Larix Publications

Journal of Pharma Research

https://jprinfo.com/

Vol. 12 Issue 01, 2023 ISSN: 2319-5622

Research Article

Structural, non-specific, and accessory proteins and variants of SARS-CoV-2: Current perspective

Dinesh Kumar Lakshmi Narayanan[1], Saminathan Kayarohanam[2], Ashok Kumar Janakiraman[3], Sinouvassane Djearamane[4], Siddharthan Selvaraj[1]

[1]Faculty of Pharmacy, AIMST University, Bedong, Kedah 08100, Malaysia
[2]Geomatika University College, Setiawangsa, Kuala Lumpur -54200
[3]Faculty of Pharmaceutical Sciences, UCSI University, Cheras 56000, Kuala Lumpur, Malaysia
[4]Department of Biomedical Science, Faculty of Science, Universiti Tunku Abdul Rahman Jalan University, Bandar Barat, Kampar 31900, Perak, Malaysia.
[1]Faculty of Dentistry, AIMST University, Bedong, Kedah 08100, Malaysia

Received on: 23-12-2022; Revised and Accepted on: 02-01-2023

ABSTRACT

Background
Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the causative agent of Covid-19. We have been fortunate enough for the swift response of the Scientific world to the various types of vaccines available today. Presently there is ongoing continuous research for more therapeutic options. Yet the worrying fact is there has been the emergence of various new mutant strains. These mutant strains are hindering the process the whole world is carrying out since the beginning of the pandemic to curb it. These mutant strains are of concern since they possess high infective abilities, neutralizing the natural as well as vaccine-induced antibodies and with immune-evading properties. As the focus present is more on the mutation and mutant strains of Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) understanding the role of their structural proteins, non-structural proteins, and accessory proteins is more significant as they do have a key role in the above-mentioned properties in mutant strains.

Conclusion
Researchers and the scientific community need the compiled data about the various variants as per WHO, and crucial information on SARS-CoV-2 structural proteins, non-structural proteins, and accessory proteins. This review is to elucidate SARS-CoV-2 structural proteins, non-structural proteins, and accessory proteins as well as the mutant variants.

Keywords: *Accessory proteins, Covid–19, non-structural proteins, SARS-CoV-2, structural proteins, variant strains*

Background

Coronaviruses (CoV) belong to the positive-stranded RNA virus type which causes infections in humans and animals. These viruses cause respiratory infections in humans and animals. The infections may be mild to lethal depending on the type of the virus from the coronavirus family. The primary target cells for the corona viruses are epithelial cells. They are known to cause respiratory and gastrointestinal infections which may be acute or may become chronic if there is lengthened shedding of the virus and have a mortality rate of 3.4% worldwide as per WHO. The coronavirus family was first reported and ratified by a team of renowned scientists comprising J. D. Almeida, D. M. Berry, C. H. Cunningham, D. Hamre, M. S. Hofstad, L. Mallucci, K. McIntosh and D. A. J. Tyrrell in 1968.1 Coronavirus belongs to the order–Nidovirales which includes the largest RNA genomes, family–Coronoviridae and subfamily-Orthocoronavirinae.[2] Nidoviruses exert a very complex RNA-synthesizing machinery. One of the significant characteristics of this order is they possess proofreading 3′-to-5′ exoribonuclease (ExoN) which makes sure there are no errors in nucleotide incorporation done by the RNA-

***Corresponding Author:**
Dinesh Kumar Lakshmi Narayanan
Faculty of Pharmacy, AIMST University, Bedong, Kedah 08100, Malaysia
Email: dineshkumar@aimst.edu.my
DOI: https://doi.org/10.5281/zenodo.7498622

dependent RNA polymerase.[3] Coronaviridae family belongs to the single-stranded RNA virus type, and they are the more prominent RNA viruses. They possess 25 to 32 kb largest genomes. They are subdivided into two subfamilies namely Coronavirinae and Torovirinae. They are differentiated based on nucleocapsids.[4] Virions of the coronavirinae are spherical and torovirinae are bent into crescent structures shown in Figure 1.

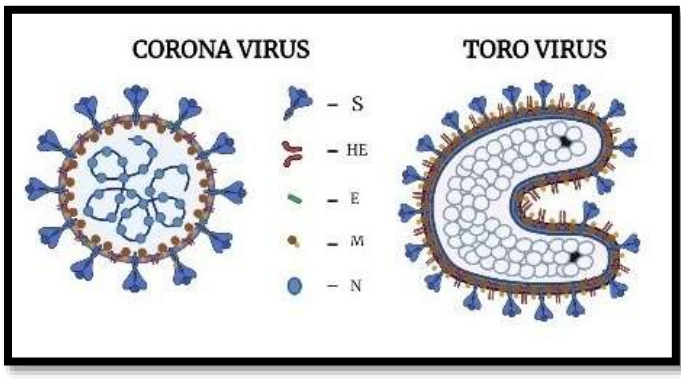


Fig 1: Virions of Coronovirinae and Toronovirinae are differentiated on basis of nucleocapsid, which is spherical in coronavirinae and crescent in toronovirinae

Subfamily of coronaviridae, includes four clusters or genera viz., Alpha coronavirus, Beta coronavirus, Gamma coronavirus, and Delta coronaviruses. Coronavirus, as mentioned earlier, is the prominent RNA virus, with a 25 to 32 kb genome and a virion with 118–136 nm in diameter. Virions of coronavirus are spherical with large spike glycoprotein. They are widespread and mostly cause respiratory and enteric infections. Till 2002 it was considered a pathogen that may not affect or cause mild infections in humans, but the emergence of severe acute respiratory syndrome (SARS) changed that perception. SARS was found to be caused by SARS-CoV and in 2014 there was another outbreak in the Middle East which was found to be caused by MERS-CoV (Middle East Respiratory Syndrome Coronavirus). These outbreaks have led to more research into their replication, transmission, morphogenesis, pathogenesis, and distribution.[4]

This review focuses on the importance of structural proteins, non-structural proteins, and accessory proteins of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and their key role in the viral machinery, pathogenesis, transmission as well as importantly how the mutation in these lead to the emergence of various mutants. Understanding the role of these proteins, and mutations in these proteins will help us to understand the variants and mechanism behind their immune escape, increased transmissibility, pathogenesis, virulence, etc., many research and review articles have been documented about any one of the structural proteins, non-structural proteins, and accessory proteins. In the present review, we have discussed all the structural proteins, non-structural proteins, and accessory proteins of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) as well as the variants with the information about the mutation.

Methods

An extensive literature review was done using the WHO database about Covid- 19, CDC Covid -19 information, ICTV website, Pubmed, Medline, Pubmed Central, Embase, and Google Scholar databases. A combination of the Keywords “COVID-19”, OR “2019- nCoV” OR “SARS-CoV-2”, OR “Coronavirus”, OR AND “virology” was used during the literature survey. From the information obtained from the above-said sources, this review was drafted.

