ESC - a comprehensive resource for SARS-CoV-2 immune escape variants

Mercy Rophina[1,2], Kavita Pandhare[1], Afra Shamnath[1], Mohamed Imran[1,2], Bani Jolly[1,2], Vinod Scaria[1,2#]

[1]CSIR Institute of Genomics and Integrative Biology (CSIR-IGIB), Mathura Road, New Delhi, India
[2]Academy of Scientific and Innovative Research (AcSIR), Ghaziabad- 201002, India

 # address for correspondence

ABSTRACT

Ever since the breakout of COVID-19 disease, ceaseless genomic research to inspect the epidemiology and evolution of the pathogen has been undertaken globally. Large scale viral genome sequencing and analysis have uncovered the functional impact of numerous genetic variants in disease pathogenesis and transmission. Emerging evidence of mutations in spike protein domains escaping antibody neutralization is reported. We have a precise collation of manually curated variants in SARS-CoV-2 from literature with potential escape mechanisms from a range of neutralizing antibodies. This comprehensive repository encompasses a total of 532 variants accounting for 146 unique variants tested against 75 antibodies and patient convalescent plasma. This resource enables the user to gain access to an extensive annotation of SARS-CoV-2 escape mutations which we hope would contribute to exploring and understanding the underlying mechanisms of immune response against the pathogen. The resource is available at http://clingen.igib.res.in/esc/.

Keywords : COVID-19, SARS-CoV-2, Neutralizing antibodies, Escape mutations, Genomes

INTRODUCTION

Genomics approaches have been instrumental in understanding the origin and evolution of SARS-CoV-2, the causative agent for the COVID-19 pandemic (Zhang and Holmes, 2020). Availability of the genome sequence of one of the earliest SARS-CoV-2 genomes from Wuhan province (Tang et al., 2020) and high throughput approaches to resequence and analyse viral genomes have facilitated the availability of numerous open genomic data sharing initiatives by the researchers worldwide. Pioneering public sources like GenBank (GenBank Overview) and Global Initiative on Sharing all Influenza Data (GISAID) (Shu and McCauley, 2017) provide access to systematically organized genomes and genomic data of SARS-CoV-2. The China National GeneBank DataBase (CNGBdb) (Chen et al., 2020), Genome Warehouse (GWH) (Genome Warehouse) and Virus Pathogen Resource (ViPR) (Pickett et al., 2012) are few other resources which provide access to viral genomes and perform analyses on phylogeny, sequence similarity and genomic variants.

There has been a significant interest in recent times in understanding the functional impact of genetic variants in SARS-CoV-2 apart from exploring the genetic epidemiology. D614G variant in Spike protein has been one the earliest and prominent examples with potential implications associated with the infectivity of the virus (Korber et al., 2020). Studies explaining the possible impact of SARS-CoV-2 variants in diagnostic primers and probes (Jain et al., 2020) have augmented the importance of analysing the variations in the pathogen and their underlying role in disease pathogenesis. Various resources have been made available to help comprehend the virus better and also to understand its evolution. Public sources exclusively documenting functionally relevant SARS-CoV-2 variants based on literature evidence are also available (Rophina et al.).

With the advent of therapies including monoclonal antibodies, convalescent plasma as well as the recent availability of vaccines have accelerated interest in genetic variants which could affect the efficacy of such modalities of therapy. The targeting of spike proteins by broad-neutralizing antibodies against SARS-CoV-2 offers a potential means of treating and preventing further infections of COVID-19 (Jiang et al., 2020). Evidence on immunodominant epitopes with significantly higher response rates have also been reported (Farrera-Soler et al., 2020). Antibody response to SARS-CoV-2 is one of the key immune responses to SARS-CoV-2 which is actively being pursued to develop therapeutic strategies as well as vaccines (Biswas et al., 2020). The recent months have seen enormous research into the structural and molecular architecture of the interactions between the spike protein in SARS-CoV-2 and antibodies. Studies have also provided insights into the genetic variants which could confer partial or complete resistance to antibodies (Weisblum et al., 2020) as well as panels of convalescent plasma. With vaccines being widely available, the evidence on the effect of genetic variants on efficacy of vaccines are also emerging (Williams and Burgers, 2021)

The lack of a systematic effort to compile genetic variants in SARS-CoV-2 associated with immune escape motivated us to compile the information in a relevant, searchable and accessible format. Towards this goal, we systematically evaluated publications for evidence on immune escape associated with genetic variants in SARS-CoV-2 and created a compendium - ESC. This compendium can be searched using a user friendly web interface to retrieve information on immune escape variants as well as their extensive functional annotations. To the best of our knowledge this is the first most comprehensive resource for immune escape variants for SARS-CoV-2. The resource can be accessed online at http://clingen.igib.res.in/esc/.

MATERIALS AND METHODS

Data and Search Strategy

Genetic variants in the SARS-CoV-2 genome and evidence suggesting association with immune escape were systematically catalogued. A significant number of variants were associated with escape or resistance to a range of neutralizing and monoclonal antibodies, while a small subset were associated with resistance to convalescent plasma. The data was compiled by manual curation of literature available from peer-reviewed publications and preprints and comprehending the evidence mentioned in the respective articles. Literature reports with relevant information on antibody escape variants were retrieved from sources including PubMed, LitCovid, Google Scholar and articles from preprint servers. The publications were systematically checked for details pertaining to the mutation, antibodies tested and experimental methods followed in the study. Collated data was organized in a pre-formatted template based on their protein positions. This comprehensive compendium was used for further functional annotations.

Variant Information and Annotations

The variant information and annotations were retrieved from annotation tables for individual features using ANNOVAR (Wang et al., 2010). Variant annotations broadly include genic features like the variant type and functional annotations related to deleteriousness and evolutionary conservation. Information on protein domains and immune epitopes were compiled and customized from various public sources. Variant sites reported to be potentially problematic including homoplasic regions, sites with recurrent sequencing errors and hypermutable sites were also marked. Variants mapping back to sites of potential SARS-CoV-2 diagnostic primers/probes were also used for annotation (Jain et al., 2020).

