SARS-CoV-2 Sequences



MIRCORE

October 24, 2020

Inhan Lee, Ph.D. inhan@mircore.org



Copyright © miRcore All Rights Reserved



Volume 181, Issue 4, 14 May 2020, Pages 914-921.e10



Resource

The Architecture of SARS-CoV-2 Transcriptome

Dongwan Kim ^{1, 2}, Joo-Yeon Lee ³, Jeong-Sun Yang ³, Jun Won Kim ³, V. Narry Kim ^{1, 2, 4} $\stackrel{\circ}{\sim}$ $\stackrel{\boxtimes}{\sim}$, Hyeshik

Chang $1, 2 \approx M$ 5000 10000 15000 20000 25000 29903 0 7a 7b 1 3a E M 6 | /8 N 10 ORF1a S 3' UTR ORF1b AAAAA... TRS-L TRS-B TRS-B gRNA S AAAAA... 3a I AAAAA... AAAAA... E M entry AAAAA... AAAAA... Figure 1. Schematic Presentation of the Μ AAAAA... 7a I SARS-CoV-2 Genome Organization, the AAAAA... 7h Canonical Subgenomic mRNAs, and the AAAAA... Ν AAAAA... **Virion Structure** 10(?) AAAAA...

Individual task

- S protein
 - What does it do?
 - Can you find the RNA sequences that S protein is transcribed from?



Let's use UCSC genome browser

https://genome.ucsc.edu/cgi-bin/hgTracks?db=wuhCor1



Alleles





Copyright © miRcore All Rights Reserved

Let's use NCBI

https://www.ncbi.nlm.nih.gov/sars-cov-2/





Contact Us

	 Severe acute respiratory syndrome coronavirus 2 dat Search, retrieve, and analyze SARS-CoV-2 GenBank data. Tree of complete SARS-CoV-2 sequences View a map with geographic distribution of SARS-CoV-2 sequences View SRA data containing coronaviruses 									 Betacoro SARS-Co NCBI SAI CDC outi 	<u>:oronavirus BLAST®</u> <u>-CoV-2 articles in PubMed</u> <u>SARS-CoV-2 Resources</u> outbreak information
	Selecte	d Results: 0						Pu	ibMed Dowr	ıload Align	Build Phylogenetic Tree
ters	Nucleotide (33,857) Protein (360		Protein (360,0	23)	RefSeq Genome (1)						Select Columns
how Fil		Accession 🜲	Release Date 🌲	Spec	es 🖨	Length 🌲	Geo Location 🖨	USA 🖨	Host 🌲	Isolation Source ≑	Collection Date 🖨
S		NC 045512 RefSeq	2020-01-13	Sever	e acute respiratory sy	29903	China		Homo sapiens		2019-12
		<u>MW161041</u>	2020-10-23	Sever	e acute respiratory sy	4043	Russia		Homo sapiens	oronasopharynx	2020-06-04
		<u>MW161259</u>	2020-10-23	Sever	e acute respiratory sy	29796	USA		Homo sapiens	oronasopharynx	2020-01-19
		<u>MW148537</u>	2020-10-22	Sever	e acute respiratory sy	29807	USA: TX	ΤX	Homo sapiens	oronasopharynx	2020-05-14
		<u>MW148538</u>	2020-10-22	Sever	e acute respiratory sy	29805	USA: TX	ТΧ	Homo sapiens	oronasopharynx	2020-05-14
		<u>MW148539</u>	2020-10-22	Sever	e acute respiratory sy	29812	USA: TX	ΤX	Homo sapiens	oronasopharynx	2020-05-13
		<u>MW148540</u>	2020-10-22	Sever	e acute respiratory sy	29815	USA: TX	ТΧ	Homo sapiens	oronasopharynx	2020-05-14
		<u>MW148541</u>	2020-10-22	Sever	e acute respiratory sy	29812	USA: TX	ТХ	Homo sapiens	oronasopharynx	2020-05-12

Group task

- Identify genome location of a protein of your choice
 Why are you interested in this protein?
- Select SARS-CoV-2 sequences (about five or so) from a location that you are interested in
 - Note the accession numbers
 - How many different SNPs did you see in the protein?
 - Use Align function
 - Look for genome locations corresponding to the protein of choice
 - Any SNPs resulting in amino acids?



Copyright © miRcore All Rights Reserved