Results

Characteristics of coronavirus virions

Coronavirus virions possess a single-stranded positive-stranded RNA (+ssRNA) genome which is unimolecular. The nucleic acid of the coronavirus is capped, and polyadenylated which makes them more infective. Virion structure (subfamily Coronavirinae) is shown in Figure 2.

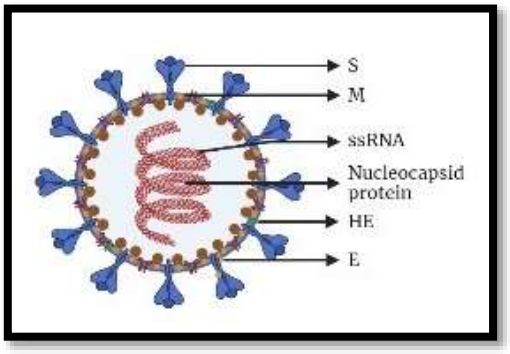


Fig 2 – Structure of Coronovirus -Virion structure (subfamily Coronavirinae) with structural proteins.

Proteins in coronaviruses and their characteristics, function, and antigenic properties are illustrated in table 1

Table 1: Proteins in coronaviruses

Protein	Characteristic	Function	Antigenic property
Spike glycoprotein (S)	Homo-trimeric type I membrane glycoprotein	It is a Fusion protein. It is responsible for the binding with the receptor of the host cell and also for the fusion. HE – Hemagglutinin-esterase	Inducer of virus-neutralizing antibodies
Membrane glycoprotein (M)	Integral type III membrane protein with N- or O-linked glycans	Thickness of envelope	They are responsible for inducing antibodies
Envelope protein (E)	pentameric integral membrane protein	Assembly and morphogenesis of virions. HE – Hemagglutinin-esterase an envelope protein facilitates S protein binding and release of virions.	Induces antibodies type which will prevent virion binding
Nucleocapsid protein (N)	RNA-binding phosphoprotein	Genome encapsidation, RNA synthesis and translation, RNA chaperone activity, and Type I interferon antagonist.	It is a dominant antigen during the natural infection N specific antibody have less immunity and it is used for seroevidence

Table data. Adapted from [2-4]

The genome of the coroaviridae consists of 5′ and 3′ untranslated regions (UTRs) and six open reading frames (ORFs) for the genome replication, encapsidation, and codes for the main structural proteins viz., S, E, M, and N. Apart from this, there are accessory genes that play a significant role in replication during infection. Then accessory genes also mutate for adapting to new hosts and niches which augurs well for the highly dynamic nature genome. Continuous RNA synthesis enables each mRNA to get transcribed into sg minus-strand RNA template.2

SARS – CoV2

Coronaviruses have been causative agents in several human infectious disease outbreaks such as SARS in 2002–2003 and the Middle East respiratory syndrome (MERS) in 2012.[5] In late December 2019 in Wuhan there was an outbreak of respiratory infections with symptoms of viral pneumonia alike the SARS and MERS6 which was found to be beta coronavirus type which has never been existed earlier.7 The complete viral genome of the beta coronavirus was sequenced by various organisations and was released. International Committee on Taxonomy of Viruses (ICVT) has tasked the Coronaviridae Study Group (CSG) for the virus classification and taxonomic nomenclature of the family Coronaviridae, The CSG after investigating the human pathogen, tentatively named it 2019-nCoV. The CSG recognized this virus as a sister clade to the prototype human and bat severe acute respiratory syndrome coronaviruses (SARS-CoVs) of the species severe acute respiratory syndrome-related coronavirus and designated it as SARS-CoV-2.8 Subsequently the World Health Organisation (WHO) named the disease as COVID-19. According to the uniprot taxonomy SARS-CoV-2 belongs to the following lineage.

Lineage

According to uniprot genome sequencing project database SARS-CoV-2 is of the realm- Riboviria, Nidovirales order, Coronavirinae suborder, Coronavirinae family, orthocoronavirinae subfamily, Betacoronavirus genus, Sarbecovirus subgenus and SARS coronavirus species.

SARS-CoV-2 is a single-stranded RNA (+ssRNA) virus. SARS-CoV-2 belongs to genus Beta-coronavirus of the Coronaviridae family. The size of the SARS-CoV- 2 genome is about 29 Kb as depicted in Figure 3, 4. It contains fourteen ORFs and encodes for more than 27 proteins.[9] Out of this there are two major ORFs namely ORF1a and ORF1b which is of the two third of the genome. Located near the 5′UTR, ORF1a encodes for polyprotein pp1a of ten non-specific proteins. ORF1b gene encodes for pp1ab (polyprotein) which can be processed into sixteen non-specific proteins. These non-specific proteins are cleaved by casein-like proteases for the viral replication complex.10 Apart from the above-mentioned two ORFs remaining viral genomes encode several accessory proteins as well as four structural proteins such as small envelope (E) protein, nucleocapsid (N) protein, spike (S) glycoprotein, and matrix (M) protein.[11] These accessory protein genes as mentioned earlier can mutate for adoption into new host and niche.

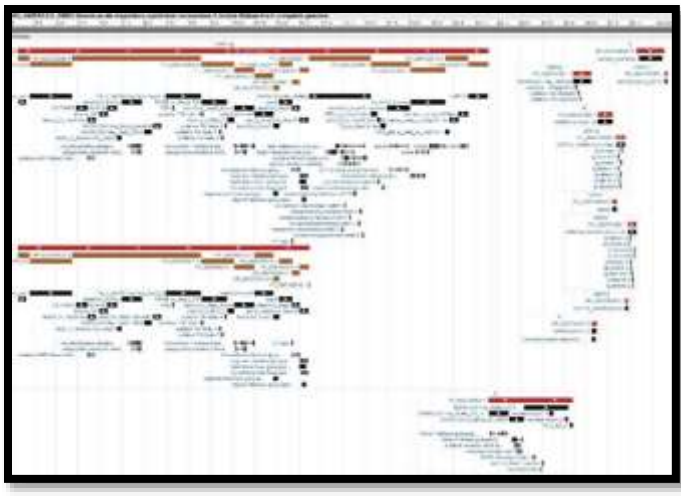


Fig 3- Complete Genome Organization of SARS-CoV-2

Obtained from
https://www.ncbi.nlm.nih.gov/projects/sviewer/?id=NC_045512&tracks

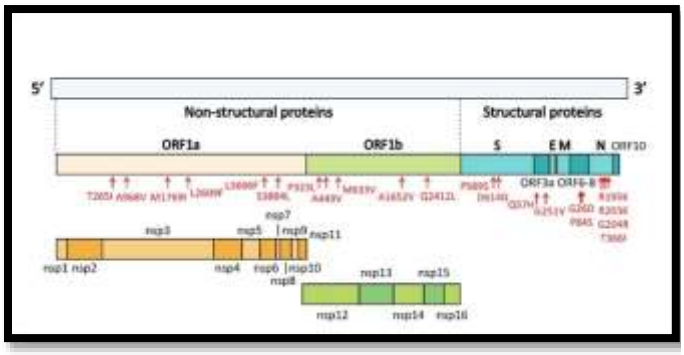


Fig 4 Genome organization of SARS-CoV-2. Credit: https://www.biorxiv.org/content/10.1101/2020.12.16.423178v1.full.pdf

SARS-CoV-2 structurally contains four important proteins namely Spike glycoprotein (S), Membrane protein (M), Small envelope glycoprotein (E) and nucleocapsid protein (N) as depicted in the figure 5. Apart from this they also possess non-structural proteins and accessory proteins which plays key role in their pathogenicity, assembly, release etc.