Compilation of B-cell epitope data

Details on B-cell epitopes spanning the protein residues of SARS-CoV-2 were retrieved from Immune Epitope Database and Analysis Resource (IEDB) (Vita et al., 2019a). All epitopes of SARS-CoV-2 (IEDB ID : 2697049) against human hosts with reported positive or negative B cell assays and any type of MHC restriction were used for analysis. Epitope information pertaining to each amino acid residue including the epitope type (Linear/discontinuous), epitope sequence with corresponding start and end positions and IEDB identifiers were systematically mapped

back and documented. With a view of analysing the impact of these mutations on B cell epitopes, potential B cell epitopes were predicted from mutated spike protein sequences using BepiPred (Jespersen et al., 2017). Curated lists of amino acid changes were incorporated in the reference spike protein sequence (YP_009724390) and potential B cell epitopes were predicted with a default threshold value of 0.5.

Antibody details and annotation

Information pertaining to the list of antibodies associated with escape mechanisms were retrieved from available public sources. Compiled antibodies were systematically mapped back to the ABCD (for AntiBodies Chemically Defined) database which provides integrated information regarding the antibodies along with its corresponding antigens and protein cross links to fetch unique antibody identifiers (Page et al., 2016; Lima et al., 2020).

Database and Web Interface

The back-end of the web interface was implemented using Apache web server and MongoDB v3.4.10. to provide a user friendly interface for variant search. The JavaScript Object Notation (JSON) file format was used to systematically store the data. PHP 7.0, AngularJS, HTML, Bootstrap 4, and CSS were used to code the web interface for querying. Highcharts javascript library was also used for improved presentation and interactivity.

RESULTS & DISCUSSION

Repository of SARS-CoV-2 escape mutations

We compiled a total of 532 entries from 19 recent publications which studied SARS-CoV-2 variants and their effect on immune escape. This encompassed a total of 146 unique variants spanning spike protein, ORF1ab and ORF3a. Of the total unique variants, 143 variants were found mapping to Spike protein while 3 variants which were reported to confer potential epitope loss were found in ORF1ab and ORF3a. The compiled list of variants were associated with 75 unique SARS-CoV-2 antibodies and patient polyclonal sera. Summary of compiled variants along with their associated antibodies are listed in **Table 1.** The variant annotations encompassed 23 unique data features retrieved from multiple databases and datasets and mapped using ANNOVAR. The data features used in the study are summarised in **Supplementary Table 1.**

Antibody association mapping

By scanning through the spike protein residues and their associations with SARS-CoV-2 neutralizing and monoclonal antibodies, we were able to compile the exact count of antibodies reported to have potential associations with the residues. Antibody - amino acid residue mapping observed in Spike protein and in Receptor binding domain is represented in **Figures 1, 2 and 3.** From our analysis we were able to observe that spike protein residues ranging from 350 to 500 exhibited potential antibody associations. 9 hotspot residues (417, 444, 445, 450, 475, 484, 486, 490,493) were found associated with escape from >10 monoclonal antibodies.

Overview of B cell epitopes and immunodominant epitope regions

There were a total of 310 experimentally validated B cell epitopes including 263 linear and 47 discontinuous epitopes in Spike protein. Reported B cell epitope information was mapped back to residues possessing antibody escape mutations. Also for every escape mutation B-cell epitopes were predicted using BepiPred-2.0 (Jespersen et al., 2017){ref} using default options to see if any existing epitopes were affected when compared with reference sequence. We observed that 18 escape mutations caused complete loss of epitope given for that mutation.

Database features

The resource offers a user friendly interface enabling the users to search for variants based on their amino acid change, gene name or the antibody name as per the specified format. The search query returns a list of matching results, whose complete functional annotations can be viewed by clicking on the displayed elements. The resource provides a list of annotation features for each variant precisely organized into 5 major sections namely Variant details, Antibody details, Domain and Epitope details, Functional annotation and Variant frequency. **Figure 3a and b** portrays the query search and the result display section of the resource

Basic details pertaining to the variant like the amino acid change, genomic variation and the mutation type are enlisted in the Variant Details section. Information on the associated neutralizing antibodies and their identifiers are provided in the Antibody details section. Domain and epitope details section exclusively comprises details on the protein domain, epitopes reported to span the protein residue through experimental validations. Computationally predicted functional annotations on deleteriousness from SIFT (Ng and Henikoff, 2003), evolutionary conservation scores provided by PhastCons (Siepel and Haussler), GERP (Cooper, 2005) and PhyloP (Pollard et al., 2010) are included in the Functional annotation section. This section also enlists protein domain information retrieved from UniProt and immune epitopes documented from IEDB (Vita et al., 2019a, 2019b), UCSC and predictions from different software packages (B cells- BepiPred 2.0, CD4-IEDB Tepitool, CD8-NetMHCpan4). Annotations of potential error prone sites including sites of sequencing errors, homoplasic and hypermutable regions (NicolaDeMaio et al., 2020) and diagnostic primer/probe sites are also mapped.

CONCLUSIONS

With evidence emerging on genetic variants in SARS-CoV-2 associated with resistance to monoclonal antibodies and convalescent plasma using *in-vitro* assays have provided a unique insight into the structural and functional mechanisms whereby the pathogen could evolve and evade antibodies. These insights could have enormous implications in efficacy of vaccines currently being used as well as under trials. A number of recent studies have reported the impact of a few immune escape variants on the efficacy of vaccines (Nelson et al., 2021). It is expected that similar studies would be extended for a wider number of variants as well as

vaccines. We therefore foresee that the ESC resource would be a central resource to enable such studies and provide a ready reference to the emerging evidence on immune escape. We also foresee more evidence to emerge on cellular immune escape mechanisms which would also be updated in the future.

