

2021-22

BIO F242
Introduction to Bioinformatics

Experiment - 6

Name : Suchismita Tripathy
ID : 2019A7PS0554P

Gene Name : HUS1 Checkpoint Clamp Component (HUS1)
transcript variant X1
Organism : Theropithecus gelada
Accession Number : XM_025380797.1

28 March, 2022

1 Aim

To perform pairwise sequence alignment of Theropithecus gelada HUS1 Checkpoint Clamp Component (HUS1) transcript variant X1 gene and protein sequence using BLASTn and BLASTp to find out five closely related organisms (either Eukaryotes and Prokaryotes)

2 Materials Required

BLASTn, BLASTp, gene sequence of Theropithecus gelada HUS1 Checkpoint Clamp Component (HUS1) transcript variant X1

3 Procedure

1. Retrieve the genes and protein sequence in FASTA format from NCBI and ORF results respectively.
2. Open the web page of BLAST under NCBI.
3. Paste the appropriate sequence (Nucleotide for BLASTn and Protein for BLASTp) in a given space or upload your respective sequence for which you want to perform PSA.
4. Select appropriate parameters like word size, threshold value, output format, Gap open, and extension penalty in respective tools.
5. Take relevant screenshots of observations.
6. Analyze the results.

4 Observations

4.1 BLASTn

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
<input checked="" type="checkbox"/>	PREDICTED: Theropithecus gelada HUS1 checkpoint clamp component (HUS1), transcript variant X...	Theropithecus...	5598	5598	100%	0.0	100.00%	3031	XM_025380797.1
<input checked="" type="checkbox"/>	PREDICTED: Theropithecus gelada HUS1 checkpoint clamp component (HUS1), transcript variant X...	Theropithecus...	5321	5321	95%	0.0	100.00%	3024	XM_025380798.1
<input checked="" type="checkbox"/>	PREDICTED: Theropithecus gelada HUS1 checkpoint clamp component (HUS1), transcript variant X...	Theropithecus...	2519	4365	77%	0.0	100.00%	2423	XR_003118722.1
<input checked="" type="checkbox"/>	PREDICTED: Papio anubis polycystin 1 like 1, transient receptor potential channel interacting (PKD1L...	Papio anubis	1576	1660	31%	0.0	99.77%	10787	XM_031665407.1
<input checked="" type="checkbox"/>	PREDICTED: Cercopithecus atys HUS1 checkpoint homolog (S. pombe) (HUS1), transcript variant X3....	Cercopithecus atys	1291	1839	33%	0.0	99.58%	1393	XM_012065500.1
<input checked="" type="checkbox"/>	PREDICTED: Macaca mulatta HUS1 checkpoint clamp component (HUS1), transcript variant X3_mR...	Macaca mulatta	1772	1772	32%	0.0	99.19%	1482	XM_015133577.2
<input checked="" type="checkbox"/>	PREDICTED: Macaca nemestrina HUS1 checkpoint clamp component (HUS1), transcript variant X4....	Macaca nemes...	1772	1772	32%	0.0	99.19%	1462	XM_011759363.2
<input checked="" type="checkbox"/>	PREDICTED: Cercopithecus atys HUS1 checkpoint homolog (S. pombe) (HUS1), transcript variant X1....	Cercopithecus atys	1546	2094	38%	0.0	99.19%	1164	XM_012065498.1
<input checked="" type="checkbox"/>	PREDICTED: Macaca nemestrina HUS1 checkpoint clamp component (HUS1), transcript variant X2....	Macaca nemes...	1888	2443	44%	0.0	99.05%	1359	XR_994196.1
<input checked="" type="checkbox"/>	PREDICTED: Macaca mulatta HUS1 checkpoint clamp component (HUS1), transcript variant X2_mis...	Macaca mulatta	1823	2375	43%	0.0	99.02%	1324	XR_001443596.2
<input checked="" type="checkbox"/>	PREDICTED: Macaca nemestrina HUS1 checkpoint clamp component (HUS1), transcript variant X3....	Macaca nemes...	1823	2382	43%	0.0	99.02%	1324	XR_994195.1
<input checked="" type="checkbox"/>	PREDICTED: Macaca nemestrina HUS1 checkpoint clamp component (HUS1), transcript variant X5...	Macaca nemes...	1310	1485	27%	0.0	98.02%	948	YP_004107.1

