2021-22

BIO F242
Introduction to Bioinformatics

Experiment - 7

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 align multiple gene and protein sequence by taking HUS1 Checkpoint Clamp Component (HUS1) gene and protein sequences using Clustal Omega and T-COFFEE

2 Materials Required

Clustal Omega, T-COFFEE, gene and protein sequences of HUS1 Checkpoint Clamp Component (HUS1)

3 Procedure

1. Retrieve the genes and protein sequence in FASTA format from NCBI and BLASTn and BLASTp results respectively.
2. Open the web page of Clustal omega and T-COFFEE under EBI.
3. Paste the appropriate sequence (Nucleotide or Protein) in a given space or upload your respective sequence for which you want to perform MSA.
4. Select appropriate parameters like output format, input sequence, guided tree etc in the tool.
5. Take relevant screenshots of observations.
6. Analyze the results.

4 Observations

4.1 Gene Sequence

4.1.1 Clustal Omega

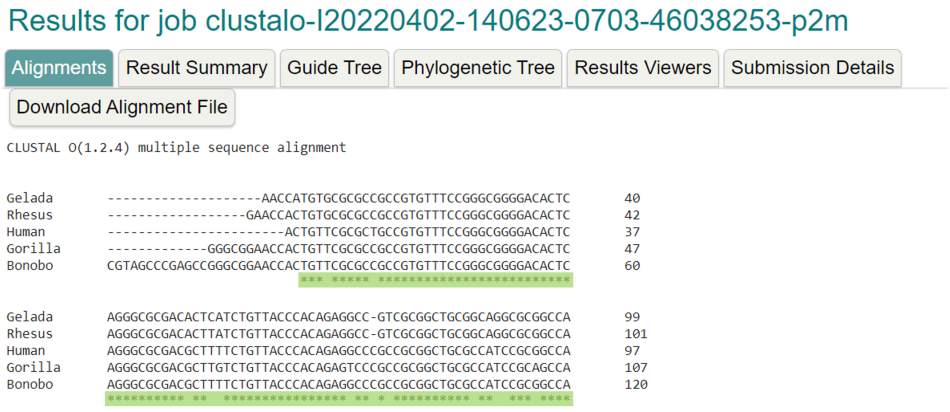

Figure 1: Alignments

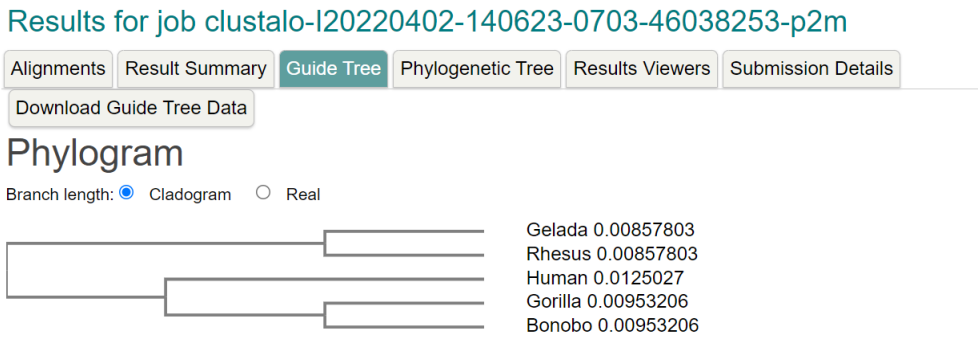

Figure 2: Phylogram : Cladogram

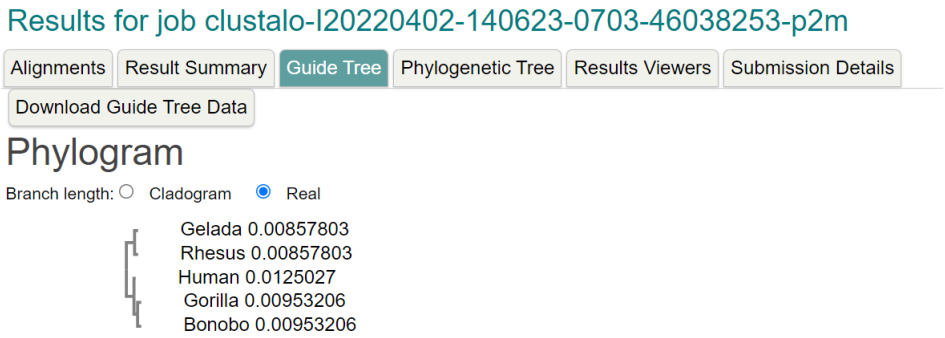

Figure 3: Phylogram : Real

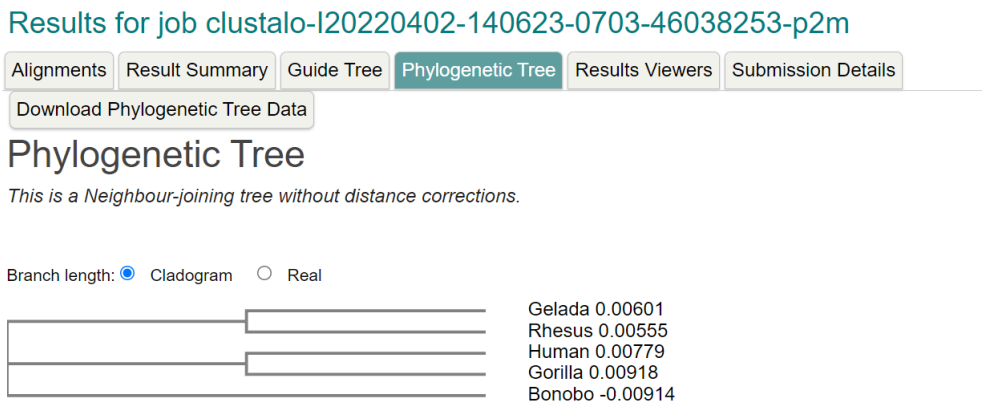

Figure 4: Phylogenetic Tree : Cladogram

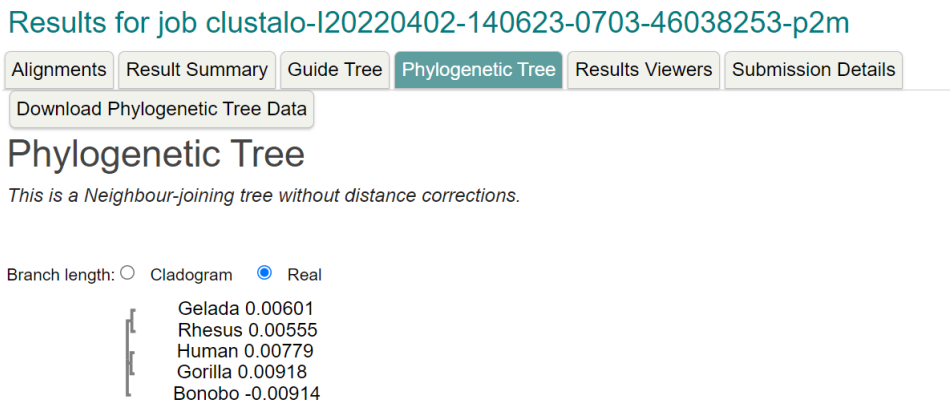

Figure 5: Phylogenetic Tree : Real

4.1.2 T-Coffee

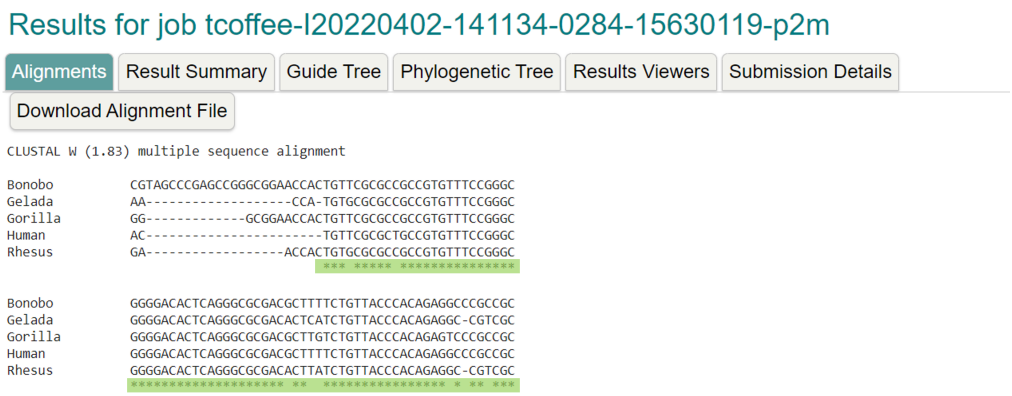

Figure 6: Alignments

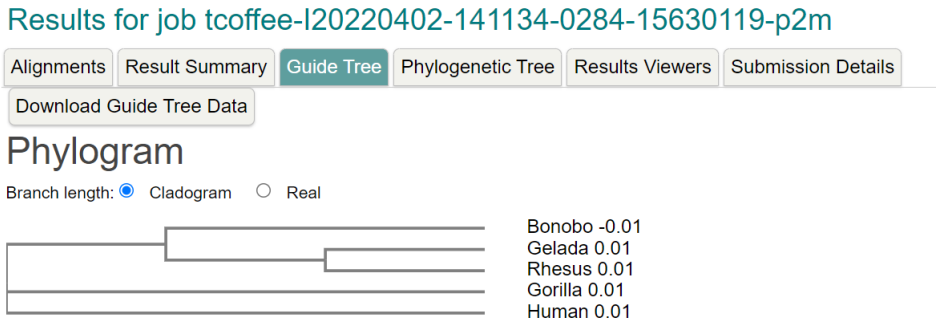

Figure 7: Phylogram : Cladogram

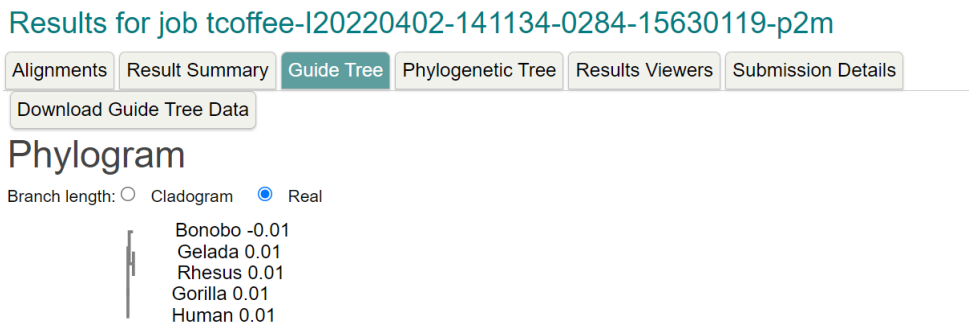