FIGURES AND TABLES

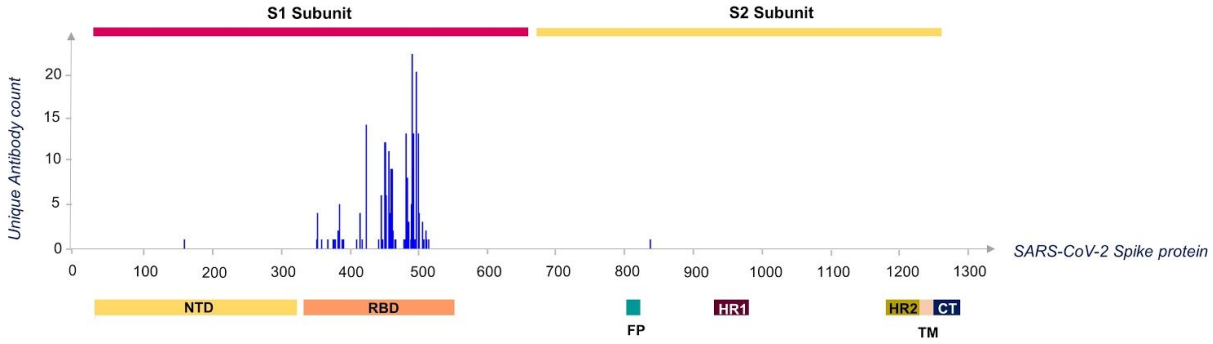

Figure 1. SARS-CoV-2 escape mutations along the spike protein residues. Illustration of mutation - antibody associations along the Spike protein residues.

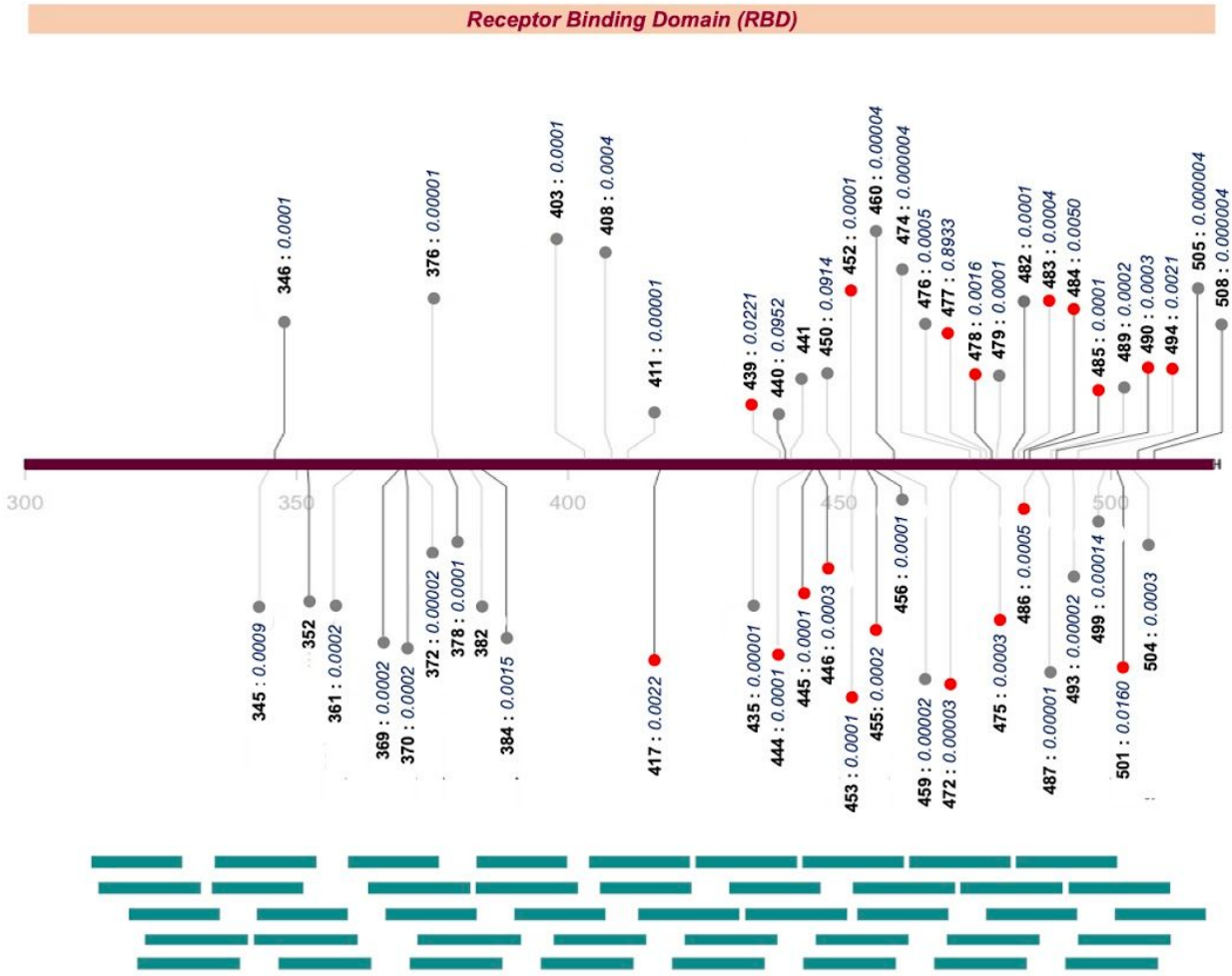

Figure 2 : Distribution of variants in Receptor binding domain of SARS-CoV-2 Spike protein. Variant sites with potential impact on neutralization of human polyclonal sera are represented in red. Cumulative frequencies of variants at RBD sites are represented along with tracks of validated epitope sites.

