

# BV-BRC Test Report

## A6. Service – Homology / BLAST

<b>Item to test</b>	Homology / BLAST Service using bacterial and viral query sequences and databases
<b>URL</b>	<a href="https://www.bv-brc.org/app/Homology">https://www.bv-brc.org/app/Homology</a>
<b>Prerequisites</b>	FASTA NA and AA query sequences for bacterial and viral proteins
<b>References</b>	<a href="https://www.bv-brc.org/docs/quick_references/services/blast.html">https://www.bv-brc.org/docs/quick_references/services/blast.html</a> <a href="https://www.bv-brc.org/docs/tutorial/blast/blast.html">https://www.bv-brc.org/docs/tutorial/blast/blast.html</a>
<b>Tester(s)</b>	Rebecca Wattam, Maulik Shukla
<b>Test date</b>	06-May-2022 (follow-up from original test)
<b>Test result</b>	<b>Passed</b>

### Overview

- Test the Homology / BLAST service using exemplar queries for both bacteria and viruses.
- Test input options, i.e. various BLAST programs, nucleotide and protein query sequences.
- Test various BLAST databases, i.e. reference genome databases, taxon level databases, and genome specific databases for both bacterial and viruses.
- For each job submitted, verify successful completion of the job, availability of the output files in the workspace, and quality of the results.
- Test result selection and action from the result table.

### Test Data

<b>Dataset</b>	<b>Rational</b>	<b>Input Format</b>	<b>Input</b>
SARS-CoV-2 Spike protein	Viral protein of interest	FASTA NA and AA sequences	sarscov2_spike_protein_na.fasta sarscov2_spike_protein_aa.fasta
MTB ahpD protein	Bacterial protein of interest	FASTA NA and AA sequences	mtb_ahpd_protein_na.fasta mtb_ahpd_protein_aa.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/BLAST>

### Test Results

- All jobs completed successfully, without any errors.
- All jobs resulted in expected output files in corresponding job output directory.
- For each of the jobs, the results were as expected, finding similar genes, proteins or genomic sequences from the selected reference genome set, taxon level, or genome level databases.
- 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/BLAST>























- Below are a series of screenshots showing successful completion of the jobs, availability of the result files in the workspace, and the result table.
- Successful completion of all jobs using sars-cov-2 spile protein NA or AA sequence as query and search using BLASTN, BLASTX, BLASTP, and TBLASTN against databases of all reference genomes, taxon level, and single genome level.

Status	ID	Service	Output Name	Submit	Start	Completed
completed	7746845	Homology	Bacteria ahpD protein - Ref - BLASTX	5/6/22, 8:04 PM	5/6/22, 8:04 PM	5/6/22, 8:05 PM
completed	7746844	Homology	Bacteria ahpD protein - Ref - BLASTN - contigs	5/6/22, 8:03 PM	5/6/22, 8:03 PM	5/6/22, 8:06 PM
completed	7746843	Homology	Bacteria ahpD protein - Ref - BLASTN	5/6/22, 8:03 PM	5/6/22, 8:03 PM	5/6/22, 8:05 PM
completed	7746842	Homology	Bacteria ahpD protein - Taxon - BLASTN - contigs	5/6/22, 8:03 PM	5/6/22, 8:03 PM	5/6/22, 8:11 PM
completed	7746841	Homology	Bacteria ahpD protein - Taxon - BLASTN	5/6/22, 8:02 PM	5/6/22, 8:03 PM	5/6/22, 8:11 PM
completed	7746839	Homology	Bacteria ahpD protein - Genome - BLASTX	5/6/22, 8:02 PM	5/6/22, 8:02 PM	5/6/22, 8:02 PM
completed	7746838	Homology	Bacteria ahpD protein - Genome - BLASTN - contigs	5/6/22, 8:01 PM	5/6/22, 8:01 PM	5/6/22, 8:02 PM
completed	7746837	Homology	Bacteria ahpD protein - Genome - BLASTN	5/6/22, 8:01 PM	5/6/22, 8:01 PM	5/6/22, 8:01 PM
completed	7746830	Homology	Bacteria ahpD protein - Taxon - TBLASTN - contigs	5/6/22, 7:24 PM	5/6/22, 7:24 PM	5/6/22, 7:25 PM
completed	7746829	Homology	Bacteria ahpD protein - Taxon - TBLASTN	5/6/22, 7:24 PM	5/6/22, 7:24 PM	5/6/22, 7:25 PM
completed	7746828	Homology	Bacteria ahpD protein - taxon - BLASTP	5/6/22, 7:24 PM	5/6/22, 7:24 PM	5/6/22, 7:24 PM
completed	7746827	Homology	Bacteria ahpD protein - Genome - TBLASTN - contigs	5/6/22, 7:18 PM	5/6/22, 7:18 PM	5/6/22, 7:25 PM
completed	7746825	Homology	Bacteria ahpD protein - Genome - BLASTP	5/6/22, 7:18 PM	5/6/22, 7:18 PM	5/6/22, 7:22 PM
completed	7746824	Homology	Bacteria ahpD protein - Ref - TBLASTN - contigs	5/6/22, 7:17 PM	5/6/22, 7:17 PM	5/6/22, 7:18 PM
completed	7746823	Homology	Bacteria ahpD protein - Ref - TBLASTN	5/6/22, 7:17 PM	5/6/22, 7:17 PM	5/6/22, 7:18 PM
completed	7746822	Homology	Bacteria ahpD protein - Ref - BLASTP	5/6/22, 7:16 PM	5/6/22, 7:16 PM	5/6/22, 7:17 PM

