

# VirusDecode

: Decoding viruses, defending humanity

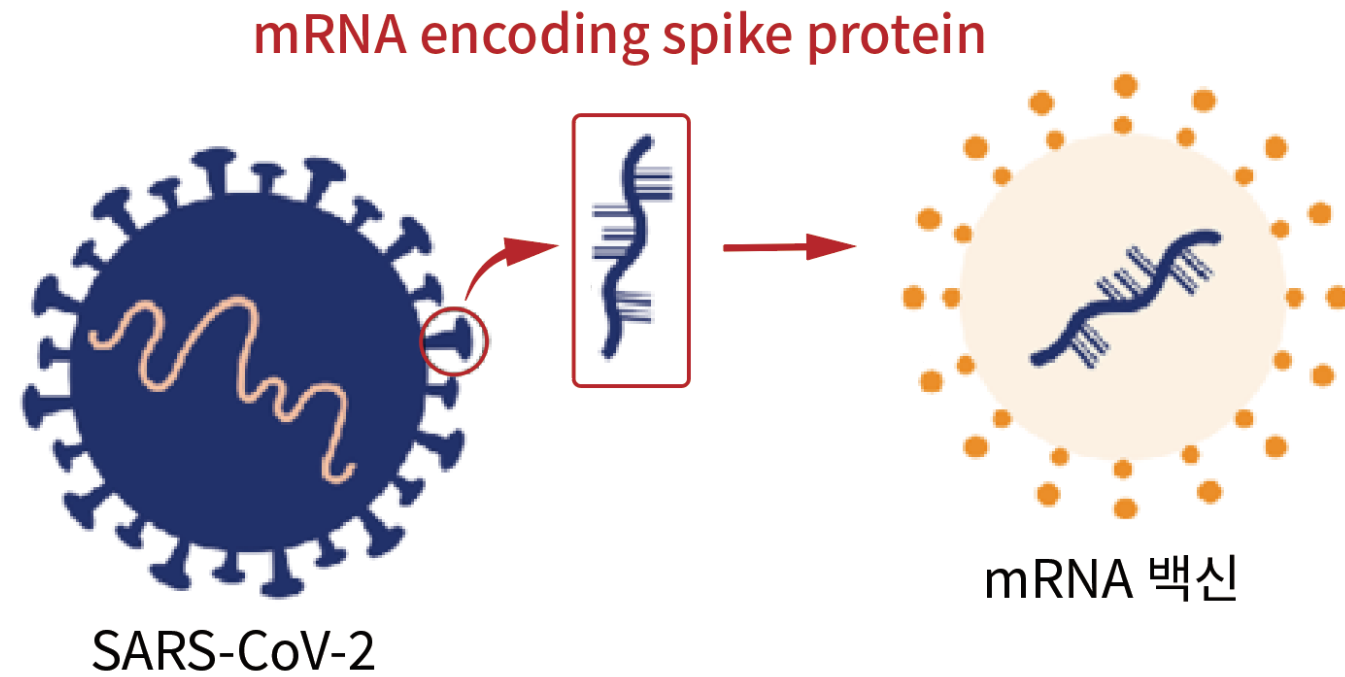
팀 구아바

개발팀: 김구호, 김경연, 박경일, 박다솔, 오성현

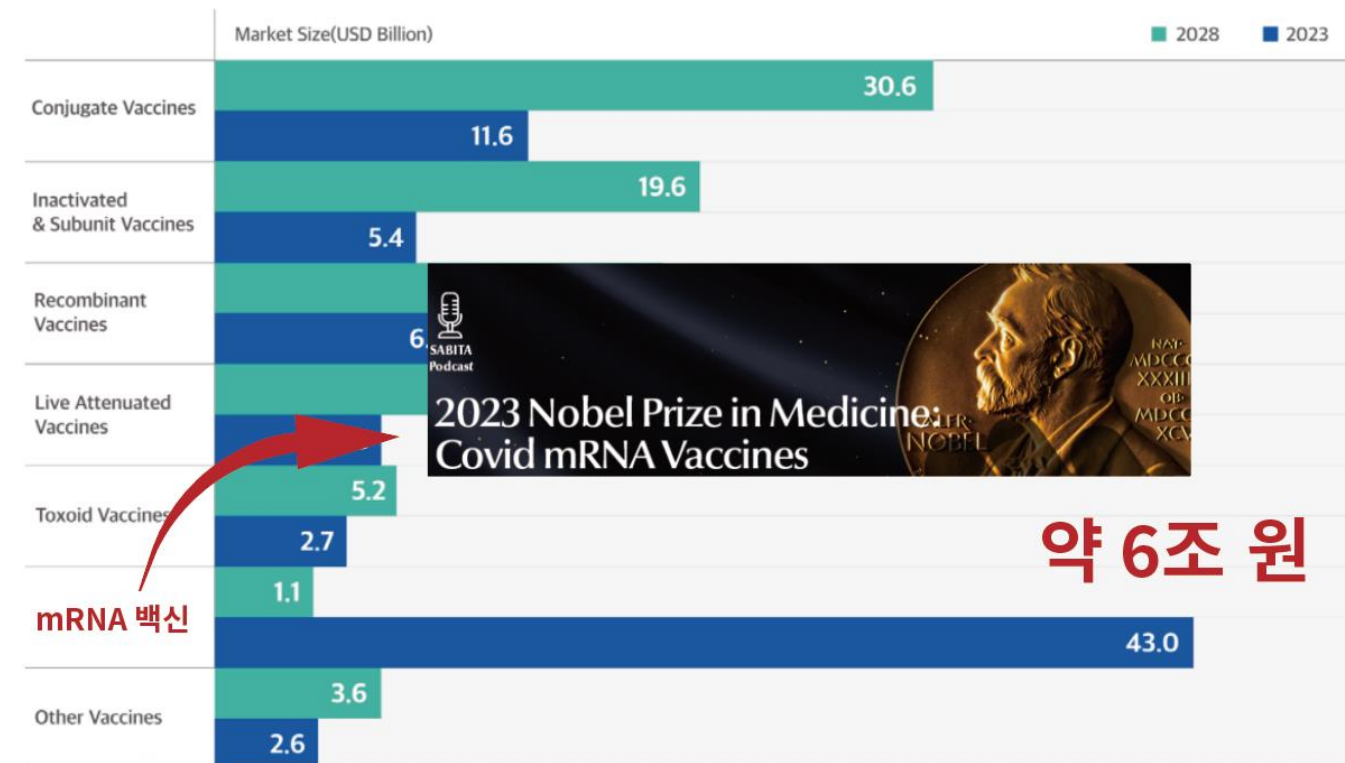
바이오팀: 김민겸, 황재현



Q 타임라인 검색			상태	범주	에픽	유형
스프린트						
<input type="checkbox"/>	<input checked="" type="checkbox"/>	VS-184 [Frontend] 도움말 기능 구현	완료			
<input type="checkbox"/>	<input checked="" type="checkbox"/>	VS-184 [Frontend] 탭 논리에 맞춰서 비활성화	완료			
<input type="checkbox"/>	<input checked="" type="checkbox"/>	VS-188 [Frontend] mRNA design 탭 배치 수정/ 제목 달아주기	완료			
<input type="checkbox"/>	<input checked="" type="checkbox"/>	VS-189 [Frontend] History tab 구현 및 백엔드 연동	완료			
<input type="checkbox"/>	<input checked="" type="checkbox"/>	VS-202 [Frontend] History에서 load시에, mRNA design, 3D viewer 관련 json 없...	완료			
<input type="checkbox"/>	<input checked="" type="checkbox"/>	VS-205 [Frontend] 새로고침시에 mRNA design, 3D viewer tab이 비활성화되는 ...	완료			
<input type="checkbox"/>	<input checked="" type="checkbox"/>	VS-220 [Frontend] Test Code - Login	완료			
<input type="checkbox"/>	<input checked="" type="checkbox"/>	VS-224 [Frontend] Test Code - InputSeq & Analysis	완료			
<input type="checkbox"/>	<input checked="" type="checkbox"/>	VS-209 [Frontend] JavaScript -> TypeScript	완료			
<input type="checkbox"/>	<input checked="" type="checkbox"/>	VS-32 Backend				
<input type="checkbox"/>	<input checked="" type="checkbox"/>	VS-33 Bioinformatics				
<input type="checkbox"/>	<input checked="" type="checkbox"/>	VS-49 [Bioinformatics] 1단계 - NCBI Parsing 알고리즘 구현	완료			
<input type="checkbox"/>	<input checked="" type="checkbox"/>	VS-47 [Bioinformatics] 2단계 - Mutation 추출 알고리즘 구현 (Based on DNA)	완료			
<input type="checkbox"/>	<input checked="" type="checkbox"/>	VS-70 [Bioinformatics] 2단계 - Alignment 알고리즘 구현 (Based on DNA)	완료			



