SuperTree - TME4

You have to send your answers to this e-mail address, subject [PHYL-2015] TME4. Your report must be a single pdf file, it should be simple, clear and well organized. The first page should be a summary indicating the page of each answer. Put your Scripts in the Annexe.

Use Forest to visualize trees, see Forest.readme.

Exercise 1: Constructing phylogenetic trees for thousands species *(Bonus)*

1. What are the problems found in the construction of phylogenetic trees having thousands species?

2. What are the strategies for reconstructing phylogenetic trees for large datasets?

3. What are the strategies for handling data fragmentation?

4. Describe the Maximum Agreement Subtree algorithm.

5. Describe the Maximum compatible trees algorithm

Exercise 2: Constructing a phylogenetic tree from a single domain family

1. Download the *Ribosomal_S27 (PF01599)* proteins from PFAM database or take it from RB_sequences.tar.gz.

2. Select sequences from the species defined in the list RB.txt.

3. Align and build the phylogenic tree using the program neighbor of phylip package.

4. Visualize. The clades were grouped together? In order to facilitate the visualization put a different color for each clade, see clades.txt. Include this tree in your report.

Exercise 3: Constructing a phylogenetic tree from a set of domain families

1. Consider the PFAM families in RB_sequences.tar.gz and select sequences from species defined in the list RB.txt

2. Align the selected sequences of each family, concat aligned sequences and build the tree by using neighbor of phylip package. Visualize. The clades were grouped together?

Exercise 4: Super Trees

1. Consider sequences in RB_clades.tar.gz, sequences in each family were splitted according to their clades. For each pair of clades do: align sequences of each family, concat them and create a phylogenetic tree. It is necessary to create a script to do these tasks, include the scripts in your report;

2. Use the program treePack to combine the trees produced in the last step. All trees must be in a single file (one per line). See treePack.readme to generate a super Tree. Visualize. The clades were grouped together?

3. Compare and discuss the three trees produced. What is your conclusion?