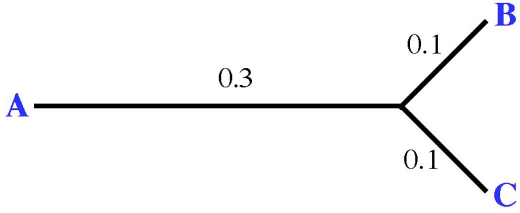


Assignment 5 - Fun with Maximum Likelihood

Part I. Likelihood methods using paper and pencil ☺.

- 1) Complete exercise 4(a and b) on page 99 of the textbook.
- 2) Calculate the likelihood for the following tree under the JC69 model

Taxon:	Sequences:	Topology:
A	AG	
B	AC	
C	AG	

Remember that to calculate the likelihood of a site, you have to consider all possible reconstructions for internal nodes and that the edge lengths are given in terms of expected number of substitutions v , where $v = 3\alpha t$ (for the JC model)

Part II. Likelihood methods using RAxML

- 3) Use RAxML to find the ML tree based on the dataset cob_nt.fasta* we used for the previous assignments. What RAxML command options did you use for your analysis? Why?
- 4) Use RAxML to find the ML tree based on the dataset cob_aa.fasta* we used for the previous assignments. What RAxML command options did you use for the analysis? Why?

Print these trees (for an extra point, print these trees “face to face” aka “tanglegram” using Dendroscope). Are they the same? If not, which tree do you think is closer to the true tree? Explain!

- 5) Create PBS job scripts to do bootstrap analyses (1000 replicates) on the HPC-class cluster for the same datasets you used in 3 and 4. Request 32 processor cores, maximum amount of RAM (you can ask for up to 62 GB per node), and 4 hr. of time. How much actual time did your analyses take? Which of the trees had higher bootstrap support? Print the bootstrap trees with the bootstrap support values and submit them with the rest of the assignment.

*Remember to align the sequences first