

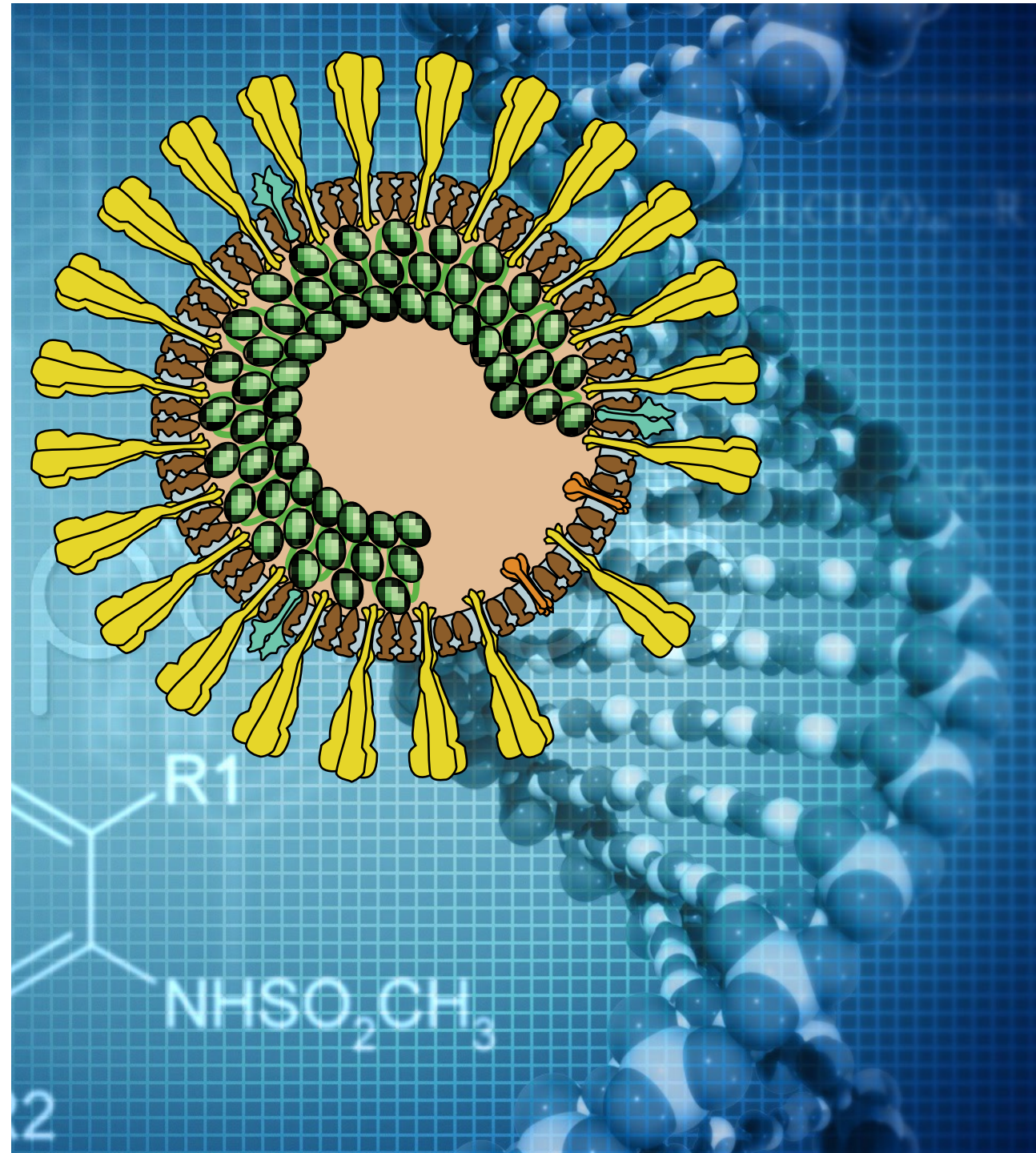


Swiss Institute of  
Bioinformatics

# VIRAL SUBSPECIES IN ViralZone and UniProtKB

Philippe Le Mercier, Swiss-Prot group

9th April 2024



# Viral subspecies in ViralZone and UniProtKB



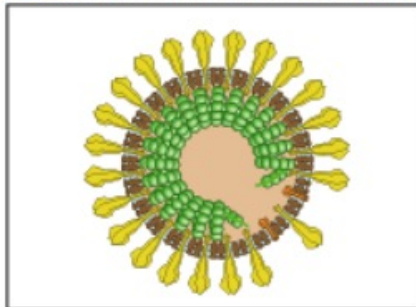
- **SARS-CoV-2 and Variants in ViralZone**
- Variants in UniProt
- Variants annotation/ metrics

# SARS-CoV-2, COVID-19 Coronavirus Resource

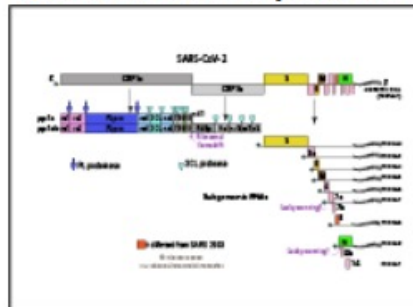


## ViralZone resources

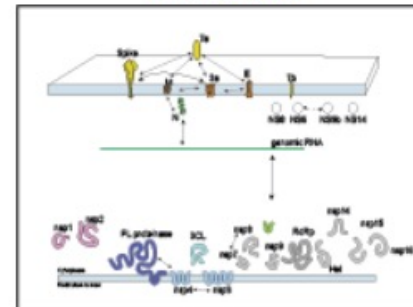
### Betacoronavirus fact sheet



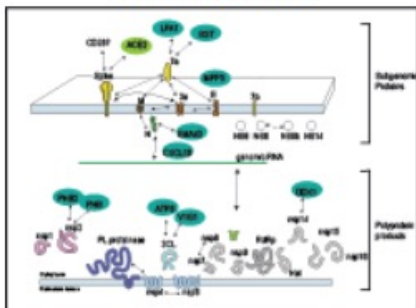
### Genome and expression



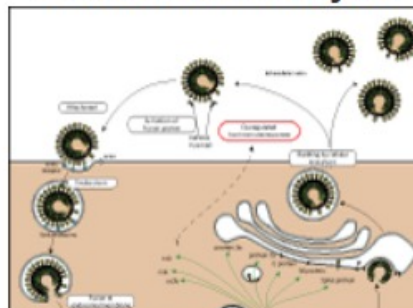
### Proteome



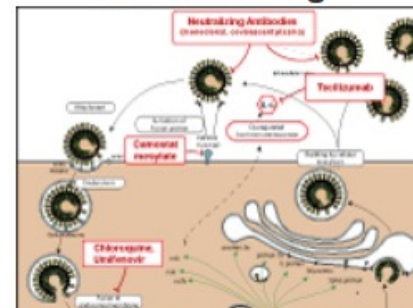
### Interactome



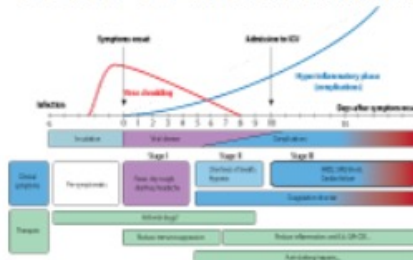
### Coronavirus life cycle



### Antiviral drugs

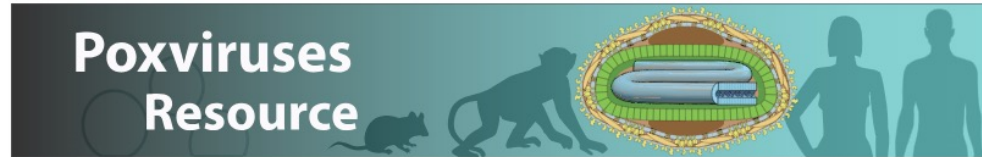


### COVID-19 and treatment



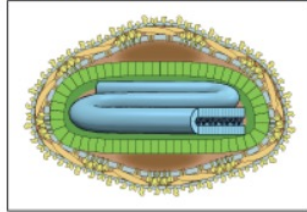


# Monkeypox resource

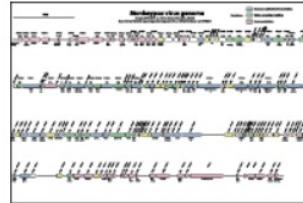


MonkeyPox is a misnomer, **the natural reservoir hosts are rodents**. The virus can spillover to monkeys and human but it is a rodent virus.  
The model organism for MonkeyPox is vaccinia virus.

