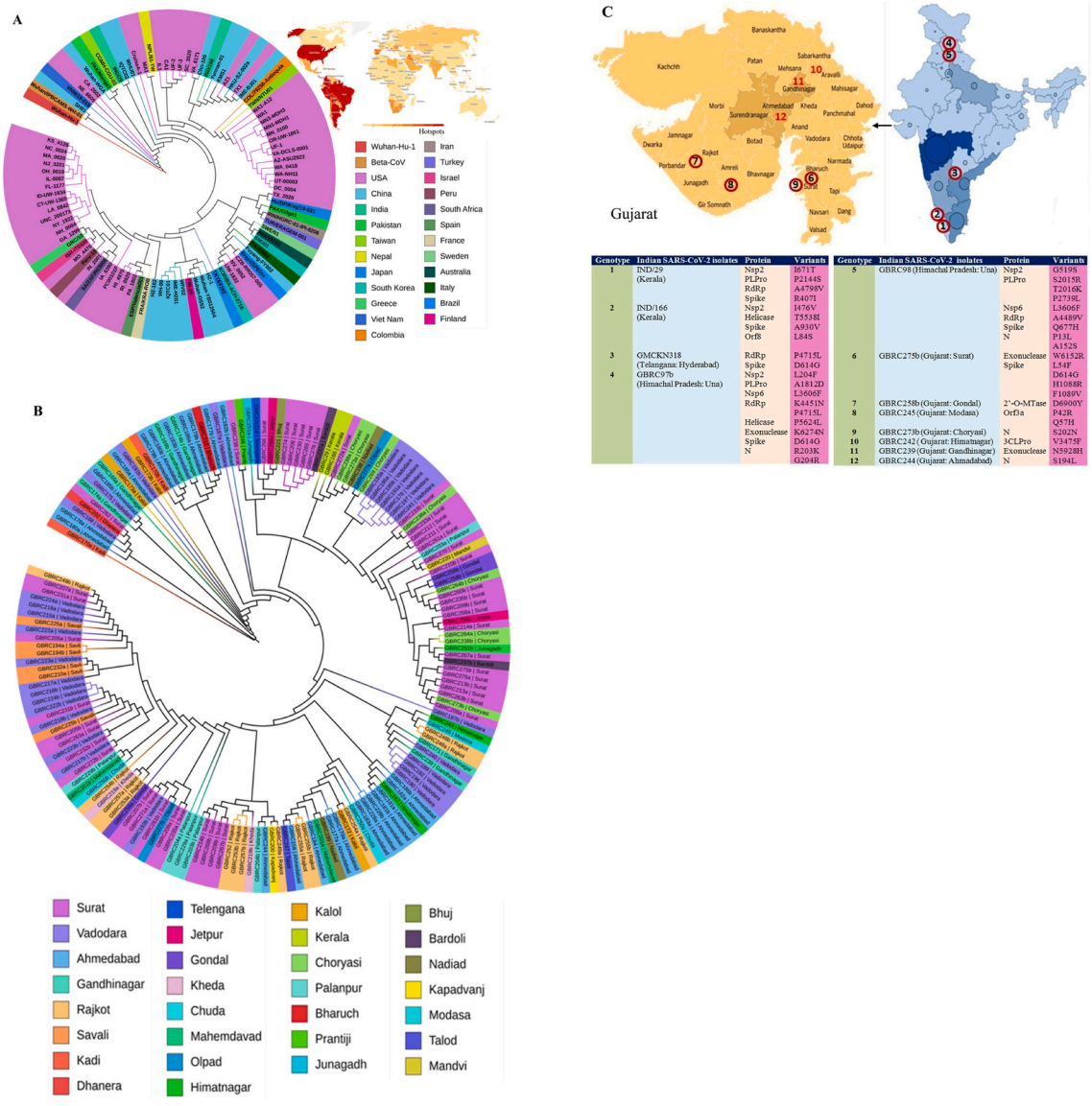
SWIS-MODEL [50] and I-TASSER [51]. The structural constraints in the modeled proteins were energetically minimized by the refineD server using machine learning-based restrained relaxation [52].

A protein-ligand docking was carried for the prediction of the binding specificity of antiviral drugs to the protein variants identified from Indian isolates using the PatchDock server [36]. The PatchDock algorithm is based on hot-spot rich surface patches matching and initial transformation with geometric hashing and poses clustering. The shape complementarity, interface shape, and sizes of the molecules determine the quality of docked models [36]). A rigid body protein-protein docking was conducted with clustering RMSD 4.0 for the prediction of the binding specificity of Spike protein to ACE-2-B^{0}AT1 and anti-SARS-CoV-2 vaccines by the PatchDock, followed by the FireDock [53]. Binding energies, hydrogen bonding, attractive and repulsive van der Waals were calculated from the resulted protein-protein complexes by optimizing side-chain conformations and rigid-body orientations. An iterative knowledge-based scoring function was used to calculate the ligand RMSD and docking score of protein-protein complexes by the HDOCK [6,54]. The scoring functions of all docking models were evaluated by the AutoDock Vina v1.1.2 software [55]. The docked models were visualized in the PyMOL v2.4 software [56]. The hydrogen and hydrophobic interactions in the drug-protein and protein-protein complexes were calculated and visualized using the LigPlot$^+$ v2.2 software [57].

3. Results

3.1. Phylogenomic analysis

To identify the genetic variations that are potentially linked to the severity of COVID-19, we constructed a large-scale phylogenetic tree with 16, 536 SARS-CoV-2 completely annotated genome sequences worldwide (Fig. S1). The majority of SARS-CoV-2 isolates from the USA and China were clustered within the country and some of them scattered across the world. As shown in Fig. 2A, the first Wuhan-Hu-1 isolate was closely clustered with isolates from Vietnam (VNM/nCoV-19-01S) and USA (USA/CA_2602, USA/NH_0004). This cluster is grouped with isolates from India (IND/Kerala/29), Pakistan (PAK/Manga1), and Nepal



Genotype	Indian SARS-CoV-2 isolates	Protein	Variants
1	IND/29 (Kerala)	Nsp2	I671T
		PLPro	P2144S
		RdRp	A4798V
		Spike	R407I
2	IND/166 (Kerala)	Nsp2	I476V
		Helicase	T5538I
		Spike	A930V
		Orf8	L84S
3	GMCKN318 (Telangana: Hyderabad)	RdRp	P4715L
		Spike	D614G
4	GBRC97b (Himachal Pradesh: Una)	Nsp2	L204F
		PLPro	A1812D
		Nsp6	L3606F
		RdRp	K4451N
			P4715L
		Helicase	P5624L
		Exonuclease	K6274N
		Spike	D614G
		N	R203K
			G204R

Genotype	Indian SARS-CoV-2 isolates	Protein	Variants
5	GBRC98 (Himachal Pradesh: Una)	Nsp2	G519S
		PLPro	S2015R
			T2016K
			P2739L
		Nsp6	L3606F
		RdRp	A4489V
		Spike	Q677H
		N	P13L
			A152S
6	GBRC275b (Gujarat: Surat)	Exonuclease	W6152R
		Spike	L54F
			D614G
			H1088R
			F1089V
7	GBRC258b (Gujarat: Gondal)	2'-O-MTase	D6900Y
8	GBRC245 (Gujarat: Modasa)	Orf3a	P42R
			Q57H
9	GBRC273b (Gujarat: Choryasi)	N	S202N
10	GBRC242 (Gujarat: Himatnagar)	3CLPro	V3475F
11	GBRC239 (Gujarat: Gandhinagar)	Exonuclease	N5928H
12	GBRC244 (Gujarat: Ahmadabad)	N	S194L

Fig. 2. Neighbor-joining trees for the representation of the phylogeographic distribution of selected SARS-CoV-2 isolates. Fig. 2A represents isolates obtained from different countries whereas Fig. 2B depicts the phylogenetic classification of isolates from different regions in India. The top-right coronavirus map in Fig. 2A represents the hotspot countries as of August 24, 2020 (https://www.worldometers.info/coronavirus/). Fig. 2C (top) shows the location of SARS-CoV-2 isolates from Indian regions where deleterious mutations were identified in this study.

(NPL/61-TW). The second isolate (IND/Kerala/166) of India forms a separate sub-clade with isolates from China (CHN/Yunnan-01/2020, CHN/KMS1) and the USA (USA-AZ1). Our analysis suggests that many Gujarat isolates are genetically more related to the isolates from Israel, Columbia, and Malaysia. The results of this study indicate a phylogenetic disparity across the Indian isolates.

3.2. Phylogenomic analysis of Indian SARS-CoV-2 isolates

Besides, we constructed a phylogenetic tree with 3 clades to know the genome sequence similarity and dissimilarity across Indian SARS-CoV-2 isolates (Fig. 2B; Fig. S2). The isolates with identical sequence coverage were removed from the tree. As shown by our analysis, SARS-CoV-2 isolates obtained from Surat, Ahmadabad, and Vadodara have genetically diverged from each other. The first SARS-CoV-2 (IND/29) was isolated from a patient of Kerala state on March 29, 2020 [58]. This isolate had only 5 genomic variants compared to the Wuhan-Hu-1 isolate. It has a distinct genetic variation with Kerala isolate (IND/166) and subsequently expanded its genetic diversity with multiple substitutions compared to the Gujarat isolates. Himachal Pradesh (GBRC97b, GBRC98) isolates are clustered with Jamnagar (GBRC108, 106) and Kodinar (GBRC100) isolates whereas Kerala (IND/29) related to Surat (GBRC272ba), Ahmadabad (GBRC157a) and Vadodara (GBRC125, 130) isolates. Telangana (GMCKN318) isolate is also grouped with Surat isolates. Phylogenetic studies of Indian SARS-CoV-2 isolates revealed that Surat isolates are the important ones to disperse genetic diversity leading to an increased virus transmission rate in India.

