BINF 730 Biological Sequence Analysis

Midterm Exam

Due October 21, 2019



Above is a figure from the book *Molecular Evolution: A Phylogenetic Approach* by Roderick D.M. Page and Edward C. Holmes. You will try to reproduce this result for the midterm exam. You may use the programs you developed or publicly available tools (please cite tool). In either case describe briefly the algorithm being used.

1. Extract sequences for the DNA or protein for a specific gene from a database for human, chimpanzee, gorilla, orangutan, and gibbon. You may pick the gene.
2. Create a distance matrix giving pairwise distances between the protein for the different species.
3. Build a phylogenetics tree to describe the relationship between these primates based on the gene of choice. Use any method you choose (it can be your code or public program).
4. Does it agree with Fig. 6.10 above. Why or why not?
5. Please email me the result by 9 am November 9th in a word document or pdf file.