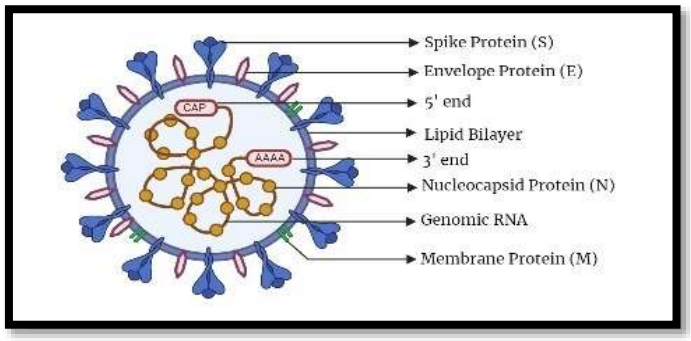


Fig 5 – Virion of SARS-CoV-2 with structural proteins and Genomic RNA

Structural proteins and their significance

The spike glycoprotein (S), membrane glycoprotein (M), envelope glycoprotein (E) and nucleocapsid protein (N) are the essential structural proteins. These proteins have an important and foremost role in the viral life cycle, pathogenesis, infectivity etc., The S protein is formed like a clove with two subunits S1 for receptor (Human ACE2) binding and S2 for membrane fusion.[12] The M protein is essential for the virion assembly inside the host cell and inducing the immune responses.[13] The N protein comprises of N Protein Terminal (NTD), serine-rich linker and C Protein Terminal (CTD).[14] It upgrades viral passage and performs post-combination cell measures fundamental for viral endurance in the host. The E protein is responsible for virion egress as well as the pathogenicity of the virus[15]. The significance of SARS-CoV-2 proteins and details are presented in table 2.

Table 2: Significance of SARS-CoV-2 proteins

Proteins	Significance
Spike glycoprotein (S) Molecular weight 180–200 kDa	The essential protein which is responsible for the entry of SARS-CoV- 2 into the host's Human Angiotensin Converting Enzyme cell 2 (hACE2). Roughly there are 15–30 freely rotating S proteins reported on the envelope of the virus.13 Spike protein consists of two subunits S1 for receptor binding and S2 for membrane fusion. Receptor Binding Domain (RBD) of the spike protein S1 will be attached with hACE2 in a form of trimer. Then it is activated / cleaved into S1 and S2 by proteolytic enzyme of the human which is called as Furin and it is found more in respiratory cells and S2 subunit fuses with the cell membrane of the host.16 Spike protein also down regulate host tethrin (which is responsible for inhibiting viruses including RNA virus) by lysosomal degradation.17 Spike protein also very important factor in mutation as discussed later in variants section.
Membrane protein (M) Molecular weight 18 kDa	M protein is responsible for assembly of virion. It inhibits the TBK1 (TNF receptor-associated factor family member-associated NF-κB activator-binding kinase 1) mediated innate antiviral immune response courtesy their ubiquitin pathway by degrading the TBK1 which in turn inhibits the phosphorylation of Interferon Regulatory transcription factor (IRF3) and thereby suppressing IFN-I production.[14] For regulating the localization of S protein in golgi apparatus where the virus budding takes place. It also aids in proliferation rapidly, evading against host immunity and in replication.[18]
Small envelope glycoprotein (E) Molecular weight 8.4 – 12 kDa	Virus morphogenesis and assembling of virions. It acts as viroporin (forming hydrophilic pores thereby participating in replication of virus and release from the host cells) thereby altering the cellular functions and calcium homeostasis. E protein also induces apoptosis in host cells. It is also responsible for increased production of IL1 beta (Interleukin-1β) by activating NLRP3 inflammasone of the host (innate immune system). They also involve in budding of virus in golgi apparatus.[19]
Nucleocapsid protein (N) Molecular weight 45 .6 kDa	For entering into the host cell N protein plays a major role. They are involved in viral particle release as well as in packaging of RNA. N protein forms the nucleoprotein core which is essential in viral replication, transcription, and translation.[20] They regulate various host cellular processes especially innate immune responses.[21]

Non- structural proteins and their significance

Non-structural proteins (NSPs) are polypeptides that lack ordinary sub-protein structure like the primary, secondary, or tertiary structures. The NSPs do have significant role in the life cycle and pathogenicity of the virus. Details of the NSPs are listed in table 3.

Table 3: Non-structural proteins (Coded by ORF1ab)

Proteins	Site of interaction in human cells	Functions
NSP 1: N-terminal product of the viral replicase	DNA replication, cytoskeleton	Inhibits translation of the host cells. Degrade host mRNAs. Inhibits the apoptosis thereby playing an important role thereby increasing the viral infectivity and growth.[20, 22-25]
NSP 2: N-terminal product	Vesicle trafficking	It binds to prohibitin 1 and prohibitin 2 and disrupts the intracellular host signalling.[22-25]
NSP 3: Papain-like proteinase	Not yet clearly understood	It is responsible for release of other non-structural proteins such as NSP1, NSP2, and NSP3. It also facilitates translation. Apart from this they also suppress the synthesis of proteins in the host cells.[22-25]
NSP 4: Membrane-spanning protein containing transmembrane domain 2	Mitochondria	It involves in the transcription by forming replication complex thereby playing an important role in replication and virion assembly.[22-25]
NSP 5: Proteinase and main proteinase	Epigenetics and gene expression regulation	Protease enzyme for cleaving the other matured and intermediate proteins. Protease activity.[22-25]
NSP 6: Putative transmembrane domain	Vesicle trafficking	It Induces auto phagosomes and double-membrane vesicles. It is involved in autophagy.[22-25]
NSP 7: RNA-dependent RNA polymerase	Vesicle trafficking	RNA polymerase activity.[22-25]
NSP 8: Multimeric RNA polymerase; replicase	Epigenetics and gene expression regulation, RNA processing Regulation, Signalling, Mitochondria	Primase activity by forming a complex with NSP 8 This NSP is significant for the replication and virulence.[22-25]
NSP 9: Single-stranded RNA-binding viral protein	Nuclear import machinery, extracellular matrix	Acts as a cofactor for exoribonuclease activity of NSP14.[22-25]
NSP 10: Growth-factor-like protein possessing two zinc binding motifs	Vesicle trafficking	Acts as a cofactor for exoribonuclease activity of NSP16, and NSP 14 .[22-25]
NSP 11: Consists of 13 amino acids (sadaqsflngfav) andidentical to the first segment of Nsp12	Not yet clearly understood	It is important for replication.[22-25]
NSP 12: RNA-dependent RNA polymerase	Not yet clearly understood	For replication and methylation activity.[22-25]

