

Name: Neha Kanojiya

Class- MSc CS – I

Roll No.- 536

Subject – Bioinformatics

**Topic – Perform a BLAST
search on any gene sequence**

Practical No: 5

Aim: Perform a BLAST search on a specific gene sequence of a specify 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

NIH National Library of Medicine
National Center for Biotechnology Information

Log in

Nucleotide Search

NCBI Home
[Database List \(A-Z\)](#)
[All Resources](#)
[Chemicals & Bioassays](#)
[Data & Software](#)
[DNA & RNA](#)
[Domains & Structures](#)
[Genes & Expression](#)
[Genetics & Medicine](#)
[Genomics & Maps](#)
[Homology](#)
[Literature](#)
[Proteins](#)
[Sequence Analysis](#)
[Taxonomy](#)
[Training & Tutorials](#)
[Variation](#)

Welcome to NCBI
 The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.

[About the NCBI](#) | [Mission](#) | [Organization](#) | [NCBI News & Blog](#)

Submit
 Deposit data or manuscripts into NCBI databases

Download
 Transfer NCBI data to your computer

Learn
 Find help documents, attend a class or watch a tutorial

Develop
 Use NCBI APIs and code libraries to build applications

Analyze
 Identify an NCBI tool for your data analysis task

Research
 Explore NCBI research and collaborative projects

Popular Resources
[PubMed](#)
[BioRxiv](#)
[PubMed Central](#)
[BLAST](#)
[Nucleotide](#)
[Genome](#)
[SNP](#)
[Gene](#)
[Protein](#)
[PubChem](#)

NCBI News & Blog
 NEW Streamlined ClinVar Submission of Assertion Criteria
 10 Nov 2022
 ClinVar is a freely available submission database for information about
 Re-evaluating the BLAST Nucleotide Database (nt)
 17 Nov 2022
 The ongoing sequencing revolution has resulted in exponential growth of the
 RefSeq Release 215
 18 Nov 2022

[Mosquito flavivirus NS5 gene for polyprotein, partial cds, strain YDFV/Sep/2013](#)
 4
 950 bp linear RNA
 Accession: AB81117.1 GI: 824555718
[Protein](#) [PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)

[Mosquito flavivirus gene for polyprotein, complete cds, strain YDFV/Oct/2013](#)
 5
 10,863 bp linear RNA
 Accession: AB81116.1 GI: 824555716
[Protein](#) [PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)

[Hangzhou flavivirus 3 isolate YWLFY92, complete genome](#)
 6
 9,531 bp linear RNA
 Accession: MZ29930.1 GI: 2150355026
[BioProject](#) [BioSample](#) [Protein](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)

[Tembusu virus flavivirus polyprotein \(flavivirus polyprotein gene\) gene, complete cds](#)
 7
 10,990 bp linear RNA
 Accession: NC_015843.2 GI: 381333626
[Assembly](#) [BioProject](#) [Protein](#) [PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)

[Culex flavivirus gene for polyprotein, partial cds, clone: cpcr_flavi_no5_24](#)
 8
 268 bp linear RNA
 Accession: LC227562.1 GI: 324717335
[Protein](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)

[Caciparore virus flavivirus polyprotein \(flavivirus polyprotein gene\) and truncated polyprotein \(flavivirus polyprotein gene\) genes, complete cds](#)
 9
 10,284 bp linear RNA
 Accession: NC_026923.1 GI: 755702596
[Assembly](#) [BioProject](#) [Protein](#) [PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)

https://www.ncbi.nlm.nih.gov/nuccore/NC_015843.2

NIH National Library of Medicine
National Center for Biotechnology Information

Log in

Nucleotide Search

Advanced Help

GenBank Send to

Tembusu virus flavivirus polyprotein (flavivirus polyprotein gene) gene, complete cds
 NCBI Reference Sequence: NC_015843.2
[FASTA](#) [Graphics](#)

[Go to:](#)

LOCUS NC_015843 10958 bp ss-RNA linear VRL 13-AUG-2015
 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
 KLVARKW RefSeq.
 SOURCE Tembusu virus
 ORGANISM Tembusu virus
 Viruses; Riboviria; Orthornavirae; Kitrinoviricota; Flaviviricetes; Amarillivirales; Flaviviridae; Flavivirus.

