

Practical 5

29 Feb 2024

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
MVEKRELRCRLLGALLMLCATLPAGAQT PADWKEQSYAYSADRTPLSTVLQDFADGHSVD
LHLGNVEDTEVTAKIRAENASAFDLRLALEHHFQWFVYNNNTLYVSPQDEQSSERLEISPD
AAPDIKQALSGIGLLDPRFGWGELPDDGVVLVTGPPQYLELVKRFSEQREKKEDRRKVM
T FPLRYASVADRTIHYRDQTVVIPGVATMLNELMNGKRAAPASASGIDSTPGGPD
TNSMMQ NTQTLLSRLSSRNKTSNRAGGRDNEIEDVSGRISADVRNNALLIRDDDKRHDEYSQ
LIAK IDVPQNLVEIDAVILDDIDRTALNRLEANWQATLGGVTGGSSLMMSGSGTLFVSD
FKRFFAD IQALEGEGTASIVANPSVLTLENQPAVIDFSQTAYITATGERVADIQPV
TAGTSLQVTPR AVGNEGHSSIQLMIDIEDGHVQTNGDGQATGVKRGTVSTQALISEN
RALVLGGFHVVEESA DRDRRIPLLGDIPWLGQLFSSKRHEISQRQLFILTPRLIGDQ
TDPTRYVTADNRQQQLSD AMGRVERRHSSVNQHHDVVENALRDLAEGQSPAGFQ
PQTSCTRLSEVCRSTPALLFESTRG QWYSSSTNGVQLSVGVVRNTSSKPLRFDEAN
CASKRTLAVAVWPHSALAPGESAEVYLAM 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 (beta) sequences in humans 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 the PDB database. Analyze the results and discuss the significance of the results.

Deadline: 13 March 2024