## 해외 시장 규모



mRNA백신 연구동향 분석 보고서 (한국화학연구원 감염병기술전략센터, 2024)

- SARS-CoV-2 바이러스로 인한 COVID-19 확진자 수: 약 7억 7천만 명
  - mRNA 백신으로 팬데믹 피해를 최소화
  - 하지만 **돌파 감염** 문제 발생
  - **바이러스 변이**가 주 원인
  - 변종 바이러스에 대한 새로운 mRNA 백신을 제작해야 함.

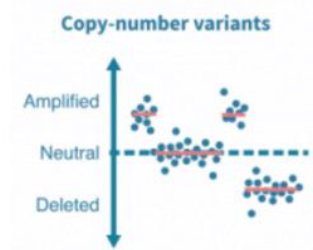
- 바이러스 변이는 빠르고 무작위적으로 발생한다.
  - Research key = **mRNA 백신 개발 속도 향상**

### Project goal

- **변이 바이러스 유전체 분석 및 mRNA 변환 기능을 통합한 소프트웨어(VirusDecode)** 개발
- mRNA 백신 개발 속도 향상

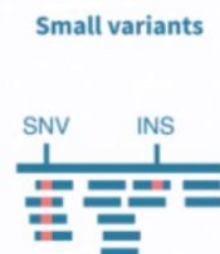
## Traditional workflow

### 바이러스 서열 분석 C



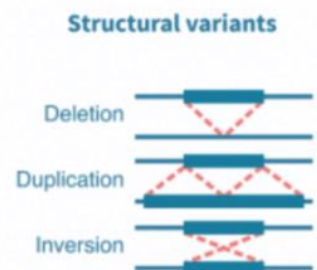
### 바이러스 서열 정보 탐색

### 바이러스 서열 분석 A



### mRNA 항원 최적화

### 바이러스 서열 분석 B



### mRNA 백신 후보

### mRNA 안정성 최적화



## VirusDecode

### 바이러스 서열 정보

NCBI

Unknown genome

### Alignment 및 변이 분석

Muscle5

라이선스: GNU  
기능: 다중 서열 정렬

Viztein

라이선스: MIT  
기능: 3차 구조 시각화

- \*변이 바이러스 판별
- \*변이 위치 파악
- \*단백질 3D 시각화

### mRNA 변환 및 분석

Linear design

라이선스: Custom License  
기능: mRNA 디자인

Fornac

라이선스: Apache 2.0  
기능: 2차 구조 시각화

- \*mRNA 변환 부위 선택
- \*mRNA 안정성 및 백신 효율 평가
- \*mRNA 2D 시각화

### 기존의 mRNA 백신 디자인 방식

- 변이 바이러스 유전체 분석, mRNA 변환, mRNA 안정성 평가를 진행하기 위해 각각의 SW 이용 필요
- 비효율적, 연구개발 속도 저하

### VirusDecode

- 바이러스 유전체 분석을 통한 변이 분석/ mRNA 변환/ 안정성 및 백신 효율 평가를 통합한 소프트웨어
- mRNA 백신 개발 속도 향상에 기여

**Muscle5** **Muscle5**  
GNU License

 **LinearDesign**  
Custom License

**fornac** v1.1.8 **fornac**  
Apache 2.0 License

**viztein** v0.1.8 **viztein**  
MIT License

**Muscle5**  
License GNU

새로운 변이 바이러스 서열

NC_045512.2	Q	P	T	N	G	V	G	Y	Q	P	Y	R	V	V	V	L	S	F	E	L	L	H
OR240434.1	Q	P	T	N	G	V	G	Y	Q	P	Y	R	V	V	V	L	S	F	E	L	L	H
PP346415.1	Q	P	T	N	G	V	G	Y	Q	P	Y	R	V	V	V	L	S	F	E	L	L	H
MT576556.1	Q	P	T	N	G	V	G	Y	Q	P	Y	R	V	V	V	L	S	F	E	L	L	H
MW642250.1	Q	P	T	Y	G	V	G	Y	Q	P	Y	R	V	V	V	L	S	F	E	L	L	H
OL672836.1	R	P	T	Y	G	V	G	H	Q	P	Y	R	V	V	V	L	S	F	E	L	L	H
	501								510													520