3.3. Analysis of genomic variations

Genome sequences of Indian SARS-CoV-2 isolates were compared with a reference genome (Wuhan-Hu-1) for the identification of gene mutation events (Supplementary file 2). A total of 291 gene mutation events were detected in 26 Indian SARS-CoV-2 isolates. Results suggested that the isolates obtained from Savali (GBRC225a), Surat (GBRC275b), and Chuda (GBRC250b) had more than 15 mutations compared to others. The number of nucleotide differences varies from a minimum of 5 nucleotides to a maximum of 18 nucleotides among Indian isolates. As shown by our analysis, the C-to T exchange has the most prominent numbers in all nucleotide change directions in Indian SARS-CoV-2 genomes. Genomic variations C199T, A2250C, C833T, C2794T, C2995T, G4258T, C11041T, C14366T, C14529T, C18526T, C18835T, C19112T, G21682T, C22402T, A23361G, G25521T, C26693T, C28812T and C29708T identified from Indian isolates introducing mutations in corresponding proteins.

3.4. Analysis of protein variants

By our protein mutation analysis, 12 hotspot regions were identified in India of which 7 from Gujarat state and 5 from other Indian regions (Fig. 2C). We identified the above 9 protein mutations from Himachal Pradesh (GBRC97b, GBRC98) and Surat (GBRC275b) isolates. The mutation events frequently occur in spike protein, RdRp, Nsp2, and PLpro. A total of 171 D614G mutations are identified from 460 Indian isolates in which 59 mutations in Surat isolates. D614G is a frequent mutation identified from Telangana, Himachal Pradesh (GBRC97b), and Surat (GBRC275b) isolates. Besides, a rare mutation (P1089V) is identified in the spike protein of Surat isolate (GBRC275b). The study found two new protein mutations W6152R and N5928H in exonuclease of Surat isolate (GBRC275b) and Gandhinagar isolate (GBRC239), respectively. There are no genetic variations in the Nsp7, Nsp8, Nsp10, Nsp11, endRNAs, orf6, and orf7a of Indian SARS-CoV-2 isolates. It suggests that these genes are highly conserved and identical to the Wuhan-Hu-1 isolate.

3.5. Impact of folding rate on mutated proteins

The folding rate is a key measure to determine the protein multiplication, virulence, host specificity, and infectivity of a pathogen. The influence of folding rates of different structural classes of mutated proteins was studied herein (Table 1). It indicates the kinetic folding rates of all mutated proteins are shown slightly faster than reference proteins. We predicted a fast-folding rate in all β-classes followed by all α-classes compared to reference proteins. It suggests that there are no drastic changes in secondary structural elements in the mutated proteins. A slow folding rate is predicted in mixed structural classes of mutated proteins. The folding rates are influenced by the unknown classes of mutated proteins, particularly in coil and turns (Fig. S3). It indicates the structure-function stabilization points to be found in such folding regions. We also analyzed the folding rate of mutated proteins from Indian isolates to know the effect on the structure-function-virulence link. The fast folding rates are predicted in the Orf1ab protein of Telengana and Kerala isolates and spike, PLPro, Nsp2, Orf8, and RdRp of many Indian isolates. A slow folding rate is noted in all structural classes of helicase from Kerala isolates, α-classes of spike protein from Surat (GBRC275b) isolate, and nucleocapsid protein of many isolates. There is a drastic change in the folding rate of exonuclease of Indian isolates as the results of W6152R and N5928H mutations. The results of this study suggest the folding rate can determine the structure-function relationships of mutated proteins, which in turn increase viral load inside hosts *via* a fitted post-translation process.

3.6. Impact of virulence property of mutated proteins

A virulence property of a protein can be determined by its amino acid composition and physicochemical characteristics of substituted amino acids. Our results show that selected Indian isolates have no major involvement in information processing but some isolates contributed to the metabolic process (Table 2). The virulence properties are predicted in PLPro from Kerala (IND/29), Himachal Pradesh (GBRC98) isolates, and 3CLPro from Himatnagar (GBRC242) isolate. Himachal Pradesh (GBRC97b), Surat (GBRC275b), and Gandhinagar (GBRC239) isolates have retained virulence in exonuclease but Wuhan-Hu-1isolate not retained so. Himachal Pradesh (GBRC97b) isolate had virulence property in the helicase. Interestingly, we found a new functional shift from virulence to metabolic process upon T5538I mutation in helicase of Kerala (IND/166) isolate, suggested that its survivability and adaptability in host systems.

3.7. Impact of binding specificity to the FDA-approved antiviral drugs

Antiviral drugs were docked with respective SARS-CoV-2 proteins and their binding affinity of them calculated to investigate drug escaping mechanisms (Table 3). Molecular docking studies show that antiviral drugs have low binding energies to many Indian isolates compared to Wuhan-Hu-1. It points out that current antivirals could not interact with proteins from Indian isolates having deleterious mutation events. Remdesivir has shown a low binding specificity to the RdRp of SARS-CoV-2 isolates reported in India, suggestive of relapsing current antiviral treatment. Emodin is a proficient antiviral targeting spike protein, which shows more binding specificity to Indian isolates except for the D614G variant. The drug specificity is predicted to decrease in Surat isolate (GBRC275b) due to the newly identified mutation (W6152R) in exonuclease. However, there is no radical change in the drug specificity of a new mutation (N5928H) in Gandhinagar isolate (GBRC239) except ribavirin. Artenimol also shows a more binding affinity to helicase in Indian isolates. Remdesivir, ribavirin, and azacitidine are no longer interactions with respective targets whereas emodin and artenimol are suggested as suitable antivirals for the treatment of patients infected with these isolates.

Table 1
Comparative analysis of folding rate and folding mechanism of protein variants identified from Indian SARS-CoV-2 isolates.

Genotype	Isolate	Protein	Mutation	Folding classes								Fold rate ln (kf)	
				All-α		All-β		α+β:α/β		Unknown			
				Ref.	Var.	Ref.	Var.	Ref.	Var.	Ref.	Var.	Ref.	Var.
1	IND/29 (Kerala)	Orf1ab		8.08	8.08	22.6	21.6	8.67	−0.349	5.9	7.13	45.25	36.46
		Nsp2	I671T	11.4	12.4	12.3	14.1	−28.7	−17.1	6.19	4.8	1.19	14.2
		PLPro	P2144S	6.26	6.49	13.9	14.9	−0.51	3.08	−3.42	8.99	16.22	33.46
		RdRp	A4798V	13.9	13.6	21.7	20.6	−13.3	−5.21	4.04	6.04	26.34	35.03
		Spike	R407I	3.4	4.76	21.5	23.7	−11	3.91	6.12	5.46	20.02	37.83
2	IND/166 (Kerala)	Orf1ab		8.08	7.86	22.6	24	8.67	14.2	5.9	11.7	45.25	57.76
		Nsp2	I476V	11.4	13	12.3	15.8	−28.7	−17.1	6.19	10.9	1.19	22.6
		Helicase	T5538I	3.56	3.84	17.8	16.4	−3.33	−24.1	11.2	1.55	29.23	−2.31
		Spike	A930V	3.4	4.38	21.5	20.4	−11	−5.92	6.12	0.496	20.02	19.35
		Orf8	L84S	0.827	0.037	42.2	42.6	−44.6	−14.4	5.47	4.22	3.897	32.45
3	GMCKN318 (Telangana: Hyderabad)	Orf1ab		8.08	7.86	22.6	24	8.67	14.2	5.9	13.5	45.25	59.56
		RdRp	P4715L	13.9	13.8	21.7	21.5	−13.3	−0.237	4.04	4.96	26.34	40.02
		Spike	D614G	3.4	4.65	21.5	24.9	−11	−0.137	6.12	5.89	20.02	35.30
4	GBRC97b (Himachal Pradesh: Una)	Orf1ab		8.08	7.91	22.6	21.8	8.67	3.56	5.9	10.1	45.25	43.37
		Nsp2	L204F	11.4	12.4	12.3	15.6	−28.7	−19.7	6.19	−0.63	1.19	7.667
		PLPro	A1812D	6.26	6.67	13.9	12.1	−0.51	8.29	−3.42	−0.85	16.22	26.20
		Nsp6	L3606F	6.6	7.5	63.8	66.9	−17.9	−3.54	0.996	−7.75	53.49	63.11
		RdRp	K4451 N P4715L	13.9	14.3	21.7	22.9	−13.3	−9.05	4.04	7.33	26.34	35.48
		Helicase	P5624L	3.56	4.38	17.8	18.4	−3.33	−3.6	11.2	1.99	29.23	21.17
		Exonuclease	K6274 N	9.05	8.53	16.9	19.9	4.29	−5.47	0.83	5.23	31.07	28.19
		Spike	D614G	3.4	4.65	21.5	24.9	−11	−0.137	6.12	5.89	20.02	35.30
		N	R203K G204R	−9.64	−10.2	−7.18	−6.32	28.6	17.2	5.53	6.23	17.31	6.91
5	GBRC98 (Himachal Pradesh: Una)	Orf1ab		8.08	8.43	22.6	20.6	8.67	7.32	5.9	11.7	45.25	48.05
		Nsp2	G519S	11.4	12.3	12.3	17.4	−28.7	−17.1	6.19	4.44	1.19	17.04
		PLPro	S2015R T2016K P2739L	6.26	6.55	13.9	14	−0.51	−9.66	−3.42	5.44	16.22	16.33
		Nsp6	L3606F	6.6	7.5	63.8	66.9	−17.9	−3.54	0.996	−7.75	53.49	63.11
		RdRp	A4489V	13.9	13.6	21.7	20.6	−13.3	−5.21	4.04	6.04	26.34	35.03
		Spike	Q677H	3.4	4.27	21.5	20.3	−11	0.228	6.12	5.16	20.02	29.95
		N	P13L A152S	−9.64	−9.41	−7.18	−1.17	28.6	23.9	5.53	9.05	17.31	22.37
6	GBRC275b (Gujarat: Surat)	Orf1ab		8.08	7.86	22.6	23.5	22.6	−2.69	5.9	7.59	59.18	36.26
		Exonuclease	W6152R	9.05	9.37	16.9	20	4.29	−14.3	0.83	1.21	31.07	16.28
		Spike	L54F D614G H1088R F1089V	3.4	4.58	21.5	21.1	−11	−10.1	6.12	9.04	20.02	24.62
7	GBRC258b (Gujarat: Gondal)	Orf1ab		8.08	7.65	22.6	22.1	22.6	1.75	5.9	14.1	59.18	45.6
		2′-O-MTase	D6900Y	6.18	5.29	25.6	24.2	10.9	−1.72	0.743	1.14	43.42	28.91
8	GBRC245 (Gujarat: Modasa)	Orf3a	P42R Q57H	−6.16	−5.58	39.4	32.6	0.335	−25.2	−4.11	−5.43	29.46	−3.61
9	GBRC273b (Gujarat: Choryasi)	N	S202 N	−9.64	−8.78	−7.18	−9.07	28.6	27.2	5.53	3.96	17.31	13.31
10	GBRC242 (Gujarat: Himatnagar)	Orf1ab		8.08	7.65	22.6	22.1	22.6	10.7	5.9	12.8	59.18	53.25
		3CLPro	V3475F	15.5	16.2	19.2	21.9	−2.24	2.58	5.22	5.19	37.68	45.87
11	GBRC239 (Gujarat: Gandhinagar)	Orf1ab		8.08	7.86	22.6	24	22.6	14.2	5.9	13.3	59.18	59.36
		Exonuclease	N5928H	9.05	8.75	16.9	19.3	4.29	−15.1	0.83	3.15	31.07	16.1
12	GBRC244 (Gujarat: Ahmadabad)	N	S194L	−9.64	−10.5	−7.18	−6.61	28.6	27.3	5.53	2.35	17.31	12.54

