

03-60-558

Computational Molecular Biology

Assignment #2

Deadline: November 19, 2003, 09:50am

Dr. Alioune Ngom

November 10, 2005

1 Multiple and Profile Alignments

1. Write a program (in Perl, Bioperl, Java or BioJava or a language of your choice) to multiply-align the following four sequences AGATAGAAACTGATATATA, ATACTACGGAGGG, GAACGTAGGCGTAT and AGAAAAAGAGT. You can use any multiple alignment algorithm discussed in class or not.
2. Generate a profile of the multiple alignment
3. Using the profile, write a program to align AGACTAAGGTACGCGAGT against the profile
4. Multiply-align the four sequences using any alignment software from the web, and compare with your results from above. (You may generate your profile from the BioEdit's result)

2 (

Web Search and Testing Alignments) For this question, go to <http://www.bioinformaticsonline.org> and then find the list of problems for chapter 5 (of our textbook), and solve problem 2. You will need to do your own web search to find the SwissProt, BALiBASE, CLUSTALW, POA and DIALIGN web sites.

3 Multiple-Alignment using Hidden Markov Model

Go to <http://www.bioinformaticsonline.org> and solve Problem 6 of chapter 5. You will need to read <http://www.bioinformaticsonline.org/ch/ch05/simg.html> (or go to Chapter 5's page and select sample image) in order to understand how to use HMM for multiple-alignment.

4 Bonus Question

Using the simple model of HMM discussed in class and Needleman-Wunsch algorithm (last Tuesday) for pairwise alignment, obtain a multiple-alignment of the 4 sequences of section. Compare your results with your results in section 1 and give your comments.