

BV-BRC Test Report

A15. Service – MSA and SNP

Item to test	MSA and SNP Service using bacterial and viral gene and protein sequences
URL	https://www.bv-brc.org/app/MSA
Prerequisites	Bacterial and viral feature groups and FASTA sequence files
References	https://www.bv-brc.org/docs/quick_references/services/msa_snp_variation_service.html https://www.bv-brc.org/docs/tutorial/msa_snp_variation/msa_snp_variation.html
Tester(s)	Christian Zmasek, Yun Zhang, Maulik Shukla
Test date	08-May-2022 (follow-up from original test)
Test result	Passed

Overview

- Test the MSA and SNP Service using exemplar bacterial and viral gene and protein sequences.
- Test input options, i.e., feature groups, fasta sequence files, and sequence input box.
- Test different MSA algorithms, MAFFT and Muscle.
- For each job submitted, verify successful completion of the job, presence of output files in various formats, review resulting MSA and SNPs.
- View MSA using interactive MSA viewer and verify all functionality.

Test Data

Dataset	Rational	Input Format	Input
SARSCoV2 spike proteins	Viral proteins of interest	Feature group, FASTA file	SARSCoV2_Spike_protein.fasta
MTB ahpD proteins	Bacterial proteins of interest	Feature group, FASTA file	MTB_ahpD_protein.fasta

- All test datasets and corresponding job results are available in the following public workspace:
<https://www.bv-brc.org/workspace/BVBRC@patricbrc.org/BVBRC%20Tests/Multiple%20Sequence%20Alignment>

Test Results

- All MSA jobs completed successfully, without any errors.
- All jobs resulted in expected output files in corresponding job output directory, including alignment in afa, nexus, phy, and pir formars and consensus sequence in fasta format.
- SNPs were present as a tsv file in the expected format and the tsv viewer allow filtering and sorting of the table based on the SNP scores.

- The link to MSA viewer loaded the MSA and all functionality worked as expected, including changing node labels, and showing / hiding positions based on conservation.
- All test datasets and corresponding job results are available in the following public workspace: <https://www.bv-brc.org/workspace/BVBRC@patricbrc.org/BVBRC%20Tests/Multiple%20Sequence%20Alignment>
- Below are a series of screenshots showing successful completion of the jobs, availability of the result files in the workspace, and the MSA viewer.

Start with: UNALIGNED SEQUENCES ALIGNED SEQUENCES

SELECT FEATURE GROUP
 MTB ahpD protein
 DNA PROTEIN

SELECT DNA OR PROTEIN FASTA FILE
 FASTA file

INPUT FASTA SEQUENCE
 Enter fasta records of sequences to align.

ALIGNER
 Mafft

OUTPUT FOLDER
 Multiple Sequence Alignment

OUTPUT NAME
 Output Name

Your job has been submitted successfully. Please visit your [Jobs List](#) to check the status of your job and access the results.

[Reset](#) [Submit](#)

Status	ID	Service	Output Name	Submit	Start	Completed
completed	7747116	MSA	SARSCov2 Spike Protein - Nonhuman Mammals - MSA	5/7/22, 5:05 PM	5/7/22, 5:05 PM	5/7/22, 5:06 PM
completed	7747808	MSA	MTB ahpD protein - MSA	5/8/22, 4:28 PM	5/8/22, 4:28 PM	5/8/22, 4:28 PM
completed	7747809	MSA	MTB ahpD protein - NA	5/8/22, 4:28 PM	5/8/22, 4:28 PM	5/8/22, 4:28 PM
completed	7747810	MSA	MTB ahpD protein - AA	5/8/22, 4:28 PM	5/8/22, 4:28 PM	5/8/22, 4:28 PM
completed	7747811	MSA	MTB ahpD protein - AA - Muscle	5/8/22, 4:29 PM	5/8/22, 4:29 PM	5/8/22, 4:29 PM
completed	7747812	MSA	SARSCov2 Spike protein - MSA	5/8/22, 4:29 PM	5/8/22, 4:29 PM	5/8/22, 4:29 PM
completed	7747813	MSA	SARSCov2 Spike protein - Muscle	5/8/22, 4:29 PM	5/8/22, 4:29 PM	5/8/22, 4:30 PM
completed	7747814	MSA	SARSCov2 Spike protein - NA	5/8/22, 4:30 PM	5/8/22, 4:30 PM	5/8/22, 4:31 PM
completed	7747815	MSA	SARSCov2 Spike protein - AA	5/8/22, 4:30 PM	5/8/22, 4:30 PM	5/8/22, 4:30 PM
completed	7747816	MSA	SARSCov2 Spike protein - AA	5/8/22, 4:30 PM	5/8/22, 4:30 PM	5/8/22, 4:30 PM
completed	7747817	MSA	SARSCov2 Spike protein - AA - Muscle	5/8/22, 4:30 PM	5/8/22, 4:30 PM	5/8/22, 4:31 PM
completed	7747827	MSA	MTB ahpD protein - AA - fasta	5/8/22, 5:00 PM	5/8/22, 5:00 PM	5/8/22, 5:01 PM
completed	7747831	MSA	SARSCov2 Spike protein - fasta	5/8/22, 5:01 PM	5/8/22, 5:01 PM	5/8/22, 5:01 PM
completed	7747832	MSA	SARSCov2 Spike protein - seq input	5/8/22, 5:04 PM	5/8/22, 5:04 PM	5/8/22, 5:04 PM