The multi-state kinetic folding mechanism was predicted for both reference and variant proteins.

3.8. Impact of binding specificity to human receptor

The binding specificity of Indian SARS-CoV-2 isolates was studied by predicting molecular interactions between spike protein and human ACE-2 receptor; spike protein and human ACE-2-B^{0}AT complex (Fig. 3; Fig. S3; Tables S1 and S2). Surat (GBRC275b) isolate is found to be effective but its interactions with this complex are predicted to low. The binding affinity of Wuhan-Hu-1 with ACE-2 is somewhat similar to Indian isolates and some variants show a reasonable binding affinity to ACE-2. Spike protein from Kerala (IND/29) isolate has a more binding affinity to ACE-2-B^{0}AT1 complex than that of Kerala (IND/166) isolate. The results of ligand RMSD indicates the binding poses of Telangana and Himachal Pradesh (GBRC97b) isolates are more favorable to ACE-2 and ACE-2-B°AT1 complex. Telangana and Himachal Pradesh (GBRC97b) isolates have shown a considerable binding affinity to ACE-2, reflecting these isolates could increase the host specificity and infectivity.

3.9. Impact of binding specificity to WHO-approved anti-SARS-CoV-2 vaccines

The crystal structures of human neutralizing antibodies were used to dock with spike proteins of Indian isolates for the prediction of their binding specificities to relapse current anti- SARS-CoV-2 vaccines (Fig. 4; Fig. S5; Table S1). The mutation events in the spike protein of these isolates have been shown to destabilize the H-bonding patterns to anti-SARS-CoV-2 vaccines. The binding poses observed for the spike protein of Indian isolates are almost similar to those observed for Wuhan-Hu-1. 80R shows to form more H-bonds with Wuhan-Hu-1 and other vaccines are more interactions with spike protein from Himachal Pradesh (GBRC97b) and Telangana isolates. The binding specificity of the CR3022 antibody to Kerala (IND/166) and Telangana isolates are relatively high to that of Wuhan-Hu-1. CC12.1 shows to have more interactions to Surat (GBRC275b) and Kerala isolates whereas CC12.1 and

Table 2
Comparative analysis of virulence properties of protein variants identified from Indian SARS-CoV-2 isolates.

Genotype	Isolate	Protein	Mutation	Cellular Process		Metabolism		Virulence factors	
				Ref.	Var.	Ref.	Var.	Ref.	Var.
1	IND/29 (Kerala)	Nsp2	I671T	−2.72165	−2.72165	3.304706	3.30470	−3.39956	−3.39956
		PLPro	P2144S	−3.81331	−3.81331	−3.09635	−3.09635	0.135194	0.135194
		RdRp	A4798V	1.227022	1.22702	−2.16893	−2.16893	0.960535	0.960535
		Spike	R407I	1.492872	1.492872	−1.25244	−1.25244	−0.69898	−0.69898
2	IND/166 (Kerala)	Nsp2	I476V	−2.72165	−3.87946	3.304706	3.62299	−3.39956	−4.20081
		Helicase	T5538I	−0.54011	−0.44787	0.931417	1.10038	1.304757	0.92286
		Spike	A930V	1.492872	1.246268	−1.25244	−0.7755	−0.69898	−0.99728
		Orf8	L84S	0.947547	0.947547	0.551035	0.551035	−0.30211	−0.30211
3	GMCKN318 (Telangana: Hyderabad)	RdRp	P4715L	1.227022	1.227022	−2.16893	−2.16893	0.960535	0.960535
		Spike	D614G	1.492872	1.492872	−1.25244	−1.25244	−0.69898	−0.69898
4	GBRC97b (Himachal Pradesh: Una)	Nsp2	L204F	−2.72165	−2.0011	3.304706	3.233755	−3.39956	−3.00631
		PLPro	A1812D	−3.81331	−3.81331	−3.09635	−3.09635	0.135194	0.135194
		Nsp6	L3606F	1.988548	1.988548	−0.29145	−0.29145	−0.91113	−0.91113
		RdRp	K4451 N P4715L	1.227022	1.227022	−2.16893	−2.16893	0.960535	0.960535
		Helicase	P5624L	−0.54011	−0.32021	0.931417	0.974854	1.304757	1.244815
		Exonuclease	K6274 N	0.205947	0.205947	−1.91813	−1.91813	1.812892	1.812892
		Spike	D614G	1.492872	1.492872	−1.25244	−1.25244	−0.69898	−0.69898
		N	R203K G204R	0.766355	0.766355	0.788503	0.788503	−0.85619	−0.85619
5	GBRC98 (Himachal Pradesh: Una)	Nsp2	G519S	−2.72165	−2.72165	3.304706	3.304706	−3.39956	−3.39956
		PLPro	S2015R T2016K P2739L	−3.81331	−3.81331	−3.09635	−3.09635	0.135194	0.135194
		Nsp6	L3606F	1.988548	1.988548	−0.29145	−0.29145	−0.91113	−0.91113
		RdRp	A4489V	1.227022	1.488159	−2.16893	−2.84141	0.960535	0.90483
		Spike	Q677H	1.492872	1.492872	−1.25244	−1.25244	−0.69898	−0.69898
		N	P13L A152S	0.766355	1.145885	0.788503	0.258568	−0.85619	−0.76989
6	GBRC275b (Gujarat: Surat)	Exonuclease	W6152R	0.205947	0.205947	−1.91813	−1.91813	1.812892	1.812892
		Spike	L54F D614G H1088R F1089V	1.492872	1.492872	−1.25244	−1.25244	−0.69898	−0.69898
7	GBRC258b (Gujarat: Gondal)	2′-O-MTase	D6900Y	1.464238	1.464238	−0.14461	−0.14461	−0.21774	−0.21774
8	GBRC245 (Gujarat: Modasa)	Orf3a	P42R Q57H	1.952861	1.952861	0.104934	0.104934	−0.83548	−0.83548
9	GBRC273b (Gujarat: Choryasi)	N	S202 N	0.766355	0.766355	0.788503	0.788503	−0.85619	−0.85619
10	GBRC242 (Gujarat: Himatnagar)	3CLPro	V3475F	0.521603	0.534629	−0.90669	−1.33277	1.521026	1.458814
11	GBRC239 (Gujarat: Gandhinagar)	Exonuclease	N5928H	0.205947	0.205947	−1.91813	−1.91813	1.812892	1.812892
12	GBRC244 (Gujarat: Ahmadabad)	N	S194L	0.766355	1.362327	0.788503	0.237112	−0.85619	−0.82981

The green shadow indicates the predicted property of each protein.

m396 had a maximum binding affinity to a reference strain.