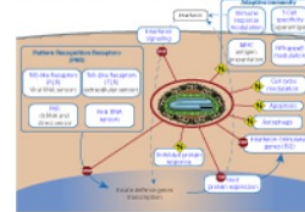
## Orthopoxvirus fact sheet



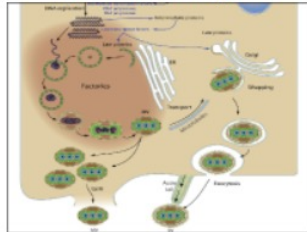
## Genome and expression



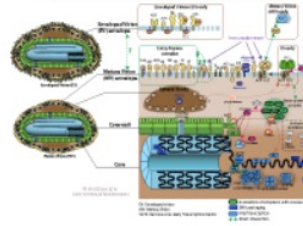
## Host-virus interactions



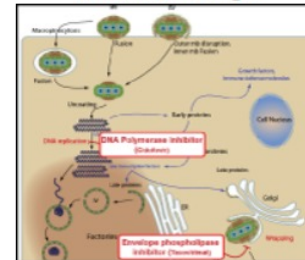
## Poxvirus life cycle



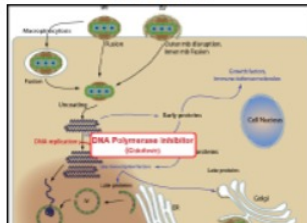
## Structural proteins



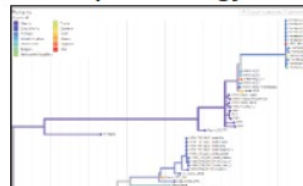
## Antiviral drugs



## Vaccines



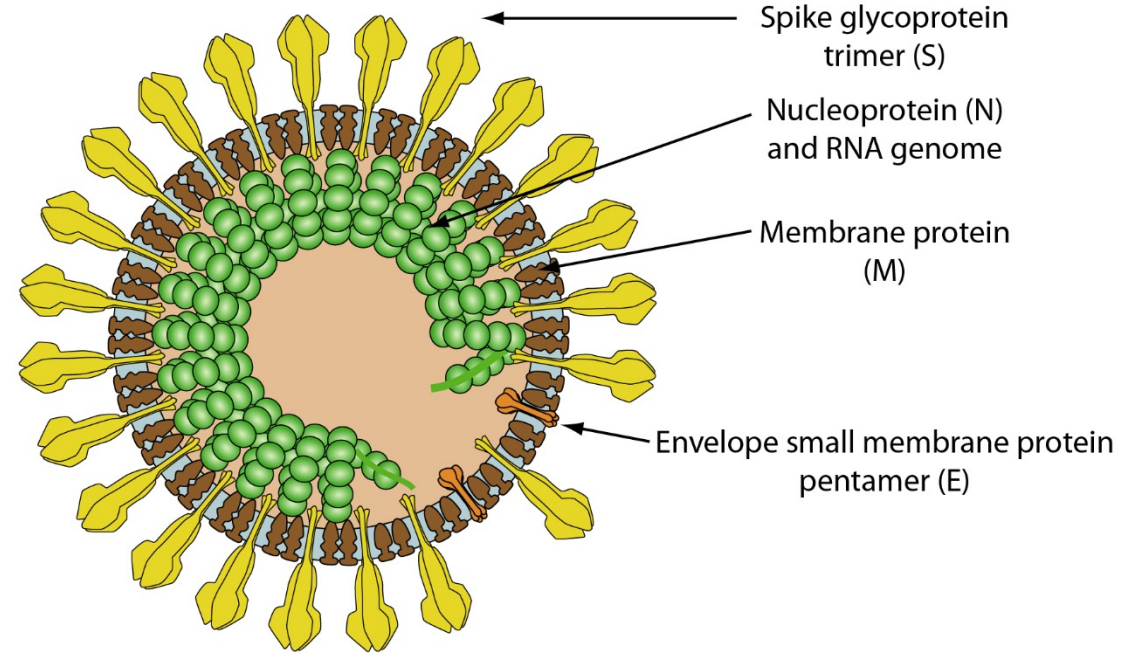
## Nextstrain: Monkeypox epidemiology





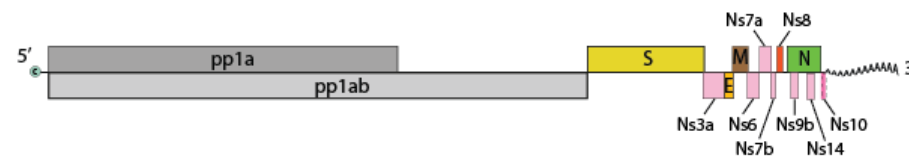
# SARS coronavirus 2

**Virion** about 120 nm in diameter



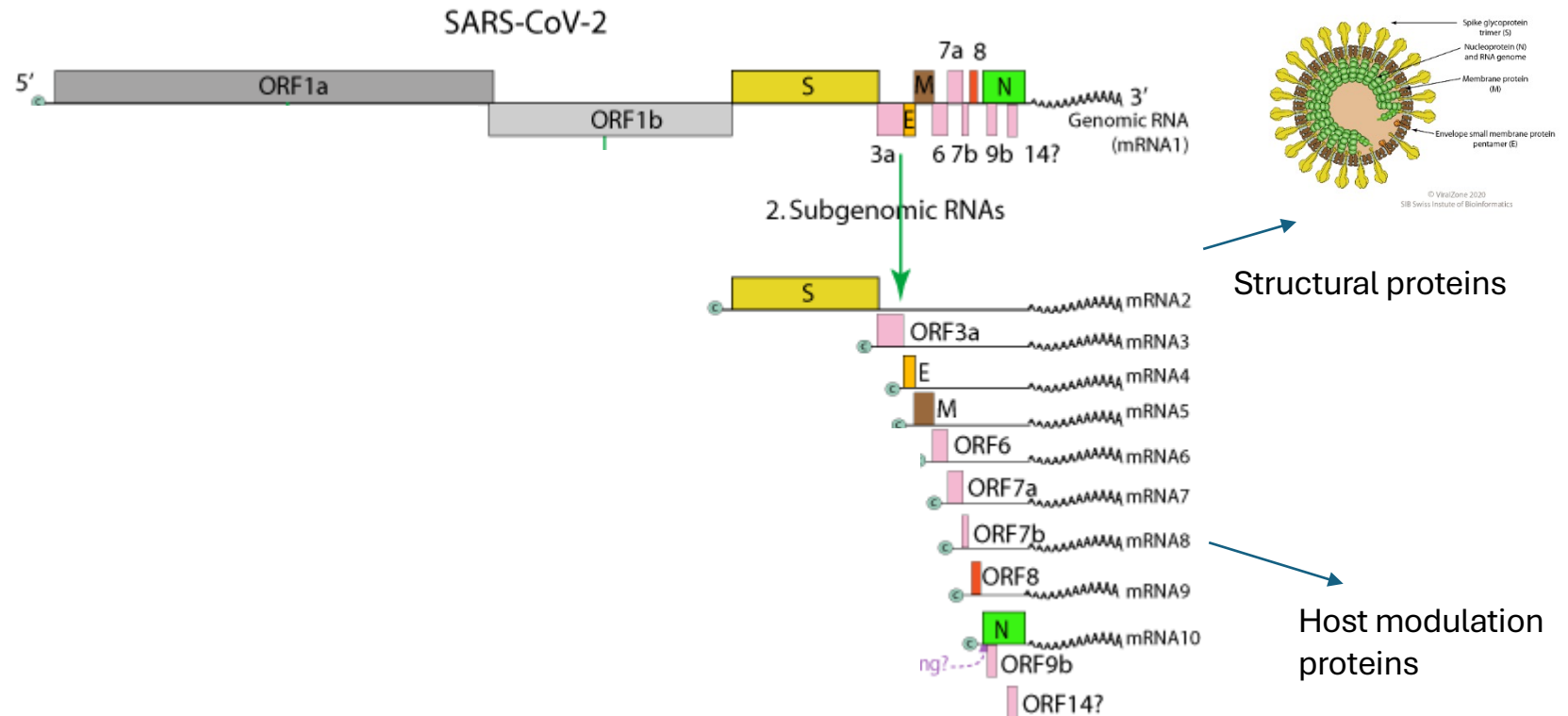
© ViralZone 2020  
SIB Swiss Institute of Bioinformatics

**Genome:** single stranded RNA messenger 29.9kb long, encoding 13 ORFs



# SARS-CoV-2 genome

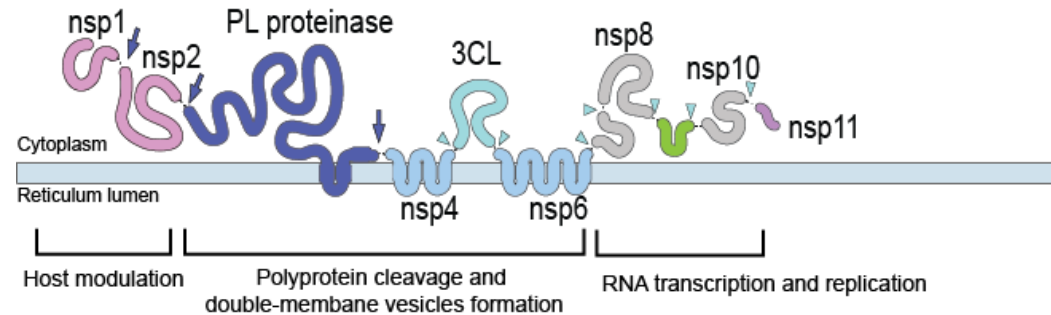
- 9 SgRNA are produced during viral replication by discontinuous transcription



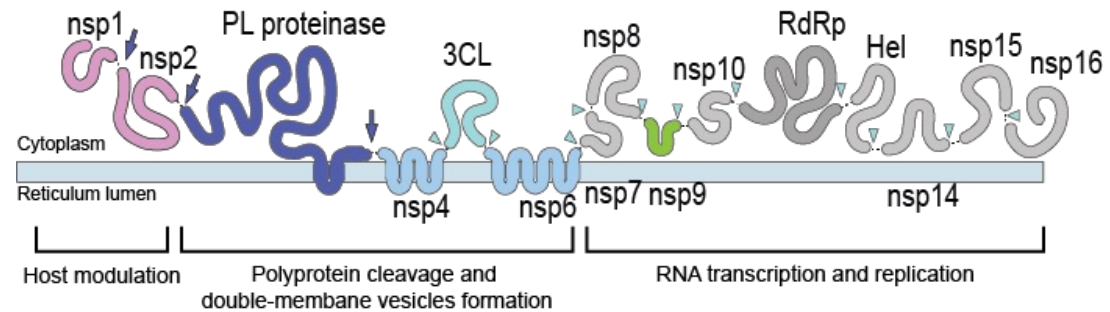
# Coronavirus polyproteins

Cleaved in 17 different chains, necessary for virus RNA transcription and replication.

ORF1a =  
polyprotein 1a (pp1a)

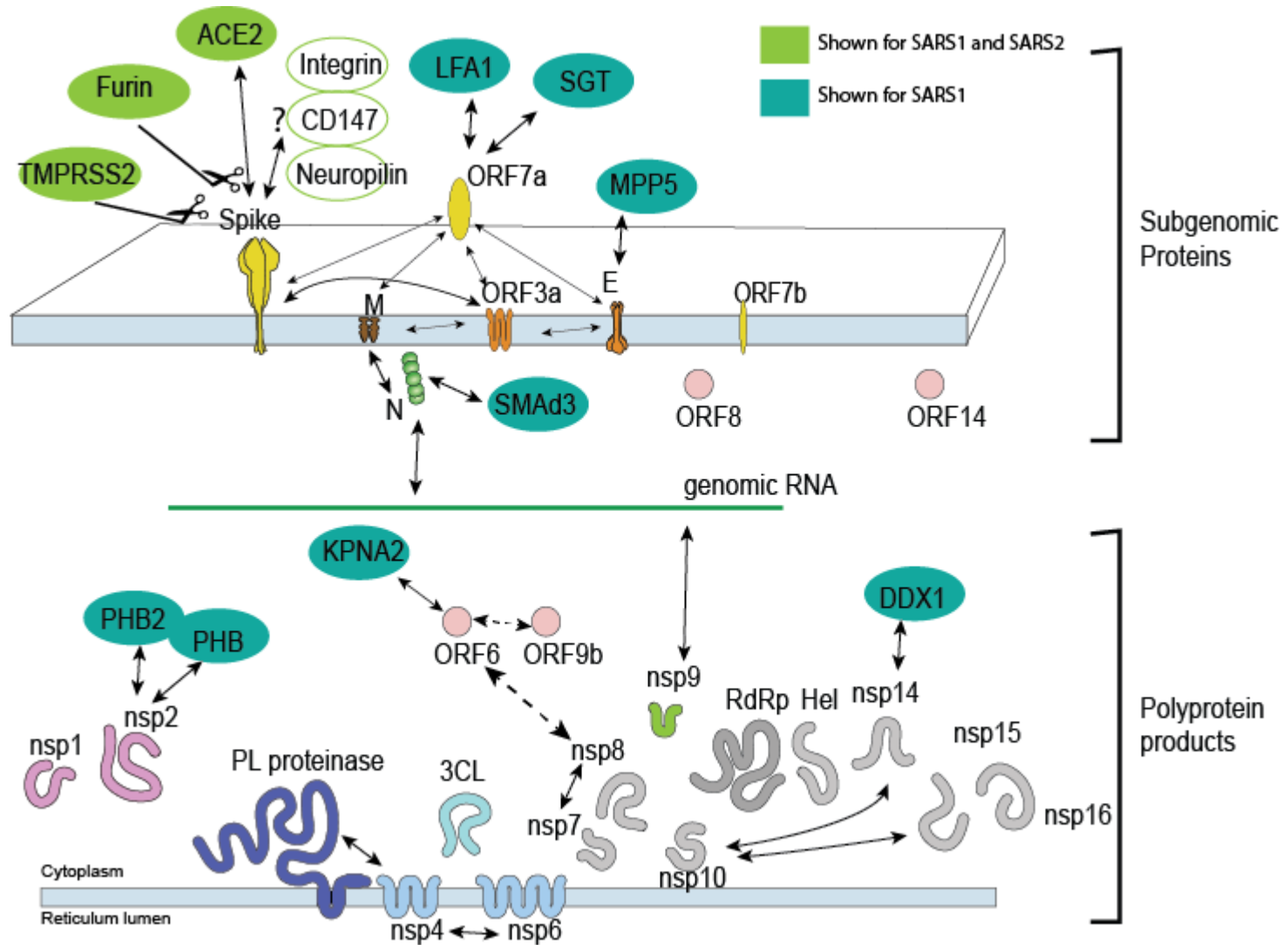


ORF1a frameshifted to ORFb =  
polyprotein 1ab (pp1ab)

