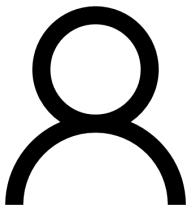


The National Center for Biotechnology Information (NCBI)

SC113603 Molecular Biology
Lect. Todsapol Techo

Introduction of NCBI

- **The National Center for Biotechnology Information (NCBI)**
- Act like a portal to access all biological database including DNA, Protein, and etc.
 - Nucleotide sequence Database
 - Sanger Sequencing
 - The International Nucleotide Sequence Database Collaboration (**INSDC**)
 - GenBank (NCBI), EMBL-EBI, and DDBJ
 - NGS
 - Sequence Read Archive (**SRA**)
 - Protein sequence and structure
 - Protein Data Bank (**PDB**) and Uniport
 - PubMed
 - Abstract and Citation of scientific articles
- Provide **web-based tools** for sequence analysis



Pub**Med**.gov

DDBJ
DNA Data Bank of Japan

ENA
European Nucleotide Archive

WORLDWIDE
PDB
PROTEIN DATA BANK

UniProt

Web-based Tools

- BLAST
- COBALT
- etc.

Analyze

NCBI provides a wide variety of data analysis tools that allow users to manipulate, align, visualize and evaluate biological data.

Selected Analysis Tools

All Tools	Literature	Health	Genomes	Genes	Proteins	Chemicals																							
Filter this table <input type="text"/>																													
<table border="1"><thead><tr><th>Tools</th><th>Description</th></tr></thead><tbody><tr><td>Amino Acid Explorer</td><td>Explores amino acid properties, substitutions and functions</td></tr><tr><td>Assembly Archive</td><td>Links the raw sequence information found in the Trace Archive with assembly information found in GenBank/EMBL/DDBJ</td></tr><tr><td>Basic Local Alignment Search Tool (BLAST)</td><td>Finds regions of local similarity between biological sequences</td></tr><tr><td>Batch Entrez</td><td>Retrieves records specified in an uploaded file of identifiers</td></tr><tr><td>BioAssay Services</td><td>Tools that summarize the biological test results in the PubChem database</td></tr><tr><td>BLAST Link (BLink)</td><td>Displays the results of a pre-computed BLAST search of a protein against all other protein sequences at NCBI</td></tr><tr><td>BLAST Microbial Genomes</td><td>Finds regions of local similarity between query sequences and sequences from complete microbial genomes</td></tr><tr><td>BLAST RefSeqGene</td><td>Finds regions of local similarity between query sequences and genomic sequences in the RefSeqGene/LRG set</td></tr><tr><td>CDTree</td><td>Classifies protein sequences and investigates their evolutionary relationships</td></tr><tr><td>Cn3D</td><td>Displays and manipulates 3-dimensional structures and alignments from the Structure database</td></tr><tr><td>COBALT</td><td>Performs protein multiple sequence alignments</td></tr></tbody></table>						Tools	Description	Amino Acid Explorer	Explores amino acid properties, substitutions and functions	Assembly Archive	Links the raw sequence information found in the Trace Archive with assembly information found in GenBank/EMBL/DDBJ	Basic Local Alignment Search Tool (BLAST)	Finds regions of local similarity between biological sequences	Batch Entrez	Retrieves records specified in an uploaded file of identifiers	BioAssay Services	Tools that summarize the biological test results in the PubChem database	BLAST Link (BLink)	Displays the results of a pre-computed BLAST search of a protein against all other protein sequences at NCBI	BLAST Microbial Genomes	Finds regions of local similarity between query sequences and sequences from complete microbial genomes	BLAST RefSeqGene	Finds regions of local similarity between query sequences and genomic sequences in the RefSeqGene/LRG set	CDTree	Classifies protein sequences and investigates their evolutionary relationships	Cn3D	Displays and manipulates 3-dimensional structures and alignments from the Structure database	COBALT	Performs protein multiple sequence alignments
Tools	Description																												
Amino Acid Explorer	Explores amino acid properties, substitutions and functions																												
Assembly Archive	Links the raw sequence information found in the Trace Archive with assembly information found in GenBank/EMBL/DDBJ																												
Basic Local Alignment Search Tool (BLAST)	Finds regions of local similarity between biological sequences																												
Batch Entrez	Retrieves records specified in an uploaded file of identifiers																												
BioAssay Services	Tools that summarize the biological test results in the PubChem database																												
BLAST Link (BLink)	Displays the results of a pre-computed BLAST search of a protein against all other protein sequences at NCBI																												
BLAST Microbial Genomes	Finds regions of local similarity between query sequences and sequences from complete microbial genomes																												
BLAST RefSeqGene	Finds regions of local similarity between query sequences and genomic sequences in the RefSeqGene/LRG set																												
CDTree	Classifies protein sequences and investigates their evolutionary relationships																												
Cn3D	Displays and manipulates 3-dimensional structures and alignments from the Structure database																												
COBALT	Performs protein multiple sequence alignments																												

Search Systems in NCBI

- To Find the interesting things in NCBI-related database (Mostly Gene)
- **Two Main Search Systems on NCBI**
 - Entrez search systems (Like Google)
 - Basic Local Alignment Search Tool (BLAST)

National Library of Medicine
National Center for Biotechnology Information

All Databases ▾

Welcome to NCBI

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

Basic Local Alignment Search Tool

BLAST finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance. [Learn more](#)

NEWS

BLAST+ 2.13.0 is here!

Starting with this release, we are including the blastn_vdb and tblastn_vdb executables in the BLAST+ distribution.

Thu, 17 March 2022

[More BLAST news...](#)

Web BLAST

Nucleotide BLAST
nucleotide ► nucleotide

blastx
translated nucleotide ► protein

tblastn
protein ► translated nucleotide

Protein BLAST
protein ► protein

Entrez Search System

- General search to obtain Interesting Information from NCBI
- Narrow down your interesting information
- Usually start with the name of interesting gene, protein, or etc.

Results found in 30 databases

GENE

G6PD – glucose-6-phosphate dehydrogenase

Homo sapiens (human)

Also known as: G6PD1

Gene ID: 2539

RefSeq transcripts (3) RefSeq proteins (3) RefSeqGene (1) PubMed (477)

Orthologs Genome Data Viewer BLAST Download

RefSeq Sequences +

Literature

Bookshelf	450
MeSH	17
NLM Catalog	9
PubMed	6,510
PubMed Central	14,676

Genes

Gene	1,905
GEO DataSets	25
GEO Profiles	6,579
HomoloGene	2
PopSet	107

Proteins

Conserved Domains	1
Identical Protein Groups	255
Protein	52,112
Protein Family Models	12
Structure	43

Cancer

Clinical

PubChem

Search Box

Results found in 30 databases

Number of database matched search text

GENE

G6PD – glucose-6-phosphate dehydrogenase*Homo sapiens* (human)

Also known as: G6PD1

Gene ID: 2539

RefSeq transcripts (3) RefSeq proteins (3) RefSeqGene (1) PubMed (477)

[Orthologs](#)[Genome Data Viewer](#)[BLAST](#)[Download](#)

Was this helpful?



RefSeq Sequences

Get transcript, protein, or gene sequences

RefSeq transcripts

G6PD – 3 of 3 transcripts

Transcript	Isoform	Len (nt)
NM_000402.4	a	2,406
NM_001042351.3	b	2,267
NM_001360016.2	b	2,223

[View full table](#)

NCBI Datasets

The information of transcripts

Results from other database

Literature

Bookshelf
MeSH
NLM Catalog
PubMed
PubMed Central

Genomes

Assembly
BioCollections
BioProject
BioSample
Genome
Nucleotide
SRA
Taxonomy

Clinical

ClinicalTrials.gov
ClinVar
dbGaP
dbSNP
dbVar
GTR
MedGen
OMIM

Genes

Gene
GEO DataSets
GEO Profiles
HomoloGene
PopSet

Proteins

Conserved Domains
Identical Protein Groups
Protein
Protein Family Models
Structure

PubChem

BioAssays
Compounds
Pathways
Substances

GENE

Was this helpful?



G6PD – glucose-6-phosphate dehydrogenase

Homo sapiens (human)

Also known as: G6PD1

Gene ID: 2539

[RefSeq transcripts](#) (3) [RefSeq proteins](#) (3) [RefSeqGene](#) (1) [PubMed](#) (477)

[Orthologs](#)

[Genome Data Viewer](#)

[BLAST](#)

[Download](#)

RefSeq Sequences

Showing 3 of 3 (by status, accession number)

Transcript	nt	Protein	aa	Isoform	Status
NM_001360016.2	2,223	NP_001346945.1	515	b	MANE SELECT
NM_000402.4	2,406	NP_000393.4	545	a	curated
NM_001042351.3	2,267	NP_001035810.1	515	b	curated

[View full table](#)

More options

Search the Nucleotide and Protein databases for more sequences

[RefSeq transcripts](#) (3)

[Archival GenBank transcript sequences](#) (13)

[RefSeq proteins](#) (3)

[Archival GenBank protein sequences](#) (148)

RefSeq transcripts

[G6PD](#) – 3 of 3 transcripts

Transcript	Isoform	Len (nt)
NM_000402.4	a	2,406
NM_001042351.3	b	2,267
NM_001360016.2	b	2,223

[View full table](#)

NCBI Datasets

- **RefSeq Database**
 - Reference sequence for each gene

- **Example sequences**
 - NM_001360016.2: Transcript
 - NP_001346945.1: Protein
 - NG_009015.2: Gene
 - NC_000023.11: Chromosome (Assembly)

Genome Data Viewer

Home Share this page Reset All More Info...

Homo sapiens (human)

Search assembly

Location, gene or phenotype

Examples ►

Assembly: GRCh38.p14 (GCF_000001405.40) • Chr X (NC_000023.11)

NC_000023.11: 154,529,769 - 154,549,205



Assemblies

Assembly

Switch organism

GCF_000001405.40 (GRCh38.p14)

Select an assembly to change view

Locations for Gene G6PD

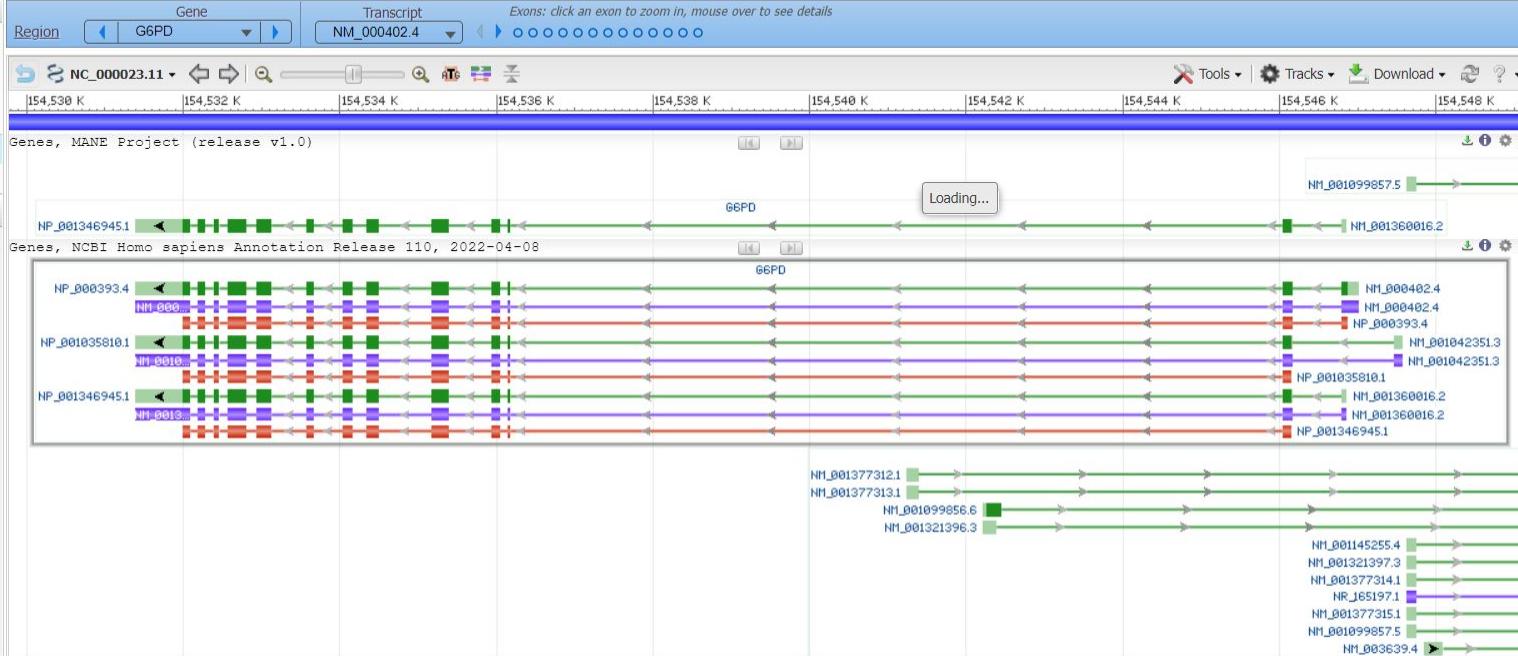
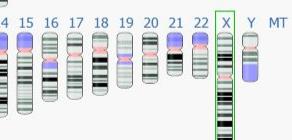
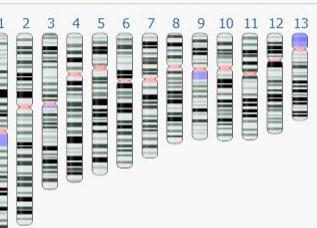
Sequence Location

NC_000023.11 154,531,390 - 154,547,569

Ideogram View

Unplaced/unlocalized scaffolds: 166

Alt loci/patches: 514



GenBank ▾

Send to: ▾

Change region shown

Homo sapiens glucose-6-phosphate dehydrogenase (G6PD), RefSeqGene on chromosome X

NCBI Reference Sequence: NG_009015.2

[FASTA](#) [Graphics](#)Go to:

LOCUS NG_009015 23182 bp DNA linear PRI 15-DEC-2022
 DEFINITION Homo sapiens glucose-6-phosphate dehydrogenase (G6PD), RefSeqGene
 on chromosome X.

RefSeqGene Accession

VERSION NG_009015.2
 KEYWORDS RefSeq; RefSeqGene.
 SOURCE Homo sapiens (human)
 ORGANISM [Homo sapiens](#)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
 Catarrhini; Hominidae; Homo.

COMMENT REVIEWED [REFSEQ](#): This record has been curated by NCBI staff in
 collaboration with Sujatha Mohan. The reference sequence was
 derived from [AC244090.3](#).
 This sequence is a reference standard in the [RefSeqGene](#) project.