Mutations in many Indian isolates did not influence the binding affinities of human neutralizing antibodies, apart from Surat isolate. It suggests current anti-SARS-CoV-2 vaccines are still effective against understudied mutations in Indian isolates. CR3022 antibody forms 9 H-bonds with spike protein of Surat isolate where Val1121 additionally interacted with Thr20 from its light chain (Table S2). Val1121 residue also forms H-bond with Arg18 from its light chain. Compared to Wuhan-Hu-1, the spike protein of Surat isolate destabilizes H-bonding between Ile28 from a light chain and Tyr1047 from spike protein. 80R and m396 antibodies are not interacted with spike protein to neutralize Surat isolate due to destabilized H-bonding patterns. A new vaccine therapeutic intervention is thus needed for Surat isolate as it harbors a potential mutation (P1089V) in its spike protein.

4. Discussion

RNA viruses evolve rapidly because their genomes can mutate one million times higher than the host genome. The genetic diversity of RNA viruses is determined by selective pressure on viral and host-dependent processes [59]). Accordingly, a phylogenetic disparity was found across the Gujarat and Kerala isolates. The individuals from Kerala states have acquired SARS-CoV-2 at least from two sources [58]. Our study revealed a phylogenetic discrepancy between two Kerala isolates as both isolates have typical mutations in spike (R407I/A930V; Nsp2: I671T/I476V), PLPro, RdRp, helicase, and Orf8. The mutational events might have reflected on the genetic diversity of Kerala isolates. Similarly, we found a phylogenetic discrepancy within Surat and Himachal Pradesh isolates. Gujarat and Kerala SARS-CoV-2 isolates are not closely related to isolates from other parts of the world and within other hotspots in India. Phylogenomic data of this study suggested that these isolates are capable of high transmission and virulence to individuals in hotspot regions of India. Also, it pointed out that current antiviral therapeutics might not work well in India as a result of increasing deleterious mutational events and genomic variations in Indian isolates.

[11] reported 353, 341 gene mutation events, and 10, 066 protein mutations from 48,635 SARS-CoV-2 genomes [11]. Spike-D614G, Nsp3-G251V, Nsp3: F106F, Nsp12b: P314L, ORF3a: Q57H, ORF8-L84S, N: RG203KR, and 5′UTR:241 were the most frequent mutations in global SARS-CoV-2 isolates similar to earlier analyses [8,11]. Spike: A23403G and RdRp: C14408T identified as the dominant mutations in Gujarat isolate, which have shown to decrease patients at a frequency of 97.67–95.35% [9]. D614G is a frequent and important mutation identified in global and Indian SARS-CoV-2 isolates. It is responsible for the initial entry of this virus into the host cell *via* ACE-2 [60]. It can bring a potential change in the structure-function-virulence link of spike protein even the mutated position located outside the receptor-binding domain. It may be resulted due to the replacement of acidic and negatively charged aspartic acid with glycine [61,62].

Interestingly, Surat (GBRC275b) and Gandhinagar (GBRC239)

Table 3
Comparative analysis of binding specificity of protein variants from Indian SARS-CoV-2 isolates with FDA-approved antiviral drugs.

Genotype	Isolate	Gene	Mutant	Drug name	DrugBank/ PubChem	PMID	Score		Area		ACE	
							Ref.	Var.	Ref.	Var.	Ref.	Var.
1	IND/29 (Kerala)	Nsp2	I671T	Ribostamycin sulfate	DB03615	31443266	5334	5334	687.7	687.7	−72.19	−72.19
				E−64	123985	31443266	4606	4606	537.8	537.8	−286.28	−286.28
		PLPro	P2144S	Simeprevir	DB06290	18678486	6956	6956	944.4	944.4	−306.29	−306.29
		RdRp	A4798V	Remdesivir	DB14761	32020029	6472	6460	862.6	858.6	−297.68	−298.62
				Hydroxychloroquine	DB01611	32020029	5546	5602	642.2	646.5	−200.07	−34.74
		Spike	R407I	Emodin	DB07715	32194980	4022	4242	507.4	518.8	−159.29	−169.22
2	IND/166 (Kerala)	Nsp2	I476V	Ribostamycin sulfate	DB03615	31443266	5334	5334	687.7	687.7	−72.19	−72.19
				E−64	123985	31443266	4606	4606	537.8	537.8	−286.28	−286.28
		Helicase	T5538I	Artenimol	DB11638	32696720	4102	4190	507.9	515.1	−250.26	−77.86
		Spike	A930V	Emodin	DB07715	32194980	4022	4220	507.4	532.9	−159.29	−157.1
		Orf8	L84S	Azacitidine	DB00928	15899874	3324	3130	367.6	361	−165.95	−51.28
3	GMCKN318 (Telangana: Hyderabad)	RdRp	P4715L	Remdesivir	DB14761	32020029	6472	6266	862.6	829.2	−297.68	−272.29
				Hydroxychloroquine	DB01611	32020029	5546	5568	642.2	665.9	−200.07	−198.71
		Spike	D614G	Emodin	DB07715	32194980	4022	4042	507.4	515.2	−159.29	−150.58
4	GBRC97b (Himachal Pradesh: Una)	Nsp2	L204F	Ribostamycin sulfate	DB03615	31443266	5334	5334	687.7	687.7	−72.19	−72.19
				E−64	123985	31443266	4606	4606	537.8	537.8	−286.28	−286.28
		PLPro	A1812D	Simeprevir	DB06290	18678486	6956	7080	944.4	884	−306.29	−259.49
		Nsp6	L3606F	Loratadine	DB00455	25318072	3896	3896	518.3	518.3	−216.58	−216.58
		RdRp	K4451 N	Remdesivir	DB14761	32020029	6472	6308	862.6	849	−297.68	−349.93
			P4715L	Hydroxychloroquine	DB01611	32020029	5546	5430	642.2	635	−200.07	−28.24
		Helicase	P5624L	Artenimol	DB11638	32696720	4102	4228	507.9	529.5	−250.26	−113.65
		Exonuclease	K6274 N	2-(N-Morpholino) ethanesulfonate	4599203	32470577	3074	3152	343.4	352.6	−112.57	−112.60
				Migalastat	DB05018	32353859	2714	2714	311.1	311.5	−81.75	−88.36
				Mycophenolic acid	DB01024	32353859	5338	5264	605	595.8	−182.69	−159.11
				Ribavirin	DB00811	32353859	3938	3784	439.8	457.7	−85.04	−153.66
		Spike	D614G	Emodin	DB07715	32194980	4022	4042	507.4	515.2	−159.29	−150.58
		N	R203K G204R	Silmitasertib	DB15408	21159648	4516	4514	487.8	582.6	−170.62	−330.15
5	GBRC98 (Himachal Pradesh: Una)	Nsp2	G519S	Ribostamycin sulfate	DB03615	31443266	5334	5334	687.7	687.7	−72.19	−72.19
				E−64	123985	31443266	4606	4606	537.8	537.8	−286.28	−286.28
		PLPro	S2015R T2016K P2739L	Simeprevir	DB06290	18678486	6956	6956	944.4	944.4	−306.29	−306.29
		Nsp6	L3606F	Loratadine	DB00455	25318072	3896	3896	518.3	518.3	−216.58	−216.58
		RdRp	A4489V	Remdesivir	DB14761	32020029	6472	6572	862.6	808.8	−297.68	−0.23
				Hydroxychloroquine	DB01611	32020029	5546	5534	642.2	629.9	−200.07	−29.61
		Spike	Q677H	Emodin	DB07715	32194980	4022	4154	507.4	518.2	−159.29	−150.90
		N	P13L A152S	Silmitasertib	DB15408	21159648	4516	4514	487.8	582.6	−170.62	−330.15
6	GBRC275b (Gujarat: Surat)	Exonuclease	W6152R	2-(N-Morpholino) ethanesulfonate	4599203	32470577	3074	2982	343.4	324	−112.57	−108.46
				Migalastat	DB05018	32353859	2714	2608	311.1	298.4	−81.75	−43.19
				Mycophenolic acid	DB01024	32353859	5338	4426	605	514.2	−182.69	−159.56
				Ribavirin	DB00811	32353859	3938	3452	439.8	414.7	−85.04	−150.4
		Spike	L54F D614G H1088R F1089V	Emodin	DB07715	32194980	4022	4170	507.4	505.3	−159.29	−166.39
7	GBRC258b (Gujarat: Gondal)	2′-O-MTase	D6900Y	Sinefungin	DB01910	17139284	1168	1074	134.6	111.7	−350.22	−377.34
8	GBRC245 (Gujarat: Modasa)	Orf3a	P42R Q57H	Tranilast	DB07615	32190290	4992	5098	583	580.9	−124.1	−120.40
9	GBRC273b (Gujarat: Choryasi)	N	S202 N	Silmitasertib	DB15408	21159648	4516	4514	487.8	582.6	−170.62	−330.15
10	GBRC242 (Gujarat: Himatnagar)	3CLPro	V3475F	Raltegravir	DB06817	32266873	5840	5786	724.2	723.2	−214.91	−205.25
11	GBRC239 (Gujarat: Gandhinagar)	Exonuclease	N5928H	2-(N-Morpholino) ethanesulfonate	4599203	32470577	3074	3084	343.4	349.9	−112.57	−110.08
				Migalastat	DB05018	32353859	2714	2792	311.1	325.9	−81.75	−110.33
				Mycophenolic acid	DB01024	32353859	5338	5328	605	605.5	−182.69	−181.98
				Ribavirin	DB00811	32353859	3938	3712	439.8	444.5	−85.04	−130.46
12	GBRC244 (Gujarat: Ahmadabad)	N	S194L	Silmitasertib	DB15408	21159648	4516	4514	487.8	582.6	−170.62	−330.15