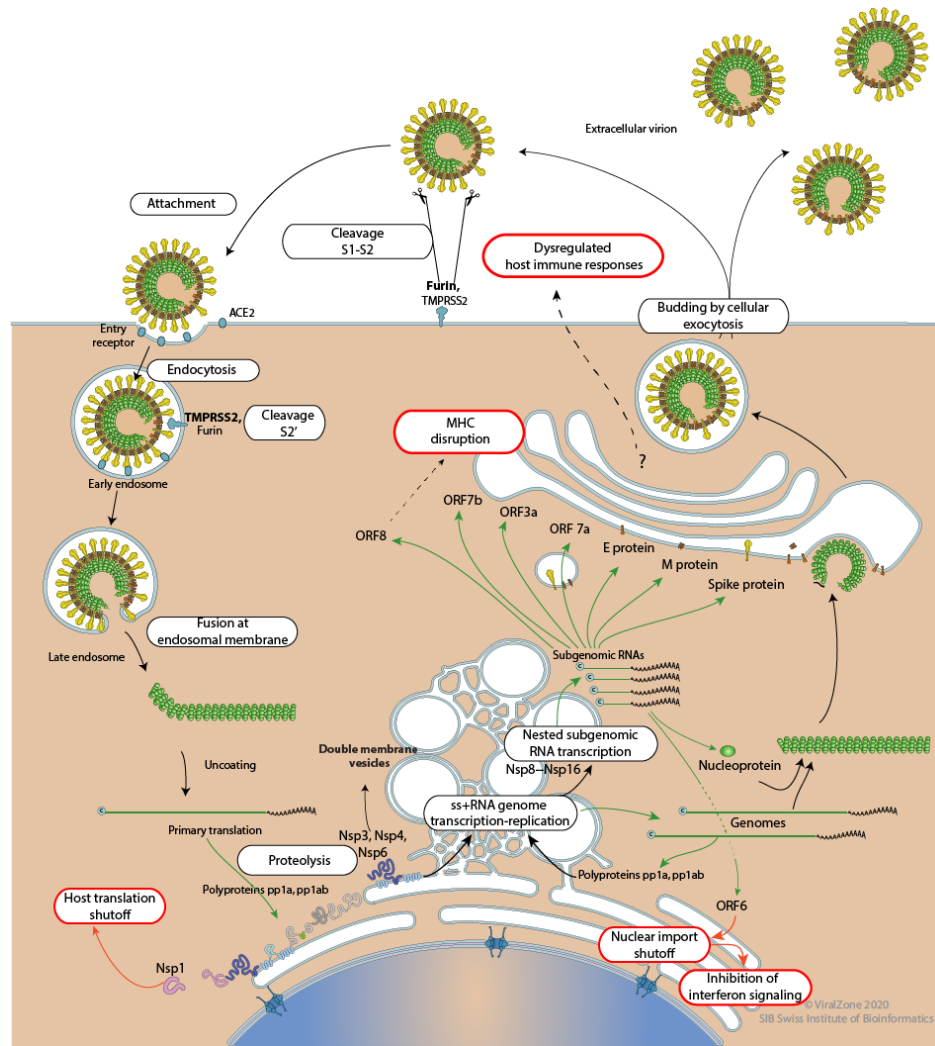




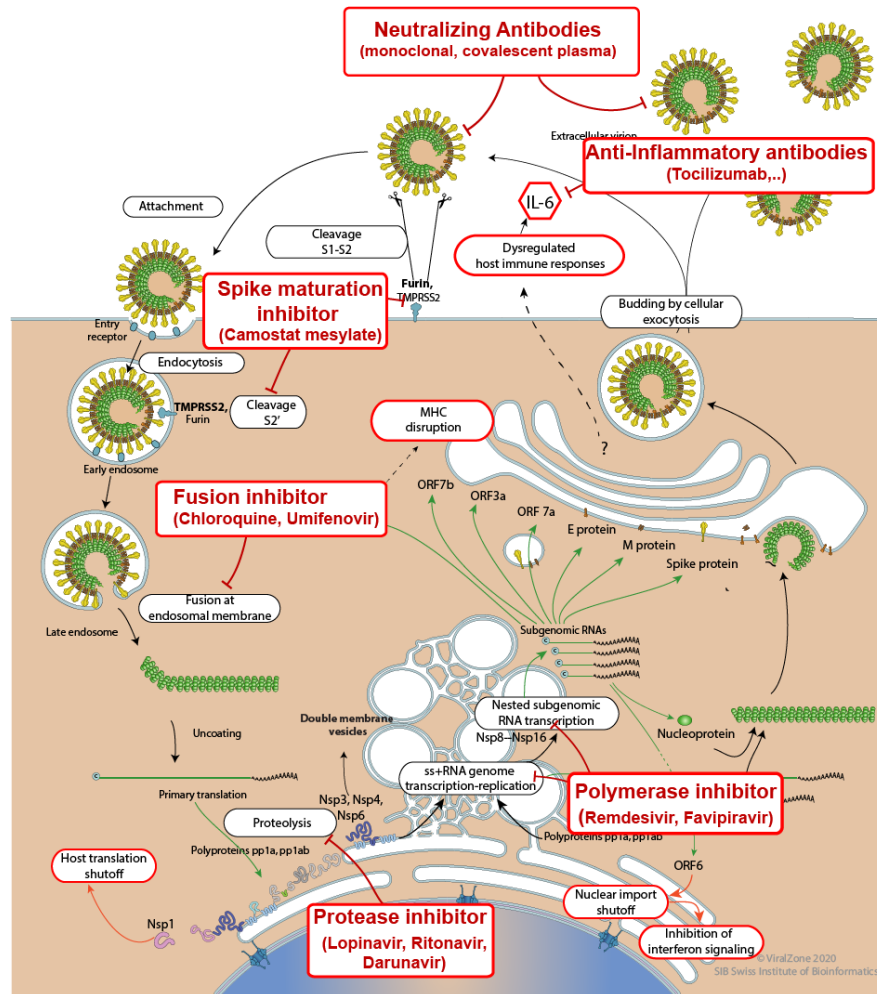
# SARS-CoV-2 curated interactome



# Coronavirus life-cycle



# Antivirals in the context of life cycle



## Compound

Camostat

Hydroxychloroquine

Umifenovir/Arbidol

Ritonavir, Lopinavir

Darunavir

Remdesivir

Favipiravir

## Ref

[\[1\]](#)

[\[2\]](#)

[\[3\]](#)

[\[4\]](#)

[\[5\]](#)

[\[6\]](#)

[\[7\]](#)

## Status for Covid-19

Under investigation

No-benefit in clinical trials

No-benefit in clinical trials

No-benefit in clinical trials

No evidence of SARS-CoV-2 inactivation

Weakly effective in clinical trials

No evidence of SARS-CoV-2 inactivation, under investigation

## Notes

TMPRSS2 inhibitor that may prevent Spike cleavage and activation. Spike could still be activated by furin

Non-specific anti-malaria drug. May prevent virus entry by modifying endosomal pH.

Anti influenza drug

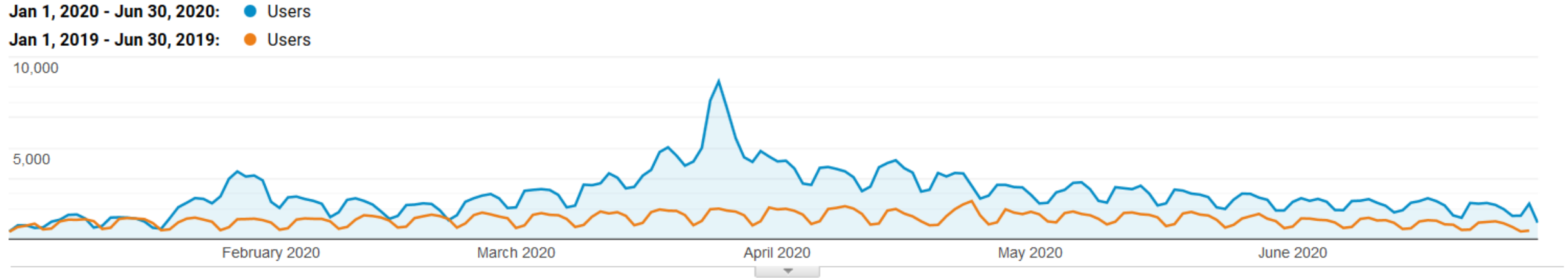
Molecules targeting HIV protease which is a ASP protease, but SARS-Cov proteases are THIOLE proteases.

Molecule under investigation against MERS and SARS coronaviruses.

Drug developed against influenza, redirected against ebola and now SARS-CoV-2

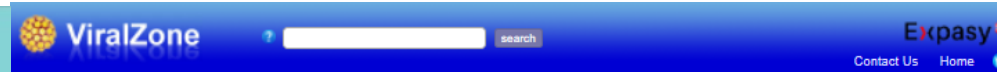


# Users found the coronavirus resource usefull



Source Google analytics

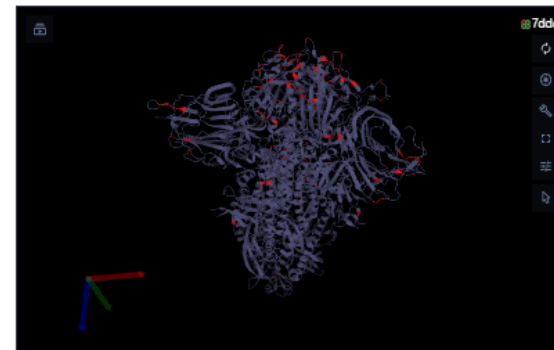
# Variants in ViralZone



## Sars-CoV-2 circulating variants

This page describes circulating SARS-CoV-2 variants - Last updated 9/Feb/2022.  
It includes all variants designated as "Variant of Concern" (VOC) or "Variant of Interest" (VOI) from WHO.  
Variant sequences can be parsed from UniProt P0DTC2 at COVID-19 pre-release portal.

Variants are lines that contain fixed mutations in their genome. Spike protein mutations affect both tropism (receptor binding) and immune evasion and are therefore the focus of surveillance. However, other viral protein mutations can also have effects on pathogenesis, cellular tropism and transmission.



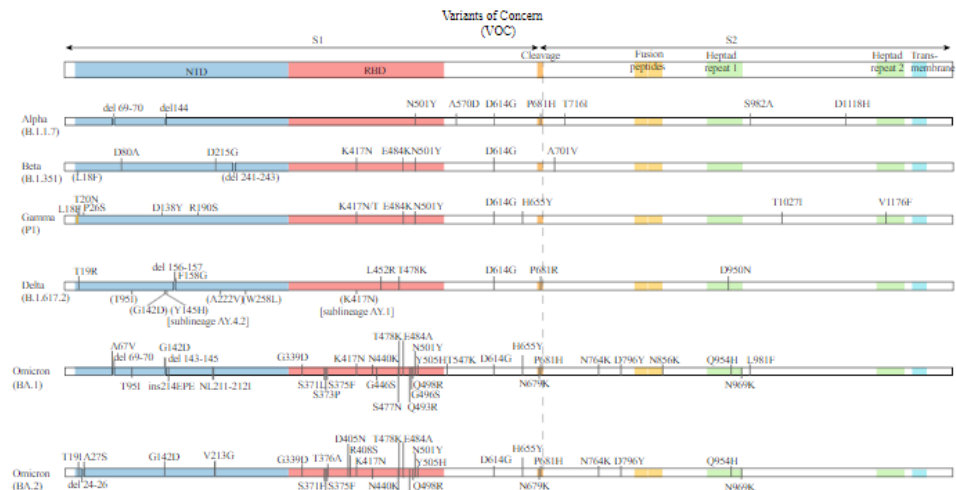
SARS-CoV-2 Spike protein: Click to highlight 3D structure regions of interest

- S1 chain
  - N-Terminal domain (NTD)-Antibody binding
  - Receptor binding domain(RBD)- Receptor and antibody binding
  - Main domains targeted by neutralizing antibodies
- S2 chain
  - Heptad repeat1 (HR1)- fusion helix
- RGD motif

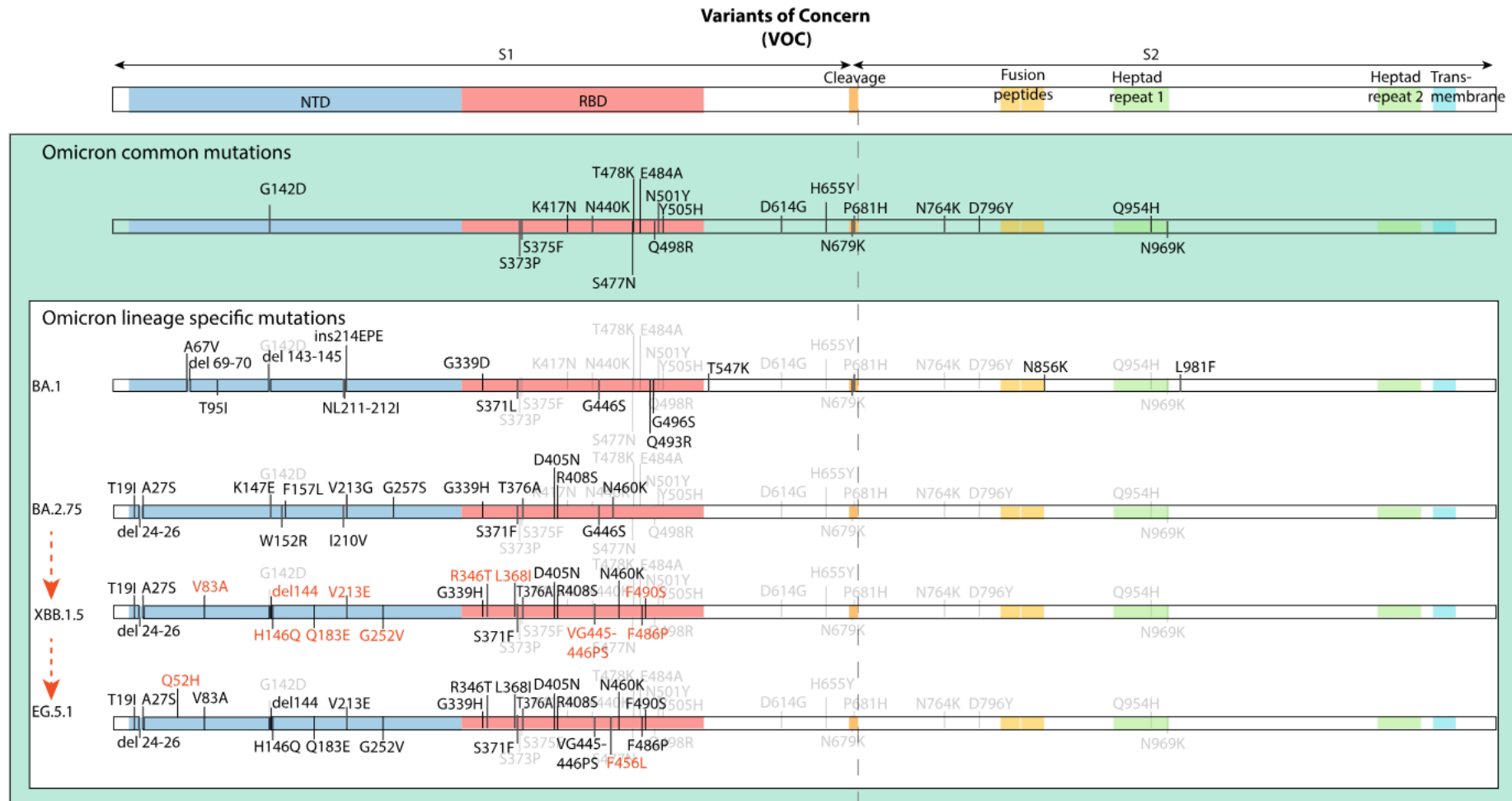
Variants of Concern (VOC): Click to display variant sites (zoom in FDB window for a better view)

