Practical Bioinformatics

Quiz 3 (TOTAL OF 15 POINTS)

Duration: 1 hour

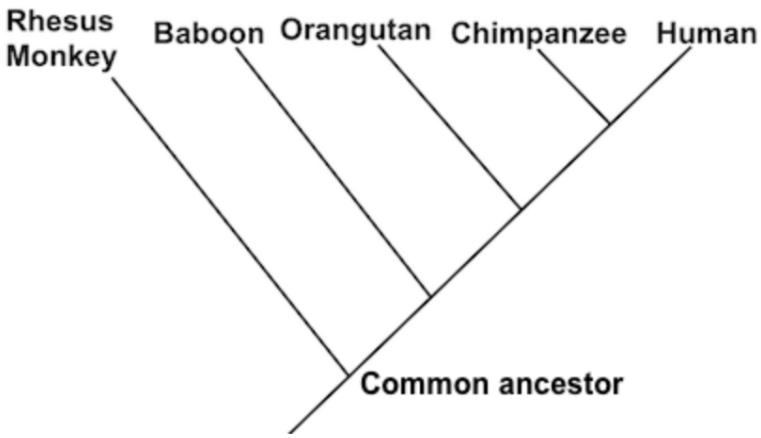
April 12 , 2021

Instructions:

- Please turn on the camera.
- Do your Quiz questions individually.
- ***Do not* zip your submissions.**
- If the solution requires you to use paper, paste a good quality image of the solution in the document that you are submitting.
- All the Queries, if any will be cleared by your respective TA.

---

1. Based on the figure, which two organisms are least closely related? [1]



2. Draw the phylogenetic tree for the following data that shows four amino acids found across three species [2]

Species	Amino Acid Sequence
Gorilla	Lys-Glu-His-Lys
Horse	Arg-Lys-His-Lys
Zebra	Arg-Lys-His-Lys

3. Choose the right set of mapping of the terms with their role. [1]
   A= Compares an amino acid query sequence against a protein sequence database
   B= Compares a nucleotide query sequence against a nucleotide sequence database
   a. A= BLASTN, B=BLASTP
   b. A=BLASTP, B=BLASTN
   c. A= BLASTX, B=BLASTN
   d. A=TBLASTN, B=BLASTX

4. What is “E value”? Usually E values smaller than a certain threshold are considered to demonstrate homology. This threshold is usually about ? If BLAST returns a match with an E-value of 2.4 e-11, what is the probability that this match represents a false positive? [1+1+1=3]

5. Differentiate between global and local alignment. Explain the various scoring matrices. What are two commonly used scoring matrices for data bank searches and for aligning protein sequences? [2+2+1=5]

6. (i) According to the molecular clock hypothesis:
   a. all proteins evolve at the same, constant rate;
   b. all proteins evolve at a rate that matches the fossil record;
   c. for every given protein, the rate of molecular evolution gradually slows down like a clock that runs down; or
   d. for every given protein, the rate of molecular evolution is approximately constant in all evolutionary lineages.

(ii) What are distance metrics used in the construction of phylogenetic trees? Name at least three of these and describe their key principle in one line each.

[1+2=3]