

Phylogeny – TP1

Contact : juliana.silva_bernardes@upmc.fr

You have to send your answers to this e-mail address. Your report must be a single pdf file, it have to be simple, clear and well organized. The first page should be a summary indicating the page of each answer.

► INSTALL PHYLIP program

<http://koti.mbnet.fi/tuimala/oppaat/phylip2.pdf>

<http://evolution.gs.washington.edu/phylip/download/phylip-3.696.tar.gz>

<http://evolution.gs.washington.edu/phylip/install.html>

► Download all sequences required for the exercises at

<http://www.lcqb.upmc.fr/julianab/teaching/PHYL/TME1.zip>

Exercise 1: UPGMA

1. Take the distance matrix below and construct a tree by using the UPGMA algorithm. Explain each step showing all values.

	A	B	C	D	E	F	G
A							
B	19.00						
C	27.00	31.00					
D	8.00	18.00	26.00				
E	33.00	36.00	41.00	31.00			
F	18.00	1.00	32.00	17.00	35.00		
G	13.00	13.00	29.00	14.00	28.00	12.00	

Exercise 2: PAH

Here we consider the phenylalanine-4-hydroxylase enzyme. Its role is to degrade phenylalanine. In human, a mutation in its gene is responsible for the phenylketonuria disease.

1. Download the **human**, **rat** (*Rattus norvegicus*), **mouse** (*Mus musculus*), **cow** (*Bos taurus*) and **Caenorhabditis elegans** sequences of phenylalanine-4-hydroxylase from SwissProt <http://www.uniprot.org/> (search for “phenylalanine-4-hydroxylase” to get them) . Align them using CLUSTAL (you can use <http://www.ebi.ac.uk/Tools/msa/clustalo/>)
Save it in Phylip format and include this file in your report.

2. Use the program **protdist** (PHYLIP package) to compute a distance matrix for the downloaded sequences. Include the matrix in your report.
3. Use the program **neighbor** to compute an UPGMA tree and a Neighbor Joining tree. Include both trees to your report. For better visualization, you can use <http://iubio.bio.indiana.edu/treeapp/treeprint-form.html>
4. Considering NJ produces an unrooted tree, are the two trees different? Justify our answer.
5. What is an additive matrix? Is the distance matrix produced in 2 additive? Use the 4 points condition.

Exercise 3: CFTR

The Cystic Fibrosis Transmembrane conductance Regulator is a protein that regulates the movement of chloride and sodium ions through epithelial cells membranes. A mutation in this protein is the cause of cystic fibrosis (*mucoviscidose* in French).

1. Consider the “CFTR in mammals” alignment having proteins of several mammals. Compute the distance matrix with **protdist** (like in previous exercise)
2. Compute both UPGMA and NJ trees. Display and compare the tree. Include both trees in your report.
3. What do you observe for **rat** and **mouse** with UPGMA? With NJ? Look at these species in the distance matrix, what do you observe?
4. Now look at the **pig** (*Sus scrofa*) in your trees. Consider its position relative to cow (*Bos taurus*), mouse (*Mus musculus*) and horse (*Equus caballus*). The correct tree of placental mammals is provided below. Is the pig nearer to cow or horse? It is what is found by UPGMA and NJ? (To help, you can write a simplified unrooted tree with these 4 species.)

Exercise 4: P53

Tumor suppressor P53 is a protein involved in cell apoptosis and growth regulation. This protein is a protection against cancer as it prevents the cell from forming a tumor. But if its gene gets mutated in somatic cells, then the protein is no longer functional and the cell can generate a tumor. Therefore, when we sequence tumor cells, we often see many mutations in this gene.

1. Consider the P53 alignment and compute the distance matrix and NJ tree (like in previous exercises)
2. Now compare the position of these 4 species between the CFTR and P53 **NJ trees**: ***Homo sapiens***, ***Felis catus*** (cat), ***Loxodonta africana*** (elephant) and ***Monodelphis domestica*** (opossum). Which species group together in which tree?
3. Look at the correct tree (below) again. Opossum is a marsupial and therefore is outside this tree. Is the CFTR NJ tree correct? Is the P53 tree correct?

Correct phylogenetic tree of placental mammals.