- D614G
- B.1.1.7 (Alpha)
- B.1.351 (Beta)
- P1 (Gamma)
- B.1.617.2 (Delta)
- B.1.427, B.1.429 (Epsilon)
- B.1.1.529 (Omicron BA.1)

- Description of amino acid variations
- Links to
  - PANGO,
  - CoVariants.org,
  - outbreak.info,
  - BV-BRC
- Links to reference genomes /proteins



# Overview of circulating variants (2023)





# Variants data and links to reference sequences

## Reference sequence

Lineage	Synonyms	Emergence	Spike protein	Other proteins	Consensus sequences	Phenotypes
19A	Wuhan-Hu-1, nCoV	China, Dec 2019	None	(ORF8: L84S in outbreak.info <a href="#">reference</a> )	<a href="#">Genome</a> <a href="#">Spike protein</a> <a href="#">All proteins</a>	Reference for variants

**Variant of concern (VOC):** A variant for which there is evidence of an increase in transmissibility, more severe disease and or escape of neutralization by antibodies

Show  entries

Search:

Lineage	Synonyms	Emergence	Spike protein	Other proteins	Consensus sequences	Phenotypes
<b>Omicron BA.1</b> <a href="#">&gt;PANGO</a> <a href="#">CoVariants.org</a> <a href="#">oubtbreak.info:</a> <a href="#">/world prevalence</a> <a href="#">/mutations</a> <a href="#">BV-BRC</a>	BA.1 (previously B.1.1.529), 21K	South Africa, Dec 2021	<input type="button" value="show mutations"/>	<input type="button" value="show mutations"/>	<a href="#">Genome</a> <a href="#">Spike Protein NCBI</a>	Less pathogenic than previous variants because of S and NSP6 mutations
<b>Omicron BA.2</b> <a href="#">&gt;PANGO</a> <a href="#">CoVariants.org</a> <a href="#">oubtbreak.info:</a> <a href="#">/world prevalence</a> <a href="#">/mutations</a> <a href="#">BV-BRC</a>	BA.2 21L	South Africa, Dec 2021	<input type="button" value="show mutations"/>	<input type="button" value="show mutations"/>	<a href="#">Genome</a> <a href="#">Spike protein NCBI</a>	Omicrons are less pathogenic than previous variants because of S and NSP6 mutations
<b>Omicron BA.2.12.1</b> <a href="#">&gt;PANGO</a> <a href="#">CoVariants.org</a> <a href="#">BV-BRC</a> <a href="#">oubtbreak.info:</a> <a href="#">/world prevalence</a> <a href="#">/mutations</a>	BA.2.12.1 22C	North-America, Dec 2021	<input type="button" value="show mutations"/>	<input type="button" value="show mutations"/>	<a href="#">Genome</a> <a href="#">Spike protein NCBI</a>	Omicrons are less pathogenic than previous variants because of S and NSP6 mutations
<b>Omicron BA.2.75</b> <a href="#">&gt;PANGO</a> <a href="#">CoVariants.org</a>	BA.2.75 22D	India, spring 2022	<input type="button" value="show mutations"/>	<input type="button" value="show mutations"/>	<a href="#">Genome</a> <a href="#">Spike protein NCBI</a>	Omicrons are less pathogenic than previous variants



# CoV3D allowed to find an early reference sequence



CoV3D  
CORONAVIRUS 3D STRUCTURE DATABASE

Home Structures Sequences About Download FAQ Links Tools

Below are polymorphisms in SARS-CoV-2 spike glycoprotein sequences, with respect to original sequence used in full spike glycoprotein structures.

Search:

Mutations	Count	Example sequence	View mutation
A67V/H69del/V70del/T95I...	102836	UGW32340	<a href="#">View</a>
A67V/H69del/V70del/T95I...	43589	UFP04971	<a href="#">View</a>
A67V/H69del/V70del/T95I...	4937	UHV59178	<a href="#">View</a>
A67V/H69del/V70del/T95I...	2374	UIX13664	<a href="#">View</a>
A67V/H69del/V70del/T95I...	2069	UHV86077	<a href="#">View</a>
A67V/H69del/V70del/T95I...	1536	UGO98052	<a href="#">View</a>
A67V/H69del/V70del/T95I...	1034	UHV74268	<a href="#">View</a>
L5F/A67V/H69del/V70del/...	960	UHT27752	<a href="#">View</a>
A67V/H69del/V70del/T95I...	959	UHW31637	<a href="#">View</a>
A67V/H69del/V70del/T95I...	555	UHW31532	<a href="#">View</a>
A67V/H69del/V70del/T95I...	394	UHW40387	<a href="#">View</a>
A67V/H69del/V70del/T95I...	386	UHV86291	<a href="#">View</a>
A67V/H69del/V70del/T95I...	385	UHW38802	<a href="#">View</a>

# Viral subspecies in ViralZone and UniProtKB



- SARS-CoV-2 and Variants in ViralZone
- **Variants in UniProt**
- Variants annotation/ metrics

# Early access to SARS-CoV-2 annotation in UniProt

UniProt releases are published every 8 weeks.  
A pre-release portal has been established to provide users with quick access to the SARS-CoV-2 annotation.

The screenshot shows the UniProt website interface. At the top, there is a search bar with 'UniProtKB' selected and a search button. Below the search bar, there are navigation links: BLAST, Align, Retrieve/ID mapping, Peptide search, SPARQL, Help, and Contact. A blue banner below the navigation links reads 'The new UniProt website is here! Take me to UniProt BETA'. Below this banner, a message states: 'The mission of UniProt is to provide the scientific community with a comprehensive, high-quality and freely accessible resource of protein sequence and functional information.' To the right of this message is a 'Basket' icon with a '2' next to it.

The main content area is divided into several sections:

- UniProtKB** (UniProt Knowledgebase):
  - Swiss-Prot (567,483)**: Manually annotated and reviewed. Records with information extracted from literature and curator-evaluated computational analysis.
  - TrEMBL (231,354,261)**: Automatically annotated and not reviewed. Records that await full manual annotation.
- UniRef**: Sequence clusters.
- UniParc**: Sequence archive.
- Proteomes**: Proteome sets.
- Supporting data**: Literature citations, Cross-ref. databases, Taxonomy, Diseases, Subcellular locations, and Keywords.

On the right side, there is a 'News' section with social media icons (Blog, Twitter, Facebook, RSS). The news items include:

- Forthcoming changes**: Planned changes for UniProt.
- UniProt release 2022\_02**: Prenylation for antiviral activity | Cross-references to AlphaFoldDB | Version numbers for identifiers in Ensembl cross-references in Uni...
- UniProt release 2022\_01**: A phospholipase for clear vision | Cross-references to MANE-Select.

At the bottom of the page, there are three main sections:

- Getting started**: Includes a 'Text search' icon.
- UniProt data**: Includes a 'Download latest release' icon and the text 'Get the UniProt data'.
- Protein spotlight**: Features a 'Sapped' article from May 2022.



# Covid-19 UniProt website

<https://covid-19.uniprot.org/>

The screenshot shows a web browser displaying the UniProtKB COVID-19 portal. The address bar shows the URL [covid-19.uniprot.org/uniprotkb?query=\\*](https://covid-19.uniprot.org/uniprotkb?query=). The page title is "COVID-19 UniProtKB 125 results".

**COVID-19 UniProtKB 125 results**

This site provides the latest available pre-release UniProtKB data for the SARS-CoV-2 coronavirus and other entries relating to the COVID-19 outbreak. Therefore, data and functionality provided here may differ from the main UniProt.org website which is updated every eight weeks. This site will be updated as new relevant information becomes available, independent of the general UniProt release schedule.

This data can also be accessed via our FTP on [ftp://ftp.uniprot.org/pub/databases/uniprot/pre\\_release/](ftp://ftp.uniprot.org/pub/databases/uniprot/pre_release/)



You can view our latest webinar about Sars-CoV-2 and this portal from [here](#), presented on the 2nd April 2020.

Please go to the [UniProt.org](https://www.uniprot.org) website for all other entries and functionalities.

You can view all COVID-19 related publications contributed by the community [here](#).

You can view all curated LitCovid COVID-19 related publications [here](#).

**Latest update:** 21-March-2022

**Download**  

**Status**

- Reviewed (Swiss-Prot) (125)

**Model organisms**

- Human (93)

**Other organisms**

- Severe acute respiratory syndrome coronavirus 2 (17)
- Severe acute respiratory syndrome coronavirus (15)

**Proteins with**






- 3D structure (97)
- Active site (16)
- Activity regulation (30)
- Alternative products(isoforms) (59)
- Binding site (13)

**More items**

**Protein Existence**

- PROTEIN\_LEVEL (119)
- UNCERTAIN (4)

**Protein Details:**

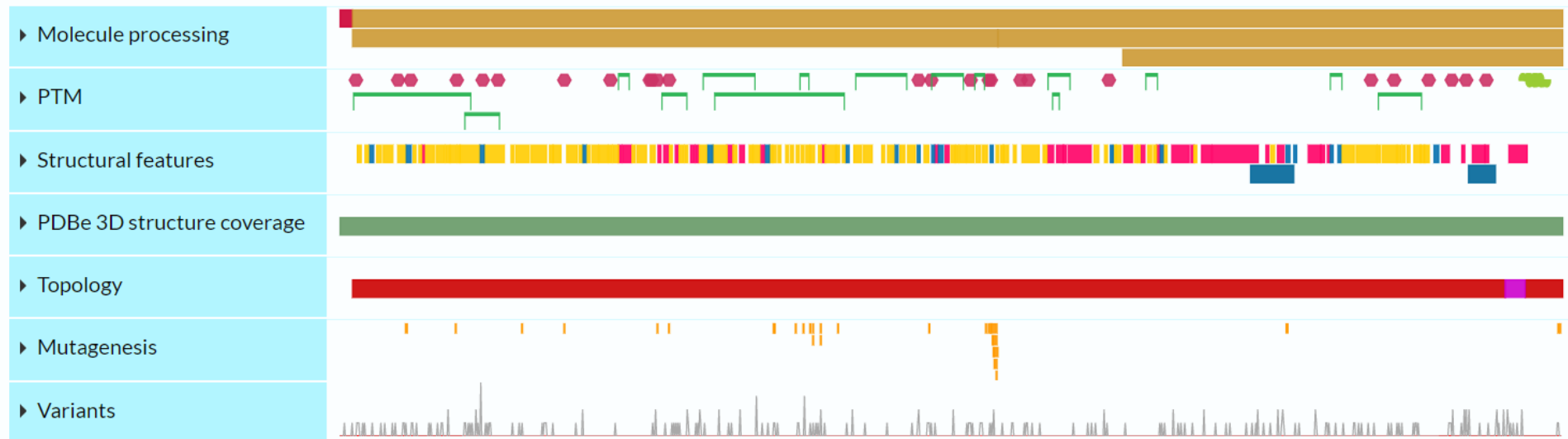
- O15393 · TMPS2\_HUMAN**  
Transmembrane protease serine 2 · Homo sapiens · Gene: TMRSS2 · 492 amino-acids ·   
[#Hydrolase](#) [#Protease](#) [#Serine protease](#)  
**3 domains · 9 reviewed variants · 3 active sites · 2 isoforms · 62 interactions · 13D structure · 28 reviewed publications**
- O95786 · DDX58\_HUMAN**  
Antiviral innate immune response receptor RIG-I · Homo sapiens · Gene: DDX58 · 925 amino-acids ·   
[#Helicase](#) [#Hydrolase](#) [#RNA-binding](#) [#Antiviral defense](#) [#Host-virus interaction](#) [#Immunity](#) [#Innate immunity](#)  
**5 domains · 6 PTMs · 4 reviewed variants · 2 isoforms · 29 interactions · 1 disease · 28 3D structures · 69 reviewed publications**
- O95992 · CH25H\_HUMAN**  
Cholesterol 25-hydroxylase · Homo sapiens · Gene: CH25H · 272 amino-acids ·   
[#Monooxygenase](#) [#Oxidoreductase](#) [#Lipid biosynthesis](#) [#Lipid metabolism](#) [#Steroid biosynthesis](#) [#Steroid metabolism](#) [#Sterol biosynthesis](#) [#Sterol metabolism](#)  
**1 domain · 1 reviewed variant · 2 interactions · 9 reviewed publications**
- P00973 · OAS1\_HUMAN**  
2'-5'-oligoadenylate synthase 1 · Homo sapiens · Gene: OAS1 · 400 amino-acids ·   
[#Nucleotidyltransferase](#) [#RNA-binding](#) [#Transferase](#) [#Antiviral defense](#) [#Immunity](#) [#Innate immunity](#)  
**6 reviewed variants · 4 isoforms · 12 interactions · 13D structure · 31 reviewed publications**
- P07711 · CATL1\_HUMAN**  
Procathepsin L · Homo sapiens · Gene: CTSL · 333 amino-acids · 

# Variant annotation in Swis-Prot

We annotate the WHO named variants in Swiss-Prot with FT VARIANT in the Wuhan 2020 reference proteome.