다중 서열 정렬 / 유전자 주석 / 변이 서열 파악

**Baidu Research**

mRNA Sequence  
AGCAACAAGCUGGACAGCAAGGUGUCAGGUAAUUAACAUAUACCUGUACCG ...show

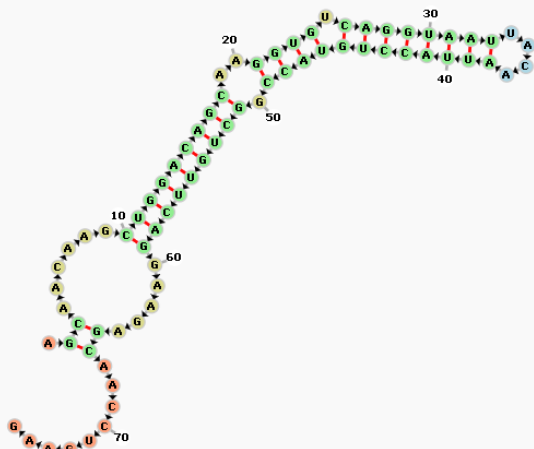
mRNA Structure  
..(((.....(((((((((((.....((((((((((((.....)))))))))))))). show

mRNA folding free energy  
-29.70 kcal/mol

mRNA CAI  
0.938

mRNA 안정성 및 백신 효율 평가

**fornac** v1.1.8  
License Apache-2.0



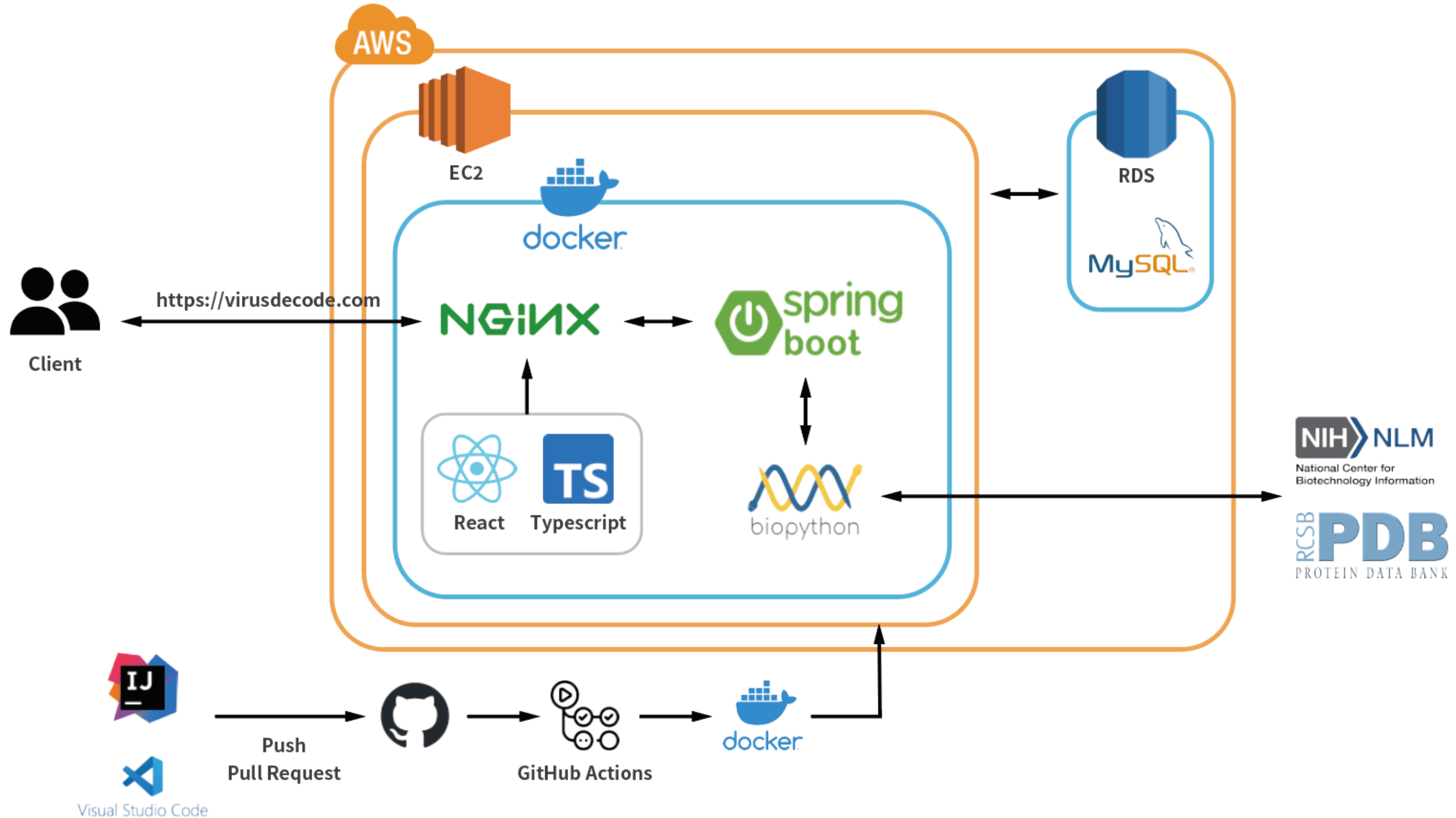
mRNA 시각화

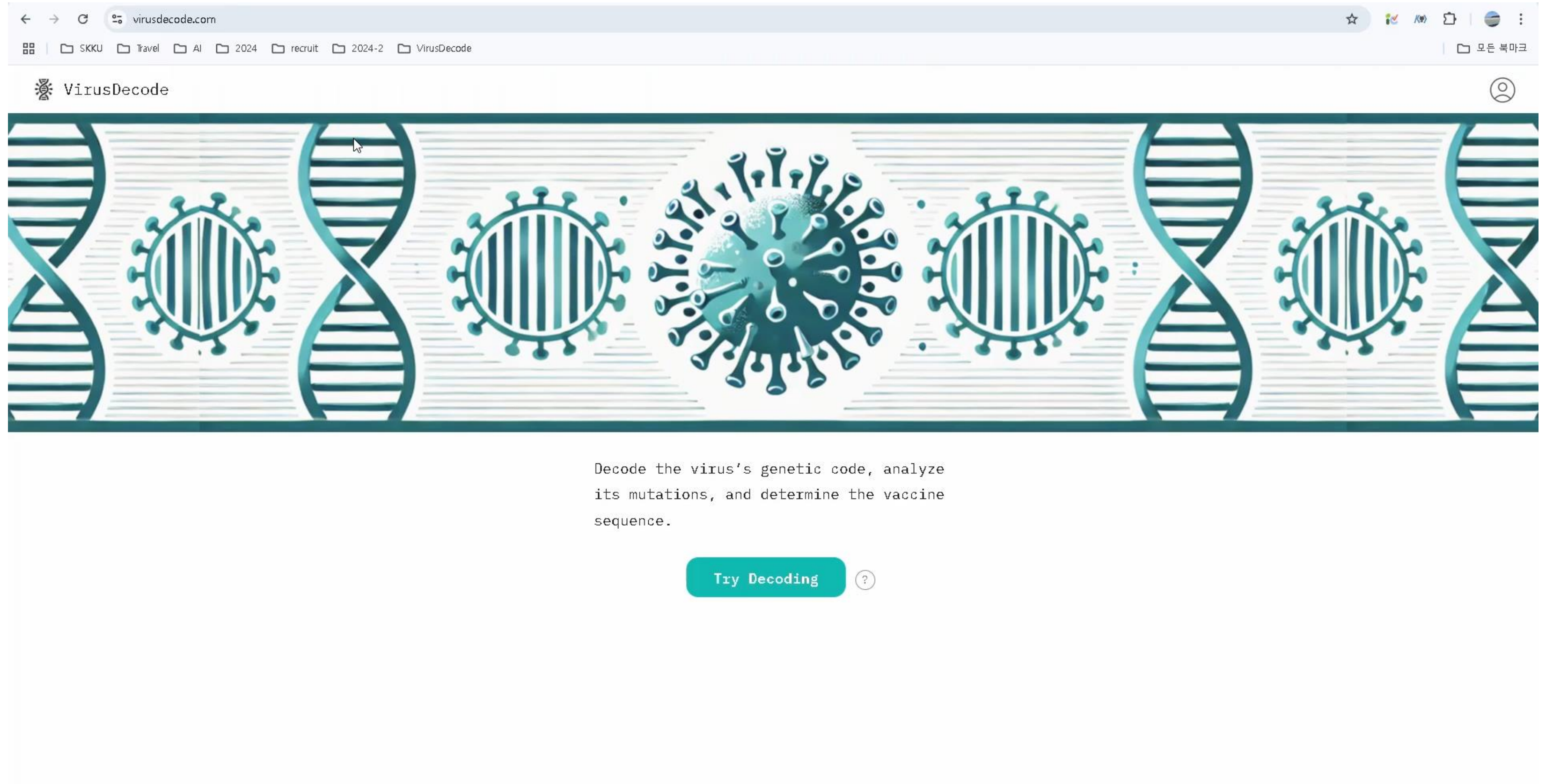
**viztein** v0.1.8  
License MIT



단백질 3D 시각화









pytest

JUnit 

```
gk@GK-LIONINE:~/VirusDecode-project/VirusDecode/bioinformatics$ pytest -v test/
===== test session starts =====
platform linux -- Python 3.10.12, pytest-8.3.3, pluggy-1.5.0 -- /usr/bin/python3
cachedir: .pytest_cache
rootdir: /home/gk/VirusDecode-project/VirusDecode/bioinformatics
collected 24 items

test/test_alignment.py::TestAlignmentSuccessCases::test_run_muscle_dna_success PASSED [ 4%]
test/test_alignment.py::TestAlignmentSuccessCases::test_validate_and_fetch_reference_sequence_success PASSED [ 8%]
test/test_alignment.py::TestAlignmentFailureCases::test_fetch_reference_sequence_failure PASSED [ 12%]
test/test_alignment.py::TestAlignmentFailureCases::test_run_muscle_dna_failure PASSED [ 16%]
test/test_alignment.py::TestAlignmentFailureCases::test_validate_sequences_parsing_error PASSED [ 20%]
test/test_mrna_design.py::TestMRNADesignSuccessCases::test_run_linear_design_success PASSED [ 25%]
test/test_mrna_design.py::TestMRNADesignSuccessCases::test_set_protParam_success PASSED [ 29%]
test/test_mrna_design.py::TestMRNADesignFailureCases::test_run_linear_design_directory_not_found PASSED [ 33%]
test/test_mrna_design.py::TestMRNADesignFailureCases::test_run_linear_design_failure PASSED [ 37%]
test/test_mrna_design.py::TestMRNADesignFailureCases::test_set_protParam_invalid_sequence PASSED [ 41%]
test/test_metadata.py::TestMetadataSuccessCases::test_print_metadata PASSED [ 45%]
test/test_metadata.py::TestMetadataSuccessCases::test_run PASSED [ 50%]
test/test_metadata.py::TestMetadataSuccessCases::test_set_metadata_http_error PASSED [ 54%]