- Successful completion of all jobs using MTB aphD protein NA or AA sequence as query and search using BLASTN, BLASTX, BLASTP, and TBLASTN against databases of all reference genomes, taxon level, and single genome level.

Status	ID	Service	Output Name	Submit	Start	Completed
completed	7746792	Homology	Viral Spike protein - Ref - BLASTP	5/6/22, 6:38 PM	5/6/22, 6:38 PM	5/6/22, 6:38 PM
completed	7746793	Homology	Viral Spike protein - Ref - TBLASTN	5/6/22, 6:38 PM	5/6/22, 6:39 PM	5/6/22, 6:39 PM
completed	7746794	Homology	Viral Spike protein - Ref - TBLASTN - contigs	5/6/22, 6:39 PM	5/6/22, 6:39 PM	5/6/22, 6:39 PM
completed	7746803	Homology	Viral Spike protein - Ref - BLASTN	5/6/22, 6:47 PM	5/6/22, 6:47 PM	5/6/22, 6:47 PM
completed	7746804	Homology	Viral Spike protein - Ref - BLASTN - contigs	5/6/22, 6:47 PM	5/6/22, 6:47 PM	5/6/22, 6:48 PM
completed	7746805	Homology	Viral Spike protein - Ref - BLASTX	5/6/22, 6:48 PM	5/6/22, 6:48 PM	5/6/22, 6:48 PM
completed	7746808	Homology	Viral Spike protein - Taxon - BLASTN - contigs	5/6/22, 6:49 PM	5/6/22, 6:49 PM	5/6/22, 6:52 PM
completed	7746807	Homology	Viral Spike protein - Taxon - BLASTN	5/6/22, 6:48 PM	5/6/22, 6:48 PM	5/6/22, 6:52 PM
completed	7746799	Homology	Viral Spike protein - Ref - BLASTP - taxon	5/6/22, 6:44 PM	5/6/22, 6:44 PM	5/6/22, 6:54 PM
completed	7746801	Homology	Viral Spike protein - taxon - TBLASTN - contigs	5/6/22, 6:45 PM	5/6/22, 6:45 PM	5/6/22, 6:56 PM
completed	7746802	Homology	Viral Spike protein - taxon - BLASTP	5/6/22, 6:46 PM	5/6/22, 6:46 PM	5/6/22, 6:57 PM
completed	7746806	Homology	Viral Spike protein - taxon - BLASTP	5/6/22, 6:48 PM	5/6/22, 6:48 PM	5/6/22, 7:03 PM
completed	7746815	Homology	Viral Spike protein - genome - BLASTP	5/6/22, 7:07 PM	5/6/22, 7:07 PM	5/6/22, 7:07 PM
completed	7746816	Homology	Viral Spike protein - genome - TBLASTN	5/6/22, 7:07 PM	5/6/22, 7:07 PM	5/6/22, 7:07 PM
completed	7746817	Homology	Viral Spike protein - genome - TBLASTN - contigs	5/6/22, 7:08 PM	5/6/22, 7:08 PM	5/6/22, 7:08 PM
completed	7746818	Homology	Viral Spike protein - genome - BLASTN - contigs	5/6/22, 7:10 PM	5/6/22, 7:10 PM	5/6/22, 7:10 PM
completed	7746819	Homology	Viral Spike protein - genome - BLASTN	5/6/22, 7:10 PM	5/6/22, 7:10 PM	5/6/22, 7:10 PM
completed	7746820	Homology	Viral Spike protein - genome - BLASTX	5/6/22, 7:10 PM	5/6/22, 7:10 PM	5/6/22, 7:10 PM

- BLAST results in the workspace directories.

