

Question 1

Name: TDP-43, Tar DNA binding protein.

Assession ref_seq gene ID: NP_031401

Species: Homo Sapiens

Question 2:

Method: TBLASTN search against ESTs using NCBI

Data base: Expressed Sequence Tags

Organism: Nematodes

The screenshot shows the NCBI BLAST search interface. The top section is titled 'Enter Query Sequence' and contains a text input field with 'NP_031401' entered. To the right of this field is a 'Query subrange' section with 'From' and 'To' input fields. Below the text input field is a section for uploading a file, with a 'Choose File' button and a 'No file chosen' status. Below that is a 'Job Title' section with a text input field containing 'NP_031401:TAR DNA-binding protein 43 [Homo...'. There is also a checkbox for 'Align two or more sequences'. The bottom section is titled 'Choose Search Set' and contains several options: 'Database' set to 'Expressed sequence tags (est)', 'Organism' set to 'nematodes (taxid:6231)' with an 'exclude' checkbox and an 'Add organism' button, 'Exclude' options for 'Models (XM/XP)' and 'Uncultured/environmental sample sequences', 'Limit to' option for 'Sequences from type material', and an 'Entrez Query' field. There are also links for 'YouTube' and 'Create custom database'.

Chosen match:

Accession number [FF098943.1](#), a 646 basepair length mRNA sequences from caenorhabditis brenneri. See below for more details.

Job Title

NP_031401:TAR DNA-binding protein 43 [Homo...

RID

XZTBZ2ES013 Search expires on 02-07 00:21 am [Download All](#) ▼

Program

TBLASTN [?](#) [Citation](#) ▼

Database

est [See details](#) ▼

Query ID

NP_031401.1

Description

TAR DNA-binding protein 43 [Homo sapiens]

Molecule type

amino acid

Query Length

414

Other reports

[?](#)

Filter Results

Organism

only top 20 will appear ☐ exclu

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

10 sequences selected [?](#)

Distribution of the top 10 Blast Hits on 10 subject sequences

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

Research Professor [shiLab-Zoom](#) [Biology of the Extr...](#) [GitHub - ShuJiaLa...](#) [M.Eng. Program |...](#) [BENG186B](#)

Query Length 414

Other reports [?](#)

Descriptions

Graphic Summary

Alignments

Taxonomy

Sequences producing significant alignments

Download ▼

Select columns ▼

Show 10 ▼ [?](#)

☒ select all 10 sequences selected

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

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	CPAD-aab47f04.b1 PB2801 EST CPAD1 Caenorhabditis brenneri cDNA similar to ref NP_495921.1 RNA binding protein...	Caenor...	130	130	41%	2e-34	41.21%	646	FF098943.1
<input checked="" type="checkbox"/>	HY709935 full-length enriched (vector-capping) cDNA library; dev stage: mixed Diploscapter coronatus cDNA clone ndv15...	Diplosc...	124	124	42%	2e-31	37.81%	824	HY709935.1
<input checked="" type="checkbox"/>	BAAC-PNP127510.g1 C.remanei EST SB146 Caenorhabditis remanei cDNA 5' mRNA sequence	Caenor...	124	124	42%	2e-31	38.05%	761	DT933781.1
<input checked="" type="checkbox"/>	HY576189 full-length enriched (oligo-capping) cDNA library; dev stage: mixed Diploscapter coronatus cDNA clone ndf110...	Diplosc...	122	122	35%	7e-31	42.50%	704	HY576189.1
<input checked="" type="checkbox"/>	HY821971 full-length enriched (vector-capping) cDNA library; dev stage: embryo-enriched Diploscapter coronatus cDNA g...	Diplosc...	108	108	38%	6e-26	36.96%	743	HY821971.1
<input checked="" type="checkbox"/>	rg03b03.y1 Meloidogyne hapla J2 pAMP1 v1 Meloidogyne hapla cDNA 5' similar to TR:O97469 O97469 TAR-BINDING P...	Meloido...	104	104	36%	3e-25	36.42%	490	CA996783.1
<input checked="" type="checkbox"/>	AGSL1p5.6.7.8 O23 082.ab1 Pristionchus pacificus SL1 library Pristionchus pacificus cDNA 5' mRNA sequence	Pristion...	105	105	41%	4e-25	34.78%	580	FG096118.1
<input checked="" type="checkbox"/>	AHA1-aaa87q05.b1 Pristionchus pacificus AHA1 EST L1 Pristionchus pacificus cDNA similar to ref NP_495921.1 RNA...	Pristion...	100	100	39%	3e-23	33.90%	593	FE934547.1
<input checked="" type="checkbox"/>	FN872627 Caenorhabditis elegans whole worm N2 all stages Caenorhabditis elegans cDNA clone abi3100 ce2 2003081...	Caenor...	100	100	42%	7e-23	34.80%	745	FN872627.1
<input checked="" type="checkbox"/>	FN872626 Caenorhabditis elegans whole worm N2 all stages Caenorhabditis elegans cDNA clone abi3100 ce2 2002100...	Caenor...	99.0	99.0	35%	8e-23	36.36%	582	FN872626.1

