

# Virus genomic data analysis with computational biology

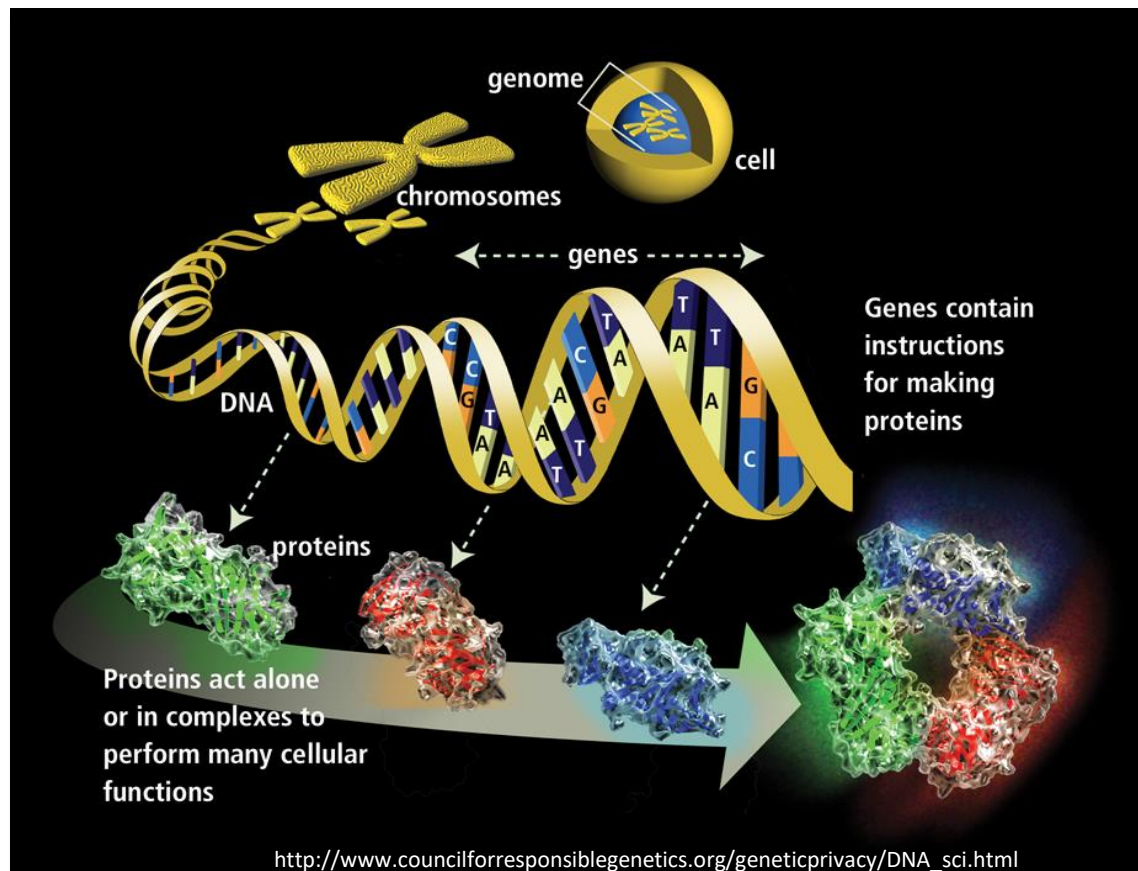
# 실험 목표

---

1. 지놈 데이터와 생물정보학의 중요성에 대한 이해
2. 지놈 데이터를 얻는 방법 익히기
3. 두 지놈 데이터를 비교하는 법 (sequence alignment) 익히기
4. 두 지놈 데이터의 비교 분석 및 분석 결과의 해석

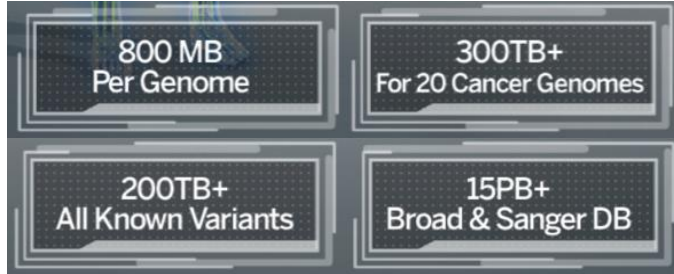
# Genome & genomic data

- ❖ All organisms have their biological information stored as genes
- ❖ The complete gene set of an organism is called genome
- ❖ Genomic data usually refers to DNA (genome) sequence data of an organism

