

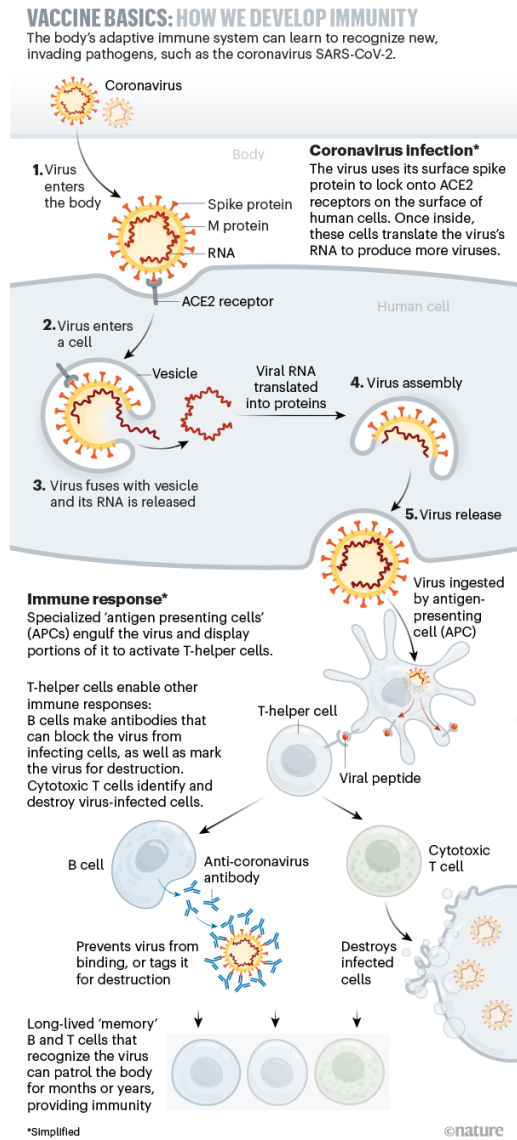
# Sequence Analysis of Pfizer and Moderna COVID-19 Vaccines



Benjamin King (benjamin.l.king@maine.edu)  
Department of Molecular and Biomedical Sciences

