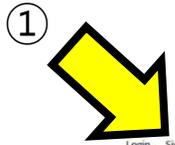


Sign up & Login



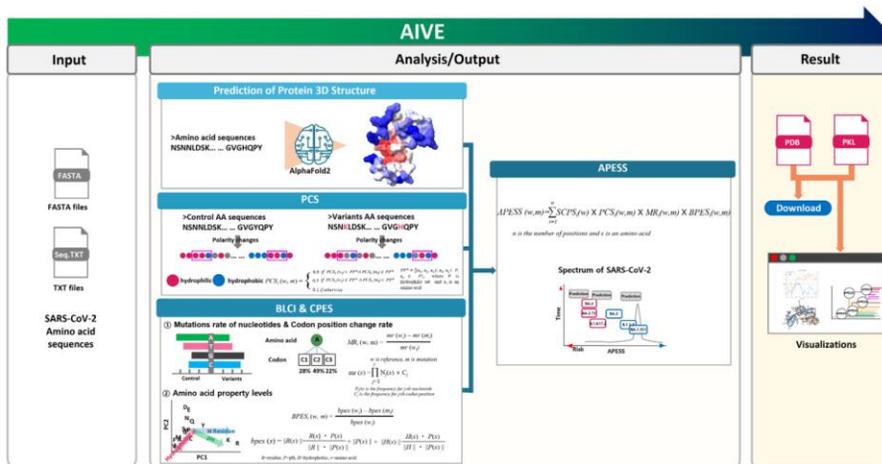
AIVE

Artificial Intelligence analytics toolkit for predicting Virus mutation in proEIn

Artificial Intelligence analytics toolkit for predicting Virus mutation in proEIn (AIVE) is a Web and GPU based analysis tool that can predict protein structures and properties from user-entered viral sequences. AIVE uses AlphaFold2 software (<https://github.com/deepmind/alphafold>) to evaluate protein structures. AIVE provides independently developed mathematical models (SCPS, PCS, MR, BPES, APES). It provides information on structural differences (SCPS, PCS), physical changes (MR), and biochemical changes (BPES) at the amino acid and nucleotide levels. These tools were calculated using various information such as amino acid chemical properties (pH, Hydrophobic, Residue), molecular structure prediction results (PAE, pLDDT), mutation and codon frequencies in viruses, and amino acid polarity features. AIVE serves optimized analysis and prediction for SARS-CoV-2 viral mutations. Analysis and prediction of other virus species will be updated later.

The information and analysis tools we provide are the following:

- A. Protein structure prediction from viral sequences using learning models Prediction of folding and docking from viral mutations
 - Comparison of folding and docking scores
- B. Polarity changes in protein sequences Measurement of repeated polarity changes
- C. Mathematical models based on amino acid and nucleotide levels (MR & BPES)
 - Scoring of rate of change for nucleotide levels
 - Scoring of rate of change for amino acid properties (Residue, Hydrophobic, and pH)
- D. Comprehensive mathematical analysis model (APES)
 - Integrating results for protein structure prediction, polarity change, and nucleotide and amino acid properties levels



AIVE

Catholic University of Korea, College of Medicine, 222 Banpo-dae-ro, Seocho-gu, Seoul 06591, Republic of Korea

AIVE is used by creating a personal account for the ease of comparing analysis results.
① Click on the "Sign Up" option to proceed with the account creation.

Sign up & Login

Register user

Name

Password

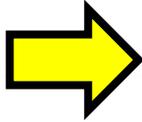
Confirm Password

E-mail

※ Email submission is optional and is used for the "Forgot password?" feature.

example@yourhost

①



Register

②



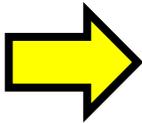
Login

Name

Password

[Forgot password?](#)

③



Login

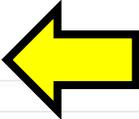
① Enter the username and password. To use the "Forgot password?" feature, enter your email and proceed with account creation.

②, ③ Access the "Login" section, enter the account information, and log in.

Demo



Virus

Project name **DEMO**  ①



Virus

Project name **DEMO** The following shows an example of how to run AIVE using the sample data (BA.5_RBM)

Target virus [?]

SARS-CoV-2 Omicron(BA.5)

※ From the V.O.C. list, variants can only be generated for SARS-CoV-2 if 'Wuhan-Hu-1' is selected.

Input virus Sequence

※ The length range for sequences that are covered is between 16 and 2700.

Alignment [?] Unable to select alignment when selecting Demo and VOCs

Amino acids with electrically charged side chain_negative

Amino acids with electrically charged side chain_positive

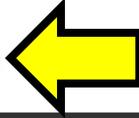
Amino acids with Polar uncharged side chain

Special cases

Amino acids with hydrophobic side chain

Keep private

Variants protein sequence

 ②

AIVE provides a demo feature that demonstrates the analysis of SARS-CoV-2 variants of concern (VOCs).
① Clicking on the Demo button will automatically select the RBM [S:437-508] region of the SARS-CoV-2 BA.5 variant.
② To predict and analyze the structure of the sequence generated by the Demo function, click the "Server Prediction" button to submit the task.

Predict VOC structure

Project name **DEMO**

tutorial

Target virus [?]

All Viruses

All Viruses

SARS-CoV-2

Target virus [?]

SARS-CoV-2

※ From the V.O.C. list, variants can only be generated for SARS-CoV-2 if 'Wuhan-Hu-1' is

Select VOC

Select VOC

Alpha(B.1.1.7)

Beta(B.1.351)

Delta(B.1.617.2)

Gamma(P.1)

Omicron(BA.1)

Omicron(BA.2)

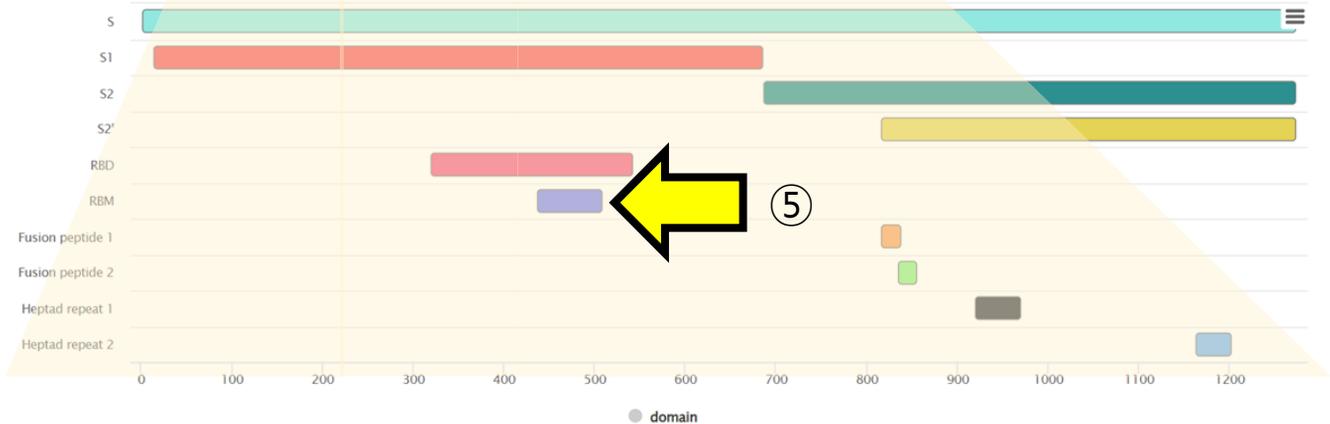
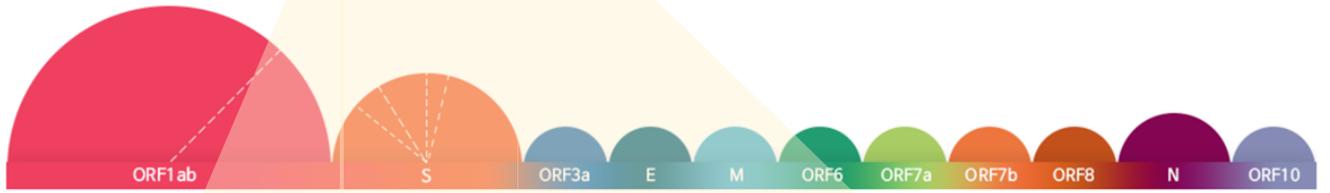
Omicron(BA.4)

Omicron(BA.5)

Wuhan-HU-1

Input virus Sequence

※ The length range for sequences that are covered is between 16 and 2700.