Name	Size	Owner	Members	Created
📄 SARSCoV2_Spike_protein.fasta	31.8 kB	me	Public	5/8/22, 4:59 PM
📄 SARSCoV2 Spike protein - seq input	39.3 kB	me	Public	5/8/22, 5:04 PM
📄 SARSCoV2 Spike protein - fasta	7.1 kB	me	Public	5/8/22, 5:01 PM
📄 SARSCoV2 Spike protein - NA	6.9 kB	me	Public	5/8/22, 4:31 PM
📄 SARSCoV2 Spike protein - Muscle	7.0 kB	me	Public	5/8/22, 4:30 PM
📄 SARSCoV2 Spike protein - MSA	6.9 kB	me	Public	5/8/22, 4:29 PM
📄 SARSCoV2 Spike protein - AA - Muscle	7.1 kB	me	Public	5/8/22, 4:31 PM
📄 SARSCoV2 Spike protein - AA	7.1 kB	me	Public	5/8/22, 4:30 PM
📄 SARSCoV2 Spike protein	22 features	me	Public	5/8/22, 4:26 PM
↑ Parent folder	-	-	-	-
📄 MTB_ahpD_protein.fasta	15.6 kB	me	Public	5/8/22, 4:59 PM
📄 MTB ahpD protein - NA	6.8 kB	me	Public	5/8/22, 4:28 PM
📄 MTB ahpD protein - MSA	6.8 kB	me	Public	5/8/22, 4:28 PM
📄 MTB ahpD protein - AA - fasta	7.0 kB	me	Public	5/8/22, 5:00 PM
📄 MTB ahpD protein - AA - Muscle	7.0 kB	me	Public	5/8/22, 4:29 PM
📄 MTB ahpD protein - AA	6.8 kB	me	Public	5/8/22, 4:28 PM
📄 MTB ahpD protein	48 features	me	Public	5/8/22, 4:27 PM

Name	Size	Owner	Members	Created
↑ Parent folder	-	-	-	-
📄 SARSCoV2 Spike protein - AA.afa	29.2 kB	me	Public	5/8/22, 4:30 PM
📄 SARSCoV2 Spike protein - AA.aln	49.3 kB	me	Public	5/8/22, 4:30 PM
📄 SARSCoV2 Spike protein - AA.consensus.fasta	1.3 kB	me	Public	5/8/22, 4:30 PM
📄 SARSCoV2 Spike protein - AA.entropy.svg	265.4 kB	me	Public	5/8/22, 4:30 PM
📄 SARSCoV2 Spike protein - AA.nexus	39.5 kB	me	Public	5/8/22, 4:30 PM
📄 SARSCoV2 Spike protein - AA.phy	46.4 kB	me	Public	5/8/22, 4:30 PM
📄 SARSCoV2 Spike protein - AA.pir	30.4 kB	me	Public	5/8/22, 4:30 PM
📄 SARSCoV2 Spike protein - AA.snp.tsv	26.6 kB	me	Public	5/8/22, 4:30 PM
📄 mafft.job.log	3.4 kB	me	Public	5/8/22, 4:30 PM
📄 muscle.job.log	105 B	me	Public	5/8/22, 4:30 PM

