BIMM143 Find-a-Gene Project

Aaron Ta | A12766546 | acta@ucsd.edu

2/12/2019

[Q1]

Tell me the name of a protein you are interested in. Include the species and the accession number. This can be a human protein or a protein from any other species as long as it's function is known. If you do not have a favorite protein, select human RBP4 or KIF11. Do not use beta globin as this is in the worked example report that I provide you with online.

Name: NRN1 (neuritin 1)

Accession: AAI58491

Species: *Xenopus tropicalis*

Function: Neuritin is a neurotrophic factor expressed by the developing nervous system. It promotes neuronal arborization and is associated with neurogenesis as well as adult neuronal plasticity.

[Q2]

Perform a BLAST search against a DNA database, such as a database consisting of genomic DNA or ESTs. The BLAST server can be at NCBI or elsewhere. Include details of the BLAST method used, database searched and any limits applied (e.g. Organism).

See *Fig 1.*

Method: NCBI tblastn 2.8.1

Database: Expressed sequence tags (est)

Organism: Include fish (taxid:7898), exclude *Danio rerio* (taxid:7955)

- As a neurotrophic factor, neuritin is known to be found in vertebrates, and has been annotated in many mammalian species. *Danio rerio* was excluded due to its usage as a model organism in neurobiology, and thus its likely already pre-existing annotation.

Chosen Match: DT227993.1, a 724 bp clone from *Pimephales promelas* adults (first result; *Fig 2*)

- **E-Value:** 4e-58
- **Query Score:** 81%
- **Score:** 185
- **Identities:** 71.79%

Alignments Download GenBank Graphics

Description	Max score	Total score	Query cover	E value	Ident	Accession
JGI_CAAT3511.fwd CAAT Pimephales promelas brain 7-8 month adults, males and females pooled (M) Pimephales promelas cDNA clone CAAT3511 5', mRN	185	185	81%	4e-58	71.79%	DT227993.1
JGI_CAAU3507.fwd CAAU Pimephales promelas brain 7-8 month adults, males and females pooled (L) Pimephales promelas cDNA clone CAAU3507 5', mRI	185	185	81%	5e-58	71.79%	DT249077.1
JGI_CAAT7544.fwd CAAT Pimephales promelas brain 7-8 month adults, males and females pooled (M) Pimephales promelas cDNA clone CAAT7544 5', mRN	185	185	81%	7e-58	71.79%	DT233298.1
JGI_CAAU5333.fwd CAAU Pimephales promelas brain 7-8 month adults, males and females pooled (L) Pimephales promelas cDNA clone CAAU5333 5', mRI	185	185	81%	7e-58	71.79%	DT252382.1
JGI_CAAT4482.fwd CAAT Pimephales promelas brain 7-8 month adults, males and females pooled (M) Pimephales promelas cDNA clone CAAT4482 5', mRN	185	185	81%	1e-57	71.79%	DT229781.1
JGI_CAAU6839.fwd CAAU Pimephales promelas brain 7-8 month adults, males and females pooled (L) Pimephales promelas cDNA clone CAAU6839 5', mRI	185	185	81%	1e-57	71.79%	DT255094.1
JGI_CAAU8249.fwd CAAU Pimephales promelas brain 7-8 month adults, males and females pooled (L) Pimephales promelas cDNA clone CAAU8249 5', mRI	185	185	81%	1e-57	71.79%	DT257663.1
omykbra004057 Oncorhynchus mykiss head Oncorhynchus mykiss cDNA, mRNA sequence	181	181	81%	5e-57	70.94%	CB492194.1
643815 NCCCWA 1RT Oncorhynchus mykiss cDNA clone 1RT145B24_D_A12 5', mRNA sequence	181	181	81%	7e-57	70.94%	CA367815.1
1310093 NCCCWA 02RT Oncorhynchus mykiss cDNA clone 02RT119B24 5', mRNA sequence	181	181	81%	7e-57	70.94%	CX257079.1
FO328086 Astyanax mexicanus whole embryos and larvae neurula to swimming larvae Astyanax mexicanus cDNA clone ARA0AEA15YD02, mRNA sequenc	180	180	84%	9e-57	70.25%	FO328086.1
CcLX06a22e08f1 Carp mixed tissue library 3 Cyprinus carpio cDNA clone 22e08 5' similar to human.pt ref|NP_057672.1|, mRNA sequence	179	179	81%	1e-56	69.49%	CA969310.2
EST_otsh_evc_5 otshevc mixed_tissue Oncorhynchus tshawytscha cDNA Oncorhynchus tshawytscha cDNA clone otsh_evc_001_004_fwd 3', mRNA sequen	177	177	81%	2e-56	70.09%	EL551836.1
EST_omor_rgc_15185 omorrgc mixed_tissue Osmerus mordax cDNA Osmerus mordax cDNA clone omor_rgc_520_060_rev 5', mRNA sequence	182	182	81%	3e-56	70.09%	EL551816.1
EST_omor_rgc_8974 omorrgc mixed_tissue Osmerus mordax cDNA Osmerus mordax cDNA clone omor_rgc_511_073_rev 5', mRNA sequence	181	181	81%	5e-56	70.09%	EL545021.1
EST_omor_rgc_2682 omorrgc mixed_tissue Osmerus mordax cDNA Osmerus mordax cDNA clone omor_rgc_502_204_rev 5', mRNA sequence	181	181	81%	8e-56	70.09%	EL538729.1
FO371109 Astyanax mexicanus whole embryos and larvae neurula to swimming larvae Astyanax mexicanus cDNA clone ARA0AGA47YJ20, mRNA sequence	177	177	84%	1e-55	70.25%	FO371109.1
FO233837 Astyanax mexicanus whole embryos and larvae neurula to swimming larvae Astyanax mexicanus cDNA clone ARA0AAA60YG12, mRNA sequenc	178	178	83%	2e-55	70.83%	FO233837.1
EST ssal rgb2 11707 rgb2 Salmo salar cDNA clone ssal rgb2 519 303 rev 5', mRNA sequence	177	177	81%	2e-55	67.80%	DW547288.1