test/test_metadata.py::TestMetadataSuccessCases::test_set_metadata_success PASSED [ 58%]
test/test_metadata.py::TestMetadataFailureCases::test_set_metadata_incomplete_data PASSED [ 62%]
test/test_metadata.py::TestMetadataFailureCases::test_set_metadata_invalid_reference_id PASSED [ 66%]
test/test_metadata.py::TestMetadataFailureCases::test_set_metadata_parsing_error PASSED [ 70%]
test/test_viewer_3d.py::TestThreeDViewerSuccessCases::test_fetch_pdb_info_success PASSED [ 75%]
test/test_viewer_3d.py::TestThreeDViewerSuccessCases::test_get_pdb_ids_by_sequence_success PASSED [ 79%]
test/test_viewer_3d.py::TestThreeDViewerSuccessCases::test_get_pdb_info_success PASSED [ 83%]
test/test_viewer_3d.py::TestThreeDViewerSuccessCases::test_run_success PASSED [ 87%]
test/test_viewer_3d.py::TestThreeDViewerFailureCases::test_get_pdb_ids_by_sequence_timeout PASSED [ 91%]
test/test_viewer_3d.py::TestThreeDViewerFailureCases::test_get_pdb_info_request_exception PASSED [ 95%]
test/test_viewer_3d.py::TestThreeDViewerFailureCases::test_get_pdb_info_timeout PASSED [100%]


===== 24 passed in 0.35s =====
```

Element ^	Class, %	Method, %	Line, %	Branch, %
✓ VirusDecode.backend	53% (14/26)	98% (53/54)	99% (314/315)	100% (90/90)
> config	100% (2/2)	100% (4/4)	100% (8/8)	100% (0/0)
> controller	100% (4/4)	100% (17/17)	100% (81/81)	100% (26/26)
> dto	0% (0/8)	100% (0/0)	100% (0/0)	100% (0/0)
> entity	0% (0/3)	100% (0/0)	100% (0/0)	100% (0/0)
> repository	100% (0/0)	100% (0/0)	100% (0/0)	100% (0/0)
> service	100% (8/8)	100% (32/32)	100% (225/225)	100% (64/64)
BackendApplication	0% (0/1)	0% (0/1)	0% (0/1)	100% (0/0)



Spec		Tests	Passing	Failing	Pending	Skipped
✓ 01_SignUp.cy.js	00:13	4	4	—	—	—
✓ 02_LoginLogout.cy.js	00:09	4	4	—	—	—
✓ 03_GuestLogin.cy.js	00:03	2	2	—	—	—
✓ 04_HeaderBar.cy.js	00:13	4	4	—	—	—
✓ 05_History.cy.js	00:10	3	3	—	—	—
✓ 06_SequenceInput.cy.js	01:11	11	11	—	—	—
✓ 07_SequenceAnalysisAndAlignment.cy.js	02:30	6	6	—	—	—
✓ 08_SequenceConverting.cy.js	00:55	5	5	—	—	—
✓ 09_mRNADesign.cy.js	00:35	3	3	—	—	—
✓ 10_Protein3DDesign.cy.js	01:12	4	4	—	—	—
✓ All specs passed!	07:15	46	46	—	—	—













