

# Comparative Assignment One

James O'Reilly  
Student Number: r0773125

Mpn515 is a beta subunit of a DNA-directed RNA polymerase found in *Mycoplasma pneumoniae*. This protein catalyses the transcription of DNA into RNA using the four ribonucleoside triphosphates as substrates. The top 25 homologs from different species were found using BLASTP with the standard arguments. The FASTA sequences for these homologs were then put in a single FASTA file for multiple alignment.

## 1 Online Multiple Alignment

The multiple alignment was first done using MUSCLE and then T-COFFEE. The phylogenetic trees were then visualised using iTol. Figures 1 and 2 show the phylogenetic trees.

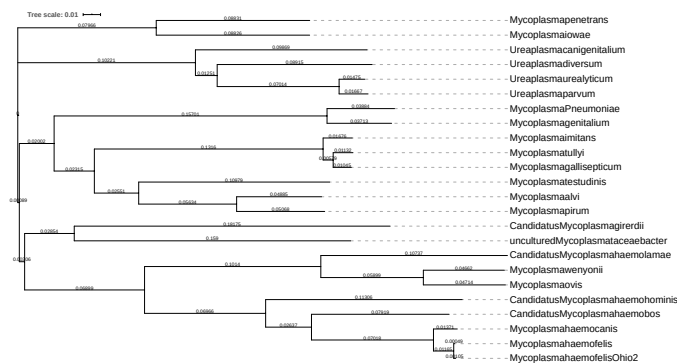


Figure 1: Phylogenetic tree with branch lengths, constructed from MUSCLE alignment

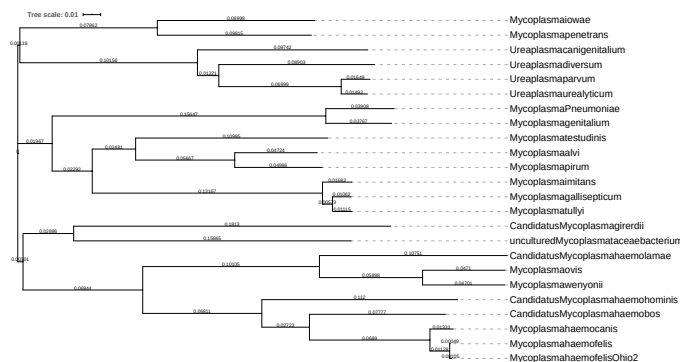


Figure 2: Phylogenetic tree with branch lengths, constructed from T-COFFEE alignment

## 2 Multiple Alignment on the Server

**How many sequences are there in the Fasta file?** The number of sequences can be found using the command `grep ">" filename | wc -l`. The `16S_Myco.fsa` file has 18 sequences and the `Rpoa_Myco.fsa` file has 18 sequences also.

Multiple alignment was performed using ClustalW with accurate alignment, 1000 bootstraps and neighbour-joining. Positions with gaps were ignored when creating the corresponding tree. The trees are given in Figures 3 and 4. Even though I set the `-BOOTLABELS=node` option in ClustalW, I could not visualise the bootstrap values in either iTol or FigTree.

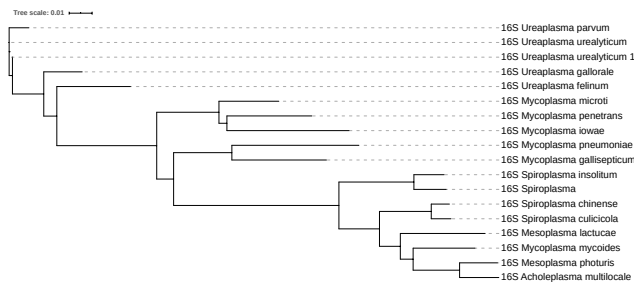


Figure 3: Phylogenetic tree for 16S, constructed from CLUSTALW alignment

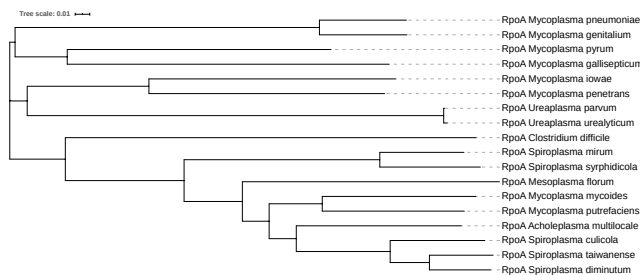


Figure 4: Phylogenetic tree for RpoA, constructed from CLUSTALW alignment

## References