Ralph Goguanco

A15937787

BIMM-143: FIND-A-GENE PROJECT

[Q1] Name: Adenosine Deaminase

Accession: NP\_000013

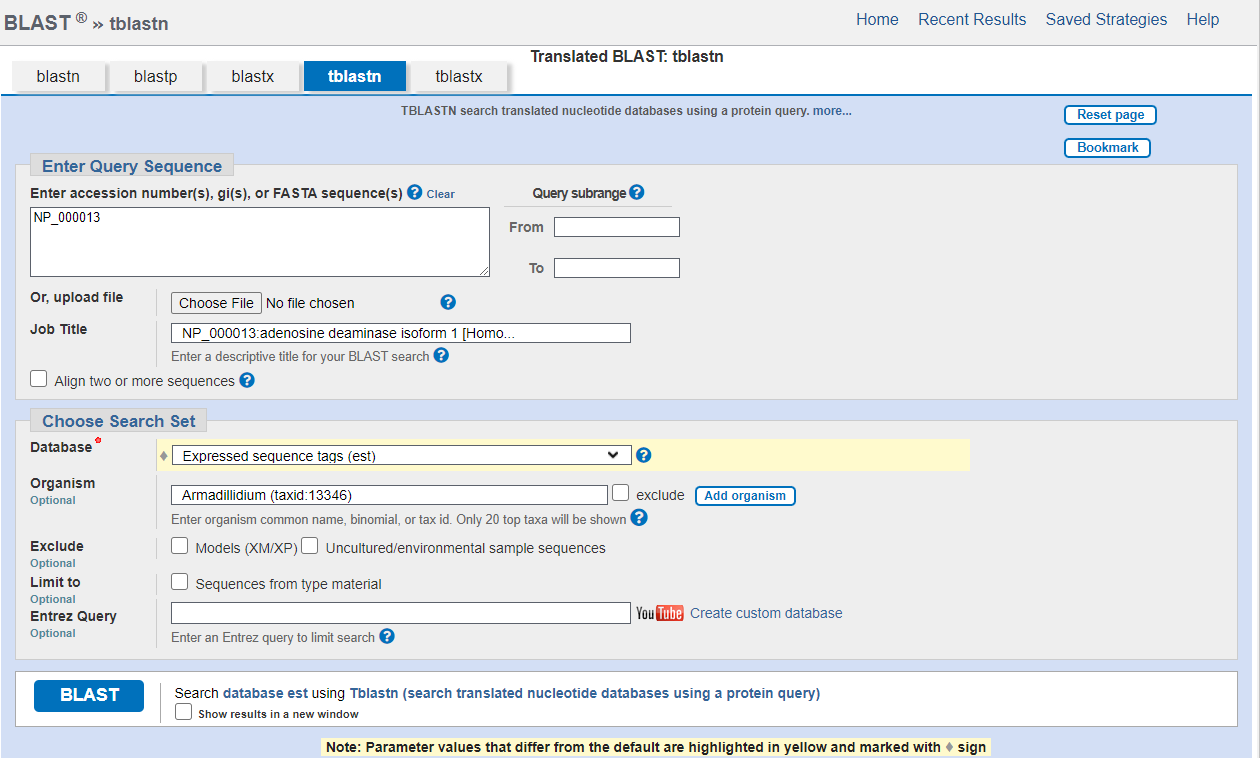
Species: Homo Sapiens

[Q2] Blast Method: TBLASTN search against Armadillidium