<https://github.com/VirusDecode-project/VirusDecode>


**VirusDecode**
Public

Edit Pins
Watch 0
Fork 6
Starred 2

main
4 Branches
Tags

Add file
Code

<div>  <b>Guho-Kim</b> <span>fix: Update deploy.yml</span> <span>87755bb · 2 days ago</span> <span>608 Commits</span> </div>		
	.github/workflows	fix: Update deploy.yml 2 days ago
	backend	fix: NCBI 조회 실패한 경우 db에 저장 안 되도록 수정 4 days ago
	bioinformatics	fix: Handle sequences starting with stop codons in translation last week
	frontend	test: Update test code 5 days ago
	.dockerignore	fix: Update .dockerignore for __pycache__ last week
	.gitignore	chore: Update .gitignore last week
	Dockerfile.backend	fix: Modify bioinformatics path last week
	Dockerfile.frontend	fix: Add default.conf in Dockerfile last week
	LICENSE	Update LICENSE from MIT to Apache-2.0 2 weeks ago
	README.md	docs: Update README.md last week
	docker-compose.yml	fix: Add default.conf in Dockerfile last week

About

VirusDecode integrates bioinformatics tools for efficient virus sequence analysis, supporting rapid mRNA vaccine development with a biopython.

[www.virusdecode.com](http://www.virusdecode.com)

bioinformatics
virus
biopython
sars-cov-2

Readme
Apache-2.0 license
Activity
Custom properties
2 stars
0 watching
6 forks

Report repository

Releases

No releases published

- 각 개발자의 다양한 도구 추가
  - 플랫폼의 지속적인 발전
  - 바이러스 변이와 백신 연구의 효율성 증폭

README
Apache-2.0 license

## VirusDecode

**VirusDecode** is an open-source project designed to streamline virus sequence analysis by integrating various bioinformatics tools, supporting rapid mRNA vaccine development. With a Java Spring backend, a React frontend, and Biopython for bioinformatics, VirusDecode addresses traditional inefficiencies in genome analysis, enabling quick responses to viral mutations and fostering scientific collaboration.

Please note that two additional open-source tools, MUSCLE for sequence alignment and LinearDesign for mRNA structure prediction, are required but not included in this repository. To set up these tools, follow the provided installation guidelines.

The project structure is organized as follows:

```

.
├── frontend
├── backend
└── bioinformatics
    └── analysis
        ├── metadata.py
        ├── alignment.py # Uses MUSCLE
        ├── mRNA_design.py # Uses LinearDesign
        └── viewer_3d.py
    
```

The bioinformatics directory contains the core components of the bioinformatics solution, with each script utilizing tools like MUSCLE and LinearDesign via predefined paths. These tools are accessed through the analysis scripts, which interface with the web API to perform sequence analysis and return results to users.



추가 기능 구현: 변이 바이러스 계통수 분석



- Home
- Gallery
- Documentation
- TreeView
- Support
- About
- Download
- latest v3.1.3

A Python framework for the analysis and visualization of trees.

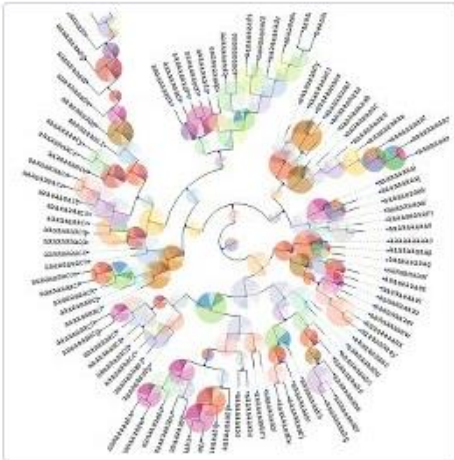
- Download
- Python API
- Cookbook
- Phylogenomic tools
- Contribute

```
from ete3 import Tree
tree = Tree('((A,B), D);')

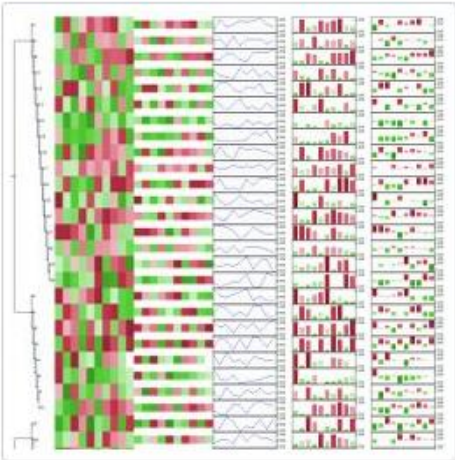
print tree
#      /-A
#     /-|
#    --| \-B
#     \-D

A = tree & "A"
A.up.show()
```