On Jan 17, 2014 this sequence version replaced [NG_009015.1](#).

Summary: This gene encodes glucose-6-phosphate dehydrogenase. This protein is a cytosolic enzyme encoded by a housekeeping X-linked gene whose main function is to produce NADPH, a key electron donor in the defense against oxidizing agents and in reductive biosynthetic reactions. G6PD is remarkable for its genetic diversity. Many variants of G6PD, mostly produced from missense mutations, have been described with wide ranging levels of enzyme activity and associated clinical symptoms. G6PD deficiency may cause neonatal jaundice, acute hemolysis, or severe chronic

Customize view

Analyze this sequence

Run BLAST

Pick Primers

[Highlight Sequence Features](#)

Find in this Sequence

Articles about the G6PD gene

Extracellular vesicles-transferred SBSN drives glioma aggressiveness by activ [Oncogene. 2022]

Glucose-6-phosphate Dehydrogenase (G6PD) A-Variant Frequency an [Am J Trop Med Hyg. 2022]

Suppression of ACE2 SUMOylation protects against SARS-CoV-2 infectio [Nat Commun. 2022]

See all...

Reference sequence information

RefSeq alternative splicing

See 3 reference mRNA sequence splice variants for the G6PD gene.

More about the G6PD gene

This gene encodes glucose 6 phosphate

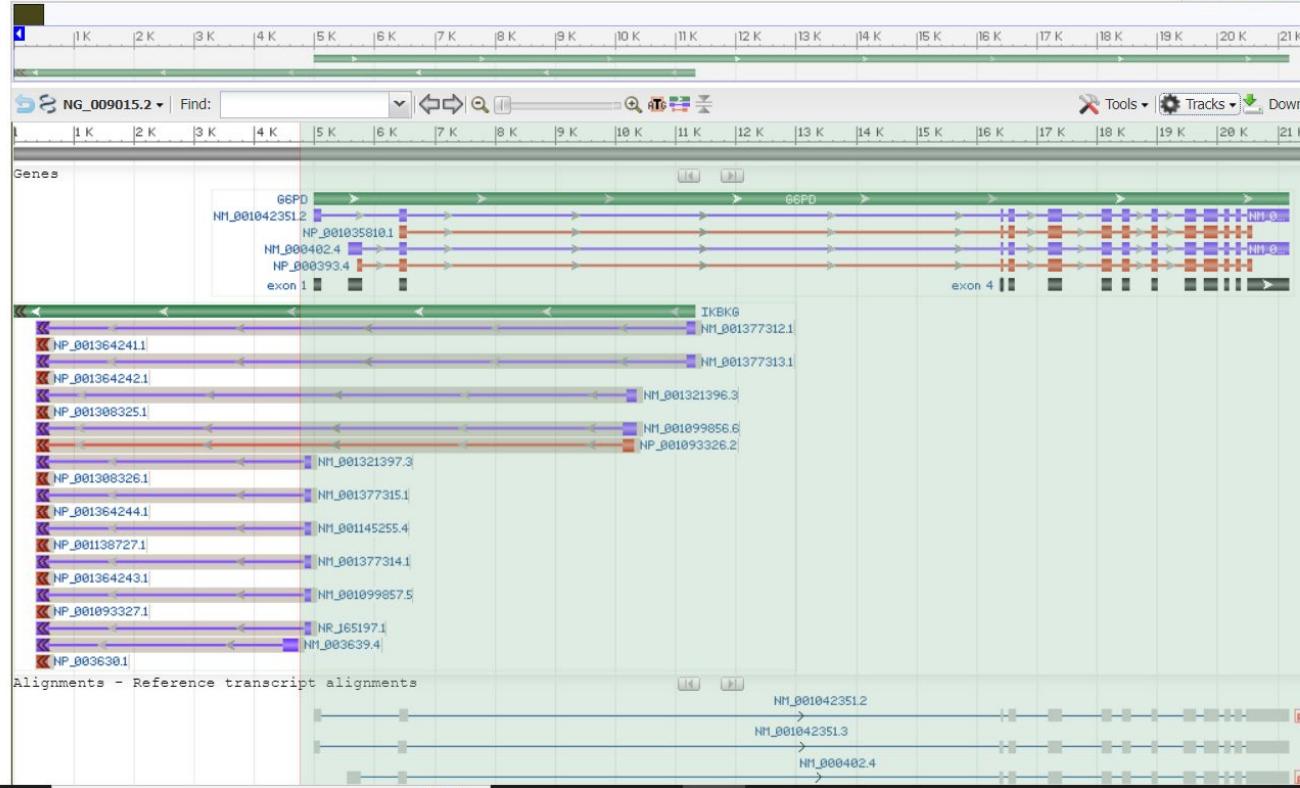
Graphics ▾

Send to: ▾

Analyze this sequence

Homo sapiens glucose-6-phosphate dehydrogenase (G6PD), RefSeqGene on chromosome X

NCBI Reference Sequence: NG_009015.2

[GenBank](#) [FASTA](#)[Link To This View](#)

- [Range: 4752..21642](#)
- [16,891 nt](#)
- [Copy Sequence \(Selection\)](#)
- [Zoom On Range](#)
- [Zoom To Sequence](#)
- [Set New Marker On Range](#)
- [Modify Range](#)
- [Add New Panel On Range](#)
- [BLAST Search \(Selection\)](#)
- [Primer BLAST \(Selection\)](#)
- [Download FASTA \(Selection\)](#)
- [Download GenBank Flat File \(Selection\)](#)

Reference sequence information

RefSeq alternative splicing
See 3 reference mRNA sequence s gene.

More about the G6PD gene

This gene encodes glucose-6-phosphate dehydrogenase, a cytosolic enzyme encoded by a linked gene whose main f...
Also Known As: G6PD1

Homologs of the G6PD gene

The G6PD gene is conserved in Rhesus monkey, mouse, rat, zebrafish, fruit fly, mosquito, *S.cerevisiae*, *K.lactis*, *E.gossypii*, *S.Neapolitana*, *S.crassa*, *A.thaliana*, rice, and frog.

Related information

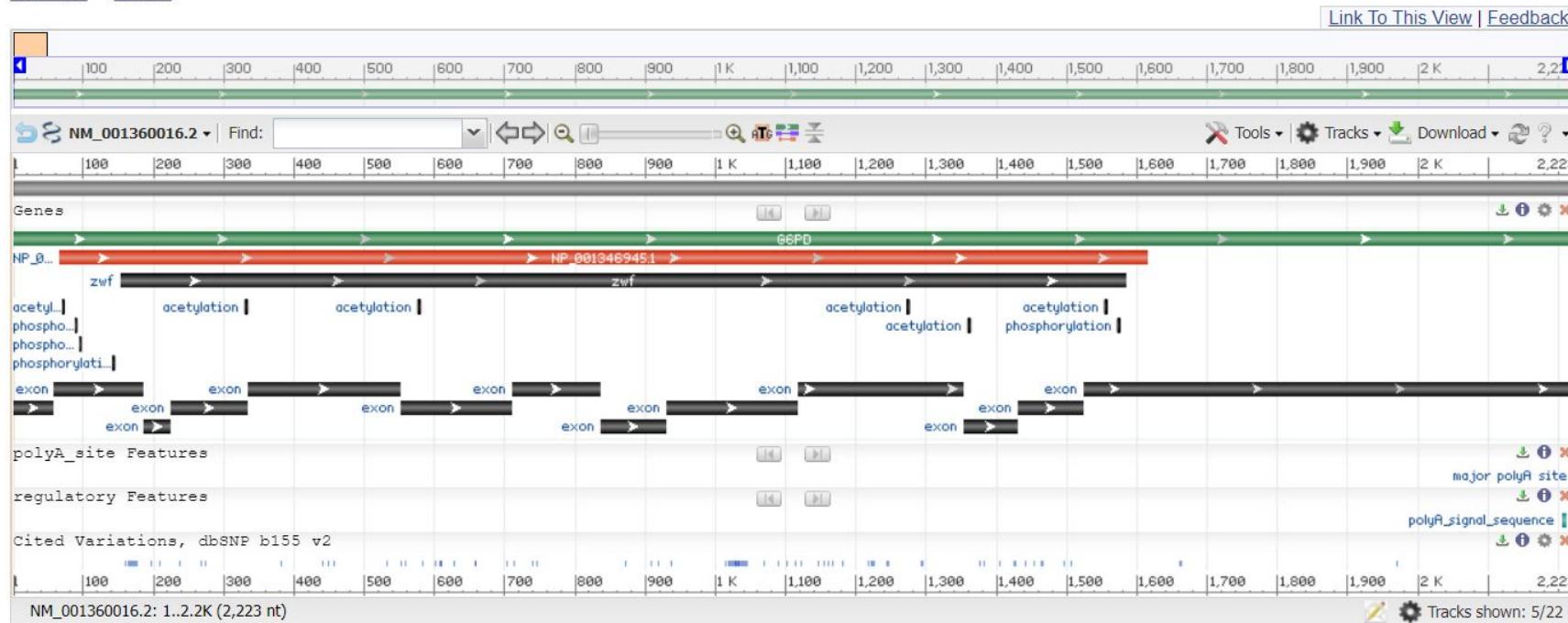
Graphics ▾

Send to: ▾

Homo sapiens glucose-6-phosphate dehydrogenase (G6PD), transcript variant 3, mRNA

NCBI Reference Sequence: NM_001360016.2

[GenBank](#) [FASTA](#)



Graphics ▾

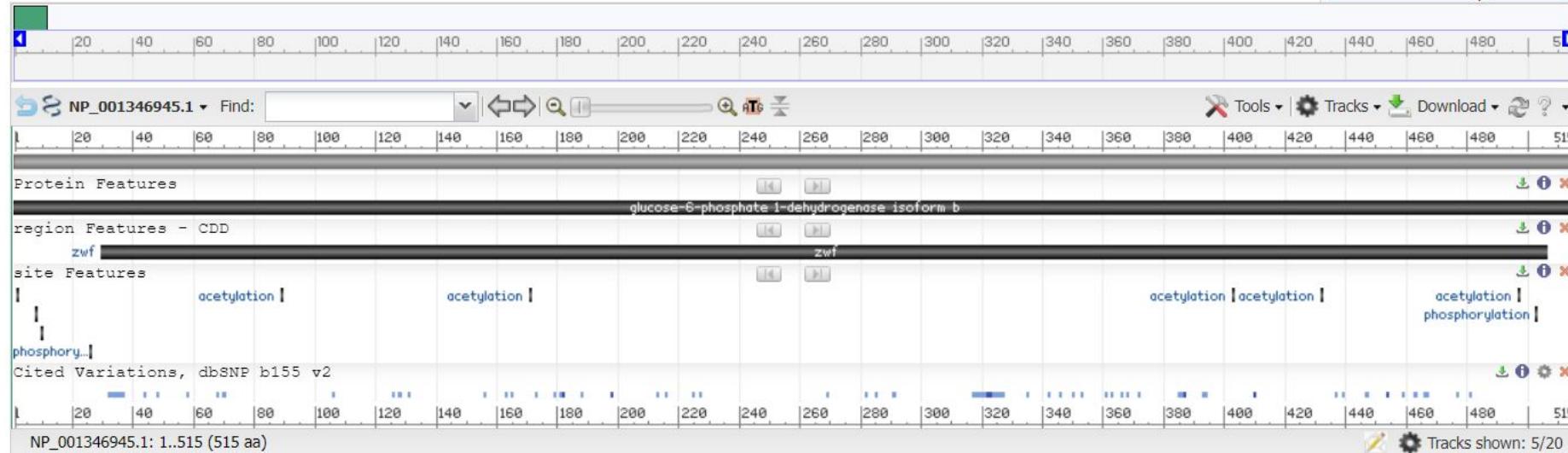
Send to: ▾

glucose-6-phosphate 1-dehydrogenase isoform b [Homo sapiens]

NCBI Reference Sequence: NP_001346945.1

[GenPept](#) [Identical Proteins](#) [FASTA](#)

[Link To This View](#) | [Feedback](#)



Homo sapiens glucose-6-phosphate dehydrogenase (G6PD), RefSeqGene on chromosome X

NCBI Reference Sequence: NG_009015.2

[GenBank](#) [Graphics](#)

>NG_009015.2 Homo sapiens glucose-6-phosphate dehydrogenase (G6PD), RefSeqGene on chromosome X

```
TGCCCTGTCCAGGATGGACAGCACTTCATGAGTGCTCAGTGCAGGGACACAGAGGCCCTGGACATGTTGG
TCCAACCTGGAAGGTTCAATTCCAGCTCAGAGCTCCCTGTAAGATCCAGCAAGGCTGTCCTGGAGCCTAC
CCGGCAGCGTCTTCCTCCCTGCCCCCTGGCTTCTCCCCCTGTCCTCCTTCCCTCCAGTACTG
GTCCTGAGAGCATGCTAAACTTCCCTCCGCTAATCTGCATCTCAGAGTCTACTTCTGAAACCAAGAGG
GAGGGCAGCACATCCCCACCCAGGCCAGGAAACGCCAGGGTGCAGGTGGGAGAACAGGCCCTCCCT
CAGCAGGGCTGGATCCCCTGACTCTTCCTCACCTGGAGCTCTGATTCTCCTCCAGGCAGCGCTGGAG
GGTCTCAGGAGCGCCCTGTTCTGAAGGCAGGTGCAAGCATGGCTGGCTCCCCAGAGGAGACTTCCGCCC
AGTACGTCCTGATCTGCTGCCGGGACCACACTGGGCTGACCACATCACAAGTTGGCTCTCCAGAGGT
GCCTATTCTACCAACAGGCAAGGGCTGAAAGAGAAAGGCCAGCAGGGAGTCACAGGAGACCCACATCT
TACCCAGCAGAAAAGCCTATGAGTTCTGGGGCTACATTTAAACAGAATGCTGATAATTACACCTAGTG
ATCAGGATGCTAAAGTGGGTTTCAGTCTAAGGGTAGCAGATACTCCATTCAAAGGGAGGTGCTGGA
CCACCCATGTGGGACTTGAAGGAGCTAGCTCCAGGCTGAAATAGAAGAATTCAACCCAGAGATGGAGA
GTCGGGACAAGACTGTTGTGGAGGGAGATGTCAGTATGCTGGAAATCTGATCCCTACACACCTCAA
GCAGGGAGAAGTGGAGGGTACCGGCCCTCATCCAGTGTAGTGGGCTGAATGGTAGCCCCCGAAAAGATA
TGCCCATGTTCAACCCCTGAAACATATGTGATTGGGAAAGGGCTTTGCAAGATGTAGGATCTCAAGA
TGAGATCATTCTGGGTGACGTGGGTGGGCCCTAAATCAAACACGTGCTTTGAGAGACAGAAGAGAG
AAGCCAGAGGGAGACGCAAGGGAGGGCATGTCGGACTGCCCTGGAGACTGGAGTGAGGCAGCCAC
AAGCCAGGAATAATGGGGCAGAACAGCTGGGAGAGGCAAGGGAGGATTGCCCTTGAAGCCTTGAGG
GTGTGTGCTTCAACCTTGGTTTGACTCTAGTCTCAGAACAGTGTGAGATTAAATGTCTGGCG
GGCGCAGTGGCTCACACCTGTAATCCAGCACTTGGGAGGCCAGGCGGGCAATCATGAGGTAGGAG
ATCGAGACCAGCCTGGCTAACACGGTAAACCTGCTACTAAAAAATGGCGGTCA
CGGTGGCTACGCCGCTGTAATCCAGCACTTGGGAGGCCAGCAGGGTGGATCATGAGGTAGGAGATTG
AGACCATCCTGGCTAACACAGTGAACACCTGCTACTAAAAAATAGCCAGGCATGGTAG
CGGGCGCCTGTAAGCCAGCTACCCAGGAGGCTGAGGCAAGGAGATGGCGTAACCCGGGAGGCAGGAGA
ATGGCGTGAACCCGGGATCGCGCACTGCACTCCAGCCTGGCAACAGAGCGAGACTCATCTCAAAAAA
AAAAAAATACAAAAAATTAGCAGAGTGTGGTGACACGCGCCTGAGTCCAGCTACTCGGGAGGCCGAAGC
AAGAGAATCGCTGAACCCAGGAGGCCAGTGGCAGTGTGAGATTGTGCCACTGCACTCCACCTT
```