FT VARIANT	796	796	D -> Y (in strain: 19B/501Y, Omicron/BA.1, Omicron/BA.2) {*CUR}
FT VARIANT	856	856	N -> K (in strain: Omicron/BA.1) {*CUR}
FT VARIANT	859	859	T -> N (in strain: Lambda/C.37) {*CUR}
FT VARIANT	888	888	F -> L (in strain: Eta/B.1.525) {*CUR}

BLAST [Download](#) [Add a Publication](#) [Feedback](#)



# Nomenclature for deletions is not consistent

Outbreak.info: N211I, del212



CoVariant.org

Del211, L212I

BV-BRC

N211-, L212I

Emma B. Hodcroft



Aminoacid changes (2)

Deletion

S: N 211 -

Substitution

S: L 212 I

Nucleotide changes nearby (1)

Deletion

22194-22196

Context

Codon

Ref. AA

Ref.

Query

Query AA

1st nuc.

210	211	212	213
	N	L	
A T T	A A T	T T A	G T G
A T T	- - -	T A G	T G
	-	I	
22198	22193	22196	22199

FT VARIANT 211..212 FT /note="NL -> I (in strain: Omicron/BA.1)"  
FT /evidence="ECO:0000305"

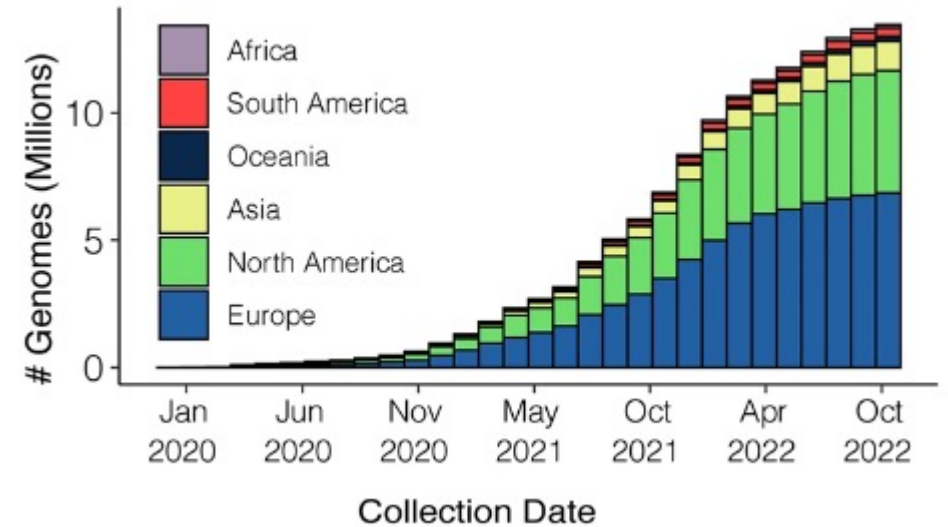
# Viral subspecies in ViralZone and UniProtKB



- SARS-CoV-2 and Variants in ViralZone
- Variants in UniProt
- **Variants annotation/ metrics**

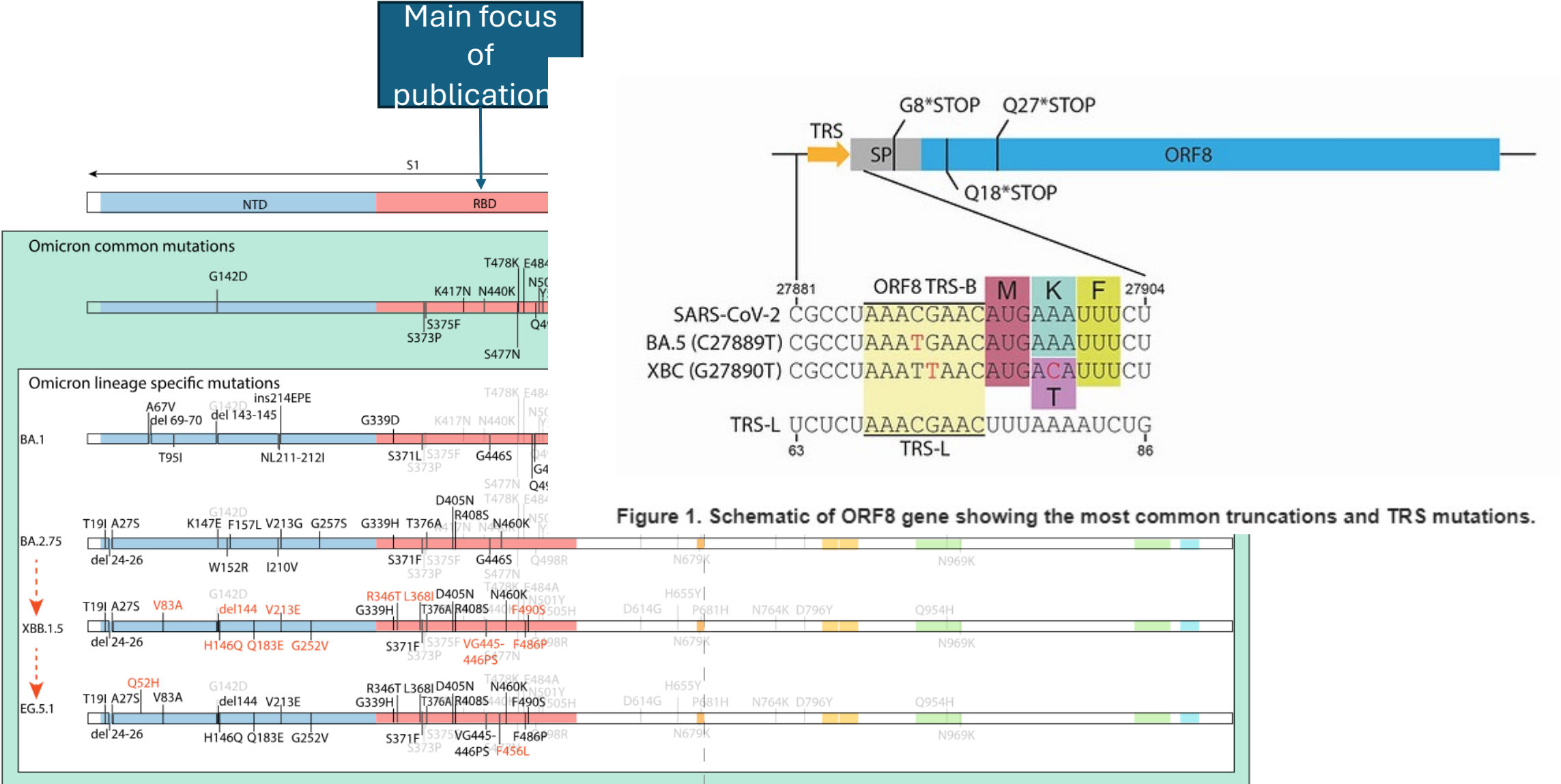
# Variants mutations: how to make sense of it

- Virus surveillance produces a lot of sequences
- But there are many more power for the production of sequences than for their analysis