KEYWORDS All Columns

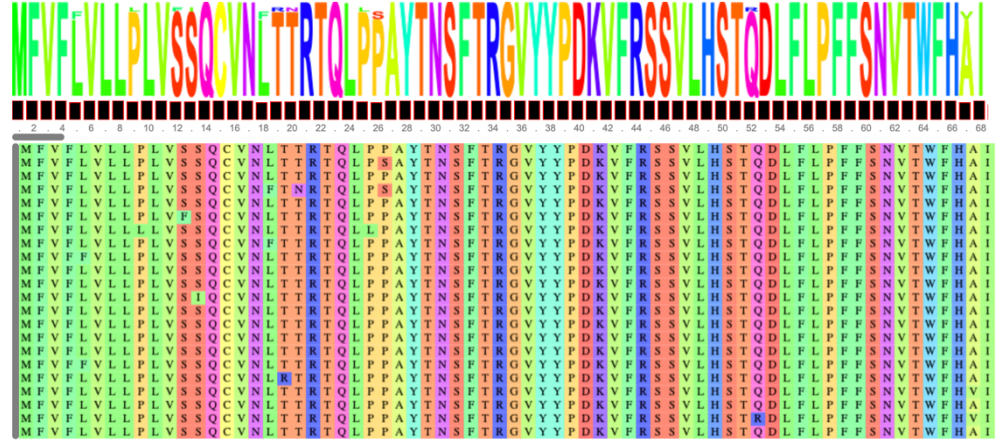
First Row Contains Column Headers

<input type="checkbox"/>	Position	Score	Consensus	Detail	NumberOfSequence
<input type="checkbox"/>	452	94	Leu	Arg=4,Gln=1,Leu=17	22
<input type="checkbox"/>	501	85	Asn	Asn=16,Tyr=6	22
<input type="checkbox"/>	152	79	Trp	Arg=1,Cys=1,Leu=1,Trp=19	22
<input type="checkbox"/>	69	77	His	Del=5,His=17	22
<input type="checkbox"/>	70	77	Val	Del=5,Val=17	22
<input type="checkbox"/>	417	70	Lys	Asn=2,Lys=19,Thr=1	22
<input type="checkbox"/>	95	68	Thr	Ile=4,Thr=18	22
<input type="checkbox"/>	1027	68	Thr	Ile=4,Thr=18	22
<input type="checkbox"/>	242	57	Leu	Del=3,Leu=19	22
<input type="checkbox"/>	243	57	Ala	Ala=19,Del=3	22
<input type="checkbox"/>	655	57	His	His=19,Tyr=3	22
<input type="checkbox"/>	716	57	Thr	Ile=3,Thr=19	22
<input type="checkbox"/>	145	53	Tyr	Del=1,Ser=1,Tyr=20	22
<input type="checkbox"/>	156	53	Glu	Del=1,Glu=20,Gly=1	22
<input type="checkbox"/>	253	53	Asp	Asn=1,Asp=20,Gly=1	22
<input type="checkbox"/>	346	53	Arg	Arg=20,Lys=1,Ser=1	22
<input type="checkbox"/>	796	53	Asp	Asp=20,His=1,Tyr=1	22
<input type="checkbox"/>	5	44	Leu	Leu=20,Phe=2	22
<input type="checkbox"/>	18	44	Leu	Leu=20,Phe=2	22
<input type="checkbox"/>	26	44	Pro	Pro=20,Ser=2	22
<input type="checkbox"/>	67	44	Ala	Ala=20,Val=2	22
<input type="checkbox"/>	142	44	Gly	Asp=2,Gly=20	22
<input type="checkbox"/>	157	44	Phe	Del=2,Phe=20	22