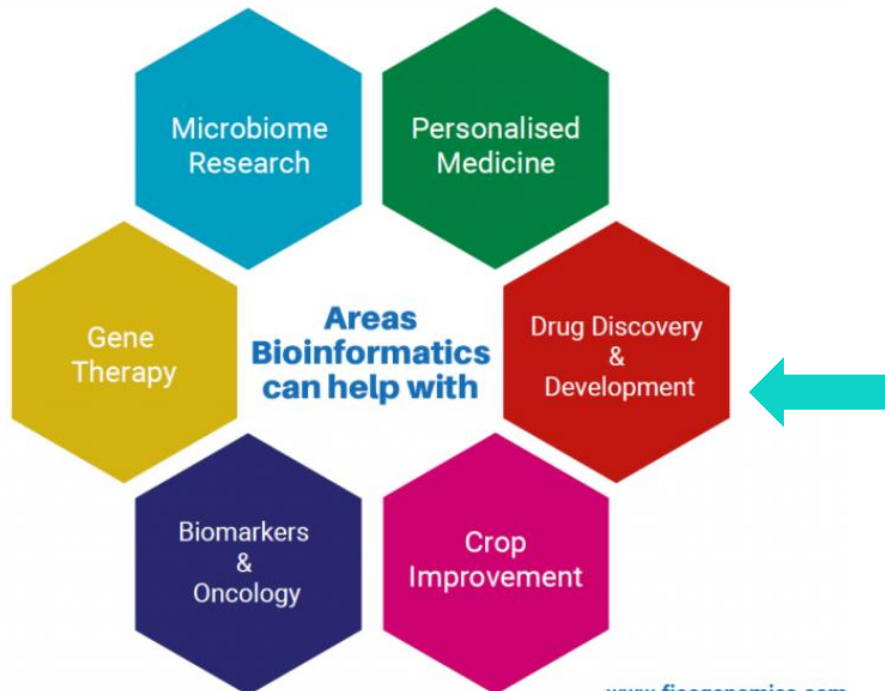


# Bioinformatics enable analysis of big data by using computers

Exponential growth of bio data



Manual



Bioinformatics



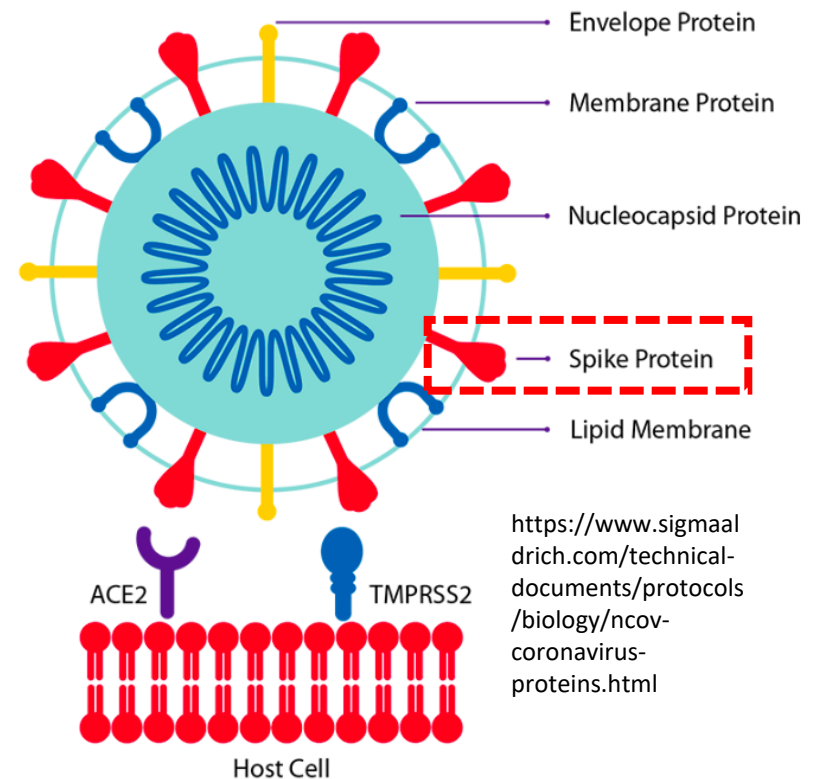
## Example: virus genomics study

---

- ❖ Identify natural reservoir of the virus and how zoonotic transmission happened
  - ❖ Prediction of transmission mechanism
  - ❖ Identification of most favorable environment for transmission
  - ❖ Prevention of initial outbreak
- ❖ Annotation of key protein needed for host infection and possible toxins
  - ❖ Antiviral treatment and invention of vaccines
  - ❖ FDA approved Ebola virus ( ~90% fatality) vaccine Ervebo in 2019
- ❖ Prediction of fatality, and affected areas (lung, skin, etc) by virus

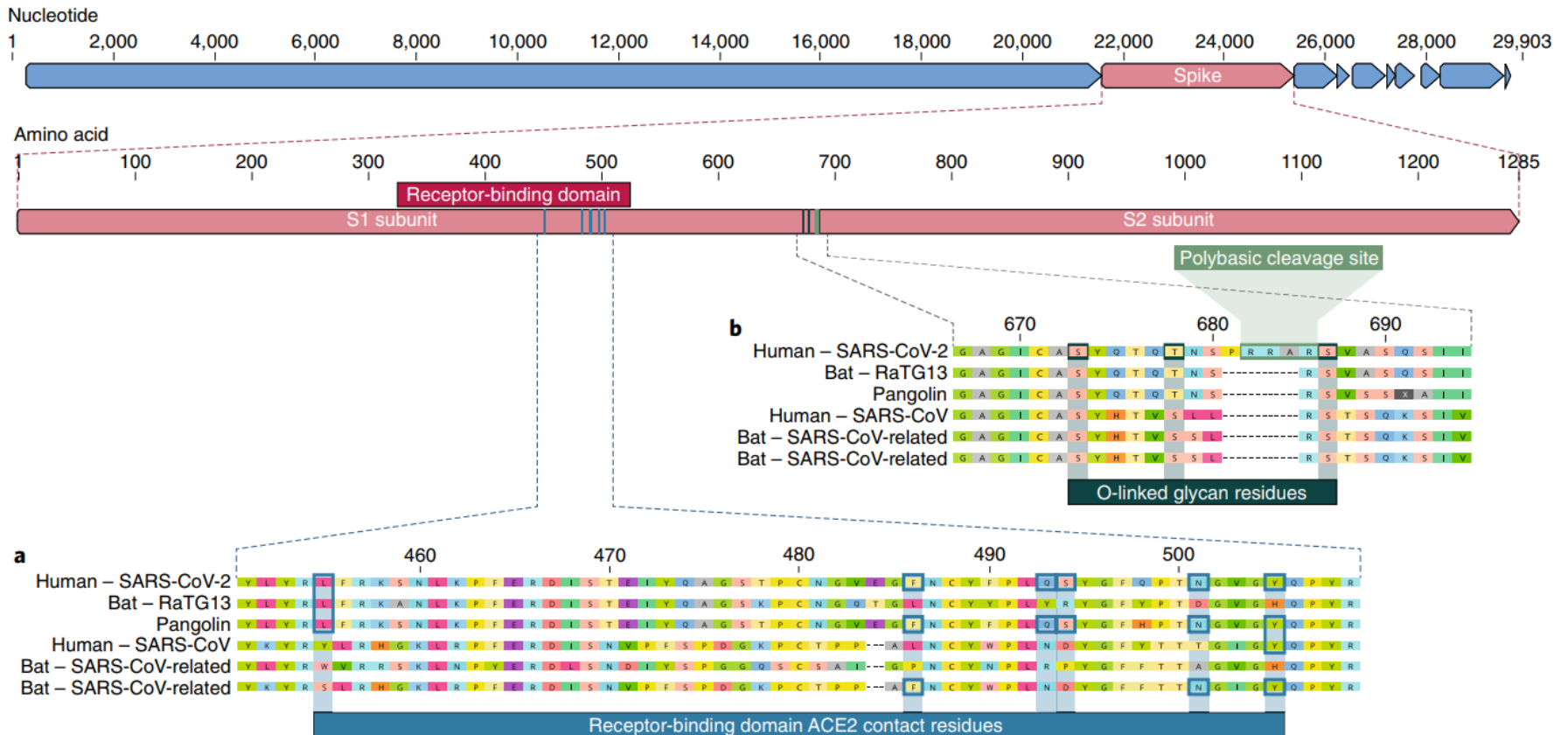
# SARS-CoV-2 (Severe acute respiratory syndrome coronavirus 2)

- ❖ SARS-CoV-2 is the cause of COVID-19 (Coronavirus disease 19)
- ❖ SARS-CoV-2 is the 7<sup>th</sup> coronavirus known to infect humans along:
  - ❖ SARS-CoV (2003)
  - ❖ MERS-CoV (Middle east respiratory syndrome coronavirus, 2012)
- ❖ SARS-CoV-2 has a spike glycoprotein on their surface that binds to ACE2 protein on the surface of host (animal) cells that gives them entry into cytoplasm
- ❖ Amino acid sequence of the spike protein determine the range of host



# Example of virus genome sequence analysis

## ❖ Comparison of amino acid sequences of several coronavirus



Andersen K G et al., Nature Medicine 2020

# How to compare “similar but different” genomes: match

---

ACGTACTA

Reference

ACGTACTA

Test



A C G T A C T A

| | | | | | | |

A C G T A C T A

perfect match



# How to compare “similar but different” genomes: mismatch

---

ACGTACTA

Reference

ACTTACTA

Test



A	C	G	T	A	C	T	A
A	C	T	T	A	C	T	A
		mismatch					

# How to compare “similar but different” genomes: gap (indel)

---

ACGTACTA

Reference

ACTACTA

Test



A C G T A C T A  
| | | | | | |  
A C - T A C T A  
gap

# How to compare “similar but different” genomes: scoring

ACGTACTA

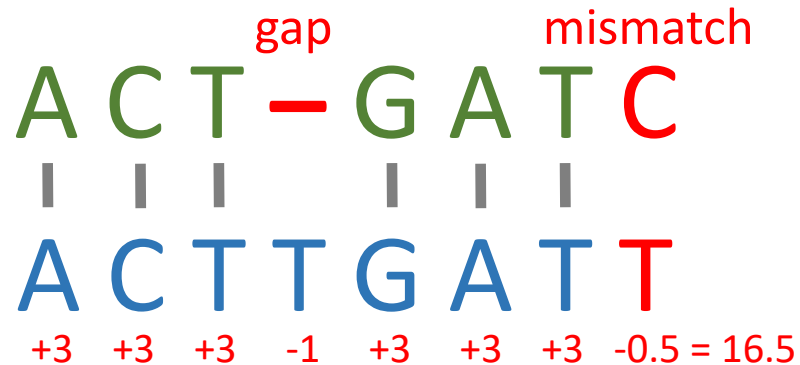
Reference

ACTACTA

Test



- Match = +3.0
- Mismatch = -0.5
- Gap = -1.0



- ❖ The **alignment score** is used to estimate how ‘similar’ two/multiple genomes are
- ❖ But genome sequence is usually too long to handle manually