Gene Orthologs

Search NCBI

G6PD

Search

G6PD - glucose-6-phosphate dehydrogenase

This gene encodes glucose-6-phosphate dehydrogenase. This protein is a cytosolic enzyme encoded by a housekeeping X-linked gene whose main function is to produce NADPH, a key electron donor in the defense against oxidizing agents and in reductive biosynthetic reactions. G6PD is remarkable for its genetic diversity. Many variants of G6PD, mostly produced from missense mutations, have been described with wide ranging levels of enzyme activity and associated clinical symptoms. G6PD deficiency may cause neonatal jaundice, acute hemolysis, or severe chronic non-spherocytic hemolytic anemia. Two transcript variants encoding different isoforms have been found for this gene. [provided by RefSeq, Jul 2008]

Genes similar to G6PD

NCBI Orthologs How was this calculated?

0 items

SEARCH THE TAXONOMY TREE

Enter taxonomic name

- Craniata
 - ▷ vertebrates

Protein alignment

one sequence per gene (3)
 all sequences per gene (24)

Align

328 genes for: ver

Add to cart Protein alignment Download

3 selected

Previous Next

Species

Gene

Architecture

aa

Feedback

COBALT

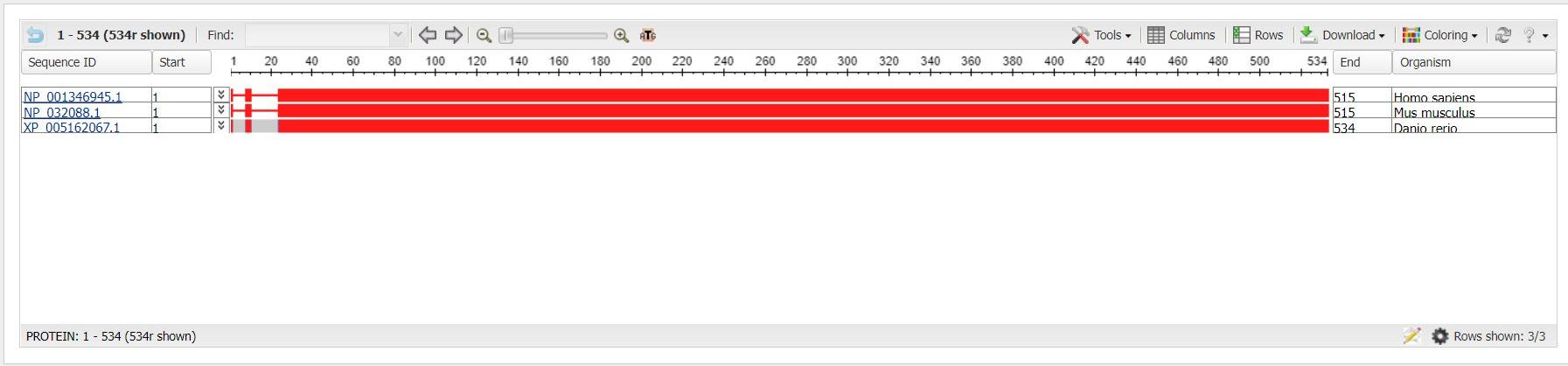
Constraint-based Multiple Alignment Tool

Home Recent Results Help

Phylogenetic Tree Edit and Resubmit ▶ Download

- Cobalt RID V5WTPGA6212 (3 seqs)

▼ Graphical Overview

▼ Descriptions Select All ▶ Alignment parameters

Accession	Description	Links
<input checked="" type="checkbox"/> NP_001346945.1	glucose-6-phosphate 1-dehydrogenase isoform b [Homo sapiens]	Related Information
<input checked="" type="checkbox"/> NP_032088.1	glucose-6-phosphate 1-dehydrogenase X [Mus musculus]	Related Information
<input checked="" type="checkbox"/> XP_005162067.1	glucose-6-phosphate 1-dehydrogenase isoform X1 [Danio rerio]	Related Information

COBALT

Constraint-based Multiple Alignment Tool

Home Recent Results Help

Phylogenetic Tree View

This tree is based on COBALT multiple alignment more...

[Reset Tree](#)

Cobalt RID V5WTPGA6212

Number of Seqs 3

Tree method

Fast Minimum Evolution

Max Seq Difference

0.85

Distance

Grishin (protein)

Sequence Label

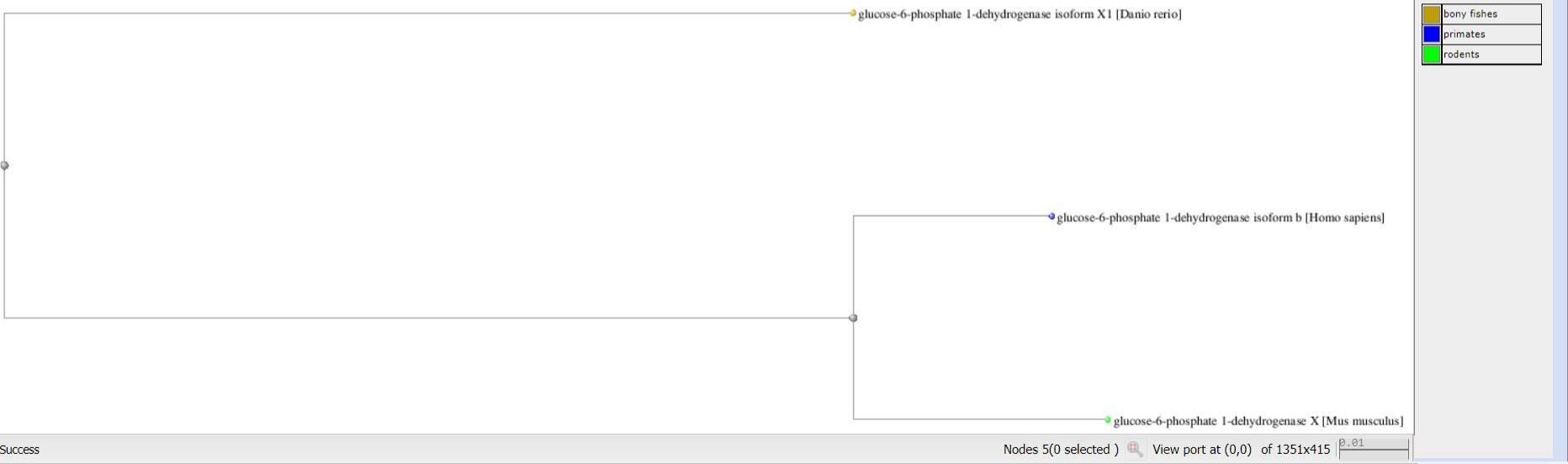
Sequence Title (if avail.)

[Hide legend](#) Find: all

Label color map

Blast names color map

bony fishes
primates
rodents



BLAST

SC113603 Molecular Biology
Lect. Todsapol Techo

BLAST System

- Search tool for **interesting sequence**
- Require **sequences** as the input to search

GAACTCGGGAAGCCGGCAGAAGTGTGAGGCCGG
TAGGGCCGCATCCCGCTCCGGAGAGAAGTCTGAG
TCCGCCAGGCTCTGCAGGCCGGAAAGCTCGGTAA
TGATAAGCACGCCGCCACTTGCAGGGCGTCAC
CGCCTACAGCCCCCTGCTCTGGACGGCGCGT
CTAGCCTGGGGCGCTGGCCCGCCCGCCCTCTC
CGGGGGAGGAATCAAGAAGAGACTGCCAATAGGC
CGGCTTGACCCGCAACAGGCAGGGTTCCCGGG
GGAGTGGCGCGCAGAAGGCCCGCCAGGAGGCCA
GGGACAGCCCAGAGGAGGCGTGGCACGCTGCCG
GCGGAAGTGGAGCCCTCCGCAGCGCGAGGCCGC
CGGGCAGGCCGGAAACCGGACAGTAGGGCGG
GGCCGGGCCGGCGATGGGGATGCGGGAGACTACGC
GGAGCTGCACCCGTGCCGCCGAATTGGGGATG
CAGAGCAGCGGCAGGGGTATGGCAGGCCAGCGCG
GGCCGGCCCTCAGCGCAGGTGCCGAGAGGCAGG
GGCTGGCCTGGGATGCGCGCACCTGCCCTGCC
CGCCCCGCCGCACGAGGGGTGGTGGCCGAGGCC
CCGCCCCGCACGCCCTGCCGTAGGCAGGTCCGCTC

MGRRGSAPGNRTRLGCERGGRRRSADSVMAEQ
VALSRTQVCGILREELFQGDAFHQSDTHIFIIMG
ASGDLAKKIIYPTIWWLFRDGLLPENTFIVGYAR
SRLTVADIRKQSEPFFKATPEEKLKLEDFFFARN
YVAGQYDDAASYQRLNSHMNALHLGSQANRLFYL
ALPPTVYEAVTKNIHESCMSQIGWNRIIVEKPF
RDLQSSDRLSNHISLFREDQIYRIDHYLGKEMV
QNLMVLRFANRIFGPIWNRDNIACVILTFKEPFG
TEGRGYFDEFGIIRDVMQNLLQMLCLVAMEKPA
STNSDDRDEKVVLKCISEVQANNVVLQYVGN
PDGEGEATKGYLDPTVPRGSTTATFAAVVLYVE
NERWDGVFILRCGKALNERKAEVRLQFHVDVAGD
IFHQQCKRNELVIRVQPNEAVYTKMMTKKPGMF
NPEESELDTYGNRYKNVKLFDAYERLILDVFC
SQMHFVRDELREWIFTPLLHQIELEPKPKIP
YIYGSRGPTEADELMKRVGFQYEGTYKWVNPHKL

>NC_000023.11:c154547569-154531390
Homo sapiens chromosome X,
GRCh38.p14 Primary Assembly
GAACTCGGGAAGCCGGCGAGAAGTGTGAGGCCGG
TAGGGCCGCATCCCGCTCCGGAGAGAAGTCTGAG
TCCGCCAGGCTCTGCAGGCCGCCGGAAAGCTCGGTAA
TGATAAGCACGCCGCCACTTGCAGGGCGTCAC
CGCCTACAGCCCCCTGCTCTGCAGGCCGCCGT
CTAGCCTGGGGCGCTGCCGCCGCCCTCTC
CGGGGGAGGAATCAAGAAGAGACTGCCAATAGGGC
CGGCTTGACCCCGCAACAGGCAGGGTTCCCGGG
GGAGTGGCGCGCAGAAGGCCGCCAGGAGGCCA
GGGACAGCCCAGAGGAGGCGTGGCCACGCTGCCG
GCGGAAGTGGAGCCCTCCGCAGCGCGAGGCCGC
CGGGCAGGCCGGAAACCGGACAGTAGGGCGG

Fasta Format

- **Fasta:** Simple Formats for Sequences

Header

- filename.**fasta**, filename.**fa**, filename.**fna**, and etc.

The diagram illustrates the Fasta format with a yellow box highlighting the header line. Two black arrows point from the left and right towards the sequence line, which starts with '>NC_000023.11:c154547569-154531390'. The sequence line contains a long string of DNA bases.

