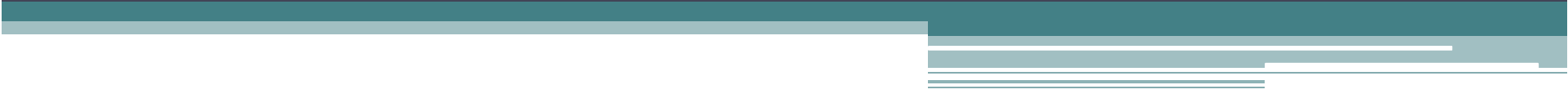


# NCBI

**BLAST-It finds region of similarity between biological sequences.**

A decorative graphic at the bottom of the slide consisting of a solid teal horizontal bar, followed by a thin white line, and then three thin, parallel teal lines of varying lengths extending to the right.

## AIM OF THE EXPERIMENT

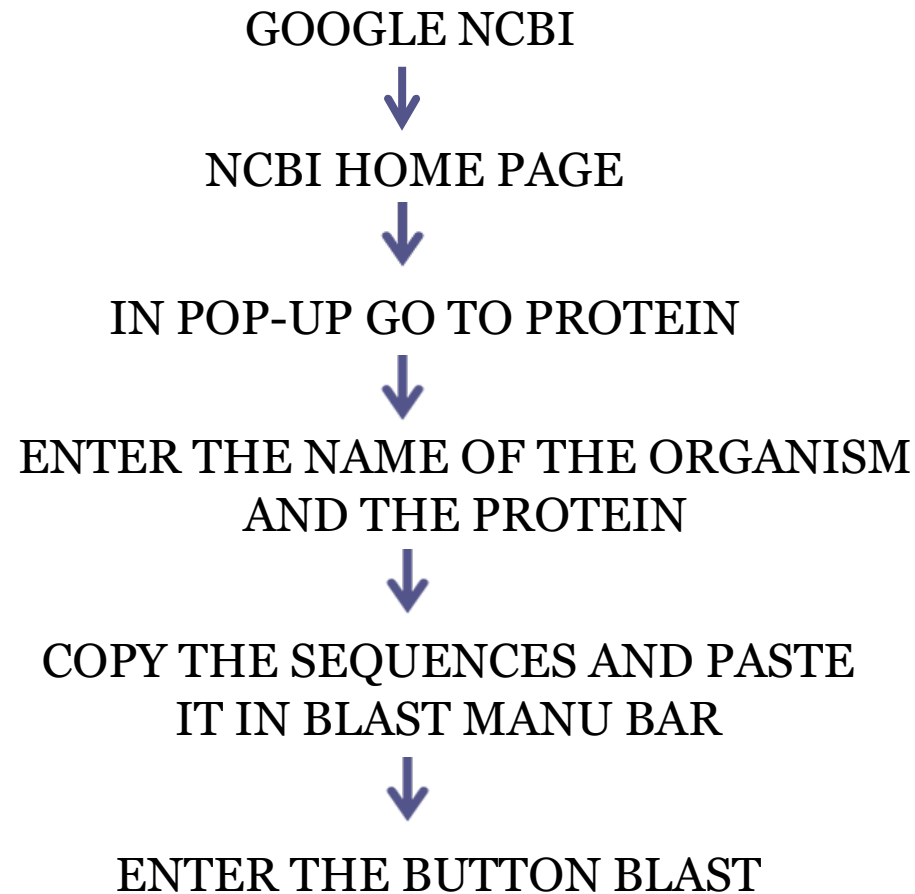
To do NCBI BLAST for comparison of sequences between various living organisms.

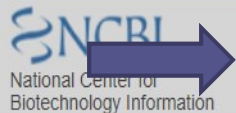
## PRINCIPLE

- BLAST can be used for several purposes. These include identifying species, locating domains, establishing phylogeny, DNA mapping, and comparison.
- When working with a known species, and looking to sequence a gene at an unknown location, **BLAST** can compare the chromosomal position of the sequence of interest, to relevant sequences in the database.
- BLAST can be used for several purposes. These include identifying species, locating domains, establishing phylogeny, DNA mapping, and comparison.
- When working with a protein sequence you can input it into BLAST, to locate known domains within the sequence of interest.
- When working with genes, BLAST can locate common genes in two related species, and can be used to map annotations from one organism to another.