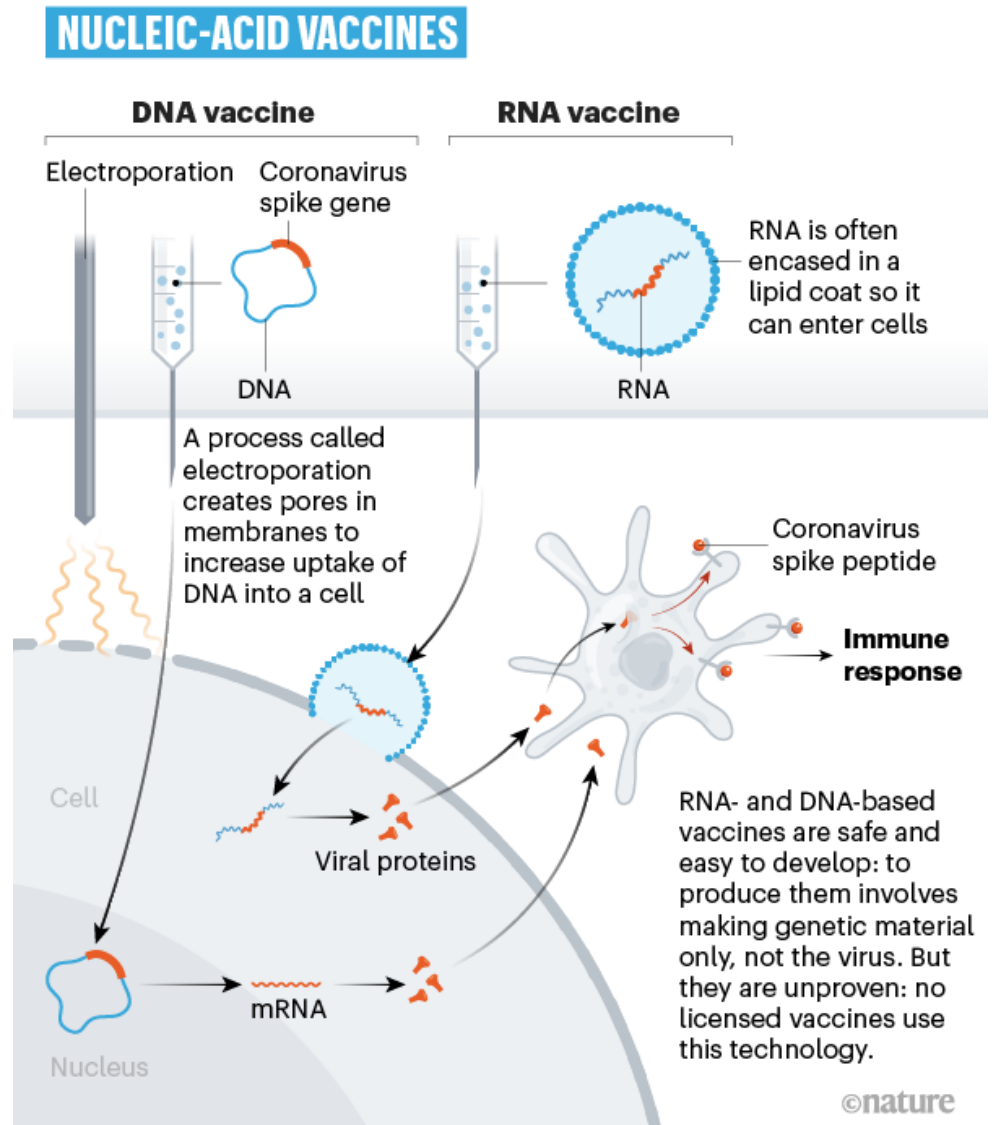


# SARS-CoV-2 Infection and Immune Response



<https://www.nature.com/articles/d41586-020-01221-y>

# RNA COVID-19 Vaccines Rely on Cell to Express Spike Protein



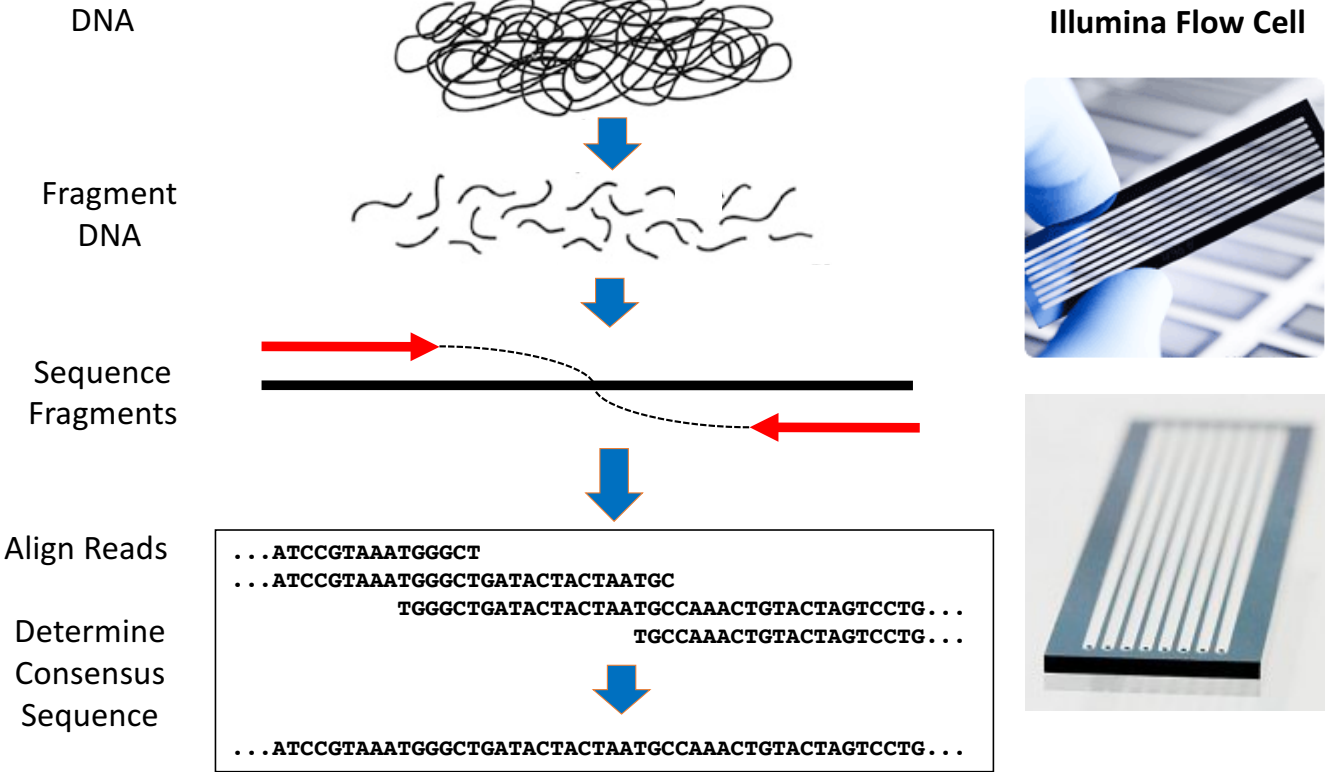
<https://www.nature.com/articles/d41586-020-01221-y>

# The \$1000 Genome Is Revolutionizing How We Study Biology and Practice Medicine

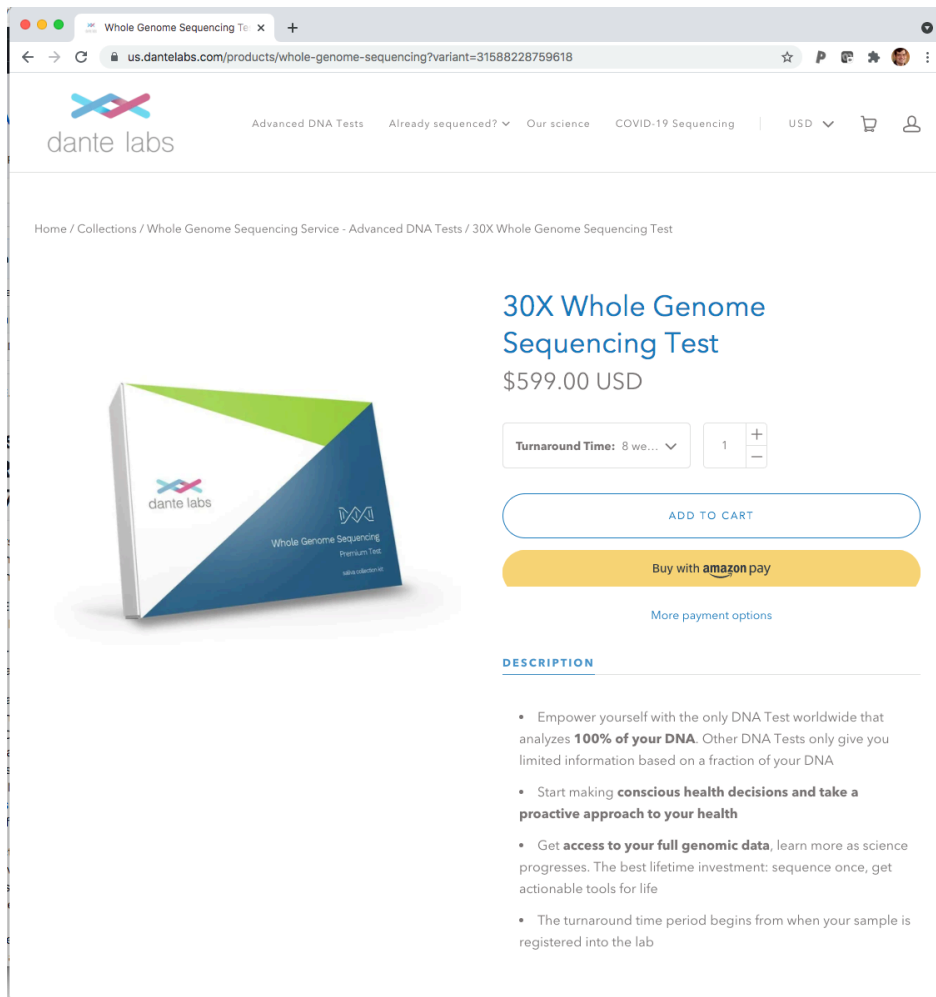
\$2.7B Over Decade → \$1000 in ~3 Days



# Whole Genome Shotgun Sequencing



# Your Own Genome Sequenced (30x Coverage)



The screenshot shows the Dante Labs website product page for the 30X Whole Genome Sequencing Test. The page includes the Dante Labs logo, navigation links, and a breadcrumb trail: Home / Collections / Whole Genome Sequencing Service - Advanced DNA Tests / 30X Whole Genome Sequencing Test. The product is displayed with a 3D rendering of the test kit box, which is white with green and blue accents and features the Dante Labs logo and the text "Whole Genome Sequencing Premium Test". The price is listed as \$599.00 USD. Below the price, there is a dropdown menu for "Turnaround Time: 8 we..." and a quantity selector set to "1". There are three buttons: "ADD TO CART", "Buy with amazon pay", and "More payment options". A "DESCRIPTION" section follows, containing four bullet points:

- Empower yourself with the only DNA Test worldwide that analyzes **100% of your DNA**. Other DNA Tests only give you limited information based on a fraction of your DNA
- Start making **conscious health decisions and take a proactive approach to your health**
- Get **access to your full genomic data**, learn more as science progresses. The best lifetime investment: sequence once, get actionable tools for life
- The turnaround time period begins from when your sample is registered into the lab

## Black Friday Sale (Nov. 25, 2020)



The banner features a red background with a white torn-paper effect. On the left and right sides, there are images of laboratory equipment, including large white machines on a counter. In the center, the Dante Labs logo is displayed above the text "dante labs". Below this, the text "Whole Genome" is written in white, followed by a large "\$149" in white. The overall design is clean and professional, emphasizing the sale price.

