Molecular Evolution & Phylogeny: practical 1

Bioinformatics Community of Practice, BecA-ILRI Hub, Nairobi, September 2018

Distance matrix methods; parsimony

Tools used in this session

For some exercises in this this practical you will use either your local install of the MEGA software or the online server at https://ngphylogeny.fr. You may also use software installed on the HPC at ILRI Nairobi.

Task 1: Distance matrix methods

Below is a distance matrix giving the expected number of substitutions for 100 sites forming an alignment of the given OTUs (apes and human, source: Wen-Hsiung Li et al, 1987).

OTU	Chimpanzee	Gorilla	Human	Orangutan	Rhesus monkey
Chimpanzee	0				
Gorilla	1.57	0			
Human	1.45	1.51	0		
Orangutan	2.94	3.04	2.98	0	
Rhesus monkey	7.55	7.39	7.51	7.10	0

Browse http://evolution.genetics.washington.edu/phylip/doc/distance.html to know how to write the distance matrix above into Phylip format, and use a pure text editor in order to do so. Save and close your file.

Let's use the BioNJ algorithm!

For this task, please use the server https://ngphylogeny.fr . Go to the tab "online programs" and then "BioNJ". Run BioNJ on the server after uploading directly that distance matrix. It builds for you a tree based on the distances in the input matrix.

Task 2: Maximum Parsimony

Consider the following nucleotide alignment:

seq1 **A A G A G T G C A**
seq2 **A G C C G T G C G**
seq3 **A G A T A T C C A**
seq4 **A G A G A T C C G**

\<site\> 1 2 3 4 5 6 7 8 9

How many internal nodes would one get in an unrooted binary tree on these four sequences? Find the sites that are informative from the viewpoint of Maximum Parsimony methods, determine the unrooted Maximum Parsimony binary tree and then a possible corresponding set of ancestral sequences.

Task 3: Maximum Parsimony on the server or using MEGA

Run the same analysis as in task 3 using parsimony analysis on the server (go to "A la carte" phylogenetic analysis and disable all steps but the construction and visualisation of the phylogenetic tree) or on your local MEGA. Compare with the result you got for task 2, including the total number of parsimony steps.

Task 4: BioNJ and Maximum Parsimony on a real dataset

Use the file "seven_seqs.fa" and perform BioNJ and Maximum Parsimony tree reconstruction methods on it. Compare the results.