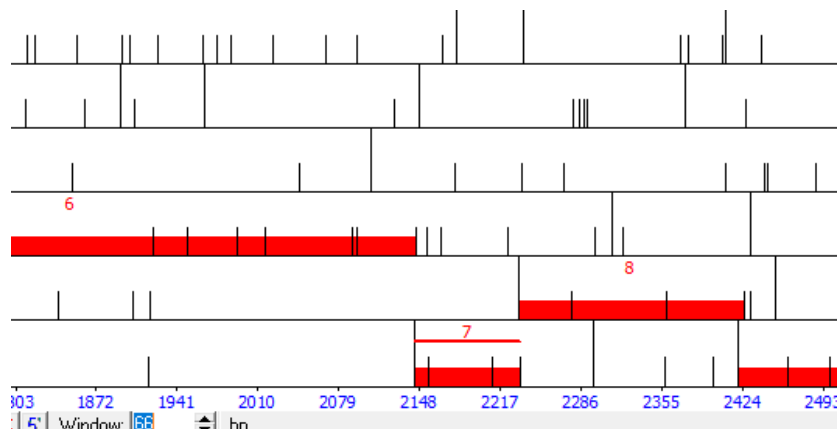


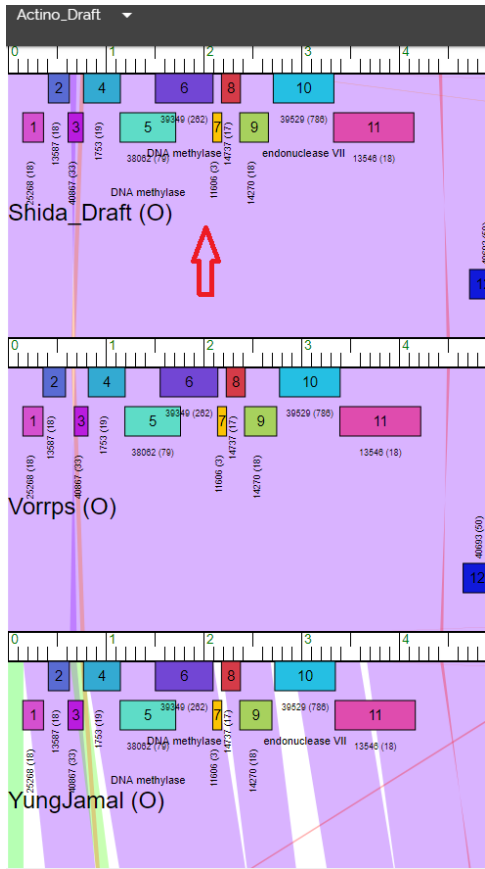
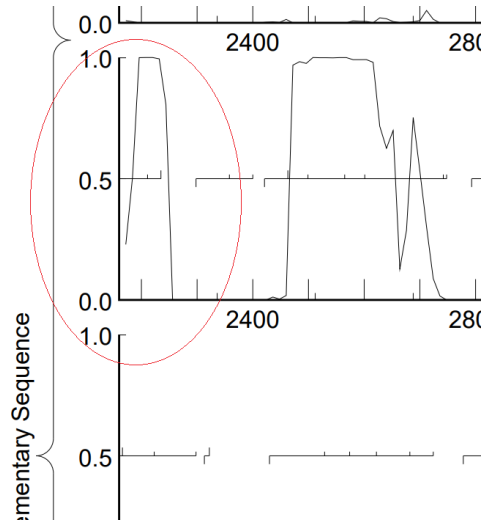
## Having trouble getting NCBI Blast data for a gene via DNA Master, but can get the data directly via phagesDb and NCBI websites: Shida gene 7

Dear phage hunters, since the very first auto-annotation till now, we've had trouble getting NCBI blast data for Shida Gene 7 when we do Blast through DNA Master, but can get the blast data when we do the blastp directly on the NCBI website

(<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), as well as in PhagesDb. This gene is present in Vorrps, YungJamal & Krili in phamerator and had q1:s1, 100% with phage Krill in phagesDb; Original GeneMark call @bp 2236

