**BT302: Bioinformatics          Lab Assignment No. 2          17.08.2022**

**Note 1:** Submit the assignment online through Moodle either in .doc or .pdf format. Your final report file should be named as "**YourName_BT302_Lab2_17082022**". Make sure that your name and roll numbers are written at the first page of your final report. Note that you can upload only one file; thus, put together all the answers in a single file.

**Note 2:** There are two parts of this assignment. In this first part, students are expected to answer the questions 1-3. In the second part, students are expected to write a script to achieve the intended results. Each student should choose only one of the parts.


-----------------------------------------------------PART 1-----------------------------------------------

**Goal of this exercise is to learn about sequence alignments using various programs.**

**Use sequence number 1 and 2 from the sequence file to answer the below questions**

1. Can you figure out the name and organism of the proteins (*Hint: use the program BLAST*)?

2. Use the program Emboss-Needle to carry out **global alignment**.
   (a) Note down the sequence identity and similarity, gaps and scores. Why is the sequence similarity higher than identity? ***Note: default set of parameters.***

   (b) Now, change the scoring matrices (PAM10, PAM250, BLOSUM30 and BLOSUM90). Note down the differences in the results, if any.

   (c) Show that by increasing Gap Open Penalty, the total alignment score decreases? ***Note: PAM160 and BLOSUM62.***

   *(d)* Are the alignments same after changing Gap Open Penalty? ***Note: PAM160 and BLOSUM62.***

   (e) Show that by changing End Gap Open Penalty, the alignments as well as the total score remain almost similar. Why do you think it is reasonably correct? ***Note: PAM160 and BLOSUM62.***

3. Use the program Emboss-Water to carry out **local alignment**.
   (a) Note down the sequence identity and similarity, gaps and scores. ***Note: default set of parameters.***

   (b) Are the alignment statistics of two alignments (**global and local**) similar? If yes, what do you conclude? ***Note: default set of parameters.***

   (c) Show that by increasing Gap Open Penalty, the total alignment score decreases up to a particular value? ***Note: PAM160 and BLOSUM62***

(d) Show that keeping the <u>Gap Open Penalty</u> its maximum value; you get the best gap-less local alignment. ***Note: PAM160 and BLOSUM62***

---------------------------------------------**END of PART 1**-------------------------------------

---------------------------------------------**PART 2**-------------------------------------------------

**1.** Write a code to align two user-input protein sequences globally using user-defined scoring scheme.

**Note:** Copy paste or attach your written code in the report and create a README file on how to use the code.

---------------------------------------------**END of PART 2**-------------------------------------