In AIVE, users can personally select and analyze the VOC and domains of SARS-CoV-2.

① Write the Project name for task categorization.

② Select SARS-CoV-2 in the Target Virus section to retrieve information on Coronaviruses.

③, ④, ⑤ Choose the VOC and regions you want to analyze to retrieve the corresponding Amino Acid sequence information.

Submit

Prediction

Virus

Project name **DEMO**

tutorial

Target virus [?]

SARS-CoV-2

Delta(B.1.617.2)

※ From the V.O.C. list, variants can only be generated for SARS-CoV-2 if 'Wuhan-Hu-1' is selected.

Input virus Sequence

※ The length range for sequences that are covered is between 16 and 2700.

NSNNLDSKVGGNYYRFLFRKSNLKPFRDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPY

Upload fasta file

Alignment [?] Unable to select alignment when selecting Demo and VOCs

437438439440441442443444445446447448449450451452453454455456457458459460461462463464465466467468469470471472473474475476477478479480481482483484

N S N N L D S K V G G N Y N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G S T P C N G V E

485486487488489490491492493494495496497498499500501502503504505506507508

G F N C Y F P L Q S Y G F Q P T N G V G Y Q P Y

Amino acids with electrically charged side chain_negative R W E

Amino acids with electrically charged side chain_positive D K

Amino acids with Polar uncharged side chain S T N Q

Special cases C G P

Amino acids with hydrophobic side chain A V I L M F Y W

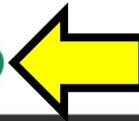
Keep private [?]

Variants protein sequence

NSNNLDSKVGGNYYRFLFRKSNLKPFRDISTEIYQAGSKPCNGVEGFNCYFPLQSYGFQPTNGVGYQPY

Copy Clipboard

Server Prediction



⑥

⑥ Click the "Server Prediction" button to start the prediction and analysis of the Project

List



Login Sign Up Sitemap

Home Prediction Report About Tutorial

List
Result viewer



wjchoi Logout Sitemap

Home Prediction Report About Tutorial



Prediction List

Job List

tutorial

Monomer
2 hour left

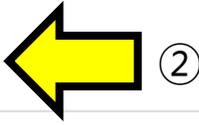
SARS-CoV-2

Process

Job List

tutorial

Monomer
[Result Info](#)



SARS-CoV-2

Complete

In the "List" section, users can check the list of submitted Projects and their progress.

① Click the "Report menu" and then select "List" from the submenu to check Projects.

② You can check the analysis results by clicking on the "Result info" section for completed tasks.

Generate SARS-CoV-2 mutated sequence

Project name **DEMO**

tutorial

Target virus [?]

All Viruses

All Viruses

SARS-CoV-2

Target virus [?]

SARS-CoV-2

※ From the V.O.C. list, variants can only be generated for SARS-CoV-2 if 'Wuhan-Hu-1' is

Select VOC

Select VOC

Alpha(B.1.1.7)

Beta(B.1.351)

Delta(B.1.617.2)

Gamma(P.1)

Omicron(BA.1)

Omicron(BA.2)

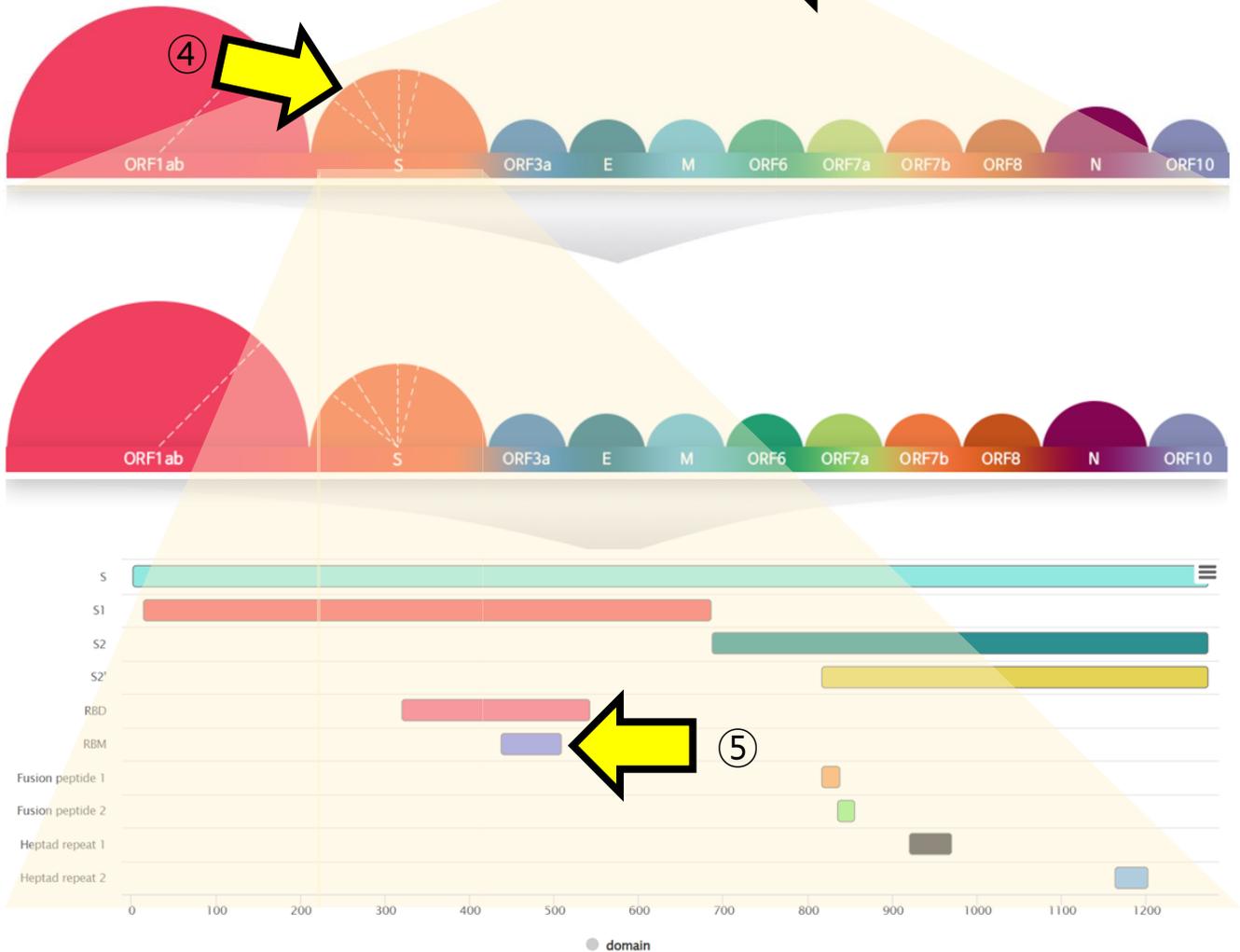
Omicron(BA.4)

Omicron(BA.5)

Wuhan-HU-1

Input virus Sequence

※ The length range for sequences that are covered is between 16 and 2700.



Users can not only access information about VOCs but also generate mutations for analysis.

① Specify a Project name to categorize the submitted task

②, ③, ④, ⑤ Choose Wuhan-HU-1 sequence of SARS-CoV-2 and retrieve the sequence of the region you want to verify.