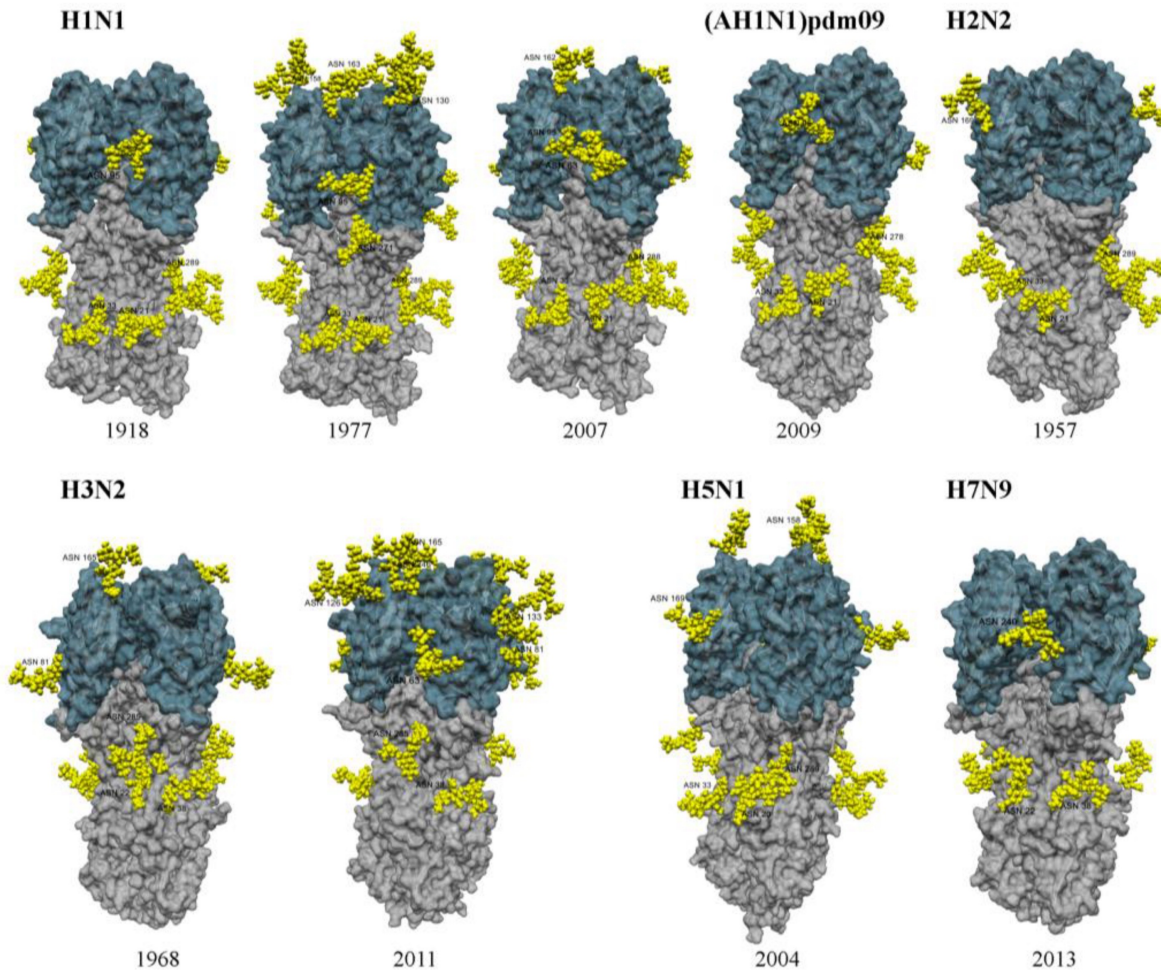


# Mutations considered in the context of domains

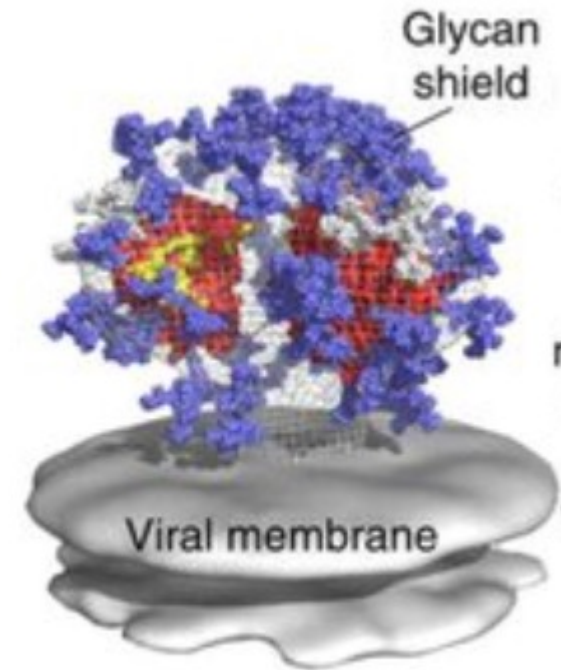


# Mutations considered in the context of glycosylation

Influenza glycosylation repositioning: immunity escape

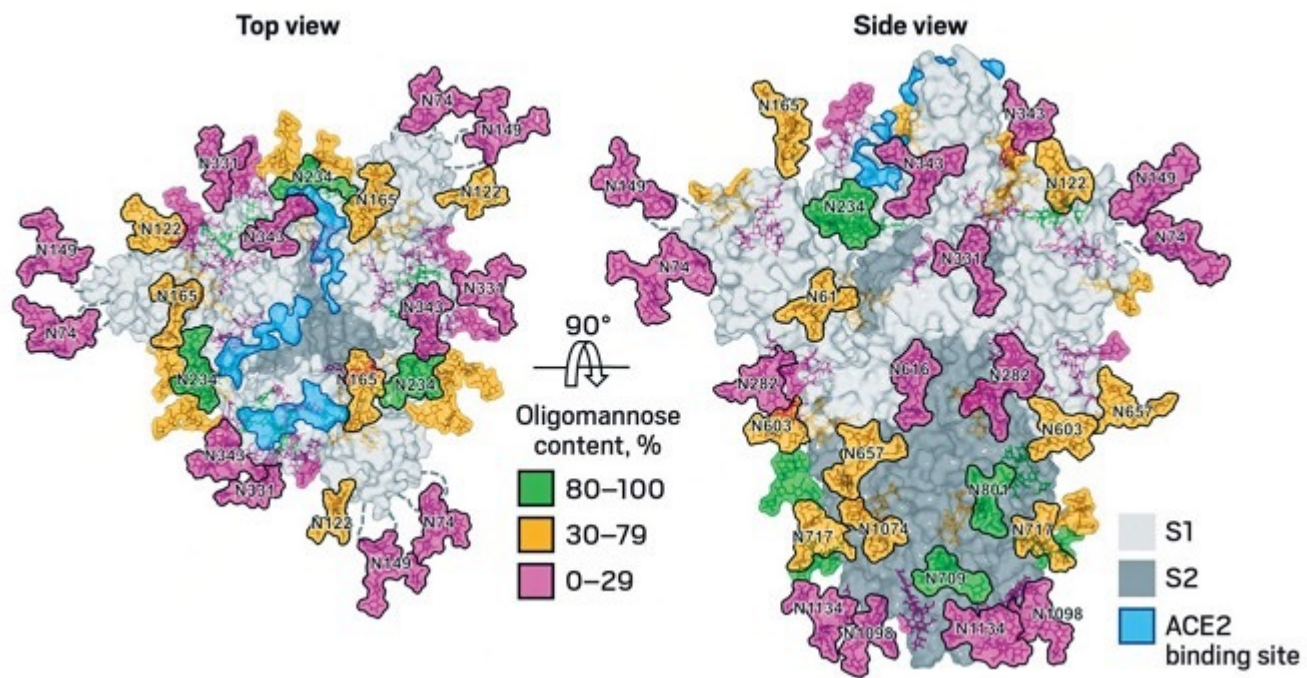


HIV glycan shield



# Do sars-cov-2 variants reposition glycosylation sites?

Sars-CoV-2 glycosylations



Glycan prediction using NetNGlyc-1.0

Output for 'UFO69279.1'

```

Name: UFO69279.1 Length: 1270
HFVFLVLLPLVSSQVNLTRTQLPPAYTMSFTRGVYQPKVFRSSVLHSTQDLFPFSSVTFVNFHSVSGTNGKRFDPN
VLPFNDGVYFASIEKSNIRGILFGTTLDDKTSQSLLVNATNVVIXKCEFOFCNDPFLDHNKNMSMESEFRVYSSANN
CTFEYVSQPLMDELKQGNKRLREFVKNIIDGYFKIVSKHTPIVREPEDLPQGSFALEPLVDLPIGINTRFQTLLA
LHRSYLPQDSSGATAGAAAYVQVLPRTFLKYNENGIITDWDICALDPLSETKCTLKSFTVEGIVQYSNFRVQPT
ESLVRPMTLMLCFDFEVNATRFASFVHNRKRSNKVADVSVLYLPPFTFKCYQVSPTKLNDLCTHMYADSVEIR
GDEVRIAPQQTGNADYNYKLPDFDTGCVIAHNSNKLDKVSQNYLYLFRKSNLKFPERDISFTEYQAGNKPNGV
AGNCVYPLRSYFRPTYGVGHQPVYRVVLSFELLHAPATVCGPKKSNLVKKNVNFENGLKGTGVLTSNKKFLPFQ
QFGRDIADTDAVRDPTLEILDITPCSFGSVITPDTNSQNAVLYQGVNCTEVPVAIHADLPTHRVYVSTGSNV
QTRAGGLIGAEVYNSYECIDIPIGAGCASYQIQTKSHRRARSQSISIVYHSLGASVASYNHSIAIPNFTISVT
TEILLPQSHTKSVKTYVEIGSSTCSNILLQSSFCQLKRALTGIAVEQGNKTEQVFAQVQCYKTPPIKRYGGAFS
QILPDPSPKRSFIEDLFFNKVTLADAGIKQYDGLDIAARDLCAQKFKGLTVLPLLLDDEIAQVTSALLAGTIT
SGITFGAGAAQLPFAMQAYRFGVGTQVNLVYENKLIANQFNSAIGKIQDLSSTASALGKLQDVNHNQAQLNTLV
KQLSSKFGAISVLENDIFSRDKVEAEVQDRILITGRLSLQTYVTQQLIRAAERASANAATKHECVLQGSKRDFPC
GKQVHLSFQSPAFHGVFLHWYVPAQKFNITADPAICHSQAHFRFREGVSKNGTHPVTQRVYEQIITTONTVS
GNCDDVIGIVNTPYDLPDSFKEELDKYKNNHTSPDVLDSGINSVWNIQEKIDRLNVEAKNLNLSLIDLQEL
GKYEQIKIPHYIWLGFJAGLTAIVMTIMLCCMTSCCCKLGGCCSCGCKFEDDSEPLVKGLKHYT
.....N.....N.....N.....N.....N.....
80
.....N.....N.....N.....N.....N.....
160
.....N.....N.....N.....N.....N.....
240
.....N.....N.....N.....N.....N.....
320
.....N.....N.....N.....N.....N.....
400
.....N.....N.....N.....N.....N.....
480
.....N.....N.....N.....N.....N.....
560
.....N.....N.....N.....N.....N.....
640
.....N.....N.....N.....N.....N.....
720
.....N.....N.....N.....N.....N.....
800
.....N.....N.....N.....N.....N.....
880
.....N.....N.....N.....N.....N.....
960
.....N.....N.....N.....N.....N.....
1040
.....N.....N.....N.....N.....N.....
1120
.....N.....N.....N.....N.....N.....
1200
GKYEQIKIPHYIWLGFJAGLTAIVMTIMLCCMTSCCCKLGGCCSCGCKFEDDSEPLVKGLKHYT
.....N.....N.....N.....N.....N.....
80
.....N.....N.....N.....N.....N.....
160
.....N.....N.....N.....N.....N.....
240
.....N.....N.....N.....N.....N.....
320
.....N.....N.....N.....N.....N.....
400
.....N.....N.....N.....N.....N.....
480
.....N.....N.....N.....N.....N.....
560
.....N.....N.....N.....N.....N.....
640
.....N.....N.....N.....N.....N.....
720
.....N.....N.....N.....N.....N.....
800
.....N.....N.....N.....N.....N.....
880
.....N.....N.....N.....N.....N.....
960
.....N.....N.....N.....N.....N.....
1040
.....N.....N.....N.....N.....N.....
1120
.....N.....N.....N.....N.....N.....
1200

```

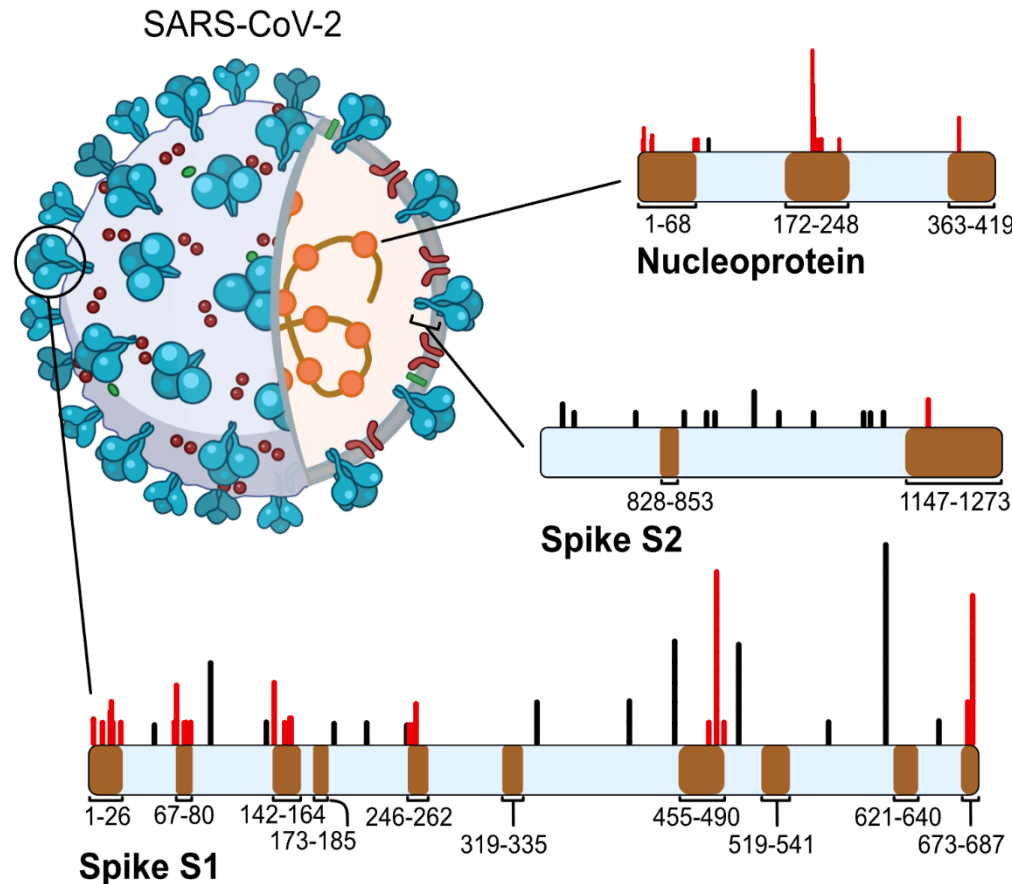
SeqName	Position	Potential	Pury	N-Glyc
UFO69279.1	17 NLTT	0.6607	(8/9)	+
UFO69279.1	61 NMTH	0.7863	(9/9)	+++
UFO69279.1	72 NGTK	0.7056	(9/9)	+++
UFO69279.1	120 NATN	0.6784	(8/9)	+
UFO69279.1	144 NKSW	0.5437	(6/9)	+
UFO69279.1	160 NCTF	0.6226	(8/9)	+
UFO69279.1	231 NIEP	0.7616	(9/9)	+++
UFO69279.1	279 NGTI	0.7381	(9/9)	+++
UFO69279.1	328 NITN	0.5888	(8/9)	+
UFO69279.1	340 NATR	0.5704	(8/9)	+
UFO69279.1	600 NTSN	0.5787	(6/9)	+
UFO69279.1	613 NCTE	0.7281	(9/9)	+++
UFO69279.1	654 NNSY	0.5141	(5/9)	+
UFO69279.1	706 NNSI	0.3533	(8/9)	-
UFO69279.1	714 NFTI	0.6427	(9/9)	+++
UFO69279.1	798 NFSO	0.6246	(9/9)	+++

(Threshold=0.5)