Figure 8: Phylogram : Real

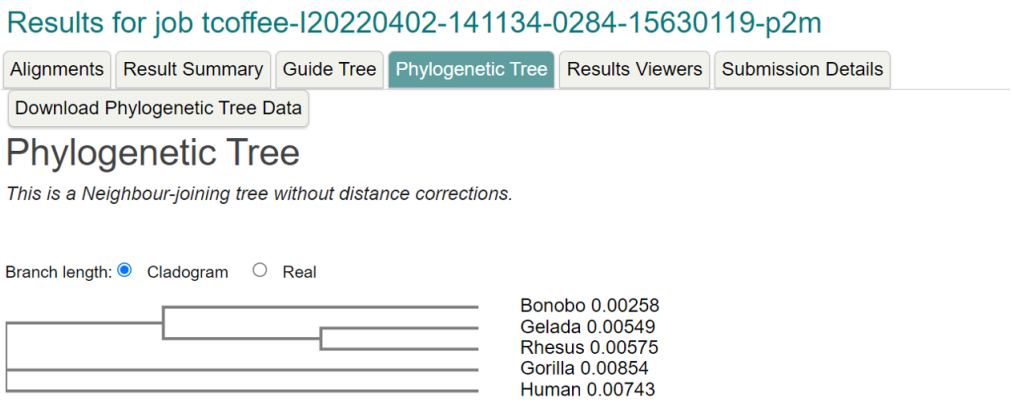

Figure 9: Phylogenetic Tree : Cladogram

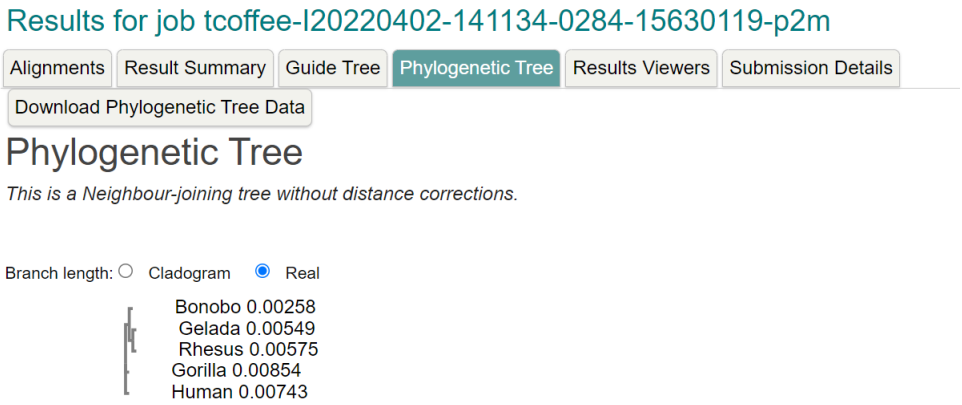

Figure 10: Phylogenetic Tree : Real

4.2 Protein Sequence

4.2.1 Clustal Omega

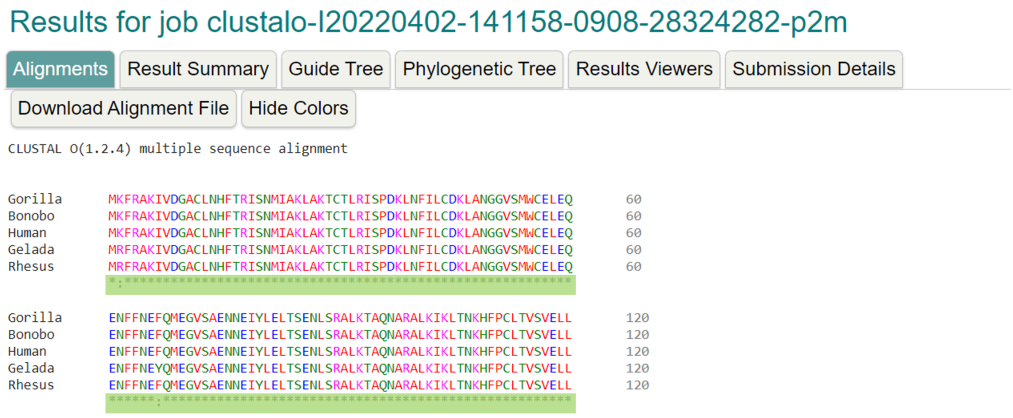

Figure 11: Alignments

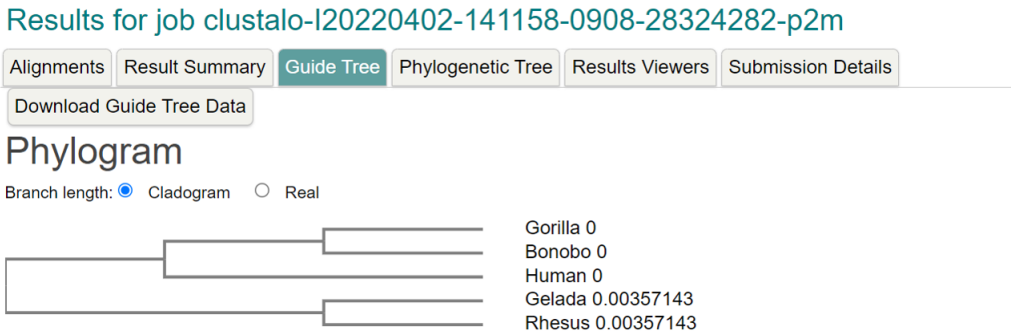

Figure 12: Phylogram : Cladogram

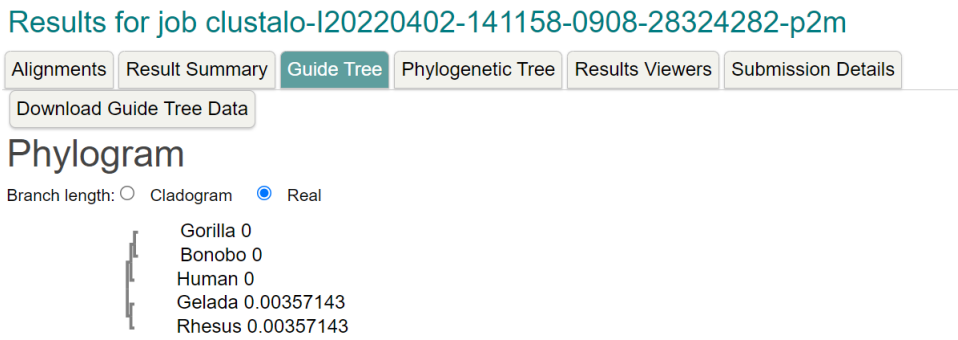

Figure 13: Phylogram : Real

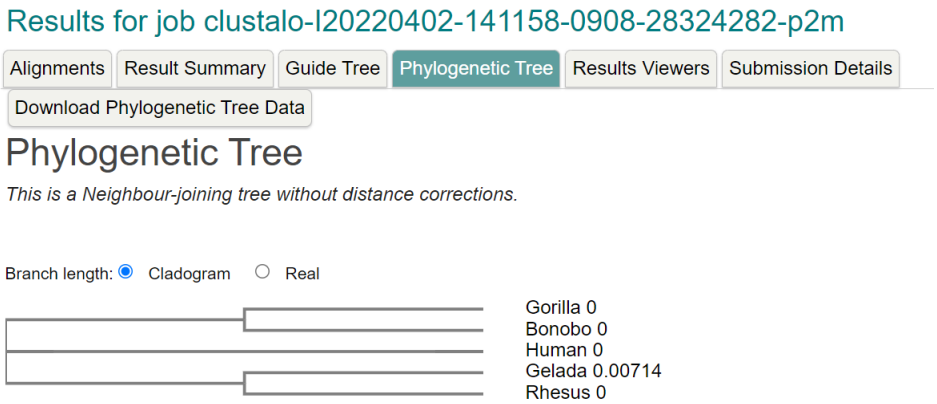

Figure 14: Phylogenetic Tree : Cladogram

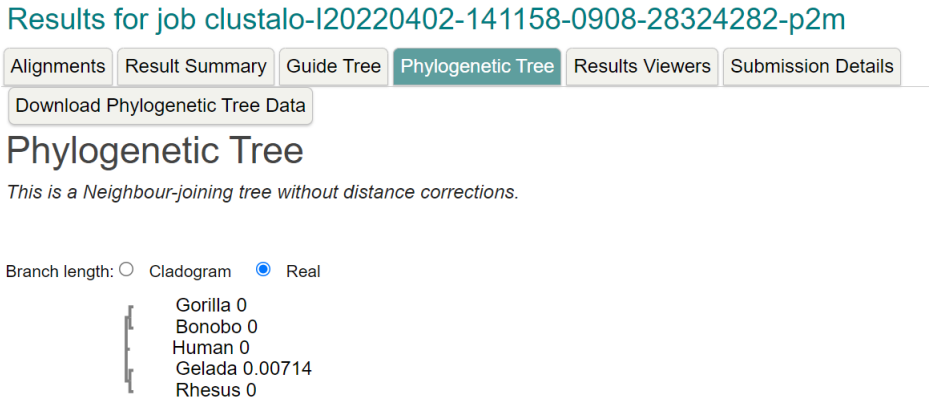

Figure 15: Phylogenetic Tree : Real

4.2.2 T-Coffee

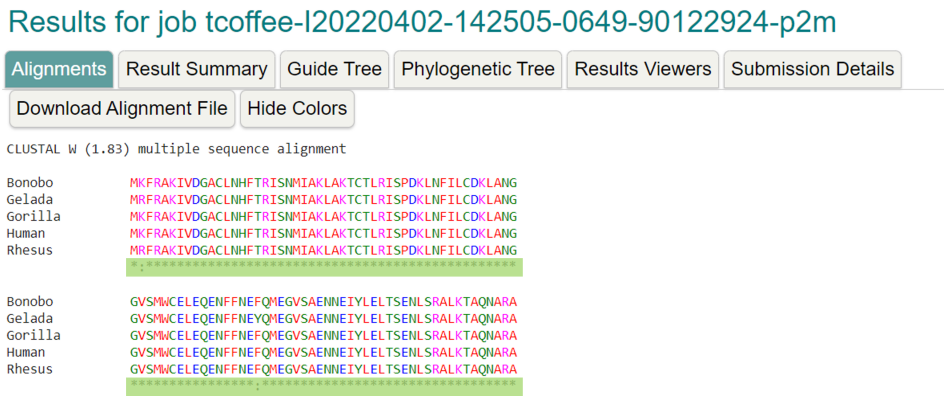

Figure 16: Alignments

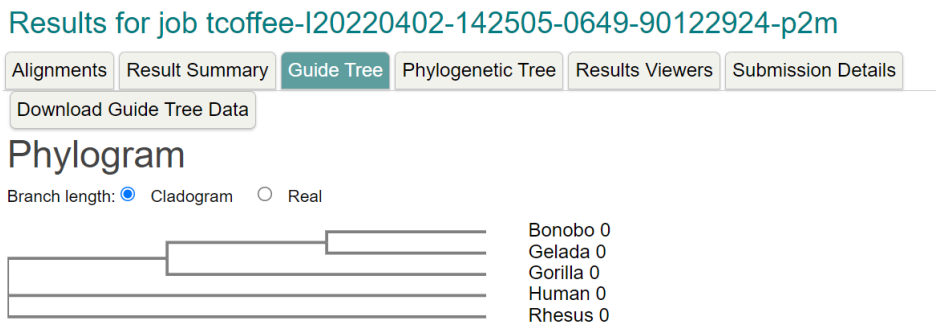

Figure 17: Phylogram : Cladogram