## ILLUMINA MiSeq in Hitchner Hall at UMaine



# What are the sequences of the RNA in the Pfizer and Moderna COVID-19 Vaccines?

The screenshot shows a GitHub repository page for the project "Assemblies-of-putative-SARS-CoV2-spike-encoding-mRNA-sequences-for-vaccines-BNT-162b2-and-mRNA-1273" by NAalytics. The repository has 98 watchers, 3.4k stars, and 529 forks. The main branch is "main" with 1 branch and 0 tags. The file list includes:

File Name	File Type	Last Modified
NAalytics Update README.md		21fac9b on Apr 14 23 commits
Assemblies of putative SARS-CoV2...	Add files via upload	4 months ago
Figure1Figure2_032321.fasta	FASTA	5 months ago
README.md	Update README.md	4 months ago

The "About" section contains the following text:

RNA vaccines have become a key tool in moving forward through the challenges raised both in the current pandemic and in numerous other public health and medical challenges. With the rollout of vaccines for COVID-19, these synthetic mRNAs have become broadly distributed RNA species in numerous human populations. Despite their ubiquity, sequences ...

The "Releases" section shows "No releases published" and the "Packages" section shows "No packages published".

<https://github.com/NAalytics/Assemblies-of-putative-SARS-CoV2-spike-encoding-mRNA-sequences-for-vaccines-BNT-162b2-and-mRNA-1273>



# Pfizer Sequence

```
1 >Figure1_032321_Spike-encoding_contig_assembled_from_BioNTech/Pfizer_BNT-162b2_vaccine
2 GAGAAATAAGTATTCTTCTGGTCCCAACAGACTCAGAGAGAACCCGACCACTGTTCTGGTCTGGTCTGCTGCC
3 TCTGGTGTCCAGCCAGTGTGAACCTGACCAACAGCAACAGCTGCTCCAGCTACACCAACAGCTTACAGAGCGC
4 TGACTACCCCAAGGTTTCAGATCCAGCGTGTGACTCTACCCAGGACCTGTTCTGCTCTTCTTCAGCAACGTG
5 ACCTGTTCCAGCCATCACGTGTCGGCCACCAATGGCAACAGAGATTGACAAACCCCTGCTGCTTCAACGACGG
6 GGTGACTTTGCCAGCAGCGAAGTCCAACATCATCAGAGGCTGGATCTTCCGCCACCACTGGACAGCAAGACCCAGA
7 GCCTGTGATCGTGAACACGCCCAACGTTGTCATCAAGTGTGGAGTTCAGTTCGCAACGACCCCTTCTGGGGC
8 GTCTACTACCAAGAACCAACAGAGCTGGATGAAAGCGAGTTCGGGTGTACAGCAGCGCCAAACACTGCACCTCGA
9 GTACTGTGCCAGCTTCTCTGATGGACTGGAAGGCAAGCAGGCAACTTCAAGAACCTGCGGAGTTCTGTTTTAAGA
10 ACATGACGGCTACTTCAAGATCTACAGCAAGCACCCTTCAACCTCGTGGGGATCTGCTCAGGGCTTCTCTGCT
11 CTGGAACCTCTGGTGGATCTGCCATCGGACTCAACATCACCGGTTTACAGACTGCTGGCTCTGACAGAACTACT
12 GACACTGGCGATAGCAGCAGCGGATGGACAGTGGTGGCCGCTACTACTGGTGGCTACTGCGACTGAGAACTTCC
13 TGTGAAGTCAACAGAACCGCACTCACCGACCCGTGGATTGTGCTCTGGATCTCTGAGCAGCAAAAGTGACC
14 CTGAAGTCTTACCGTGGAAAAGGACTCTACAGCAGCAAGCACTTCCGGGTGAGCCACCGAATCATCTGCGGTT
15 CCCAATATCACCAATCTGTGCCCTTGGGAGGTGTTCAATGACCACGATTGCTCTGTGTAGCCCTGAAACGGG
16 AGCGGATCAGCAATTCGGTGGCGACTACTCGTGTGTAACACTCCGCGACTTACAGCACTTCAAGTCTACGGCGT
17 TCCCTCAAGCTGAACGACTGTGCTTCAAAAGCTGTACGGCAGCAGCTGTGATCTCGGGAGATGAAGTGGCGGA
18 GATGGCCCTGGACAGCAGGCAAGATCGGCACTACAACCTACAAGCTGCGCAGCACTTCCGGCTGTGATGTTGCT
19 GGACAGCAACCACTGGACTCAAGTCCGGCGCACTCAACTACTGCTACCGGCTGTCTCGGAAGTCAACTGGA
20 CCTCTGAGCGGGACATCTCACAGAGATCTACGGCCAGCAGCCCTTGAACGGCTGGAAGCTTCACTGTA
21 CTTCCCACTGACTCTACGGCTTCAAGCCAAATGGCTGGCTATGACCTACAGAGTGTGGTCTGACTGCTG
22 AACTGCTGATGCTCCCTGACAGAGTGGGGCTTGAAGAACCAACTCTGTAAGAACAAATGGTGAACCTTCAAC
23 TTTCAACGGCTGACCGGCAACCGGCTGTGACAGAGAGCAACAGAAAGTTCTGCGATTCCAGCAATTTGGCGGAT
24 CGCGATACACAGAGCGCTTGAAGTCCCAAGCACTCGAAATCTGGACATCACCTTGCAGCTTCCGGGAGTGT
25 CTGATACCCCTGGACCAACAGCAAGCAATCAGGTGGAGTGTGACAGGACGTGAACGTGACCGAAGTGGCGT
26 GCAATCAGCGGATCAGCTGACACTACATGGGGGTGACTCAACCGGCAAGTGTGTTTACAGCAGAGCGGCTG
27 TCTGATCGGAGCGAGCAGTGAACAAATAGCTACGAGTGGACATCCCACTCGGCGTGGAAATCGCGCAGCTAC
28 CAGACAAACAGCCCTCGGAGCGCAGAGCGTGGCCAGCAGAGCATATTGCTACAAATGTCTTGGCCGCGAG
29 AACAGCTGGCTACTCCAACACTATCGCTATCCCAACCACTCACATCAGCGTGCACCAAGATCTCGCTGCT
30 GCTCAAGCAAGACAGCGTGGACTGACCATGTACATCTGGCGGATTCCACGAGTGTCCAACCTGCTGTGAGT
31 ACGGAGCTTCTGCAACCGAGTGAATGAGGCTGACAGGATGGCTGGCAAGGACAGCAAGAACCCAAAGAGTGT
32 GCCCAAGTGAAGCAGATCTACAAGACCCCTCTAACAAGACTTCCGGCGCTTCAATTCAGCCAGATTTCTGCCG
33 TAGCAAGCCAGCAAGCGGAGCTTCACTGAGGACTGCTGTTCAACAAGTGAACACTGGCCAGCGGCTTCAACG
34 AGTATGGCGATTGTCTGGCGGACTTGGCCAGGGATCTGATTTGGCCAGAAGTTTAAAGGACTGACAGTCTGCT
35 CCTCTGCTGACCGATGAGATGATCGCCAGTACACATCTGCTGCTGGCCGCAACATCAAGCGGCTGGCAATTGG
36 ACGAGCGCCGCTCTGAGATCCCTTTGCTATGAGATGGCTACCGGTTCAACGGCATCGAGTGCACCAAGATGTC
37 TGTACAGAACAGAACTGATGCCAACAGTTCACAGCCATCGGCAAGTCCAGGACAGCTGACAGCAGCAGCA
38 AGCGCCCTGGAAAGTGCAGGAGTGTGCAACAGAAAGTCCAGGACACTGAACACCTGGTCAAGCAGTCTGCTCCA
39 TTTGGCGCATCAGCTCTGCTGAAAGATATCTGAGCAGACTGGACCTCTGAGGCGGAGGTGCAGATCGACAGAC
40 TGTACAGCGCAGACTGCAGAGCTCCAGACATACGTGACCCAGCAGCTGACAGAGCGCGGAGATTAGAGCTTCC
41 AATCTGGCCGCAACAGATGCTGAGTGTGTTGGGCGAGAGCAAGAGTGGACTTTGGCGCAAGGCTTACCACT
42 GATGAGCTTCTCAGTCTGCCCTCACGGCTGGTGTCTTCTGACGTGACATATGTCGCCCTCAAGGAAGAAATTTCA
43 CACCGCTCCAGCATCTGCCAGCAGCGCAAGGCACTTCTAGAGAAGCGTGTGCTGTTCAACGGCAACCTATTTG
44 TCTGACAGCGGAACTTCTACAGGCTCAGATCATCAACAGCAGCAACACTCTGCTGCTGGCACTGCGAGTCTG
45 GATGGCAATTTGAACAACCTGCTAGACCTCTGACCGCCAGCTGGACAGCTTCAAAGGAACTGGAAAGTACT
46 TTAAGAACACCAAGCCCGGAGCTGGACTGGCGGATACAGCGAATCAATGCGAGCTGCTGAACATCCAGAAAG
47 ATCCAGCGGCTGACAGGCTGGCAAGAATCTGAACGAGAGCTGACTGACCTGCTCAAGAACTGGGAAAGTACGAGCAG
48 CATCAAGTGGCCCTGTACATCTGGCTGGCTTATCGCCGACTGATTGCCATGATGTTGCAACATCATGCTGTT
49 GCATGACAGCTGCTGATGCTGCTGAAGGCTGTTGATGCTGTCAGCTGCTGCAAGTTCGACGAGGAACTTCTG
50 CCGCTGCTGAAGGCGTGAACCTGCACTACACATGATGACTGAGCTGGTACTGATGACGCAAGTCTAGTGGCCCT
51 TCCGCTCTGGGCTACCAGTCTCCCGACTCGGGTCCAGGATGCTCCCACTCCACTGCCCCTACCACTCA
52 CTGACTTCCAGACCTCCCAAGCAGCGCAATGAGCTCAAAAGCTTAGCTAGCACACCTCCAGCGGAAACAG
53 CAGTAACTTACCTTAGCAATAAAGCAAGTTTAACTAAGCTATAACCCAGGGTGGTCAATTTCTGCTGACGAC
54 ACCCTGGAGTACGA
```