## Procedure

- **Procedure for comparison of the sequences between various living organism.**





Protein ▾

human prion



Search



COVID-19 is an emerging, rapidly evolving situation.  
 Get the latest public health information from CDC: <https://www.coronavirus.gov>.  
 Get the latest research from NIH: <https://www.nih.gov/coronavirus>.  
 Find NCBI SARS-CoV-2 literature, sequence, and clinical content: <https://www.ncbi.nlm.nih.gov/sars-cov-2/>.

NCBI Home

Resource List (A-Z)

All Resources

Chemicals & Bioassays

Data & Software

DNA & RNA

Domains & Structures

Genes & Expression

Genetics & Medicine

Genomes & 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

Bookshelf

PubMed Central

BLAST

Nucleotide

Genome

SNP

Gene

Protein

PubChem

## NCBI News & Blog

New annotations in RefSeq: white-tufted-ear marmoset, ruddy duck, and more  
 12 Aug 2020

In June and July, the NCBI Eukaryotic Genome Annotation Pipeline released

Recent enhancements in Genome Workbench version 3.4.1  
 06 Aug 2020

New Features Version 3.4.1 of Genome Workbench: NCBI's sequence annotation

New interaction data, downloads and track hub available for RefSeq Functional

Protein

Protein

Advanced

Search

Help



COVID-19 is an emerging, rapidly evolving situation.  
 Get the latest public health information from CDC: <https://www.coronavirus.gov>.  
 Get the latest research from NIH: <https://www.nih.gov/coronavirus>.  
 Find NCBI SARS-CoV-2 literature, sequence, and clinical content: <https://www.ncbi.nlm.nih.gov/sars-cov-2/>.

FASTA

Send to:

Change region shown

## prion [Homo sapiens]

GenBank: BAG32276.1

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



>BAG32276.1 prion [Homo sapiens]

```
MANLGCWMLVLFVATWSDLGLCKRKPGGWNTGGSRYPGQSPGGNRYPPQGGGGWGQPHGGGWGQPHG
GGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGGLGGYMLGSAMSRPIIH
FGSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERVV
EQMCITQYERESQAYYQRGSSMVLFSPPVILLISFLIFLIVG
```

### Analyze this sequence

Run BLAST

Identify Conserved Domains

Highlight Sequence Features

Find in this Sequence

### Protein 3D Structure



Cryo-EM structure of amyloid fibril formed by full-length human prion protein  
 PDB: 6LNI  
 Source: Homo sapiens  
 Method: Electron Microscopy

Resolution: 2.702 Å

See all 47 structures...

### Articles about the PRNP gene

Prion protein transcription is auto-regulated through dyn [Biochim Biophys Acta Gene Regu...]

Prion protein-Semisynthetic prion protein (PrP)

This is the amino acid sequence of Human prion protein, I am copying this to do protein blast.

All Databases ▾

Search



COVID-19 is an emerging, rapidly evolving situation.  
Get the latest public health information from CDC: <https://www.coronavirus.gov>.  
Get the latest research from NIH: <https://www.nih.gov/coronavirus>.  
Find NCBI SARS-CoV-2 literature, sequence, and clinical content: <https://www.ncbi.nlm.nih.gov/sars-cov-2/>.

NCBI Home

Resource List (A-Z)

All Resources

Chemicals & Bioassays

Data & Software

DNA & RNA

Domains & Structures

Genes & Expression

Genetics & Medicine

Genomes & 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

Bookshelf

PubMed Central

**BLAST**

Nucleotide

Genome

SNP

Gene

Protein

PubChem

### NCBI News & Blog

New annotations in RefSeq: white-tufted-ear marmoset, ruddy duck, and more  
12 Aug 2020

In June and July, the NCBI Eukaryotic Genome Annotation Pipeline released

Recent enhancements in Genome Workbench version 3.4.1  
06 Aug 2020

New Features Version 3.4.1 of Genome Workbench: NCBI's sequence annotation



COVID-19 is an emerging, rapidly evolving situation.

