

## Key for Web Exploration Questions, Chapter 7

---

1. In what important way is a tree computed using the UPGMA algorithm different from a tree computed by the neighbor-joining algorithm? Which do you think better models evolution, and why?

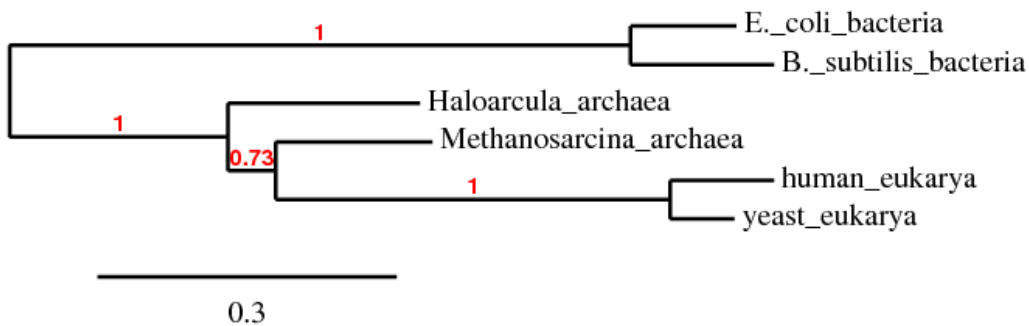
UPGMA assumes that when two clusters are grouped, showing a relationship between two species or two groups of species, that the lengths of the branches between each species (or group) and the common ancestor will be the same. That is, UPGA assumes a constant rate of evolution: since the two species have been evolving for the same amount of time since they split, they should have the same number of mutations.

Neighbor-joining assumes that the rate of evolution will not necessarily be the same along both branches from a common ancestor and allow for branch lengths to be calculated properly even if they're different. It is generally assumed that the basic rate of mutation should be very similar for all species, but there are certainly conditions that could increase the rate of mutation for one species over another. Or, perhaps even more likely, strong selective pressure on one species not felt by the other could result in more mutations becoming “fixed” in the genome in a shorter time. So, the assumptions of neighbor-joining are a better match for real-world biology even though UPGMA can work in many situations.

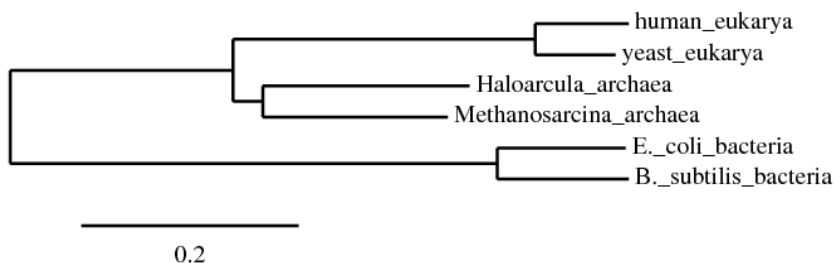
2. Summarize concisely what you learned about the relationships among the three domains from your trees. Were the trees you developed by different methods consistent in terms of branching orders and evolutionary pathways? How consistent were they in terms of branch length?

Below are some trees drawn by phylogeny.fr using different methods of tree-building:

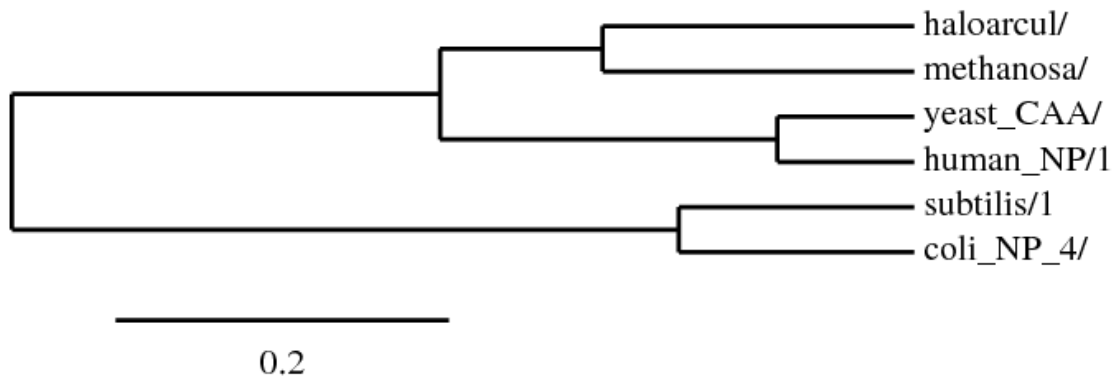
PhyML (default):