# How to compare “similar but different” genomes: clustering

## Example: Hierarchical Agglomerative Clustering

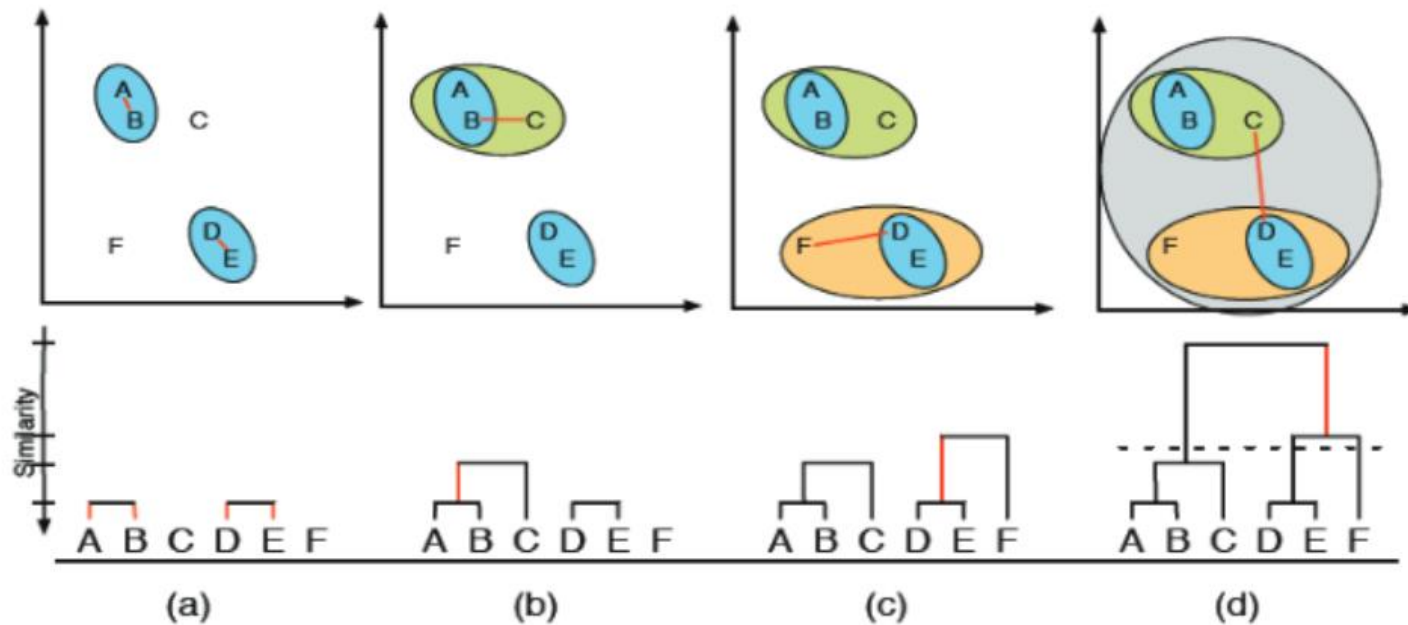
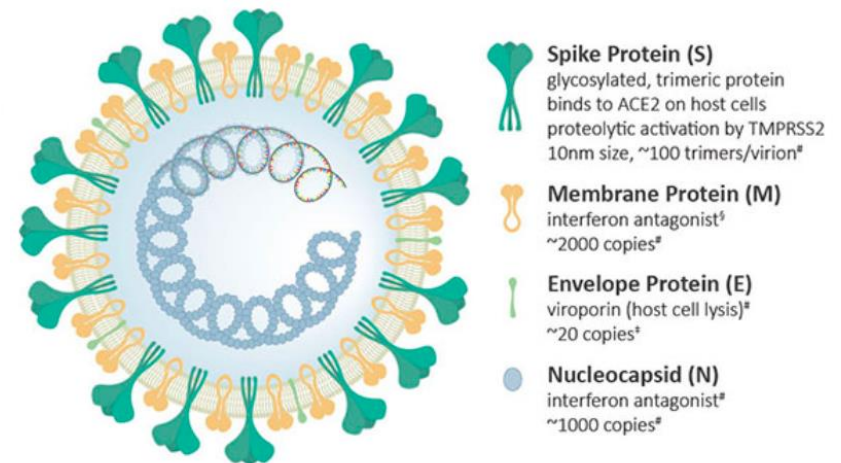
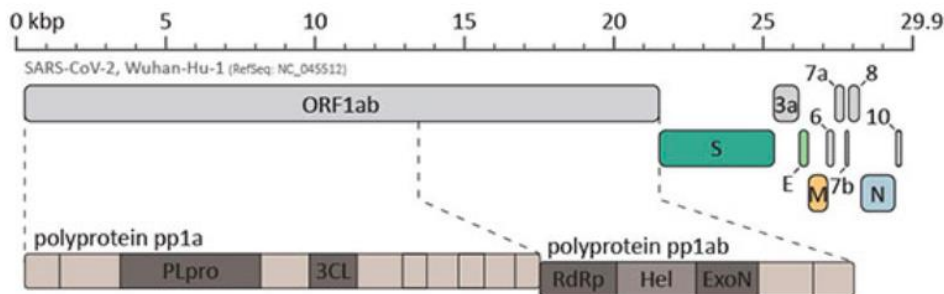


Figure 4: Example of hierarchical clustering: clusters are consecutively merged with the most nearby clusters. The length of the vertical dendrogram-lines reflect the nearness.

Janssen P et al., 2012;

# Analysis of SARS-Cov2 genome and protein sequence data

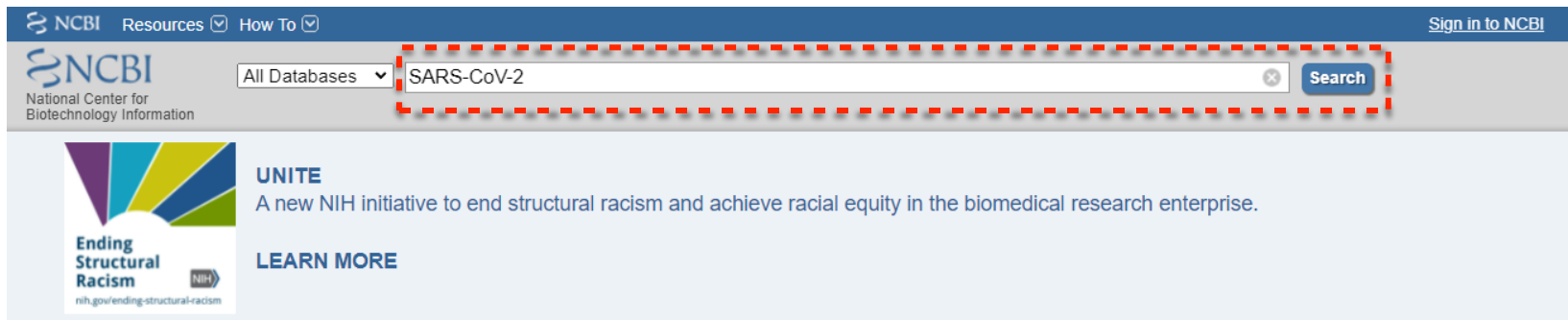
- ❖ Comparing various SARS-Cov2 genome will show how the virus spread through the world.
- ❖ Currently, most institutes diagnose COVID-19 by real-time PCR (RT-PCR), which amplifies specific regions of virus genome sequence
  - ❖ WHO recommends usage of **envelope gene** and RdRp region for initial screening