REFERENCE 1 (bases 1 to 10958)
 AUTHORS 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): e002682 (2013)
 PUBMED 23545128
 REFNOTE Publication Status: Online-Only
 REFERENCE 2 (bases 1 to 10958)
 AUTHORS Huang,X., Han,K., Zhao,D., Liu,Y., Zhang,J., Liu,H., Zhang,K.,

Change region shown
Customize view
Analyze this sequence
 Run BLAST
 Pick Primers
 Highlight Sequence Features
 Find in this Sequence

Articles about the flavivirus polyprotein gene
 Substantial Attenuation of Virulence of Tembusu Virus Strain PS Is Determined by a (J Virol. 2021)
 Identification of a Neutralizing Monoclonal Antibody That Recognizes a Unique (Viruses. 2020)
 A Single Mutation at Position 156 in the Envelope Protein of Tembusu Virus Is Respon (J Virol. 2016)
 See all

Reference sequence information
 NC_015843.2

Run BLAST option we have to select

Enter accession number(s), gis(s), or FASTA sequence(s) [?](#) [Clear](#) Query subrange [?](#)

NC_012043.2 From To

Or, upload file No file chosen [?](#)

Job Title

Enter a descriptive title for your BLAST search [?](#)

☐ Align two or more sequences [?](#)

Choose Search Set

Database ☒ Standard databases (in etc.), ☐ rRNA/ITS databases ☐ Genomic + transcript. databases ☐ RefSeq/GenBank

Organism Enter organism name or id—completions will be suggested ☐ exclude [Add organism](#)

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

☐ Models (XMP) ☐ Uncultured/environmental sample sequences

Limit to ☐ Sequences from type material

Entrez Query [YouTube](#) Create custom database

Enter an Entrez query to limit search [?](#)

Program Selection

Optimize for ☒ Highly similar sequences (megablast) ☐ More dissimilar sequences (discontiguous megablast) ☐ Somewhat similar sequences (blastn)

Choose a BLAST algorithm [?](#)

BLAST Search database Nucleotide collection (nr) using Megablast (Optimize for highly similar sequences)

☐ Show results in a new window

[+ Algorithm parameters](#)

BLAST

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Tembusu virus strain JS24 complete genome	Tembusu virus	20295	20295	100%	0.0	100.00%	10990	JF895923.2
Tembusu virus isolate SD2010 complete genome	Tembusu virus	20098	20098	100%	0.0	99.67%	10990	MN649260.1
Tembusu virus strain JS2010 complete genome	Tembusu virus	20064	20064	99%	0.0	99.64%	10990	JX273163.1
Duck egg-drop syndrome virus strain hvd1 complete genome	Duck egg-drop syndrome virus	20048	20048	99%	0.0	99.61%	10990	JQ929420.1
Tembusu virus isolate Tembusu virus strain complete genome	Tembusu virus	20026	20026	99%	0.0	99.67%	10989	KF162981.1
Duck Tembusu virus isolate H-2 complete genome	Duck Tembusu virus	20020	20020	99%	0.0	99.56%	10990	KJ485355.1
Duck egg-drop syndrome virus strain JXSD complete genome	Duck egg-drop syndrome virus	20020	20020	99%	0.0	99.56%	10990	JQ929421.1
Tembusu virus isolate HB2010 complete genome	Tembusu virus	20018	20018	100%	0.0	99.55%	10990	MN649262.1
Tembusu virus isolate YYS complete genome	Tembusu virus	20015	20015	99%	0.0	99.55%	10990	JF272406.1
Tembusu virus isolate SDMS complete genome	Tembusu virus	20009	20009	99%	0.0	99.54%	10990	KC333637.1
Tembusu virus isolate Z-48 complete genome	Tembusu virus	20009	20009	99%	0.0	99.54%	10990	JF459991.1
Tembusu virus strain AH-F11 from China complete genome	Tembusu virus	20004	20004	99%	0.0	99.54%	10990	KM112528.1
Duck egg-drop syndrome virus strain pison complete genome	Duck egg-drop syndrome virus	20004	20004	99%	0.0	99.54%	10990	JQ929425.1
Tembusu virus genomic RNA complete genome strain TMUV-YY10u	Tembusu virus	19998	19998	99%	0.0	99.53%	10990	AB917038.1
Duck Tembusu virus strain B7_2010 complete genome	Duck Tembusu virus	19998	19998	99%	0.0	99.53%	10990	KC950540.1
Duck egg-drop syndrome virus strain duan complete genome	Duck egg-drop syndrome virus	19998	19998	99%	0.0	99.53%	10990	JQ929421.1
Duck Tembusu virus strain GD1101 complete genome	Duck Tembusu virus	19989	19989	99%	0.0	99.51%	10990	KT624876.1
Tembusu virus isolate pY150202 polyprotein gene complete cds	Tembusu virus	10001	10001	99%	0.0	99.60%	10990	MF522175.1

