BIL458: Bioinformatic (Bioinformatik)
Homework 1

Spring 2020   assigned March 05, 2020

Arzu Kakisim, Instructor   due March 26, 2020 06.00pm

Homework policy

- Write your name and student id clearly at the top of your assignment.

- You have to send your code as a zip file. Just write your school number in the zip file and folder.

- You have to add a "read me" file that gives some details about your code.

- You have to prepare a report that gives test results of your code.

- You are to complete this assignment individually (not explicitly tell each other the answers, not to copy code, not to allow your code to be copied).

- Plagiarism is prohibited.

- You have to code the algorithm using Python language.

- No credit will be given for assignments after the deadline.

Problems (Homework grade will be calculated over 130.)

1. (30) Implement the Needleman-Wunsch algorithm to find global alignments. Use the data set and score matrices provided by me. Test your algorithm by aligning first sequence with the other sequences, and put your pairwise alignment scores and results to your report. (Dont forget that you must do backtracking.)

2. (20) Implement Smith-Waterman Algorithm to find local alignments. Follow the same instructions as above.

3. (30) Implement BLAST Algorithm to find local alignments. Test your algorithm using first three sequences when window size=4, and put your pairwise alignment scores and results to your report. (The threshold is determined by you.)

4. (30) Implement Star Aligment Algorithm to find multi alignment. Test your algorithm using first four sequences by selecting the first sequence as a center sequence. Put your alignment result and scores (consensus, entropy and pairwise) to your report.

5. (20) Implement CLUSTALW Algorithm (UPGMA) to create a tree for multiple alignment. Use hamming distance score to construct a distance matrix (H(X,Y)=the number of mismatch (not use indel)). Draw your tree as a figure and put the figure to your report.