Figure 1: Descriptions

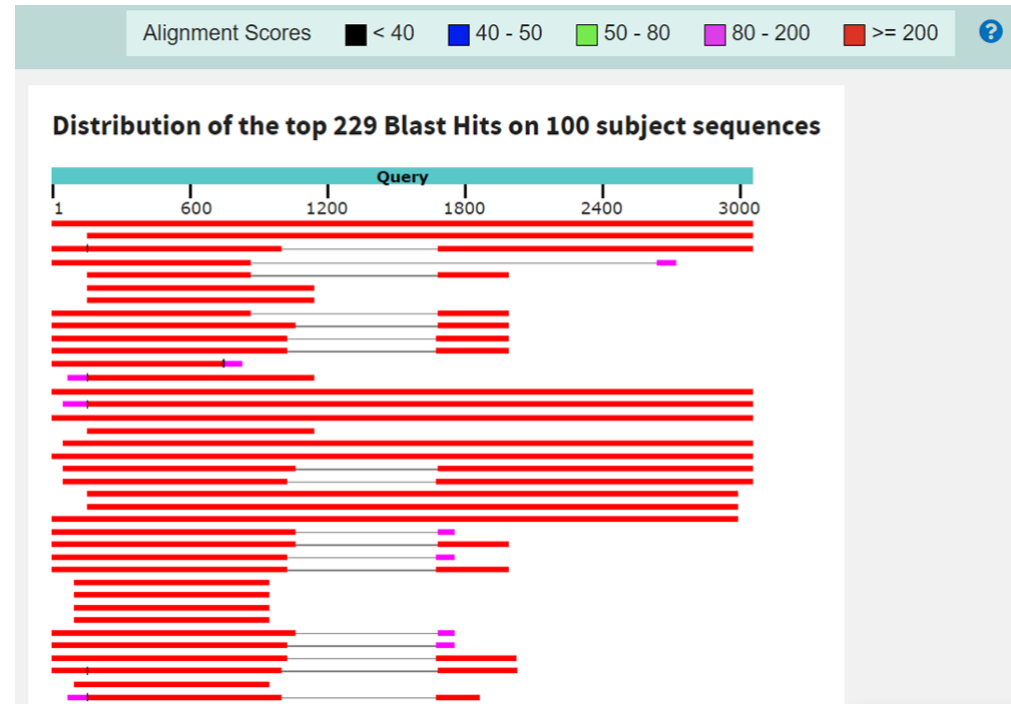


Figure 2: Graphic Summary

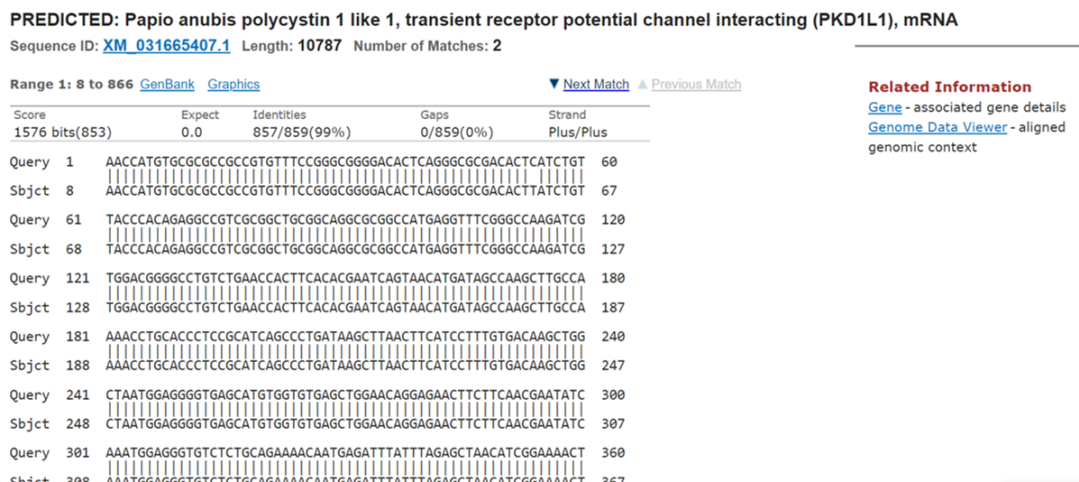


Figure 3: Alignment with sequence showing highest similarity

4.2 BLASTp

Sequences producing significant alignments			Download	Select columns	Show	100	?		
<input checked="" type="checkbox"/> 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"/>	PREDICTED: checkpoint protein HUS1 isoform X2 [Cercopithecus atys]	Cercopithecus atys	553	553	100%	0.0	100.00%	280	XP_011920889.1
<input checked="" type="checkbox"/>	checkpoint protein HUS1 isoform X2 [Theropithecus gelada]	Theropithecus gelada	508	508	92%	0.0	100.00%	259	XP_025236583.1
<input checked="" type="checkbox"/>	PREDICTED: checkpoint protein HUS1 isoform X1 [Cercopithecus atys]	Cercopithecus atys	499	499	90%	7e-177	100.00%	288	XP_011920888.1
<input checked="" type="checkbox"/>	checkpoint protein HUS1 isoform X1 [Chlorocebus sabaeus]	Chlorocebus sabaeus	551	551	100%	0.0	99.64%	280	XP_007979561.1
<input checked="" type="checkbox"/>	checkpoint protein HUS1 isoform X2 [Chlorocebus sabaeus]	Chlorocebus sabaeus	506	506	92%	4e-180	99.61%	259	XP_007979562.1