>NC_000023.11:c154547569-154531390 Homo sapiens chromosome X, GRCh38.p14 Primary Assembly
GAACTCGGGAAGCCGGCGAGAAAGTGTGAGGCCGCGTAGGGCCGCATCCCGCTCCGGAGAGAAAGTCTGAG
TCCGCCAGGCTCTGCAGGCCCGCGGAAGCTCGTAATGATAAGCACGCCGCCACTTGCAGGGCGTCAC
CGCCTACACGCCCTCGTCTCGGACGGCGCGTCTAGCCTGGGGCGCTGGCCGCCCGCCCTCTC
CGGGGGAGGAATCAAGAAGAGACTGCCAATAGGGCCGGCTTGACCCGCGAACAGGGCAGGGTTCCCGGG
GGAGTGCGCGGCAGAAGGCCGCCAGGAGCCGAGGGACAGCCCAGAGGAGGCGTGGCACGCTGCCG
GCGGAAGTGGAGGCCCTCCGCGAGCGCGAGGGCGCCGGGCAGGCGGGAAACCGGACAGTAGGGCGG

Sequence

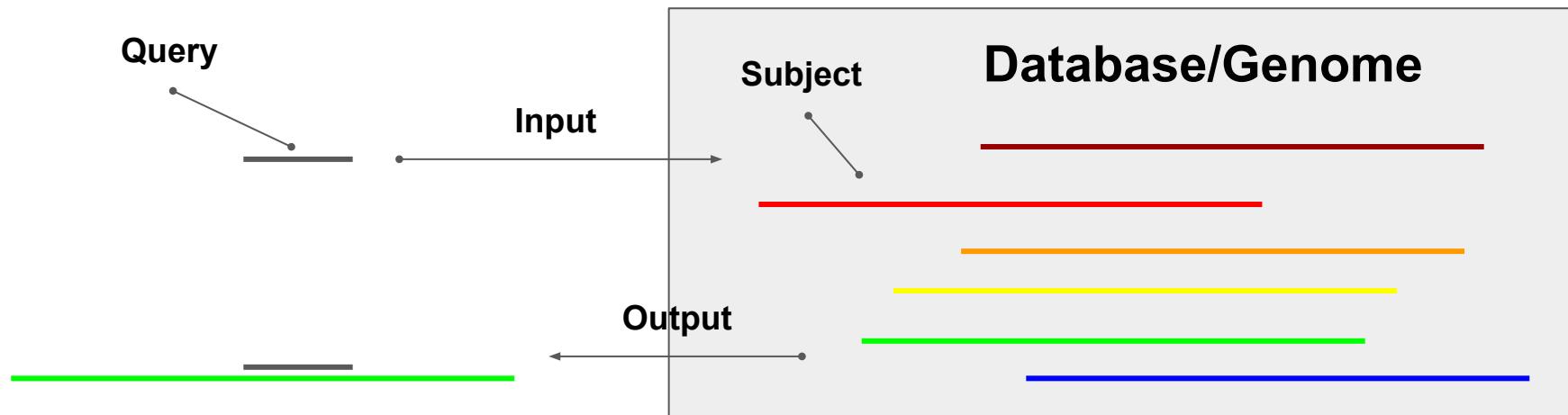
```
>Seq1 [organism=Carpodacus mexicanus] [clone=6b] actin (act) mRNA, partial cds
CCTTATCTAATTTGGAGCATGAGCTGGCATAGTTGAAACGCCCTCAGCCTCCTCATCCGTGCAGAATAATAATTCTTATAGTAATACCAATCATGATCGGTGGT
TTCGGAAACTGACTAGTCCCCTACTCATAAT
>Seq2 [organism=uncultured bacillus sp.] [isolate=A2] corticotropin (CT) gene, complete cds
GGTAGGTACCGCCCTAAGNCCTTAATCCGAGCAGAACTANGCCAACCCGGAGCCCTCTGGGAGACGACTCAACACCACCTCTTGACCCAGCAGGAGGAGACCCA
GTACTATAACCAGCACCTATTCTGATTCTT
>Seq3 [organism=Phalaenopsis equestris var. leucaspis]
CCTATACTTAATTTGGCGCATGAGCCGAATGGTGGGTACCGCTCTAACGCCCTCATTGAGCAGAACTAGGCCAACCCGGAGCCCTCTGGGAGACGACCAAGTCTA
CAACGTGGTTGTACGGCCCATGCCTCG
>Seq9 [organism=Petunia integrifolia subsp. inflata]
TAGTTGAAACGCCCTCAGCTACTCATCCGAGCAGAACTAGGCCAACCCGGAAACCCCTCTGGGAGATGACCAAATCTACAATGTAATCGTCACTGCCCATGCCTCGAA
TAATCTTCTTATAGTAATACCAAGTCATA
```

BLAST

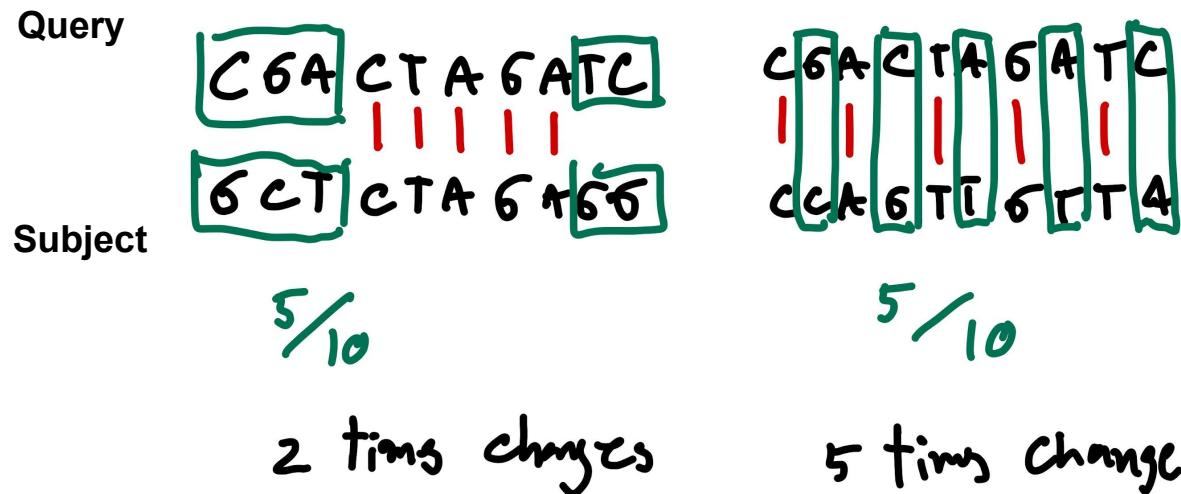
- BLAST tells you about non-chance similarities between biological sequences
- If similarity do not due to chance then they must be due to something else
 - Homology
 - Simple identification
- When you have a sequence, the possible questions are
 - What is relate to?
 - Homology
 - Function
 - Is it new or already in database?
 - Find the mach sequences in database
 - Organism of origin
 - Where is it located or how is it organized
 - In genome

BLAST Algorithm

- Seed and Extend
- Local Alignment

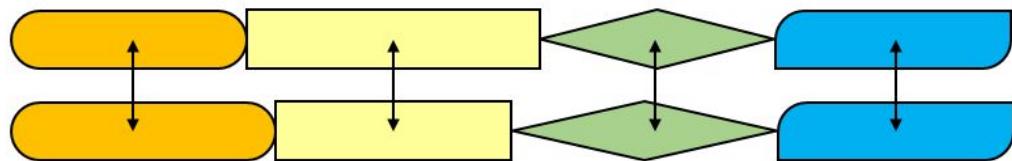


Sequence Alignment

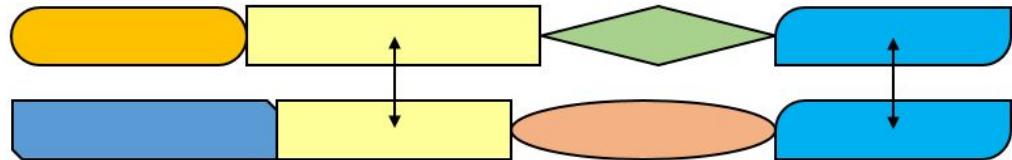


Pairwise Alignment

- Similarity between Two sequences
 - Query
 - Subject
- Global alignment Try to map all sequence from quarry to subject
- Local alignment try to map some sequences in query to subject



Global Alignment



Local Alignment

Global alignment and Local alignment

SEQUENCE 1 ATGACTTTACCAGGTGGTGGTATTAGTTCTATTTGGTCGGTTGGCTTGATTGCCATTATT
SEQUENCE 2 ATGATTTGTCAGGTACCATTGCT

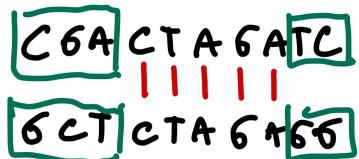
Global alignment

1	ATGACTTTACCAGGTGGTGGTATTAGTTCTATTTGGTCGGTTGGCTTGATTGCCATTATT	66
2 ATGATTTGTCAGG-----TACCATT-----GCT-----	24

Local alignment

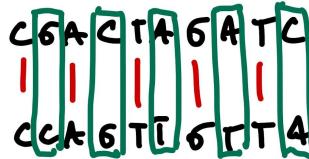
1	ATGACTTTACCAGGT 15
1	. ..
1	ATGATTTGTCAGGT 15

Seed and Extent



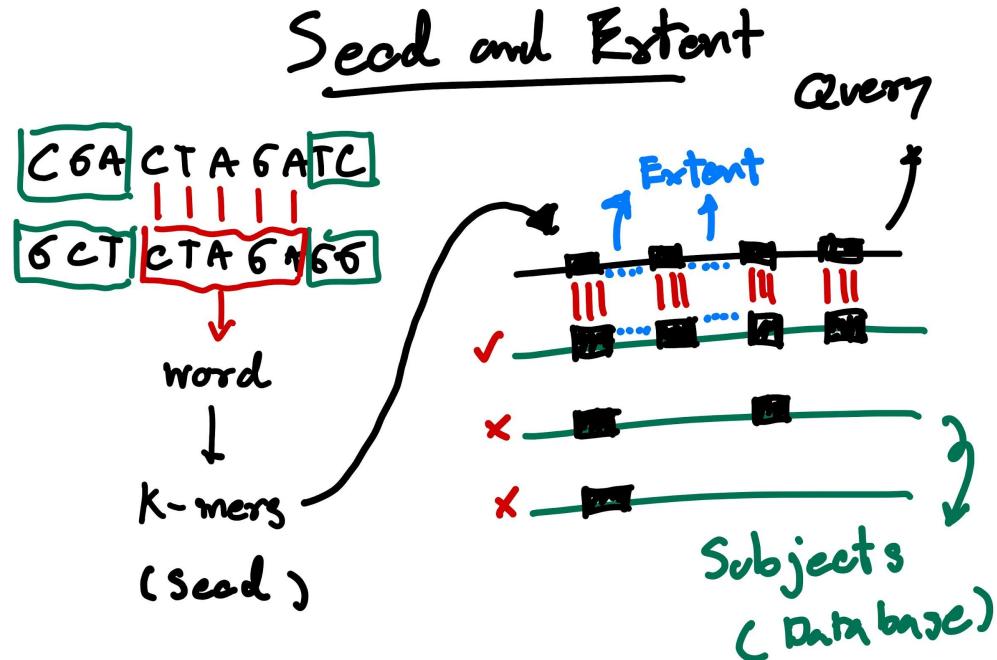
5/10

2 times changes



5/10

5 times change



BLAST Scoring Systems

- Local alignment Based on **Smith–Waterman algorithm for scoring systems**
- High score mean highly similar between 2 sequences

Initialize the scoring matrix

	T	G	T	T	A	C	G	G
G	0							
G	0							
T	0							
T	0							
G	0							
A	0							
C	0							
T	0							
A	0							

Substitution matrix:

$$S(a_i, b_j) = \begin{cases} +3, & a_i = b_j \\ -3, & a_i \neq b_j \end{cases}$$

Gap penalty: $W_k = kW_1$
 $W_1 = 2$

Statistical Calculation: Expectation Value (E-value)

- This value tell you about the possibility that your query sequence similar to subject sequence by chance.
- the number of hits one can "expect" to see by chance when searching a database of a particular size

BLAST Types

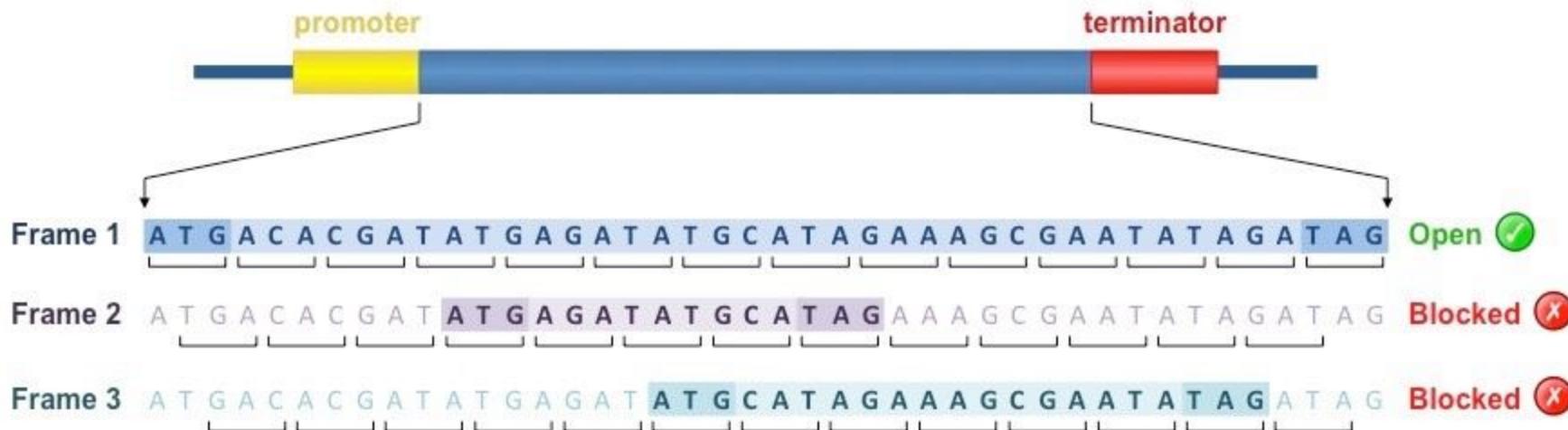
- **Basic Search**

- blastn
- blastp

	Program	Query Type	Subject Type	Computation
	blastn	N ——————>	— N	~ 1X
● Translated Search	blastp	P ——————>	— P	~ 1X
	blastx	N	— P	~ 6X
● Translational Frame	tblastn	P ——————>	N	~ 6X
	tblastx	N	N	~36X

(other BLAST types not listed: psiblast, deltablast, rpsblast)

Open Reading Frame (ORF)