NSP 13: RNA-dependent RNA polymerase (Pol/RdRp)	Epigenetics and gene expression regulation, vesicle trafficking, signalling, Cytoskeleton	For the replication (Helicase) and transcription (RNA TPase activity). [22-25]
NSP 14: Proofreading Exoribonuclease domain	Not yet clearly understood	Exoribonuclease activity. for the stability of viral mRNA it should evade the host immunity.This is made possible by mRNA capping activity of Nsp14. [22-25]
NSP 15: EndoRNAse; nsp15-A1 and nsp15B-NendoU	Vesicle trafficking, nuclear import machinery	Prevents the detection of viral dsRNA by the host, which is made possible by Endoribonuclease activity, thereby helping the virus to evade immune system. [22-25]
NSP 16: 2′-O-ribose methyltransferase	Not yet clearly understood	It serves to protect from innate immune response of host cell by exoribonuclease activity response. [22-25]

Accessory proteins and their significance

There are twelve accessory proteins known so far. They do a key role in viral pathogenicity, evasion of immune response by host and other significant functions. The accessory proteins and their significance listed in table 4.

Table 4: Significances of accessory proteins

Proteins	Significance
ORF3a	Act as a viroporin, forming pores to facilitate the virus release20 and confers for maximal replication and virulence22 as well as induces apoptosis.[24-25]
ORF3b	it acts an interferon (IFN) antagonist by suppressing the induction of type I interferon. [24-25]
ORF 3c	Signalling activities as well as disrupting the membranes of the host.[24-25]
ORF 3d	Strongest antibody responses24-25 host modulation. [26]
ORF6	IFN antagonist. [24-25]
ORF7a	Antagonize IFN-I. [24-25]
ORF 7b	Interfere with some cellular processes and may cause heart rate dysregulation. [24-25]
ORF8	Involved in transmissibility and pathogenesis .24-25
ORF9b	By inhibiting the MAVS signalosome, it Suppresses innate immunity. [24-25]
ORF9c	Impairs the interferon signalling, antigen processing and presentation and complement signalling, thereby helping in evading the immune response.[24-25]
ORF 10	Not yet clearly understood
ORF 14	Not yet clearly understood

Pathogenesis

The pathogenesis starts with the binding of SARS-CoV-2 and human ACE2 cells of Type II alveolar epithelial cells which are present in respiratory system. Apart from the above location ACE2 receptors are also found in other organs. They are found in upper oesophagus, enterocytes from the ileum, myocardial cells, proximal tubular cells of the kidney as well as the urothelial cells of the bladder.17 It is also found in both male and female reproductive tissues.[27] For the binding of virus with host cells as mentioned earlier S - spike protein plays a key role. It resembles like a crown as depicted in the figure 5 and it is located the outer surface. As mentioned earlier it is of two subunits S1 and S2. The S1 subunit is further subdivided into a receptor-binding domain (RBD) and an N-terminal domain (NTD), which is involved in viral entry into the host cell which is the hACE2. TM protease serine 2 (TMPRSS2 protease which is present in the hACE2 activating the S protein and thereby promoting the viral entry into the cell. It is then followed by activation of the spike protein. The activation is facilitated by furin proteolytic cleavage by a host protease at S1 and S2 domains junction. The newly released S2 domain N-terminus into the cell membrane will fuse the viral and cellular membranes of the host. It is then followed by transfer of the viral RNA into the host cell cytoplasm where viral replication will occur.28 Virus enters the cell and releases viral RNA, translation of the RNA genome results into polyproteins. The replication and transcription of the viral RNA genome occur via protein cleavage and assembly of the replicase–transcriptase complex. Viral RNA is replicated in the host cell, and structural proteins are synthesised, assembled, and packaged before viral particles are released.[29] They enter host cells through one of two distinct pathways: (i) TMPRSS2 aided cell surface pathway or (ii) lysosomal cathepsins aided endocytic pathway as depicted in the figure 6.[30-32]

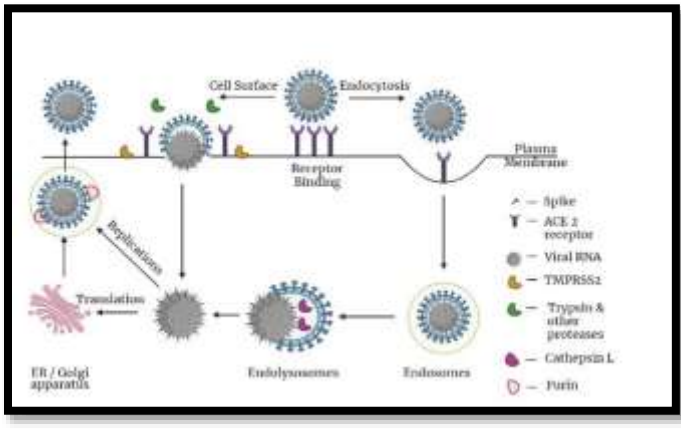


Fig -6 Entry of SARS-CoV-2 into the host cell

ACE2, angiotensin converting enzyme 2, TMPRSS2 – Trans Membrane Protein Serine Protease 2, ER, Endoplasmic reticulum, SARS-CoV-2, Severe Acute Respiratory Syndrome Coronavirus 2.

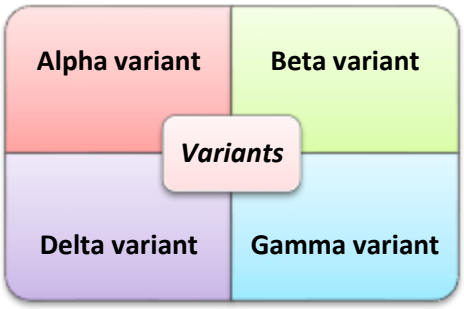
Variants

Mutations in the viral genome happens to adapt into new hosts and environments. It has been always for the part and parcel of the virus which increases their pathogenicity as well as to evade the immune response of the host. This can be termed as adaptive mutation. RNA viruses are more capable of this mutation than DNA viruses. If there is even a small change in the amino acid like tyrosine replacing the asparagine in N501Y it will impart in their immune evasion property and it will also affect the various therapeutics development like vaccines obviously. RNA viruses are known for its genetic evolution and also mutating themselves to adapt in their new human hosts. SARS-CoV -2 which is an RNA virus is no exception for such adaptive mutation for its survival and it is said to be done by horizontal transfer. Such mutations are a concern for health care professionals and scientific community all over the world. They result in the emergence of multiple variants. Such variants will differ from their ancestral strains. Continuous genomic sequencing will help us to identify such variants and act according to that.

