

Home Work 1

Introduction to Bioinformatics Algorithms, Winter 2012

Given on Wednesday, January 11, 2010. Due on Wednesday, January 18 (E.mail to TA).

State your answers legibly and concisely. Your solutions will be graded on correctness, elegance, **clarity** and originality. Your proofs should **avoid getting bogged down in too much detail**. Your algorithms should be as efficient as possible (inefficient algorithms will result in lower grades). Please note that the work handed in must be your own. No late assignments will be accepted without prior arrangement.

Problem Write a program for computing the 2-break distance between genomes (consisting of circular chromosomes). Use the program to answer the following questions (each question is 25 points):

- Given circular permutations $+1, -7, +3, +2, -6, +4, -5$ and $+2, +4, +7, +1, +3, -6, -5$ find the minimum number of 2-breaks required to transform one permutation into another.
- What is the expected 2-break distance $distance(n)$ between two *random* genomes on n genes? Use the program you wrote to construct the plot of $distance(n)$ for $10 \leq n \leq 100$ using simulations. You will need to rigorously define the notion of a random genome and write a program for generating random genomes.
- Find the *reversal distance* between the linear versions of these permutations (feel free to use <http://grimm.bioprospects.org/GRIMM>). Find minimal number of breakpoint re-uses in the transformations of $+1, -7, +3, +2, -6, +4, -5$ into $+2, +4, +7, +1, +3, -6, -5$. Please note that this problem asks you to find the reversal distance rather than the 2-break distance (you don't need the program you wrote to answer this question).
- Rigorously define the notion of breakpoint between signed circular genomes π and σ and prove that the number of breakpoints between π and σ equals to the number of breakpoints between σ and π .