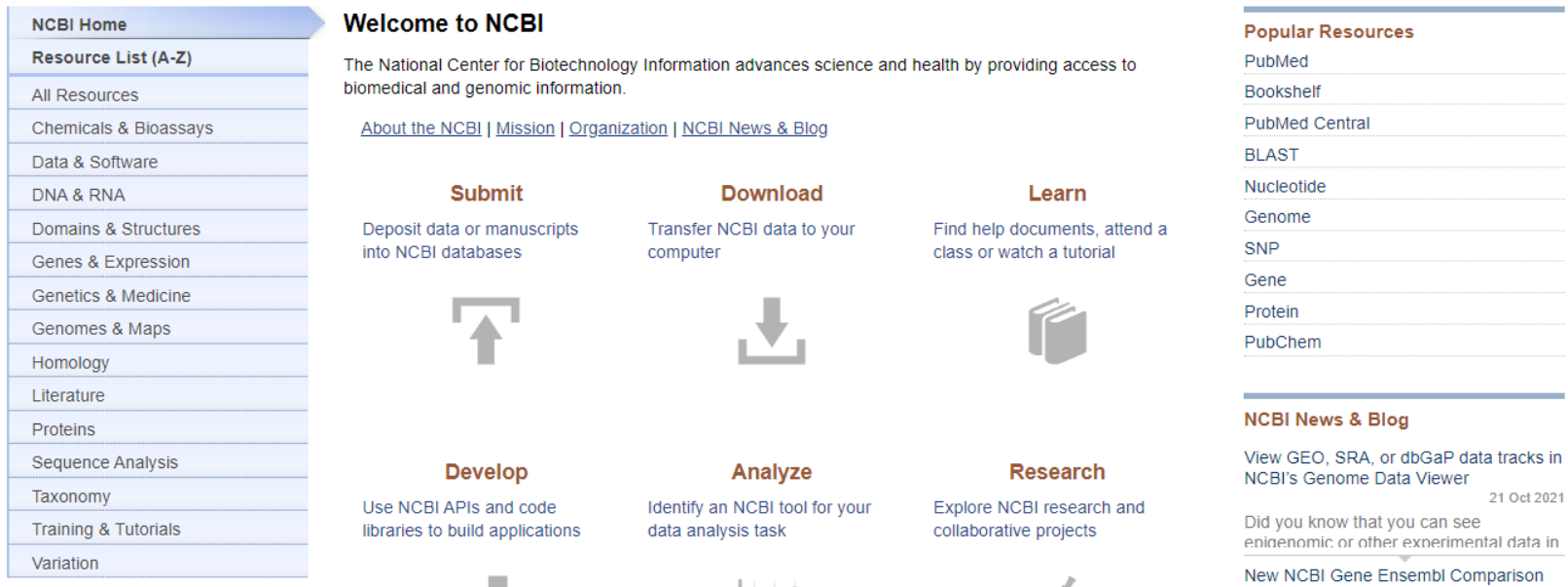
<https://www.enzolifesciences.com/science-center/technotes/2020/july/how-do-coronavirus-disease-covid-19-tests-work?/>

# Get access to SARS-CoV-2 sequence data

- ❖ Open <https://www.ncbi.nlm.nih.gov/> with Chrome
- ❖ Search SARS-CoV-2




The screenshot shows the top navigation bar of the NCBI website. The search bar is highlighted with a red dashed box and contains the text "SARS-CoV-2". The search button is labeled "Search".



The screenshot shows the main content area of the NCBI homepage. On the left is a "Resource List (A-Z)" menu. The main content area is titled "Welcome to NCBI" and includes a description of the center's mission, a list of links (About the NCBI, Mission, Organization, NCBI News & Blog), and three main action buttons: "Submit" (Deposit data or manuscripts into NCBI databases), "Download" (Transfer NCBI data to your computer), and "Learn" (Find help documents, attend a class or watch a tutorial). Below these are three more buttons: "Develop" (Use NCBI APIs and code libraries to build applications), "Analyze" (Identify an NCBI tool for your data analysis task), and "Research" (Explore NCBI research and collaborative projects). On the right is a "Popular Resources" list (PubMed, Bookshelf, PubMed Central, BLAST, Nucleotide, Genome, SNP, Gene, Protein, PubChem) and an "NCBI News & Blog" section with a link to "View GEO, SRA, or dbGaP data tracks in NCBI's Genome Data Viewer" dated 21 Oct 2021.

# Get access to SARS-CoV-2 sequence data


**National Library of Medicine**  
 National Center for Biotechnology Information
 Log in

Search NCBI  x Search

Results found in 30 databases

TAXONOMY Was this helpful?


*Severe acute respiratory syndrome coronavirus 2*

*Severe acute respiratory syndrome coronavirus 2* is a below species classification of *Severe acute respiratory syndrome-related coronavirus*

Taxonomy ID: 2697049

[NCBI SARS-CoV-2 resources](#)


**NCBI Virus**  
 Browse and download


**BLAST**  
 Use our new Betacoronavirus database for SARS-CoV-2 genome sequence analysis

Genomes	
Assembly	116
BioCollections	0
BioProject	1,562
BioSample	5,671,881
Genome	1
Nucleotide	4,684,923
SRA	4,347,788
Taxonomy	1

Clinical	
ClinicalTrials.gov	8,759
ClinVar	4
dbGaP	0
dbSNP	0
dbVar	12,788
GTR	85
MedGen	98
OMIM	16



## SARS-CoV-2 protein structures

View 3D structures and conserved domains of novel coronavirus proteins, including (S)pike, (E)nvelope, (M)embrane, and (N)ucleocapsid

# Get access to SARS-CoV-2 sequence data

[Viruses](#) / [Orthornavirae](#) / [Pisuviricota](#) / [Pisoniviricetes](#) / [Nidovirales](#) / [Coronaviridae](#) / [Betacoronavirus](#) / Severe acute respiratory syndrome-related coronavirus

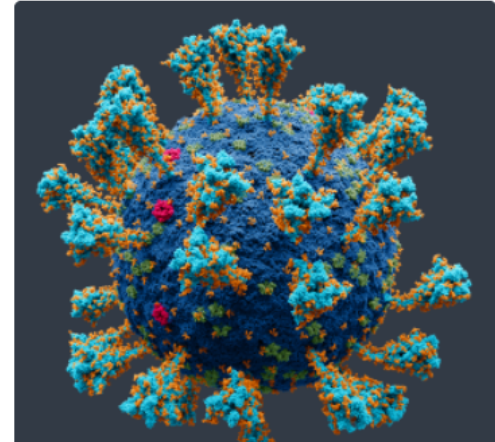
BETA

## Severe acute respiratory syndrome coronavirus 2 ☆

*Severe acute respiratory syndrome coronavirus 2* is a below-species classification of *Severe acute respiratory syndrome-related coronavirus*