The classification frameworks for naming SARS-CoV-2 hereditary ancestries are done by organisations such as GISAID, Nextstrain and Pango, it will be still used by researchers worldwide. WHO organized a meeting of scientific fraternity from GISAID, Nextstrain, Pango, virus researchers, experts in microbiological nomenclature etc., and proposed an idea of naming the variants in a way, that can be understood even by the non-scientific people. Such proposal was aimed at simple to-articulate and non-defaming marks for Variants of concern (VOI) and Variants of interest (VOC). Right now, the scientific fraternity proposed to use the Greek alphabets namely alpha, beta, gamma and delta. This nomenclature is aimed will be easier and more practical to be discussed by non-scientific audiences.33 Therefore, in this article we use the WHO nomenclature for easy understanding as well as the GISAID, Nextstrain and Pango nomenclature while discussing the variants.

Variants of concern

According to WHO a strain will be placed under VOC if that strain transmissible is increased manifold/if the virulence of the strain is increased/if the effectiveness of vaccines and therapeutics is decreased against that particular strain.



According to WHO as of September 12th, 2021 there are four variants of concern as follows.

Alpha variant

Nomenclature -Pango lineage – B.1.1.7 #, GISAID clade-GRY, Next strain clade -201 (V1). The variant was documented initially from U.K according to WHO. This Variant possess a mutation in their RBD of the spike protein. The amino acid asparagine is replaced with tyrosine which is termed as N501Y. This particular mutation increases the affinity of the spike proteins to ACE2 cells by introducing π–π interaction that enhances RBD binding to ACE2 which in turn enhancing viral entry. The mutation also stops binding of neutralizing antibody with virus.[34] This variant harbours Spike Protein Substitutions namely 69del, 70del, 144del, E484K*, S494P*, N501Y, A570D, D614G, P681H, T716I, S982A, D1118H.35

Beta variant

Nomenclature - Pango lineage – B.1.351, GISAID clade – GH/501Y.V2, Nextstrain clade-20H (V2). The variant was documented initially from South Africa according to WHO. This variant three mutations at key sites in the RBD (K417N, E484K and N501Y). Out of this two N501 Y and E484 K mutations enhances the binding with human ACE 2 cells.[36] The other mutation which is known as K417 increases the binding with Human ACE2 cells by forming a salt bridge interaction across the central contact region with D30 of Human ACE2. E484K mutation has been the reason for neutralization of natural antibodies.[37] These variant harbours spike protein substitutions namely D80A, D215G, 241del, 242del, 243del, K417N, E484K, N501Y, D614G, A701V.35

Gamma variant

Nomenclature - Pango lineage – P.1, GISAID clade – GR/501Y.V3, Nextstrain clade - 20J (V3). The variant was documented initially from Brazil according to WHO. This variant also has three changes in the RBD (K417T, E484K, and N501Y) like the beta variant which has mentioned before. These changes enhance the bonding of human ACE2 cells with RBD which in turn leading to increased transmissibility. Apart from that these residues do also have the capability of neutralizing the natural antibodies as well as the antibodies from the vaccines.38 This variant harbours Spike Protein Substitutions namely L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, D614G, H655Y, T1027I.35

Delta variant

Nomenclature - Pango lineage – B.1.617.2§, GISAID clade – G/478K.V1, Next strain clade - 21A. The variant was documented initially from India according to WHO. This variant was found to be having thirteen 13 amino acid changes. The three changes in spike proteins namely E484K, L452R, and P681R are attributed to its virulence, transmissibility, neutralize the antibodies produced naturally and artificially as well its resistance to various therapeutics. E484 K as described earlier attributes to their evasion to immune responses.[39] The L452 is the residue found on the spike protein within the RBD and enables it to bind, The L452 R mutation increases this binding to many folds.[40] P681R mutation enhanced the cleavage of the full-length spike to S1 and S2 which leads to increased infection via cell surface entry.[41] This variant harbour spike protein substitutions namely T19R, (V70F*), T95I, G142D, E156-, F157-, R158G, (A222V*), (W258L*), (K417N*), L452R, T478K, D614G, P681R, D950N.[35]

Omicron variant

Nomenclature - Pango lineage – B.1.1.529, GISAID clade – GRA, Next strain clade – 21K, 21L, 21M.The variant which is spreading, and dominating was documented from many countries according to WHO. The spike protein of this variant according to CDC harbours 30 amino acid substitutions namely A67V, del69-70, T95I, del142-144, Y145D, del211, L212I, ins214EPE, G339D, S371L, S373P, S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493R, G496S, Q498R, N501Y, Y505H, T547K, D614G, H655Y, N679K, P681H, N764K, D796Y, N856K, Q954H, N969K, L981F (Receptor Binding Proteins substitutions-G339D, S371L, S373P, S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493R, G496S, Q498R, N501Y, Y505H).The mutations N501Y and Q498R increases the binding with ACE-2 receptor thereby increasing infectious ability while H655Y, N679K, P681H mutations confers for the increased transmission according to CDC.51 K417N and N501Y mutations confers for the immune evasion – antibody escape property, while R346 confers to the neutralization of therapeutic antibodies.52 The mutations are also found in ORFs namely K856R, L2084I, A2710T, T3255I, P3395H, I3758V, P314L, I1566V, P10S substitution making the total mutations of around [60.53]

Variants of interest

A SARS-CoV-2 variant is termed as VOI, if any genetic changes in the strain do modify their evasion of immune response, transmissibility, severity of disease, decreased response to therapeutics and observed to cause significant

community transmission or multiple COVID-19 clusters in multiple countries, with increasing relative prevalence alongside increasing number of cases over time, or other obvious epidemiological impacts clearly showing an emerging risk to global public health.

Lambda variant

Nomenclature - Pango lineage – C.37, GISAID clade – GR/452Q.V1, Nextstrain clade – 21G.

