

2021-22

BIO F242

Introduction to Bioinformatics

Experiment - 4

Name : Suchismita Tripathy

ID : 2019A7PS0554P

Gene Name : HUS1 Checkpoint Clamp Component (HUS1)
transcript variant X1

Organism : Theropithecus gelada

Accession Number : XM_025380797.1

22 February, 2022

1 Aim

To predict the number of Open Reading Frames (ORF) present in your gene sequence along with maximum length of ORF and its protein sequence in all possible frames using ORF Finder.

2 Materials Required

ORF Finder, gene sequence of Theropithecus gelada HUS1 Checkpoint Clamp Component (HUS1) transcript variant X1

3 Procedure

1. Open the home page of NCBI and retrieve the Theropithecus gelada HUS1 Checkpoint Clamp Component (HUS1) transcript variant X1 gene sequence in Fasta format having “Accession No” (XM_025380797.1).
2. Open the home page of ORF Finder (<https://www.ncbi.nlm.nih.gov/orffinder/>)
3. Select the parameters such as like ORF length, genetic code, type of start codon and nested ORFs etc.
4. Take appropriate screenshots of observations.
5. Analyze the results.

4 Observations

Enter accession number, gi, or nucleotide sequence in FASTA format:

XM_025380797.1

From: To:

Choose Search Parameters

Minimal ORF length (nt):

Genetic code:

ORF start codon to use:

☒ "ATG" only

☐ "ATG" and alternative initiation codons

☐ Any sense codon

Figure 1: ORF Finder Input Screenshot

Open Reading Frame Viewer

PREDICTED: Theropithecus gelada HUS1 checkpoint clamp component (HUS1), transcript variant X1, mRNA

ORFs found: 2 Genetic code: 1 Start codon: 'ATG' only

XM_025380797.1 Find:

ORF2 (280 aa) Display ORF as... Mark

>1c1|ORF2
HRFRKIVDGLAHFTRISNIAKLAKTCTLRSPDKLHIFILCKLAWG
QVSNLELGEHFFHVEYQEGVSAGMIEYLEITSENLRAALKTKMABA
LKIKLTNKHFPCLTVSVLLSHSSSRIVTHDPIKVIPIKWLQDEPV
VPDPDVSYLPVLKTKNSVVERKHISNHLVIEANLDELNLKIELELC
VTTHFDLGGPPLASESTKHQRMH#AEVHIDIRKLQFLAQQQVNPTR
ALCNIVNNKMHFDLLHEDVSLQVFIPALS

Mark subset... Marked: 0 Download marked set as Protein FASTA

Label	Strand	Frame	Start	Stop	Length (nt aa)
ORF2	+	3	99	941	843 280
ORF1	+	2	2357	2710	354 117

Figure 2: ORF Finder Output Screenshot 1/3

TEDE: Theropithecus gelada HUS1 checkpoint clamp component (HUS1), transcript variant X1, mRNA

Found: 2 Genetic code: 1 Start codon: 'ATG' only

1 AACCATGTGCGCCGCGCTGTTTCGGGCGGGGACACTCAGGGCGGACACTCATCTGTTACCCACAGAGGCCGTCGCGGTCGCGGCAG

ORF2 91 GCGCGGCCATGAGGTTTCGGGCCAAGATCGTGGACGGGGCTGTCTGAACCACTTCACACGAATCAGTAACATGATAGCCAAGCTTGCCA

ORF2 181 AAACCTGCACCTCCGCATCAGCCCTGATAAGCTTAACCTTCATCTTTGTGACAAGCTGGCTAATGGAGGGGTGAGCATGTGTGTGAGC

ORF2 271 TGGACAGGAGAACTTCTTCAACGAATATCAATGGAGGGGTGTCTCTGCAGAAACAAATGAGATTATTAGAGCTAACATCGGAAACT

ORF2 361 TATCTCGAGCCTTGAAACTGCCAGAAATGCCAGAGCCTTGAAACTCAAACTGACTAATAAACACTTCCCTGCCTCAGAGTCTCCGTGG

ORF2 451 AGCTGTTATCTATGTCAAGCAGTAGCCGATTGTGACACATGACATCCCTATAAAGGTGATTCTAGGAAATTGTGGAGGACTTGCAAG

ORF2 541 AACCGTGGTCCAGATCCTGATGTTAGTATTATTACCACTTGAAGACTATGAAGAGTGTGTGGAAAAATGAAAAACATCAGCA

ORF2 631 ATCACCTGTTATTGAAGCAAACTAGATGGAGAAATGAATTTGAAATAGAAACTGAATTAGTATGTGTACAACTCATTTTAAAGATC

Figure 3: ORF Finder Output Screenshot 2/3

ORF2 (280 aa)

Display ORF as...

Mark

```
>lc1|ORF2
MRFRAKIVDGAACLNHFTRISNMIAKLAKTCTLRISPDKLNFI LCDKLANG
GVSMWCELEQENFFNEYQMEGVSAENNEIYLELTSENLSRALKTAQNARA
LKIKLTNKHFPCLTVSVELLSMSSSSSRIVTHDIPKVIKPRKLWKDLQEPV
VPDPDVSIIYLPVLKTMKSVVEKMKNISNHLVIEANLDGELNLKIETELVC
VTTHFKDLGNPPLASESTHQDRNMEHMAEVHIDIRKLLQFLAGQQVNPTK
ALCNIVNNKMMVHFDLLHEDVSLQYFIPALS
```

Figure 4: ORF Finder Output Screenshot 3/3

5 Inferences

The given gene sequence of Theropithecus gelada HUS1 Checkpoint Clamp Component (HUS1) transcript variant X1 consists of :

2 Number of ORFs

The ORF that has Maximum Length **ORF2**

The length of ORF that has Maximum Length **843**

The location of START Codon in Maximum Length of ORF **99**

The location of STOP Codon in Maximum Length of ORF **941**

The Direction of STRAND in Maximum Length of ORF **+**

The Length of Amino Acid encoded by BIGGEST ORF **280 amino acids**