Get the latest public health information from CDC: <https://www.coronavirus.gov>.

Get the latest research from NIH: <https://www.nih.gov/coronavirus>.

Find NCBI SARS-CoV-2 literature, sequence, and clinical content: <https://www.ncbi.nlm.nih.gov/sars-cov-2/>.

## 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+ 2.10.1 is released – Fix for TBLASTN Multi-Threading issue.

This version supports pulling databases from our FTP site as well from [cloud providers](#) or our [BLAST+Docker solution](#).

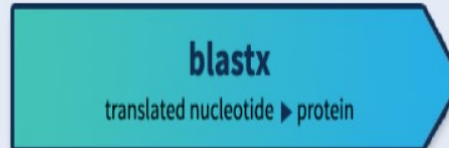
Thu, 18 June 2020 12:00:00 EST

[More BLAST news...](#)

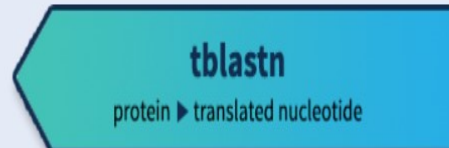
## Web BLAST



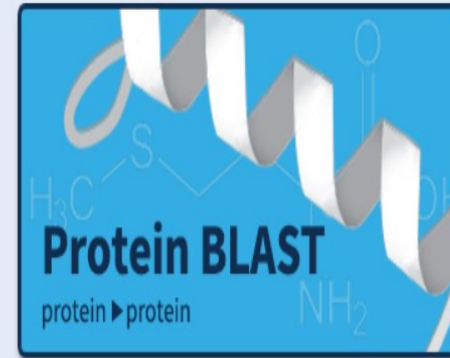
**Nucleotide BLAST**  
nucleotide ▶ nucleotide



**blastx**  
translated nucleotide ▶ protein



**tblastn**  
protein ▶ translated nucleotide



**Protein BLAST**  
protein ▶ protein

Paste the sequence that we have copied/sequenced.

## Translated BLAST: tblastn

blastn blastp blastx **tblastn** tblastx

TBLASTN search translated nucleotide 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

```
MANLGCNMLVLFVATWSDLGLCKKRPKGGWINTGGSRYPGQGSPPGNRYPPQGGGGWQPHGGGWQPHG
GGWQPHGGGWQPHGGGWQGGGTHSQWINKPSKPKTNWIKHMAGAAAAGAVVGGGGYMLGSMSRPIIH
FGSDYEDRYRENIHRYPNQVYRPMDEYSNQNFFVHDCVNIITIKQHTVTTTTKGENFTETDVKMHERV
EQWCITQYERESQAYYORGSSHWLFSPPVILLISFLTELTVG
```

From

To

**BLAST results will be displayed  
in a new format by default**

You can always switch back to the  
Traditional Results page.



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

Nucleotide collection (nr/nt)

Organism

Optional

Enter organism name or id—completions will be suggested

exclude

Enter 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

Entrez Query

Optional

[YouTube](#) [Create custom database](#)

Enter an Entrez query to limit search

**BLAST**

Search database Nucleotide collection (nr/nt) using Tblastn (search translated nucleotide databases using a protein query)

Show results in a new window

[+ Algorithm parameters](#)



# Results

BLAST® » blastp suite » results for RID-KD1DJ6EJ016

[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	dbj BAG32276.1
RID	<a href="#">KD1DJ6EJ016</a> Search expires on 08-15 20:38 pm <a href="#">Download All</a> ▾
Program	BLASTP <a href="#">?</a> <a href="#">Citation</a> ▾
Database	nr <a href="#">See details</a> ▾
Query ID	<a href="#">BAG32276.1</a>
Description	prion [Homo sapiens]
Molecule type	amino acid
Query Length	253
Other reports	<a href="#">Distance tree of results</a> <a href="#">Multiple alignment</a> <a href="#">MSA viewer</a> <a href="#">?</a>

## 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](#) ▾

[Manage Columns](#) ▾

Show

100 ▾

[?](#)

select all 100 sequences selected

[GenPept](#)

[Graphics](#)

[Distance tree of results](#)

[Multiple alignment](#)

