

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

About ESC

This comprehensive repository encompasses a total of **4545** variants accounting for **2058** unique variants tested against **149** 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.

A Quick start to ESC

The Home Page

Contents and highlights of ESC

The homepage introduces the platform enlisting a brief summary of the data along with an option for users to download the same. We also provide user friendly search options with examples provided for reference.

The screenshot displays the ESC homepage with the following sections:

- Header:** Comprehensive resource of immune escape-variants in SARS-CoV-2
- Logo:** IGIB (Institute of Genetics & Integrative Biology)
- About esc:** A brief description of the resource, noting it is a comprehensive and manually curated compendium of genetic variants in SARS-CoV-2 associated with immune escape. It includes variants and associations from published literature and preprints, and a variety of antibodies ranging from monoclonal to oligoclonal and convalescent plasma panels.
- Search Bar:** Variant / Antibody name
- Example Search:** Variant: A475V, N440K, E484K, K417N; Gene: S, ORF3a, ORF1ab; Antibody/Vaccine Name: BNT162b2, AZD1222; Study Type: Experimental methods, Computational predictions; VCs/Vols: B.1.1.7, B.1.617.2
- Statistics:** Last Updated: 27/07/2021; Total number of entries compiled: 4545; Number of unique variants: 2058; Unique antibodies compiled: 149; Number of vaccine studies: 6
- Buttons:** DOWNLOAD, USER MANUAL
- Acknowledgment:** We sincerely acknowledge the availability of genome sequence data, authors from a range of originating and submitting labs on GISAID for collaborative efforts. ACKNOWLEDGEMENT TABLE CAN BE DOWNLOADED HERE
- Pie Chart:** A pie chart showing the distribution of variants across different genes. Labels include ORF1ab: 6, ORF3a: 2, and S: 4325.

Search options

ESC is designed in such a way to provide the users a friendly search environment.

- *Initial result display*

B.1.1.7

The screenshot shows a search results page with a header "B.1.1.7" and a search bar with a magnifying glass icon. Below the header is a teal bar labeled "Search results". The main content is a table with three columns: "Gene", "Variant", and "Antibody". The table has 10 visible rows, each corresponding to a different variant of the S gene (N501Y) and its associated antibody treatment. At the bottom of the table, it says "Showing 1 to 10 of 32 entries" and includes a navigation bar with buttons for "Previous", "1", "2", "3", "4", and "Next".

Gene	Variant	Antibody
S	N501Y	Convalescent plasma from the first wave of SARS-CoV-2 infection in early 2020
S	N501Y	AZD1222/Covishield
S	N501Y	BBV152/Covaxin
S	N501Y	CC12.1
S	D614G	Bamlanivimab/LY-CoV555
S	N501Y	B38
S	N501Y	COVA1-18
S	N501Y	COVA2-15
S	N501Y	S309
S	N501Y	C102

Showing 1 to 10 of 32 entries

Previous 1 2 3 4 Next

- Variant annotation display

This page is categorized into **8** major sections as follows

1. Variant Details

Provides the basic details pertaining to the variant including the Gene name, Genomic variation, type of variation, reference and alternate bases and amino acids. Linkouts to genomic references are also duly provided.

VARIANT DETAILS : K417N

Variant :	K417N
Gene Name :	S
NCBI Gene ID :	43740568
Gene Location :	NC_045512.2:21563-25384
Ensembl Gene ID :	ENSSASG00005000004
Variant Position :	22813
Reference Base :	G
Alternate Base :	T
Amino Acid Position :	417
Reference Amino Acid :	K (Lys)
Alternate Amino Acid :	N (Asn)
Genomic Variation :	22813G>T
Mutation Type :	Antibody Escape
Ensembl Transcript ID :	ENSSAST00005000004.1
CDS Position :	1251
Codons :	aaG/aaT
Ensembl Variant ID :	MN908947.3:22813:G:T
HGVS Nomenclature :	None

2. Antibody Details

Provides information regarding the associated antibody, antibody identifiers (if any) and briefs the experimental procedures.