Possible Amino Acid Sequences (Forward) {

I S G V I L V A D V G G R L I L K G C T V H E P G V E P A T A D G E

Nucleotide Sequence {

CGATCTCGGGCGTTCGGTCGCCATGTCGGCGGCCGACTCATCTGAAAGGCTGCACCGCTCACGGAAACCGGGCGTCGAACCCGCAACCGCGGACGGCGA

.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|

GCTAGAGCCCGCAAGACCAGCGGTACAGCCGCCGGCTGAGTAGAACTTCCGACGTGGCAAGTGCTGGCCCGCAGCTTGGCGTTGGCGCCTGCCGCT

Possible Amino Acid Sequences (Reverse) {

R D R A N Q D G I D A A S E D Q F A A G N V F R A D F R C G R V A

S R P R E P R R H R R G V * R S L S C R E R V P R R V P L R P R R

A I E P T R T A S T P P R S M K F P Q V T * S G P T S G A V A S P S

Gene 1

S * S T K Q M W T T C R F P E R R C R * V A F V A S S G T V R G L

N R D R Q S K C G L H A D S L R G G V G K W L L S P R R E P F A G C

I V I D K A N V D Y M Q I P * E A V S V S G F C R L V G N R S R V

AATCGTGATCGACAAAGCAAATGTGGACTACATGCAGATTCCCCTGAGAGGGCGGTGTCGGTAAGTGGCTTTGTCGCCCTCGTGGGAACCGTICGCGGGTT

.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|

TTAGCACTAGCTGTTCTGTTAACCTGATGTACGTCAAGGGACTCTCCGCCACAGCCATTACCGAAAACAGCGGAGCAGCCCTGGCAAGCGCCCAA

F D H D V F C I H V V H L N G S L R H R Y T A K T A E D P V T R P N

F R S R C L L H P S C A S E R L P P T P L H S K D G R R S G N A P

I T I S L A F T S * M C I G Q S A T D T L P K Q R R T P F R E R T

Gene 2

S R P N I P A P Q R I A L S P W V S M E Y Y E I W R R Q P A V R R

R A P T F Q R R N A S R C R R G C R W S I T R F G A V S P R C G A

V A P Q H S S A A T H R A V A V G V D G V L R D L A P S A R G A A R

GTCGCGCCCCAACATTCAGCGCCGCAACGCGATCGCCTGCGCCGTGGGTGTCGATGGAGTATTACGAGATTGGCGCCTGAGCCCGCGTGCAGCGCG

.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|

CAGCGCGGGGGTGTAGGTCGCGCGTTCGCTAGCGCGACAGCGGCCACAGCTACCTCATAATGCTCTAAACCGCGCAGTCGGCGCCACGCGCG

D R G L M G A G C R M A S D G H T D I S Y * S I Q R R * G A T R R

Q R A G V N W R R L A D R Q R R P H R H L I V L N P A T L G R H P A

T A G W C E L A A V C R A T A T P T S P T N R S K A G D A R P A A

A V S F L A R N I A Q L G L H L F E R K D D A D R K R L T D H P L A

R C R S W R A T S R N S V C T C S S A R T M P T A S G * P T T R S

G V V P G A Q H R A T R S A P V R A Q G R C R P Q A V D R P P A R

GCGGTGTCGTTCTGGCGCGCAACATCGCGCAACTCGGTCTGCACTGAGCGCAAGGACGATGCCGACCGCAAGCGTTGACCGACCCGCTCG

.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|

CGCCACAGCAAGGACCGCGCGTGTAGCGCGTGTAGCCAGACGTGGACAAGCTCGCGTCTGCTACGGCTGGCGTTGCCAACACTGGCTGGCTGGCGAGC

A T D N R A R L M A C S P R C R N S R L S S A S R L R N V S W G S

R H R E Q R A V D R L E T Q V Q E L A L V I G V A L P Q G V V R E

R P T T G P A C C R A V R D A G T R A C P R H R G C A T S R G G A R

Six-Frame Translation



 todste@orcid

BLAST®

Home Recent Results Saved Strategies Help

Basic Local Alignment Search Tool

BLAST finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance. [Learn more](#)

NEWS

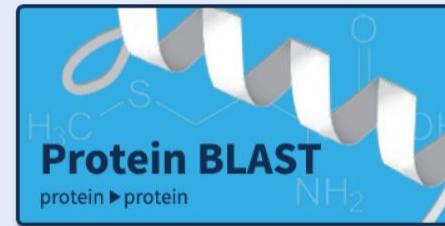
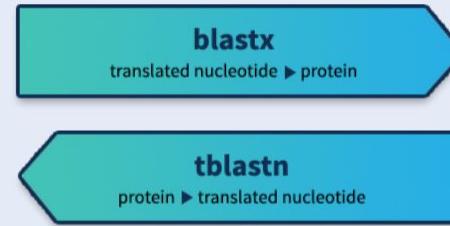
BLAST+ 2.13.0 is here!

Starting with this release, we are including the blastn_vdb and tblastn_vdb executables in the BLAST+ distribution.

Thu, 17 March 2022

 [More BLAST news...](#)

Web BLAST



[blastn](#)[blastp](#)[blastx](#)[tblastn](#)[tblastx](#)

Standard

BLASTN programs search nucleotides

Enter Query Sequence

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

NC_045512

Query subrange

From To

Or, upload file

 Choose 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 (nr etc.) rRNA/ITS databases Genomic + transcript databases Betacoronavirus Nucleotide collection (nr/nt) [?](#)

Organism

Optional

 Enter organism name or id—completions will be suggested exclude Add organismEnter organism common name, binomial, or tax id. Only 20 top taxa will be shown [?](#)

Exclude

Optional

 Models (XM/XP) Uncultured/environmental sample sequences

Limit to

Optional

 Sequences from type material  Create custom databaseEnter 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 [?](#)

Choose Search Set

Database

-
- Standard databases (nr etc.)
-
- rRNA/ITS databases
-
- Genomic + transcript databases

Organism
Optional

-
- Nucleotide collection (nr/nt)
-
- RefSeq Select RNA sequences (refseq_select)
-
- Reference RNA sequences (refseq_rna)
-
- RefSeq Representative genomes (refseq_representative_genomes)
-
- RefSeq Genome Database (refseq_genomes)
-
- Whole-genome shotgun contigs (wgs)
-
- Expressed sequence tags (est)
-
- Sequence Read Archive (SRA)
-
- Transcriptome Shotgun Assembly (TSA)
-
- Targeted Loci(TLS)
-
- High throughput genomic sequences (HTGS)
-
- Patent sequences(pat)
-
- PDB nucleotide database (pdb)
-
- Human RefSeqGene sequences(RefSeq_Gene)
-
- Genomic survey sequences (gss)
-
- Sequence tagged sites (dbsts)

Exclude
OptionalLimit to
OptionalEntrez Query
Optional

Program Selection

Optimize for

Choose a BLAST algorithm [?](#)

BLAST

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

 Show results in a new window

 NCBI 

Entrez PubMed Nucleotide Protein Genome Structure PMC Taxonomy BioCollections

Search for as lock

Display levels using filter:

Nucleotide Protein Structure Genome Popset SNP Conserved Domains GEO Datasets PubMed Central
 Gene HomoloGene SRA Experiments LinkOut BLAST GEO Profiles Protein Clusters Identical Protein Groups SPARCLE
 BioProject BioSample Assembly dbVar Genetic Testing Registry Host Viral Host PubChem BioAssay

Lineage (full): [cellular organisms](#); [Eukaryota](#); [Opisthokonta](#); [Metazoa](#); [Eumetazoa](#); [Bilateria](#); [Deuterostomia](#); [Chordata](#); [Craniata](#); [Vertebrata](#); [Gnathostomata](#); [Teleostomi](#); [Euteleostomi](#); [Sarcopterygii](#); [Dipnotetrapodomorpha](#); [Tetrapoda](#); [Amniota](#); [Mammalia](#); [Theria](#); [Eutheria](#); [Euarchontoglires](#)

• **Primates** Click on organism name to get more information.

- [Haplorrhini](#)
 - [Simiiformes](#)
 - [Catarrhini](#)
 - [Platyrrhini](#) (New World monkeys)
 - [Tarsiiformes](#)
 - [Tarsiidae](#) (tarsiers)
- [Strepsirrhini](#) (prosimians)
 - [Chiromyiformes](#)
 - [Daubentoniidae](#) (aye-ayes)
 - [Lemuriformes](#)
 - [Cheirogaleidae](#) (dwarf and mouse lemurs)
 - [Indriidae](#)
 - [Lemuridae](#) (lemurs)
 - [Lepilemuridae](#)
 - [Palaeopropithecidae](#)
 - [Lorisiformes](#)
 - [Galagidae](#) (galagos)
 - [Lorisidae](#) (lorises)
- [unclassified Primates](#)
 - [marmosets](#)
 - [Primates sp. BOLD:AAA0001](#)
 - [Primates sp. BOLD:AAC7735](#)
 - [Primates sp. BOLD:AAC9009](#)
 - [Primates sp. BOLD:AAC9015](#)

Algorithm parameters

General Parameters

Max target sequences

100

Select the maximum number of aligned sequences to display 

Short queries

Automatically adjust parameters for short input sequences 

Expect threshold

0.05 

Word size

28 

Max matches in a query range

0 

Scoring Parameters

Match/Mismatch Scores

1,-2 

Gap Costs

Linear 

Filters and Masking

Filter

Low complexity regions 

Species-specific repeats for: [Ashbya] aceris (nom. inval.) 

Mask

Mask for lookup table only 

Mask lower case letters 

BLAST

Search **database Nucleotide collection (nr/nt)** using **Megablast** (Optimize for highly similar sequences)

Show results in a new window

blastn

blastp

blastx

tblastn

tblastx

Enter Query SequenceEnter accession number(s), gi(s), or FASTA sequence(s) [?](#) [Clear](#)Query subrange [?](#)From To

Or, upload file

 Choose File No file chosen [?](#)

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

Compare

 Select to compare standard and experimental database [?](#)**Standard**

Database

 Non-redundant protein sequences (nr) [?](#)Organism
OptionalPrimates (taxid:9443) exclude Add organismEnter organism common name, binomial, or tax id. Only 20 top taxa will be shown. [?](#)Exclude
Optional Models (XM/XP) Non-redundant RefSeq proteins (WP) Uncultured/environmental sample sequences**Program Selection**

Algorithm

- Quick BLASTP (Accelerated protein-protein BLAST)
 - blastp (protein-protein BLAST)
 - PSI-BLAST (Position-Specific Iterated BLAST)
 - PHI-BLAST (Pattern Hit Initiated BLAST)
 - DELTA-BLAST (Domain Enhanced Lookup Time Accelerated BLAST)
- Choose a BLAST algorithm [?](#)

BLASTP programs search protein databases using a pr

Algorithm parameters**General Parameters**Max target
sequences100 [?](#)Select the maximum number of aligned sequences to display [?](#)

Short queries

Automatically adjust parameters for short input sequences [?](#)

Expect threshold

0.05 [?](#)

Word size

6 [?](#)Max matches in a
query range0 [?](#)**Scoring Parameters**

Matrix

BLOSUM62 [?](#)

Gap Costs

Existence: 11 Extension: 1 [?](#)Compositional
adjustmentsConditional compositional score matrix adjustment [?](#)**Filters and Masking**

Filter

 Low complexity regions [?](#)

Mask

 Mask for lookup table only [?](#) Mask lower case letters [?](#)

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



todste@orcid

BLAST® » blastp suite » RID-V3KUK8DX01N

Home Recent Results Saved Strategies Help

Format Request Status

[Formatting options]

Job Title: Protein Sequence

Request ID	V3KUK8DX01N
Status	Searching
Submitted at	Sun Jan 1 13:22:11 2023
Current time	Sun Jan 01 13:22:43 2023
Time since submission	00:00:32

This page will be automatically updated in 6 seconds

FOLLOW NCBI



pbio.1001014.pdf

drive-download-20....zip

ตารางเรียนวิชา SC 1....pdf

cmb.2012.0022.pdf

Show all



Type here to search



18°C ແລ້ມນີສ 1/22 AM
ENG 1/2/2023



todste@orcid

BLAST® » blastp suite » results for RID-V3KUK8DX01N

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 Protein Sequence

RID V3KUK8DX01N Search expires on 01-03 01:22 am Download All

Program BLASTP ? Citation

Database nr See details

Query ID lcl|Query_784607

Description unnamed protein product

Molecule type amino acid

Query Length 146

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

to

E value

to

to

Query Coverage

to

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

?

Feedback

pbio.1001014.pdf

^

drive-download-20....zip

^

ตารางเรียนวิชา SC 1....pdf

^

cmb.2012.0022.pdf

^

Show all



Type here to search



18°C ແລ້ມນີສ



ENG

12:3 AM

1/2/2023

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

Descriptions Graphic Summary Alignments Taxonomy

Sequences producing significant alignments

Download Select columns Show 100 ?

select all 100 sequences selected

GenPept Graphics Distance tree of results Multiple alignment MSA Viewer

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	RecName: Full=Hemoglobin subunit beta; AltName: Full=Beta-globin; AltName: Full=Hemoglobin beta chain [...] Macaca mulatta	Macaca mulatta	299	299	100%	4e-102	100.00%	146	P02026.1
<input checked="" type="checkbox"/>	hemoglobin subunit beta [Cercocebus atys]	Cercocebus atys	298	298	100%	8e-102	99.32%	147	NP_001292888.1
<input checked="" type="checkbox"/>	hemoglobin beta chain - stump-tailed macaque [Macaca arctoides]	Macaca arctoides	298	298	100%	1e-101	99.32%	146	C24693
<input checked="" type="checkbox"/>	hemoglobin subunit beta [Miopithecus talaponi]	Miopithecus tala...	296	296	100%	5e-101	98.63%	147	AYO89367.1
<input checked="" type="checkbox"/>	hemoglobin subunit beta [Macaca mulatta]	Macaca mulatta	295	295	100%	2e-100	97.95%	147	NP_001157900.1
<input checked="" type="checkbox"/>	RecName: Full=Hemoglobin subunit beta; AltName: Full=Beta-globin; AltName: Full=Hemoglobin beta chain [...] Mandrillus sphinx	Mandrillus sphinx	294	294	100%	3e-100	97.95%	146	P08259.1
<input checked="" type="checkbox"/>	hemoglobin subunit beta [Erythrocebus patas]	Erythrocebus pa...	294	294	100%	3e-100	97.95%	147	AYO89363.1
<input checked="" type="checkbox"/>	hemoglobin subunit beta isoform X2 [Theropithecus gelada]	Theropithecus g...	293	293	100%	4e-100	97.95%	147	XP_025213811.1
<input checked="" type="checkbox"/>	RecName: Full=Hemoglobin subunit beta; AltName: Full=Beta-globin; AltName: Full=Hemoglobin beta chain [...] Theropithecus g...	Theropithecus g...	293	293	100%	5e-100	97.95%	146	P02029.1
<input checked="" type="checkbox"/>	hemoglobin subunit beta [Chlorocebus sabaeus]	Chlorocebus sa...	293	293	100%	5e-100	97.95%	147	NP_001316847.1
<input checked="" type="checkbox"/>	RecName: Full=Hemoglobin subunit beta; AltName: Full=Beta-globin; AltName: Full=Hemoglobin beta chain [...] Chlorocebus aet...	Chlorocebus aet...	293	293	100%	7e-100	97.95%	146	P02028.1
<input checked="" type="checkbox"/>	RecName: Full=Hemoglobin subunit beta; AltName: Full=Beta-globin; AltName: Full=Hemoglobin beta chain [...] Cerocebus atys	Cerocebus atys	293	293	100%	1e-99	97.95%	146	P02031.1
<input checked="" type="checkbox"/>	hemoglobin subunit beta [Papio anubis]	Papio anubis	292	292	100%	2e-99	97.26%	147	NP_001162318.1
<input checked="" type="checkbox"/>	RecName: Full=Hemoglobin subunit beta; AltName: Full=Beta-globin; AltName: Full=Hemoglobin beta chain [...] Papio cynoceph...	Papio cynoceph...	292	292	100%	2e-99	97.26%	146	P02030.1
<input checked="" type="checkbox"/>	hemoglobin subunit beta [Trachypithecus francoisi]	Trachypithecus f...	290	290	100%	1e-98	96.58%	147	XP_033062959.1
<input checked="" type="checkbox"/>	RecName: Full=Hemoglobin subunit beta; AltName: Full=Beta-globin; AltName: Full=Hemoglobin beta chain [...] Semnopithecus...	Semnopithecus...	290	290	100%	1e-98	96.58%	146	P02032.1
<input checked="" type="checkbox"/>	RecName: Full=Hemoglobin subunit beta; AltName: Full=Beta-globin; AltName: Full=Hemoglobin beta chain [...] Hylobates lar...	Hylobates lar...	290	290	100%	1e-98	95.89%	146	P02025.1

Feedback



pbio.1001014.pdf



drive-download-20...zip



ตารางเรียนรู้ SC 1...pdf



cmb.2012.0022.pdf



Show all



Type here to search



18°C แจ่มใส

12:3 AM
1/2/2023

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

Descriptions Graphic Summary Alignments Taxonomy

hover to see the title click to show alignments Show Conserved Domains Alignment Scores < 40 40 - 50 50 - 80 80 - 200 >= 200 ?