<input checked="" type="checkbox"/>	checkpoint protein HUS1 isoform X1 [Macaca nemestrina]	Macaca nemestrina	551	551	100%	0.0	99.29%	280	XP_011757664.1
<input checked="" type="checkbox"/>	checkpoint protein HUS1 isoform X1 [Macaca mulatta]	Macaca mulatta	550	550	100%	0.0	99.29%	280	XP_001088000.1
<input checked="" type="checkbox"/>	checkpoint protein HUS1 isoform X2 [Macaca nemestrina]	Macaca nemestrina	506	506	92%	2e-180	99.23%	259	XP_011757665.1
<input checked="" type="checkbox"/>	checkpoint protein HUS1 isoform X2 [Macaca mulatta]	Macaca mulatta	504	504	92%	2e-179	99.23%	259	XP_014989063.1
<input checked="" type="checkbox"/>	checkpoint protein HUS1 isoform X1 [Ptilocolobus tephrosceles]	Ptilocolobus tephrosceles	548	548	100%	0.0	98.93%	280	XP_023082175.1
<input checked="" type="checkbox"/>	checkpoint protein HUS1 isoform X2 [Pongo abelii]	Pongo abelii	503	503	92%	8e-179	98.84%	259	XP_024106000.1

Figure 4: Descriptions

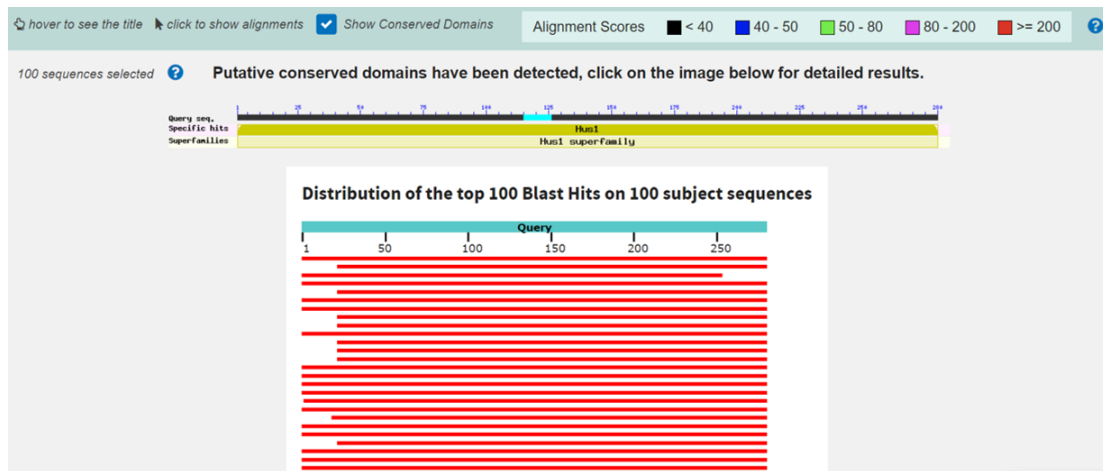


Figure 5: Graphic Summary

PREDICTED: checkpoint protein HUS1 isoform X2 [Cercopithecus atys]

Sequence ID: [XP_011920889.1](#) Length: 280 Number of Matches: 1

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

Range 1: 1 to 280 [GenPept](#) [Graphics](#)