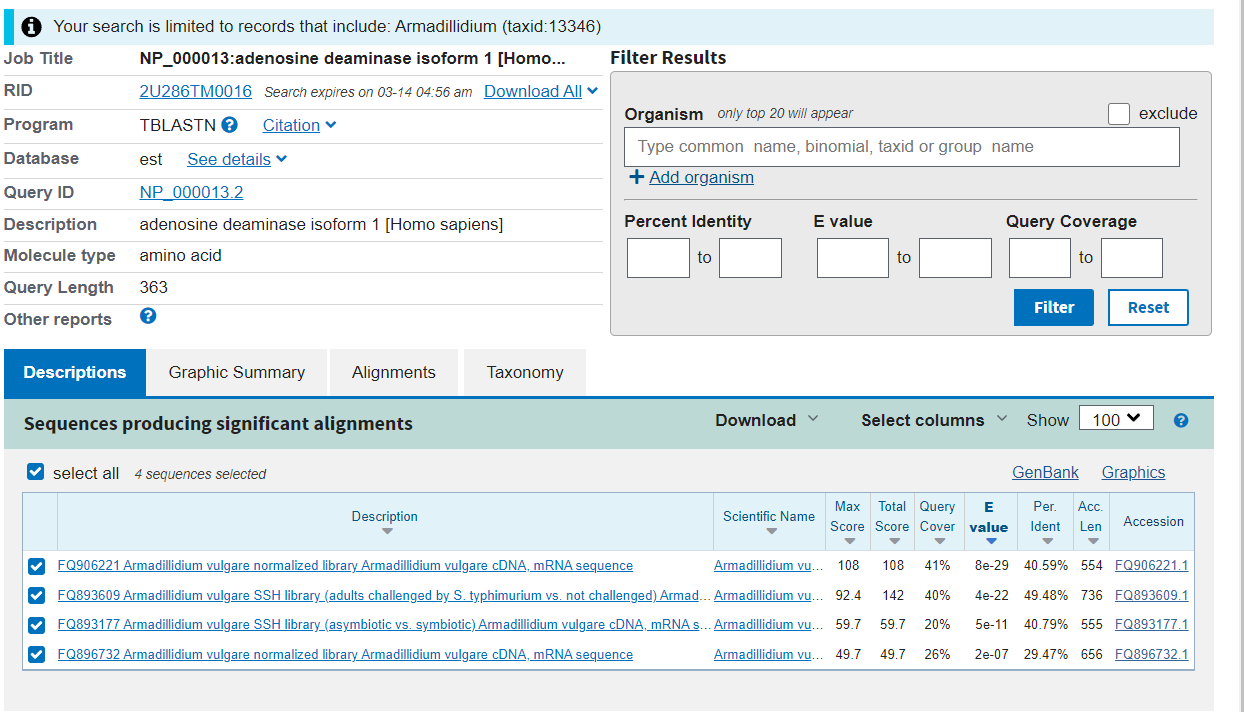
Database: Expressed Sequence Tags (est)

Organism: Armadillidium (taxid:13346)

Limits Applied:

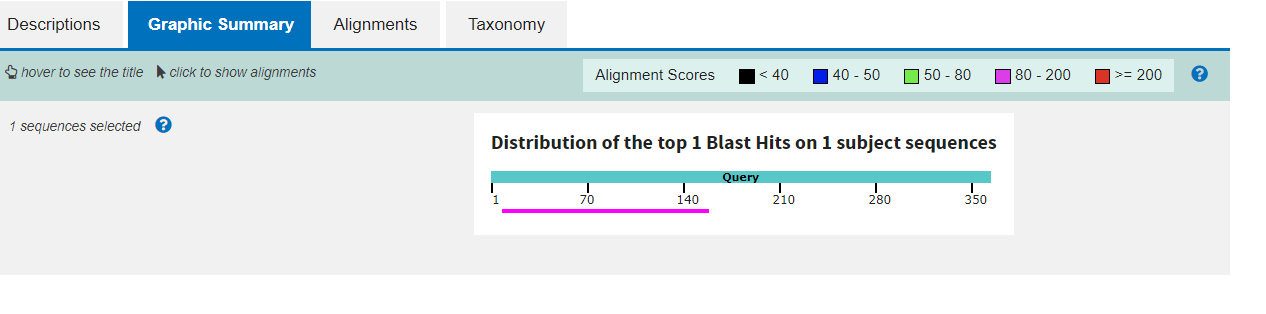


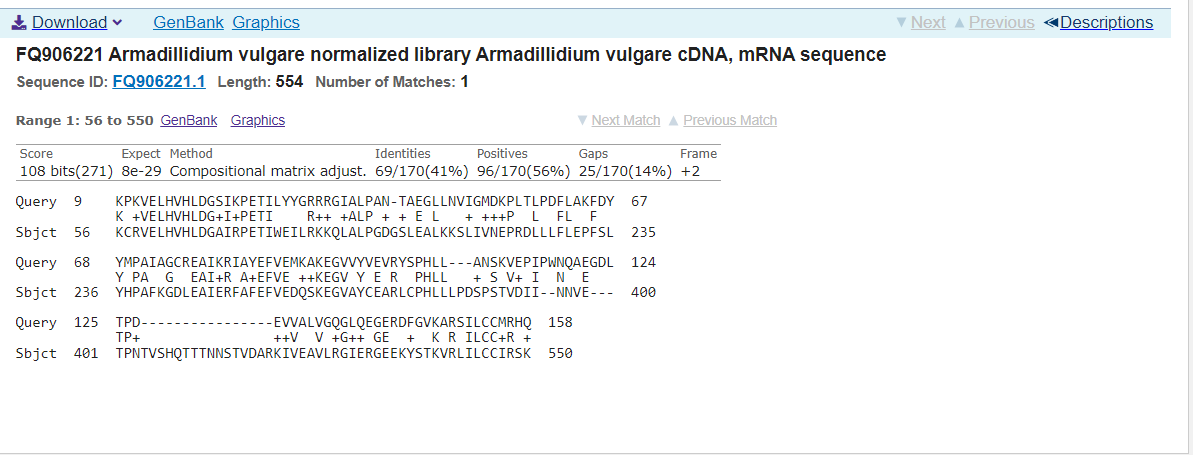
Search output list (top hits):