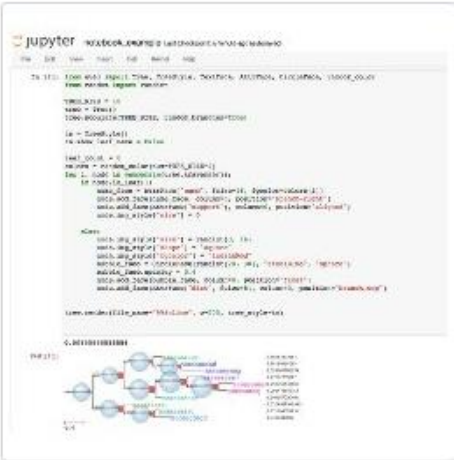
Trees as Python objects



Programmatic tree visualization



Tree annotation  
Custom node attributes can be



Jupyter notebook support

추가 기능 구현: 오픈 소스 개발자들의 참여

etePublic

SponsorWatch34

ete412 Branches56 Tags

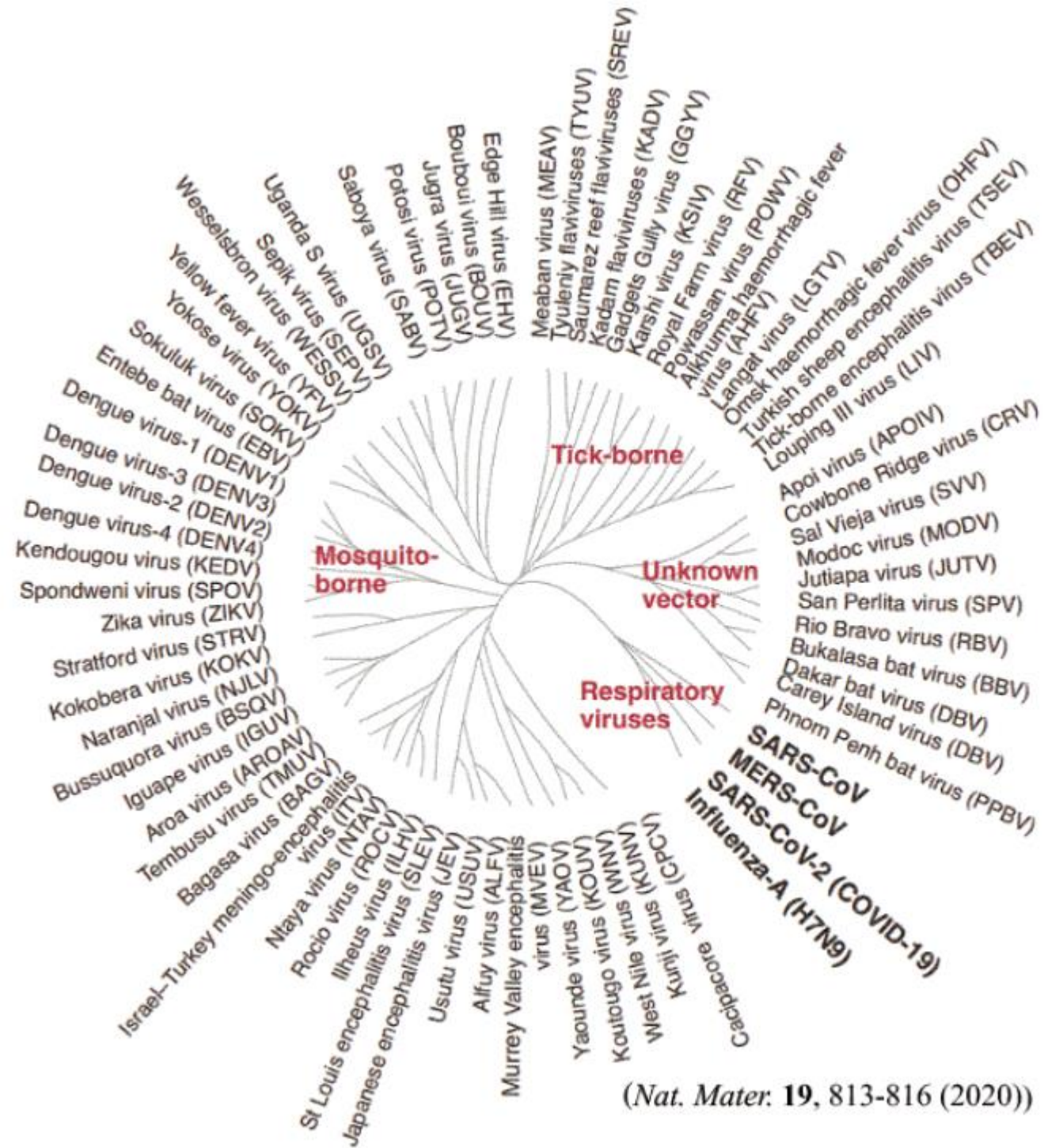
Go to file

Add file

Code

jordibc test\_orthologs\_group\_delineation.py: Simplify and use smaller te...a88743b · 6 months ago3,963 Commits

.github	+ run_tests.py	last year
doc	Fixes in t.resolve_polytomy().	10 months ago
ete4	Make update_ete_data() check by default if the file exists.	6 months ago
examples	cosmetics	last year
tests	test_orthologs_group_delineation.py: Simplify and use sm...	6 months ago
utils	Remove sdoc/ directory in preparation for an up-to-date ...	last year
.gitignore	Put temporary test files under ETE_DATA_HOME/tests for t...	last year
.travis.yml	disables python 3.4 travis test	4 years ago
CODE_OF_CONDUCT.md	Update CODE_OF_CONDUCT.md	3 years ago
CONTRIBUTING.md	Small fixes in CONTRIBUTING.md	last year
LICENSE	added new information files, renamed unittest dir	14 years ago
README.md	README.md: Remove pyqt6 from the conda packages - it i...	last year
THANKS.md	Update THANKS file.	last year
VERSION	Bump version to 4.0.0-beta2.	last year
pyproject.toml	Bump version to 4.0.0-beta2.	last year
run_tests.py	run_tests.py: Add --list option.	6 months ago
setup.py	setup.py: Automatically get all the pyx files for cythonize.	last year



# VirusDecode

: Decoding viruses, defending humanity



**Q&A**

 README  Apache-2.0 license

## VirusDecode

**VirusDecode** is an open-source project designed to streamline virus sequence analysis by integrating various bioinformatics tools, supporting rapid mRNA vaccine development. With a Java Spring backend, a React frontend, and Biopython for bioinformatics, VirusDecode addresses traditional inefficiencies in genome analysis, enabling quick responses to viral mutations and fostering scientific collaboration.

Please note that two additional open-source tools, MUSCLE for sequence alignment and LinearDesign for mRNA structure prediction, are required but not included in this repository. To set up these tools, follow the provided installation guidelines.

The project structure is organized as follows:

```
.
├── frontend
├── backend
└── bioinformatics
    └── analysis
        ├── metadata.py
        ├── alignment.py # Uses MUSCLE
        ├── mRNA_design.py # Uses LinearDesign
        └── viewer_3d.py
```



The bioinformatics directory contains the core components of the bioinformatics solution, with each script utilizing tools like MUSCLE and LinearDesign via predefined paths. These tools are accessed through the analysis scripts, which interface with the web API to perform sequence analysis and return results to users.

## 2. Alignment Tool

- Download and install [MUSCLE v3.8.1551](#):

```
mkdir muscle
cd muscle
wget https://www.drive5.com/muscle/muscle_src_3.8.1551.tar.gz
tar -xvzf muscle_src_3.8.1551.tar.gz
make
sudo cp muscle /usr/local/bin/
cd ..
rm -r muscle
```

## 3. Clone LinearDesign

- VirusDecode uses the LinearDesign open-source software for specific bioinformatics analysis. You must clone LinearDesign separately as per its license terms:

```
git clone https://github.com/LinearDesignSoftware/LinearDesign.git
cd LinearDesign
make
cd ..
```

- Note: The LinearDesign code is free for academic, non-profit, and research use. Redistribution of the code with or without modification is not permitted without explicit written permission from the lead corresponding author. If you intend to use this software for commercial purposes, please contact the lead corresponding author for licensing.