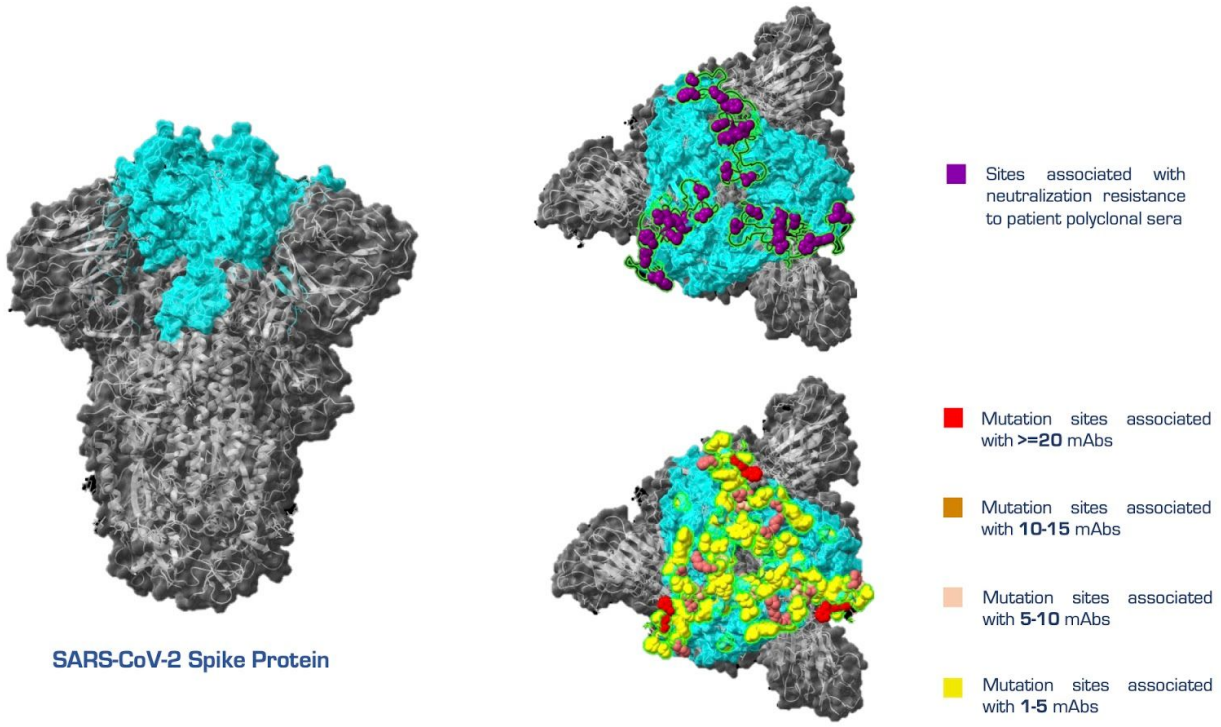

Figure 3. Receptor Binding Domain of Spike protein with mutation hotspot sites associated with decreased neutralization against patient polyclonal sera and monoclonal antibodies.

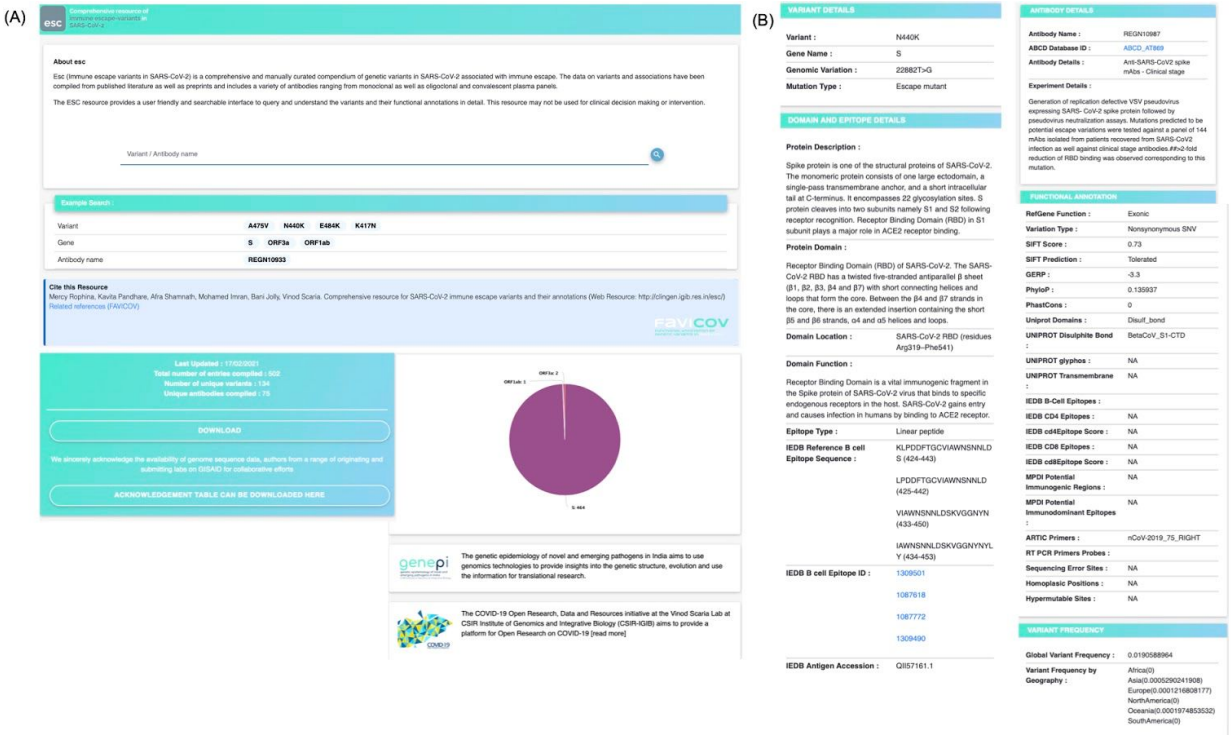

Figure 3. Panel illustrating the query search and display features in esc resource