Generate SARS-CoV-2 mutated sequence

Virus

Project name **DEMO**

tutorial

Target virus

SARS-CoV-2

Wuhan-HU-1

※ From the V.O.C. list, variants can only be generated for SARS-CoV-2 if 'Wuhan-Hu-1' is selected.

Input virus Sequence

※ The length range for sequences that are covered is between 16 and 2700.

NSNNLDSKVGGNYYLRLFRKSNLKPFDISTEIQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYPY

Upload fasta file

Alignment

437438439440441442443444445446447448449450451452453454455456457458459460461462463464465466467468469470471472473474475476477478479480481482483484

N S N N L D S K V G G N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G S T P C N G V E

485486487488489490491492493494

G F N C Y F P L Q S Y G F Q P

⑥

⑦

Ami S charged side chain_negative R H K

Ami T acids with electrically charged side chain_positive D E

Q Amino acids with Polar uncharged side chain S T N Q

C Special cases C G P

P Amino acids with hydrophobic side chain A V I L M F Y W

Keep private

Variants protein sequence

NSNNLDSKVGGNYYLRLFRKSNLKPFDISTEIQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYPY

⑧

Copy Clipboard

Server Prediction

⑨

- ⑥ In Alignment, click on the positions in the sequence of Wuhan-HU-1 to select the mutated Amino Acid.
- ⑦ Additionally, choose which codon to mutate into from the selected Amino Acid. This function is used to assess the impact of mutations at the codon level.
- ⑧ The sequence with the selected mutations is displayed.
- ⑨ Submit the task to proceed with the structural prediction and analysis of the mutated sequence.

User sequence - Monomer

Virus

Project name **DEMO**

tutorial

Target virus (?)

All Viruses

①

※ From the V.O.C. list, variants can only be generated for SARS-CoV-2 if 'Wuhan-Hu-1' is selected.

Input virus Sequence

※ The length of the sequence that are covered is between 16 and 12700.

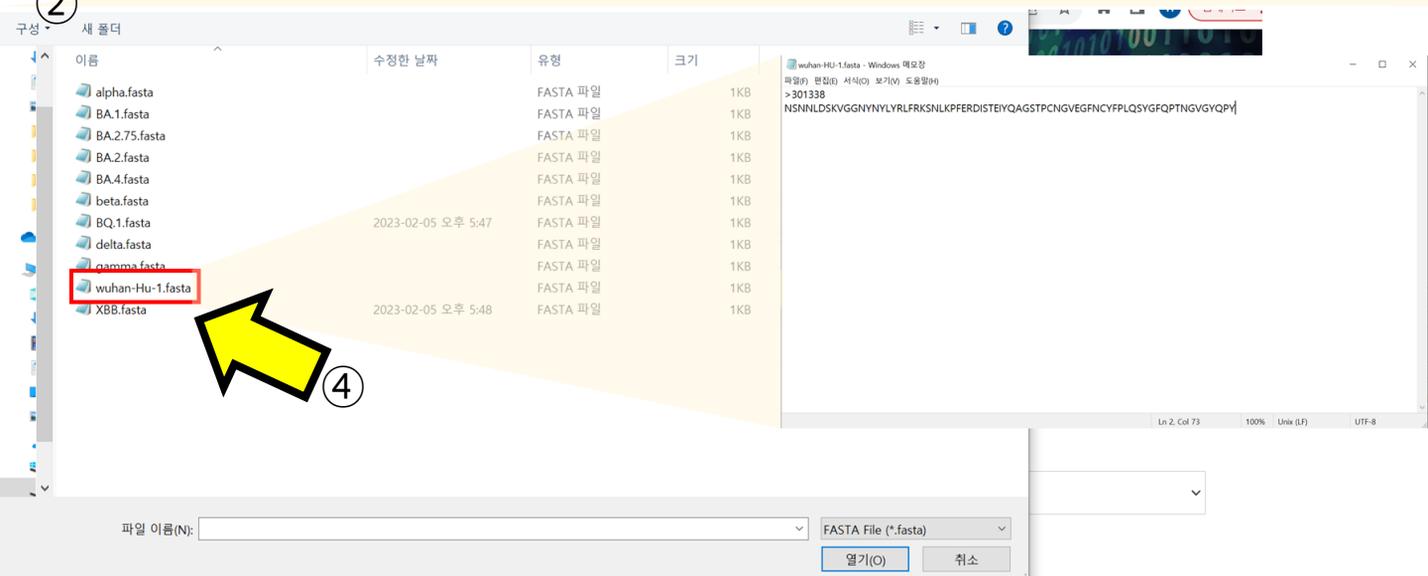
+ Add - Del

Upload fasta file

②

③

④



In AIVE, users can predict the structure of sequences, not limited to SARS-CoV-2. Let's take a look at monomer structure prediction

- ① Select "All viruses" in the Target Virus section.
- ② Enter the Amino Acid sequence you want to check in the "Input virus Sequence" box.
- ③ Alternatively, you can upload a fasta file instead of entering it directly.
- ④ When the upload window appears, select the fasta file you want to check to pull up the sequence.

User sequence

Virus

Project name **DEMO**

tutorial

Target virus [?]

All Viruses

※ From the V.O.C. list, variants can only be generated for SARS-CoV-2 if 'Wuhan-Hu-1' is selected.

Input virus Sequence

※ The length range for sequences that are covered is between 16 and 2700.

NSNNLDSKVGGNYNLYRLFRKSNLKPFRDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYPY

+ Add - Del

Upload fasta file

Alignment [?]

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48

N S N N L D S K V G G N Y N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G S T P C N G V E

49 50 51 52 53 54 55 56 57 58 59

G F N C Y F P L Q S Y Q P Y

Select codon

69 70 71 72

Y P Y

R CGT
H CGC
K CGA
D CGG
E AGA
S AGG

charged side chain_negative R H K

Acids with electriccally charged side chain_positive D E

Amino acids with Polar uncharged side chain S T N Q

Special cases C G P

Amino acids with hydrophobic side chain A V I L M F Y W

Keep private [?]

Variants protein sequence

NSNNLDSKVGGNYNLYRLFRKSNLKPFRDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYPY

Copy Clipboard

Server Prediction

- ⑥ The Amino Acid sequence recorded in the fasta file you uploaded is displayed in "Alignment".
- ⑦ Click on the positions in the displayed sequence where you want to introduce mutations and select the mutated Amino Acid.
- ⑧ Additionally, choose which codon to mutate into from the selected Amino Acid. This function is used to analyze the impact of mutations at the codon level.
- ⑨ The sequence with the selected mutations is displayed.
- ⑩ Submit the task to proceed with the structural prediction and analysis of the mutated sequence.

User sequence - Multimer

Project name **DEMO**

tutorial

Target virus [?]

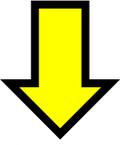
All Viruses

※ From the V.O.C. list, variants can only be generated for SARS-CoV-2 if 'Wuhan-Hu-1' is selected.

Input virus Sequence

※ The length range for sequences that are covered is between 16 and 2700.

