

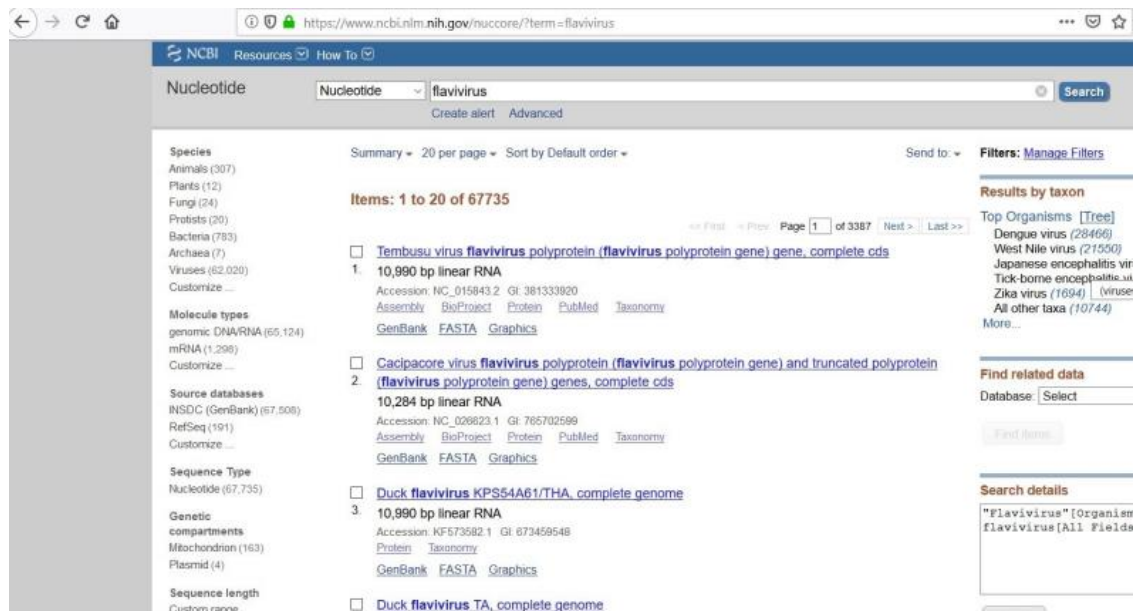
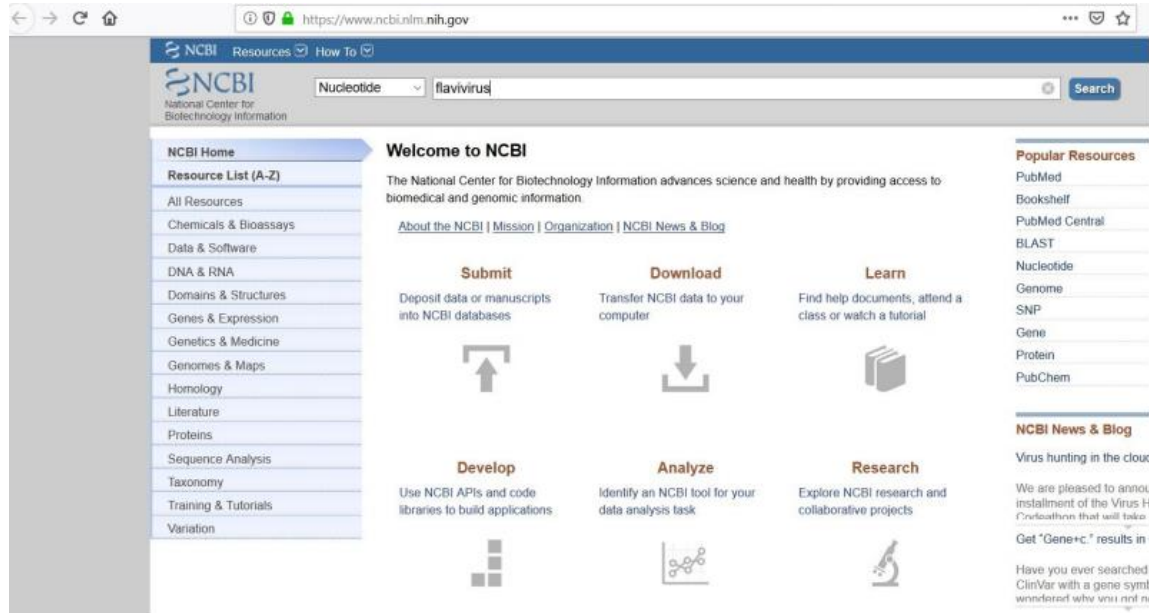
Aim: Perform a BLAST search on a specific gene sequence of a specific organism.

Steps:

Go to the National Center for Biotechnology Information Site

<https://www.ncbi.nlm.nih.gov/>

Select Nucleotide from All Databases and find any organism in a search bar



NCBI Nucleotide database page for NC\_015843.2. The page displays the following information:

- GenBank** (Send to)
- Change region shown**
- Customize view**
- Analyze this sequence** (Run BLAST, Pick Primers, Highlight Sequence Features, Find in this Sequence)
- Related information** (Assembly, BioProject, Protein, PubMed, Taxonomy, Full text in PMC, Functional Class)
- FASTA** (Graphics)
- Go to** (Go to)
- LOCUS** NC\_015843 10990 bp ss-RNA linear VRL 13-AUG-2018
- DEFINITION** Tembusu virus flavivirus polyprotein (flavivirus polyprotein gene) gene, complete cds.
- ACCESSION** NC\_015843 NC\_016958 NC\_018670
- VERSION** NC\_015843.2
- DBLINK** BioProject: PRJNA485481
- KEYWORDS** RefSeq.
- SOURCE** Tembusu virus (TMUV)
- ORGANISM** Tembusu virus
- REFERENCE** Viruses; Riboviria; Flaviviridae; Flavivirus.
- AUTHORS** 1 (bases 1 to 10990) Han, K., Huang, X., Li, Y., Zhao, D., Liu, Y., Zhou, X., You, Y. and Xie, X.
- TITLE** Complete genome sequence of goose tembusu virus, isolated from jiangnan white geese in jiangsu, china
- JOURNAL** Genome Announc 1 (2), E0023612 (2013)
- PUBMED** 23516233
- REMARK** Publication Status: Online-Only