Gene name	Variant	Antibodies	Ref
S	Δ146	4A8	(McCarthy et al., 2021)
S	V483G	SARS2-23	(Liu et al.)
S	N460K	LY-CoV016	(Shang and Axelsen)
S	F456S	2B04	(Liu et al.)
S	V483I	P2B-2F6	(Shang and Axelsen)
S	F456V	SARS2-58	(Liu et al.)
S	P384S	CoV2-2677	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020)
S	T345A	2H04	(Liu et al.)
S	N460S	LY-CoV016	(Shang and Axelsen)
S	V483P	HA001	(Yi et al., 2020)
S	N460T	LY-CoV016	(Shang and Axelsen)
S	N439K	H00S022,C135,LY-CoV555,Patient sera polyclonal ,REGN10933,REGN10987,S309	(Barnes et al., 2020) (Shang and Axelsen) (Thomson et al.)(Greaney et al.)
S	N487H	LY-CoV016	(Shang and Axelsen)
S	Y508H	H014	(The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity, 2020)
S	V445A	CoV2-2096,CoV2-2499,P2B-2F6,REGN10933,REGN10934,REGN10954,REGN10964,REGN10984,REGN10987,REGN10988,REGN10989	(Baum et al., 2020) (Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020) (Shang and Axelsen)
S	N487K	LY-CoV016,REGN10933	(Shang and Axelsen)
S	Δ243-244	4A8	(McCarthy et al., 2021)
S	T345N	2H04	(Liu et al.)
S	K378E	SARS2-31	(Liu et al.)
S	V445F	CoV2-2096,CoV2-2499,REGN10987	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020) (Shang and Axelsen)
S	V445G	SARS2-22	(Baum et al., 2020)

S	G485D	REGN10933,REGN10934,REGN10954,REGN10964,REGN10984,REGN10987,REGN10988,REGN10989	(Liu et al.)
S	T345S	2H04	(Liu et al.)
S	V445I	CoV2-2096,CoV2-2499	(Baum et al., 2020)(Shang and Axelsen)(Greaney et al.)
S	L455F	CC12.1,Patient polyclonal sera ,REGN10933,REGN10934,REGN10954,REGN10964,REGN10984,REGN10987,REGN10988,REGN10989	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020)
S	K417E	REGN10933,REGN10934,REGN10954,REGN10964,REGN10984,REGN10987,REGN10988,REGN10989	(Baum et al., 2020)
S	N501Y	B38,C102,C105,CC12.1,COVA1-18,COVA2-15,Patient polyclonal sera,S309	(Shang and Axelsen),(Tegally et al., 2020)(Greaney et al.),(Shen et al., 2021)
S	K378N	CoV2-2082,CoV2-2094,CoV2-2677,rCR3022	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020)
S	T478I	SARS2-19,SARS2-21,SARS2-71	(Liu et al.)
S	K378Q	SARS2-31	(Liu et al.)
S	K378R	CoV2-2094,CoV2-2677,rCR3022	
S	Y369C	CoV2-2677	(Andreano et al.)
S	K417N	C102,C105,CC12.1,CoV2-2082,CoV2-2094,LY-CoV016,Patient polyclonal sera	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020) (Shang and Axelsen) (Tegally et al. 2020),(Greaney et al.),(Cele et al.)
S	248aKTRNKSTSRRE248k	Convalescent Plasma,D14,F05,F10,F20,G12,H20,I15,I21,J13	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020)
S	K444E	2H04,SARS2-21	(Liu et al.)
S	G485R	P2B-2F6,Patient polyclonal sera	(Shang and Axelsen),(Greaney et al.)
S	Δ141-144	4A8	(McCarthy et al., 2021)
S	T376I	CoV2-2082,CoV2-2094	(Liu et al.)
S	P479L	SARS2-34,SARS2-71	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020)
S	T478P	SARS2-71	(Liu et al.)

S	S459Y	CC12.1	(Shang and Axelsen)
S	R346G	2H04,SARS2-01	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020)
S	K417R	CoV2-2082,CoV2-2094	(Liu et al.)
S	W152L	4A8	(Cele et al.)
S	A701V	Patient Polyclonal sera	(Ip et al.)
S	N370K	CoV2-2677	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020)
S	R346K	C135,P2B-2F6	(Thomson et al.)
S	K417V	LY-CoV555,REGN10933,REGN10987,S309	(Shang and Axelsen),(Weisblum et al., 2020)
S	A352D	SARS2-01	(Liu et al.)
S	R346M	C135	(Liu et al.)
S	P479S	SARS2-21	(Liu et al.) (Shang and Axelsen)
S	K444N	P2B-2F6,REGN10987,SARS2-21,SARS2-22	(Weisblum et al. 2020)
S	C361T	rCR3022	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020)
S	A475P	HA001	(Yi et al., 2020)
S	K444Q	REGN10933,REGN10934,REGN10954,REGN10964,REGN10984,REGN10987,REGN10988,REGN10989	(Baum et al., 2020)
S	S494D	HA001	(Yi et al., 2020)
S	K444R	P2B-2F6,SARS2-22	(Liu et al.) (Shang and Axelsen)
S	R408I	CC12.1,CoV2-2082,CoV2-2094	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020),(Shang and Axelsen)
S	N370S	CoV2-2677	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020)
S	R346S	C135	(Barnes et al., 2020)(Weisblum et al. 2020)
S	R408K	CoV2-2082,CoV2-2094,SARS2-31	(Complete Mapping of Mutations to the

			SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020)(Barnes et al., 2020) (The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity, 2020), (Shang and Axelsen)
S	A475V	157,247,B38,C102,C105,CA1,CB6,CC12.1,CoV2-2165,CoV2-2832,LY-CoV016,P2C-1F11	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020) (Liu et al.)
S	F140del	Convalescent sera ,D14,F05,F10,F20,G12,H20,I15,I21,J13	(Andreano et al.)
ORF3a	S1498F	Epitope loss	(Non-synonymous mutations of SARS-CoV-2 leads epitope loss and segregates its variants, 2020)
S	E484A	1B07,2B04,CC12.1,CoV2-2050,CoV2-2479,CoV2-2832,P2B-2F6	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020) (Liu et al.) (Shang and Axelsen)
S	S494L	CC12.1,P2B-2F6	(Shang and Axelsen)
S	E484D	1B07,CoV2-2050,CoV2-2479,CoV2-2832,SARS2-23,SARS2-66	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020) (Liu et al.)
S	S494P	CC12.1,P2B-2F6,Patient polyclonal sera ,SARS2-01	(Liu et al.) (Shang and Axelsen),(Greaney et al.)
S	R408T	CoV2-2082,CoV2-2094	((Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020)
S	L18F	Patient polyclonal sera	(Cele et al.)
S	K450Q	SARS2-66	(Liu et al.)
S	G476D	SARS2-21,SARS2-34,SARS2-71	(Liu et al.)
S	E484G	1B07	(Liu et al.)
S	A435S	CoV2-2094	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020)
S	Y505C	CC12.1	(Shang and Axelsen)