- Using LinearDesign with VirusDecode

In our project, we used LinearDesign with a specific option, --lambda 3, for optimal performance. The command we used is as follows:

```
"/lineardesign --lambda 3"
```

## 3. Usage

To run the VirusDecode application:


### 1. Clone our Repository

- Clone the project

```
git clone https://github.com/VirusDecode-project/VirusDecode.git
cd VirusDecode
```

- Integration with Project Directory: After cloning the repository, it should appear in the project directory structure as follows:

```
.
├── LinearDesign
└── VirusDecode
    ├── frontend
    ├── backend
    ├── bioinformatics
    └── ...
```


 **muscle** Public

Watch 6


main 3 Branches 4 Tags











Add file

<> Code

 **rcedgar** Working on workflows

80cd944 · 3 months ago

 204 Commits

 .github/workflows	Working on workflows	3 months ago
 src	Working on workflows	3 months ago
 test_data	Add test_data/scripts	3 months ago
 test_results	Add test_results	3 months ago
 test_scripts	Add test_scripts/check_results.py and RUN_TESTS.bash	3 months ago
 .gitattributes	Fix .gitattributes and .gitignore.	3 years ago
 .gitignore	Add test_data/scripts	3 months ago
 CONTRIBUTING.md	Create CONTRIBUTING.md	3 years ago
 LICENSE	Initial commit	3 years ago
 README.md	Working on workflows	3 months ago

README License

GNU GENERAL PUBLIC LICENSE

Version 3, 29 June 2007

Copyright (C) 2007 Free Software Foundation, Inc. <<https://fsf.org/>>

Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed.

Preamble

The GNU General Public License is a free, copyleft license for software and other kinds of works.

LinearDesign

Public

Watch 8

main 1 Branch 1 Tag

Go to file

Add file

Code

lianghuang3 Update README.md f0126ca · 5 months ago 25 Commits

pic	logo	last year
src	.so	last year
.gitattributes	init	last year
Makefile	.so	last year
README.md	Update README.md	5 months ago
coding_wheel.txt	init	last year
codon_usage_freq_table_human.csv	init	last year
codon_usage_freq_table_yeast.csv	init	last year
gflags.py	init	last year
license.txt	Update license.txt	last year
lineardesign	init	last year
testseq	init	last year


README License

The LinearDesign code is freely accessible to all interested parties.  
It is free for academic, non-profit, and research use, and can be licensed for commercial use.

To use this software for the development of a commercial product, including but not limited to software, service, or pharmaceuticals, please contact the lead corresponding author.

Redistribution of the code with or without modification is not permitted without explicit written permission by the lead corresponding author.



 viztein

Public

Watch 1

master

39 Branches

Tags

Go to file

Add file

<> Code

thetechnocrat-dev



Update README.md

2753013 · 4 years ago

35 Commits

example	add package lock	6 years ago
src	commit	6 years ago
.babelrc	init create-react-library@2.6.7	6 years ago
.editorconfig	init create-react-library@2.6.7	6 years ago
.eslintrc	init create-react-library@2.6.7	6 years ago
.gitignore	commit	6 years ago
CNAME	Update CNAME	6 years ago
README.md	Update README.md	4 years ago
package-lock.json	commit	6 years ago
package.json	switch to new deploy and example site	6 years ago
rollup.config.js	init create-react-library@2.6.7	6 years ago

README

# Viztein


[Viztein](#) is a React library component for visualizing **proteins** and other macromolecules. The code is a light wrapper around [NGL](#) which uses [WebGL](#) for rendering graphics.

npm package

0.1.8

License

MIT

 **fornac** Public

Watch 6

master


2 Branches

5 Tags

Go to file

Add file

Code

 **bad-ants-fleet** reset d3 dependency to 3.5.13


36df3c5 · 2 years ago


767 Commits

dist	Update dist/ directory	2 years ago
doc/img	Added cofolded sequences to the documentation	8 years ago
examples	Fix transition plot and other examples	5 years ago
src	Changes behaviour on transition if structure does not cha...	5 years ago
test	Collect all the code from different repos into this project i...	5 years ago
.babelrc	update dev and build tools	5 years ago
.editorconfig	make linting a separate task	5 years ago
.eslintrc.json	make linting a separate task	5 years ago
.gitignore	update dev and build tools	5 years ago
.gitmodules	cleaned up, generated makefile and fixed several issues	9 years ago
LICENSE	Create LICENSE	9 years ago
README.md	API change to better understandable terms and better ex...	5 years ago
package.json	reset d3 dependency to 3.5.13	2 years ago
webpack.config.js	Update to version 1.2.0 and upgrade style-loader version...	2 years ago


README

Apache-2.0 license

 An official website of the United States government [Here's how you know](#) ▼

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

Log in

 Some portions of our website are experiencing intermittent outages, and E-utilities requests may take longer than normal to complete. We apologize for the inconvenience and thank you for your patience as we resolve the issue.

Nucleotide

Nucleotide ▼

Search

Advanced

Help

GenBank ▼

Send to: ▼

Change region shown ▼

Customize view ▼

Analyze this sequence ▲

Run BLAST

Pick Primers

Highlight Sequence Features

Find in this Sequence

NCBI Virus ▲


Retrieve, view, and download SARS-CoV-2 coronavirus genomic and protein sequences.

Related information ▲

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; Coronaviridae; Orthocoronavirinae; Betacoronavirus; Sarbecovirus; Severe acute respiratory syndrome-related coronavirus.

RCSB PDB

Deposit

Search

Visualize

Analyze

Download

Learn

About

Careers

COVID-19

Help

Contact us

MyPDB

RCSB PDB

PROTEIN DATA BANK

226,707

Structures from the PDB

1,068,577

Computed Structure Models (CSM)

Enter search term(s), Entry ID(s), or sequence

Include CSM

Q

Advanced Search | Browse Annotations

Help

PDB-101

wwPDB

EMDataResource

NAKB

wwPDB Foundation

PDB-Dev

f t v o in

Access Computed Structure Models (CSMs) of available model organisms

Learn more

Welcome

Deposit

Search

Visualize

Analyze

Download

Learn

RCSB Protein Data Bank (RCSB PDB) enables breakthroughs in science and education by providing access and tools for exploration, visualization, and analysis of:

Experimentally-determined 3D structures from the **Protein Data Bank (PDB)** archive

Computed Structure Models (CSM)

from AlphaFold DB and ModelArchive

These data can be explored in context of external annotations providing a structural view of biology.

Explore NEW Features

PDB

PROTEIN DATA BANK

NEW!