Run BLAST option we have to select

NCBI BLAST search interface. The 'Choose Search Set' section shows the following options:

- Database**: ☐ Human genomic + transcript ☐ Mouse genomic + transcript ☒ Others (nr etc.)
- Organism**:
- Exclude**: ☐ Models (XM/XP) ☐ Uncultured/environmental sample sequences
- Limit to**: ☐ Sequences from type material
- Entrez Query**:

The 'Program Selection' section shows the following options:

- Optimize for**: ☒ Highly similar sequences (megablast) ☐ More dissimilar sequences (discontiguous megablast) ☐ Somewhat similar sequences (blastn)
- Choose a BLAST algorithm**:

The 'BLAST' button is highlighted. Below it, the search parameters are displayed: Search database Nucleotide collection (nr/nt) using Megablast (Optimize for highly similar sequences). The 'Show results in a new window' checkbox is checked.

Algorithm parameters

BLAST is a registered trademark of the National Library of Medicine

BLAST shows in screen

Sequences producing significant alignments		Download	Manage Columns	Show
select all 99 sequences selected		GenBank	Graphics	Dis
Description	Max Score	Total Score	Query Cover	E value
Tembusu virus strain JS804, complete genome	20064	20064	99%	0.0
Tembusu virus strain JS/2010, complete genome	20064	20064	99%	0.0
Duck egg-drop syndrome virus strain byd1, complete genome	20048	20048	99%	0.0
Tembusu virus isolate Tembusu virus strain, complete genome	20026	20026	99%	0.0
Duck Tembusu virus isolate df-2, complete genome	20020	20020	99%	0.0
Duck egg-drop syndrome virus strain JXSP, complete genome	20020	20020	99%	0.0
Tembusu virus isolate YY5, complete genome	20015	20015	99%	0.0
Tembusu virus isolate SDMS, complete genome	20009	20009	99%	0.0
Tembusu virus isolate ZJ-6, complete genome	20009	20009	99%	0.0
Tembusu virus strain AH-F10 from China, complete genome	20004	20004	99%	0.0
Duck egg-drop syndrome virus strain pigeon, complete genome	20004	20004	99%	0.0
Tembusu virus genomic RNA, complete genome, strain TMUV-YY1Du	19998	19998	99%	0.0
Duck Tembusu virus strain BZ, 2010, complete genome	19998	19998	99%	0.0
Duck egg-drop syndrome virus strain duan, complete genome	19998	19998	99%	0.0
Duck Tembusu virus strain GD1H01, complete genome	19989	19989	99%	0.0

Here the result will be display

Tembusu virus strain JS804, complete genome		Download	GenBank	Graphics	Next Match	Previous Match	Related Information
Sequence ID: JF895923.2 Length: 10990 Number of Matches: 1							
Range 1: 1 to 10990							
Score	Expect	Identities	Gaps	Strand			
20295 bits(10990)	0.0	10990/10990(100%)	0/10990(0%)	Plus/Plus			
Query 1	AGAAATTGCGCTGTGTGAACCTTATCCAAACAGCTTTTGGAGTAGTGCCTGTGAACGTAA	60					
Sbjct 1	AGAAATTGCGCTGTGTGAACCTTATCCAAACAGCTTTTGGAGTAGTGCCTGTGAACGTAA	60					
Query 61	ACACAGTTTGAACGTTTTTGGATAGAGACAACTATGTCTAACAAAAAACAGGAAGACC	120					
Sbjct 61	ACACAGTTTGAACGTTTTTGGATAGAGACAACTATGTCTAACAAAAAACAGGAAGACC	120					
Query 121	CGGCTCAGGCCGGGTTGTCAATATGCTAAAGCGCGGAACGTCGCCGGAATCCGCTAGC	180					
Sbjct 121	CGGCTCAGGCCGGGTTGTCAATATGCTAAAGCGCGGAACGTCGCCGGAATCCGCTAGC	180					
Query 181	GCGGATAAAGAGGACGATTGATGGGGTCTGAGAGGAGCAGGACCCATAAGGTTTGTGCT	240					
Sbjct 181	GCGGATAAAGAGGACGATTGATGGGGTCTGAGAGGAGCAGGACCCATAAGGTTTGTGCT	240					
Query 241	GGCTCTACTGACTTTCTCAAGTTTACAGCCCTGAGGCCAACATTGGAATGCTGAAGAG	300					
Sbjct 241	GGCTCTACTGACTTTCTCAAGTTTACAGCCCTGAGGCCAACATTGGAATGCTGAAGAG	300					
Query 301	ATGGAAGCTGCTGGAGTTAATGAGGCGACCAACATCTGAAAAGCTTCAAGCGTGACAT	360					
Sbjct 301	ATGGAAGCTGCTGGAGTTAATGAGGCGACCAACATCTGAAAAGCTTCAAGCGTGACAT	360					
Query 361	TGGACAGATGCTCGACGGACTGAATAAGCGGAAGGCCAAACGTCGGGGGGGAGTTGCTC	420					
Sbjct 361	TGGACAGATGCTCGACGGACTGAATAAGCGGAAGGCCAAACGTCGGGGGGGAGTTGCTC	420					