S	N440D	REGN10987	(Shang and Axelsen)
S	D215G	Patient polyclonal sera	(Cele et al.)
S	E484K	1B07,2B04,ABCD_AT480,ABCD_AT483,C121,C144,CC12.1,CoV2-2050,CoV2-2479,CoV2-2832,Convalescent plasma,D14,F05,F10,F20,G12,H20,I15,I21,J13,P2B-2F6,REGN10933,REGN10934,REGN10954,REGN10964,REGN10984,REGN10987,REGN10988,REGN10989,SARS2-02,SARS2-32,SARS2-55	(Andreano et al.),(Barnes et al., 2020),(Baum et al., 2020) (Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020) (Liu et al.) (Shang and Axelsen),(Weisblum et al. 2020),(Tegally et al. 2020),(Greaney et al.),(Cele et al.)
S	Y489H	LY-CoV016,REGN10933	(Shang and Axelsen)
S	S477G	CC12.1,SARS2-07,SARS2-16,SARS2-19,SARS2-34,SARS2-58,SARS2-71	(Liu et al.) (Shang and Axelsen)
S	S477I	CC12.1,SARS2-58	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020)(Shang and Axelsen)(Greaney et al.)
S	N440K	C135,REGN10987	(Liu et al.) (Shang and Axelsen)
S	E484Q	CC12.1,CoV2-2050,CoV2-2479,CoV2-2832,Patient polyclonal sera,P2B-2F6	(Barnes et al., 2020) (Shang and Axelsen)(Weisblum et al. 2020)
S	F486I	REGN10933	(Shang and Axelsen)
S	Q493K	C121,C144,LY-CoV016,REGN10933,REGN10934,REGN10954,REGN10964,REGN10984,REGN10987,REGN10988,REGN10989	(Baum et al., 2020) (Shang and Axelsen) (Weisblum et al. 2020)
S	S477N	CC12.1,Patient polyclonal sera,SARS2-07,SARS2-16,SARS2-19,SARS2-34,SARS2-58	(Shang and Axelsen).
S	Q493L	CC12.1,P2B-2F6,REGN10933	(Liu et al.) (Shang and Axelsen),(Greaney et al.)
S	G476S	CC12.1,SARS2-21	(Liu et al.) (Shang and Axelsen)
S	F486L	CoV2-2832,HA001,Patient polyclonal sera,REGN10933,SARS2-21	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020) (Liu et al.) (Shang and Axelsen) ,(Yi et al., 2020)(Greaney et al.)
S	V382L	rCR3022	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding

			Domain that Escape Antibody Recognition, 2020)
S	S477R	CC12.1,SARS2-07,SARS2-16,SARS2-23,SARS2-34	(Liu et al.) (Shang and Axelsen)
S	N440T	REGN10987	(Shang and Axelsen)
S	Q493R	C121,C144,LY-CoV016	(Barnes et al., 2020) (Shang and Axelsen) (Weisblum et al. 2020)
S	L452M	P2B-2F6,Patient polyclonal sera	(Baum et al., 2020),(Du et al., 2020) (Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020) (The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity, 2020) (Liu et al.) (Shang and Axelsen),(Weisblum et al. 2020)
S	F490L	261-262,BD-368-2,C121,CC12.1,CoV2-2050,CoV2-2479,H4,P2B-2F6,P2B2F6,REGN10933,REGN10934,REGN10954,REGN10964,REGN10984,REGN10987,REGN10988,REGN10989,SARS2-66,X593	(Liu et al.)
S	F486S	SARS2-21	(Shang and Axelsen),(Greaney et al.)
S	N440Y	REGN10987	(Shang and Axelsen)
S	Y453F	CC12.1,Patient polyclonal sera,REGN10933,REGN10934,REGN10954,REGN10964,REGN10984,REGN10987,REGN10988,REGN10989	(Baum et al., 2020)(Shang and Axelsen). (Greaney et al.)
S	F486V	REGN10933,REGN10934,REGN10954,REGN10964,REGN10984,REGN10987,REGN10988,REGN10989,SARS2-71	(Liu et al.)
S	242-244del	Patient polyclonal sera	(Cele et al.)
S	P499H	CoV2-2096,CoV2-2499,REGN10987	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020) (Shang and Axelsen)
S	L452R	BD-368-2,Patient polyclonal sera,SARS2-01,SARS2-32	(Du et al., 2020) (Liu et al.) (Greaney et al.)
S	F486Y	1B07	(Cele et al.)
S	D80A	Patient polyclonal sera	(Liu et al.)

S	G482S	P2B-2F6	(Shang and Axelsen)
S	Q474P	SARS2-34	(Baum et al., 2020) (Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020) (Liu et al.) (Shang and Axelsen),(Greaney et al.)
S	F490S	CC12.1,CoV2-2050,CoV2-2479,P2B-2F6,Patient,REGN10933,REGN10934,REGN10954,REGN10964,REGN10984,REGN10987,REGN10988,REGN10989,SARS2-32	(Liu et al.)
S	N450D	REGN10933,REGN10934,REGN10954,REGN10964,REGN10984,REGN10987,REGN10988,REGN10989,SARS2-07	(Baum et al., 2020) (Liu et al.)
S	P499L	SARS2-07	(Liu et al.)
ORF3a	G251V	Epitope Loss	(Non-synonymous mutations of SARS-CoV-2 leads epitope loss and segregates its variants, 2020)
S	R403K	CC12.1	(Non-synonymous mutations of SARS-CoV-2 leads epitope loss and segregates its variants, 2020)
S	P4715L	NA	(Shang and Axelsen)
S	R403M	CC12.1	(Liu et al.) (Shang and Axelsen)
S	P499R	CoV2-2096,CoV2-2499,REGN10987	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020) (Shang and Axelsen)
S	N450K	SARS2-32	(Liu et al.)
S	N448K	REGN10987	(Shang and Axelsen)
S	G504D	LY-CoV016,SARS2-31	(Shang and Axelsen)
S	P499S	CoV2-2096,CoV2-2499,REGN10987	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020) (Shang and Axelsen)
S	G446A	CoV2-2096,CoV2-2499,REGN10987	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020) (Shang and Axelsen)
S	G446D	SARS2-02,SARS2-22,SARS2-32	(Liu et al.)