BioNJ



UPGMA:

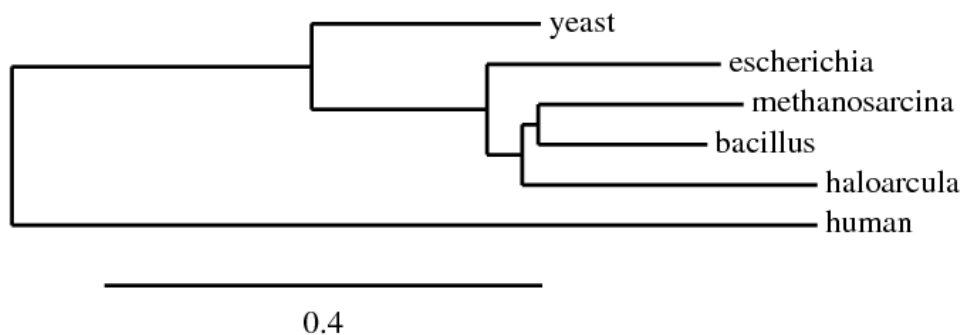


Bayesian:

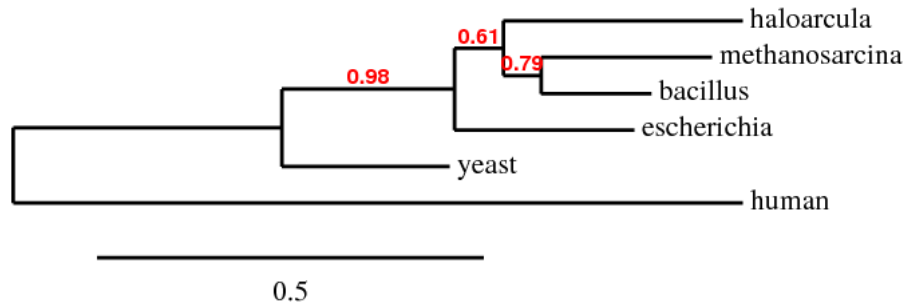
All the trees give the same grouping, with the two bacteria clustered together, the two archaea clustered together and the two eukaryotes clustered together. Since we already know these relationships based on many other data, this grouping serves to validate the tree. Consistently, all the trees suggest a closer relationship between the archaea and the eukarya than between either and the bacteria. As expected, UPGMA assumes equal branch lengths from any node to the leaves that extend from it, while the others allow for differences in rates of evolution. The branch lengths to the major groups are generally consistent among the trees, though small differences can be observed, particularly when looking at two terminal nodes descending from the same internal node.

3. It would make sense that if one highly conserved protein works as a “molecular clock,” then any other similarly conserved protein would give the same results. To test that assumption, generate a phylogeny with a different highly conserved protein, the heat-shock protein Hsp70 (also known as DnaK in bacteria). Download the amino-acid sequence of the Hsp70 protein for the same six organisms (NP\_002145, AET14830, DNAK\_ECOLI, DNAK\_BACSU, YP\_306886, DNAK\_HALMA), align the sequences, examine and curate the alignment and produce trees using neighbor-joining and maximum likelihood methods. Summarize the results of this analysis and discuss anomalies between the two molecular clocks. What did you learn about the reliability of evolutionary hypotheses based on molecular data from this exercise?

BioNJ:



PhyML:



These trees look very different from the ones drawn using EF-1 $\alpha$ /EF-Tu as the molecular clock. Although the groupings are a little less clear, the pairs of species from each domain are still more similar to each other than to the other domains. However, now the bacteria and the archaea group together, leaving the eukarya noticeably farther away. Furthermore, the human gene is much farther from the yeast gene than was the case for EF-1 $\alpha$ . This suggests a need for caution in identifying “clock” sequences and the need to confirm hypotheses using many different clocks as well as other kinds of data.