ANTIBODY DETAILS

Antibody or Vaccine Name :	24-11K
Antibody or Vaccine Category & Description :	A SARS-CoV-2 NAb
ABCD Database ID :	None

3. Variants of Concern/Interest

Summarizes the lineage information of variants which are characteristic of Variants of Concern/Interests categorized by CDC

VARIANTS OF CONCERN/INTEREST

VoCs/VoIs :	B.1.351
VoCs/VoIs Aliases :	20H/501Y.V2
VoCs/VoIs Description :	~50% increased transmission. Significantly reduced susceptibility to the combination of bamlanivimab and etesevimab monoclonal antibody treatment, but other EUA monoclonal antibody treatments are available. Reduced neutralization by convalescent and post-vaccination sera

4. Protein Domain details

Exclusive details regarding the protein domain, domain location, domain functions.

PROTEIN DOMAIN DETAILS

PROTEIN NAME :	Spike glycoprotein
UNIPROT PROTEIN ID :	P0DT2
PROTEIN LENGTH :	1273 amino acids
Protein Description :	Spike protein is one of the structural proteins of SARS-CoV-2. The monomeric protein consists of one large ectodomain, a single-pass transmembrane anchor, and a short intracellular tail at C-terminus. It encompasses 22 glycosylation sites. S protein cleaves into two subunits namely S1 and S2 following receptor recognition. Receptor Binding Domain (RBD) in S1 subunit plays a major role in ACE2 receptor binding.
Protein Domain :	Receptor Binding Domain (RBD) of SARS-CoV-2. The SARS-CoV-2 RBD has a twisted five-stranded antiparallel β sheet (β_1 , β_2 , β_3 , β_4 and β_7) with short connecting helices and loops that form the core. Between the β_4 and β_7 strands in the core, there is an extended insertion containing the short β_5 and β_6 strands, α_4 and α_5 helices and loops.
Domain Position :	SARS-CoV-2 RBD (residues Arg319–Phe541)
Function of the Domain :	Receptor Binding Domain is a vital immunogenic fragment in the Spike protein of SARS-CoV-2 virus that binds to specific endogenous receptors in the host. SARS-CoV-2 gains entry and causes infection in humans by binding to ACE2 receptor.

5. Epitope Details

Information on reference B and T cell epitopes and related details from IEDB

EPITOPE DETAILS	
Epitope Type :	Linear peptide
IEDB Reference B cell Epitope Sequence :	VIRGDEVRQIAPGQTGKI (401-418) QIAPGQTGKIADYNKYKL (409-426) QTGKIADYNKYKLPPDDFTGCV (414-433) KIADYNKYKLPPDDFTGCVI (417-434)
IEDB B cell Epitope ID :	1087774 1087691 1309563 1087572
IEDB B cell Antigen Accession :	QHD43416.1 QHD43416.1 QII57161.1 QHD43416.1
IEDB Reference T cell epitope Sequence :	EVRQIAPGQTGKIAD (406-420) RQIAPGQTGKIADYNYKL (408-425) APGQTGKIADYNYKL (411-425) GKIADYNYKLPPDDFT (416-430)
T cell Epitope ID :	1310377 1075039 1310279 1310447
IEDB T cell Antigen Accession :	QKE11719.1 YP_009724390.1 QKE11719.1 QKE11719.1

6. Literature Evidence

This section briefs the methods and workflow adapted in the reported literature evidence