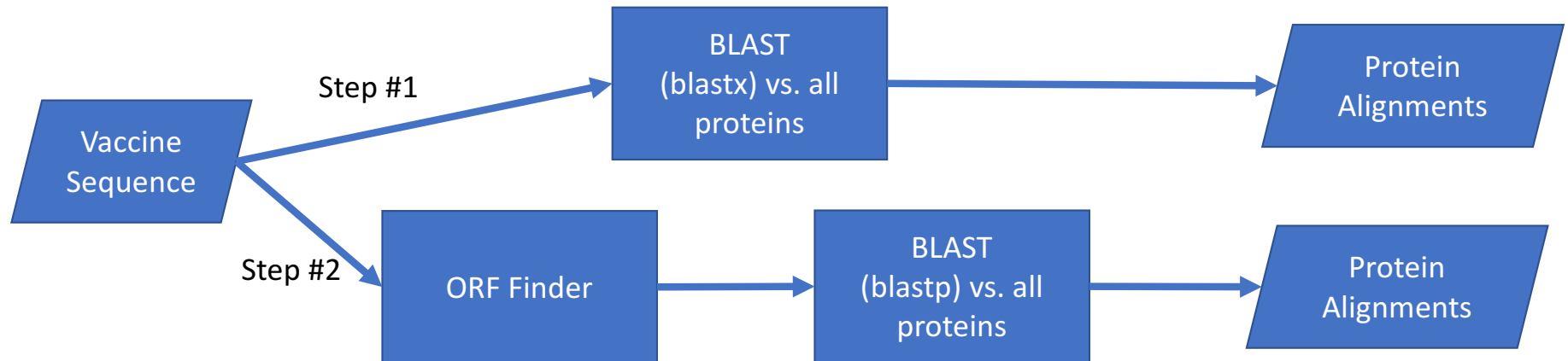
# Moderna Sequence

```
55 >Figure_2_32321_Spike-encoding_contig_assembled_from_Moderna_mRNA-1273_vaccine
56 GGGAAATAAGAGAGAAAGAGTAAGAAATAAAGACCCGGCGCGCCACCTGCTGTGCTGGTCTGCTGCC
57 GCCCTGTGGAGCAGCGAGTGAACCTGACCCCGGACCCAGCTGCCACCGCACTCAACAGCTTACCCCGGG
58 GGGTCTACTACCCGCAAGTGTTCGGAGCAGGCTGTCGACAGCAGCCAGGACCTGTTCTGCTCTTTCAGCAAC
59 GTGACTGGTCCAGCCATCACGTGAGCGGCCAACCGCACCAAGCGGTTGACAAACCCCTGCTGCTTCAACG
60 CGCGGTGACTTCCGACAGCAGAGAGCAACATCATCCGGGGCTGGATCTTCCGACCCACCTGGACAGCAAGACCC
61 AGAGCTGCTGATGGTAAACGCCAACCTGGTGTGATCAAGTGTGGAGTCCAGTTCGCAACGACCCCTTCTG
62 GCGTGTACTACCAAGAACCAAGAGCTGGATGGAGGAGGAGTTCGGGTGTACAGCAGCGCCAAACACTGCACCT
63 CGATACGTGAGCAGCGCTTCTGATGGACTGGAGGCAAGCAGGCAACTTCAAGAACTGCGGGAGTCTGTGTTCA
64 AGAACACTGACGGCTACTCAAGATCTACAGCAAGCACCACTCAACCTGTTGGGGATCTGCCAGGGCTTCTCA
65 GCCTGGAGCCCTGTGGACTGCCATCGGATCAACATCACCGGTTCCAGACCTGCTGGCCCTGACCGGAGCTA
66 CTGACCCAGGCGACAGCAGCGGGTGGACAGCAGGCGGCTGTTACTACGTGGCTACTCGACGCCCGGACT
67 TCTGCTGAAGTCAACAGAGACCGCACTCACCGACCGCTGGACTGCCCTGGACCTCTGAGCAGCAAGTGT
68 ACCCTGAAGAGTTCACCTGGAGAGGGCTACTACAGCAGCAGCACTTCCGGGTGAGCCACAGAGAGATCTGG
69 GTTCCCAACATCACCACTGCTCCCTTCCGGGAGGTTCAACGCCACCCGGTTCGCGAGGCTGACTGCGAAC
70 GGAAGCGGATCAGCACTGCTGGCCGACTACAGGCTGTGTAACAACGCGCCAGCTTCAACACTTCAAGTCTAC
71 GTGAGCCACCAAGCTGAACGCTGTGCTTACCAAGCTGTACCGCCAGCAGCTGTGATCGTGGCAGCAGAGTGG
72 GCGAGCTGACCCCGGACAGCAAGTGGCAGCTACACTACAAGCTGCGCCAGGATTCACCGGCTGCTGATG
73 CTGGAAAGCAGCAACACTCGACAGCAAGTGGCGGCACTCAACTACTCTGACCGGCTTCCGAAAGCAACCTG
74 AAGCCCTGAGCGGAGTACAGCAGCAGTACTCAAGCGGCTCCACCTTGAACGGCTGGAGGCTTCAACTG
75 CTACTCTCTGAGAGTACGGCTTCCAGCCCAACGCGGCTGGCTACAGCCCTACCGGCTGGTGTGAGCT
76 TGAGCTGCTGACCGCCAGCAGCGTGTGGCCCAAGAGAGCAACACTGCTGTAAGAACAAAGTGGGAACTT
77 AACTTCAAGGCTTACCGGCAACCGGCTGTGACAGAGCAACAGAAATTTCTGCCCTTCAAGACTTCCGCGG
78 CATCGCCAGCACCAGCAGCTGTGGGATCCCAAGCCTGGAGATCTGGACATCACCTTTCAGCTTCCGGCGG
79 TGAGGCTGATCACCAGGCAACCAACAGCAGCAACAGTGGCGGCTGTGACAGGAGTGAATGACAGGAGTGGCC
80 TTGCGCTACCAAGCGCAGCTGACACCACTGCGGGTGTACAGCAGCAGCGGCAAGTGTTCAGACCCGGCG
81 TTGCGTGTGCGCGGAGCAGTGAACAGAGTACAGTGGCAGTCCCATGGCCGCGGATCTGTGCACTGAC
82 AGACAGCAGCAATTCACCGGAGGCAAGGAGCTGGCCAGCAGAGCATCTCGCTACACTGAGCCTGGGCGCC
83 GAGAACAGGCTGGCTCAGCAACAGCAGTCCGCTCCCAACCACTTCACTACAGCTGACACCGGAGATCTG
84 CGTAGCATGACCAAGCAGCGGACTGACATGTACATCTGCGCGACAGCAGGAGTGAAGCAACTGCTGTGCT
85 AGTACGGCAGCTTCTGCAACAGCTGAACCGGCTGACCGGCTGCGCGTGGAGCAGCAAGAACCCAGGAGGTT
86 TTTGCCAGGTGAAGCAGTCTACAAGACCCCTCCATCAAGGACTTCCGGCGCTTCACTTCAAGCAGATCTGCC
87 CCCAGCAAGCCAGCAAGCGGAGCTTACAGGAGCTGCTGTTCAACAAGTGAACCTAGCAGCAGCGGCTTAC
88 AGCAGTACGGCGACTGCTCGGAGATAGCGCCGGGACTGATCTGCGCCAGAAGTTCAACGGCTGACCGTGTG
89 CCTCTGCTGACAGCAGATGATCGCCAGTACACATCTGCTGCTGGCCGCAACATCAAGCGGCTGGCAATTGG
90 CCTCTGCTGACAGCAGATGATCGCCAGTACACCGGCTTACAGCAGCAGCTGTTAGCAGCACTCAACGGCG
91 CGGCTGGAGCGCTCTGAGATCCCTTCCGATGAGATGGCTACCGGTTCAACGGCTGCGGCTGACCCAGAAC
92 TGCTGACAGAACAGAAAGTGTGCGCAACAGTTCACAGCGCATCGGCAAGTCCAGGACAGCTGAGCAGCAGC
93 GCTAGCGCCCTGGCAAGCTGACAGAGCTGGTGAACAGAACCCAGGCTTGAACACTGTTGAAGCAGCTGACAG
94 CAACTTGGCGCATCAGCAGGCTGCTGAAGCAGCTCTGAGCGGCTGGACCTCCCGAGGCGGAGTGCAGATG
95 GCGTACTGCTGGCGGCTGACAGCCTGACAGCTACGTGACCCAGCAGCTGATCGGGCCCGGAGATTCGGG
96 CCGCACTGGCCGCAAGTGAAGTGGCTGCTGGCCAGAGCAAGCGGTTGACTTCCGGAGGAGGCTACCA
97 CCTGATGAGCTTCCCAAGAGCGCAACCCAGGAGTGGTCTCTGCACTGACTGCTGCCCGCAGCAAGAACT
98 TCAACCGCCCGGCACTCTGCGACAGCGCAAGGCCACTTCCCGGGAGGCGTGTGTTGAGCAAGCAGCCAC
99 TGGTGTGAGCCAGGGAATCTTACAGCGCCAGATCATCACCAAGCAACACTTCTGAGCGGCAAGTGCAGCT
100 GGTGATCGGATGTGAACAACCGTGTGACTTCCCTGAGCGCCAGCTGCAAGCTTCAAGAGAGGCTGGACAG
101 ACTCAAGAATCACAGCCGCGACTGACCTGGCGGACTCAGCGGCTCAAGCGAGCTGGTGAACATCAAGAA
102 GAGATGATCGGCTGACAGGCTGGCAAGCACTGACAGGAGCTGATGACTGACAGCTGGCAGAGCTGGCAG
103 GTACATCAAGTGGCTGATCATCTGGCTGGCTTATCGCGGCTGATGCTGATGATGATGATGATGATGATGAT
104 GCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
105 GCGCGTGTGAAGGCGTGAACCTGCACTACACATGATGACTGAGCTGGTACTGATGACGCAAGTCTAGTGG
106 GCGCTTCCGCGTGGGCTACCAGTCTCCCGACTCGGGTCCAGGATGCTCCCACTCCACTGCCCCTACCACT
107 CTGACTTCCAGACCTCCCAAGCAGCGCAATGAGCTCAAAAGCTTAGCTAGCACACCTCCAGCGGAAACAG
108 CAGTAACTTACCTTAGCAATAAAGCAAGTTTAACTAAGCTATAACCCAGGGTGGTCAATTTCTGCTGACGAC
109 ACCCTGGAGTACGA
```