Figure 1:

Download GenBank Graphics ▼ Next ▲ Previous Descriptions

JGI_CAAT3511.fwd CAAT Pimephales promelas brain 7-8 month adults, males and females pooled (M) Pimephales promelas cDNA clone CAAT3511 5', mRNA sequence

Sequence ID: DT227993.1 **Length:** 724 **Number of Matches:** 1

Related Information

UniGene - clustered expressed sequence tags

Range 1: 162 to 512 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
185 bits(469)	4e-58	Compositional matrix adjust.	84/117(72%)	99/117(84%)	0/117(0%)	+3

```
Query  1    MGLKLSGRYIFLVLAVHLAYLLQAVRAAGKCDAVFKGLSDCMLTLGDRVANYPQDLEEKK  60
            MGL LSGRYI L LAV +AYLLQAVRAAGKC+ VFKG SDC+L LGD +ANYPQ+L+E++
Sbjct  162  MGLTLSGRYISLFLAVQIAYLLQAVRAAGKCETVFKGFSDCLLHLGDNMANYPQELDEQE  341

Query  61   NLDTICSYWDDFHVCTVTALADCQEGAADIWEKLKRQSKNLNIQGSLFELCPGSTGA  117
            NL TIC+YWDDFH C  TALADCQEGA D+WEKLK++S++L  +GSLFELC G  GA
Sbjct  342  NLKTICTYWDDFHSCATTALADCQEGATDLWEKLKKESRSLEFRGSLFELCAGGNGA  512
```

Figure 2:

[Q3]

Chosen Sequence (Fig 2):

- >*P. promelas* protein (from BLAST result)
- MGLTLSGRYISLFLAVQIAYLLQAVRAAGKCETVFKGFSDCLLHLGDNMANYPQELDEQENLKTICTYWD DFHSCATTALADCQEGATDLWEKLKKESRSLEFRGSLFELCAGGNGA

Name: *Pimephales* neuritin

Species: *Pimephales promelas* ('Fathead minnow')

- *Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Pimephales.*

[Q4]

See *Fig 3*.

There are no 100% matches; the closest matches, however, are all to species in the carp family (*Cyrprinidae*), which *P. promelas* also belongs to. The results of this blastp search were not filtered specifically to include or exclude any species and were run against the *nr* database.

