

All BLAST algorithms are available on the web site of ncbi at: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>

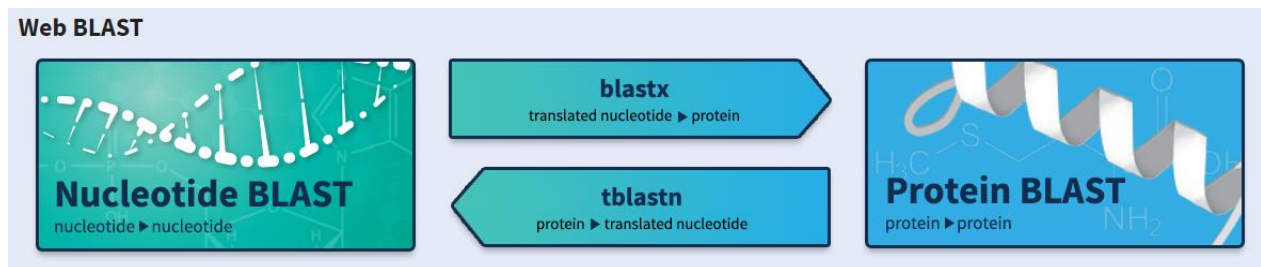


Fig 1: Main BLAST programs.

To use, BLASTx (nucleotide against protein), one must double-click.

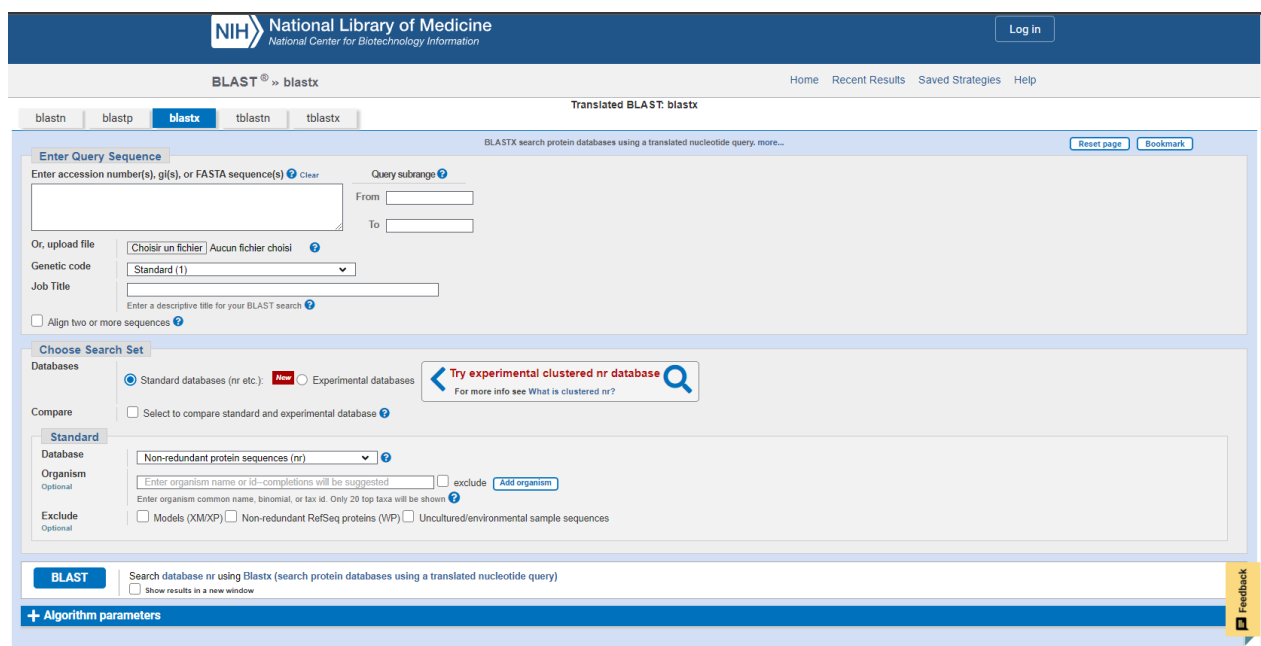


Fig 2: BLAST sequence submission interface.

