

Practical 1

24 Jan 2022

Questions

Instructions

1. Upload your answers (along with codes) in moodle with naming “roll no._p1”.
2. Do not copy codes, it will be checked for plagiarism

1. Download the EMBOSS package (<http://emboss.sourceforge.net/download/>) and copy to your Windows system. In case of **Linux** use the command: `sudo apt-get install jemboss`
2. Using EMBOSS, find the complementary strand for the sequence:
CTCGGATTTGTAAAGATCATGATCTCATACATAGTACCTAGCCATTG
3. Write a program to find complementary strand for the sequence given in Q2.
4. Using EMBOSS, find the protein sequence for
GACATTGTGAACAGTAAAAAGTCCATGCAATGCGCAAGGAGCAGAAGAGGAAGCAGGGC
AAGCAGCGCTCCATGGGCTCTCCCATGGACTACTCTCCTCTGCCCATCGACAAGCATGAG
CCTGAATTTGGTCCATGCAGAAGAAAAGTGGATGGG
(ii) Identify the reading frame equivalent to the following sequence.
PIQFSSAWTKFRLMLVDGQRRVVHGRAHGALLALLPLLLLAHCMDFFTVHNV
5. Write a program to translate the given DNA sequence (refer Q4) to protein sequence.
6. Write a code to search the DNA sequence provided in the Q4 for a specific string, where the string is user-specified string like ‘AAG’, ‘ACTA’, etc. The program should print the total number of matches for the given string and the positions of the matches. Run the code with different length of input string and observe the trend in no of matches
7. Familiarize with other applications in EMBOSS. For example, melting temperature, bending, curvature etc.
8. Write a program to compute the average base stacking energy for the sequence given in Q2 (AA: -4; AT: -7; AC: -5; AG: -11; TA: -7; TT: -2; TC: -3; TG: -4; CA: -9; CT: -5; CC: -6; CG: -7; GA: -9; GT: -6; GC: -4; GG: 11).
9. Compute the average melting temperature of the following sequences using Seq2Feature tool (<https://www.iitm.ac.in/bioinfo/SBFE/index.html>) and comment on the results (Enter one sequence at a time in fasta format)
(i) ATATATATAT ii) GCGCGCGCGC
10. Calculate the AT and GC content of the sequence AAATGGCCCTAA using Seq2Feature tool

Deadline: 30-01-2022