The variant was documented initially from Peru according to WHO. It harbours seven nonsynonymous mutations in the Spike gene (Δ247-253, G75V, T76I, L452Q, F490S, and T859N) and a deletion in the ORF1a gene (Δ3675-3677). The mutations L452Q, F490S and L452R are attributed to increased affinity with hACE2 cells. The F490S also attributes to the neutralization of antibodies in vitro.[45]

Mu variant

Nomenclature - Pango lineage – B.1.621, GISAID clade – GH, Nextstrain clade – 21H. Earliest documented samples were from Colombia according to WHO. These variant harbours various spike protein mutations. There is amino acid changes I95I, Y144T, Y145S, there is also insertion 146N in the N-terminal domain. There are R346K, E484K and N501Y mutations in the RBD as well as the P681H in the S1/S2 cleavage site of the spike protein. There are also E484K and N501Y mutations which is present in previous strains. All these changes may attribute to neutralizing antibodies capability.[46]

Variants under Monitoring

The variants of whose phenotypic or epidemiological impact is not clear/unknown and may have a possibility of pose risk in the future are placed under the Variants under Monitoring as per World Health Organisation. The variants namely Eta, Iota and Kappa were placed among the Variants of Interest

Eta variant

Nomenclature - Pango lineage – B.1.525, GISAID clade – G/484K.V3, Nextstrain clade - 21D. Earliest documented samples were from multiple countries according to WHO. This variant harbours S: Q677H, F888L mutation (leucine replaces the phenylalanine) S: Del 69-70, S: Del 144, and S: E484K. This variant courtesy its mutations is thought to be capability of increased risk of disease severity and can evade the immune responses produced by vaccines [42] The spike protein substitutions reported for this strain are A67V, 69del, 70del, 144del, E484K, D614G, Q677H, F888L.[35]

Iota variant

Nomenclature- Pango lineage – B.1.526, GISAID clade – GH/253G.V1, Nextstrain clade - 21F. The variant was documented initially from United States of America according to WHO. This variant harbours E484K or S477N as well as the D253G change. The combination of two mutations484K or 477N may attribute to potential immune response evasion.43 The spike protein substitutions reported for this strain L5F, (D80G*), T95I, (Y144-*), (F157S*), D253G, (L452R*), (S477N*), E484K, D614G, A701V, (T859N*), (D950H*), (Q957R*).[35]

Kappa variant

Nomenclature - Pango lineage – B.1.617.1, GISAID clade – G/452R.V3, Nextstrain clade – 21B. The variant was documented initially from India according to WHO. These variant harbours combination of specific mutations L452R, E484Q and P681R which attributes their neutralizing ability against the natural antibodies as well as vaccine induced antibodies and also increased transmissibility.44 The Spike Protein Substitutions in this variant are (T95I), G142D, E154K, L452R, E484Q, D614G, P681R, Q1071H.35

AZ.5 variant

Nomenclature - Pango lineage – B.1.1.318 (A.Z.5), GISAID clade – GR, Nextstrain clade – Not assigned. The variant was documented in multiple countries according to WHO. These variant harbours combination of amino acid mutations T95I, E484K, D614G, P681H, D796H and one deletion in the Spike glycoprotein which is Y144del.These mutations are associated with the variant's ability of immune escape especially the mutations namely E484K, Y144del and also for the increased transmissibility.47-48

C.1.2 variant

Nomenclature - Pango lineage – C.1.2, GISAID clade – GR, Nextstrain clade – Not assigned. The variant was earlier documented in South Africa according to WHO. These variant harbours combination of amino acid mutations D614G, E484K, H655Y, N501Y, N679K, Y449H.It is one of the most mutated variants which has mutations predominantly in spike protein, nucleocapsid proteins which are associated with variant's ability of immune escape and for the increased transmissibility.[48-49]

B.1.630 variant

Nomenclature - Pango lineage – B.1.617.1, GISAID clade – G/452R.V3, Nextstrain clade – 21B. The variant was earlier documented in India according to WHO. These variant harbours combination of amino acid mutations L452R, T478K, D614G, P681R, E484X (d). According to European Centre for Disease Prevention and Control there are B.1.617.2 variant which harbours L452R, T478K, D614G, P681R and E484X (d) mutation, B.1.617.2 + Q613H variant which harbours L452R, T478K, D614G,

P681R and Q613H mutation, B.1.617.2 + Q677H variant which harbours L452R, T478K, D614G, P681R and Q677H mutation among this variant group. The mutations confer for the immune escape, enhanced ACE2 binding, neutralization of antibodies and also increased transmissibility.[48, 50]

Discussion

Covid – 19 caused by SARS-CoV-2, continues to be challenge to the world even though scientific world has come up with vaccines as there is an emergence of various variants of concern and various of interest. In order to understand about the new strains and to develop newer effective therapeutics comprehensive understanding about the molecular mechanism is need of the hour. As mentioned in the article the SARS-CoV-2 is an RNA virus which is capable of adopting to the new environment by mutation. Apart from the structural proteins which has been the centre of attraction SARS-CoV-2 accessory proteins and non-structural proteins also plays a very significant role in terms of immune evasiveness, pathogenicity and neutralization of antibodies which has been discussed in the article. The mutations in the SARS-CoV-2 discussed in the article has been mostly with the spike protein attributing for the increased affinity with the host, pathogenicity, immune evasiveness. Etc. There are other mutations which attributes to the neutralizing capability against the natural and vaccine induced antibodies. There is a continuous investigation and updating regarding the details of SARS-CoV-2 which need to be considered for the precautionary guidelines and in the development of therapeutics.

Conclusion

We have reviewed the information and compiled it for better understanding of the SARS-CoV-2 structural proteins, non-structural proteins and accessory proteins as well as about the mutant variants. More research about the SARS-CoV-2 structure, proteins and the molecular mechanism will provide us more insights in the future. Regarding the variants, we will come to know more about immune evasiveness, transmissibility, pathogenicity of the variants which have been already reported in coming days.

Abbreviations

C protein terminal (CTD), Coronaviridae study group (CSG), Coronaviruses (CoV), Envelope glycoprotein (E), Exoribonuclease (ExoN), Human angiotensin converting enzyme cell 2 (hACE2), International committee on taxonomy of viruses (ICVT), Membrane glycoprotein (M), Middle East Respiratory Syndrome Coronavirus (MERS-CoV), N protein terminal (NTD), Non-structural proteins (NSPs), Nucleocapsid protein (N), Open reading frames (ORFs), Receptor binding domain (RBD), Severe acute respiratory syndrome (SARS), Spike glycoprotein (S), Untranslated regions (UTRs), Variants of concern (VOI) , Variants of interest (VOC), World Health Organisation (WHO).