Q

P

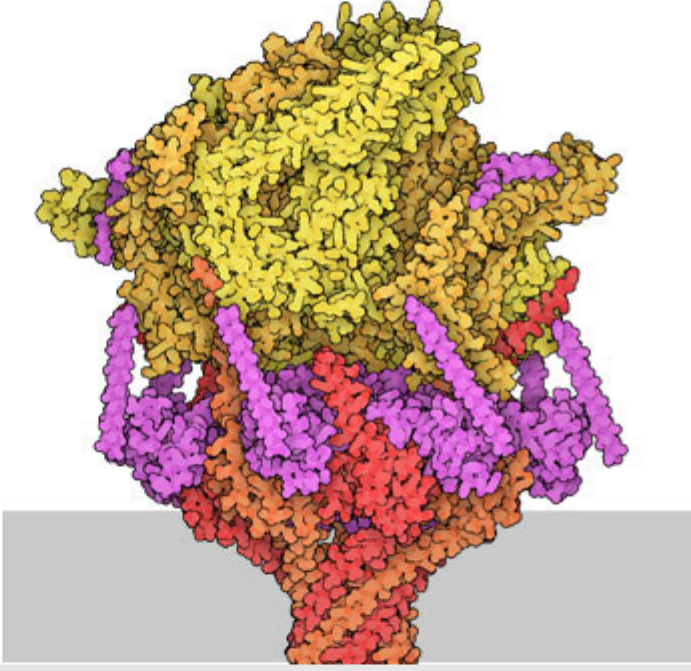
C

D

PDB-101

Training Resources

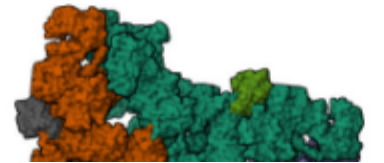
November Molecule of the Month




Malaria Parasite PTEX

Latest Entries

As of Tue Oct 29 2024




Features & Highlights



Deprecation of FTP File Download Protocol on November 1, 2024  
wwPDB to deprecate FTP download protocol in the PDB archive

News

Publications



Meet RCSB PDB at SACNAS  
Learn about summer internship opportunities and more at the Rutgers booth #648  
» 10/27/2024

VirusDecode

NC\_045512.2

DONE

Variant Sequence

Upload File

+

Drag your FASTA files here

Paste Sequence

Sequence1

ATGTTTGTTTTCTTGTATTGCCACTAGTACC.....

+ Add Sequence

① Enter a valid NCBI reference sequence ID, like NC\_045512.2.

② Once a valid NCBI reference sequence ID is entered, the following details will appear:

nc\_045512.2

DONE

Metadata

Sequence ID: NC\_045512.2

Name: NC\_045512

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

Length: 29903

③ Enter the sequence for analysis in FASTA format or as an ATCG sequence.

④ You can edit or delete the name of the input sequence.

⑤ After processing the reference and input sequences, the 'Next' button will be enabled, allowing you to proceed to the next analysis step.

Next →

You can hide the left panel or restart to manage the history.

You can restart the session.



[illegible]

VirusDecode

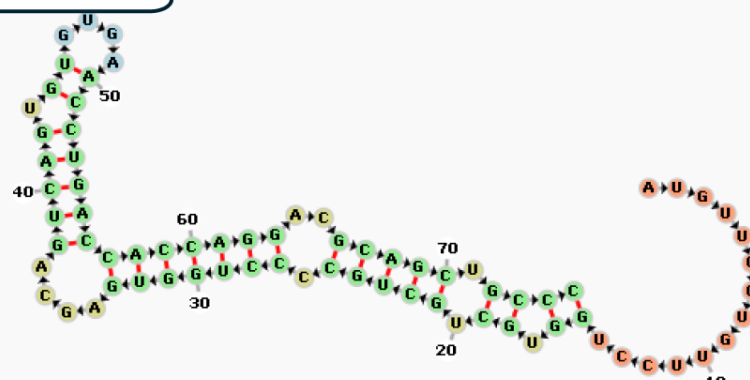
AlignmentmRNA design3D viewer

The detailed description of mRNA conversion can be found [here.](#)

mRNA Parameters

mRNA Localization

① Review key mRNA parameters like sequence, secondary structure, folding energy, and CAI (Codon Adaptation Index).



Amino Acid Sequence

MEFVLVLLPLVSSQCVNLTTRTQLP ...[show](#)

mRNA Sequence

AUGUUCGUGUUCUGGUGCUGCGCCUGGUGAGCAGUCAUGUGUGAA ...[show](#)

mRNA Structure

.....((((((((((((((((((((((((((((((((.....) show

mRNA folding free energy

-32.40 kcal/mol

mRNA CAI

0.930

③ Click the 'show' button to display the full sequence or structure.

② You can also check protein parameters for the sequence.

Protein Parameters

Molecular Weight

2820.4148000000005 Da

Isoelectric Point(PI)

7.999833488464356

	Amino Acid Count	Amino Acid Percent (%)
A	0	0.00
C	1	4.00
D	0	0.00

VirusDecode

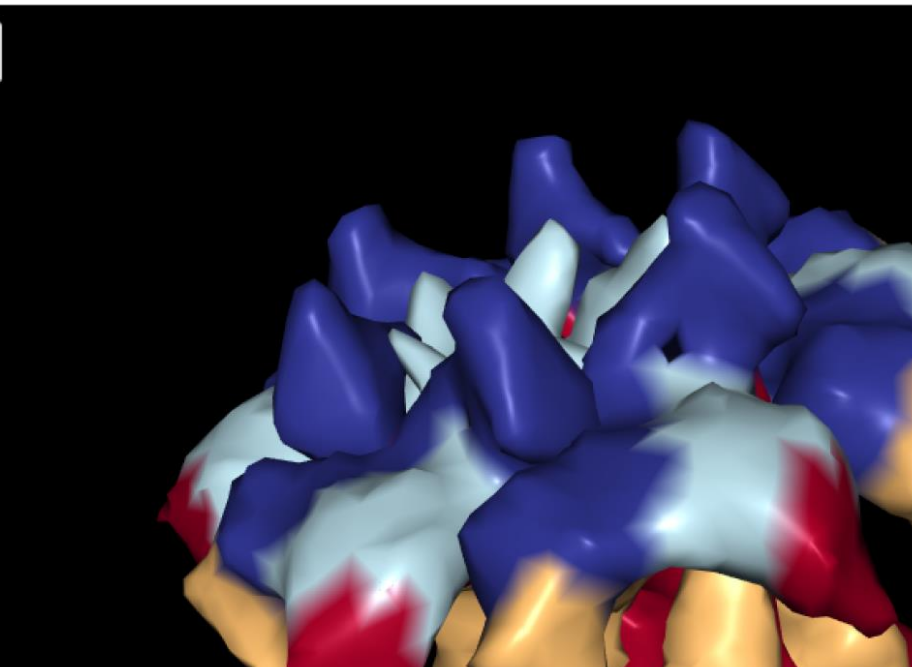
Alignment

mRNA design

3D viewer

Default ▾

① Choose the 3D protein image representation, such as Cartoon or Ball format.



② Interact with the 3D image by scrolling or moving the mouse.

8YAX.pdb

PDB ID

8YAX

8YB5

8YB7

707Y

707Z

7080

7081

6M71

6XEZ

6XQB

③ Select a PDB ID to view the 3D structure of the protein associated with the input sequence.