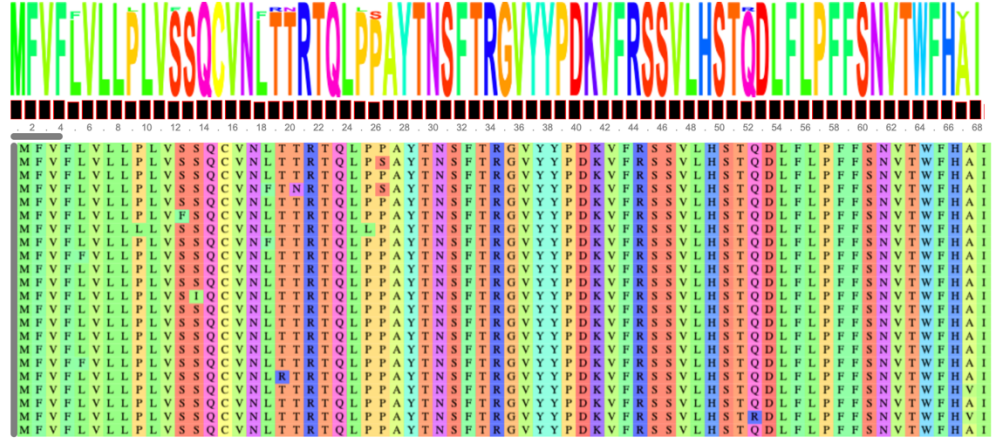
1 - 200 of 1278 results

< 1 2 3 ... 7 >

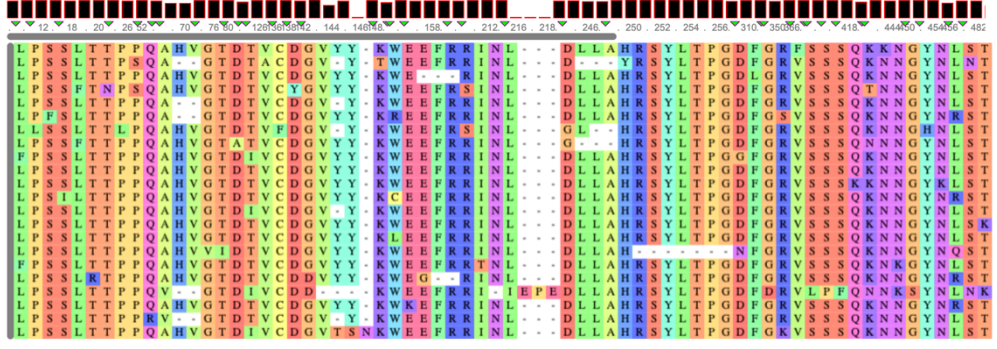
- ID Label
- 1 fig|2697049.446809.CDS.3
- 2 fig|2697049.1040807.CDS.3
- 3 fig|2697049.1329549.CDS.3
- 4 fig|2697049.331381.CDS.3
- 5 fig|2697049.300955.CDS.3
- 6 fig|2697049.1840333.CDS.3
- 7 fig|2697049.1259397.CDS.3
- 8 fig|2697049.293736.CDS.3
- 9 fig|2697049.333080.CDS.3
- 10 fig|2697049.107626.CDS.3
- 11 fig|2697049.1032747.CDS.4
- 12 fig|2697049.333200.CDS.3
- 13 fig|2697049.374079.CDS.3
- 14 fig|2697049.335130.CDS.6
- 15 fig|2697049.293793.CDS.3
- 16 fig|2697049.407547.CDS.3
- 17 fig|2697049.889036.CDS.3
- 18 fig|2697049.1272479.CDS.3
- 19 fig|2697049.2862877.peg.3
- 20 fig|2697049.433713.CDS.3
- 21 fig|2697049.120701.CDS.3
- 22 fig|2697049.1877076.peg.3



- ID Label
- 1 B.1.466.2
- 2 B.1.620
- 3 B.1.1.523
- 4 P.1
- 5 B.1.1.7
- 6 C.36.3
- 7 C.1.2
- 8 B.1.351
- 9 B.1.526
- 10 B
- 11 B.1
- 12 B.1.429
- 13 B.1.1.1318
- 14 B.1.1.519
- 15 R.1
- 16 C.37
- 17 B.1.619
- 18 B.1.617.3
- 19 BA.1
- 20 B.1.617.1
- 21 B.1.525
- 22 B.1.621



- ID Label
- 1 fig|2697049.446809.CDS.3
- 2 fig|2697049.1040807.CDS.3
- 3 fig|2697049.1329549.CDS.3
- 4 fig|2697049.331381.CDS.3
- 5 fig|2697049.300955.CDS.3
- 6 fig|2697049.1840333.CDS.3
- 7 fig|2697049.1259397.CDS.3
- 8 fig|2697049.293736.CDS.3
- 9 fig|2697049.333080.CDS.3
- 10 fig|2697049.107626.CDS.3
- 11 fig|2697049.1032747.CDS.4
- 12 fig|2697049.333200.CDS.3
- 13 fig|2697049.374079.CDS.3
- 14 fig|2697049.335130.CDS.6
- 15 fig|2697049.293793.CDS.3
- 16 fig|2697049.407547.CDS.3
- 17 fig|2697049.889036.CDS.3
- 18 fig|2697049.1272479.CDS.3
- 19 fig|2697049.2862877.peg.3
- 20 fig|2697049.433713.CDS.3
- 21 fig|2697049.120701.CDS.3
- 22 fig|2697049.1877076.peg.3



References

- [MSA and SNP Service Quick Reference Guide](#)
- [MSA and SNP Service Tutorial](#)