References

1. Mallucci L, McIntosh K, Tyrrell D. Virology: coronaviruses. Nature. 1968; 220:650. https://doi.org/10.1038/220650b0

2. De Groot R.J., Cowley, J.A, Enjuanes, L., Faaberg, K.S., Perlman, S., Rottier, P.J.M., Snijder, E.J., Ziebuhr, J. and Gorbalenya, A.E. Virus taxonomy (2020). https://talk.ictvonline.org/ictv-reports/ictv_9th_report/positive-sense-rna-viruses-2011/w/posrna_viruses/219/nidovirales

3. Ogando NS, Ferron F, Decroly E, Canard B, Posthuma CC, Snijder EJ. The curious case of the nidovirus exoribonuclease: its role in RNA synthesis and replication fidelity. Frontiers in microbiology. 2019 Aug 7; 10:1813. https://doi.org/10.3389/fmicb.2019.01813

4. Payne S. Family coronaviridae. Viruses. 2017:149. https://doi.org/10.1016/B978-0-12-803109-4.00017-9

5. Bermingham A, Chand MA, Brown CS, Aarons E, Tong C, Langrish C, Hoschler K, Brown K, Galiano M, Myers R, Pebody RG. Severe respiratory illness caused by a novel coronavirus, in a patient transferred to the United Kingdom from the Middle East, September 2012. Eurosurveillance. 2012 Oct 4; 17(40):20290.

6. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, Zhao X, Huang B, Shi W, Lu R, Niu P. T. Research; China Novel Coronavirus Investigating and Research Team. A novel coronavirus from patients with pneumonia in China. 2019:727-33.

7. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, Hu Y, Tao ZW, Tian JH, Pei YY, Yuan ML. A new coronavirus associated with human respiratory disease in China. Nature. 2020 Mar; 579(7798):265-9. https://doi.org/10.1038/s41586-020-2008-3

8. Gorbalenya AE, Baker SC, Baric RS, De Groot RJ, Drosten C, Gulyaeva AA, Haagmans BL, Lauber C, Leontovich AM, Neuman BW, Penzar D. The species severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nat. Microbiol. 5, 536–544. Link: https://go. Nature. Com/3cW9qJR. 2020.

9. Silva SJ, Silva CT, Mendes RP, Pena L. Role of nonstructural proteins in the pathogenesis of SARS-CoV-2. J Med Virol. 2020. 92: 1427-29.

10. Wu A, Peng Y, Huang B, Ding X, Wang X, Niu P, Meng J, Zhu Z, Zhang Z, Wang J, Sheng J. Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. Cell host & microbe. 2020 Mar 11; 27(3):325-8.

11. Cui J, Li F, Shi ZL. Origin and evolution of pathogenic coronaviruses. Nature Reviews Microbiology. 2019 Mar; 17(3):181-92. doi: 10.1038/s41579-018-0118-9

12. Shang J, Wan Y, Luo C, Ye G, Geng Q, Auerbach A, Li F. Cell entry mechanisms of SARS-CoV-2. Proceedings of the National Academy of Sciences. 2020 May 26; 117(21):11727-34. https://doi.org/10.1073/pnas.2003138117

13. Sui L, Zhao Y, Wang W, Wu P, Wang Z, Yu Y, Hou Z, Tan G, Liu Q. SARS-CoV-2 membrane protein inhibits type I interferon production through ubiquitin-mediated degradation of TBK1. Frontiers in immunology. 2021; 12. https://doi.org/10.3389/fimmu.2021.662989

14. Shepley-McTaggart A, Sagum CA, Oliva I, Rybakovsky E, DiGuilio K, Liang J, Bedford MT, Cassel J, Sudol M, Mullin JM, Harty RN. SARS-CoV-2 Envelope (E) protein interacts with PDZ-domain-2 of host tight junction protein ZO1. Plos one. 2021 Jun 9; 16(6):e0251955. https://doi.org/10.1371/journal.pone.0251955

15. Satarker S, Nampoothiri M. Structural proteins in severe acute respiratory syndrome coronavirus-2. Archives of medical research. 2020 Aug 1; 51(6):482-91. doi:10.1016/j.arcmed.2020.05.012

16. Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, Graham BS, McLellan JS. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science. 2020 Mar 13; 367(6483):1260-3.

17. Redondo N, Zaldívar-López S, Garrido JJ, Montoya M. SARS-CoV-2 Accessory Proteins in Viral Pathogenesis: Knowns and Unknowns. Frontiers in Immunology. 2021; 12. https://doi.org/10.3389/fimmu.2021.708264

18. Thomas S. The structure of the membrane protein of sars-cov-2 resembles the sugar transporter semisweet. Pathogens and Immunity. 2020; 5(1):342. Published 2020 Oct 19. doi:10.20411/pai.v5i1.37738

19. UniProtKB - P59637 (VEMP-SARS). https://www.uniprot.org/uniprot/P59637

20. Gordon DE, Jang GM, Bouhaddou M, Xu J, Obernier K, White KM, O'Meara MJ, Rezelj VV, Guo JZ, Swaney DL, Tummino TA. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature. 2020 Jul; 583(7816):459-68.

21. Bai Z, Cao Y, Liu W, Li J. The SARS-CoV-2 Nucleocapsid Protein and Its Role in Viral Structure, Biological Functions, and a Potential Target for Drug or Vaccine Mitigation. Viruses. 2021 Jun; 13(6):1115.

22. Yadav R, Chaudhary JK, Jain N, Chaudhary PK, Khanra S, Dhamija P, Sharma A, Kumar A, Handu S. Role of Structural and Non-Structural Proteins and Therapeutic Targets of SARS-CoV-2 for COVID-19. Cells. 2021 Apr; 10(4):821. https://doi.org/10.3390/cells10040821

23. Raj R. Analysis of non-structural proteins, NSPs of SARS-CoV-2 as targets for computational drug designing. Biochemistry and biophysics reports. 2021 Mar 1; 25:100847. https://doi.org/10.1016/j.bbrep.2020.100847.

24. Arya R, Kumari S, Pandey B, Mistry H, Bihani SC, Das A, Prashar V, Gupta GD, Panicker L, Kumar M. Structural insights into SARS-CoV-2 proteins. Journal of molecular biology. 2021 Jan 22; 433(2):166725. doi: 10.1016/j.jmb.2020.11.024

25. Gasmalbari E, Abbadi OS. Non-Structural Proteins of SARS-CoV-2 as potential sources for vaccine synthesis. Infectious Diseases & Tropical Medicine 2020; 6: e667. DOI: 10.32113/idtm_202010_667

26. UniProtKB - P0DTG0 (ORF3D-SARS2). https://www.uniprot.org/uniprot/P0DTG0

27. Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, Schiergens TS, Herrler G, Wu NH, Nitsche A, Müller MA. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. Cell. 2020 Apr 16; 181(2):271-80.

28. Murgolo N, Therien AG, Howell B, Klein D, Koeplinger K, Lieberman LA, Adam GC, Flynn J, McKenna P, Swaminathan G, Hazuda DJ. SARS-CoV-2 tropism, entry, replication, and propagation: Considerations for drug discovery and development. PLoS Pathogens. 2021 Feb 17;17(2):e1009225. https://doi.org/10.1371/journal.ppat.1009225

29. Huang Y, Yang C, Xu XF, Xu W, Liu SW. Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. Acta Pharmacologica Sinica. 2020 Sep; 41(9):1141-9. https://doi.org/10.1038/s41401-020-0485-4

30. Yang N, Shen HM. Targeting the endocytic pathway and autophagy process as a novel therapeutic strategy in COVID-19. International journal of biological sciences. 2020; 16(10):1724.