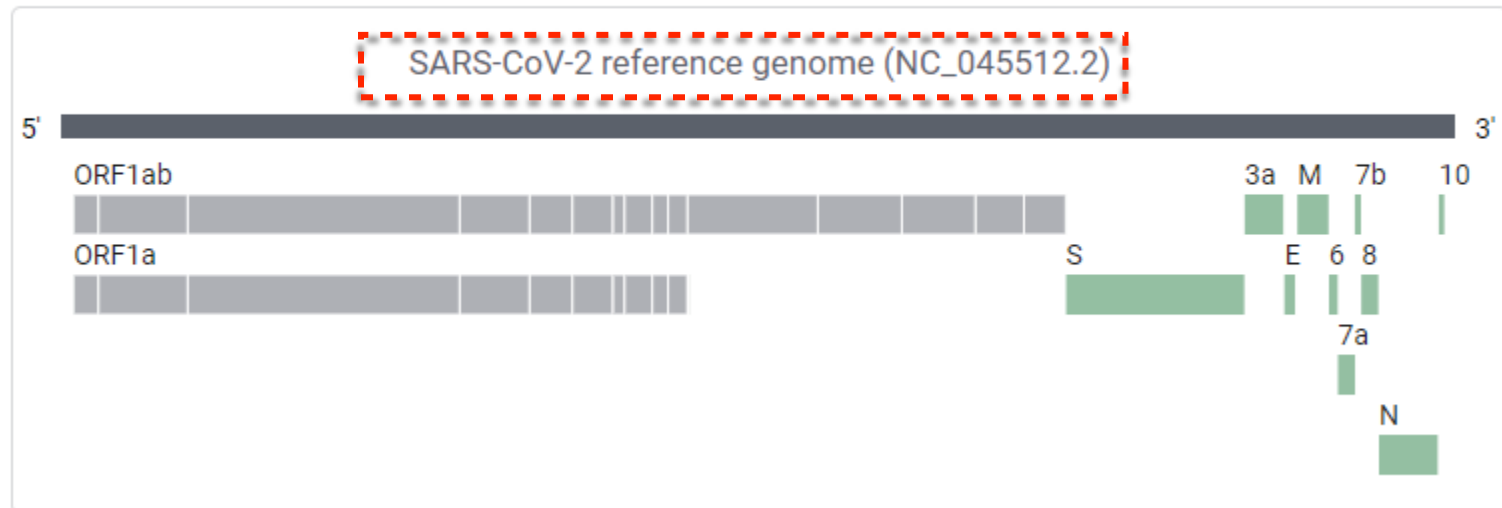
[Browse taxonomy](#)

Current scientific name	<i>Severe acute respiratory syndrome coronavirus 2</i>
Acronym	SARS-CoV-2
Genome type	ssRNA(+)
NCBI Taxonomy ID	2697049



Annotation from NCBI RefSeq

↓ scroll down





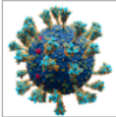
# Get access to SARS-CoV-2 sequence data

NIH National Library of Medicine  
National Center for Biotechnology Information

Search NCBI  Search


Results found in 30 databases

TAXONOMY Was this helpful?



**Severe acute respiratory syndrome coronavirus 2**  
*Severe acute respiratory syndrome coronavirus 2* is a below-species classification of *Severe acute respiratory syndrome-related coronavirus*  
Taxonomy ID: 2697049  
[NCBI SARS-CoV-2 resources](#)

NCBI Virus  
Browse and download



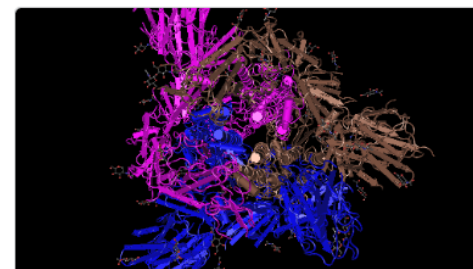
**BLAST**  
Use our new Betacoronavirus database for SARS-CoV-2 genome sequence analysis

scroll down



Genomes	
Assembly	116
BioCollections	0
BioProject	1,562
BioSample	5,671,881
Genome	1
Nucleotide	4,684,923
SRA	4,347,788
Taxonomy	1

Clinical	
ClinicalTrials.gov	8,759
ClinVar	4
dbGaP	0
dbSNP	0
dbVar	12,788
GTR	85
MedGen	98
OMIM	16



## SARS-CoV-2 protein structures

View 3D structures and conserved domains of novel coronavirus proteins, including (S)pike, (E)nvelope, (M)embrane, and (N)ucleocapsid

# Get access to SARS-CoV-2 sequence data

Organism Overview ; [Genome Assembly and Annotation report \[92\]](#)

ID: 86693

## Severe acute respiratory syndrome coronavirus 2

Severe acute respiratory syndrome coronavirus 2 Genome sequencing

Lineage: [Viruses\[24936\]](#); [Riboviria\[5409\]](#); [Orthornavirae\[4146\]](#); [Pisuviricota\[1086\]](#); [Pisoniviricetes\[651\]](#); [Nidovirales\[136\]](#); [Cornidovirineae\[72\]](#); [Coronaviridae\[72\]](#); [Orthocoronavirinae\[69\]](#); [Betacoronavirus\[24\]](#); [Sarbecovirus\[9\]](#); [Severe acute respiratory syndrome-related coronavirus\[8\]](#); [Severe acute respiratory syndrome coronavirus 2\[1\]](#)

SARS-CoV-2 sequencing from San Diego county

### Summary

**Sequence data:** genome assemblies: 92; sequence reads: 41

**Statistics:** median total length (Mb): 0.029882  
median protein count: 10  
median GC%: 38

### Representative (genome information for reference and representative genomes)

Reference genome:

- [Severe acute respiratory syndrome coronavirus 2 ASM985889v3](#)

Type	Name	RefSeq	INSDC	Size (Kb)	GC%	Protein	Gene
Chr	-	NC_045512.2	MN908947.3	29.9	38.0	12	11

# Get access to SARS-CoV-2 sequence data

## Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome


NCBI Reference Sequence: NC\_045512.2

[FASTA](#) [Graphics](#)

Go to:



LOCUS NC\_045512 29903 bp ss-RNA linear VRL 18-JUL-2020  
DEFINITION Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome.  
ACCESSION NC\_045512  
VERSION NC\_045512.2  
DBLINK BioProject: [PRJNA485481](#)  
KEYWORDS RefSeq.  
SOURCE Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)  
ORGANISM [Severe acute respiratory syndrome coronavirus 2](#)  
Viruses; Riboviria; Orthornavirae; Pisuviricota; Pisoniviricetes; Nidovirales; Coronidovirineae; Coronaviridae; Orthocoronavirinae; Betacoronavirus; Sarbecovirus.  
REFERENCE 1 (bases 1 to 29903)  
AUTHORS Wu, F., Zhao, S., Yu, B., Chen, Y. M., Wang, W., Song, Z. G., Hu, Y., Tao, Z. W., Tian, J. H., Pei, Y. Y., Yuan, M. L., Zhang, Y. L., Dai, F. H., Liu, Y., Wang, Q. M., Zheng, J. J., Xu, L., Holmes, E. C. and Zhang, Y. Z.  
TITLE A new coronavirus associated with human respiratory disease in China  
JOURNAL Nature 579 (7798), 265-269 (2020)  
PUBMED [32015508](#)  
REMARK Erratum:[Nature. 2020 Apr;580(7803):E7. PMID: 32296181]  
REFERENCE 2 (bases 13476 to 13503)  
AUTHORS Baranov, P. V., Henderson, C. M., Anderson, C. B., Gesteland, R. F., Atkins, J. F. and Howard, M. T.  
TITLE Programmed ribosomal frameshifting in decoding the SARS-CoV genome  
JOURNAL Virology 332 (2), 498-510 (2005)  
PUBMED [15680415](#)

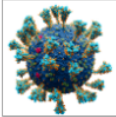
# Get access to all available SARS-CoV-2 sequence data

**National Library of Medicine**  
National Center for Biotechnology Information Log in


Search NCBI  × Search

Results found in 30 databases


TAXONOMY Was this helpful?  