Name	Size	Owner	Members	Created
 sarscov2_spike_protein_na.fasta	4.0 kB	me	Public	5/7/22, 6:51 AM
 sarscov2_spike_protein_aa.fasta	1.4 kB	me	Public	5/7/22, 6:52 AM
 mtb_ahpd_protein_na.fasta	690 B	me	Public	5/7/22, 6:51 AM
 mtb_ahpd_protein_aa.fasta	327 B	me	Public	5/7/22, 6:52 AM
 Viral Spike protein - taxon - TBLASTN - contigs	8.8 kB	me	Public	5/6/22, 6:56 PM
 Viral Spike protein - taxon - TBLASTN	7.7 kB	me	Public	5/6/22, 6:45 PM
 Viral Spike protein - taxon - BLASTP	11.4 kB	me	Public	5/6/22, 7:03 PM
 Viral Spike protein - genome - TBLASTN - contigs	8.9 kB	me	Public	5/6/22, 7:08 PM
 Viral Spike protein - genome - TBLASTN	8.8 kB	me	Public	5/6/22, 7:07 PM
 Viral Spike protein - genome - BLASTX	11.4 kB	me	Public	5/6/22, 7:10 PM
 Viral Spike protein - genome - BLASTP	8.8 kB	me	Public	5/6/22, 7:07 PM
 Viral Spike protein - genome - BLASTN - contigs	11.5 kB	me	Public	5/6/22, 7:10 PM
 Viral Spike protein - genome - BLASTN	11.4 kB	me	Public	5/6/22, 7:10 PM
 Viral Spike protein - Taxon - BLASTN - contigs	11.5 kB	me	Public	5/6/22, 6:52 PM
 Viral Spike protein - Taxon - BLASTN	11.4 kB	me	Public	5/6/22, 6:52 PM
 Viral Spike protein - Ref - TBLASTN - contigs	8.8 kB	me	Public	5/6/22, 6:39 PM
 Viral Spike protein - Ref - TBLASTN	8.7 kB	me	Public	5/6/22, 6:39 PM
 Viral Spike protein - Ref - BLASTX	11.3 kB	me	Public	5/6/22, 6:48 PM
 Viral Spike protein - Ref - BLASTP - taxon	8.8 kB	me	Public	5/6/22, 6:54 PM
 Viral Spike protein - Ref - BLASTP	8.7 kB	me	Public	5/6/22, 6:38 PM
 Viral Spike protein - Ref - BLASTN - contigs	11.4 kB	me	Public	5/6/22, 6:48 PM
 Viral Spike protein - Ref - BLASTN	11.3 kB	me	Public	5/6/22, 6:47 PM

Bacteria ahpD protein - taxon - BLASTP	7.7 kB	me	Public	5/6/22, 7:24 PM
Bacteria ahpD protein - Taxon - TBLASTN - contigs	7.7 kB	me	Public	5/6/22, 7:25 PM
Bacteria ahpD protein - Taxon - TBLASTN	7.7 kB	me	Public	5/6/22, 7:25 PM
Bacteria ahpD protein - Taxon - BLASTN - contigs	8.1 kB	me	Public	5/6/22, 8:11 PM
Bacteria ahpD protein - Taxon - BLASTN	8.0 kB	me	Public	5/6/22, 8:11 PM
Bacteria ahpD protein - Ref - TBLASTN - contigs	7.7 kB	me	Public	5/6/22, 7:18 PM
Bacteria ahpD protein - Ref - TBLASTN	7.6 kB	me	Public	5/6/22, 7:18 PM
Bacteria ahpD protein - Ref - BLASTX	8.0 kB	me	Public	5/6/22, 8:05 PM
Bacteria ahpD protein - Ref - BLASTP	7.6 kB	me	Public	5/6/22, 7:17 PM
Bacteria ahpD protein - Ref - BLASTN - contigs	8.0 kB	me	Public	5/6/22, 8:06 PM
Bacteria ahpD protein - Ref - BLASTN	8.0 kB	me	Public	5/6/22, 8:05 PM
Bacteria ahpD protein - Genome - TBLASTN - contigs	7.7 kB	me	Public	5/6/22, 7:25 PM
Bacteria ahpD protein - Genome - BLASTX	8.0 kB	me	Public	5/6/22, 8:02 PM
Bacteria ahpD protein - Genome - BLASTP	7.7 kB	me	Public	5/6/22, 7:22 PM
Bacteria ahpD protein - Genome - BLASTN - contigs	8.1 kB	me	Public	5/6/22, 8:02 PM
Bacteria ahpD protein - Genome - BLASTN	8.0 kB	me	Public	5/6/22, 8:01 PM