31. Millet JK, Whittaker GR. Physiological and molecular triggers for SARS-CoV membrane fusion and entry into host cells. Virology. 2018 Apr 1; 517:3-8.

32. Tang T, Bidon M, Jaimes JA, Whittaker GR, Daniel S. Coronavirus membrane fusion mechanism offers a potential target for antiviral development. Antiviral research. 2020 Jun 1; 178:104792.

33. Tracking SARS-CoV-2 variants. https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/

34. Yang TJ, Yu PY, Chang YC, Liang KH, Tso HC, Ho MR, Chen WY, Lin HT, Wu HC, Hsu ST. Effect of SARS-CoV-2 B. 1.1. 7 mutations on spike protein structure and function. Nature Structural & Molecular Biology. 2021 Aug 12:1-9. https://doi.org/10.1038/s41594-021-00652-z

35. SARS-CoV-2 Variant Classifications and Definitions. https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-info.html

36. Starr TN, Greaney AJ, Hilton SK, Ellis D, Crawford KH, Dingens AS, Navarro MJ, Bowen JE, Tortorici MA, Walls AC, King NP. Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding. Cell. 2020 Sep 3; 182(5):1295-310.

37. Baum A, Fulton BO, Wloga E, Copin R, Pascal KE, Russo V, Giordano S, Lanza K, Negron N, Ni M, Wei Y. Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies. Science. 2020 Aug 21; 369(6506):1014-8.

38. Dejnirattisai W, Zhou D, Supasa P, Liu C, Mentzer AJ, Ginn HM, Zhao Y, Duyvesteyn HM, Tuekprakhon A, Nutalai R, Wang B. Antibody evasion by the P. 1 strain of SARS-CoV-2. Cell. 2021 May 27; 184(11):2939-54. https://doi.org/10.1016/j.cell.2021.03.055.

39. Wise J. Covid-19: The E484K mutation and the risks it poses. BMJ 2021; 372:n359. https://doi.org/10.1136/bmj.n359

40. Mor O, Mandelboim M, Fleishon S, Bucris E, Bar-Ilan D, Linial M, Nemet I, Kliker L, Lustig Y, Mendelson ES, Zuckerman NS. The Rise and fall of a Local SARS-CoV-2 Variant with the Spike Protein Mutation L452R. Vaccines. 2021 Aug; 9(8):937. https://doi.org/10.3390/vaccines9080937

41. Frazier LE, Lubinski B, Tang T, Daniel S, Jaimes JA, Whittaker GR. Spike protein cleavage-activation mediated by the SARS-CoV-2 P681R mutation: a case-study from its first appearance in variant of interest (VOI) A. 23.1 identified in Uganda. bioRxiv. 2021 Jul 1.

42. Pereira F, Tosta S, Lima MM, Reboredo de Oliveira da Silva L, Nardy VB, Gómez MK, Lima JG, Fonseca V, de Oliveira T, Lourenço J, Alcantara LC. Genomic surveillance activities unveil the introduction of the SARS-CoV-2 B. 1.525 variant of interest in Brazil: Case Report. Journal of Medical Virology. 2021 May 15. https://doi.org/10.1002/jmv.27086

43. Lasek-Nesselquist E, Lapierre P, Schneider E, George KS, Pata J. The localized rise of a B. 1.526 variant containing an E484K mutation in New York State. medRxiv. 2021 Jan 1. https://doi.org/10.1101/2021.02.26.21251868

44. Edara VV, Lai L, Sahoo M, Floyd K, Sibai M, Solis D, Flowers MW, Hussaini L, Ciric CR, Bechnack S, Stephens K. Infection and vaccine-induced neutralizing antibody responses to the SARS-CoV-2 B. 1.617. 1 variant. bioRxiv. 2021 Jan 1. https://doi.org/10.1101/2021.05.09.443299

45. Romero PE, Dávila-Barclay A, Salvatierra G, González L, Cuicapuza D, Solis L, Marcos-Carbajal P, Huancachoque J, Maturrano L, Tsukayama P. The emergence of SARS-CoV-2 variant lambda (C. 37) in South America. medRxiv. 2021 Jan 1. https://doi.org/10.1101/2021.06.26.21259487

46. Laiton-Donato K, Franco-Munoz C, Alvarez-Diaz DA, Ruiz-Moreno H, Usme-Ciro J, Prada D, Reales J, Corchuelo S, Herrera-sepulveda M, Naizaque J, Santamaria G. Characterization of the emerging B. 1.621 variant of interest of SARS-CoV-2. medRxiv. 2021 Jan 1. https://doi.org/10.1101/2021.05.08.21256619

47. Genomic epidemiology of SARS-CoV-2 in Mauritius reveals a new wave of infections dominated by the B.1.1.318, a variant under investigation. Houriiyah Tegally et al doi: https://doi.org/10.1101/2021.06.16.21259017

48. https://www.ecdc.europa.eu/en/covid-19/variants-concern

49. SARS-COV-2 C.1.2 variant is highly mutated but may possess reduced affinity for ACE2 receptor Xiang-Jiao Yang https://doi.org/10.1101/2021.10.16.464644

50. Structure-activity relationships of B.1.617 and other SARS-CoV-2 spike variantsTzu Jing Yang, Pei-Yu Yu, Yuan-Chih Chang, Ning-En Chang, Yu-Xi Tsai, Kang Hao Liang, Piotr Draczkowski, Bertina Lin, Yong-Sheng Wang, Yu Chun Chien, Kay-Hooi Khoo, Han-Chung Wu, Shang-Te Danny Hsu https://doi.org/10.1101/2021.09.12.459978

51. https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/scientific-brief-omicron-variant.html

52. Miller, N. L., Clark, T., Raman, R., & Sasisekharan, R. (2021). Insights on the mutational landscape of the SARS-CoV-2 Omicron variant. BioRxiv: the preprint server for biology, 2021.12.06.471499. https://doi.org/10.1101/2021.12.06.471499

53.Xuemei He,Weiqi Hong,Xiangyu Pan,Guangwen Lu,Xiawei Wei, SARS-CoV-2 Omicron variant:

Characteristics and prevention.
https://doi.org/10.1002/mco2.110

How to cite this article:

Dinesh Kumar Lakshmi Narayanan, et.al; Structural, non-specific, and accessory proteins and variants of SARS-CoV-2: Current perspective. **J Pharm Res, 2023; 12(01):01-14**. **DOI: https://doi.org/10.5281/zenodo.7498622**

Conflict of interest: The authors have declared that no conflict of interest exists.
Source of support: Nil