By switching from one tab (i.e. 'blastx', on the figure) to another, one can access the other BLAST programs.

Many databases are proposed. By default, nt nucleotides collections as nucleotide database and nr as protein database. New experimental databases are present and can be relevant, depending on use cases. Some of these databases are smaller and can reduce BLAST processing duration.

BLAST searches can also be restricted to given taxa making it faster. Some other can be excluded from search.

Advanced parameters can be used to adjust searches.

All parameters are described in the corresponding window to get more detailed information.

Once a search is successfully finished, we get a new page with results. If homologous sequences are found, they will be sorted, in increased order of e-value, by default. Many filters are present to focus on part of these results.

Multiple sequences alignments or a distance-based tree can be obtained from **BLAST** outputs.

A detailed “how to” is available at [https://github.com/stirera/d-chimer\\_v1?tab=readme-ov-file#6-appendix--install-blast-programs-and-databases](https://github.com/stirera/d-chimer_v1?tab=readme-ov-file#6-appendix--install-blast-programs-and-databases) to get blast programs and databases for a command line usage.

--

Let's illustrate with a use case:

```
>set3_133_k141_72529_flag_1_multi_11.0000_len_1871
```

```
TTTTTTTTATTTTTTAAATTTCCCTCCCCCTGCGCGCCTTTGGGGTTTTCTCGCGCGCGTGGGTGTCTGCTGGTGG
GGGGTCCCGTGGGCGCGCCCCCTTTGGGTTCGCCGCGGCCGCGAGGGCGTTGCACCAGGGCAGAGGGT
CGCACTCCTCGCTGACGCTCGGAGTGCTTAGTGACCTTAAGAGGTAGACACCGACCACTGAGACTCCCACGGCCTT
GTTTATTTGAAACCGAGCTTGAAGCTCACCTTTATTTACAGGGTTTGGGGTGGGGAGCCAGGGGTAGATAGGTGTGT
CTGTGAGGTACTTTCTGGGGGGCTGTGAAGGCCTGGCAGGCCTGAACTCAGTGTTCATCCGTGGTATCCTTG
TATGGGAGTAGGTTTGTGGTGGAAATAGGAACACTTCGTTTATTGCGGGAAGGATAACAGGGGGTTTACGTGG
AGATTCGCCTGGGTCTTGCGAGTTGGAGCGCGAACTGCTGGAGTTGGAGCTGGAGTGCTCGCTGCTCTCGGAGC
TTTCGCTGGAGGAGCTCCCTCGTCGTCTCCGCCGGTCTCTCCGAGGAGAGGCTCGGCGTCTGGGTCTCTTCGA
ACGAGTACCGGGGTCTTTGACCGGTGAACCTGAGCCGCTTTCGACAGCGCCGTCTTCTCGTTGGTATTCTGAGGG
TAGAAAGAGGAGGGGTGCCCTGAGGGTTGAAGCAACATACTCGCCAGGTATTGGTTGCTTGAGCAGACGTCTGA
GAGTTGTCTCAGTAAAGAGGCCTCTTCTCTGGTGAAGTTGTGGAGGACCCACGGGGGCGTGACGAGCTGCGGGT
CAACGACTTGTAATTCGCGAGGCTGTGACGGGGTCCGGATACTTGGTTGTCTTTCACGGGTTTCTGATGATCTG
AGAGTGGACAGTGTTACCTCCCCACATGAACCTGAAGGAGTACTTGGCCACTAGCTGGCTGTTTTGTTATCGTCTC
TGTAAGTCCAGGGGGCGCAGGTTACTATGTGATGGAATACGTGAGTCTGGTATCTTAGCATGACATACCACTTTGTC
CTCTCGTGCAGGGGGATGTAGCCCCCTCCGTCTATCCATTTGCCCTGCCAAAATTTCTGTTGTAGACTACGAAGCC
TGCAAGTGTGCTTTTCTGTCTACGAGGGGCGGCTTGGTGTAGGGACACCTGATGAGAACTTTCCTTCTTGCTGA
GGTCGCCCCACTGTATTTGAGTCTCTAGGTAGTCTGCATACCCGTTTAGGAGGGGCCACAGGGGCATGTCTTTGAT
CTCCAGTGGCACTCTGGTTTTTGAAGTGTGTGCTTTAGTGTGGTACTGGAACCATACCATGTTACCTCGGC
CTGTGCTAAGAAGGGATTGTATACTATGTCTCTGTAGGCAGTAAAGAATTCAATGTTGGATCTGCCATCTGAGAGG
```