100 sequences selected ? Putative conserved domains have been detected, click on the image below for detailed results.

Query seq. VHLTPPEKKNVTTLNGKVNVDVVGGEALGRLLVYPTKRFESFGDSSPDRVHGPNPKVRHKGKVLGAFSDGLNHLDNLKGTFRQSELNCQDHLVHPDEWFKLLQWLVCVLRAHIFGKEFTPQVQHRYQKVVAQGVANALRKYH

Specific hits Hb-beta-like
Superfamilies Globin-like superfamily

Distribution of the top 100 Blast Hits on 100 subject sequences

Query 1 20 40 60 80 100 120 140

Feedback

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

Descriptions Graphic Summary Alignments Taxonomy

Alignment view Pairwise ? Restore defaults Download

100 sequences selected

[Download](#) [GenPept](#) [Graphics](#)

[Next](#) [Previous](#) [Descriptions](#)

RecName: Full=Hemoglobin subunit beta; AltName: Full=Beta-globin; AltName: Full=Hemoglobin beta chain [Macaca mulatta]

Sequence ID: [P02026.1](#) Length: 146 Number of Matches: 1

[See 1 more title\(s\)](#) [See all Identical Proteins\(IPG\)](#)

Range 1: 1 to 146 [GenPept](#) [Graphics](#)

[Next Match](#) [Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
299 bits(765)	4e-102	Compositional matrix adjust.	146/146(100%)	146/146(100%)	0/146(0%)

Query 1 VHLTPEEKNAVTLWGKVNDEVGGEALGRLLLVYPWTQRFESFGDLSSPAVMGNPKV 60
Sbjct 1 VHLTPEEKNAVTLWGKVNDEVGGEALGRLLLVYPWTQRFESFGDLSSPAVMGNPKV 60

Query 61 KAHGKKVLAGFS DGLNHL DNLKG TFAQL SELH CDKL LHVD PENFK LLLGN VL CVAL AH FGK 120
KAHGKKVLAGFS DGLNHL DNLKG TFAQL SELH CDKL LHVD PENFK LLLGN VL CVAL AH FGK 120
Sbjct 61 KAHGKKVLAGFS DGLNHL DNLKG TFAQL SELH CDKL LHVD PENFK LLLGN VL CVAL AH FGK 120

Query 121 EFTPQVQAAYQKV VAGVAN ALAH KYH 146
Sbjct 121 EFTPQVQAAYQKV VAGVAN ALAH KYH 146

Related Information

[Gene](#) - associated gene details

[Identical Proteins](#) - Identical proteins to P02026.1

[Download](#) [GenPept](#) [Graphics](#)

[Next](#) [Previous](#) [Descriptions](#)

hemoglobin subunit beta [Cercocebus atys]

Sequence ID: [NP_001292888.1](#) Length: 147 Number of Matches: 1

[See 12 more title\(s\)](#) [See all Identical Proteins\(IPG\)](#)

Feedback



pbio.1001014.pdf



drive-download-20....zip



ตารางเรียนพืชฯ SC 1....pdf



cmb.2012.0022.pdf



Show all



Type here to search



19°C แจ่มใส



138 AM

1/2/2023

Descriptions	Graphic Summary	Alignments	Taxonomy																																																																				
Reports	Lineage	Organism	Taxonomy																																																																				
100 sequences selected ?																																																																							
<table border="1"><thead><tr><th>Description</th><th>Score</th><th>E value</th><th>Accession</th></tr></thead><tbody><tr><td>Macaca mulatta (Rhesus monkey) [primates]</td><td>▼ Next</td><td>▲ Previous</td><td>◀ First</td></tr><tr><td><u>RecName: Full=Hemoglobin subunit beta; AltName: Full=Beta-globin; AltName: Full=Hemoglobin beta chain_[Macaca mulatta]</u></td><td>299</td><td>4e-102</td><td>P02026</td></tr><tr><td><u>hemoglobin beta_[Macaca mulatta]</u></td><td>299</td><td>4e-102</td><td>681081B</td></tr><tr><td><u>Hemoglobin beta chain_[Macaca mulatta]</u></td><td>298</td><td>8e-102</td><td>EHH23170</td></tr><tr><td><u>TPA: globin A1_[Macaca mulatta]</u></td><td>298</td><td>8e-102</td><td>SAI82170</td></tr><tr><td><u>hemoglobin subunit beta_[Macaca mulatta]</u></td><td>295</td><td>2e-100</td><td>NP_001157900</td></tr><tr><td><u>beta hemoglobin_[Macaca mulatta]</u></td><td>295</td><td>2e-100</td><td>ACS94046</td></tr><tr><td>Cercocebus atys (sooty mangabey) [primates]</td><td>▼ Next</td><td>▲ Previous</td><td>◀ First</td></tr><tr><td><u>hemoglobin subunit beta_[Cercocebus atys]</u></td><td>298</td><td>8e-102</td><td>NP_001292888</td></tr><tr><td><u>beta hemoglobin_[Cercocebus atys]</u></td><td>298</td><td>8e-102</td><td>ACS94047</td></tr><tr><td><u>RecName: Full=Hemoglobin subunit beta; AltName: Full=Beta-globin; AltName: Full=Hemoglobin beta chain_[Cercocebus atys]</u></td><td>293</td><td>1e-99</td><td>P02031</td></tr><tr><td>Macaca nemestrina (pig-tailed macaque) [primates]</td><td>▼ Next</td><td>▲ Previous</td><td>◀ First</td></tr><tr><td><u>hemoglobin subunit beta_[Macaca nemestrina]</u></td><td>298</td><td>8e-102</td><td>XP_011717123</td></tr><tr><td><u>RecName: Full=Hemoglobin subunit beta; AltName: Full=Beta-globin; AltName: Full=Hemoglobin beta chain_[Macaca nemestrina]</u></td><td>298</td><td>8e-102</td><td>P68225</td></tr><tr><td>Macaca fascicularis (crab-eating macaque) [primates]</td><td>▼ Next</td><td>▲ Previous</td><td>◀ First</td></tr><tr><td><u>hemoglobin subunit beta_[Macaca fascicularis]</u></td><td>298</td><td>8e-102</td><td>XP_045227322</td></tr></tbody></table>				Description	Score	E value	Accession	Macaca mulatta (Rhesus monkey) [primates]	▼ Next	▲ Previous	◀ First	<u>RecName: Full=Hemoglobin subunit beta; AltName: Full=Beta-globin; AltName: Full=Hemoglobin beta chain_[Macaca mulatta]</u>	299	4e-102	P02026	<u>hemoglobin beta_[Macaca mulatta]</u>	299	4e-102	681081B	<u>Hemoglobin beta chain_[Macaca mulatta]</u>	298	8e-102	EHH23170	<u>TPA: globin A1_[Macaca mulatta]</u>	298	8e-102	SAI82170	<u>hemoglobin subunit beta_[Macaca mulatta]</u>	295	2e-100	NP_001157900	<u>beta hemoglobin_[Macaca mulatta]</u>	295	2e-100	ACS94046	Cercocebus atys (sooty mangabey) [primates]	▼ Next	▲ Previous	◀ First	<u>hemoglobin subunit beta_[Cercocebus atys]</u>	298	8e-102	NP_001292888	<u>beta hemoglobin_[Cercocebus atys]</u>	298	8e-102	ACS94047	<u>RecName: Full=Hemoglobin subunit beta; AltName: Full=Beta-globin; AltName: Full=Hemoglobin beta chain_[Cercocebus atys]</u>	293	1e-99	P02031	Macaca nemestrina (pig-tailed macaque) [primates]	▼ Next	▲ Previous	◀ First	<u>hemoglobin subunit beta_[Macaca nemestrina]</u>	298	8e-102	XP_011717123	<u>RecName: Full=Hemoglobin subunit beta; AltName: Full=Beta-globin; AltName: Full=Hemoglobin beta chain_[Macaca nemestrina]</u>	298	8e-102	P68225	Macaca fascicularis (crab-eating macaque) [primates]	▼ Next	▲ Previous	◀ First	<u>hemoglobin subunit beta_[Macaca fascicularis]</u>	298	8e-102	XP_045227322
Description	Score	E value	Accession																																																																				
Macaca mulatta (Rhesus monkey) [primates]	▼ Next	▲ Previous	◀ First																																																																				
<u>RecName: Full=Hemoglobin subunit beta; AltName: Full=Beta-globin; AltName: Full=Hemoglobin beta chain_[Macaca mulatta]</u>	299	4e-102	P02026																																																																				
<u>hemoglobin beta_[Macaca mulatta]</u>	299	4e-102	681081B																																																																				
<u>Hemoglobin beta chain_[Macaca mulatta]</u>	298	8e-102	EHH23170																																																																				
<u>TPA: globin A1_[Macaca mulatta]</u>	298	8e-102	SAI82170																																																																				
<u>hemoglobin subunit beta_[Macaca mulatta]</u>	295	2e-100	NP_001157900																																																																				
<u>beta hemoglobin_[Macaca mulatta]</u>	295	2e-100	ACS94046																																																																				
Cercocebus atys (sooty mangabey) [primates]	▼ Next	▲ Previous	◀ First																																																																				
<u>hemoglobin subunit beta_[Cercocebus atys]</u>	298	8e-102	NP_001292888																																																																				
<u>beta hemoglobin_[Cercocebus atys]</u>	298	8e-102	ACS94047																																																																				
<u>RecName: Full=Hemoglobin subunit beta; AltName: Full=Beta-globin; AltName: Full=Hemoglobin beta chain_[Cercocebus atys]</u>	293	1e-99	P02031																																																																				
Macaca nemestrina (pig-tailed macaque) [primates]	▼ Next	▲ Previous	◀ First																																																																				
<u>hemoglobin subunit beta_[Macaca nemestrina]</u>	298	8e-102	XP_011717123																																																																				
<u>RecName: Full=Hemoglobin subunit beta; AltName: Full=Beta-globin; AltName: Full=Hemoglobin beta chain_[Macaca nemestrina]</u>	298	8e-102	P68225																																																																				
Macaca fascicularis (crab-eating macaque) [primates]	▼ Next	▲ Previous	◀ First																																																																				
<u>hemoglobin subunit beta_[Macaca fascicularis]</u>	298	8e-102	XP_045227322																																																																				

Feedback



Blast Tree View

This tree was produced using BLAST pairwise alignments. [more...](#)

Reset Tree

BLAST RID V3KUK8DX01N Query ID Icl|Query_784607 Database nr

Tree method Fast Minimum Evolution Max Seq Difference 0.85 Distance Grishin (protein) Sequence Label Sequence Title (if avail.)

Mouse over an internal node for an internal node for a subtree or alignment. Click on tree label to select sequence to download

Find: all Tools Upload Hide legend

Label color map query from type material

Blast names color map carnivores primates other sequences unknown

hemoglobin subunit beta [Zalophus californianus]
PREDICTED: hemoglobin subunit beta [Colobus angolensis palliatus]
hemoglobin subunit beta [Trachypithecus francoisi]
RecName: Full=Hemoglobin subunit beta; AltName: Full=Beta-globin; AltName: Full=Hemoglobin beta chain [Semnopithecus entellus]
primates and other sequences | 77 leaves
PREDICTED: hemoglobin subunit beta [Rhinopithecus bieti]
hemoglobin subunit beta [Rhinopithecus roxellana]
primates | 2 leaves
hemoglobin subunit beta isoform X2 [Theropithecus gelada]
RecName: Full=Hemoglobin subunit beta; AltName: Full=Beta-globin; AltName: Full=Hemoglobin beta chain [Theropithecus gelada]
hemoglobin subunit beta [Papio anubis]
RecName: Full=Hemoglobin subunit beta; AltName: Full=Beta-globin; AltName: Full=Hemoglobin beta chain [Papio cynocephalus]
hemoglobin subunit beta [Lophocebus albigena]
hemoglobin subunit beta [Macaca mulatta]
RecName: Full=Hemoglobin subunit beta; AltName: Full=Beta-globin; AltName: Full=Hemoglobin beta chain [Mandrillus sphinx]
hemoglobin subunit beta [Miopithecus talapoin]
hemoglobin subunit beta [Erythrocebus patas]
primates | 2 leaves
RecName: Full=Hemoglobin subunit beta; AltName: Full=Beta-globin; AltName: Full=Hemoglobin beta chain [Cercocebus atys]
hemoglobin subunit beta [Cercocebus atys]
primates and unknown | 3 leaves