SSC: 2144-2236 rev; CP: Yes; SCS: Both; ST:SS; BLAST-Start: Krili, gp 7, phagesDb, q1:s1, 100%, 8e-12; Gap:4 bp overlap. TMHMM and SOSUI analysis confirmed this to be a **membrane protein**, with at least one transmembrane domain. Since we have tried several times without success getting Blast data for this gene via DNA master and its needed for submission of the final files, what would you advise us to do? This is the only gene in Shida that has not budged! Below are the Frames, Coding potential profile, phamerator snapshot, Caption of the results (actually no result) when Blasted Via DNA Master, PhagedDb BLASTp results, NCBI BLASTp results (when done directly via the website), TMHMM & SOSUI analysis results. Please advise, because we have tried to get blast data from this gene directly via DNA Master as required, but all in vain.





NO Blast Data when BLASTed directly via DNA Master (see below)

The screenshot shows the DNA Master interface. A window titled "BLAST search for complement (2144 - 2236) (7)" is open. The window has tabs for "Retrieve", "XML Results", and "Text Results". The main area is empty, indicating no results were found. A message at the top right of the window says "There are no BLAST results for this feature". The background shows a table with columns for Tag, Name, 5' End, 3' End, and Length. The table contains two rows: SEA\_SHIDA\_1 (1, 153, 359, 207) and SEA\_SHIDA\_2 (2, 359, 676, 318).

The screenshot shows the "Text Results" tab of the BLAST search window. The output text is as follows:

```

Length=31

<b>No significant similarity found.</b> For reasons why, <A HREF = "Blast.cgi?
CMD=Web&PAGE_TYPE=BlastDocs&DOC_TYPE=FAQ"><b>click here</b></A>.</b><br><br>
Database: All non-redundant GenBank CDS translations+PDB+SwissProt+PIR+PRF
excluding environmental samples from WGS projects
Posted date: May 7, 2020 12:42 AM
Number of letters in database: 101,765,175,729
Number of sequences in database: 282,858,011

Lambda      K          H
0.332      0.144     0.670
Gapped
Lambda      K          H
0.267      0.0410    0.140
Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Sequences: 282858011
Number of Hits to DB: 23668805
Number of extensions: 146570
Number of successful extensions: 151
Number of sequences better than 100: 0
Number of HSP's better than 100 without gapping: 0
Number of HSP's gapped: 151
Number of HSP's successfully gapped: 0
Length of query: 31
Length of database: 101765175729
    
```

## PhagedDb BLASTp results

← → ↻ 🔒 phagesdb.org/blast/results/blast.cgi

Sequences producing significant alignments:	Score (bits)	E Value
Vorrps_7, function unknown, 30	<a href="#">67</a>	8e-12
Shida_Draft_7, function unknown, 30	<a href="#">67</a>	8e-12
Krili_7, function unknown, 30	<a href="#">67</a>	8e-12

>Vorrps\_7, function unknown, 30  
Length = 30

Score = 67.4 bits (163), Expect = 8e-12  
Identities = 30/30 (100%), Positives = 30/30 (100%)

Query: 1 MIGWNAEFVLRGLAATAILTALACLGMNPL 30  
MIGWNAEFVLRGLAATAILTALACLGMNPL  
Sbjct: 1 MIGWNAEFVLRGLAATAILTALACLGMNPL 30

>Shida\_Draft\_7, function unknown, 30  
Length = 30

Score = 67.4 bits (163), Expect = 8e-12  
Identities = 30/30 (100%), Positives = 30/30 (100%)

Query: 1 MIGWNAEFVLRGLAATAILTALACLGMNPL 30  
MIGWNAEFVLRGLAATAILTALACLGMNPL  
Sbjct: 1 MIGWNAEFVLRGLAATAILTALACLGMNPL 30

>Krili\_7, function unknown, 30  
Length = 30

Score = 67.4 bits (163), Expect = 8e-12  
Identities = 30/30 (100%), Positives = 30/30 (100%)