AATATGGGGCTGTAAAATCCCAGGTAGTAGTCTAGCATGTGAGTGTTGCCTGCATATTGAGCTGTGCTGTCTCCAT  
 ATAGGTCATTGGGGCTGTGCGCCAGAGCGAAGTTGTTATCTCTCATCGTTTTAATTGTTGATTTTATTTTGTGAGGGT  
 GATAGGTGGCAAAAAGAAACAGGCTGTCTCTTGTGTGTCAGTGAAGCTGTTTTCCAACATTTAACTGGTCCTTAGTA  
 GGTTTTGTGGGAAAGGGCTCAGTGTTATAGGTAGGTGAAGTGGCATCAGTTTGAGCCTTCTTTGCTTTGTCTGCAT  
 CTATGTGGGCTCTACCATGCTGTTAGTTATGAAGGTTGCCAGTAGTGGCCAGATATTTTTTTAGTGAATAGTTCG  
 TTATATAATTCTTTTAAGTTGTTTTTATTGGGCTTATCATAG

- Copy paste the DNA sequence into the appropriate window.
- Choose a database conveniently, here standar nr database.
- Launch the BLASTx search with the “BLAST” labelled blue button.

BLAST<sup>®</sup> » blastx

Translated BLAST: blastx

blastn blastp **blastx** tblastn tblastx

BLASTX search protein databases using a translated nucleotide query. more...

Reset page Bookmark

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) [Clear](#)

GCATGTGAGTGTGGCTGCATATTGAGCTGTGCTCTCCATAGGTCAT  
 TGGGGCCTGTGCGCCAGAGCGAAGTTGTTATCTCTCATCGTTTTAATTGTT  
 GATTTTATTTGTCAAGGTGATAGGTGCAAAAAGAAACAGGCTGTCTCTT  
 GTGTCAGTGAAGCTGTTTTCCAACTATAAAGTGTCTTAGTAGGTTTT  
 GTGGCAAGGGCTCAGTGTATAGTGAAGTGGCATCAGTTTGAG  
 CCTCTTTGCTTTGTCTGCATCTATGGGCTCACCATGCTGTAGTTA  
 TGAAGGTTGCCAGTAGTGGCCAGATATTTTTAGTGAATAGTTCGTATA  
 TAATCTTTTAAGTTGTTTTATTGGGCTTATCATAG

Query subrange

From

To

Or, upload file  No file chosen

Genetic code

Job Title

Enter a descriptive title for your BLAST search

☐ Align two or more sequences

Choose Search Set

Databases ☒ Standard databases (nr etc.) [New](#) ☐ Experimental databases [Try experimental clustered nr database](#)  
 For more info see [What is clustered nr?](#)

Compare ☐ Select to compare standard and experimental database

Standard

Database

Organism  ☐ exclude [Add organism](#)  
 Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown

Exclude ☐ Models (XM/XP) ☐ Non-redundant RefSeq proteins (WP) ☐ Uncultured/environmental sample sequences

**BLAST** Search database nr using Blastx (search protein databases using a translated nucleotide query)  
☐ Show results in a new window

+ Algorithm parameters

After the BLASTx search completed, a new window with results is brought. Generally the closest homologous sequence is at top; here the sequence with accession ‘UGV43885.1’.