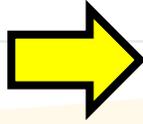
1



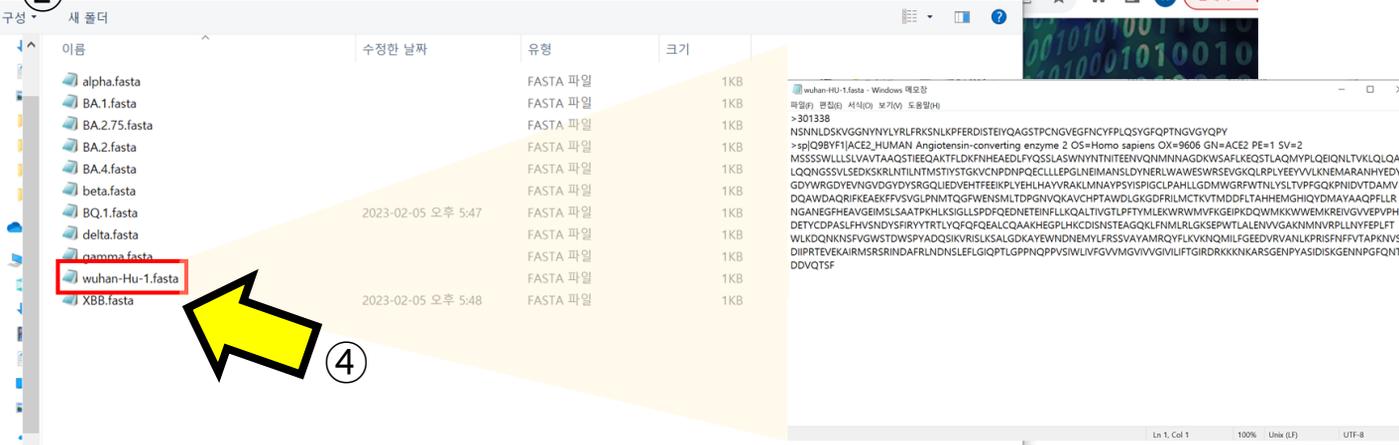
+ Add - Del

Upload fasta file

3



2



4



파일 이름(N):

FASTA File (*.fasta)

열기(O)

취소

Input virus Sequence

※ The length range for sequences that are covered is between 16 and 2700.

+ Add - Del

Upload fasta file

Input virus Sequence

※ The length range for sequences that are covered is between 16 and 2700.

NSNNLDSKVGGNYYLRLFRKSNLKPFRDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPY

+ Add - Del

MSSSSWLLLSLVAVTAAQSTIEEQAKTFLDKFNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAAGDKWSAFLKEQSTLAQMYPQLQEIQLNVLKQLQALQQNGS

+ Add - Del

Upload fasta file

5



Let's look at the case of predicting a Protein Complex structure:

- 1 Click the +Add button to create as many sequence input boxes as there are chains in the protein complex you want to predict.
- 2 Enter the Amino Acid sequence of each chain in the generated "Input virus Sequence" boxes.
- 3 Alternatively, a fasta file can be uploaded without entering the sequence directly
- 4 When the upload window appears, select the fasta file to retrieve the sequence.
- 5 The sequences of each chain, as stored in the fasta file, are inputted.

User sequence - Multimer

Alignment

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
N	S	N	N	L	D	S	K	V	G	G	N	Y	N	Y	L	Y	R	L	F	R	K	S	N	L	K	P	F	E	R	D	I	S	T	E	I	Y	Q	A	G	S	T	P	C	N	G	V	E

49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96
G	F	N	C	Y	F	P	L	Q	S	V	Q	P	S	V	I	E	E	Q	A	K	T	F	L	D	K	F	N	H	E	A	E	D	L	F	Y	Q	S	S	L	A	S	W					

97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144
L	Q	A	L	Q	Q	N	G	S	S	V	L	S	E	D	S	K	R	L	N	T	I	L	N	T	M	S	T	I	Y	S	T	G	K	V	C	N	P	D	N	P	Q	E	C	L	L	L	

145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192
E	P	G	L	N	E	I	M	A	N	S	L	D	Y	N	R	L	W	A	W	E	S	W	R	S	E	V	G	K	Q	L	R	P	L	Y	E	E	Y	V	V	L	K	N	E	M	A	R	

193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240
A	N	H	Y	E	D	Y	G	D	Y	W	R	G	D	Y	W	V	N	G	V	D	G	Y	D	Y	S	R	G	Q	L	I	E	D	V	E	H	T	F	E	E	I	K	P	L	Y	E	H	L

⋮

721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768
S	L	E	F	L	G	I	Q	P	T	L	G	P	P	N	Q	P	P	V	S	I	W	L	I	V	F	G	V	V	M	G	V	I	V	V	G	I	V	I	L	I	F	T	G	I	R	D	R

769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805
K	K	K	N	K	A	R	S	G	E	N	P	Y	A	S	I	D	I	S	K	G	E	N	N	P	G	F	Q	N	T	D	D	V	Q	T	S	F

Amino acids with electriccally charged side chain_negative R H K

Amino acids with electriccally charged side chain_positive D E

Amino acids with Polar uncharged side chain S T N Q

Special cases C G P

Amino acids with hydrophobic side chain A V I L M F Y W

Keep private 

Variants protein sequence

NSNNLDSKVGGNYYLRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPY

Copy Clipboard

MSSSSWLLLSLVAVTAAQSTIEEQAKTFLDKFNHEAEDLFYQSSLASWNYNTNITEENVQNMNAGDKWSAFLKEQSTLAQMYPLQEIQLTVKLQLQALQQI

Copy Clipboard

Server Prediction

- ⑥ Click on the positions in the displayed sequences where you want to introduce mutations and select the mutated Amino Acid.
- ⑦ Additionally, choose which codon to mutate into from the selected Amino Acid. This function is used to analyze the impact of mutations at the codon level.
- ⑧ The sequence with the selected mutations is displayed.
- ⑨ Submit the task to proceed with the structural prediction and analysis of the mutated sequence.

Result report page - Structure

AIVE analysis results

tutorial

① 3D Structure Prediction ②

- The PAE is a value that estimates the difference between the relative locations of two residues of the model and the real model. The value has a negative correlation with the accuracy of the pairwise position of two residues. As the PAE value decreases, the accuracy increases and vice versa.
- The pLDDT is a value that estimates the reliability of the model. The value indicates the likelihood of folding of the protein structure at that location, with higher values indicating a greater likelihood of folding.

1 2 3 4 5 ? Download all file

Protein structure (SARS-CoV-2, Reference) Protein structure (Mutation)

Predicted Aligned Error Predicted Aligned Error

Predicted LDDT box-plot box-plot

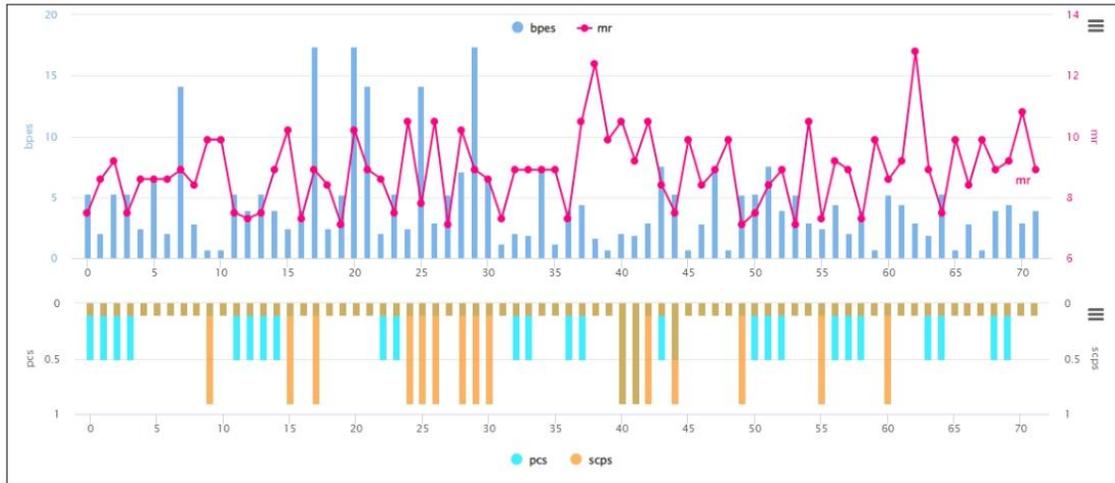
From the prediction and analysis results of SARS-CoV-2, the 3D structure prediction results can be accessed.