Alignment of Choice: Accession FQ906221.1, a 108-alignment score match. See below for alignment details.

Alignment Statistics: E value = 8e-29, Total Score = 108, Percent Identity = 40.59%, Length 554





**Alignment details:**

>FQ906221 Armadillidium vulgare normalized library Armadillidium vulgare cDNA, mRNA sequence

Sequence ID: FQ906221.1 Length: 554

Range 1: 56 to 550

Score:108 bits(271), Expect:8e-29,

Method:Compositional matrix adjust.,

Identities:69/170(41%), Positives:96/170(56%), Gaps:25/170(14%)

Query 9 KPKVELHVHLDGSIKPETILYYGRRRGIALPAN-TAEGLLNVIGMDKPLTLPDFLAKFDY 67

K +VELHVHLDG+I+PETI R++ +ALP + + E L + +++P L FL F

Sbjct 56 KCRVELHVHLDGAIRPETIWEILRKKQLALPGDGSLEALKKSLIVNEPRDLLLFLEPFSL 235

Query 68 YMPAIAGCREAIKRIAYEFVEMKAKEGVVYVEVRYSPHLL---ANSKVEPIPWNQAEGDL 124

Y PA G EAI+R A+EFVE ++KEGV Y E R PHLL + S V+ I N E

Sbjct 236 YHPAFKGDLEAIERFAFEFVEDQSKEGVAYCEARLCPHLLLPDSPSTVDII--NNVE--- 400

Query 125 TPD----------------EVVALVGQGLQEGERDFGVKARSILCCMRHQ 158

TP+ ++V V +G++ GE + K R ILCC+R +

Sbjct 401 TPNTVSHQTTTNNSTVDARKIVEAVLRGIERGEEKYSTKVRLILCCIRSK 550

[Q3] Chosen sequence:

>56-550\_1 Armadillidium vulgare normalized library Armadillidium vulgare cDNA, mRNA sequence