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

An official website of the United States government [Here's how you know](#)

**NIH** National Library of Medicine  
National Center for Biotechnology Information

Log in

BLAST® » blastx » results for RID-YBGZJWB0016

Home Recent Results Saved Strategies Help

[< Edit Search](#) [Save Search](#) [Search Summary](#) [How to read this report?](#) [BLAST Help Videos](#) [Back to Traditional Results Page](#)

Job Title: **Test\_sequence\_AnssBin**  
 RID: **YBGZJWB0016** Search expires on 03-05 19:52 pm [Download All](#)  
 Program: BLASTX [Citation](#)  
 Database: nr [See details](#)  
 Query ID: lclQuery\_7073021  
 Description: Test\_sequence\_AnssBin  
 Molecule type: dna  
 Query Length: 1871  
 Other reports: [?](#)

**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](#)

Compare these results against the new Clustered nr database [BLAST](#)

**Descriptions** Graphic Summary Alignments Taxonomy

**Sequences producing significant alignments** Download Select columns Show 100 [?](#)

☒ select all 100 sequences selected [GenPept](#) [Graphics](#)

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	ORF1 [Torque teno virus]	Torque teno virus	849	849	77%	0.0	96.06%	779	UGV43885.1
<input checked="" type="checkbox"/>	ORF1 [Torque teno virus]	Torque teno virus	843	843	77%	0.0	95.02%	779	UGV43895.1
<input checked="" type="checkbox"/>	ORF1 [Torque teno virus]	Torque teno virus	843	843	77%	0.0	95.23%	779	UGV43913.1
<input checked="" type="checkbox"/>	ORF1 [Torque teno virus]	Torque teno virus	840	840	77%	0.0	94.81%	779	UGV43909.1

Many options are possible. Results are downloadable in many csv format for example or aligned sequences in fasta format among others.

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

BLAST® » blastx » results for RID-YBGZJWB0016

Home Recent Results Saved Strategies Help

[< Edit Search](#) [Save Search](#) [Search Summary](#) [How to read this report?](#) [BLAST Help Videos](#) [Back to Traditional Results Page](#)

Job Title: **Test\_sequence\_AnssBin**  
 RID: **YBGZJWB0016** Search expires on 03-05 19:52 pm [Download All](#)  
 Program: BLASTX [Citation](#)  
 Database: nr [See details](#)  
 Query ID: lclQuery\_7073021  
 Description: Test\_sequence\_AnssBin  
 Molecule type: dna  
 Query Length: 1871  
 Other reports: [?](#)

**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](#)

Compare these results against the new Clustered nr database [BLAST](#)

**Descriptions** **Graphic Summary** **Alignments** Taxonomy

Alignment view: Pairwise [Restore defaults](#) Download

100 sequences selected [Download](#) [GenPept](#) [Graphics](#)

**MAG: ORF1 [Torque teno virus]**  
 Sequence ID: [UGV43885.1](#) Length: 779 Number of Matches: 1

Range 1: 298 to 779 [GenPept](#) [Graphics](#)

Score	Expect	Method	Identities	Positives	Gaps	Frame
849 bits(2193)	0.0	Compositional matrix adjust.	463/482(96%)	470/482(97%)	0/482(0%)	-2

Query 1870 YDKPknknlkelyneLFTKKISGHYWQTFITNSMVRahIdadkakkqtdaTSPTYNTEP 1691  
 Y+KPNK NLKELY ELFTKKISGHYWQTFITNSMVRahIDADKAK+AQT+ATSPTYNTEP  
 Sbjct 298 YNKPKNKELKELYKELFTKKISGHYWQTFITNSMVRahIDADKAKEQTNATSPTYNTEP 357  
 Query 1690 FPTKPTKQDFNSKNSFTDTRDSLFLFATYHPDKIKSTIKTMRDINFALATGPNLDYGS 1511

FASTA (completesequences)  
 FASTA (aligned sequences)  
 GenBank (completesequences)  
 Hit Table (text)  
 Hit Table (CSV)  
 Text  
 XML  
 ASN.1