- ① You can select and view the predicted 3D structures by choosing from the five available options.
- ② Use the "Download all file" button to download the result files of the predicted structures to your device.
- ③ You can visualize the predicted 3D structure for inspection and comparison with SARS-CoV-2 Wuhan-HU-1. ④, ⑤ Clicking on the highlighted regions in positions allows you to inspect them.
- ④ Predicted aligned error (PAE) is a value that estimates the difference between the relative locations of two residues of the model and the real model. A low PAE value indicates that the accuracy of the relative location of the two residues is high.
 - The color at (x, y) indicates AlphaFold's expected position error at residue x if the predicted and true structures were aligned on residue y.
 - If the PAE is generally low for residue pairs x, y from two different domains, it indicates that AlphaFold predicts well-defined relative positions and orientations for them. (Explanation from AlphaFold FAQ)
- ⑤ Predicted LDDT (pLDDT) is a value that estimates the reliability of the model. It estimates how well the actual model residue and predicted model residue match. At the same time, it indicates how well the protein structure folds in the corresponding location. A low pLDDT value indicates that the reliability of the corresponding position is low and that it possesses a disordered structure.

Result report page - APESS

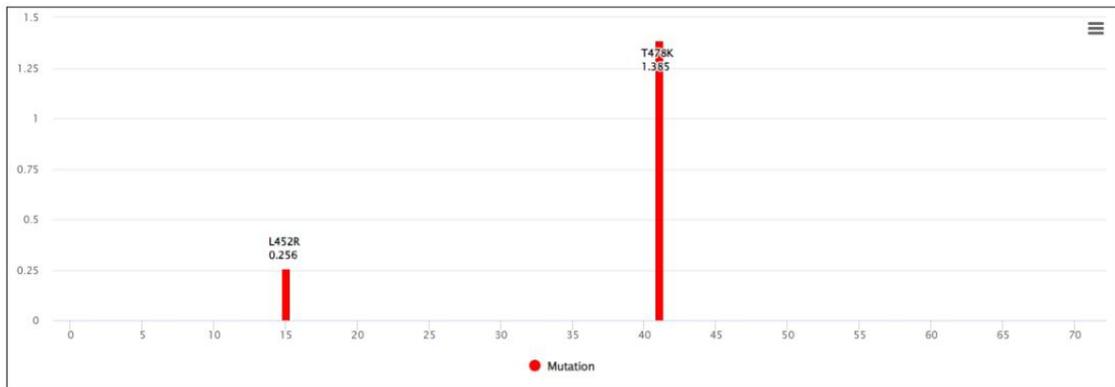
② Structure difference graph according to position

- Protein structure prediction characteristics (SCPS: SubClustering of Protein Structure in ①) and polarity (PCS: Polarity Change Score) are shown.
- SCPS divides the residues constituting the 3D protein structure into several groups using K-means clustering. The indicated groups are those with a high proportion of WHO name variants.
- PCS assigns weight to a position having a specific structure (where P appears consecutively) by polarity features of amino acid.
- Ratio change in frequencies of amino acids sequences (mr: mutation rate) and Biochemical properties of amino acid sequences (bpes: biochemical properties eigen score) are shown.
- MR is the rate of change of Amino Acid and its constituent nucleotides; the higher the rate, the higher the rate of change.
- BPES measures changes in the biochemical properties of amino acids at the site of a mutation. A high rate of change results in a high BPES value. Biochemical properties are measured by integrating amino acid residue, pH, and hydrophobic information.



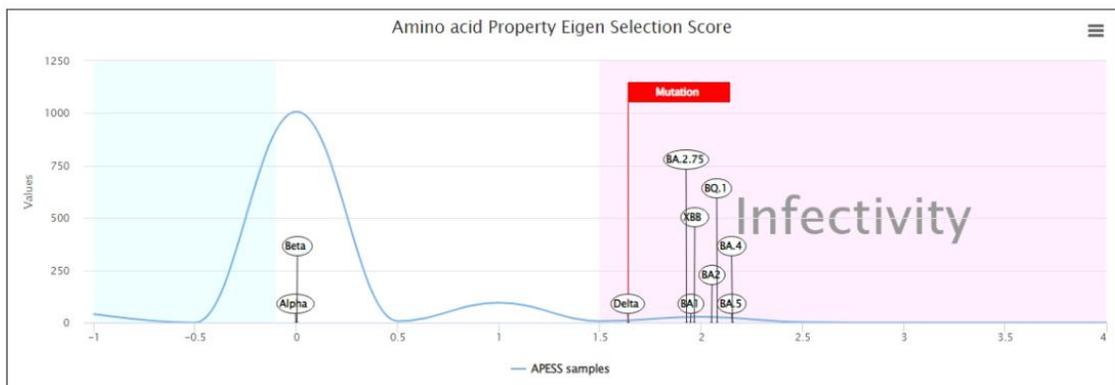
③ APES according to position

- APES is an amino acid property eigen selection score calculated by multiplying the values of SCPS, PCS, MR, and BPES.
- APES is a value that integrates the results of ② and ③. The higher the APES value, the more dangerous the mutation is.
- The graph shows the positions of mutations with structural differences (SCPS, PCS) in viral proteins. It also shows the positions and magnitudes of relatively large physical changes (MR) and biochemical changes (BPES).



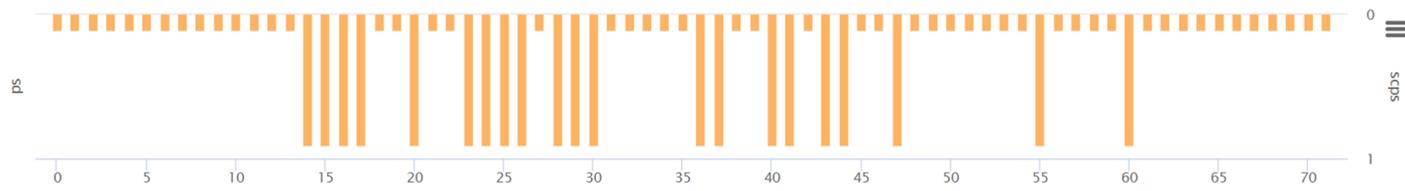
④ APES distribution graph

- The following graph shows the APES distribution of mutations generated through random sampling. The red area belongs to the quantile set of 0.05, signifying that the mutation is risky if it belongs to this area. A lettered balloon indicates the score position of the WHO VOC variant, and the red flag indicates the APES position of the protein predicted in ①.