Query: 1 MIGWNAEFVLRGLAATAILTALACLGMNPL 30  
MIGWNAEFVLRGLAATAILTALACLGMNPL  
Sbjct: 1 MIGWNAEFVLRGLAATAILTALACLGMNPL 30

Database: Actinobacteriophage Proteins as of May 09, 2020  
Posted date: May 9, 2020 4:15 AM  
Number of letters in database: 64,772,587  
Number of sequences in database: 335,569

# NCBI Blast results when done directly via the website, not via DNA Master

blast.ncbi.nlm.nih.gov/blast.cgi

COVID-19 is an emerging, rapidly evolving situation.  
Get the latest public health information from CDC: <https://www.cdc.gov/coronavirus/>.  
Get the latest research from NIH: <https://www.nih.gov/coronavirus/>.

BLAST » blastp suite » results for RID-BGFS7RF101R Home Recent Results Saved Strategies Help

[Edit Search](#) [Save Search](#) [Search Summary](#)

**Job Title** Protein Sequence

**RID** [BGFS7RF101R](#) Search expires on 05-12 01:51 am [Download All](#)

**Program** BLASTP [Citation](#)

**Database** nr [See details](#)

**Query ID** lcl|Query\_11633

**Description** None

**Molecule type** amino acid

**Query Length** 31

**Other reports** [Distance tree of results](#) [Multiple alignment](#) [MSA viewer](#)

**Filter Results**

**Organism** only top 20 will appear  exclude

Type common name, binomial, taxid or group name

[Add organism](#)

**Percent Identity**  to  **E value**  to  **Query Coverage**  to

[Filter](#) [Reset](#)

**Descriptions** [Graphic Summary](#) [Alignments](#) [Taxonomy](#)

**Sequences producing significant alignments** [Download](#) [Manage Columns](#) Show

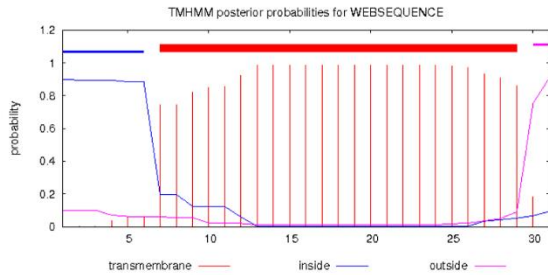
select all 1 sequences selected [GenPept](#) [Graphics](#) [Distance tree of results](#) [Multiple alignment](#)

Description	Max Score	Total Score	Query Cover	E value	Per Ident	Accession
<input checked="" type="checkbox"/> hypothetical protein SEA_VORRPS_7 [Mycobacterium phage Vorrps]	59.7	59.7	96%	5e-11	100.00%	AYG98844.1

## TMHMM result

[HELP](#) with output formats

```
# WEBSEQUENCE Length: 31
# WEBSEQUENCE Number of predicted THMs: 1
# WEBSEQUENCE Exp number of AAs in THMs: 21.7729
# WEBSEQUENCE Exp number, first 60 AAs: 21.7729
# WEBSEQUENCE Total prob of N-in: 0.89853
# WEBSEQUENCE POSSIBLE N-term signal sequence
# WEBSEQUENCE TMHMM2.0 Inside 1 6
# WEBSEQUENCE TMHMM2.0 THelix 7 29
# WEBSEQUENCE TMHMM2.0 Outside 30 31
```



# [plot](#) in postscript, [script](#) for making the plot in gnuplot, [data](#) for plot

← → ↻ ⓘ Not secure | harrier.nagahama-i-bio.ac.jp/sosui/cgi-bin/adv\_sosui.cgi

## SOSUI Result

Query title : None

Total length : 30 A. A.

Average of hydrophobicity : 1.36667

**This amino acid sequence is of a MEMBRANE PROTEIN which have 1 transmembrane helix.**

No.	N terminal	transmembrane region	C terminal	type	length
1	5	WAEFVLRGLAATAILTALACLG	27	PRIMARY	23