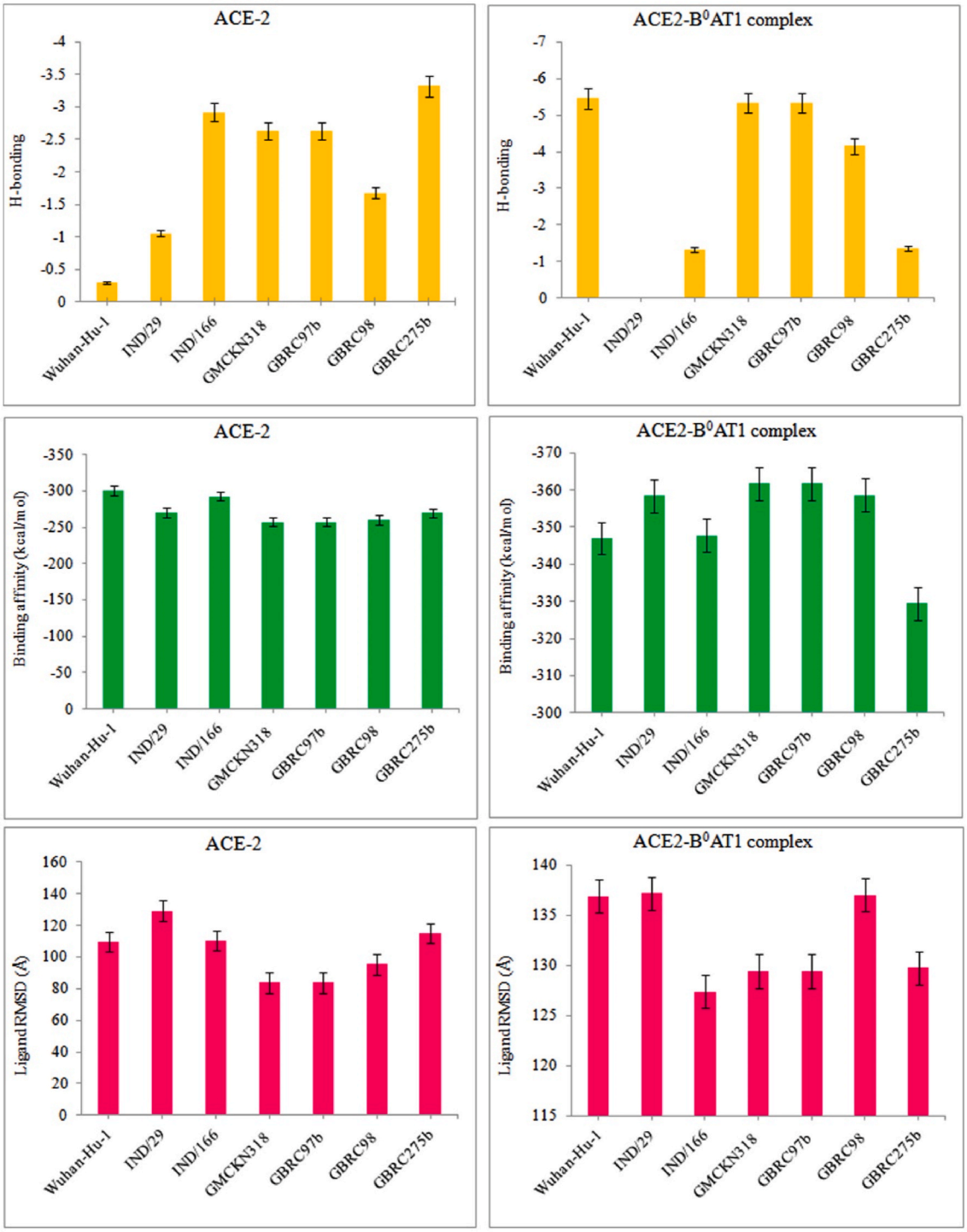


Fig. 3. Comparative analysis of binding specificities of protein variants from Indian SARS-CoV-2 isolates with a human receptor (ACE-2).

isolates have two unique mutations W6152R and N5928H in exonuclease, respectively. Mutations in this enzyme can confer resistance to ribavirin and other nucleotide analogs *via* an error-correcting mechanism [63]). This enzyme also interacts with RdRp and associated Nsp7 and Nsp8 to perform proofreading activity [64]. It is a particularly challenging task to current antivirals targeting not only exonuclease and also RdRp in Surat and Gandhinagar isolates, as reported earlier on other coronaviruses [65].

The folding rate is a molecular measure to understand the sequence-structure-function- virulence link of a protein [18–27]. The population of properly folded viral proteins is directly proportional to the level of viral spread and the lethality of the infection [32,66]. However, the kinetic rate of folding directly relies on the types of linear amino acids in a protein, the reduced hydrophobic density, and enhanced polar-polar repulsive interactions [31]. The substituted branched-chain amino acids were shown to interfere with the folding rate, shape, or conformation of a protein, which agreed to the previous studies [28,30,67]; Liji, 2020; [29]. We observed altered drug specificity in the spike protein of Surat isolates when glycine was replaced with valine similar to West Bengal isolates [68]. G14S mutation can influence the catalytic activity and folding rate of PLPro worldwide [69]. Our study found P2144S, A1812D, and S2015R mutations in PLPro affecting the structural

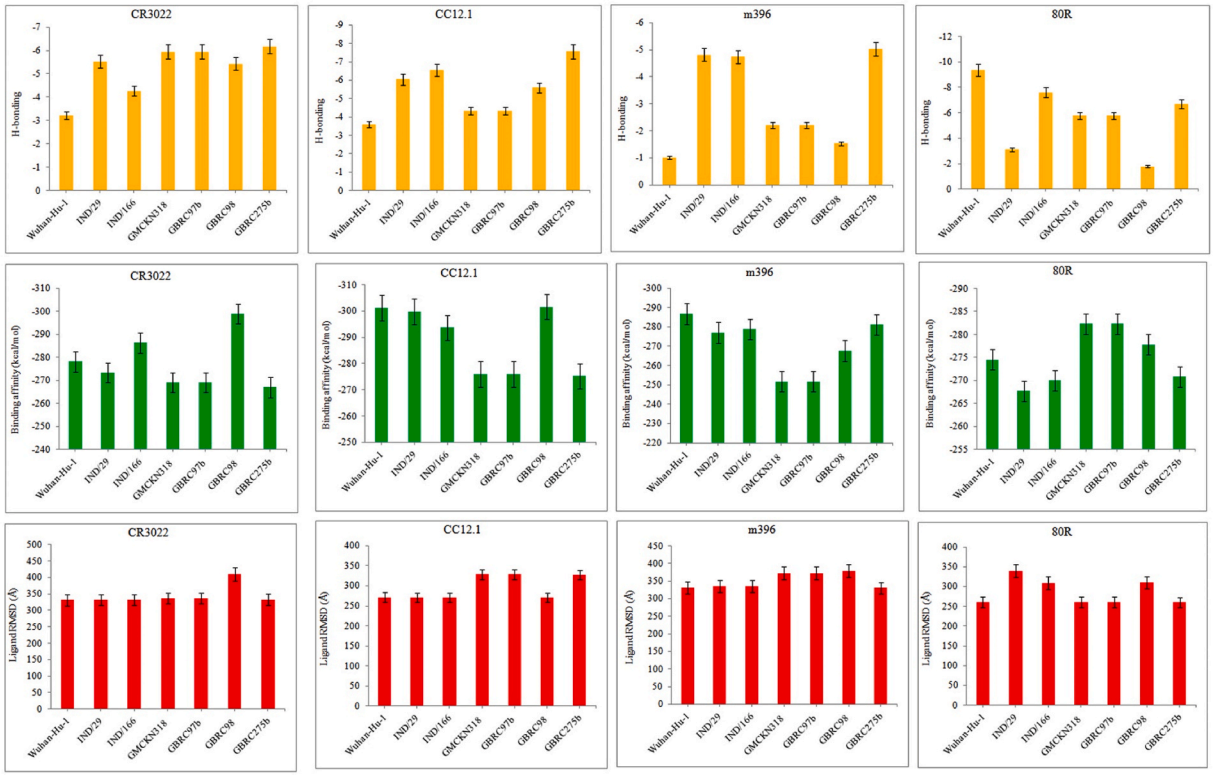


Fig. 4. Comparative analysis of binding specificities of protein variants from Indian SARS-CoV-2 isolates with anti-SARS-CoV-2 antibodies.

characteristics of this enzyme in Kerala (IND/29) and Himachal Pradesh isolates. HIV relies on a host-encoded peptidyl-prolyl cis-trans isomerase and cyclophilin A for the maturation of its coat proteins [70]. SARS-CoV-2 proteins similarly mediate folding and maturation processes with aid of host chaperones to hijack from many antivirals interfaces.