Success

BLAST is a registered trademark of the National Library of Medicine

Nodes 201(0 selected) View port at (0,0) of 1348x373 0.006

Support center Mailing list YouTube

NCBI

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

NATIONAL LIBRARY OF MEDICINE NIH USA.gov

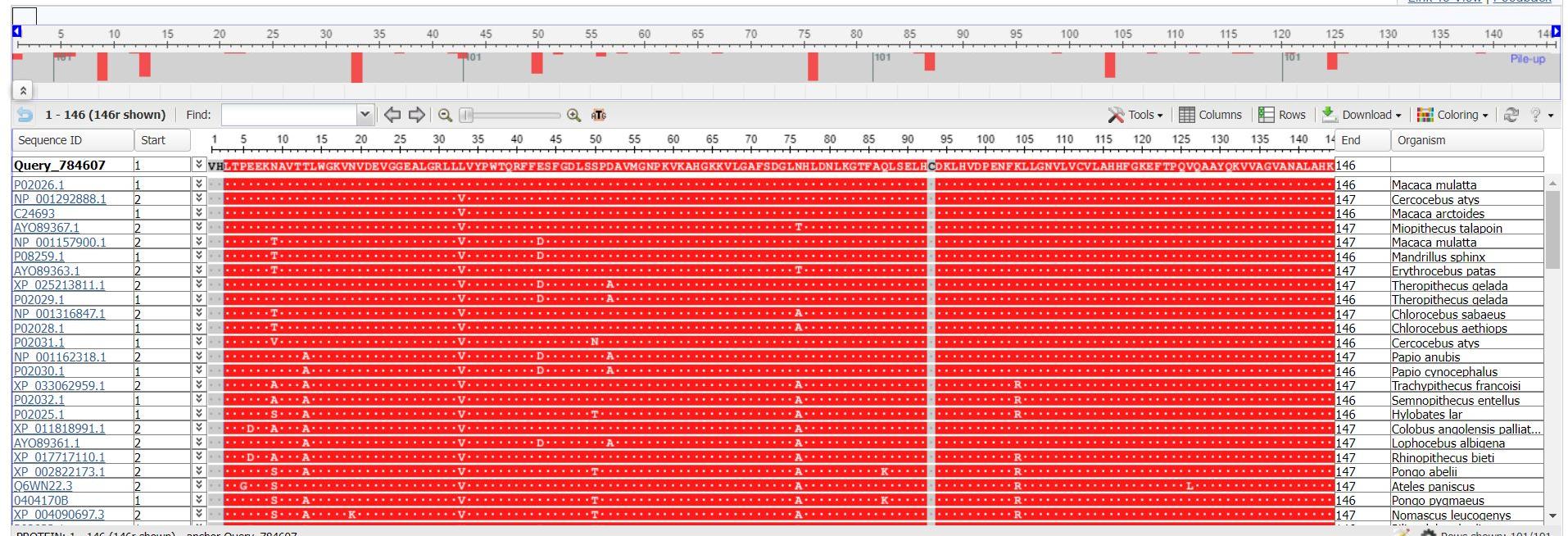


Home PubMed GenBank BLAST

Multiple Sequence Alignment Viewer 1.22.2

Alignment

Link To View | Feedback



Contact | Copyright | Disclaimer | Privacy | Accessibility

National Center for Biotechnology Information, US National Library of Medicine
8600 Rockville Pike, Rockville, MD USA 20894

pbio.1001014.pdf



drive-download-20....zip



ตารางเรียนรักษา SC 1....pdf



cmb.2012.0022.pdf



Type here to search



Show all



18°C

ແລ້ມນີສ

1:46 AM
1/2/2023

BLAST Results

RecName: Full=Hemoglobin subunit beta; AltName: Full=Beta-globin; AltName: Full=Hemoglobin beta chain [Macaca mulatta]

Sequence ID: [P02026.1](#) Length: 146 Number of Matches: 1

[See 1 more title\(s\) ▾](#) [See all Identical Proteins\(IPG\)](#)

Range 1: 1 to 146 [GenPept](#) [Graphics](#)

▼ [Next Match](#) ▲ [Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
299 bits(765)	4e-102	Compositional matrix adjust.	146/146(100%)	146/146(100%)	0/146(0%)
Query 1	VHLTPEEKNAVTLWGKVNVDEVGGEALGRLLLVYPWTQRFFESFGDLSSPDAVMGNPKV		60		
Sbjct 1	VHLTPEEKNAVTLWGKVNVDEVGGEALGRLLLVYPWTQRFFESFGDLSSPDAVMGNPKV		60		
Query 61	KAHGKKVLGAFSDGLNHLDNLKGTFAQLSELHCDKLHVDPENFKLLGNVLVCVLAHHFGK		120		
Sbjct 61	KAHGKKVLGAFSDGLNHLDNLKGTFAQLSELHCDKLHVDPENFKLLGNVLVCVLAHHFGK		120		
Query 121	EFTPQVQAAYQKVVAGVANALAHKYH	146			
Sbjct 121	EFTPQVQAAYQKVVAGVANALAHKYH	146			

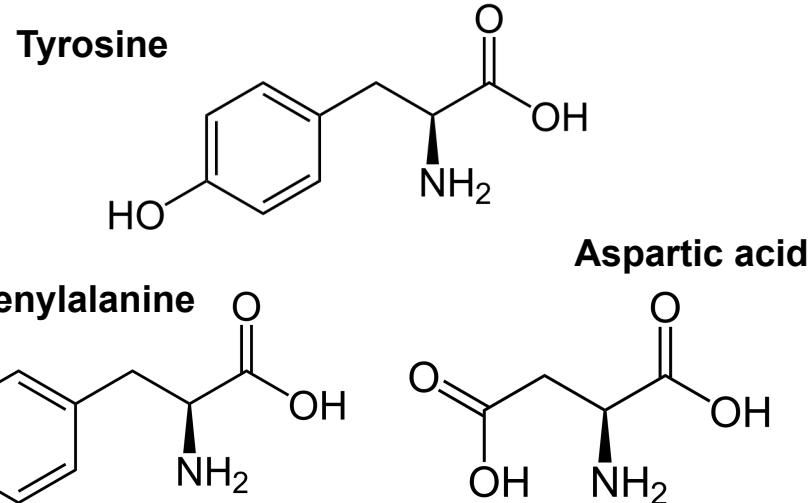
Related Information

[Gene](#) - associated gene details

[Identical Proteins](#) - Identical proteins to P02026.1

BLAST Results

- Score: Higher is better
 - Bit score is a normalized score in search space.
- Expect (E-value): Lower is better
 - Tell us about this match is occurred by chance or others
- Identity
 - Percentage of Matched Sequence
- Positive
 - Indicate a conservative substitution or substitutions that are often observed in related proteins.
- Gap
 - Percentage of GAP



Score	Expect	Method	Identities	Positives	Gaps
281 bits(718)	6e-95	Compositional matrix adjust.	137/146(94%)	141/146(96%)	0/146(0%)
Query 1	VHLTPEEKNAVTTLWGKVNVDEVGGEALGRLLL VYPWTQRFFESFGDLS+PDAVMGNPKV				60
Sbjct 1	VHLTPEEK+AVT LWGKVNVDEVGGEALGRLL+VYPWTQRFFESFGDLS+PDAVMGNPKV				60
Query 61	KAHGKKVGLGAFSDGLNHLNDNLKGTF AQLSELHCDKLHVDPENFKLLGNVLVCVLAHHFGK				120
Sbjct 61	KAHGKKVGLGAFSDGL HLNDNLKGTF LSELHCDKLHVDPENF+LLGNLV VLAAHFGK				120
Query 121	EFTPQVQAAYQKVVAGVANALAHKYH	146			
Sbjct 121	EFTP VQAAYQKVVAGVANALAHKYH				
	EFTPPVQAAYQKVVAGVANALAHKYH	146			

Examples Using BLAST

- Make Phylogenetic Tree by **mitochondrial genome** and **protein sequence**
- Annotate a Metagenomic Contig and SARS-CoV-2 Genome

Species and Gene Phylogenetic Tree

- Species Phylogenetic Tree by Mitochondrial Genome
- NC_012920.1

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s)

Query subrange

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 (nr etc.) rRNA/ITS databases Genomic + transcript databases Betacoronavirus

Organism RefSeq Genome Database (refseq_genomes)

Optional Mammalia (taxid:40674) exclude

Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown

Exclude Models (XM/XP) Uncultured/environmental sample sequences

Optional Limit to Sequences from type material

Entrez Query Create custom database
Optional 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

Algorithm parameters

General Parameters

Max target sequences Select the maximum number of aligned sequences to display

Short queries Automatically adjust parameters for short input sequences

Expect threshold

Word size

Max matches in a query range

Scoring Parameters

Match/Mismatch Scores

Gap Costs Existence: 5 Extension: 2

Filters and Masking

Filter Low complexity regions

Mask Species-specific repeats for: Homo sapiens (Human)

Mask for lookup table only

Mask lower case letters

Discontiguous Word Options

Template length

Template type



todste@orcid

BLAST® » blastn suite » results for RID-VBCMNM2Z013

[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](#)

Your search is limited to records that include: Mammalia (taxid:40674), and entrez query: mitochondrion[Filter]

Job Title **NC_012920:Homo sapiens mitochondrion, complete...**RID **VBCMNM2Z013** Search expires on 01-06 00:08 am [Download All](#) ▾Program BLASTN [Citation](#) ▾Database refseq_genomes (197 databases) [See details](#) ▾Query ID **NC_012920.1**

Description Homo sapiens mitochondrion, complete genome

Molecule type nucleic acid

Query Length 16569

Other reports [Distance tree of results](#) [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

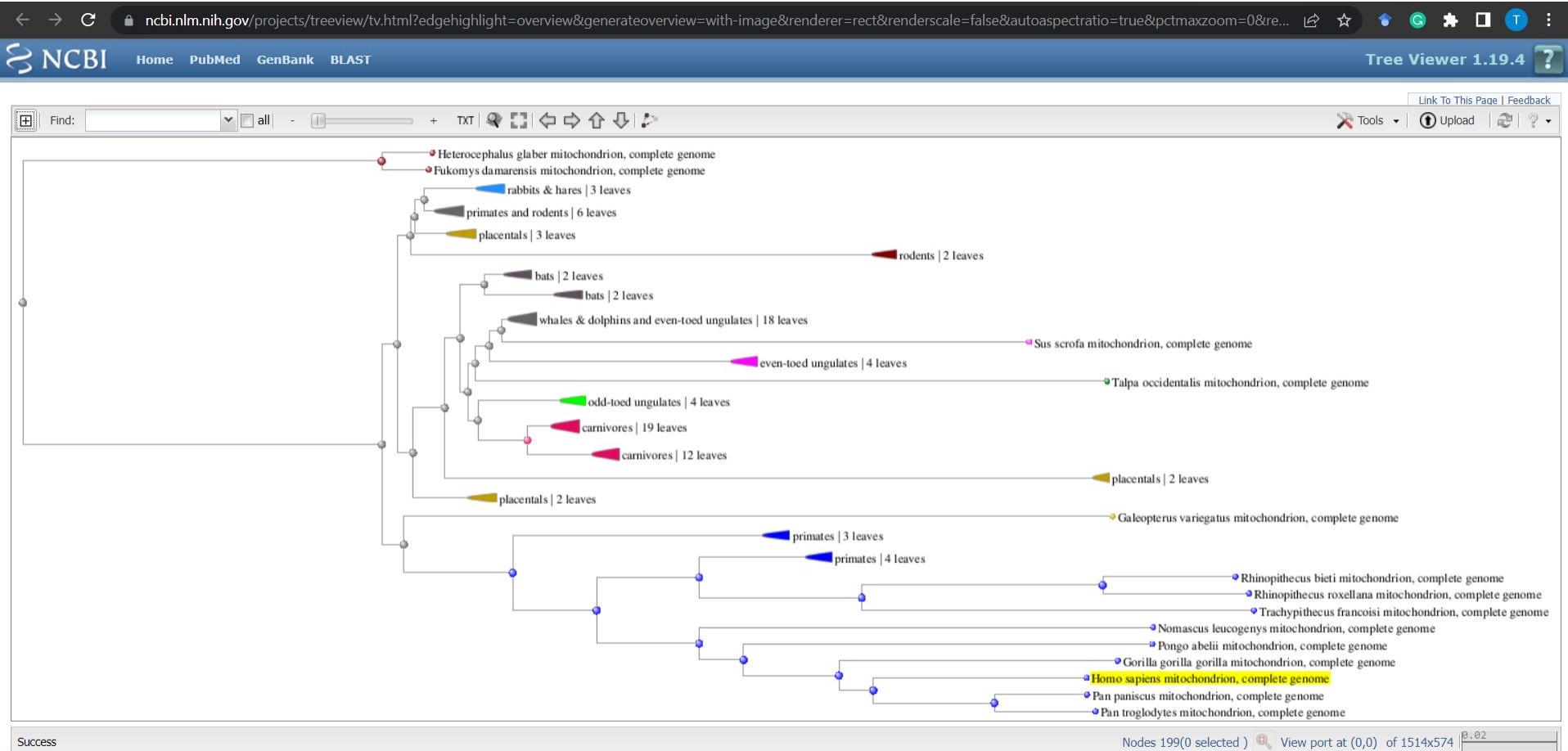
 to [Filter](#)[Reset](#)**Descriptions**[Graphic Summary](#)[Alignments](#)[Taxonomy](#)**Sequences producing significant alignments**[Download](#) ▾[Select columns](#) ▾

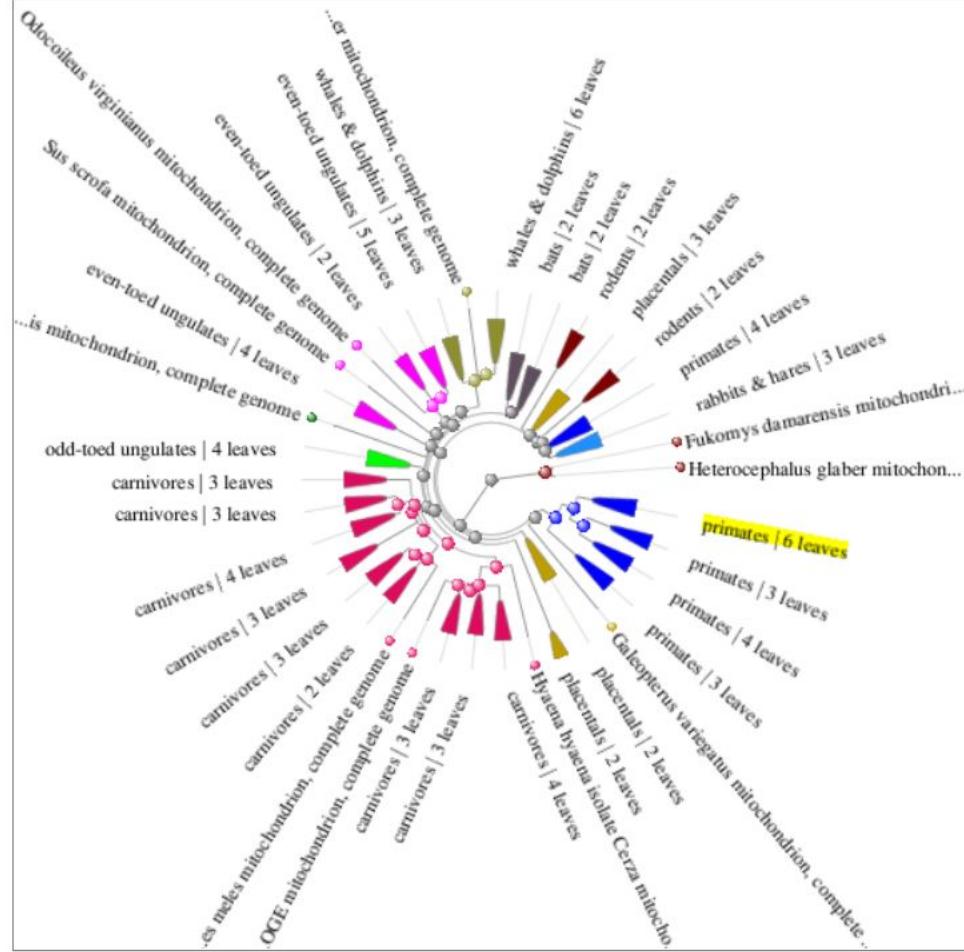
Show 100 ▾

 select all 100 sequences selected[GenBank](#) [Graphics](#) [Distance tree of results](#) [MSA Viewer](#)