-> There was no significant changes in any variant spike glycosylation



# SARS-CoV-2 variants preferentially emerge at intrinsically disordered protein sites

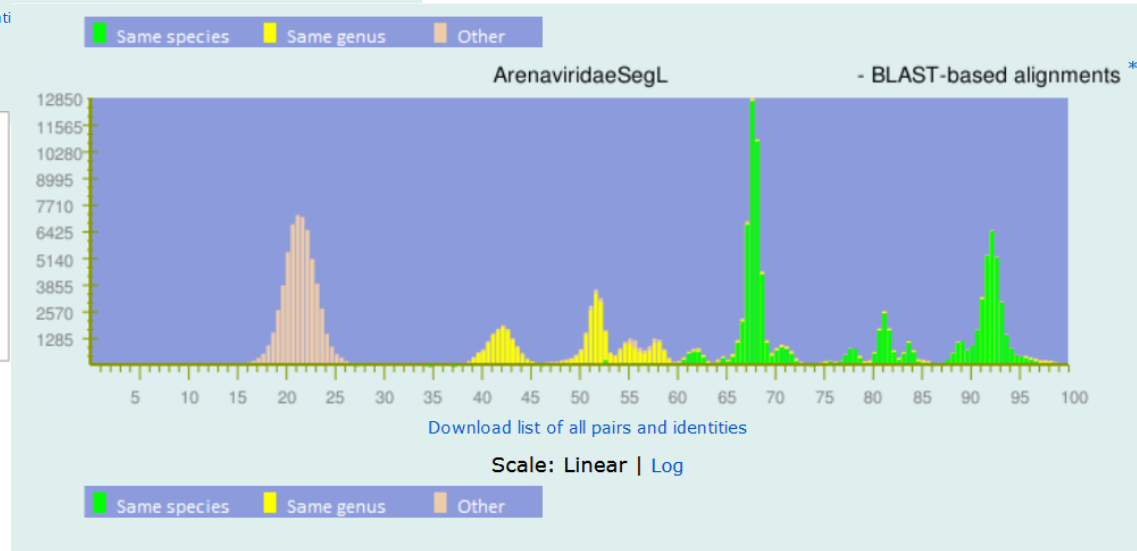
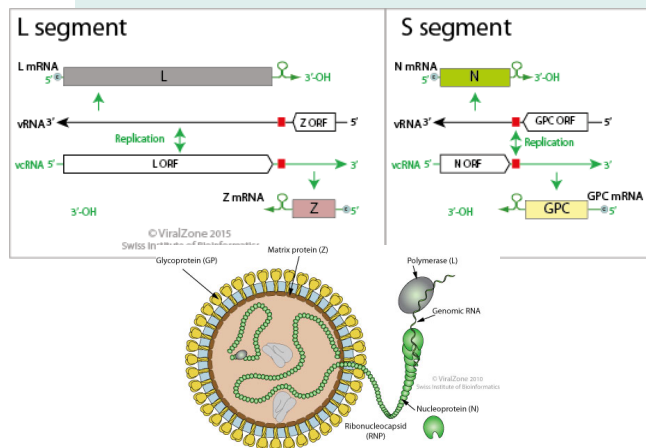
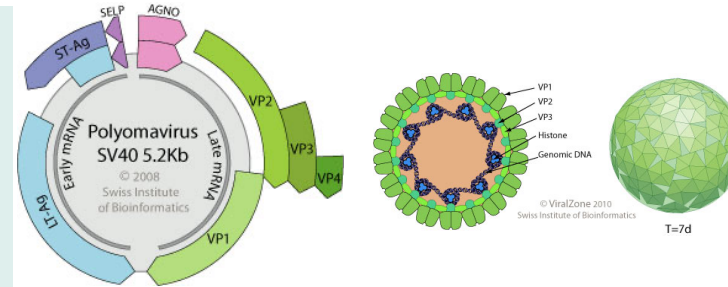
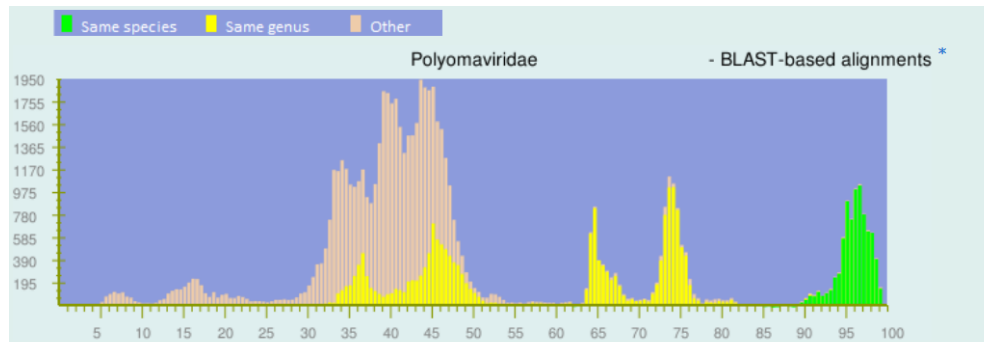


With the disprot group,  
We have established that variants  
occur preferentially at disordered  
regions in Spike S1 and  
Nucleoprotein

■ Disordered region      || Cumulative mutations on variants

# Genomic heterogeneity

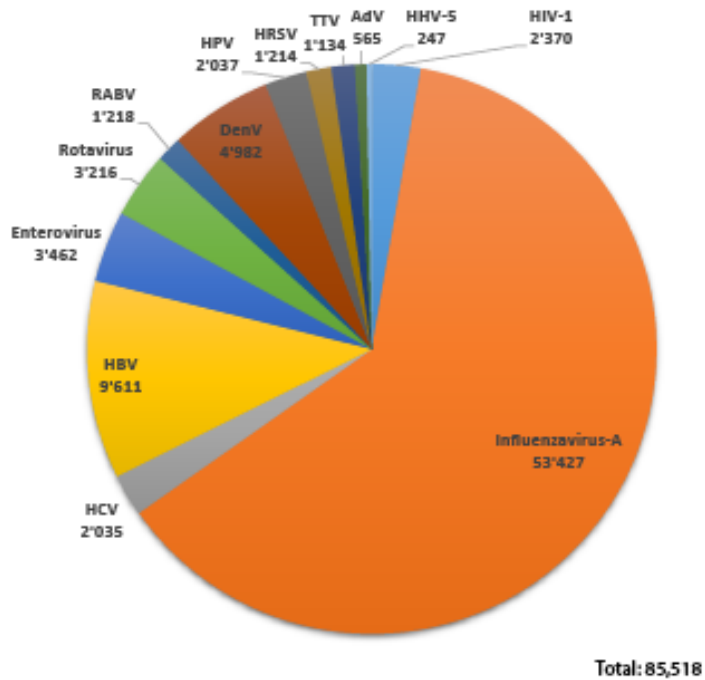
- **PASC (PAirwise Sequence Comparison)** is a web tool for analysis of pairwise identity distribution within viral families.