Virulence is tightly coupled to virus load in the host leading to an increase in the transmission of this virus [71]. The virulence and metabolic properties of SARS-CoV-2 increased at a fast-folding rate of a virus as well as host proteins. The peptidyl-prolyl cis-trans isomerase is a protein-folding catalyst that may process virulence determinants of SARS-CoV-2 for cell-to-cell spread [72] and antiviral resistance development [73]. Mutations can assist in the understanding mechanism of pathogenesis and virulence of SARS-CoV-2. Interestingly, a new functional shift from virulence to the metabolic process was predicted when a neutral polar threonine was exchanged with hydrophobic isoleucine in helicase of Kerala (IND/166) isolate. It is known that pathogens produce macromolecules to scavenge for nutrients to feed metabolism boosts virulence [74–76]. According to the 'survival of the fittest' theories of evolution, virulence, and metabolism of SARS-CoV-2 might be intimately linked to survival in a certain niche, particularly in host respiratory tracts.

The RdRp is vital for the replication machinery of RNA viruses. The P4715L mutation was located in a hydrophobic cleft of this protein providing the ability to hijack from remdesivir by the earlier work [77]. Our study also detected consecutive alterations of R203K and G204R in the viral nucleocapsid protein. It was found to bring a change in the local structural environment of this protein to escape from antivirals [78]. The ACE-2-B^{0}AT1 complex is assembled with a collectrin-like domain of ACE-2 that is recognized by a receptor-binding domain of spike glycoprotein in SARS-CoV-2 [6]). Mutations on this complex or ACE-2 alone can alter the binding specificity of viral surface protein. Telangana and Himachal Pradesh (GBRC97b) isolates have shown increased host specificity and infectivity as the spike protein them are strongly bound with the ACE-2 receptor. Genome variation analysis found a rare mutation, P1089V in the spike protein of Surat (GBRC275b) isolate, reflecting its clinical importance in COVID-19 patients. Some potential mutations (R407I, D614G, and F1089V) also identified in spike protein that can generate new antigenic variants to evade host immune surveillance. The results of our study revealed that the substitutions of branched amino acids in a protein have changed the local structural environments, folding rate of mixed structural classes, and virulence. The evolutionary constraints in the mutated proteins from Indian isolates might have influenced the sequence-structure-function-virulence link leading to the failure of antiviral drugs and vaccines. It was agreed to the previous analyses [79–82]. Accordingly, our study suggested a repurposing antiviral therapeutics targeting new antigenic variants.

5. Conclusions

The structural and functional characteristics of frequent protein variants in global and Indian SARS-CoV-2 isolates can help for a better understanding of the infection biology of this virus. Genomic variance analysis suggested the development of new diagnostic probes for the detection of emerging genotypes that arose upon frequent mutations after SARS-CoV-2 confirmation. Our study described the effect of mutations on the protein folding rate, virulence, and binding specificity to drugs, vaccines, and ACE-2. The folding rate is a key measure to determine a sequence-structure-function-virulence link of the SARS-CoV-2 proteome, which in turn increases the viral load in a host *via* the rapid copying of matured (folded) proteins. The new mutations identified in a proofreading exonuclease can alter the drug specificity of not only this enzyme but also the RdRp complex, suggesting the failure of ribavirin and remdesivir to Indian isolates. Mutations in antigenic determinants of spike protein confer the escaping ability to Surat isolate from human neutralizing antibodies. Most antivirals and vaccines are still more effective for Indian isolates. However, emodin and artenimol are shown better molecular interactions with the respective targets, suggesting repurposable antivirals for the treatment of COVID-19 in India. Finally,

we conclude that further *in vitro* and *in vivo* studies are required to evaluate the efficacy of current antiviral therapies on SARS-CoV-2 infectivity.

Declaration of competing interest

The authors confirm that this article's content has no Competing interest.

Acknowledgments

The author would like to thank the Science and Engineering Research Board, Department of Science and Technology (MSC/2020/000438), New Delhi, India, for financial support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2021.104315.

Authors' contributions

CP conceived of the presented idea and verified the analytical methods and manuscript. PR collected the data and performed the analysis. SS contributed to data collection and analysis. BM contributed to the analysis. All authors discussed the results and contributed to the final manuscript.

Data availability

Supplementary file 1.

Data set used for constructing phylogenetic trees of global and Indian isolates of SARS-CoV-2.

Supplementary file 2.

List of genomic variations identified from Indian SARS-CoV-2 isolates.

Ethical declaration

The present study does not contain any studies with human participants or animals performed by any of the authors.

For this type of study formal consent is not required.

References

[1] F. Wu, S. Zhao, B. Yu, Y.M. Chen, W. Wang, Z.G. Song, Y. Hu, Z.W. Tao, J.H. Tian, Y.Y. Pei, M.L. Yuan, Y.L. Zhang, F.H. Dai, Y. Liu, Q.M. Wang, J.J. Zheng, L. Xu, E. C. Holmes, Y.Z. Zhang, A new coronavirus associated with human respiratory disease in China, Nature 579 (2020) 265–269.

[2] R. Ralph, J. Lew, T. Zeng, M. Francis, B. Xue, M. Roux, A. Toloue Ostadgavahi, S. Rubino, N.J. Dawe, M.N. Al-Ahdal, D.J. Kelvin, C.D. Richardson, J. Kindrachuk, D. Falzarano, A.A. Kelvin, 2019-nCoV (Wuhan virus), a novel coronavirus: human-to-human transmission, travel-related cases, and vaccine readiness, J Infect Dev Ctries 14 (2020) 3–17.

[3] J.H. Beigel, K.M. Tomashek, L.E. Dodd, A.K. Mehta, B.S. Zingman, A.C. Kalil, E. Hohmann, H.Y. Chu, A. Luetkemeyer, S. Kline, D. Lopez de Castilla, R. W. Finberg, K. Dierberg, V. Tapson, L. Hsieh, T.F. Patterson, R. Paredes, D. A. Sweeney, W.R. Short, G. Touloumi, D.C. Lye, N. Ohmagari, M.D. Oh, G.M. Ruiz-Palacios, T. Benfield, G. Fätkenheuer, M.G. Kortepeter, R.L. Atmar, C.B. Creech, J. Lundgren, A.G. Babiker, S. Pett, J.D. Neaton, T.H. Burgess, T. Bonnett, M. Green, M. Makowski, A. Osinusi, S. Nayak, H.C. Lane, Remdesivir for the treatment of Covid-19 - preliminary report, N. Engl. J. Med. 22 (2020), NEJMoa2007764.

[4] D.E. Gordon, G.M. Jang, M. Bouhaddou, J. Xu, K. Obernier, K.M. White, M. J. O'Meara, V.V. Rezelj, J.Z. Guo, D.L. Swaney, T.A. Tummino, R. Hüttenhain, R. M. Kaake, A.L. Richards, B. Tutuncuoglu, H. Foussard, J. Batra, K. Haas, M. Modak, M. Kim, P. Haas, B.J. Polacco, H. Braberg, J.M. Fabius, M. Eckhardt, M. Soucheray, M.J. Bennett, M. Cakir, M.J. McGregor, Q. Li, B. Meyer, F. Roesch, T. Vallet, A. Mac Kain, L. Miorin, E. Moreno, Z.Z.C. Naing, Y. Zhou, S. Peng, Y. Shi, Z. Zhang, W. Shen, I.T. Kirby, J.E. Melnyk, J.S. Chorba, K. Lou, S.A. Dai, I. Barrio-Hernandez, D. Memon, C. Hernandez-Armenta, J. Lyu, C.J.P. Mathy, T. Perica, K.B. Pilla, S. J. Ganesan, D.J. Saltzberg, R. Rakesh, X. Liu, S.B. Rosenthal, L. Calviello, S. Venkataramanan, J. Liboy-Lugo, Y. Lin, X.P. Huang, Y. Liu, S.A. Wankowicz, M. Bohn, M. Safari, F.S. Ugur, C. Koh, N.S. Savar, Q.D. Tran, D. Shengjuler, S. J. Fletcher, M.C. O'Neal, Y. Cai, J.C.J. Chang, D.J. Broadhurst, S. Klippsten, P. P. Sharp, N.A. Wenzell, D. Kuzuoglu-Ozturk, H.Y. Wang, R. Trenker, J.M. Young, D. A. Cavero, J. Hiatt, T.L. Roth, U. Rathore, A. Subramanian, J. Noack, M. Hubert, R. M. Stroud, A.D. Frankel, O.S. Rosenberg, K.A. Verba, D.A. Agard, M. Ott, M. Emerman, N. Jura, M. von Zastrow, E. Verdin, A. Ashworth, O. Schwartz, C. d'Enfert, S. Mukherjee, M. Jacobson, H.S. Malik, D.G. Fujimori, T. Ideker, C. S. Craik, S.N. Floor, J.S. Fraser, J.D. Gross, A. Sali, B.L. Roth, D. Ruggero, J. Taunton, T. Kortemme, P. Beltrao, M. Vignuzzi, A. García-Sastre, K.M. Shokat, B. K. Shoichet, N.J. Krogan, A SARS-CoV-2 protein interaction map reveals targets for drug repurposing, Nature 583 (2020) 459–468.

[5] P. Chellapandi, S. Saranya, Genomics insights of SARS-CoV-2 (COVID-19) into target-based drug discovery, Med. Chem. Res. 1–15 (2020).

[6] R. Yan, Y. Zhang, Y. Li, L. Xia, Y. Guo, Q. Zhou, Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2, Science 367 (2020) 1444–1448.

[7] J. Díaz, SARS-CoV-2 molecular network structure, Front. Physiol. 11 (2020) 870.

[8] P. Forster, L. Forster, C. Renfrew, M. Forster, Phylogenetic network analysis of SARS-CoV-2 genomes, Proc. Natl. Acad. Sci. U. S. A. 117 (2020) 9241–9243.