Top alignment is a predicted neuritin-like homolog for *Cyprinus carpio* (*Fig 4*).

- **E-Value:** 1e-77
- **Query Score:** 100%
- **Score:** 234
- **Identity:** 94.87%

Description	Max score	Total score	Query cover	E value	Ident	Accession
PREDICTED: neuritin-like [Cyprinus carpio]	234	234	100%	1e-77	94.87%	XP_018922093.1
neuritin-like [Carassius auratus]	231	231	100%	2e-76	94.02%	XP_026093433.1
neuritin [Carassius auratus]	231	231	100%	2e-76	94.02%	XP_026056648.1
PREDICTED: neuritin [Cyprinus carpio]	231	231	100%	2e-76	94.02%	XP_018922095.1
neuritin precursor [Danio rerio]	231	231	100%	4e-76	94.02%	NP_001002507.1
PREDICTED: neuritin [Sinocyclocheilus rhinocerous]	230	230	100%	5e-76	93.16%	XP_016374848.1
PREDICTED: neuritin [Sinocyclocheilus grahami]	228	228	100%	3e-75	92.31%	XP_016138196.1
PREDICTED: neuritin-like [Sinocyclocheilus rhinocerous]	227	227	100%	7e-75	92.31%	XP_016401820.1
neuritin [Electrophorus electricus]	227	227	100%	9e-75	91.45%	XP_026869098.1
neuritin [Oncorhynchus mykiss]	222	222	100%	7e-73	89.74%	XP_021418750.1
PREDICTED: neuritin [Esox lucius]	222	222	100%	1e-72	88.89%	XP_010899235.1
neuritin [Salvelinus alpinus]	221	221	100%	2e-72	88.89%	XP_023829168.1
PREDICTED: neuritin [Salmo salar]	221	221	100%	2e-72	88.89%	XP_014036114.1
unnamed protein product [Oncorhynchus mykiss]	221	221	100%	3e-72	88.89%	CDQ56010.1
PREDICTED: neuritin [Pygocentrus nattereri]	220	220	100%	6e-72	89.74%	XP_017559428.1
neuritin [Salvelinus alpinus]	219	219	100%	1e-71	87.18%	XP_023856162.1
PREDICTED: neuritin-like [Salmo salar]	219	219	100%	2e-71	87.18%	XP_014013044.1
neuritin [Oncorhynchus kisutch]	219	219	100%	2e-71	88.03%	XP_020350443.1
unnamed protein product [Oncorhynchus mykiss]	218	218	100%	5e-71	86.32%	CDQ56600.1
neuritin [Astyanax mexicanus]	217	217	100%	6e-71	88.03%	XP_007238696.2

Figure 3:

Download GenPept Graphics ▼ Next ▲ Previous ▲ Descriptions

PREDICTED: neuritin-like [Cyprinus carpio]

Sequence ID: XP_018922093.1 **Length:** 142 **Number of Matches:** 1

See 1 more title(s)

Range 1: 1 to 117 GenPept Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps
234 bits(598)	1e-77	Compositional matrix adjust.	111/117(95%)	116/117(99%)	0/117(0%)

```
Query  1    MGLTLSGRYISLFLAVQIAYLLQAVRAAGKCETVFKGFSDCLLHLGDNMANYPQELDEQE  60
            MGLTLSGRYISLFLAVQIAYLLQAVRAAGKCETVFKGFSDCLLHLG+NMANYPQELDE+E
Sbjct  1    MGLTLSGRYISLFLAVQIAYLLQAVRAAGKCETVFKGFSDCLLHLGENMANYPQELDEKE  60

Query  61   NLKTICTYWDDFHSCATTALADCQEGATDLWEKLKKESRSLEFRGSLFELCAGGNGA  117
            NL+TICTYWDDFHSCATTALADCQEGATDLWEKLKKESR+L+FRGSLFELCA GNGA
Sbjct  61   NLQTICTYWDDFHSCATTALADCQEGATDLWEKLKKESRNLDFRGSLFELCAAGNGA  117
```

Related Information

Gene - associated gene details

New Genome Data Viewer - aligned genomic context

Identical Proteins - Identical proteins to XP_018922093.1

Figure 4: