

Total 20 pts

1 (10pts). Use the dynamic programming algorithm to compute the optimal global alignments of the following two sequences:

$x = \text{TTCATA}$

$y = \text{TGCTCGTA}$

(using scoring function: $w(\text{match})=+5$, $w(\text{mismatch})=-2$, and $w(\text{indel})=-6$). Please give details rather than just the result.

2 (10pts). Title “Use of BLAST algorithm and interpretation of results”

Tools to be used

- protein blast, is a program which searches a protein database using a protein query and can be found at <https://blast.ncbi.nlm.nih.gov/Blast.cgi>, by clicking on the protein blast icon under the Web BLAST.
- NCBI protein database, is a collection of sequences from several sources, including GenBank, RefSeq, TPA, and PDB. This database can be found at <http://www.ncbi.nlm.nih.gov/sites/entrez?db=protein>.

Objectives

- Search the protein database and retrieve the sequence of particular proteins.
- Search the protein database using protein blast to find matching sequences.
- Examine the results and identify which of those sequences are highly related to the query sequence.

Procedure

1. Go to the NCBI protein database and search for Pannexin is a protein which acts as a channel allowing the transfer of ions and other small molecules to the cell. This particular protein (pannexin 1) is found in cattle.
2. Copy the sequence of this protein and paste it into the protein blast program. Restrict the search domain to humans. This can be done by typing human into the Organism cell in the program.
3. Observe the description of the results and identify which of these matching sequences are considered related to pannexin 1 based on the E-value. Also check the actual alignments of the corresponding matches. In what percentages are the related matches identical and related to the original sequence?
4. Perform the same procedure but search instead for the RNAase using the nucleotide database. RNAase is an RNA component which is used to break down the RNA into smaller components and can be found in humans. Before submitting the query, select Nucleotide in the search bar. You will need to use the nucleotide blast which can be also

found under the Web BLAST. This time restrict the search set to the Mouse genomic plus transcript.

For extra guidelines and information please follow this link for an online tutorial <http://www.youtube.com/watch?v=HXEpBnUbAMo> on how to use NCBI BLAST tools.