KCRVELHVHLDGAIRPETIWEILRKKQLALPGDGSLEALKKSLIVNEPRDLLLFLEPFSL

YHPAFKGDLEAIERFAFEFVEDQSKEGVAYCEARLCPHLLLPDSPSTVDIINNVETPNTV

SHQTTTNNSTVDARKIVEAVLRGIERGEEKYSTKVRLILCCIRSK

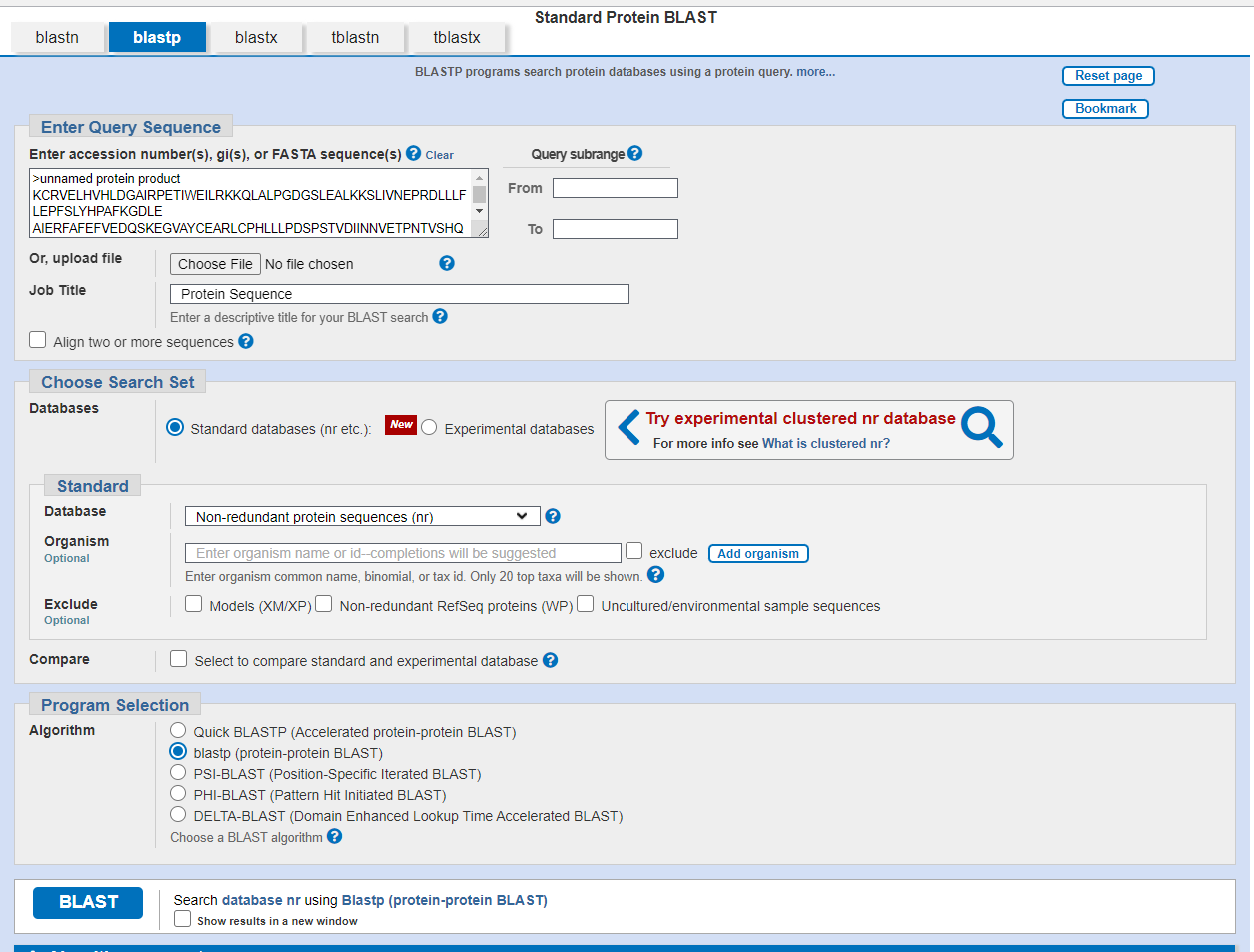
Name: Adenosine deaminase

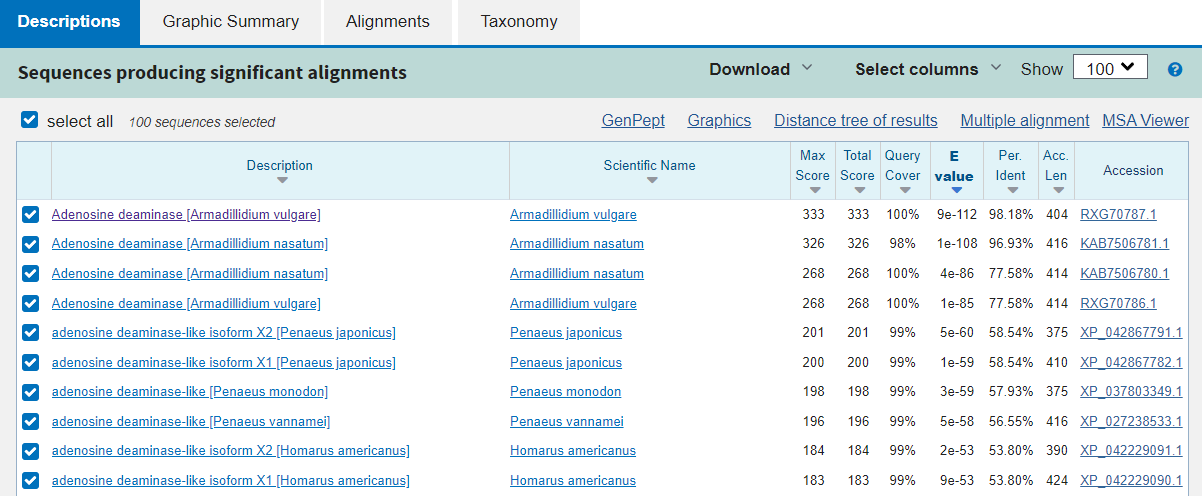
Species: Armadillidium vulgare

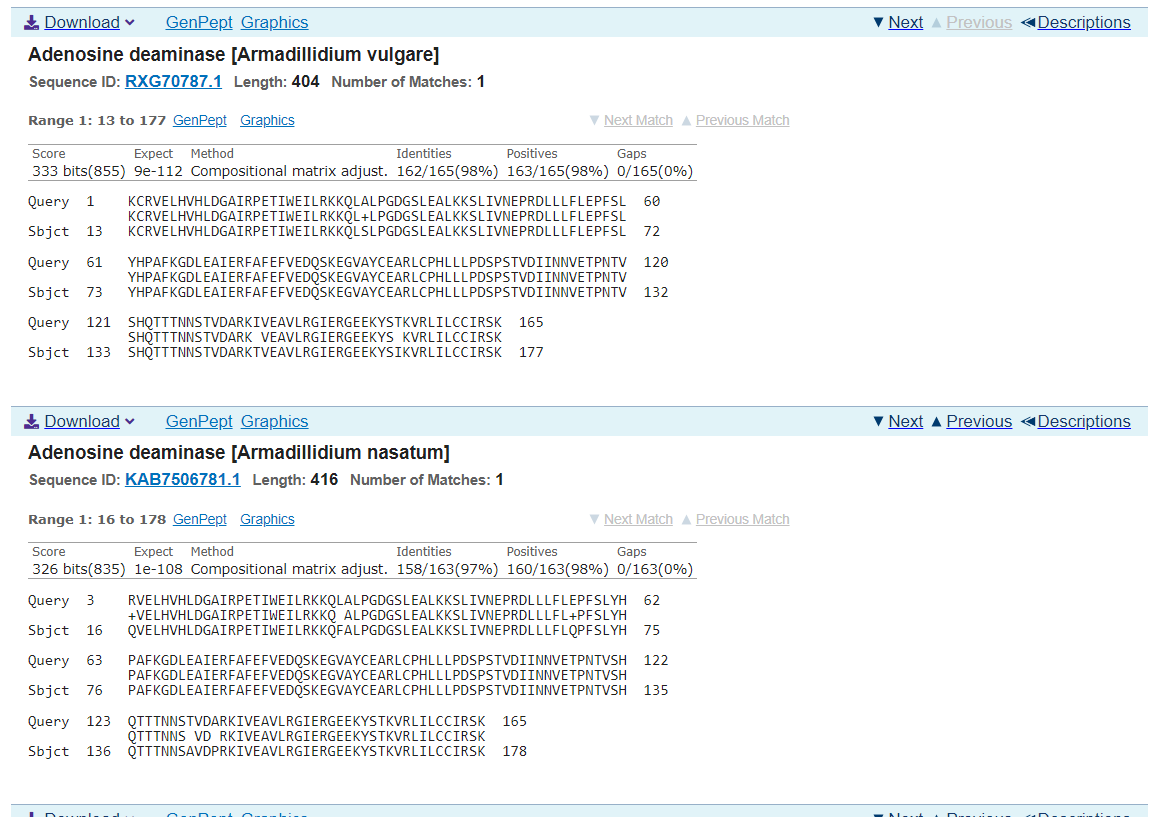
Eukaryota: Metazoa; Ecdysozoa; Arthropoda; Crustacea; Multicrustacea; Malacostraca; Eumalacostraca; Peracarida; Isopoda; Oniscidea; Crinocheta; Armadillidiidae; Armadillidium

[Q4] Details:

After A BLASTP search against the NR database of the protein we have produced from translating our results from the TBLASTN search of NP\_000013 on the Armadillidium organisms as shown on the figure below, we got a top results of only 98.18% Percent Identity.







This isn’t 100% identity, so results indicate that our match is a novel protein.

[Q5] Sequence for alignment:

>56-550\_1 Armadillidium vulgare normalized library Armadillidium vulgare cDNA, mRNA sequence

KCRVELHVHLDGAIRPETIWEILRKKQLALPGDGSLEALKKSLIVNEPRDLLLFLEPFSL

YHPAFKGDLEAIERFAFEFVEDQSKEGVAYCEARLCPHLLLPDSPSTVDIINNVETPNTV

SHQTTTNNSTVDARKIVEAVLRGIERGEEKYSTKVRLILCCIRSK

>RXG70787.1:13-177 Adenosine deaminase [Armadillidium vulgare]

KCRVELHVHLDGAIRPETIWEILRKKQLSLPGDGSLEALKKSLIVNEPRDLLLFLEPFSLYHPAFKGDLEAIERFAFEFV

EDQSKEGVAYCEARLCPHLLLPDSPSTVDIINNVETPNTVSHQTTTNNSTVDARKTVEAVLRGIERGEEKYSIKVRLILC

CIRSK

Re-labeled sequences for alignment:

>Caridean Shrimp | XP\_042867791.1:13-171 adenosine deaminase-like isoform X2 [Penaeus japonicus]

KCRVQLHIHLDGAIRHETIWELLRKKKLPLPGSGSLEDLKNALKVQEPQDLCFFLHGFTIFMPAFVGDLEAIERISHEFV

EDQAKESVAYCETRFCPHLFLPDSAHQPDYLTSEVNGTAEGNGGNVTIDDVLKAVLKGLKRGEEEFGTKVRVILCCIRG

>Giant tiger prawn | XP\_037803349.1:13-171 adenosine deaminase-like [Penaeus monodon]

KCRVQLHIHLDGAIRHETIWELLRKKKLPLPGSGSLEDLKNALKVQEPQDLGFFLHGFTIFMPAFAGDLEAIERISHEFV

EDQAKESVAYCETRFCPHLFLPDIAHQPDYLTSEVNGTAEGNSGSVTIDDVLKAVLKGLRRGEEEFGTKVRVILCCIRG

>American Lobster | XP\_042229091.1:13-183 adenosine deaminase-like isoform X2 [Homarus americanus]

KCRVQLHVHLDGCIRHETIWEVMRRKGLKLPGSGSLADLKLALQVQEPEDLLFFLSGFKIFLPAIKGDLAVIERIAHEFV

EDQAKEYVAYCEARFCPHLLLPNDTTQFSVVKCLKGKSRVNGTTDTTTEDGPVKDGEVTVDSILNAVLKGFSRGEEDFGT

KVRVILCCIHG

>Louisiana crawfish | XP\_045625303.1:13-179 adenosine deaminase-like isoform X3 [Procambarus clarkii]

KCRVQLHVHLDGAIRHETLWEVMRHKGMKLPGRGSLADLKTAVQVQEPKDLGLFLRGFQIFLPAIVGDLAVIERIAYEFV

EDQANDSVAYCEARFSPHLLLPSEQSQPNLHNEAEVQLNGTVVGSNGDSINNEVTVDSILIAVLKGLARGEEDFGTKVRV

IICCIRG

>>Gooseneck Barnacle | XP\_037080455.1:12-151 adenosine deaminase-like [Pollicipes pollicipes]

KCRVHLHVHLDGAVRPSTVWELSRQKKLLLPGNGSLKDLEEALQIHSPRDLAHFLQKFRWISPAIVGDLAAIERIAYEFV

EDEARQAVLYTEARFSPHLLVGPGGQVTARQVTEAVMRGMARGERQFGVTARLLLCCIRG

Alignment:

Obtained using MUSCLE (version 3.8) at EBI:

Gooseneck KCRVHLHVHLDGAVRPSTVWELSRQKKLLLPGNGSLKDLEEALQIHSPRDLAHFLQKFRW

Armadillidium KCRVELHVHLDGAIRPETIWEILRKKQLSLPGDGSLEALKKSLIVNEPRDLLLFLEPFSL

Caridean KCRVQLHIHLDGAIRHETIWELLRKKKLPLPGSGSLEDLKNALKVQEPQDLCFFLHGFTI

Giant KCRVQLHIHLDGAIRHETIWELLRKKKLPLPGSGSLEDLKNALKVQEPQDLGFFLHGFTI

Louisiana KCRVQLHVHLDGAIRHETLWEVMRHKGMKLPGRGSLADLKTAVQVQEPKDLGLFLRGFQI

American KCRVQLHVHLDGCIRHETIWEVMRRKGLKLPGSGSLADLKLALQVQEPEDLLFFLSGFKI

\*\*\*\* \*\*:\*\*\*\*.:\* .\*:\*\*: \*.\* : \*\*\* \*\*\* \*: :: ::.\* \*\* \*\* \*

Gooseneck ISPAIVGDLAAIERIAYEFVEDEARQAVLYTEARFSPHLLV-------------------

Armadillidium YHPAFKGDLEAIERFAFEFVEDQSKEGVAYCEARLCPHLLLPDSPSTVDIINNVETPNTV

Caridean FMPAFVGDLEAIERISHEFVEDQAKESVAYCETRFCPHLFLPDSAHQPDYLT-----SEV

Giant FMPAFAGDLEAIERISHEFVEDQAKESVAYCETRFCPHLFLPDIAHQPDYLT-----SEV

Louisiana FLPAIVGDLAVIERIAYEFVEDQANDSVAYCEARFSPHLLLPSEQSQPNLHN--EAEVQL

American FLPAIKGDLAVIERIAHEFVEDQAKEYVAYCEARFCPHLLLPNDTTQFSVVKCLKGKSRV

\*\*: \*\*\* .\*\*\*::.\*\*\*\*\*::.: \* \* \*:\*:.\*\*\*::