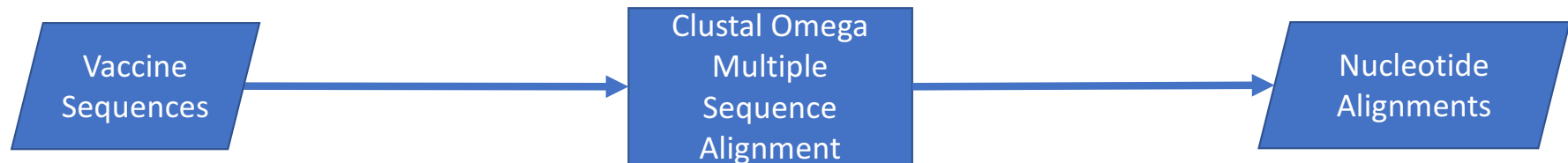


# Sequence Analysis Workflow

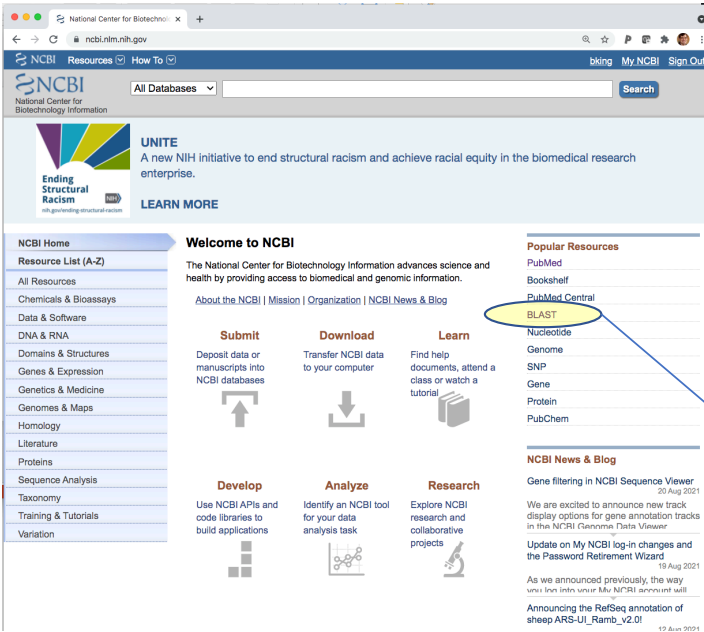
What viral protein(s) could be encoded by the sequences?



How similar are the sequences?

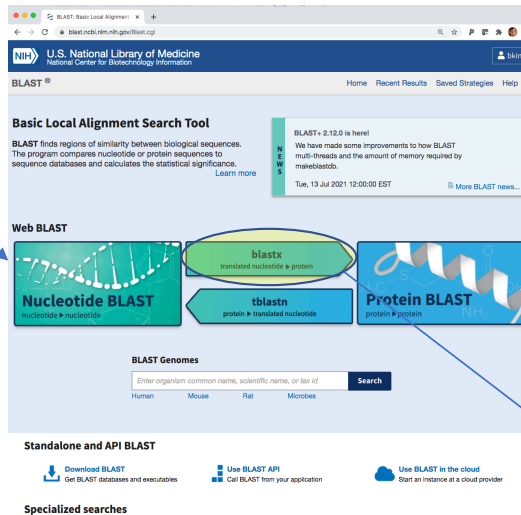


# BLASTX vs. All Proteins @ NCBI



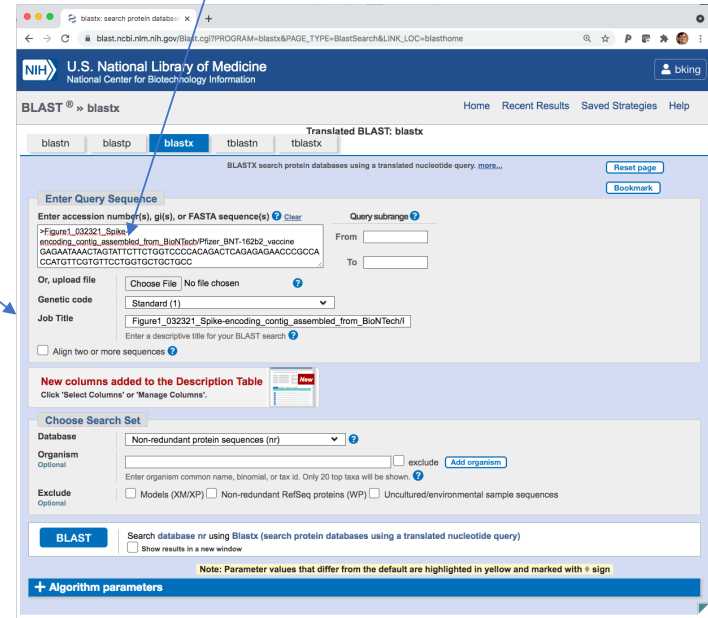
The screenshot shows the NCBI homepage with a navigation menu on the left. The 'Proteins' link is highlighted in red. The main content area features a 'Welcome to NCBI' message and several action buttons: 'Submit', 'Download', 'Learn', 'Develop', 'Analyze', and 'Research'. A yellow circle highlights the 'BLAST' link in the 'Popular Resources' section.

<https://www.ncbi.nlm.nih.gov/>



The screenshot shows the BLAST Basic Local Alignment Search Tool interface. The 'Web BLAST' section is highlighted with a yellow box. Below it, four options are shown: 'Nucleotide BLAST', 'blastx', 'tblastn', and 'Protein BLAST'. The 'blastx' option is highlighted with a yellow circle and a blue arrow pointing to the 'blastx translated nucleotide > protein' label. Below the options is a search box for 'BLAST Genomes' with a 'Search' button.

Paste the Pfizer sequence here



The screenshot shows the BLASTX search results page. The 'blastx' option is selected in the 'Translated BLAST' section. The 'Enter Query Sequence' section contains the Pfizer sequence: `>Figure1_032321_Spike-encoding_contig_assembled_from_BioNTech/Pfizer_BNT-162b2_vaccine  
GAGATATAAAGTATATCTCTCTGGTCCGACAGAGCTGAGAGAGAAKCCCGCCA  
CCATGTCGTCTCTCGTGGCTGCTGCC`. The 'Database' is set to 'Non-redundant protein sequences (nr)'. The 'BLAST' button is highlighted in blue.