LITERATURE EVIDENCE

Study Type :	Experimental methods
Details of the study :	Individual mutations associated with RBD antibody escape were introduced into a full-length SARS-CoV S plasmid using homologous recombination. The affinity of RBD mutants for ACE2-FC was measured using Surface plasmon resonance (SPR)
Experiment Type :	Mutations in isolation
Neutralization Quantification :	Reactivity of the mutant relative to wildtype = 44
Antibody Generation :	Human mAbs were generated from human memory B cells by single-cell RT-PCR. The gene sequence analysis of mAbs were performed by IMGT and IgBLAST. The S309, CR3022, CB6 and B38 VH and VL sequences were synthesized and cloned into expression vectors
SARS-CoV-2 Mutants Generation :	The DNA sequences encoding the SARS-CoV-2 RBD soluble fragments encompassing amino acids 319- 589 of the S protein were fused in-frame with an N-terminal human IgE signal peptide and a C-terminal human IgG1 Fc and 8x His tag. Global alanine scanning of the RBD was performed using site directed mutagenesis of RBD residues (330-521) to alanine (natural mutations for alanine residues). Substitutions of the residues at the antigenic sites selected for mutagenesis were based on the 2019 Novel Coronavirus Resource (2019nCoVR) released by the China National Center for Bioinformation (https://bigd.big.ac.cn/ncov/protein). For pseudovirus variants, individual mutations associated with RBD antibody escape were introduced into a full-length SARS-CoV S plasmid using homologous recombination. All of the mutations were confirmed by DNA sequence analysis
Antibody Binding escape Profile :	The SARS-CoV-2 clinical isolate nCoV-SH01 (GenBank: MT121215.1) was amplified in Vero E6 cells and used for authentic virus neutralization

7. Functional annotation

This section provides the global frequency of the variant along with its frequency across various geographical regions.

FUNCTIONAL ANNOTATION

RefGene Function :	Protein Coding Region
Variation Type :	Nonsynonymous SNV
SIFT Score :	0.31
SIFT Prediction :	Tolerated
GERP :	-3.3
PhyloP :	-0.613165
PhastCons :	0.984252
Uniprot Domains :	Disulf_bond
UNIPROT Disulphite Bond :	BetaCoV_S1-CTD
UNIPROT glyphs :	NA
UNIPROT Transmembrane :	NA
IEDB B-Cell Epitopes :	
IEDB CD4 Epitopes :	NA
IEDB cd4Epitope Score :	NA
IEDB CD8 Epitopes :	Surface glycoprotein_RQIAPGQTGK
IEDB cd8Epitope Score :	0.03
MPDI Potential Immunogenic Regions :	NA
MPDI Potential Immunodominant Epitopes :	NA
ARTIC Primers :	nCoV-2019_76_LEFT
RT PCR Primers Probes :	
Sequencing Error Sites :	NA
Homoplasic Positions :	NA
Hypermutable Sites :	NA

8. Variant Frequency

Lists the frequency of the particular variation on global and geographical scales

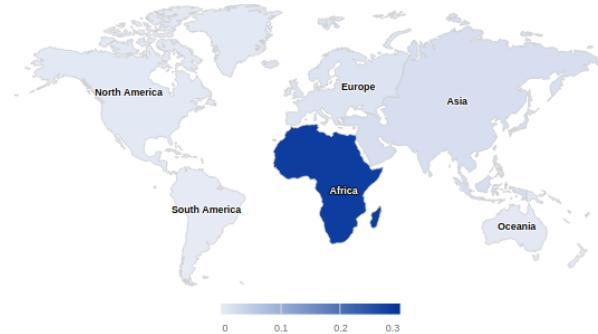
VARIANT FREQUENCY

Global Variant Frequency :	0.32667824
Variant Frequency by Geography :	Africa(0.283376) Asia(0.0210838) Europe(0.0121013) NorthAmerica(0.00472294) Oceania(0.0043945) SouthAmerica(0.0009997)
Variant Population genetics :	None
Variant Sample Genotype :	None

REFERENCES

Yi, C., Sun, X., Lin, Y., Gu, C., Ding, L., Lu, X., Yang, Z., Zhang, Y., Ma, L., Gu, W., Qu, A., Zhou, X., Li, X., Xu, J., Ling, Z., Xie, Y., Lu, H., & Sun, B. (n.d.). Comprehensive Mapping of Binding Hot Spots of SARS-CoV-2 RBD-specific Neutralizing Antibodies for Tracking Immune Escape Variants. <https://doi.org/10.21203/rs.3.rs-497595/v1>

Global Distribution



Variant frequencies were calculated using GISAID sequence variations data.

*Date of last update of variant frequencies - **27.07.2021***

