

Biological Data Analysis (CSE 182) : Assignment 2

Logistics

Submit your assignment electronically to Ali Bashir. The assignment is optional for students who wish to boost their grades as the best two assignment scores will be chosen. Due on Tuesday, December 7 (at the time of the final exam).

Perfect Phylogeny

1. Design and implement an algorithm *perf* which takes as input a binary character matrix, and constructs a perfect phylogeny if one exists. The output of the program should be a tree (described as a set of parent child relationships). Test your algorithm on the two data-sets provided. Submit a hand-drawn phylogeny on a piece of paper, with the leaves labeled with individuals, and edges labeled with columns.
2. Design an algorithm that takes a collection of SNPs from diseased (D) and normal (N) individuals, and identifies the region(s) likely to contain the mutant SNP, by plotting the linkage disequilibrium. Apply your algorithm on the supplied data.
3. (2pts.) What language did you use? How much time did you take to do the assignment? Who did you discuss your homework with?