

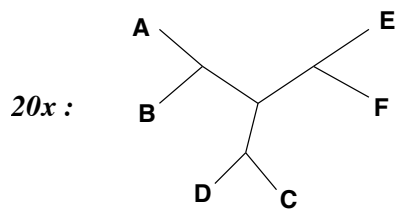
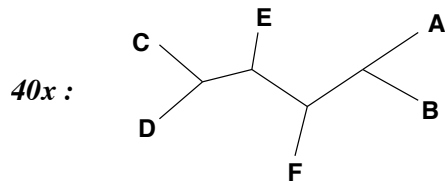
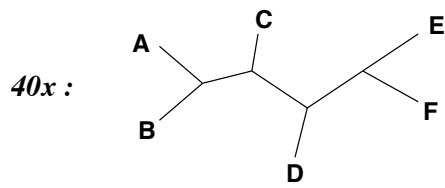
Bioinformatik für Biologen SS 2019

Complementary Homework 10

This homework will not be collected or graded, but its contents and the papers to read may be relevant in an upcoming lecture, test or assignment.

1. Describe briefly the three approaches for phylogeny reconstruction you learned? What do they take as input? What criteria/models are used to evaluate the trees?
2. Assuming unrooted, bifurcating (i.e. all internal nodes have degree 3) trees. Compute for rising numbers of leafs $n > 2$, how many of the following entities exist: internal nodes, nodes total, external branches, internal branches, branches total and different possible trees. Continue the computation for rising n until the number of trees passes 500 million for the first time. Use the approach and formulas we discussed in class.
3. Describe briefly how the concept of neighbor-joining works. Why is it not enough just to look for the two sequence (or nodes) with the smallest pairwise distance? Show an example.
4. Describe briefly with your own words the difference between probability and likelihood. What is fixed, what is variable, what do you compute with the one what with the other?
5. Describe briefly with your own words how maximum likelihood distances are constructed.
6. Describe how you optimize the branch lengths of a given tree in the maximum likelihood approach. Why is the Pulley Principle helpful for this?

7. You have computed 100 bootstrap trees:



- (a) Construct the strict consensus tree and the M_{50} majority rule consensus tree.
 - (b) Explain shortly, why the splits chosen for the M_{50} majority rule consensus tree have to occur in more than half the bootstrap trees.
 - (c) Use the splits extracted from the two "40x" trees. Which pairs of these are compatible and which are not? Use the intersection method you learned in the lecture.
8. Explain the three tree rearrangement operations presented in the lecture (NNI, SPR, TBR) to produce a neighborhood around a given tree T . How can these neighborhood trees help to evade local optima (i.e. wrong 'best' trees).
9. Use the "20x" tree from the bootstrap task above. Generate the full NNI and the full SPR neighborhood of that tree. How many trees are constructed per neighborhood by NNI and SPR respectively?