	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
<input checked="" type="checkbox"/>	<a href="#">major prion protein preproprotein Prp_precursor [Homo sapiens]</a>	520	520	100%	0.0	100.00%	<a href="#">NP_000302.1</a>
<input checked="" type="checkbox"/>	<a href="#">prion protein [synthetic construct]</a>	520	520	100%	0.0	99.60%	<a href="#">AAX37089.1</a>
<input checked="" type="checkbox"/>	<a href="#">prion protein (p27-30) (Creutzfeld-Jakob disease, Gerstmann-Strausler-Scheinker syndrome, fatal familial insomnia) [synthetic construct]</a>	520	520	100%	0.0	99.60%	<a href="#">AAV38282.1</a>
<input checked="" type="checkbox"/>	<a href="#">major prion protein [Gorilla gorilla gorilla]</a>	519	519	100%	0.0	99.60%	<a href="#">XP_004061812.1</a>
<input checked="" type="checkbox"/>	<a href="#">prion protein PrP [Homo sapiens]</a>	519	519	100%	0.0	99.60%	<a href="#">ABD63004.1</a>

[Feedback](#)

These are the sequences which are similar to the sequence that we have provided.

### major prion protein preproprotein Prp precursor [Homo sapiens]

Sequence ID: [NP\\_000302.1](#) Length: **253** Number of Matches: **1**

[See 10 more title\(s\)](#) [See all Identical Proteins\(IPG\)](#)

Range 1: 1 to 253 [GenPept](#) [Graphics](#)

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

Score	Expect	Method	Identities	Positives	Gaps
520 bits(1340)	0.0	Compositional matrix adjust.	253/253(100%)	253/253(100%)	0/253(0%)
Query 1		MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQSPGGNRYPPQGGGGWGQP			60
Sbjct 1		MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQSPGGNRYPPQGGGGWGQP			60
Query 61		HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAGA			120
Sbjct 61		HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAGA			120
Query 121		VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENMHRYPNQVYYRPMDEYSNQNNFVHDCV			180
Sbjct 121		VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENMHRYPNQVYYRPMDEYSNQNNFVHDCV			180
Query 181		NITIKQHTVTTTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV			240
Sbjct 181		NITIKQHTVTTTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV			240
Query 241		ILLISFLIFLIVG		253	
Sbjct 241		ILLISFLIFLIVG		253	

#### Related Information

[Gene](#) - associated gene details  
[Genome Data Viewer](#) - aligned genomic context  
[Identical Proteins](#) - Identical proteins to NP\_000302.1

### prion protein, partial [synthetic construct]

Sequence ID: [AAX37089.1](#) Length: **254** Number of Matches: **1**

Range 1: 1 to 253 [GenPept](#) [Graphics](#)

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

Score	Expect	Method	Identities	Positives	Gaps
520 bits(1338)	0.0	Compositional matrix adjust.	252/253(99%)	253/253(100%)	0/253(0%)
Query 1		MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQSPGGNRYPPQGGGGWGQP			60
Sbjct 1		MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQSPGGNRYPPQGGGGWGQP			60

