MULTIPLE SEQUENCE ALIGNMENT:

Exercise 1: MA using NCBI BLAST online (<https://www.ncbi.nlm.nih.gov>)

Exercise 2: 1) Domain analysis (Interpro, https://www.ebi.ac.uk/)

* 2) Pfam (protein families) seed (contains seed alignment which are quality checked)

Exercise 3: 2) Clustal Omega and html files

* 3) Muscle command line
* 4)HMMAlign and transformation to standard format (stoTransform.py)

Exercise 4: Cut out domain of interest

* 1) identify domains with hmmsearch
* 2) use python to cut domains out of protein sequence (cut\_domains\_from\_proteins.py)

Exercise 5:

NJ TREES:

Exercise 3: Use MAFFT web server to construct phylogenetic trees (python file to modify FASTA file -> input is aligned sequence file)

4) pyhlo.io to visualize and compare trees

Maximum Likelihood:

Exercise 1) Newick format and visualization

* 2) RAxML for computing tree in command line using existing alignment