Gooseneck -----GP-------GGQVTARQVTEAVMRGMARGERQFGVTARLLLCCIRG-

Armadillidium SHQTTTN-------NSTVDARKTVEAVLRGIERGEEKYSIKVRLILCCIRSK

Caridean NGTAEGN-------GGNVTIDDVLKAVLKGLKRGEEEFGTKVRVILCCIRG-

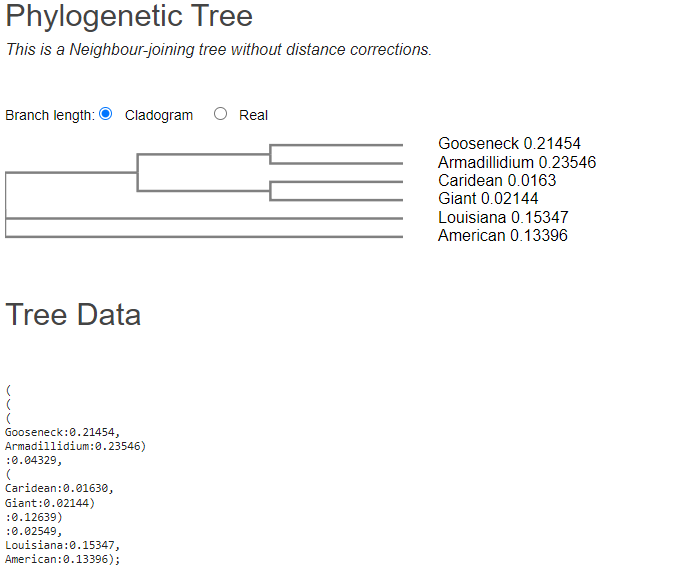
Giant NGTAEGN-------SGSVTIDDVLKAVLKGLRRGEEEFGTKVRVILCCIRG-

Louisiana NGTVVGS--NGDSINNEVTVDSILIAVLKGLARGEEDFGTKVRVIICCIRG-

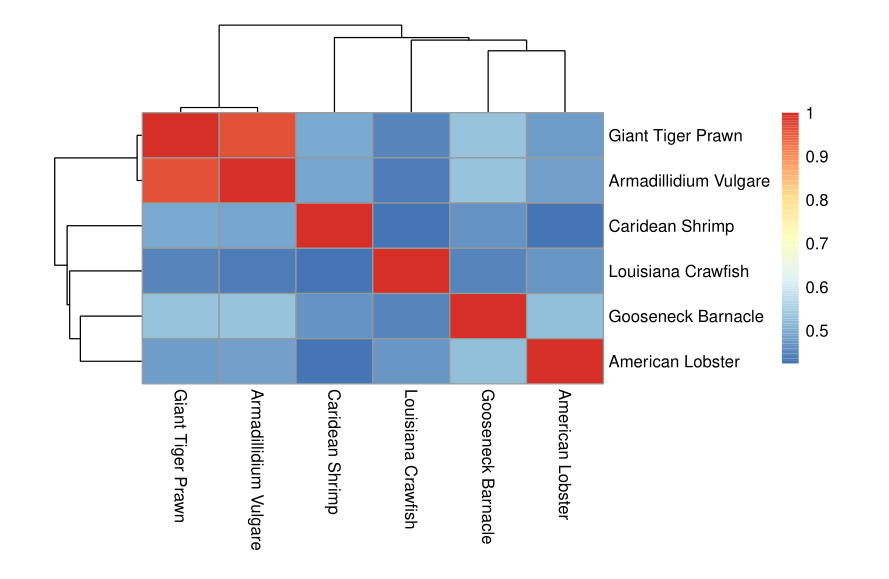
American NGTTDTTTEDGPVKDGEVTVDSILNAVLKGFSRGEEDFGTKVRVILCCIHG-

.. \* . \*\*:.\*: \*\*\* .:. ..\*:::\*\*\*..