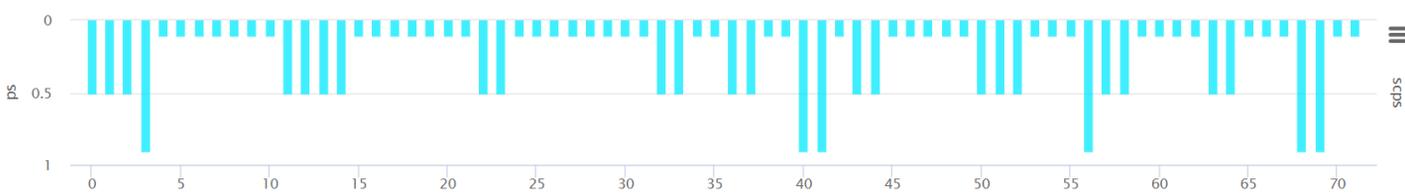


Result report page – APESs subscore

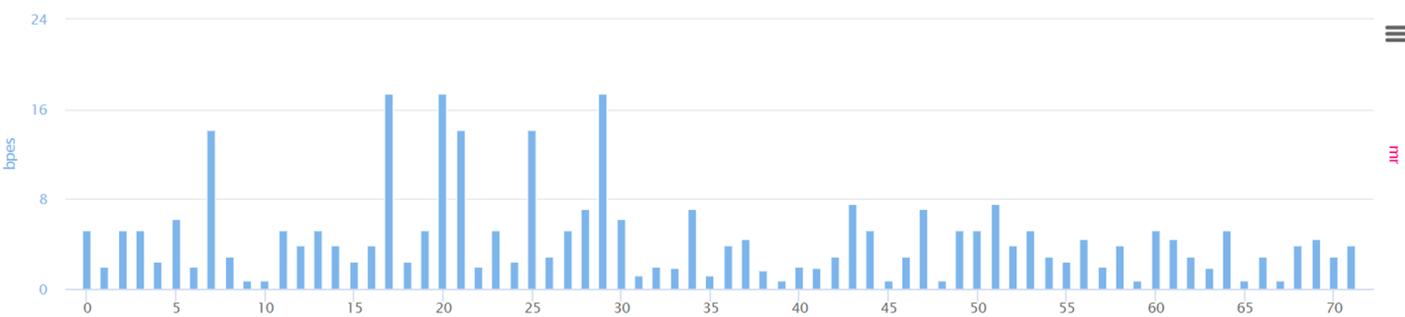
① SCPS



② PCS



③ BPES



④ MR

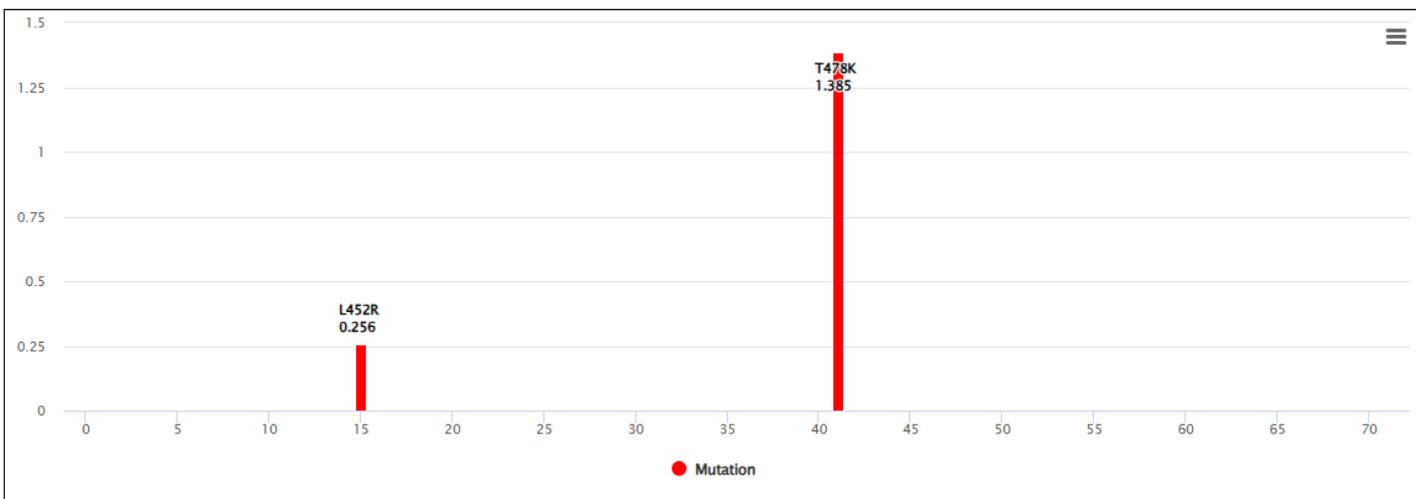


For the amino acid sequence entered by the user, the AIVE system provides a total of 6 evaluation charts.

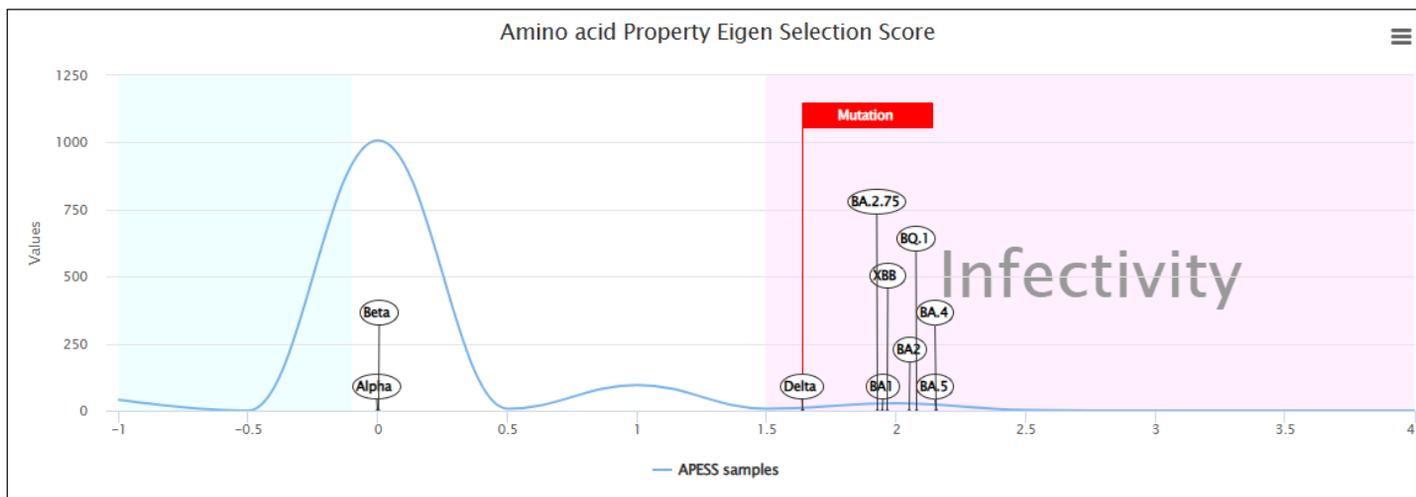
- ① AIVE predicts protein structures due to mutations in each gene for coronavirus lineages or sub-lineages. From the predicted result, it carries out grouping of amino acids (components of 3D protein structure) utilizing K-means clustering to report SCPS results.
- ② through repeated pattern analysis of polar amino acids in amino acid sequences, AIVE reports PCS results.
- ③ AIVE figures out amino acid properties to measure BPES through measurement of changes in biochemical properties of amino acid.
- ④ AIVE calculates MR through rate of change for nucleotide frequencies due to mutations.

Result report page – APRESS & distribution

⑤ APRESS



⑥ APRESS distribution graph



⑤ APRESS, the result of comprehensive mathematical model (SCP*PCS*MR*BPES) of measured analysis results is provided.

⑥ the APRESS distribution graph provides risk and spread results of the amino acid sequence entered by the user by comparing to VOCs' APRESS evaluation metric.

Result report page - Polarity

⑤ Virus amino acid info

- Visualizes and tabulates changes in repeated polarity structure sequence for wild-type sequences and mutated type sequences used in ①.

■ Sequence without mutations (Reference)

Reference amino acid sequence: N S N N L D S K V G G N Y N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G S T P C N G V E

Polarity characteristics (Reference): P P P P N A P B N N N P P P P N P B N N B B P P N B N N A B A N P P P N N P P N P P N N A

Mutation positions (Reference): P P P P H O P E H S S P H P H H E H H E E P P H E S H O E O H P P O H H P H S P E S S P S H O

Reference properties: G F N C Y F P L Q S Y G F Q P T N G V G Y Q P Y

Reference properties: N N P P P N N N P P P N N P P N N N P P N P

Reference properties: S H P S H H S H P P H S H P S P P S H S H P S H

■ Sequence with mutations (Mutation)

Mutated amino acid sequence: N S N N L D S K V G G N Y N Y R Y R L F R K S N L K P F E R D I S T E I Y Q A G S K P C N G V E

Polarity characteristics (Mutation): P P P P N A P B N N N P P P P B P B N N B B P P N B N N A B A N P P A N P P N N P B N P P N N A

Mutation positions (Mutation): P P P P H O P E H S S P H P H E H E H H E E P P H E S H O E O H P P O H H P H S P E S S P S H O