Descriptions Graphic Summary **Alignments** Taxonomy

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

10 sequences selected

[Download](#) [GenBank](#) [Graphics](#) [Next](#) [Previous](#) [Descriptions](#)

CPAD-aab47f04.b1 PB2801_EST_CPAD1 Caenorhabditis brenneri cDNA similar to ref|NP_495921.1| RNA binding protein like (45.5 kD) [Caenorhabditis elegans] pir|T22207 hypothetical protein F44G4.4 - Caenorhabditis elegans emb|CAA90120.1| Hypothetical protein F44G4.4 [Caenorhabditis elegans], mRNA sequence

Sequence ID: [FF098943.1](#) Length: 646 Number of Matches: 1

Range 1: 78 to 617 [GenBank](#) [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps	Frame
130 bits(328)	2e-34	Compositional matrix adjust.	75/182(41%)	113/182(62%)	14/182(7%)	+3
Query 97	KRAVQKTS---	DLIVLGLPWKTTEQDLKEYFSTFGEVLMVQVKKDLKTGHSKGFVFRT	153			
Sbjct 78	KRAV++ S DLIVLG+ +KT++ ++YF G V+ ++K+ G+SKGFGFVR +	KRAVERDSQPVDLIVLGVDFKTSDDCFQKYFEEIGTVVFCEIKRK-SDGNSKGFVFVRMS	254			
Query 154	EYETQVKVMS-QRHMIDGRWCDCKLPNSKQSQDEPLRSRKVFVGRCTEDMTDELREFFS	212				
Sbjct 255	Q KV++ +HMIDGR CD K+P+ + Q P SR +FVGR T+ + E +LR+ F	SVGEQNKVLAIPQHMDGRRCDVKVPDGRDKQGRPSISR-IFVGRRLTDKVDEYQLRKVFG	431			
Query 213	-----QYGDVMDVFIPKPPRAFAFVTFADDQIAQSLCGE-DLIIGISVHISNAEPKH	264				
Sbjct 432	+ V DVFIPKPPR FAFV+ + + A+ + + L + G+SV +S A+P+	DEAKSYIESAVVTDVFIKPPFRGFAFVSLSSAAEAERIVSKGSLTVNGLSVGLSIAQPRE	611			
Query 265	NS 266					
Sbjct 612	EN 617					

Question 3:

>C. elegans protein (sequence taken from BLAST result)

```
KRAVERDSQPVDLIVLGVDFKTSDDCFQKYFEEIGTVVFCEIKRK-SDGNSKGFVFVRMS
SVGEQNKVLAIPQHMDGRRCDVKVPDGRDKQGRPSISR-IFVGRRLTDKVDEYQLRKVFG
DEAKSYIESAVVTDVFIKPPFRGFAFVSLSSAAEAERIVSKGSLTVNGLSVGLSIAQPRE
EN
```