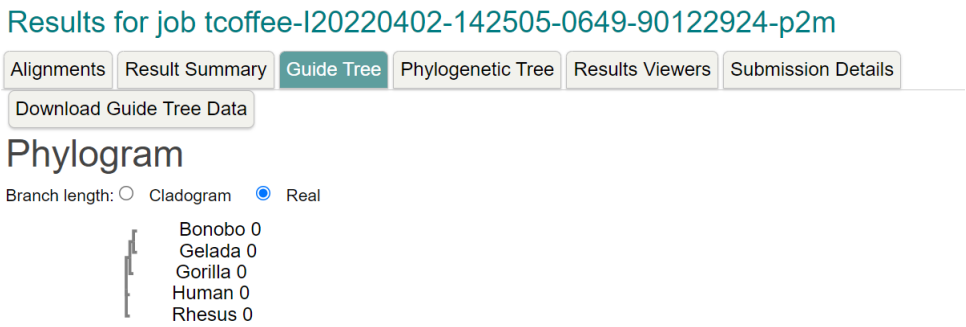

Figure 18: Phylogram : Real

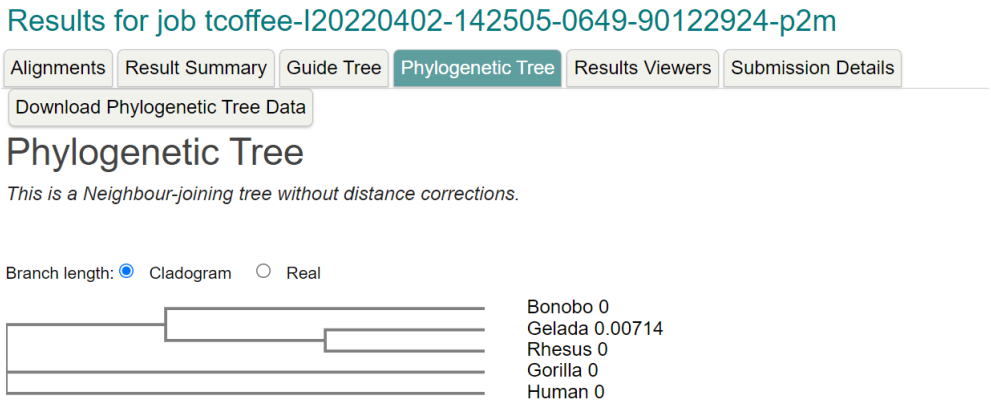

Figure 19: Phylogenetic Tree : Cladogram

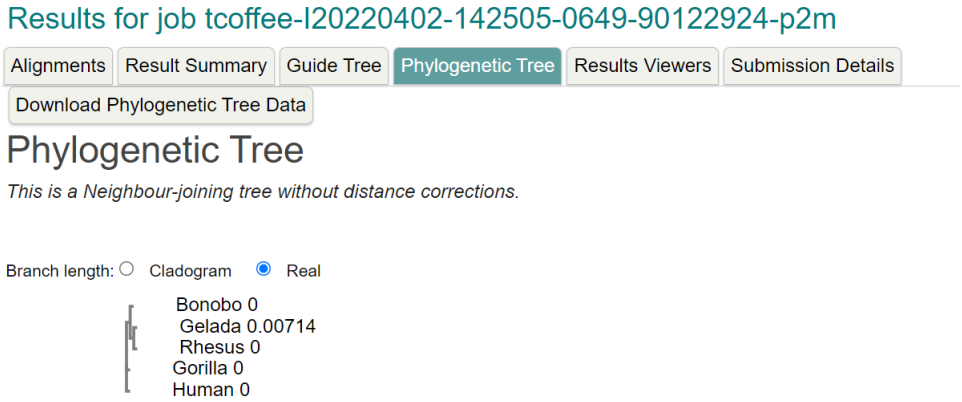

Figure 20: Phylogenetic Tree : Real

5 Inferences

1. The Multiple Sequence Alignment (MSA) using Clustal Omega of Theropithecus gelada HUS1 Checkpoint Clamp Component (HUS1) transcript variant X1 gene sequence with other 4 sequences shows presence of Conserved regions (*). The guided tree in the form of cladogram (using distance matrices) showed that Gelada and Rhesus are closely related with a minimum distance of 0.00857803. The Real tree shows that the number of mutational steps are highest in Gorilla and Bonobo having 0.00953206 value. The Phylogenetic tree (Neighbor Joining method) shows that Gelada and Rhesus are recent descenders while