Paste the Pfizer sequence here

# ORF Finder @ NCBI

ORF Finder

Open Reading Frame Finder

Enter Query Sequence

Choose Search Parameters

ORFs found: 20

ORF	Start	Stop	Length (aa)
ORF1	1	1273	1273
ORF10	1	3764	3381
ORF11	1	437	165
ORF14	1	1814	1578
ORF15	1	2972	2763
ORF16	1	2585	2418
ORF7	2	3896	4054
ORF19	2	3253	3107

ORF1 (1273 aa)

BLAST

Select "ORF1"

<https://www.ncbi.nlm.nih.gov/orffinder/>

BLAST® » blastp suite

Standard Protein BLAST

Enter Query Sequence

Choose Search Set

Database: Non-redundant protein sequences (nr)

BLAST

BLAST® » blastp suite » results for RID-JGCGERE016

Job Title: Protein Sequence

Filter Results

Description	Score	Score Expect	Query Cover	Ident	Len	Accession	
SARS-CoV-2 spike in infection state (Severe acute respiratory syndrome coronavirus 2)	2638	2638	100.0%	1273	62001_A		
S system of SARS-CoV-2 in complex bound with 4M8 (Severe acute respiratory syndrome coronavirus 2)	2637	2637	100.0%	1273	JC2L_A		
Chain A, Spike glycoprotein (Severe acute respiratory syndrome coronavirus 2)	2635	2635	100.0%	1283	70X1_A		
Spike glycoprotein (Severe acute respiratory syndrome coronavirus 2)	2633	2633	100.0%	99.84%	1282	BCN88353.1	
Distinct conformational states of SARS-CoV-2 spike protein (Severe acute respiratory syndrome coronavirus 2)	2632	2632	100.0%	99.84%	1310	6J95A.1	
Spike glycoprotein (Severe acute respiratory syndrome coronavirus 2)	2631	2631	100.0%	99.76%	1273	QJ25HE.1	
Structure of SARS-CoV-2 3D-2P in heparin-conjugated state (Severe acute respiratory syndrome coronavirus 2)	2631	2631	100.0%	99.76%	1273	J1L1_A	
Spike glycoprotein (Severe acute respiratory syndrome coronavirus 2)	2630	2630	100.0%	99.84%	1273	XP_050724390.1	
Spike glycoprotein (Severe acute respiratory syndrome coronavirus 2)	2630	2630	100.0%	99.76%	1273	QJ25HE.1	
Spike glycoprotein (Severe acute respiratory syndrome coronavirus 2)	2630	2630	100.0%	99.76%	1273	QJ25HE.1	
Spike glycoprotein (Severe acute respiratory syndrome coronavirus 2)	2630	2630	100.0%	99.76%	1273	QJ25HE.1	
Spike glycoprotein (Severe acute respiratory syndrome coronavirus 2)	2630	2630	100.0%	99.76%	1273	QJ25HE.1	

Run BLAST with "Non-redundant protein sequences (nr)"

Viewer - ORFfinder - NCBI x Protein BLAST: search protein x NCBI Blast:Protein Sequence x +

blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr\_6ZOW\_A

Download GenPept Graphics Next Previous Descriptions

### SARS-CoV-2 spike in prefusion state [Severe acute respiratory syndrome coronavirus 2]

Sequence ID: [6ZOW\\_A](#) Length: 1273 Number of Matches: 1  
[See 9 more title\(s\)](#) [See all Identical Proteins\(IPG\)](#)

Range 1: 1 to 1273 [GenPept](#) [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
2638 bits(6837)	0.0	Compositional matrix adjust.	1273/1273(100%)	1273/1273(100%)	0/1273(0%)
Query 1		MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS			60
Sbjct 1		MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS			60
Query 61		NVTWFHAIHVSQTNGTKRFDNPFVLPFNDGVYFASTEKSNIIIRGWIFGTTLDLSTQSLIV			120
Sbjct 61		NVTWFHAIHVSQTNGTKRFDNPFVLPFNDGVYFASTEKSNIIIRGWIFGTTLDLSTQSLIV			120
Query 121		NNATNVVIKVCDFQPCNDPFLGVYHKNKSWMESEFRVYSSANNCTFEYVSQPFLMDLE			180
Sbjct 121		NNATNVVIKVCDFQPCNDPFLGVYHKNKSWMESEFRVYSSANNCTFEYVSQPFLMDLE			180
Query 181		GKQGNFKNLREPVFKNIDGYFKIYSKHTPINLVRDLPGFSALEPLVDLPIGINITRFQT			240
Sbjct 181		GKQGNFKNLREPVFKNIDGYFKIYSKHTPINLVRDLPGFSALEPLVDLPIGINITRFQT			240
Query 241		LLALHRSYLTSGDSSSGWTAGAAAAYVGYLQPRTEFLKYNENGTITDAVDCALDPLSETK			300
Sbjct 241		LLALHRSYLTSGDSSSGWTAGAAAAYVGYLQPRTEFLKYNENGTITDAVDCALDPLSETK			300
Query 301		CTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPPGEVFNATRFASVYAWNRKRISN			360
Sbjct 301		CTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPPGEVFNATRFASVYAWNRKRISN			360
Query 361		CVADYSVLVNSASFSTFKCYGVSPTKLNDLCFTNVDYADSFVIRGDEVQRQIAPGQTGKIAD			420
Sbjct 361		CVADYSVLVNSASFSTFKCYGVSPTKLNDLCFTNVDYADSFVIRGDEVQRQIAPGQTGKIAD			420
Query 421		YNYKLPDDFTGCVIAWNSNLDKSVGGNYNYLRLFRKSNLKPFFERDISTEYIQAQSTPC			480
Sbjct 421		YNYKLPDDFTGCVIAWNSNLDKSVGGNYNYLRLFRKSNLKPFFERDISTEYIQAQSTPC			480
Query 481		NGVEGFNCYFPLQSYGFPQTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVN			540
Sbjct 481		NGVEGFNCYFPLQSYGFPQTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVN			540
Query 541		FNFNGLTGTGVLTSNKKFLPFQGFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITP			600
Sbjct 541		FNFNGLTGTGVLTSNKKFLPFQGFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITP			600
Query 601		GTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGNSVVFQTRAGCLIGAEHVNSY			660
Sbjct 601		GTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGNSVVFQTRAGCLIGAEHVNSY			660
Query 661		ECDIPIGAGICASYQTQTSNPRRARSVASQSI IAYTMSLGAENSVAYSNNNSIAIPTNFTI			720
Sbjct 661		ECDIPIGAGICASYQTQTSNPRRARSVASQSI IAYTMSLGAENSVAYSNNNSIAIPTNFTI			720
Query 721		SVTTEILFVSMTKTSVDCTMYICGDSSTECNSLLQYGSFCTQLNRALTGIAVEQDKNTQE			780
Sbjct 721		SVTTEILFVSMTKTSVDCTMYICGDSSTECNSLLQYGSFCTQLNRALTGIAVEQDKNTQE			780
Query 781		VFAQVKQIYKTPPIKDFGGFNFSQILPDPSPKSKRSFIEDLLFNKVTLADAGFIKQYGDC			840
Sbjct 781		VFAQVKQIYKTPPIKDFGGFNFSQILPDPSPKSKRSFIEDLLFNKVTLADAGFIKQYGDC			840