Name:

```
CPAD-aab47f04.b1 PB2801_EST_CPAD1 Caenorhabditis brenneri cDNA
similar to ref|NP_495921.1| RNA binding protein like (45.5 kD)
[Caenorhabditis elegans] pir|T22207 hypothetical protein F44G4.4
-
Caenorhabditis elegans emb|CAA90120.1| Hypothetical protein
F44G4.4
[Caenorhabditis elegans], mRNA sequence
```

ORGANISM [Caenorhabditis brenneri](#)

Eukaryota; Metazoa; Ecdysozoa; Nematoda; Chromadorea; Rhabditida;
Rhabditina; Rhabditomorpha; Rhabditoidea; Rhabditidae;
Peloderinae;Caenorhabditis.

Question 4:

A BLASTP search against NR database (see setup in first screen-shot below) yielded a top hit result is to a protein from C.elegans

Top result is to a protein from C. elegans. Tar DNA-binding protein homolog 1 was identified.

Research Professor shiLab-Zoom Biology of the Extr... GitHub - ShuJiaLa... M.Eng. Program |... BENG186B

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

Job Title C. elegans protein (sequence taken from BLAST...

RID XZUR3FNN013 [Search expires on 02-07 00:44 am](#) [Download All](#)

Program BLASTP [Citation](#)

Database nr [See details](#)

Query ID lcl|Query_23318

Description C. elegans protein (sequence taken from BLAST result)

Molecule type amino acid

Query Length 180

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 **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 [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"/>	Tar DNA-binding protein homolog 1 [Caenorhabditis elegans]	Caenorhabditis elegans	357	357	100%	9e-121	96.67%	411	NP_001022166.1
<input checked="" type="checkbox"/>	Tar DNA-binding protein homolog 1 [Caenorhabditis elegans]	Caenorhabditis elegans	352	352	100%	1e-118	95.08%	414	NP_001254189.1
<input checked="" type="checkbox"/>	hypothetical protein B9Z55_005513 [Caenorhabditis nigoni]	Caenorhabditis nigoni	348	348	100%	5e-117	94.44%	414	PIC45529.1

[Download](#) [GenPept](#) [Graphics](#) [Next](#) [Previous](#) [Descriptions](#)

Tar DNA-binding protein homolog 1 [Caenorhabditis elegans]

Sequence ID: [NP_001022166.1](#) Length: 411 Number of Matches: 1

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

Range 1: 163 to 342 [GenPept](#) [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
357 bits(917)	9e-121	Compositional matrix adjust.	174/180(97%)	180/180(100%)	0/180(0%)
Query 1	KRAVERDSQPVDLIVLGVDFTSDDCFOKYFEEIGTVVFCEIKRSDGNSKGFGFVRMSS				60
Sbjct 163	KRAVERDSQPVDLIVLGVDFTD+D+CFQKYFE+IGTVVFCEIKRSDGNSKGFGFVRMSS				222
Query 61	VGEQNKVLAIPQHMIDGRRCDVKVPDGRDKQGRPSISRIFVGRGLTDKVDYQLRKVFGE				120
Sbjct 223	VGEQNKVLAIPQHMIDGRRCDVKVPDGRDKQGRPSISRIFVGRGLTDKVDYQLRKVFGE				282
Query 121	AKSYIESAVVTDVFIPKPFGRGFAFVLSLSSAAEAERIVSKGSLTVNGLSVGLSIAQPREEN				180
Sbjct 283	AKSYIETAVVTDVFIPKPFGRGFAFVLSLSSAAEAERIVSKGSLTVNGLSVGLSIAQPREEN				342

Related Information

[Gene](#) - associated gene details

[Genome Data Viewer](#) - aligned genomic context

[Identical Proteins](#) - Identical proteins to NP_001022166.1