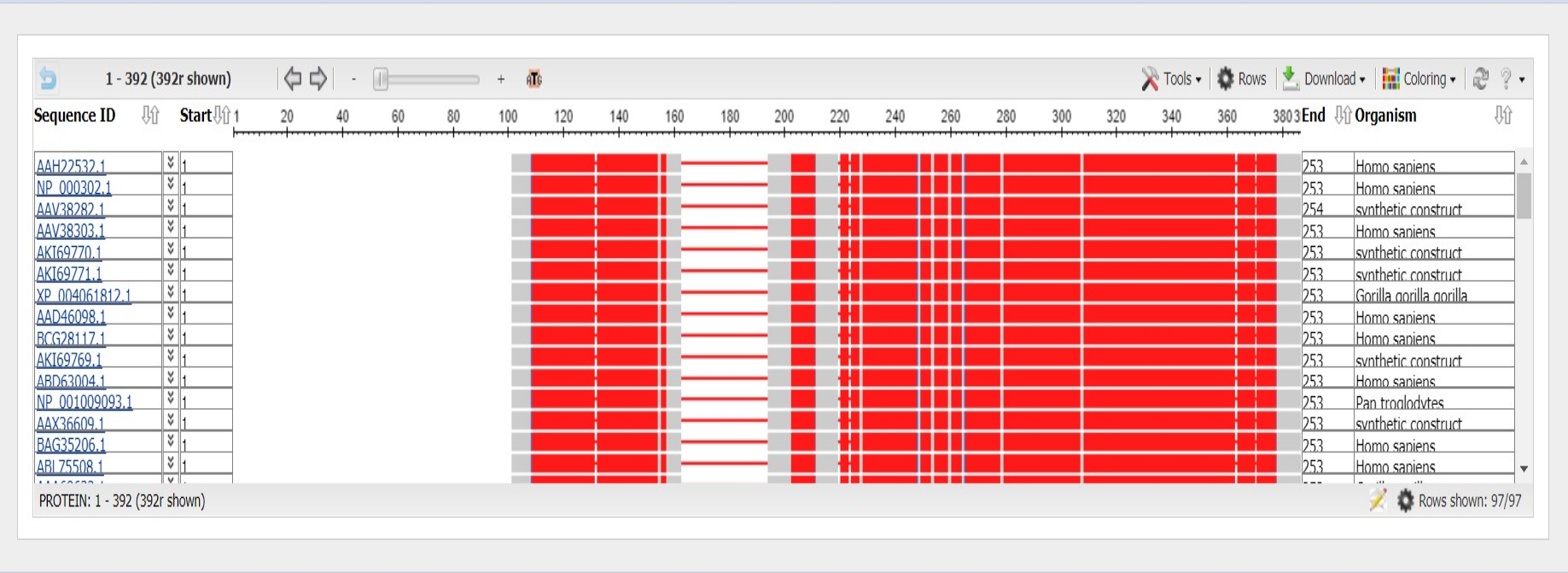




Phylogenetic Tree Edit and Resubmit Back to Blast Results Download

### Multiple Alignment Results - gb|AAH22532.1| - Cobalt RID PVS6A6RX212 (97 seqs)

#### Graphical Overview



#### Descriptions Select All Alignment parameters

Accession	Description	Links
<input checked="" type="checkbox"/> <a href="#">AAH22532.1</a>	Prion protein [Homo sapiens]	<a href="#">Related Information</a>
<input checked="" type="checkbox"/> <a href="#">NP_000302.1</a>	major prion protein preproprotein Prp precursor [Homo sapiens]	<a href="#">Related Information</a>
<input checked="" type="checkbox"/> <a href="#">AAV38282.1</a>	prion protein (p27-30) (Creutzfeld-Jakob disease, Gerstmann-Strausler-Scheinker syndrome, fatal familial insomnia) [synthetic construct]	<a href="#">Related Information</a>
<input checked="" type="checkbox"/> <a href="#">AAV38303.1</a>	prion protein (p27-30) (Creutzfeld-Jakob disease, Gerstmann-Strausler-Scheinker syndrome, fatal familial insomnia) [Homo sapiens]	<a href="#">Related Information</a>
<input checked="" type="checkbox"/> <a href="#">AKI69770.1</a>	PRNP [synthetic construct]	
<input checked="" type="checkbox"/> <a href="#">AKI69771.1</a>	PRNP [synthetic construct]	
<input checked="" type="checkbox"/> <a href="#">XP_004061812.1</a>	major prion protein [Gorilla gorilla gorilla]	<a href="#">Related Information</a>
<input checked="" type="checkbox"/> <a href="#">AAD46098.1</a>	prion protein [Homo sapiens]	<a href="#">Related Information</a>

## Conclusion

- The list of hits starts with the best match (most similar).
- E-value: expected number of chance alignments; the smaller the E-value, the better the match.
- First in the list is the query sequence itself, which obviously has the best score.
- The Accession Number: a unique code which identifies a sequence in a database (in this case it is the GenBank number).
- **E[expect] Value:** the number of alignments expected by chance with the calculated score or better. The expect value is the default sorting metric; for significant alignments the E value should be very close to zero.
- **Identity:** the highest percent identity for a set of aligned segments to the same subject sequence.