

---

## Phylogenomics

### Work locally on your laptop

The following instructions have been tested on Ubuntu. Mac and Windows users can use Ubuntu in the Virtual Machine.

1. Download MEGA 7:
  - Go to the website <http://www.megasoftware.net/> and download the appropriate MEGA7
  - To get help:
    - **GUI:** <http://www.megasoftware.net/docs> or menu Help → Contents
    - **CC:** <http://www.megasoftware.net/pdfs/MEGA7-CC-Quick-Start-Tutorial.pdf> or  
type `man megacc`
      1. use `megaproto` to create analysis options files
      2. run `megacc -a megaOptionFile.mao -d dataFile.meg [options]` to run MEGA analyses
2. Download RAXML:
  - Download RAXML under <https://github.com/stamatak/standard-RAXML>  
or type `sudo apt-get install raxml`

### Exercise 1: MEGA

In this exercise you will work with an example file, which is included in the MEGA distribution: `mtCDNA.meg` (can be found in the Examples folder of MEGA, which is located in your home (Mac) or Documents (Windows) directory). It can also be downloaded from <http://www.megasoftware.net/examples/mtCDNA.zip>. This data set contains mitochondrial cDNA sequences of several primate species.

- a) Open the `mtCDNA.meg` file with MEGA
- b) Find the appropriate substitution model for this data set with MEGA. Try to understand the result table.
- c) Create a NJ, maximum parsimony and maximum likelihood tree. Compare the different phylogenies

- d) Try to time calibrate the maximum likelihood tree by clicking on “Compute” → “Timetree”.  
You know that Gibbon is the outgroup and that chimpanzee and bonobo diverged between 1.2 – 5.0 Ma. Additionally you know that orangutan and Sumatran diverged latest 13 Ma.

## Exercise 2: RAxML

- a) Convert the .meg file to a RAxML input file (relaxed PHYLIP format) using my converter program PGDSpider (<http://www.cmpg.unibe.ch/software/PGDSpider/>).
- b) Estimate a maximum likelihood tree using RAxML by specifying the input file (-s), the substitution model (-m) and the output file name (-n) (for further options type `raxmlHPC -h`):  
`raxmlHPC -s mtCDNA.txt -n out.nwk -m GTRCATI`
- c) Run a bootstrap analysis in RAxML:  
`raxmlHPC -s mtCDNA.txt -n out.nwk -m GTRCATI -b 1234 -# 100`  
(the `-b` turns on bootstrapping with 1234 as seed and `-#` specifies the number of runs)
- d) This can now be used to but bootstrap values on the maximum likelihood tree:  
`raxmlHPC -m GTRCATI -f b -t RAxML_bestTree.raxmlOut -z RAxML_bootstrap.raxmlOut -n out_bootstrap.nwk`
- e) Open the resulting file (e.g. with MEGA) and compare it with the maximum likelihood tree of MEGA