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

Score	Expect	Method	Identities	Positives	Gaps
553 bits(1424)	0.0	Compositional matrix adjust.	280/280(100%)	280/280(100%)	0/280(0%)
Query 1	MRFRKIVDGA	CLNHFRISNMIAKLAKTCTLRISDPKLN	FILCDKLANGGVSMWCELEQ	60	
Sbjct 1	MRFRKIVDGA	CLNHFRISNMIAKLAKTCTLRISDPKLN	FILCDKLANGGVSMWCELEQ	60	
Query 61	ENFFNEYQMEG	SAENNEIYLELTSENLSRALKTAQNARALKIKLTNKHFPCLT	vsve11	120	
Sbjct 61	ENFFNEYQMEG	SAENNEIYLELTSENLSRALKTAQNARALKIKLTNKHFPCLT	VSVELL	120	
Query 121	SMSSSSRIVTH	DIPIKVIPRKLWKDLQEPVVPDPVSIYLPVLKTMKSVVEKMKNISNHL	180		
Sbjct 121	SMSSSSRIVTH	DIPIKVIPRKLWKDLQEPVVPDPVSIYLPVLKTMKSVVEKMKNISNHL	180		
Query 181	VIEANLDGELN	LKIEITELVCVTHFKDLGNPLASESTHQDRNMEHMAEVHIDIRKLLQF	240		
Sbjct 181	VIEANLDGELN	LKIEITELVCVTHFKDLGNPLASESTHQDRNMEHMAEVHIDIRKLLQF	240		
Query 241	LAGQQVNPTK	ALCNIVNNKMHFDLLHEDVSLQYFIPALS	280		
Sbjct 241	LAGQQVNPTK	ALCNIVNNKMHFDLLHEDVSLQYFIPALS	280		

Figure 6: Alignment with sequence showing highest similarity

5 Inferences

5.1 BLASTn

The Pairwise Sequence Alignment (PSA) using BLASTn of Theropithecus gelada HUS1 Checkpoint Clamp Component (HUS1) transcript variant X1

nucleotide sequence shows maximum alignment with **Papio anubis having E-value 0 and percentage similarity of 99.77%**, respectively. The top five organisms that are showing maximum similarity with Theropithecus gelada HUS1 Checkpoint Clamp Component (HUS1) transcript variant X1 are:

1. **Papio anubis polycystin 1 like 1, transient receptor potential channel interacting (PKD1L1), mRNA** having E-value 0.0 and percentage similarity 99.77% respectively
2. **Cercopithecus atys HUS1 checkpoint homolog (S. pombe) (HUS1), transcript variant X3, mRNA** having E-value 0.0 and percentage similarity 99.58% respectively
3. **Macaca mulatta HUS1 checkpoint clamp component (HUS1), transcript variant X3, mRNA** having E-value 0.0 and percentage similarity 99.19% respectively
4. **Macaca nemestrina HUS1 checkpoint clamp component (HUS1), transcript variant X4, mRNA** having E-value 0.0 and percentage similarity 99.19% respectively
5. **Cercopithecus atys HUS1 checkpoint homolog (S. pombe) (HUS1), transcript variant X1, mRNA** having E-value 0.0 and percentage similarity 99.19% respectively

5.2 BLASTp

The Pairwise Sequence Alignment (PSA) using BLASTp of Theropithecus gelada HUS1 Checkpoint Clamp Component (HUS1) transcript variant X1 amino acid sequence shows maximum alignment with checkpoint protein HUS1 isoform X2 [Cercopithecus atys] having E-value 0 and percentage similarity of 100%, respectively. The top five organisms that are showing maximum similarity with Theropithecus gelada HUS1 Checkpoint Clamp Component (HUS1) transcript variant X1 protein sequence are:

1. **checkpoint protein HUS1 isoform X2 [Cercopithecus atys]** having E-value 0.0 and percentage similarity 100% respectively
2. **checkpoint protein HUS1 isoform X1 [Cercopithecus atys]** having E-value 7e-177 and percentage similarity 100% respectively
3. **checkpoint protein HUS1 isoform X1 [Chlorocebus sabaeus]** having E-value 0.0 and percentage similarity 99.64% respectively

4. **checkpoint protein HUS1 isoform X2** [*Chlorocebus sabaeus*] having E-value 4e-180 and percentage similarity 99.61% respectively
5. **checkpoint protein HUS1 isoform X1** [*Macaca nemestrina*] having E-value 0.0 and percentage similarity 99.29% respectively