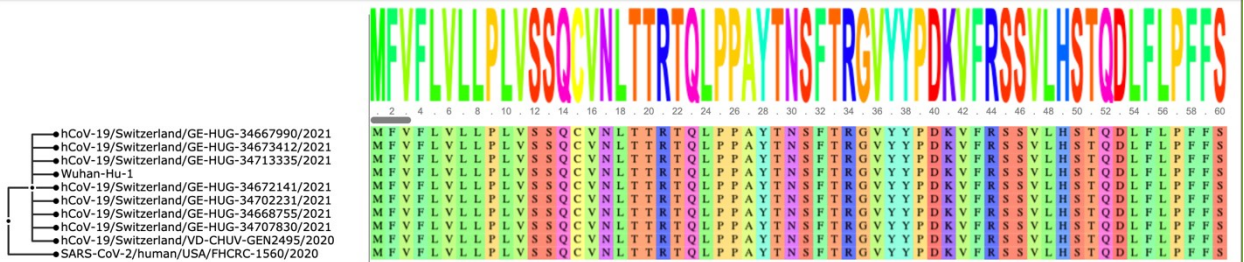
- BLAST Result table

/BVBRC@patricbrc.org/BVBRC%20Tests/BLAST/Viral%20Spike%20protein%20-%20taxon%20-%20BLASTP



Query Genome ID	Subject ID	RefSeq Locus Tag	Product	Length (NT)	Length (AA)	ALN Length	Identity (%)	Query cover (%)	Subject cover (%)	Score	E value
fig 265 Severe acute respiratory syndrome coronavirus 2 SAR  fig 2697049.115212.CDS.3	fig 2697049.107626.CDS.3	GU280_gp02	surface glycoprotein	3822	1273	1273	100	99.06	99.14	2553	0
fig 265 Severe acute respiratory syndrome coronavirus 2 Wuh  fig 2697049.1521182.CDS.	fig 2697049.1501832.CDS.		surface glycoprotein	3822	1273	1273	100	99.06	99.14	2553	0
fig 265 Severe acute respiratory syndrome coronavirus 2 strair  fig 2697049.1496554.CDS.	fig 2697049.1496554.CDS.		surface glycoprotein	3822	1273	1273	100	99.06	99.14	2553	0
fig 265 Severe acute respiratory syndrome coronavirus 2 strair  fig 2697049.1497405.CDS.	fig 2697049.1497405.CDS.		surface glycoprotein	3822	1273	1273	100	99.06	99.14	2553	0
fig 265 Severe acute respiratory syndrome coronavirus 2 strair  fig 2697049.1501832.CDS.	fig 2697049.1501832.CDS.		surface glycoprotein	3822	1273	1273	100	99.06	99.14	2553	0
fig 265 Severe acute respiratory syndrome coronavirus 2 strair  fig 2697049.1496601.CDS.	fig 2697049.1496601.CDS.		surface glycoprotein	3822	1273	1273	100	99.06	99.14	2553	0
fig 265 Severe acute respiratory syndrome coronavirus 2 strair  fig 2697049.1496501.CDS.	fig 2697049.1496501.CDS.		surface glycoprotein	3822	1273	1273	100	99.06	99.14	2553	0
fig 265 Severe acute respiratory syndrome coronavirus 2 strair  fig 2697049.1496822.CDS.	fig 2697049.1496822.CDS.		surface glycoprotein	3822	1273	1273	100	99.06	99.14	2553	0
fig 265 Severe acute respiratory syndrome coronavirus 2 strair  fig 2697049.1497168.CDS.	fig 2697049.1497168.CDS.		surface glycoprotein	3822	1273	1273	100	99.06	99.14	2553	0

- Result selection and generation of MSA



## References

- [BLAST Service Quick Reference Guide](#)
- [BLAST Service Tutorial](#)