S	I472V	CoV2-2479	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020)
S	A831V	B38	(The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity, 2020)
S	N450Y	SARS2-32	(Liu et al.)
S	A372S	CoV2-2677	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020)
S	A372T	CoV2-2677	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020)
S	A372V	CoV2-2677	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020)
S	G446S	CoV2-2096,CoV2-2499,P2B-2F6,REGN10987	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020) (Shang and Axelsen)
S	A411S	CoV2-2082	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020)
S	V483A	BD-368-2,C121,P2B-2F6	(Barnes et al., 2020),(Du et al., 2020) (Shang and Axelsen)
S	L441R	2H04	(Liu et al.)
S	G446V	CoV2-2096,CoV2-2499,Patient polyclonal sera,REGN10933,REGN10987,SARS2-02,P2B-2F6	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020) (Liu et al.) (Shang and Axelsen) (Greaney et al.)
S	P384L	CoV2-2677	(Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition, 2020)
S	N460I	LY-CoV016	(Shang and Axelsen)
S	V483F	P2B-2F6,SARS2-23	(Liu et al.)

			(Shang and Axelsen)

Table 1. Summary of variants compiled in the study along with their associated list of anti SARS-CoV-2 antibodies

SUPPLEMENTARY DATASETS

Custom Datasets
Variation Type
Non synonymous variant
Synonymous variant
SIFT_Score
SIFT_Prediction (*Deleterious/Tolerable)*
GERP
PhyloP
PhastCons
Uniprot_Domains
UNIPROT_Disulphite_Bond
UNIPROT_glyphos
UNIPROT_Transmembrane
IEDB_B-Cell_Epitopes
IEDB_CD4_Epitopes
IEDB_cd4Epitope_Score
IEDB_CD8_Epitopes
IEDB_cd8Epitope_Score
MPDI_Potential_Immunogenic_Regions
MPDI_Potential_Immunodominant_Epitopes
ARTIC_Primers
RT-PCR_Primers/Probes
Sequencing_Error_Sites
Homoplasic_Positions
Hypermutable_Sites

Supplementary Table 1. Summary of variants mapping to various annotation features.

REFERENCES

Barnes, C. O., Jette, C. A., Abernathy, M. E., Dam, K.-M. A., Esswein, S. R., Gristick, H. B., et al.

(2020). Structural classification of neutralizing antibodies against the SARS-CoV-2 spike receptor-binding domain suggests vaccine and therapeutic strategies. *bioRxiv*. doi:10.1101/2020.08.30.273920.

Baum, A., Fulton, B. O., Wloga, E., Copin, R., Pascal, K. E., Russo, V., et al. (2020). Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies. *Science* 369, 1014–1018.

Biswas, A., Bhattacharjee, U., Chakrabarti, A. K., Tewari, D. N., Banu, H., and Dutta, S. (2020). Emergence of Novel Coronavirus and COVID-19: whether to stay or die out? *Critical Reviews in Microbiology* 46, 182–193. doi:10.1080/1040841x.2020.1739001.

Chen, F. Z., You, L. J., Yang, F., Wang, L. N., Guo, X. Q., Gao, F., et al. (2020). CNGBdb: China National GeneBank DataBase. *Yi Chuan* 42, 799–809.

Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition (2020). *Cell Host Microbe*. doi:10.1016/j.chom.2020.11.007.

Cooper, G. M. (2005). Distribution and intensity of constraint in mammalian genomic sequence. *Genome Research* 15, 901–913. doi:10.1101/gr.3577405.

Du, S., Cao, Y., Zhu, Q., Yu, P., Qi, F., Wang, G., et al. (2020). Structurally Resolved SARS-CoV-2 Antibody Shows High Efficacy in Severely Infected Hamsters and Provides a Potent Cocktail Pairing Strategy. *Cell* 183, 1013–1023.e13.

Farrera-Soler, L., Daguer, J.-P., Barluenga, S., Vadas, O., Cohen, P., Pagano, S., et al. (2020). Identification of immunodominant linear epitopes from SARS-CoV-2 patient plasma. *PLoS One* 15, e0238089.

GenBank Overview Available at: https://www.ncbi.nlm.nih.gov/genbank/ [Accessed December 19, 2020].

Genome Warehouse Available at: https://www.re3data.org/repository/r3d100013428 [Accessed December 19, 2020].

Greaney, A. J., Loes, A. N., Crawford, K. H. D., Starr, T. N., Malone, K. D., Chu, H. Y., et al. Comprehensive mapping of mutations to the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human serum antibodies. doi:10.1101/2020.12.31.425021.

Ip, J. D., Kok, K.-H., Chan, W.-M., Chu, A. W.-H., Wu, W.-L., Yip, C. C.-Y., et al. Intra-host non-synonymous diversity at a neutralizing antibody epitope of SARS-CoV-2 spike protein N-terminal domain. *Clin. Microbiol. Infect.* doi:10.1016/j.cmi.2020.10.030.

Jain, A., Rophina, M., Mahajan, S., Krishnan, B. B., Sharma, M., Mandal, S., et al. (2020). Analysis of the potential impact of genomic variants in SARS-CoV-2 genomes from India on molecular diagnostic assays. *Cold Spring Harbor Laboratory*, 2020.08.05.238618. doi:10.1101/2020.08.05.238618.

Jespersen, M. C., Peters, B., Nielsen, M., and Marcatili, P. (2017). BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids*

Res. 45, W24–W29.

Jiang, S., Hillyer, C., and Du, L. (2020). Neutralizing Antibodies against SARS-CoV-2 and Other Human Coronaviruses. *Trends in Immunology* 41, 545. doi:10.1016/j.it.2020.04.008.