[9] M. Joshi, A. Puvar, D. Kumar, A. Ansari, M. Pandya, J. Raval, Z. Patel, P. Trivdi, M. Gandhi, L. Pandya, K. Patel, N. Savaliya, S. Bagatharia, S. Kumar, C. Joshi, Genomic variations in SARS-CoV-2 genomes from Gujarat: underlying role of variants in disease epidemiology, bioRxiv Preprint (2020), https://doi.org/10.1101/2020.07.10.197095.

[10] R. Kumar, H. Verma, N. Singhvi, U. Sood, V. Gupta, M. Singh, R. Kumari, P. Hira, S. Nagar, C. Talwar, N. Nayyar, S. Anand, C.D. Rawat, M. Verma, R.K. Negi, Y. Singh, R. Lal, Comparative genomic analysis of rapidly evolving SARS-CoV-2 reveals mosaic pattern of phylogeographical distribution, mSystems 5 (2020) e00505–e00520.

[11] D. Mercatelli, F.M. Giorgi, Geographic and genomic distribution of SARS-CoV-2 mutations, Front. Microbiol. 11 (2020) 1800.

[12] D.R. Chopera, Z. Woodman, K. Mlisana, M. Mlotshwa, D.P. Martin, C. Seoighe, F. Treurnicht, D.A. de Rosa, W. Hide, S.A. Karim, C.M. Gray, C. Williamson, Transmission of HIV-1 CTL escape variants provides HLA-mismatched recipients with a survival advantage, PLoS Pathog. 4 (2008), e1000033.

[13] L. Bracq, M. Xie, S. Benichou, J. Bouchet, Mechanisms for cell-to-cell transmission of HIV-1, Front. Immunol. 9 (2018) 260.

[14] C.Z. Buffalo, Y. Iwamoto, J.H. Hurley, X. Ren, How HIV Nef proteins hijack membrane traffic to promote infection, J. Virol. 93 (2019) e01322-19.

[15] M.P. Robertson, H. Igel, R. Baertsch, D. Haussler, M. Ares, W.G. Scott, The structure of a rigorously conserved RNA element within the SARS virus genome, PLoS Biol. 3 (2005) e5.

[16] F. Reggiori, I. Monastyrska, M.H. Verheije, T. Calì, M. Ulasli, S. Bianchi, R. Bernasconi, C.A. de Haan, M. Molinari, Coronaviruses hijack the LC3-I-positive EDEMosomes, ER-derived vesicles exporting short-lived ERAD regulators, for replication, Cell Host Microbe 7 (2010) 500–508.

[17] Y. Hu, B.A. Desimmie, H.C. Nguyen, S.J. Ziegler, T.C. Cheng, J. Chen, J. Wang, H. Wang, K. Zhang, V.K. Pathak, Y. Xiong, Structural basis of antagonism of human APOBEC3F by HIV-1 Vif, Nat. Struct. Mol. Biol. 26 (2019) 1176–1183.

[18] P. Chellapandi, Structural-functional integrity of hypothetical proteins identical to ADP-ribosylation Superfamily upon point mutations, Protein Pept. Lett. 21 (2014) 722–735.

[19] P. Chellapandi, S. Sakthi Shree, M. Bharathi, Phylogenetic approach for inferring the origin and functional evolution of bacterial ADP-ribosylation superfamily, Protein Pept. Lett. 20 (2013) 1054–1065.

[20] P. Chellapandi, R. Prathiviraj, A. Prisilla, Deciphering structure, function and mechanism of Plasmodium IspD homologs from their evolutionary imprints, J. Comput. Aided Mol. Des. 33 (2019) 419–436.

[21] P. Chellapandi, R. Prathiviraj, A. Prisilla, Molecular evolution and functional divergence of IspD homologs in malarial parasites, Infect. Genet. Evol. 65 (2018) 340–349.

[22] A. Prisilla, R. Prathiviraj, R. Sasikala, P. Chellapandi, Structural constraints-based evaluation of immunogenic avirulent toxins from *Clostridium botulinum* C2 and C3 toxins as subunit vaccines, Infect. Genet. Evol. 44 (2016) 17–27.

[23] A. Prisilla, R. Prathiviraj, P. Chellapandi, Molecular evolutionary constraints that determine the avirulence state of *Clostridium botulinum* C2 toxin, J. Mol. Evol. 86 (2017) 174–186.

[24] A. Prisilla, P. Chellapandi, Cloning and expression of immunogenic *Clostridium botulinum* C2I mutant proteins designed from its evolutionary imprints, Comp. Immunol. Microbiol. Infect. Dis. 65 (2019) 207–212.

[25] R. Prathiviraj, A. Prisilla, P. Chellapandi, Structure-function discrepancy in *Clostridium botulinum* C3 toxin for its rational prioritization as a subunit vaccine, J. Biomol. Struct. Dyn. 34 (2015) 1317–1329.

[26] R. Prathiviraj, P. Chellapandi, Evolutionary genetic analysis of unassigned peptidase clan-associated microbial virulence and pathogenesis, Biologia (2020), https://doi.org/10.2478/s11756-020-00529-4.

[27] R. Prathiviraj, P. Chellapandi, Deciphering molecular virulence mechanism of *Mycobacterium tuberculosis* Dop isopeptidase based on its sequence-structure-function link, Protein J. 39 (2020) 33–45.

[28] C.S. Wylie, E.I. Shakhnovich, A biophysical protein folding model accounts for most mutational fitness effects in viruses, Proc. Natl. Acad. Sci. U.S.A. 108 (2011) 9916–9921.

[29] W.M. Ng, A.J. Stelfox, T.A. Bowden, Unraveling virus relationships by structure-based phylogenetic classification, Virus Evol 6 (2020) veaa003.

[30] N. Chéron, A.W. Serohijos, J.M. Choi, E.I. Shakhnovich, Evolutionary dynamics of viral escape under antibodies stress: a biophysical model, Protein Sci. 25 (2016) 1332–1340.

[31] B. Khanppnavar, A. Roy, K. Chandra, V.N. Uversky, N.C. Maiti, S. Datta, Deciphering the structural intricacy in virulence effectors for proton-motive force mediated unfolding in type-III protein secretion, Int. J. Biol. Macromol. 159 (2020) 18–33.

[32] D.J. Opstelten, P. de Groote, M.C. Horzinek, P.J. Rottier, Folding of the mouse hepatitis virus spike protein and its association with the membrane protein, Arch. Virol. Suppl. 9 (1994) 319–328.

[33] L. Denzer, H. Schroten, C. Schwerk, From gene to protein-how bacterial virulence factors manipulate host gene expression during infection, Int. J. Mol. Sci. 21 (2020) 3730.

[34] M. Shakya, S.A. Ahmed, K.W. Davenport, M.C. Flynn, C.C. Lo, P.S.G. Chain, Standardized phylogenetic and molecular evolutionary analysis applied to species across the microbial tree of life, Sci. Rep. 10 (2020) 1723.

[35] K. Katoh, J. Rozewicki, K.D. Yamada, MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization, Briefings Bioinf. 20 (2019) 1160–1166.

[36] D. Schneidman-Duhovny, Y. Inbar, R. Nussinov, H.J. Wolfson, PatchDock and SymmDock: servers for rigid and symmetric docking, Nucleic Acids Res. 33 (2005) W363–W367.

[37] T. Mailund, C.N. Pedersen, QuickJoin-fast neighbour-joining tree reconstruction, Bioinformatics 20 (2004) 3261–3262.

[38] T.H. Jukes, C.R. Cantor, Evolution of Protein Molecules, Academic Press, New York, 1969, pp. 21–132.

[39] I. Letunic, P. Bork, Interactive Tree of Life (iTOL) v4: recent updates and new developments, Nucleic Acids Res. 47 (2019). W256-W259.

[40] R. Bouckaert, T.G. Vaughan, J. Barido-Sottani, S. Duchêne, M. Fourment, A. Gavryushkina, J. Heled, G. Jones, D. Kühnert, N. De Maio, M. Matschiner, F. K. Mendes, N.F. Müller, H.A. Ogilvie, L. du Plessis, A. Popinga, A. Rambaut, D. Rasmussen, I. Siveroni, M.A. Suchard, C.H. Wu, D. Xie, C. Zhang, T. Stadler, A. J. Drummond, Beast 2.5: an advanced software platform for Bayesian evolutionary analysis, PLoS Comput. Biol. 15 (2019), e1006650.

[41] T.A. Hall, BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT, Nucleic Acids Symp. Ser. 41 (1999) 95–98.

[42] S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D. J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.

[43] Y. Xing, X. Li, X. Gao, Q. Dong, MicroGMT: a mutation tracker for SARS-CoV-2 and other microbial genome sequences, Front. Microbiol. 11 (2020) 1502.

[44] A.G. de Brevern, A. Bornot, P. Craveur, C. Etchebest, J.C. Gelly, PredyFlexy: flexibility and local structure prediction from sequence, Nucleic Acids Res. 40 (2012) W317–W322.

[45] M.M. Gromiha, A.M. Thangakani, S. Selvaraj, FOLD-RATE: prediction of protein folding rates from amino acid sequence, Nucleic Acids Res. 34 (2006) W70–W74.

[46] E. Capriotti, R. Casadio, K-Fold: a tool for the prediction of the protein folding kinetic order and rate, Bioinformatics 23 (2007) 385–386.