Here the result will be display

Download ▾ GenBank Graphics

▼ Next ▲ Previous ◀ Descriptions

Tembusu virus isolate SD2010, complete genome

Sequence ID: [MN649260.1](#) Length: 10990 Number of Matches: 1

Range 1: 1 to 10990 [GenBank](#) [Graphics](#) [Hard Mask](#) [Low-complex Mask](#)

Score	Expect	Identities	Gaps	Strand
20098 bits(10893)	0.0	10054/10990(99%)	0/10990(0%)	Plus/Plus
Query 1	AGAAGTTCCCTGTGTGAACATTATCCAAACAGCTTTTSGAGTAGTGTGTGAACGTAA	60		
Subject 1	AGAAGTTTCATCTGTGTGAACATTATCCAAACAGCTTTTSGAGTAGTGTGTGAACGTAA	60		
Query 61	ACACAGCTTTGAACGTTTTTGGATAGAGCACTATGCTTAACAAAACACAGGAAGACC	120		
Subject 61	ACACAGCTTTGAACGTTTTTGGATAGAGCACTATGCTTAACAAAACACAGGAAGACC	120		
Query 121	CAGCTCAGGCGGGTTGTCAATATGCTAAGGCGGGAACGTCCCGGGAATCCCTCAGC	180		
Subject 121	CAGCTCAGGCGGGTTGTCAATATGCTAAGGCGGGAACGTCCCGGGAATCCCTCAGC	180		
Query 181	GCGGATAAAGAGGACGATTCATGGGCTCTGAGAGGACAGGACCATTAAGCTTTGTCT	240		
Subject 181	GCGGATAAAGAGGACGATTCATGGGCTCTGAGAGGACAGGACCATTAAGCTTTGTCT	240		
Query 241	GGCTCTACTGACTTTTTCAGTTTACAGCCCTGAGGCCAACCATTTGGAATGCTSAAGAG	300		
Subject 241	GGCTCTACTGACTTTTTCAGTTTACAGCCCTGAGGCCAACCATTTGGAATGCTSAAGAG	300		
Query 301	ATGGAAGCTGGTTGGAGTTAATGAGCGGACCAACATCTGAAGAGCTTCAAGCGTGACAT	360		
Subject 301	ATGGAAGCTGGTTGGAGTTAATGAGCGGACCAACATCTGAAGAGCTTCAAGCGTGACAT	360		
Query 361	TGGACAGATGCTCGACGGACTGAATAGCGGAAGGCGAAGCTCGGGGGAGTTGCTC	420		
Subject 361	TGGACAGATGCTCGACGGACTGAATAGCGGAAGGCGAAGCTCGGGGGAGTTGCTC	420		
Query 421	TTGGATCATTATGTTACTCCCGATAGTTGCTGGGCTGAAGCTTGGAACTATAATGGTG	480		
Subject 421	TTGGATCATTATGTTACTCCCGATAGTTGCTGGGCTGAAGCTTGGAACTATAATGGTG	480		
Query 481	AGTTTGGGCTATTTAAATAGAGCGGATCTATCAGACTTCTAGTCACTCCAAACAGCC	540		
Subject 481	AGTTTGGGCTATTTAAATAGAGCGGATCTATCAGACTTCTAGTCACTCCAAACAGCC	540		
Query 541	TGGCAGCAATGGATGGCTCTAGCTGCTCTAGATGGGACTAATGTTGAGGATGACAT	600		