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession

Descriptions		Graphic Summary	Alignments	Taxonomy	Download		Select columns		Show	100	?
Sequences producing significant alignments					GenBank		Graphics		Distance tree of results		MSA Viewer
	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession		
<input type="checkbox"/>	Homo sapiens mitochondrion complete genome	Homo sapiens	29877	29877	100%	0.0	100.00%	16569	NC_012920.1		
<input checked="" type="checkbox"/>	Pan paniscus mitochondrion complete genome	Pan paniscus	22610	23265	100%	0.0	91.41%	16563	NC_001644.1		
<input checked="" type="checkbox"/>	Pan troglodytes mitochondrion complete genome	Pan troglodytes	22529	23179	99%	0.0	91.29%	16554	NC_001643.1		
<input checked="" type="checkbox"/>	Gorilla gorilla gorilla mitochondrion complete genome	Gorilla gorilla gorilla	20712	21436	99%	0.0	89.48%	16412	NC_011120.1		
<input checked="" type="checkbox"/>	Pongo abelii mitochondrion complete genome	Pongo abelii	18295	18521	97%	0.0	85.44%	16499	NC_002083.1		
<input checked="" type="checkbox"/>	Nomascus leucogenys mitochondrion complete genome	Nomascus leucogenys	17158	17461	97%	0.0	83.97%	16478	NC_021957.1		
<input checked="" type="checkbox"/>	Theropithecus gelada mitochondrion complete genome	Theropithecus gelada	14022	14022	96%	0.0	79.65%	16546	NC_019802.1		
<input checked="" type="checkbox"/>	Macaca fascicularis mitochondrion complete genome	Macaca fascicularis	13858	13858	100%	0.0	78.89%	16575	NC_012670.1		
<input checked="" type="checkbox"/>	Macaca mulatta mitochondrion complete genome	Macaca mulatta	13692	13692	96%	0.0	79.17%	16564	NC_005943.1		
<input checked="" type="checkbox"/>	Papio anubis isolate east mitochondrion complete genome	Papio anubis	13584	13584	96%	0.0	79.08%	16516	NC_020006.2		
<input checked="" type="checkbox"/>	Rhinopithecus bieti mitochondrion complete genome	Rhinopithecus bieti	13130	13130	96%	0.0	78.51%	16551	NC_015486.1		
<input checked="" type="checkbox"/>	Trachypithecus francoisi mitochondrion complete genome	Trachypithecus francoisi	13084	13261	97%	0.0	78.37%	16554	NC_023970.1		
<input checked="" type="checkbox"/>	Rhinopithecus roxellana mitochondrion complete genome	Rhinopithecus roxellana	12819	12819	96%	0.0	78.06%	16549	NC_008218.1		
<input checked="" type="checkbox"/>	Callithrix jacchus mitochondrion complete genome	Callithrix jacchus	11352	11352	93%	0.0	76.77%	16499	NC_025586.1		
<input checked="" type="checkbox"/>	Aotus nancymaae mitochondrion complete genome	Aotus nancymaae	11212	11212	93%	0.0	76.53%	16469	NC_018116.1		
<input checked="" type="checkbox"/>	Saimiri boliviensis boliviensis mitochondrion complete genome	Saimiri boliviensis boliviensis	10883	10883	93%	0.0	76.08%	16473	NC_018096.1		
<input checked="" type="checkbox"/>	Ceratotherium simum mitochondrion complete genome	Ceratotherium simum	10559	10559	93%	0.0	75.68%	16832	NC_001808.1		
<input checked="" type="checkbox"/>	Equus przewalskii mitochondrial DNA complete genome isolate NOUMA1	Equus przewalskii	10238	10238	93%	0.0	75.20%	16592	NC_024030.1		
			10167	10167	93%	0.0	75.10%	16260	NC_014626.1		





Gene Phylogenetic Tree of ADH1A

BLASTP programs search protein databases using a protein query. more... Reset page Bookmark

Enter Query Sequence

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

NP_000658.1

From To

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

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 Optional Mammalia (taxid:40674) exclude [Add organism](#)
Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown [?](#)

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

Program Selection

Algorithm blastp (protein-protein BLAST)
 PSI-BLAST (Position-Specific Iterated BLAST)
 PHI-BLAST (Pattern Hit Initiated BLAST)
 DELTA-BLAST (Domain Enhanced Lookup Time Accelerated BLAST)
Choose a BLAST algorithm [?](#)

BLAST Show results in a new window

Am x | Inb x | Dat x | Intr x | BLA x | NCI x | We x | NCI x | BLA x | We x | Boo x | Nu x | Bla x | NC x | Mu x | +

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

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

National Library of Medicine
National Center for Biotechnology Information

todste@orcid

BLAST® » blastp suite » results for RID-VBGM5F2T016

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

Your search is limited to records that exclude: uncultured/environmental sample sequences, non-redundant RefSeq proteins (WP)

Job Title ref|NP_000658.1|

RID VBGM5F2T016 Search expires on 01-06 01:16 am [Download All](#)

Program BLASTP [Citation](#)

Database landmark [See details](#)

Query ID NP_000658.1

Description alcohol dehydrogenase 1A [Homo sapiens]

Molecule type amino acid

Query Length 375

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

E value

Query Coverage

[] to []

[] to []

[] to []

Filter

Reset

Descriptions

Graphic Summary

Alignments

Taxonomy

Sequences producing significant alignments

Download

Select columns

Show

100

select all 91 sequences selected

[GenPept](#) [Graphics](#) [Distance tree of results](#) [Multiple alignment](#) [MSA Viewer](#)

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	alcohol dehydrogenase 1A [Homo sapiens]	Homo sapiens	764	764	100%	0.0	100.00%	375	NP_000658.1
<input checked="" type="checkbox"/>	all-trans-retinol dehydrogenase [NAD(+)]ADH1B isoform 1 [Homo sapiens]	Homo sapiens	722	722	100%	0.0	93.60%	375	NP_000659.2
<input checked="" type="checkbox"/>	alcohol dehydrogenase 1C [Homo sapiens]	Homo sapiens	715	715	100%	0.0	92.80%	375	NP_000660.1

Feedback

undefined.pdf

current_Archaea....fa.gz

current_Bacteria....fa.gz

Booklet_Teaching....pdf

influenza_1163 (2).fasta

beta_globin (1).fasta

Show all



21°C ແລ້ວນີ້ 1/5/2023

COBALT

Constraint-based Multiple Alignment Tool

Home Recent Results Help

Phylogenetic Tree Edit and Resubmit Back to Blast Results ▶ Download

Multiple Alignment Results - ref|NP_000658.1| - Cobalt RID VBGP4MYM212 (100 seqs)

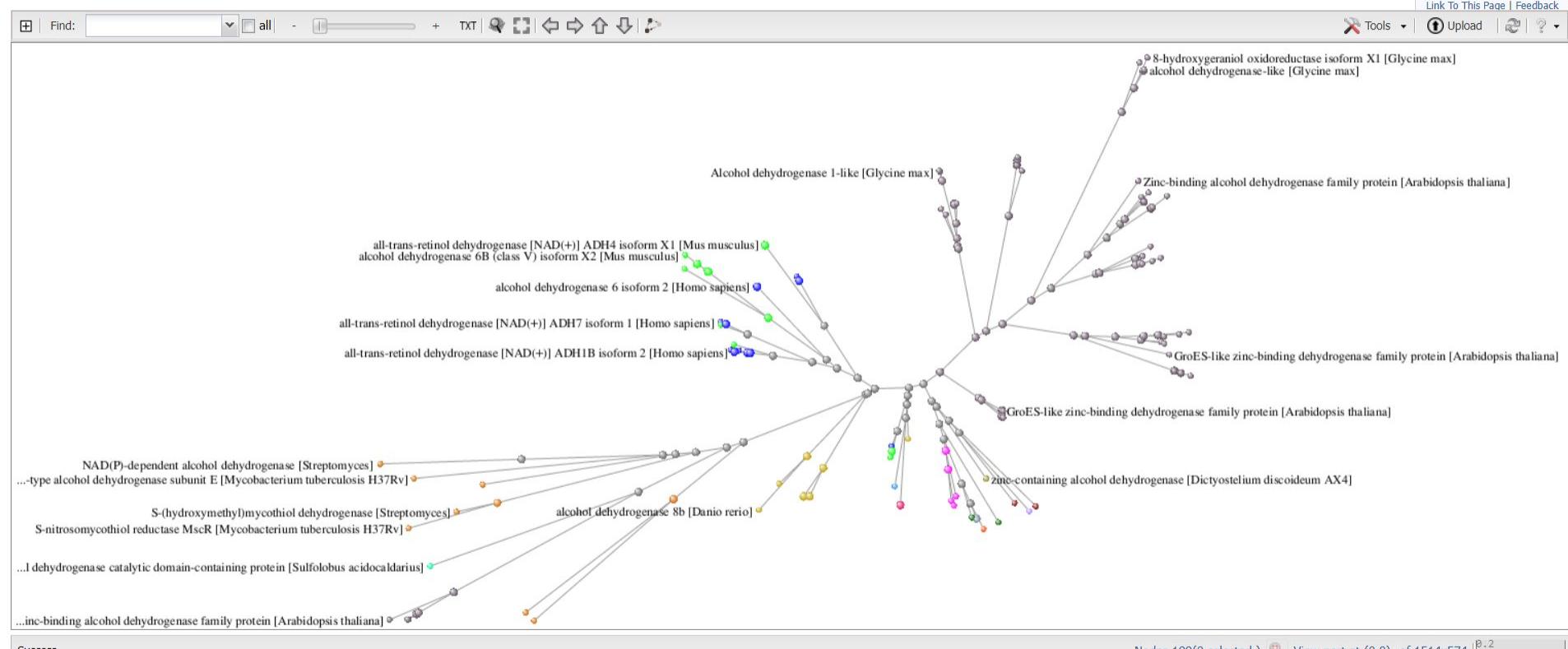
▼ Graphical Overview

373 - 413 (41r shown) | Find: 374 376 378 380 382 384 386 388 390 392 394 396 398 400 402 404 406 408 410 412 End | Organism

Sequence ID Start 374 376 378 380 382 384 386 388 390 392 394 396 398 400 402 404 406 408 410 412 End Organism

NP_000658.1	1	A K E L G A T E C I N P Q D	Y K K P I Q E V L K E M T	375	Homo sapiens
NP_000659.2	1	A K E L G A T E C I N P Q D	Y K K P I Q E V L K E M T	375	Homo sapiens
NP_000660.1	1	A K E L G A T E C I N P Q D	Y K K P I Q E V L K E M T	375	Homo sapiens
NP_001273579.1	1	A K E L G A T E C I N P Q D	Y K K P I Q E V L K E M T	335	Homo sapiens
NP_031435.1	1	A K E L G A T E C I N P Q D	Y S K P I Q E V L Q E M T	375	Mus musculus
NP_000664.3	1	A M A V G A T E C I S P K D	S T K P I S E V L S E M T	374	Homo sapiens
NP_001159976.1	1	A M A V G A T E C I S P K D	S T K P I S E V L S E M T	394	Homo sapiens
NP_033756.2	1	A L A V G A T E C I S P K D	S T K P I S E V L S D M T	374	Mus musculus
NP_571924.2	1	A K F G A T E F V N P K D	H S K P I Q E V L V E L T	376	Danio rerio
NP_001095940.1	1	A Q E L G A T E C L N P Q D	L K K P I Q E V L F D M T	375	Homo sapiens
NP_000663.1	1	A Q E L G A T E C L N P Q D	L K K P I Q E V L F D M T	368	Homo sapiens
NP_000662.3	1	A K E F G A T E C I N P Q D	F S K P I Q E V L I E M T	374	Homo sapiens
NP_031436.2	1	A K E F G A S E C I S P Q D	F S K S I Q E V L V E M T	374	Mus musculus
NP_000661.2	1	A K A L G A T D C L N P R D	L H K P I Q E V I I E L T	380	Homo sapiens

PROTEIN: 373 - 413 (41r shown)



Annotate a Metagenomic Contig

- MIZB01000007.1
- Use blastx

BLAST® » blastx » results for RID-VBJ32449016

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 **gb|MIZB01000007.1|**

RID [VBJ32449016](#) Search expires on 01-06 01:41 am [Download All](#) ▾

Program BLASTX Citation ▾

Database landmark See details ▾

Query ID **MIZB01000007.1**

Description Marine metagenome, whole genome shotgun sequence

Molecule type dna

Query Length 13728

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

Descriptions Graphic Summary Alignments Taxonomy

hover to see the title click to show alignments

40 sequences selected ?

Alignment Scores ■ < 40 ■ 40 - 50 ■ 50 - 80 ■ 80 - 200 ■ >= 200 ?

Distribution of the top 40 Blast Hits on 40 subject sequences

Score Range	Approximate Number of Hits
< 40	1
40 - 50	1
50 - 80	1
80 - 200	1
>= 200	~35

Annotate coronavirus genome

BLAST® > blastx

Translated BLAST: blastx

Home Recent Results Saved Strategies Help

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

OL989059.1

Query subrange [?](#)

From
To

Or, upload file Choose 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 Optional Severe acute respiratory syndrome coronavirus 2 (taxid:2697049) exclude [Add organism](#) [?](#)
Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown [?](#)

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

BLAST

Search database refseq_protein using Blastx (search protein databases using a translated nucleotide query)
 Show results in a new window

Note: Parameter values that differ from the default are highlighted in yellow and marked with * sign

Algorithm parameters

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

i Your search is limited to records that include: Severe acute respiratory syndrome coronavirus 2 (taxid:2697049)

Job Title **gb|OL989059.1|**RID **VBK4HZPK013** Search expires on 01-06 01:59 am [Download All](#) ▾Program **BLASTX** [?](#) [Citation](#) ▾Database **refseq_protein** [See details](#) ▾Query ID **OL989059.1**

Description Severe acute respiratory syndrome coronavirus 2 isolate S...

Molecule type rna

Query Length 29903

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](#)[Descriptions](#)[Graphic Summary](#)[Alignments](#)[Taxonomy](#)[hover to see the title](#)[click to show alignments](#)

Alignment Scores

 < 40 40 - 50 50 - 80 80 - 200 >= 200[?](#)25 sequences selected [?](#)**Distribution of the top 26 Blast Hits on 25 subject sequences**