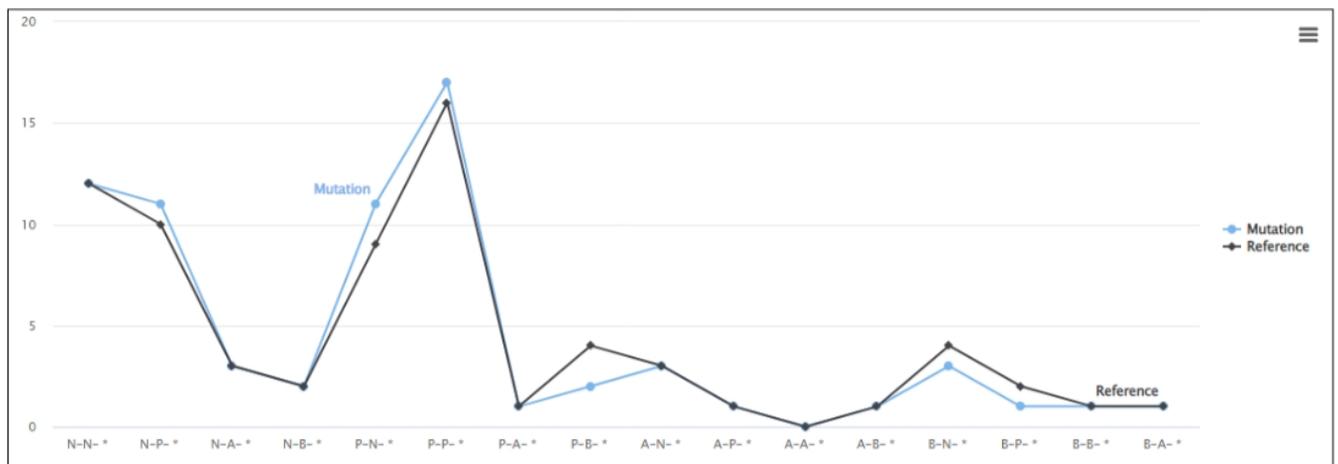
Mutation properties: G F N C Y F P L Q S Y G F Q P T N G V G Y Q P Y

Mutation properties: N N P P P N N N P P P N N P P N P P N N N P P N P

Mutation properties: S H P S H H S H P P H S H P S P P S H S H P S H

■ Polarity features

Polarity structure	Count	
	Reference	Mutation
ETC	41	42



Amino acid polarity affects protein structure and stability. As a result, the amino acid polarity due to mutation of the amino acid sequence input by the user can be observed. We found repeated polarity patterns in the coronavirus and observed changes in the properties of amino acid sequence polarity due to mutation. Therefore, we provide visualization and table view of polarity pattern changes to the user.

- ① amino acid sequence
- ② 4 polarity characteristics
- ③ 5 amino acid properties
- ④ Mutated positions are indicated in red.

Result report page – All viruses



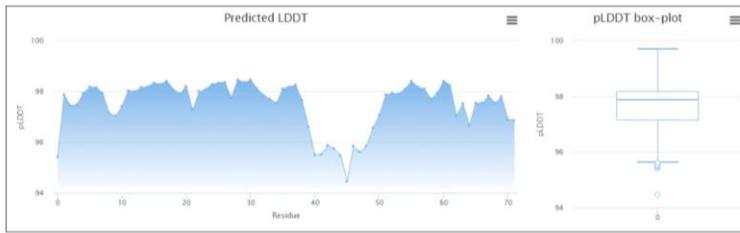
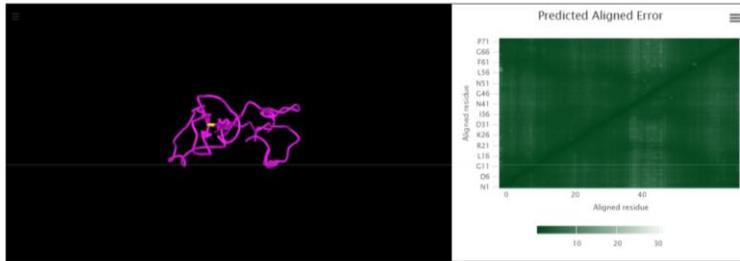
XBB

1 3D Structure Prediction

- The PAE is a value that estimates the difference between the relative locations of two residues of the model and the real model. The value has a negative correlation with the accuracy of the pairwise position of two residues. As the PAE value decreases, the accuracy increases and vice versa.
- The pLDDT is a value that estimates the reliability of the model. The value indicates the likelihood of folding of the protein structure at that location, with higher values indicating a greater likelihood of folding.

1 2 3 4 5 [Download all file](#)

Protein structure (Mutation) [Compare](#)



2 Virus amino acid info

- Visualizes and tabulates changes in repeated polarity structure sequence for wild-type sequences and mutated type sequences used in (3).

Sequence without mutations (Reference)

N S N K L D S K P S G N Y N Y L Y R L F R K S K L K P P F E R D I S T E I Y Q A G N K P C N G V A
 P P P B N A P D N P N P P P P P N P B N N B B P D B N B N A B A N P P A N P P P N N P B N P P N N N
 P P P E H O P E S P S P S P H P H H E H E H E E P E H E S H O E O H P P O H H P H S P E S S P S H H

G S N C Y S P L Q S Y G F R P T Y G V G H Q P Y
 N P P P P P N N P P P P N B N P P N N N B P N P
 S P P S H P S H P P H S H E S P H S H S E P S H

Sequence with mutations (Mutation)

N S N K L D S K P S G N Y N Y L Y R L F R K S K L K P P F E R D I S T E I Y Q A G N K P C N G V A
 P P P B N A P D N P N P P P P P N P B N N B B P D B N B N A B A N P P A N P P P N N P B N P P N N N
 P P P E H O P E S P S P S P H P H H E H E H E E P E H E S H O E O H P P O H H P H S P E S S P S H H

G S N C Y S P L Q S Y G F R P T Y G V G H Q P Y
 N P P P P P N N P P P P N B N P P N N N B P N P
 S P P S H P S H P P H S H E S P H S H S E P S H

	Polarity feature		Amino acid
Polarity features	Non-Polar	■	Ala (A), Val (V), Leu (L), Gly (G), Ile (I), Met (M), Trp (W), Phe (F), Pro (P)
	Polar	■	Ser (S), Cys (C), Asn (N), Gln (Q), Thr (T), Tyr (Y)
	Acidic	■	Asp (D), Glu (E)
	Basic	■	Lys (K), Arg (R), His (H)
Five amino acid properties	Amino acids with electrically charged side chain_negative	1	Lys (K), Arg (R), His (H)
	Amino acids with electrically charged side chain_positive	2	Asp (D), Glu (E)
	Amino acids with Polar uncharged side chain	■	Ser (S), Asn (N), Gln (Q), Thr (T)
	Special cases	4	Cys (C), Gly (G), Pro (P)
	Amino acids with hydrophobic side chain	5	Ala (A), Val (V), Leu (L), Ile (I), Met (M), Trp (W), Phe (F)

Polarity features

Polarity structure	Count	
	Reference	Mutation
ETC	45	45

Result report page –Compare

XBB

① 3D Structure Prediction ?

- The PAE is a value that estimates the difference between the relative locations of two residues of the model and the real model. The value has a negative correlation with the accuracy of the pairwise position of two residues. As the PAE value decreases, the accuracy increases and vice versa.
- The pLDDT is a value that estimates the reliability of the model. The value indicates the likelihood of folding of the protein structure at that location, with higher values indicating a greater likelihood of folding.

1 2 3 4 5 ?

Download all file

■ Protein structure (Mutation) Compare

①



Select predictions for comparison

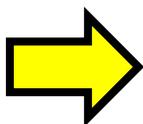


Projectname	Prediction	Targetvirus
alpha	Monomor	SARS-CoV-2
beta	Monomor	SARS-CoV-2
delta	Monomor	SARS-CoV-2
gamma	Monomor	SARS-CoV-2
BA.1	Monomor	SARS-CoV-2
BA.2	Monomor	SARS-CoV-2
BA.4	Monomor	SARS-CoV-2
		전체

ai-ve.org의 메시지

Do you want to compare with the selected results?