Gorilla has been diverged during early phase of evolution having 0.00918 value.

2. The Multiple Sequence Alignment (MSA) using T-Coffee of Theropithecus gelada HUS1 Checkpoint Clamp Component (HUS1) transcript variant X1 gene sequence with other 4 sequences shows presence of Conserved regions (*). The guided tree in the form of cladogram (using distance matrices) showed that Gelada and Rhesus are closely related with a minimum distance of 0.01. The Real tree shows that the number of mutational steps are highest in Gorilla and Human having 0.01 value. The Phylogenetic tree (Neighbor Joining method) shows that Gelada and Rhesus are recent descenders while Gorilla has been diverged during early phase of evolution having 0.00854 value.

3. The Multiple Sequence Alignment (MSA) using Clustal Omega of Theropithecus gelada HUS1 Checkpoint Clamp Component (HUS1) transcript variant X1 amino acid sequence with other 4 sequences shows presence of Conserved (*), semi-conserved(:) and un-conserved regions(.). The guided tree in the form of cladogram (using distance matrices) showed that Gorilla and Bonobo are closely related with a minimum distance of 0. The Real tree shows that the number of mutational steps are highest in Gelada and Rhesus having 0.00357143 value. The Phylogenetic tree (Neighbor Joining method) shows that Gorilla and Bonobo are recent descenders while Gelada has been diverged during early phase of evolution having 0.00714 value.

4. The Multiple Sequence Alignment (MSA) using T-Coffee of Theropithecus gelada HUS1 Checkpoint Clamp Component (HUS1) transcript variant X1 amino acid sequence with other 4 sequences shows presence of Conserved (*), semi-conserved(:) and un-conserved regions(.). The guided tree in the form of cladogram (using distance matrices) showed that Bonobo and Gelada are closely related with a minimum distance of 0. The Real tree shows that the number of mutational steps are highest in Bonobo and Gelada having 0 value. The Phylogenetic tree (Neighbor Joining method) shows that Gelada and Rhesus are recent descenders while Human has been diverged during early phase of evolution having 0 value.