### Phylogenetic trees and alignment tools

**Exercise 1: Recreating the phylogeny of great apes using the Toll-like receptor proteins**

1. Collect sequences from UniProt (<http://www.uniprot.org/>)
   1. Find the human toll-like receptor 5. Search for: “Toll-like receptor 5”.
   2. Click the first entry (should be O60602)
   3. Go to ‘Sequence’
   4. Click on ‘FASTA’
   5. Copy/Paste this sequence.
   6. Search for: “tlr5 Pan troglodytes”.
   7. Click on the first entry (should be B9VH32)
   8. Modify the URL so that it says: ‘http://www.uniprot.org/uniprot/B9VH32.fasta’
   9. Search for: “tlr5 taxonomy:"Hominidae (great apes) [9604]"”.
   10. Select the toll-like receptor 5 of *Gorilla gorilla*, *Pan paniscus* and *Pongo pygmaeus*. Modify the ‘Show’ to see more sequences. Download the FASTA Uncompressed.
   11. Collect a Toll-lie receptor 5 from a non-great ape species to use it as an outgroup (for example, look for tlr5 from *Mus musculus*).
2. Save all the sequences in a file with an extension ‘.fasta’ or ‘.fa’.
3. Use your favourite online alignment tool:

* Clustal Omega: <http://www.ebi.ac.uk/Tools/msa/clustalo/>
* Muscle: <http://www.ebi.ac.uk/Tools/msa/muscle/>
* T-Coffee: <http://www.ebi.ac.uk/Tools/msa/tcoffee/>
  + You can convert the output a sequence format with: <http://www.ebi.ac.uk/Tools/sfc/emboss_seqret/>
* Or download Jalview: <http://www.jalview.org/> (Click on ‘Launch Jalview Desktop’)

1. Examine the alignment, all sofwares will always give you an ‘alignment’ no matter how different (unrelated) the input sequences are.
2. Infer the phylogeny using Maximum likelihood algorithm implemented in <http://www.phylogeny.fr/>
   1. Click on “Advanced”.
   2. Unclick “Multiple Alignment MUSCLE”, since your sequences are already aligned.
   3. Unclick “Alignment curation: Gblocks”.
   4. Click on “Create workflow”.
   5. Upload your alignment.
   6. Click on submit.
   7. Make sure that the tree is correctly rooted. Click on “Reroot (outgroup)” and then click on the name of your outgroup.
3. What do you think it may happen if you use some paralogues for your alignment? Replace the sequences of *P. pygmaeus*, *P. paniscus* and *P. troglodytes* for Toll-like receptor 4 sequences.
4. Re-do the phylogeny. Can you explain this?