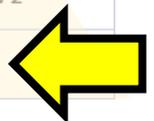
③



확인

취소

②



Structures predicted by "All viruses" can be compared with other structures using the compare feature.

- ① Click the "Compare" button to load the list of other tasks submitted by the user.
- ② Select the task you want to compare and click on it.
- ③ Click the "Confirm" button to navigate to the comparison page.

Result report page –Compare

XBB compare to test

① 3D Structure Prediction ②

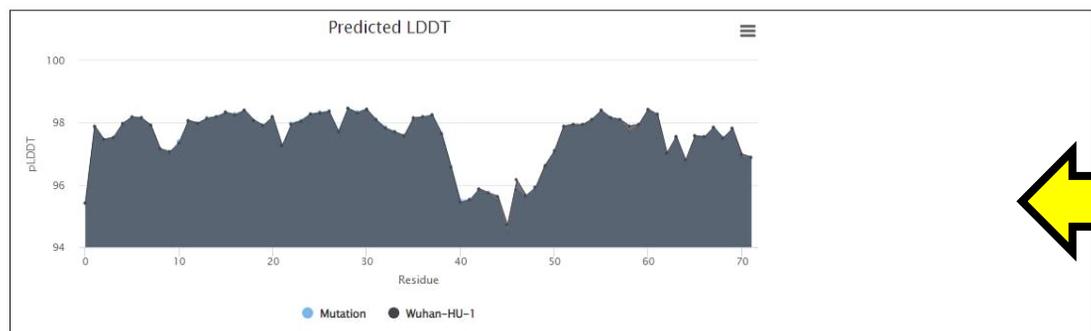
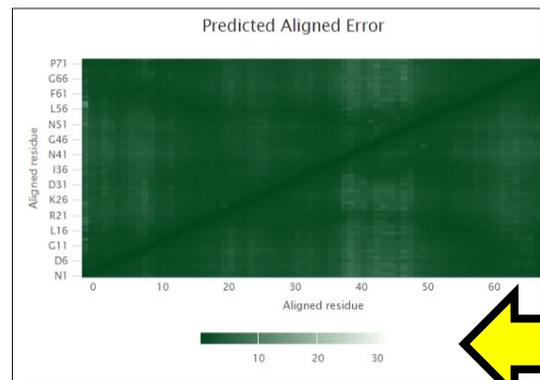
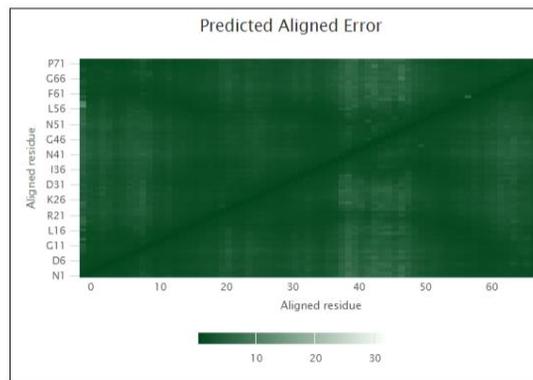
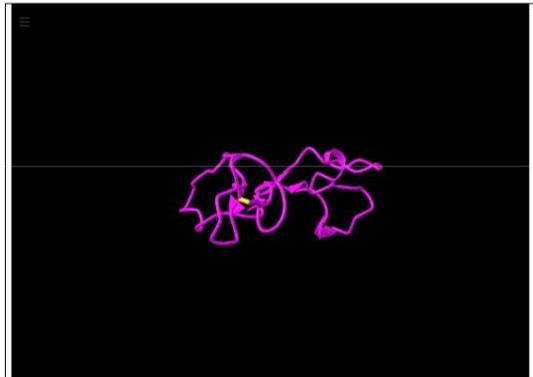
- The PAE is a value that estimates the difference between the relative locations of two residues of the model and the real model. The value has a negative correlation with the accuracy of the pairwise position of two residues. As the PAE value decreases, the accuracy increases and vice versa.
- The pLDDT is a value that estimates the reliability of the model. The value indicates the likelihood of folding of the protein structure at that location, with higher values indicating a greater likelihood of folding.



■ Protein structure (Original)



■ Protein structure (Compare target)



Through the "Compare" feature, you can compare two predicted structures in the list:

- ① You can visually inspect the PAE (Predicted Alignment Error) of the two structures using a plot.
- ② You can compare the pLDDT values of the two structures.

Result report page –Compare

② Virus amino acid info



- Visualizes and tabulates changes in repeated polarity structure sequence for wild-type sequences and mutated type sequences used in ①.

■ Sequence without mutations (Original)

N	S	N	K	L	D	S	K	P	S	G	N	Y	N	Y	L	Y	R	L	F	R	K	S	K	L	K	P	F	E	R	D	I	S	T	E	I	Y	Q	A	G	N	K	P	C	N	G	V	A
P	P	P	B	N	A	P	B	N	P	N	P	P	P	P	N	P	B	N	N	B	B	P	B	N	B	N	N	A	B	A	N	P	P	A	N	P	P	N	N	P	B	N	P	P	N	N	N
P	P	P	E	H	O	P	E	S	P	S	P	H	P	H	H	H	E	H	H	E	E	P	E	H	E	S	H	O	E	O	H	P	P	O	H	H	P	H	S	P	E	S	S	P	S	H	H



G	S	N	C	Y	S	P	L	R	S	Y	G	F	R	P	T	Y	G	V	G	H	Q	P	Y
N	P	P	P	P	N	N	B	P	P	N	N	B	N	P	P	N	N	N	B	P	N	P	
S	P	P	S	H	P	S	H	E	P	H	S	H	E	S	P	H	S	H	S	E	P	S	H

■ Sequence with mutations (Compare target)

N	S	N	K	L	D	S	K	P	S	G	N	Y	N	Y	L	Y	R	L	F	R	K	S	K	L	K	P	F	E	R	D	I	S	T	E	I	Y	Q	A	G	N	K	P	C	N	G	V	A
P	P	P	B	N	A	P	B	N	P	N	P	P	P	P	N	P	B	N	N	B	B	P	B	N	B	N	N	A	B	A	N	P	P	A	N	P	P	N	N	P	B	N	P	P	N	N	N
P	P	P	E	H	O	P	E	S	P	S	P	H	P	H	H	H	E	H	H	E	E	P	E	H	E	S	H	O	E	O	H	P	P	O	H	H	P	H	S	P	E	S	S	P	S	H	H



G	S	N	C	Y	S	P	L	Q	S	Y	G	F	R	P	T	Y	G	V	G	H	Q	P	Y
N	P	P	P	P	N	N	B	P	P	N	N	B	N	P	P	N	N	N	B	P	N	P	
S	P	P	S	H	P	S	H	P	H	S	H	E	S	P	H	S	H	S	E	P	S	H	

	Polarity feature		Amino acid
Polarity features	Non-Polar	R	Ala (A), Val (V), Leu (L), Gly (G), Ile (I), Met (M), Trp (W), Phe (F), Pro (P)
	Polar	P	Ser (S), Cys (C), Asn (N), Gln (Q), Thr (T), Tyr (Y)
	Acidic	A	Asp (D), Glu (E)
	Basic	B	Lys (K), Arg (R), His (H)
Five amino acid properties	Amino acids with electrically charged side chain_negative	1	Lys (K), Arg (R), His (H)
	Amino acids with electrically charged side chain_positive	2	Asp (D), Glu (E)
	Amino acids with Polar uncharged side chain	3	Ser (S), Asn (N), Gln (Q), Thr (T)
	Special cases	4	Cys (C), Gly (G), Pro (P)
	Amino acids with hydrophobic side chain	5	Ala (A), Val (V), Leu (L), Ile (I), Met (M), Trp (W), Phe (F)

■ Polarity features

Polarity structure	Count	
	Original	Compare target
ETC	46	45

③ The sequence and polar structures of the two structures can be compared and analyzed.