[Related Information](#)  
[Identical Proteins](#) - Identical proteins to 6ZOW\_A







# SARS-CoV-2 Genome Sequencing Update

- 3,035,314 million genomes worldwide
- 812,520 genomes from US
- 5,537 genomes from Maine
  - **583 are Variant of Concern(VOC) Delta**
  - 939 VOC Alpha (“UK Variant”)
  - 296 VOC Iota
  - 121 VOC Gamma
  - 108 VOC Zeta
  - 8 VOC Eta
  - 4 VOC Beta
  - 1 VOC Lambda
  - 0 VOC Kappa

VOC Alpha 202012/01 GRY (B.1.1.7) first detected in the UK

VOC Beta GH/501Y.V2 (B.1.351+B.1.351.2+B.1.351.3) first detected in South Africa

VOC Gamma GR/501Y.V3 (P.1+P.1.x) first detected in Brazil/Japan

VOC Delta G/478K.V1 (B.1.617.2+AY.x) first detected in India

VOI Zeta GR/484K.V2 (P.2) first detected in Brazil

VOI Eta G/484K.V3 (B.1.525) first detected in UK/Nigeria

VOI Iota GH/253G.V1 (B.1.526) first detected in USA/New York

VOI Kappa G/452R.V3 (B.1.617.1) first detected in India

VOI Lambda GR/452Q.V1 (C.37) first detected in Peru

Aug. 25, 2021



Registered Users EpiFlu™ EpiCoV™ My profile

EpiCoV™ Search Downloads Upload

Search

Accession ID  Virus name   complete  high coverage   
 Location  Host   low coverage excl  w/Patient status   
 Collection  to  Submission  to   collection date compl   
 Clade  Lineage  Substitutions  Variants

<input type="checkbox"/>	Virus name	Passage dt	Accession ID	Collection da	Submission C	<input type="checkbox"/>	Length	Host	Location	Originating
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5855/2021	Original	EPI_ISL_3544587	2021-08-12	2021-08-21	<input type="checkbox"/>	29,727	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5854/2021	Original	EPI_ISL_3544586	2021-08-12	2021-08-21	<input type="checkbox"/>	29,727	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5826/2021	Original	EPI_ISL_3544585	2021-08-12	2021-08-21	<input type="checkbox"/>	29,564	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5825/2021	Original	EPI_ISL_3544584	2021-08-12	2021-08-21	<input type="checkbox"/>	29,588	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5824/2021	Original	EPI_ISL_3544583	2021-08-12	2021-08-21	<input type="checkbox"/>	29,518	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5823/2021	Original	EPI_ISL_3544582	2021-08-12	2021-08-21	<input type="checkbox"/>	29,518	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5822/2021	Original	EPI_ISL_3544581	2021-08-12	2021-08-21	<input type="checkbox"/>	29,518	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5864/2021	Original	EPI_ISL_3544580	2021-08-12	2021-08-21	<input type="checkbox"/>	29,676	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5663/2021	Original	EPI_ISL_3544579	2021-08-12	2021-08-21	<input type="checkbox"/>	29,676	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5657/2021	Original	EPI_ISL_3544578	2021-08-12	2021-08-21	<input type="checkbox"/>	29,728	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5656/2021	Original	EPI_ISL_3544577	2021-08-12	2021-08-21	<input type="checkbox"/>	29,676	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5655/2021	Original	EPI_ISL_3544576	2021-08-12	2021-08-21	<input type="checkbox"/>	29,518	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5653/2021	Original	EPI_ISL_3544575	2021-08-12	2021-08-21	<input type="checkbox"/>	29,728	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5853/2021	Original	EPI_ISL_3544574	2021-08-11	2021-08-21	<input type="checkbox"/>	29,676	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5848/2021	Original	EPI_ISL_3544573	2021-08-11	2021-08-21	<input type="checkbox"/>	29,728	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5847/2021	Original	EPI_ISL_3544572	2021-08-11	2021-08-21	<input type="checkbox"/>	29,726	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5845/2021	Original	EPI_ISL_3544571	2021-08-11	2021-08-21	<input type="checkbox"/>	29,676	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5851/2021	Original	EPI_ISL_3544570	2021-08-10	2021-08-21	<input type="checkbox"/>	29,714	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5844/2021	Original	EPI_ISL_3544569	2021-08-10	2021-08-21	<input type="checkbox"/>	29,725	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5843/2021	Original	EPI_ISL_3544568	2021-08-10	2021-08-21	<input type="checkbox"/>	29,727	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5841/2021	Original	EPI_ISL_3544566	2021-08-10	2021-08-21	<input type="checkbox"/>	29,676	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5840/2021	Original	EPI_ISL_3544565	2021-08-10	2021-08-21	<input type="checkbox"/>	29,728	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5768/2021	Original	EPI_ISL_3544564	2021-08-10	2021-08-21	<input type="checkbox"/>	29,676	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5767/2021	Original	EPI_ISL_3544563	2021-08-10	2021-08-21	<input type="checkbox"/>	29,676	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5766/2021	Original	EPI_ISL_3544562	2021-08-10	2021-08-21	<input type="checkbox"/>	29,728	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5765/2021	Original	EPI_ISL_3544561	2021-08-10	2021-08-21	<input type="checkbox"/>	29,714	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5662/2021	Original	EPI_ISL_3544560	2021-08-10	2021-08-21	<input type="checkbox"/>	29,676	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5661/2021	Original	EPI_ISL_3544559	2021-08-10	2021-08-21	<input type="checkbox"/>	29,727	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5660/2021	Original	EPI_ISL_3544558	2021-08-10	2021-08-21	<input type="checkbox"/>	29,714	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5551/2021	Original	EPI_ISL_3544557	2021-08-10	2021-08-21	<input type="checkbox"/>	29,728	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5550/2021	Original	EPI_ISL_3544556	2021-08-10	2021-08-21	<input type="checkbox"/>	29,728	Human	North America / U	Maine Heal
<input type="checkbox"/>	hCoV-19/USA/ME-HETL-J5839/2021	Original	EPI_ISL_3544555	2021-08-09	2021-08-21	<input type="checkbox"/>	29,200	Human	North America / U	Maine Heal

Total: 583 viruses

<< < 1 2 3 4 5 > >>

Important note: In the GISAID EpiFlu™ Database Access Agreement, you have accepted certain terms and conditions for viewing and using data regarding influenza viruses. To the extent the Database contains data relating to non-influenza viruses, the viewing and use of these data is subject to the same terms and conditions, and by viewing or using such data you agree to be bound by the terms of the GISAID EpiFlu™ Database Access Agreement in respect of such data in the same manner as if they were data relating to influenza viruses.