Korber, B., Fischer, W. M., Gnanakaran, S., Yoon, H., Theiler, J., Abfalterer, W., et al. (2020). Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell* 182, 812–827.e19.

Lima, W. C., Gasteiger, E., Marcatili, P., Duek, P., Bairoch, A., and Cosson, P. (2020). The ABCD database: a repository for chemically defined antibodies. *Nucleic Acids Res.* 48, D261–D264.

Liu, Z., VanBlargan, L. A., Rothlauf, P. W., Bloyet, L.-M., Chen, R. E., Stumpf, S., et al. Landscape Analysis of Escape Variants Identifies SARS-CoV-2 Spike Mutations that Attenuate Monoclonal and Serum Antibody Neutralization. *SSRN Electronic Journal*. doi:10.2139/ssrn.3725763.

McCarthy, K. R., Rennick, L. J., Nambulli, S., Robinson-McCarthy, L. R., Bain, W. G., Haidar, G., et al. (2021). Recurrent deletions in the SARS-CoV-2 spike glycoprotein drive antibody escape. *Science*. doi:10.1126/science.abf6950.

Nelson, G., Buzko, O., Spilman, P., Niazi, K., Rabizadeh, S., and Soon-Shiong, P. (2021). Molecular dynamic simulation reveals E484K mutation enhances spike RBD-ACE2 affinity and the combination of E484K, K417N and N501Y mutations (501Y.V2 variant) induces conformational change greater than N501Y mutant alone, potentially resulting in an escape mutant. *Cold Spring Harbor Laboratory*, 2021.01.13.426558. doi:10.1101/2021.01.13.426558.

Ng, P. C., and Henikoff, S. (2003). SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* 31, 3812–3814.

NicolaDeMaio, n_j_loman, and goldman (2020). Issues with SARS-CoV-2 sequencing data. Available at: https://virological.org/t/issues-with-sars-cov-2-sequencing-data/473 [Accessed December 24, 2020].

Non-synonymous mutations of SARS-CoV-2 leads epitope loss and segregates its variants (2020). *Microbes Infect.* 22, 598–607.

Page, A. J., Taylor, B., Delaney, A. J., Soares, J., Seemann, T., Keane, J. A., et al. (2016). : rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb Genom* 2, e000056.

Pickett, B., Greer, D., Zhang, Y., Stewart, L., Zhou, L., Sun, G., et al. (2012). Virus Pathogen Database and Analysis Resource (ViPR): A Comprehensive Bioinformatics Database and Analysis Resource for the Coronavirus Research Community. *Viruses* 4, 3209–3226. doi:10.3390/v4113209.

Pollard, K. S., Hubisz, M. J., Rosenbloom, K. R., and Siepel, A. (2010). Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res.* 20, 110–121.

Rophina, M., Pandhare, K., Mangla, M., Shamnath, A., Jolly, B., Sethi, M., et al. FaviCoV -a comprehensive manually curated resource for functional genetic variants in SARS-CoV-2. doi:10.31219/osf.io/wp5tx.

Shang, E., and Axelsen, P. The Potential for SARS-CoV-2 to Evade Both Natural and Vaccine-induced Immunity. doi:10.1101/2020.12.13.422567.

Shen, X., Tang, H., McDanal, C., Wagh, K., Fischer, W., Theiler, J., et al. (2021). SARS-CoV-2 variant B.1.1.7 is susceptible to neutralizing antibodies elicited by ancestral Spike vaccines. *bioRxiv*. doi:10.1101/2021.01.27.428516.

Shu, Y., and McCauley, J. (2017). GISAID: Global initiative on sharing all influenza data - from vision to reality. *Euro Surveill.* 22. doi:10.2807/1560-7917.ES.2017.22.13.30494.

Siepel, A., and Haussler, D. Phylogenetic Hidden Markov Models. *Statistical Methods in Molecular Evolution*, 325–351. doi:10.1007/0-387-27733-1_12.

Tang, X., Wu, C., Li, X., Song, Y., Yao, X., Wu, X., et al. (2020). On the origin and continuing evolution of SARS-CoV-2. *National Science Review* 7, 1012–1023. doi:10.1093/nsr/nwaa036.

Tegally, H., Wilkinson, E., Giovanetti, M., Iranzadeh, A., Fonseca, V., Giandhari, J., et al. (2020). Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. *medRxiv*, 2020.12.21.20248640.

The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity (2020). *Cell* 182, 1284–1294.e9.

Thomson, E. C., Rosen, L. E., Shepherd, J. G., Spreafico, R., da Silva Filipe, A., Wojcechowskyj, J. A., et al. The circulating SARS-CoV-2 spike variant N439K maintains fitness while evading antibody-mediated immunity. doi:10.1101/2020.11.04.355842.

Vita, R., Mahajan, S., Overton, J. A., Dhanda, S. K., Martini, S., Cantrell, J. R., et al. (2019a). The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res.* 47, D339–D343.

Vita, R., Mahajan, S., Overton, J. A., Dhanda, S. K., Martini, S., Cantrell, J. R., et al. (2019b). The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res.* 47, D339–D343.

Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research* 38, e164–e164. doi:10.1093/nar/gkq603.

Weisblum, Y., Schmidt, F., Zhang, F., DaSilva, J., Poston, D., Lorenzi, J. C. C., et al. (2020). Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants. doi:10.7554/eLife.61312.

Williams, T. C., and Burgers, W. A. (2021). SARS-CoV-2 evolution and vaccines: cause for concern? *Lancet Respir Med*. doi:10.1016/S2213-2600(21)00075-8.

Yi, C., Sun, X., Ye, J., Ding, L., Liu, M., Yang, Z., et al. (2020). Key residues of the receptor binding

motif in the spike protein of SARS-CoV-2 that interact with ACE2 and neutralizing antibodies. *Cell. Mol. Immunol.* 17, 621–630.

Zhang, Y.-Z., and Holmes, E. C. (2020). A Genomic Perspective on the Origin and Emergence of SARS-CoV-2. *Cell* 181, 223–227.