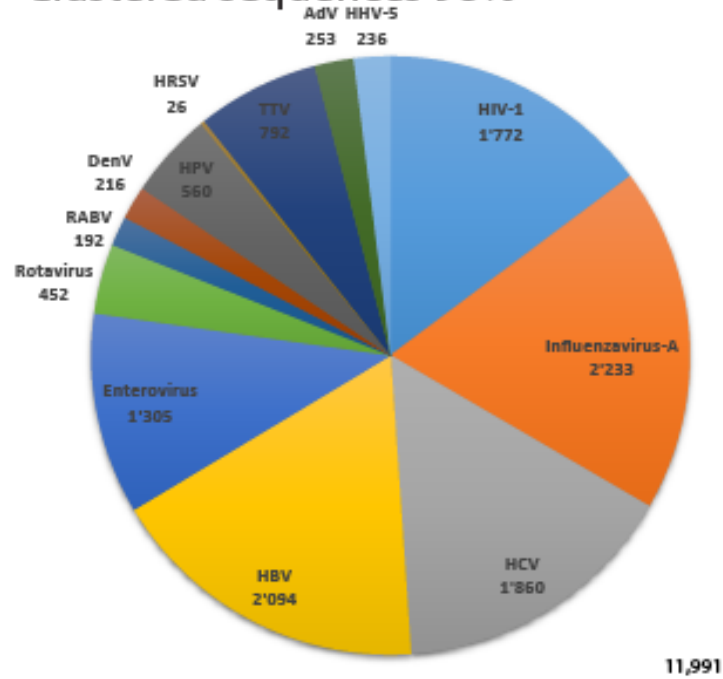


# Developing a metric for the diversity of variants?

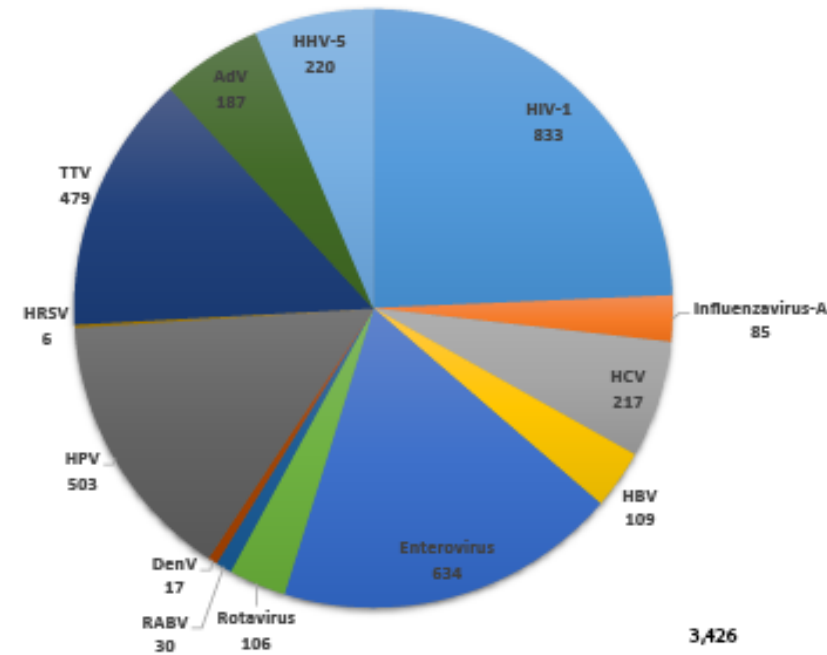
Complete sequences



Clustered sequences 98%

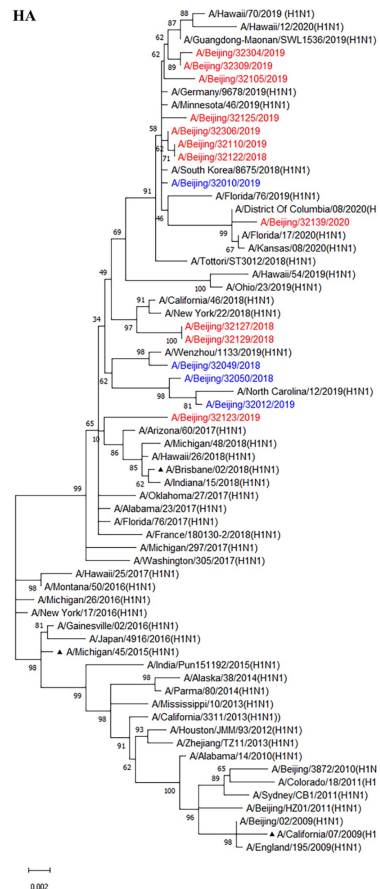


Clustered sequences 90%

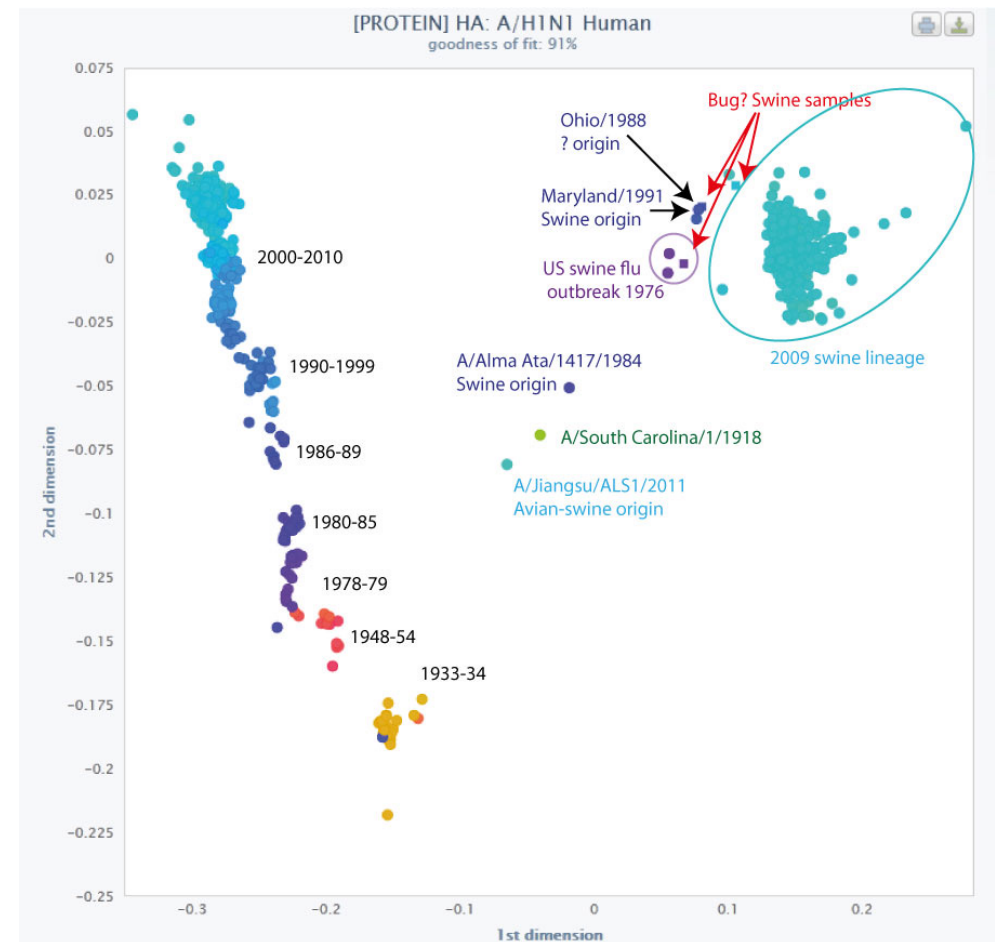


# Genome landscape representation

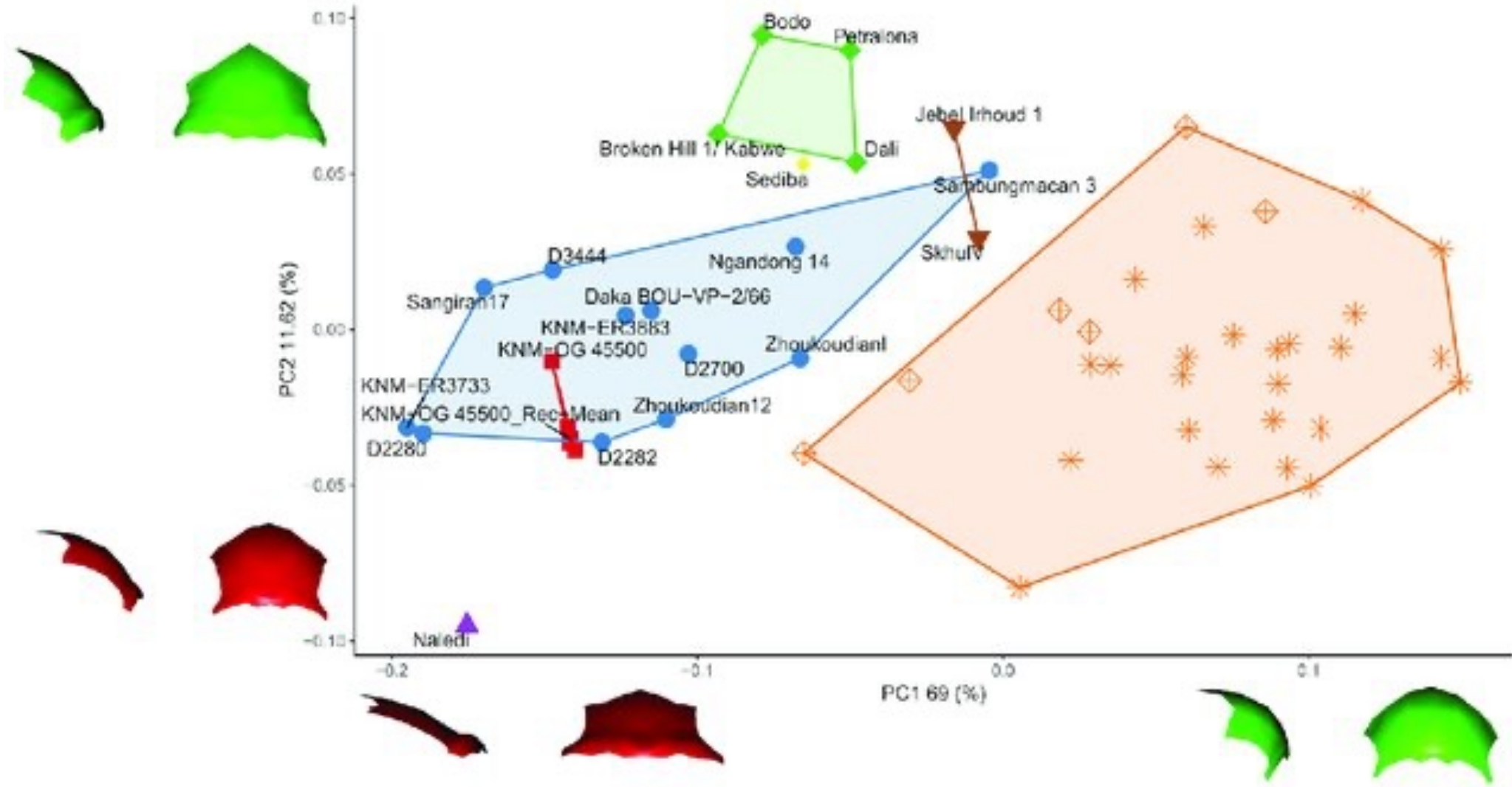
- Phylogenetic trees



- PCA analysis



# PCA plot of fossil specimens and recent modern humans

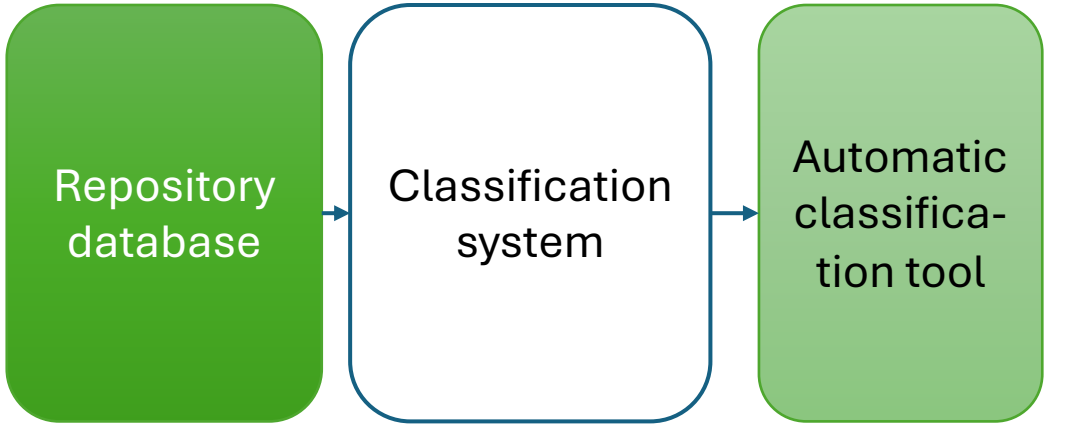


# Conclusion

---

# Monitoring below-species sequence: users needs

## Classification



INSDC, GISAID,...

- Influenza
- HIV
- Ebola
- SARS
- Rotavirus
- Monkeypox
- ...

PANGOLIN, Nextclade,...

Genotyping: HIV, HTLV, HBV, HCV, CMV, HPV, Dengue,...



## Pathology

- Drug resistance
- Tools for HIV, Influenza, ...
- Experimental
- Antigenicity: Vaccine, antibody therapy escape
- Experimental
- 3D modeling glycosylation
- Antigenicity Predictions
- tropism
- Experimental
- 3D modeling
- Receptor/tropism Predictions

## Analysis

- Epidemiology
- Nextstrain, Usher, Beast, BV-BRC, ...
- Epidemic models
- Functional analysis
- Experimental
- predictions
- Knowledge
- Pubmed, UniProt, ViralZone, BV-BRC...



# Thanks to



- Arnaud Kerhornou,
- Edouard de Castro,
- Cristina Casals-Casas,
- Patrick Masson,
- Chantal Hulo,
- Sylvain Poux,
- Ivo Pedruzzi,
- Nicole Redaschi,
- Alan Bridge

## Collaborations:



Samuel Cordey,

Pauline Vetter,

Laurent Kaiser



Richard Neher



Emma Hodcroft



Anna Maria Niewiadomska

Christian Zmasek

Richard H. Scheuermann

## Los Angeles County Health Laboratories

Alejandro R. Gener

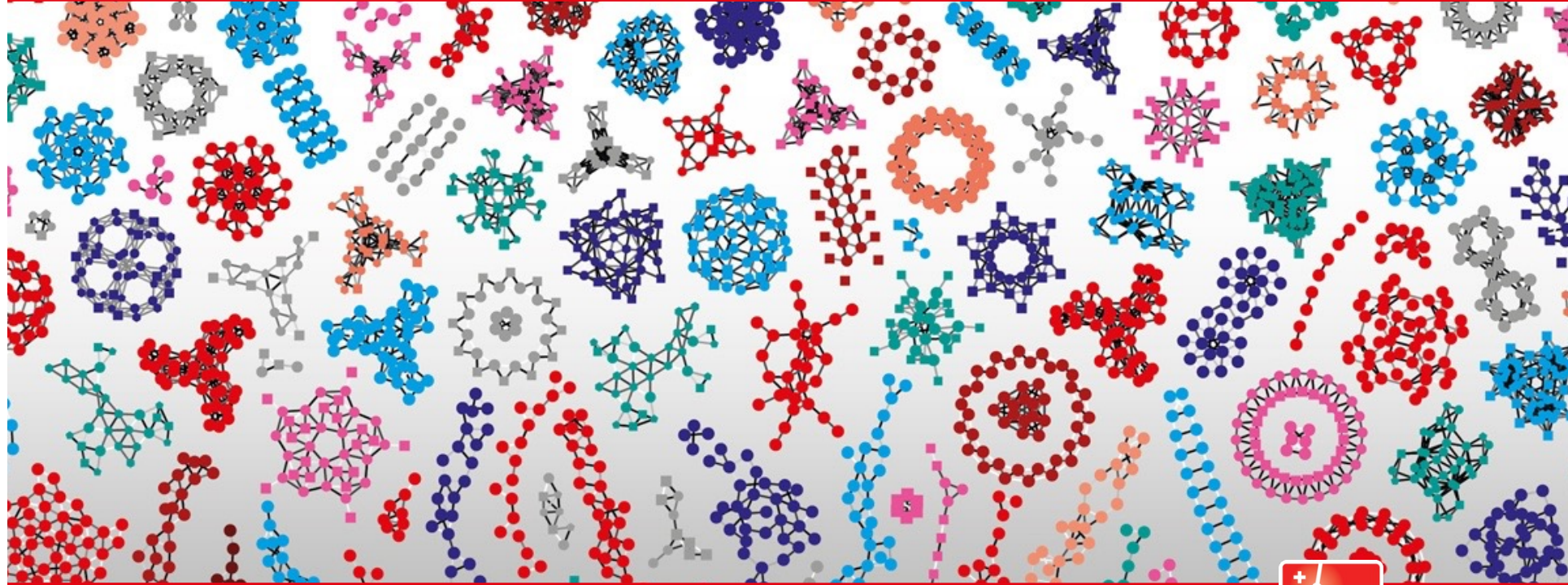


Bernard Moss

Ryan Connor



Rich Condit



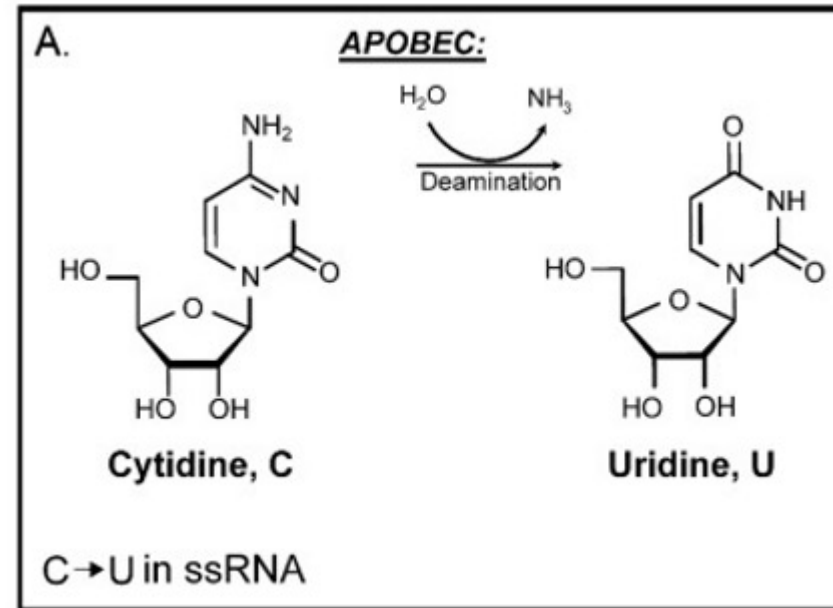
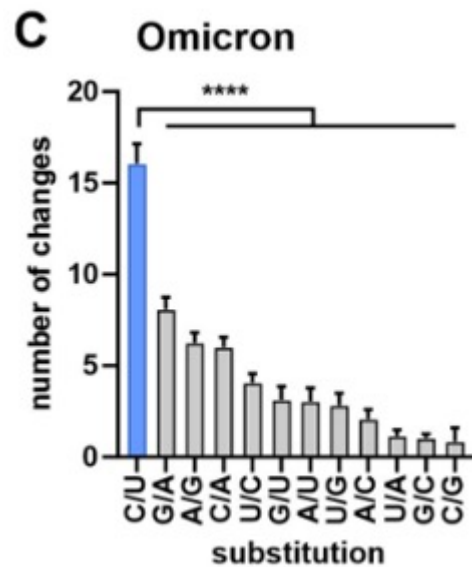
Swiss Institute of  
Bioinformatics

# Thank you

# The dynamic of SARS-CoV-2 mutations

Mutations occur due to:

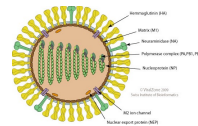
- Replication errors of the viral polymerase
- Cellular editing enzymes in the cytoplasm (immune cells)





# Isolate nomenclature

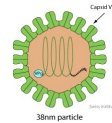
Few isolates nomenclature exist today, unfortunately there is no standart  
= no way to manage virus data with common bioinformmatics tools



## Influenza (1953):

Antigenic type/ **Isolation host**/**Location of sampling**/Isolate ID/**Year of sampling**

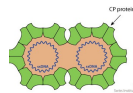
Ex: Influenza A/Puerto Rico/8/1934 H1N1



## Caliciviridae (2000):

**Isolation host**/Genus/**Species**/Isolate ID/**Year of sampling**/**Location of sampling**

Ex: Hu/NLV/I/Norwalk/1968/US



## Geminiviridae (2005):

**Species** [**Location of sampling**: Isolate ID: **Isolation host**: **Year of sampling**]

Ex: Maize streak virus, A [Uganda:Kas42:2005]