Practical 5
21 Feb 2022

Questions

1. Analyze the occurrence of similar proteins in "nr" and SWISS-PROT database for the sequence given below:
   ```
   >1336093|Genbank|Outer membrane integral membrane protein|HrcC
   MVEKRELRCRLLGALLMLCATLPAGAQTPADWKEQSYAYSADRTPLSTVLQDFADGHSVD
   LHLGNVEDTEVTAKIRAENASAFLDRLALEHHFQWFVYNNTLYVSPQDEQSSERLEISPD
   AAPDIKQALSGIGLLDPRFGWGELPDDGVVLVTGPPQYLELVKRFSEQREKKEDRRKVMT
   FPLRYASVADRTIHYRDQTVVIPGVATMLNELMNGKRAAPASASGIDSTPGGPDTNSMMQ
   NTQTLLSRLSSRNKTSNRAGGRDNEIEDVSGRISADVRNNALLIRDDDKRHDEYSQLIAK
   IDVPQNLVEIDAVILDIDRTALNRLEANWQATLGGVTGGSSLMSGSGTLFVSDFKRFFAD
   IQALEGEGTASIVANPSVLTLENQPAVIDFSQTAYITATGERVADIQPVTAGTSLQVTPR
   AVGNEGHSSIQLMIDIEDGHVQTNGDGQATGVKRGTVSTQALISENRALVLGGFHVEESA
   DRDRRIPLLGDIPWLGQLFSSKRHEISQRQRLFILTPRLIGDQTDPTRYVTADNRQQLSD
   AMGRVERRHSSVNQHDVVENALRDLAEGQSPAGFQPQTSGTRLSEVCRSTPALLFESTRG
   QWYSSSTNGVQLSVGVVRNTSSKPLRFDEANCASKRTLAVAVWPHSALAPGESAEVYLAM
   DPSRVLHASRESLLNR
   ```

2. List the algorithm parameters used for the search (Q1).

3. What is the sequence identity of the query sequence (given in Q1) with AAK81929.1 ?

4. How far are hemoglobin sequences in human and chicken similar?

5. Write a program to list all the matching pentapeptides (which occur in both the sequences) and their frequency of occurrence in given sequences.

6. Write a program to compute sequence identity, similarity, query coverage and gap percentage from the alignment of human and chicken hemoglobin sequences (refer Q4).

7. Obtain the multiple sequence alignment for TIM barrel proteins from different organisms (select 20 proteins, for example). Compare the results obtained with Clustal Omega, MAFFT, and MUSCLE. List 5 residue positions which are aligned differently in these three methods.

8. Blast the below sequence 'EPDMRTPIAHTMAW' against PDB database. Analyze the results and discuss the significance of the results.

Deadline: 27 Feb 2022