[47] S. Saha, G.P.S. Raghava, VICMpred: SVM-based method for the prediction of functional proteins of gram-negative bacteria using amino acid patterns and composition, Dev. Reprod. Biol. 4 (2006) 42–47.

[48] D.S. Wishart, Y.D. Feunang, A.C. Guo, E.J. Lo, A. Marcu, J.R. Grant, T. Sajed, D. Johnson, C. Li, Z. Sayeeda, N. Assempour, I. Iynkkaran, Y. Liu, A. Maciejewski, N. Gale, A. Wilson, L. Chin, R. Cummings, D. Le, A. Pon, C. Knox, M. Wilson, DrugBank 5.0: a major update to the DrugBank database for 2018, Nucleic Acids Res. 46 (2018) D1074–D1082.

[49] M. Yuan, H. Liu, N.C. Wu, C.D. Lee, X. Zhu, F. Zhao, D. Huang, W. Yu, Y. Hua, H. Tien, T.F. Rogers, E. Landais, D. Sok, J.G. Jardine, D.R. Burton, I.A. Wilson, Structural basis of a public antibody response to SARS-CoV-2. bioRxiv Preprint. https://doi.org/10.1101/2020.06.08.141267, 2020.

[50] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T.A.P. de Beer, C. Rempfer, L. Bordoli, R. Lepore, T. Schwede, SWISS-MODEL: homology modelling of protein structures and complexes, Nucleic Acids Res. 46 (2018) W296–W303.

[51] W. Zheng, C. Zhang, E.W. Bell, Y. Zhang, I-TASSER gateway: a protein structure and function prediction server powered by XSEDE, Future Generat. Comput. Syst. 99 (2019) 73–85.

[52] D. Bhattacharya, refineD: improved protein structure refinement using machine learning based restrained relaxation, Bioinformatics 35 (2019) 3320–3328.

[53] E. Mashiach, D. Schneidman-Duhovny, N. Andrusier, R. Nussinov, H.J. Wolfson, FireDock: a web server for fast interaction refinement in molecular docking, Nucleic Acids Res. 36 (2008) W229–W232.

[54] S.-Y. Huang, X. Zou, An iterative knowledge-based scoring function for protein-protein recognition, Proteins 72 (2008) 557–579.

[55] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, J. Comput. Chem. 31 (2010) 455–461.

[56] S. Yuan, H.C.S. Chan, S. Filipek, H. Vogel, PyMOL and Inkscape bridge the data and the data visualization, Structure 24 (2016) 2041–2042.

[57] R.A. Laskowski, M.B. Swindells, LigPlot^{+}: multiple ligand-protein interaction diagrams for drug discovery, J. Chem. Inf. Model. 51 (2011) 2778–2786.

[58] P.D. Yadav, V.A. Potdar, M.L. Choudhary, D.A. Nyayanit, M. Agrawal, S.M. Jadhav, T.D. Majumdar, A. Shete-Aich, A. Basu, P. Abraham, S.S. Cherian, Full-genome sequences of the first two SARS-CoV-2 viruses from India, Indian J. Med. Res. 151 (2020) 200–209.

[59] R. Sanjuán, P. Domingo-Calap, Mechanisms of viral mutation, Cell. Mol. Life Sci. 73 (2016) 4433–4448.

[60] P.H. Guzzi, D. Mercatelli, C. Ceraolo, F.M. Giorgi, Master regulator analysis of the SARS-CoV-2/human interactome, J. Clin. Med. 9 (2020) 982.

[61] A.K. Banerjee, F. Begum, U. Ray, Mutation hot spots in Spike protein of COVID-19, Preprints 2020 (2020), https://doi.org/10.20944/preprints202004.0281.v1, 2020040281.

[62] C. Wang, Z. Liu, Z. Chen, X. Huang, M. Xu, T. He, Z. Zhang, The establishment of reference sequence for SARS-CoV-2 and variation analysis, J. Med. Virol. 92 (2020) 667–674.

[63] A. Shannon, N.T. Le, B. Selisko, C. Eydoux, K. Alvarez, J.C. Guillemot, E. Decroly, O. Peersen, F. Ferron, B. Canard, Remdesivir and SARS-CoV-2: structural requirements at both nsp12 RdRp and nsp14 Exonuclease active-sites, Antivir. Res. 178 (2020) 104793.

[64] F. Ferron, L. Subissi, A.T. Silveira De Morais, N.T.T. Le, M. Sevajol, L. Gluais, E. Decroly, C. Vonrhein, G. Bricogne, B. Canard, I. Imbert, Structural and molecular basis of mismatch correction and ribavirin excision from coronavirus RNA, Proc. Natl. Acad. Sci. U. S. A. 115 (2018) E162–E171.

[65] L.D. Eckerle, M.M. Becker, R.A. Halpin, K. Li, E. Venter, X. Lu, S. Scherbakova, R. L. Graham, R.S. Baric, T.B. Stockwell, D.J. Spiro, M.R. Denison, Infidelity of SARS-CoV Nsp14-exonuclease mutant virus replication is revealed by complete genome sequencing, PLoS Pathog. 6 (2010), e1000896.

[66] F. Bergasa-Caceres, H.A. Rabitz, Interdiction of protein folding for therapeutic drug development in SARS CoV-2, J. Phys. Chem. B 124 (2020) 8201–8208.

[67] M. Neinast, D. Murashige, Z. Arany, Branched chain amino acids, Annu. Rev. Physiol. 81 (2019) 139–164.

[68] F. Begum, D. Mukherjee, D. Thagriki, S. Das, P.P. Tripathi, A.K. Banerjee, U. Ray, Analyses of spike protein from first deposited sequences of SARS-CoV2 from West Bengal, India F1000Res 9 (2020) 371.

[69] L. Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, L. Sauerhering, S. Becker, K. Rox, R. Hilgenfeld, Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors, Science 368 (2020) 409–412.

[70] M. Thali, A. Bukovsky, E. Kondo, B. Rosenwirth, C.T. Walsh, J. Sodroski, H. G. Göttlinger, Functional association of cyclophilin A with HIV-1 virions, Nature 372 (1994) 363–365.

[71] B. Longdon, J.D. Hadfield, J.P. Day, S.C. Smith, J.E. McGonigle, R. Cogni, C. Cao, F. M. Jiggins, The causes and consequences of changes in virulence following pathogen host shifts, PLoS Pathog. 11 (2015), e1004728.

[72] R.P. Smith, J.J. Paxman, M.J. Scanlon, B. Heras, Targeting bacterial Dsb proteins for the development of anti-virulence agents, Molecules 21 (2016) 811.

[73] A.M. Phillips, L.O. Gonzalez, E.E. Nekongo, A.I. Ponomarenko, S.M. McHugh, V. L. Butty, S.S. Levine, Y.S. Lin, L.A. Mirny, M.D. Shoulders, Host proteostasis modulates influenza evolution, Elife 6 (2017), e28652.

[74] G.A. Somerville, R.A. Proctor, At the crossroads of bacterial metabolism and virulence factor synthesis in Staphylococci, Microbiol. Mol. Biol. Rev. 73 (2009) 233–248.

[75] T.M. Fuchs, W. Eisenreich, J. Heesemann, W. Goebel, Metabolic adaptation of human pathogenic and related nonpathogenic bacteria to extra- and intracellular habitats, FEMS Microbiol. Rev. 36 (2012) 435–462.

[76] M. Diard, W.D. Hardt, Evolution of bacterial virulence, FEMS Microbiol. Rev. 41 (2017) 679–697.

[77] M. Pachetti, B. Marini, F. Benedetti, F. Giudici, E. Mauro, P. Storici, C. Masciovecchio, S. Angeletti, M. Ciccozzi, R.C. Gallo, D. Zella, R. Ippodrino, Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant, J. Transl. Med. 18 (2020) 179.

[78] A. Maitra, M.C. Sarkar, H. Raheja, N.K. Biswas, S. Chakraborti, A.K. Singh, S. Ghosh, S. Sarkar, S. Patra, R.K. Mondal, T. Ghosh, A. Chatterjee, H. Banu, A. Majumdar, S. Chinnaswamy, N. Srinivasan, S. Dutta, S. Das, Mutations in SARS-CoV-2 viral RNA identified in Eastern India: possible implications for the ongoing outbreak in India and impact on viral structure and host susceptibility, J Biosci 45 (2020) 76.

[79] D. Topalis, S. Gillemot, R. Snoeck, G. Andrei, Distribution and effects of amino acid changes in drug-resistant α and β herpesviruses DNA polymerase, Nucleic Acids Res. 44 (2016) 9530–9554.

[80] J. Mokaya, A.L. McNaughton, M.J. Hadley, A. Beloukas, A.M. Geretti, D. Goedhals, P.C. Matthews, A systematic review of hepatitis B virus (HBV) drug and vaccine escape mutations in Africa: a call for urgent action, PLoS Neglected Trop. Dis. 12 (2018), e0006629.

[81] J. Li, K. Zhang, Q. Chen, X. Zhang, Y. Sun, Y. Bi, S. Zhang, J. Gu, J. Li, D. Liu, W. Liu, J. Zhou, Three amino acid substitutions in the NS1 protein change the virus replication of H5N1 influenza virus in human cells, Virology 519 (2018) 64–73.

[82] B. Robson, COVID-19 Coronavirus spike protein analysis for synthetic vaccines, a peptidomimetic antagonist, and therapeutic drugs, and analysis of a proposed achilles' heel conserved region to minimize probability of escape mutations and drug resistance, Comput. Biol. Med. 121 (2020) 103749.