[Q6] Phylogenetic Tree



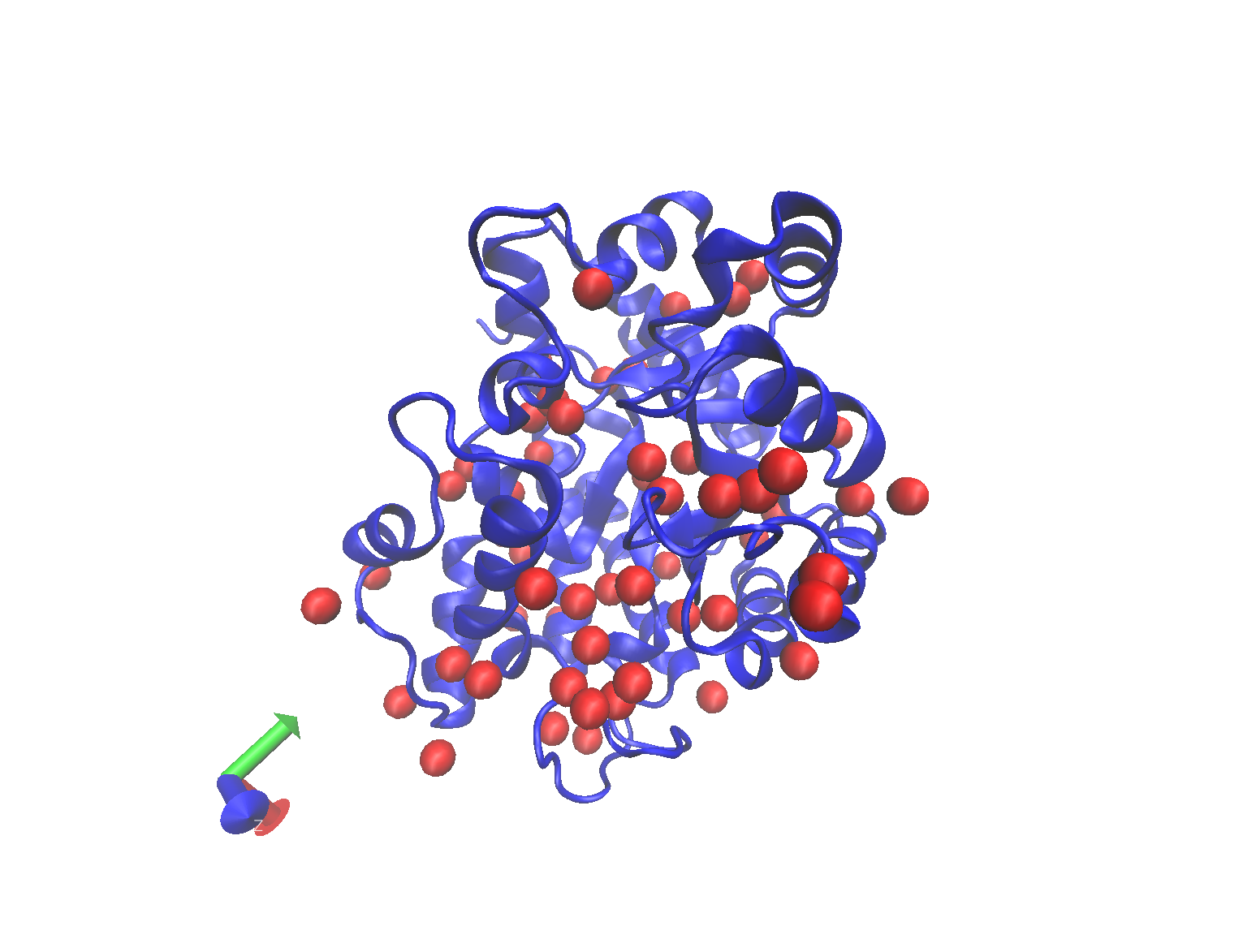
[Q7] Heatmap figure



[Q8]

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **ID** | **Technique** | **Resolution** | **Source** | **Evalue** | **Percent**  **Identity** |
| 1FKX | X-RAY  DIFFRACTION | 2.4 | Mus Musculus | 3E-28 | 42.33 |
| 1NDV | X-RAY  DIFFRACTION | 2.3 | Bos Taurus | 2E-24 | 41.07 |
| 6N91\_A | X-RAY  DIFFRACTION | 2.05 | Vibrio Cholerae | 2E-07 | 37.11 |

[Q9] Molecular Figure of One Identified PDB Structure Using VMD



Not likely to be similar because the structure of the Mus musculus protein is only 42.33% similar to our novel protein.

[Q10]

A target search of our novel protein details one single protein type named Serine/Threonine-protein kinase Sgk3 that belongs to Homo sapiens. Activity Charts in <https://www.ebi.ac.uk/chembl/target_report_card/CHEMBL6186/> states that the protein is mostly used for Inhibition, Residual Activity and Control. Associated Assays suggest that 100% of it is for B – Binding. Ligand Efficiency Plot for Target CHEMBL6186 is provided in a figure below