***Severe acute respiratory syndrome coronavirus 2***  
*Severe acute respiratory syndrome coronavirus 2* is a below-species classification of *Severe acute respiratory syndrome-related coronavirus*  
Taxonomy ID: 2697049  
[NCBI SARS-CoV-2 resources](#)



**NCBI Virus**  
Browse and download



**BLAST**  
Use our new Betacoronavirus database for SARS-CoV-2 genome sequence analysis

# Get access to all available SARS-CoV-2 sequence data

[About Us](#) [Find Data](#) [Help](#) [How to Participate](#) [Submit Sequences](#) [Contact Us](#)

## SARS-CoV-2 Data Hub

Download ▾

Quick Links

[Betacoronavirus BLAST](#)    [SARS-CoV-2 Articles in PubMed](#)    [NCBI SARS-CoV-2 Resources](#)  
[CDC Outbreak Information](#)    [SRA Data](#)    [Datasets command line](#)

[Results Table](#)    [Dashboard Visualizations](#)    [Variants Overview](#) New!    Selected Results: 0    [Align](#)    [Build Phylogenetic Tree](#)

**📢 New! Submitters' Information now available**

We appreciate the effort from all involved in collecting samples and making sequence data publicly available. In order to facilitate citations and acknowledgements, we are adding information describing submitters, including affiliated organizations or institutions provided during submission.

Although table columns may display abbreviated information, complete entries can be viewed by hovering over entries and are available by downloading the Result Table (metadata table). We are continuing to refine the availability of this information and feedback is appreciated.

Refine Results Reset

Virus +

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), taxid:2697049 ×

Accession +

Sequence Length +

Ambiguous Characters +

Nucleotide (6,875,578)    Protein (40,401,595)    RefSeq Genome (1)    [Select Columns](#)

<input type="checkbox"/>	Accession	Organism Name <span>New!</span>	Submitters	Organization	Release Date	Pangolin
<input type="checkbox"/>	<a href="#">NC_045512</a> <span>RefSeq</span>	Severe acute respiratory s...	Wu,F., et al.	National Center for Biotec...	2020-01-13	B
<input type="checkbox"/>	<a href="#">QQ867531</a>	Severe acute respiratory s...	Howard,D., et al.	Centers for Disease Contr...	2023-04-25	XBB.1.5.35
<input type="checkbox"/>	<a href="#">QQ867532</a>	Severe acute respiratory s...	Howard,D., et al.	Centers for Disease Contr...	2023-04-25	XBB.1.5
<input type="checkbox"/>	<a href="#">QQ867533</a>	Severe acute respiratory s...	Howard,D., et al.	Centers for Disease Contr...	2023-04-25	XBB.1.5
<input type="checkbox"/>	<a href="#">QQ867534</a>	Severe acute respiratory s...	Howard,D., et al.	Centers for Disease Contr...	2023-04-25	XBB.1.5
<input type="checkbox"/>	<a href="#">QQ867535</a>	Severe acute respiratory s...	Howard,D., et al.	Centers for Disease Contr...	2023-04-25	XBB.1.5

Expand Table Feedback

# Select filters (1)

Accession	+
Sequence Length	+
Ambiguous Characters	+
Sequence Type	+
RefSeq Genome Completeness	+
Nucleotide Completeness	-
<input checked="" type="checkbox"/> complete (1,772,235) (1)	
<input type="checkbox"/> partial (5,103,343)	

Provirus	+
Geographic Region	-
Search All Geo Locations	
How to filter by the U.S. states? (2)	
<input type="checkbox"/> Africa (5,990) >	
Ethiopia x	
<input type="checkbox"/> Asia (20,026) >	
Iran x	
<input type="checkbox"/> Europe (433,445) >	
Denmark x	

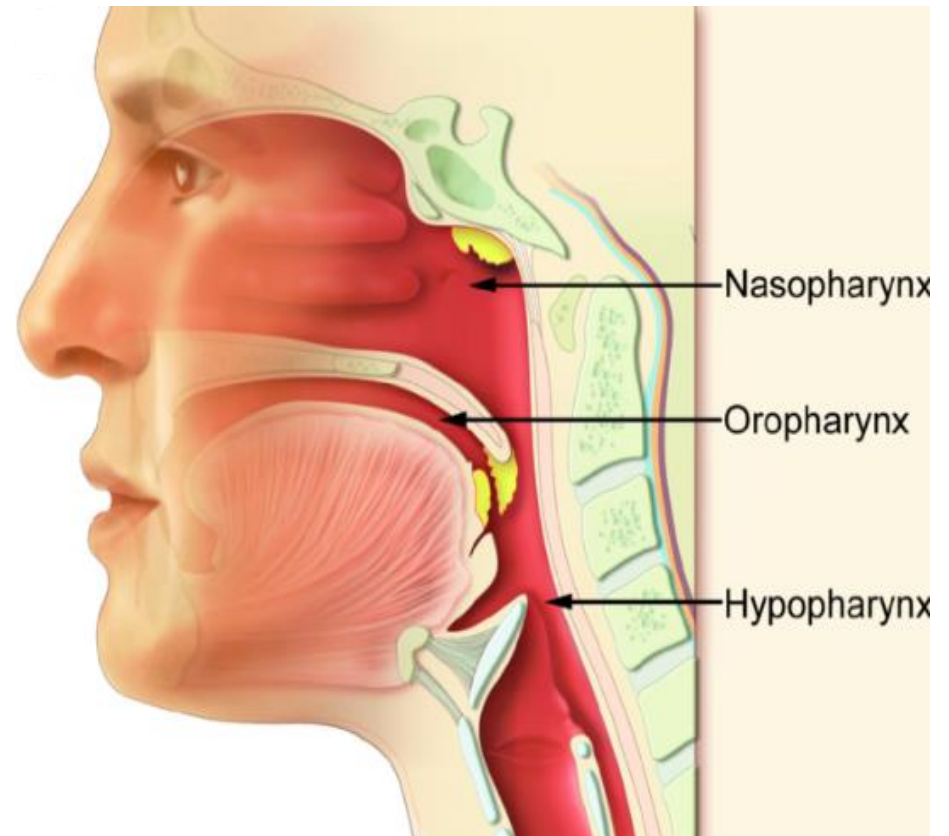
Geographic Region	-
<input type="checkbox"/> Europe (433,445) >	
Denmark x	
<input type="checkbox"/> North America (1,289,171) >	
Canada x	
<input type="checkbox"/> Oceania (12,041) >	
Palau x	
<input type="checkbox"/> South America (11,473) >	
Ecuador x	

## Select filters (2)

❖ Select 'oronasopharynx' tab of Isolation Source.

(3)

Host	+
Submitters	+
Isolation Source	-
<input checked="" type="checkbox"/> oronasopharynx (69)	



# Perform multiple alignment to compare sequences

The screenshot displays the NCBI Virus SARS-CoV-2 Data Hub interface. At the top, the NCBI Virus logo and navigation links are visible. The main content area shows a 'SARS-CoV-2 Data Hub' section with a 'Download' button and 'Quick Links' for various data sources. Below this, a navigation bar includes 'Results Table', 'Dashboard Visualizations', and 'Variants Overview'. The 'Results Table' section shows 'Selected Results: 69' and two buttons: 'Align' and 'Build Phylogenetic Tree', both highlighted with red dashed boxes and labeled with '(2)'. A notification banner indicates 'New! Submitters' Information now available'. On the left, a 'Refine Results' sidebar shows the current filter: 'Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), taxid:2697049'. The main table is titled 'Nucleotide (69)' and has columns for 'Accession', 'Organism Name', 'Submitters', 'Organization', and 'Release Date'. A red dashed box labeled '(1)' highlights the 'Accession' column header. The table lists several sequences, each with a checkbox in the left margin. A vertical 'Expand Table' button is on the left side of the table. A 'Feedback' button is located at the bottom right.

NCBI Virus  
Sequences for discovery

About Us ▾ Find Data ▾ Help ▾ How to Participate ▾ Submit Sequences ▾ [Contact Us](#)

SARS-CoV-2 Data Hub  
Download ▾

Quick Links  
Betacoronavirus BLAST  
CDC Outbreak Information  
SARS-CoV-2 Articles in PubMed  
SRA Data  
NCBI SARS-CoV-2 Resources  
Datasets command line

Results Table Dashboard Visualizations Variants Overview **New!**

Selected Results: 69

**(2)** [Align](#) [Build Phylogenetic Tree](#)

**New!** Submitters' Information now available +

Refine Results [Reset](#)

Virus +  
Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), taxid:2697049 x

Accession +  
Sequence Length +  
Ambiguous Characters +  
Sequence Type +  
RefSeq Genome Completeness +

**(1)**

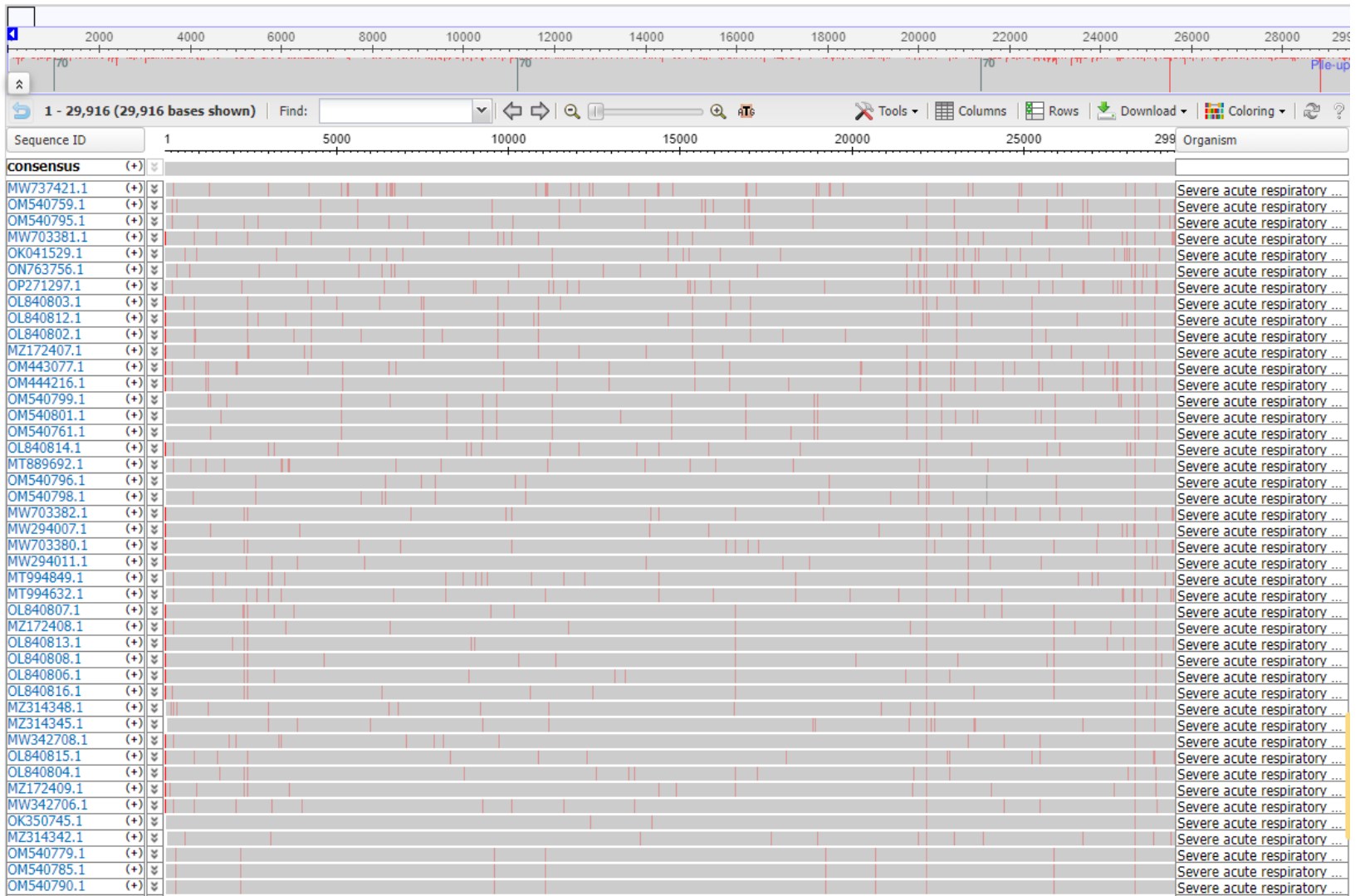
	Nucleotide (69)	Protein (828)	RefSeq Genome (0)	Select Columns		
Expand Table	<input checked="" type="checkbox"/> Accession	Organism Name <b>New!</b>	Submitters	Organization	Release Date	
<input checked="" type="checkbox"/>	<a href="#">OQ843561</a>	Severe acute respiratory s...	Binderup,A,...	University of Copenhagen...	2023-0	
<input checked="" type="checkbox"/>	<a href="#">OQ446561</a>	Severe acute respiratory s...	Sheth,M., et al.	Centers for Disease Contr...	2023-0	
<input checked="" type="checkbox"/>	<a href="#">OP722492</a>	Severe acute respiratory s...	Binderup,A,...	University of Copenhagen...	2022-3	
<input checked="" type="checkbox"/>	<a href="#">OP722493</a>	Severe acute respiratory s...	Binderup,A,...	University of Copenhagen...	2022-3	
<input checked="" type="checkbox"/>	<a href="#">OP271296</a>	Severe acute respiratory s...	Fernandez-...	University of Copenhagen...	2022-0	
<input checked="" type="checkbox"/>	<a href="#">OP271297</a>	Severe acute respiratory s...	Fernandez-...	University of Copenhagen...	2022-0	
<input checked="" type="checkbox"/>	<a href="#">ON809567</a>	Severe acute respiratory s...	Polacek,C.		2022-04	

Feedback



# Sequence comparison results

## Multiple Alignment



# Build phylogenetic tree

**SARS-CoV-2 Data Hub**  
Download ▾

Quick Links  
Betacoronavirus BLAST  
SARS-CoV-2 Articles in PubMed  
NCBI SARS-CoV-2 Resources  
CDC Outbreak Information  
SRA Data  
Datasets command line

Results Table | Dashboard Visualizations ⓘ | Variants Overview **New!**

Selected Results: 69

Align | **Build Phylogenetic Tree** (2)

**New! Submitters' Information now available** +

Refine Results Reset

Virus +  
Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), taxid:2697049 x

Accession +  
Sequence Length +  
Ambiguous Characters +  
Sequence Type +  
RefSeq Genome Completeness +

(1) **Nucleotide (69)** | Protein (828) | RefSeq Genome (0) Select Columns

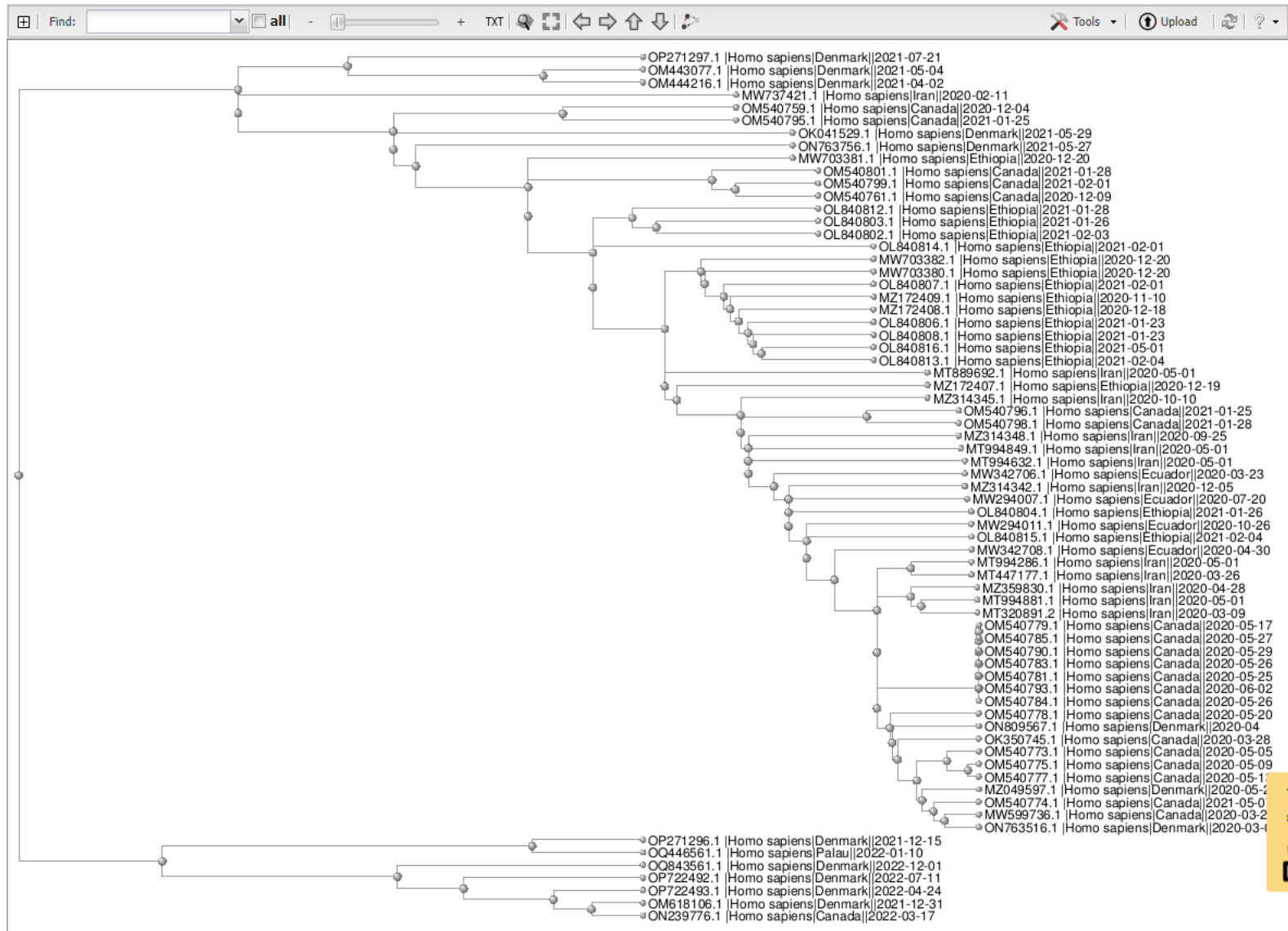
Expand Table <

<input checked="" type="checkbox"/>	Accession ▾	Organism Name <b>New!</b> ▾	Submitters ▾	Organization ▾	Release Date
<input checked="" type="checkbox"/>	<a href="#">OQ843561</a>	Severe acute respiratory s...	Binderup,A,...	University of Copenhagen...	2023-0...
<input checked="" type="checkbox"/>	<a href="#">OQ446561</a>	Severe acute respiratory s...	Sheth,M., et al.	Centers for Disease Contr...	2023-0...
<input checked="" type="checkbox"/>	<a href="#">OP722492</a>	Severe acute respiratory s...	Binderup,A,...	University of Copenhagen...	2022-0...
<input checked="" type="checkbox"/>	<a href="#">OP722493</a>	Severe acute respiratory s...	Binderup,A,...	University of Copenhagen...	2022-0...
<input checked="" type="checkbox"/>	<a href="#">OP271296</a>	Severe acute respiratory s...	Fernandez-...	University of Copenhagen...	2022-0...
<input checked="" type="checkbox"/>	<a href="#">OP271297</a>	Severe acute respiratory s...	Fernandez-...	University of Copenhagen...	2022-0...
<input checked="" type="checkbox"/>	<a href="#">ON809567</a>	Severe acute respiratory s...	Polacek,C.		2022-0...

Feedback

# Resulting phylogenetic tree from 6 countries

Phylogenetic Tree



# Get access to SARS-CoV-2 ORF3a protein sequence data

Proteins +  
ORF3a protein x

Provirus +

Geographic Region +  
China x South Korea x

(1)

SARS-CoV-2 Data Hub

Download v

Quick Links

- Betacoronavirus BLAST
- CDC Outbreak Information
- SARS-CoV-2 Articles in PubMed
- SRA Data
- NCBI SARS-CoV-2 Resources
- Datasets command line

Results Table Dashboard Visualizations Variants Overview New!

Selected Results: 190

(4)

Align Build Phylogenetic Tree

New! Submitters' Information now available +

(2)

Refine Results Reset

Virus +  
Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), taxid:2697049 x

Accession +

Sequence Length +

Ambiguous Characters +

(3)

Nucleotide (190) Protein (190) RefSeq Genome (0) Select Columns

Accession	Organism Name	Submitters	Release Date	Pangolin	Isolate
<input checked="" type="checkbox"/> <a href="#">WED41271</a>	Severe acute respiratory s...	Yang,W., et al.	2023-03-08	B.1.1.317	LA-504
<input checked="" type="checkbox"/> <a href="#">WED41283</a>	Severe acute respiratory s...	Yang,W., et al.	2023-03-08	B.1.1.317	LA-730
<input checked="" type="checkbox"/> <a href="#">WED41295</a>	Severe acute respiratory s...	Yang,W., et al.	2023-03-08	B.1.1.317	LA-779
<input checked="" type="checkbox"/> <a href="#">WED41472</a>	Severe acute respiratory s...	Yang,W., et al.	2023-03-08	BA.2.2.1	LA-234
<input checked="" type="checkbox"/> <a href="#">WED41484</a>	Severe acute respiratory s...	Yang,W., et al.	2023-03-08	BA.2.2.1	LA-234

Feedback

# Multiple alignment & comparison of 127 ORF3a proteins

## Multiple Alignment



# Assignment

---

❖ LMS 강의컨텐츠 해당 주차 “과제: 생물정보학 실험“에 첨부한 보고서 파일을 작성하여 제출합니다.

❖ 제출 기한:

❖ 월요일 반: 5/2 (화) 22:00 까지

❖ 수요일 반: 5/4 (목) 22:00 까지

(1) 6개 국가에서 3개씩 바이러스 서열을 선택하여 phylogenetic tree를 그리고, 그 결과를 캡처하여 넣으시오.  
(가장 서로 유사한 서열 pair와 가장 서로 다른 서열 pair가 무엇인지 서술하기)

(2) Multiple alignment와 scoring에 의해 어떻게 phylogenetic tree를 그리는지 설명하시오.

- Multiple alignment의 개념, scoring의 원리, tree로의 변환을 포함하여 설명