TME4 - PHYG

Witold PODLEJSKI & Jérémie PERRIN

December 5, 2019

1 Exercise 1

1.1 Question 1
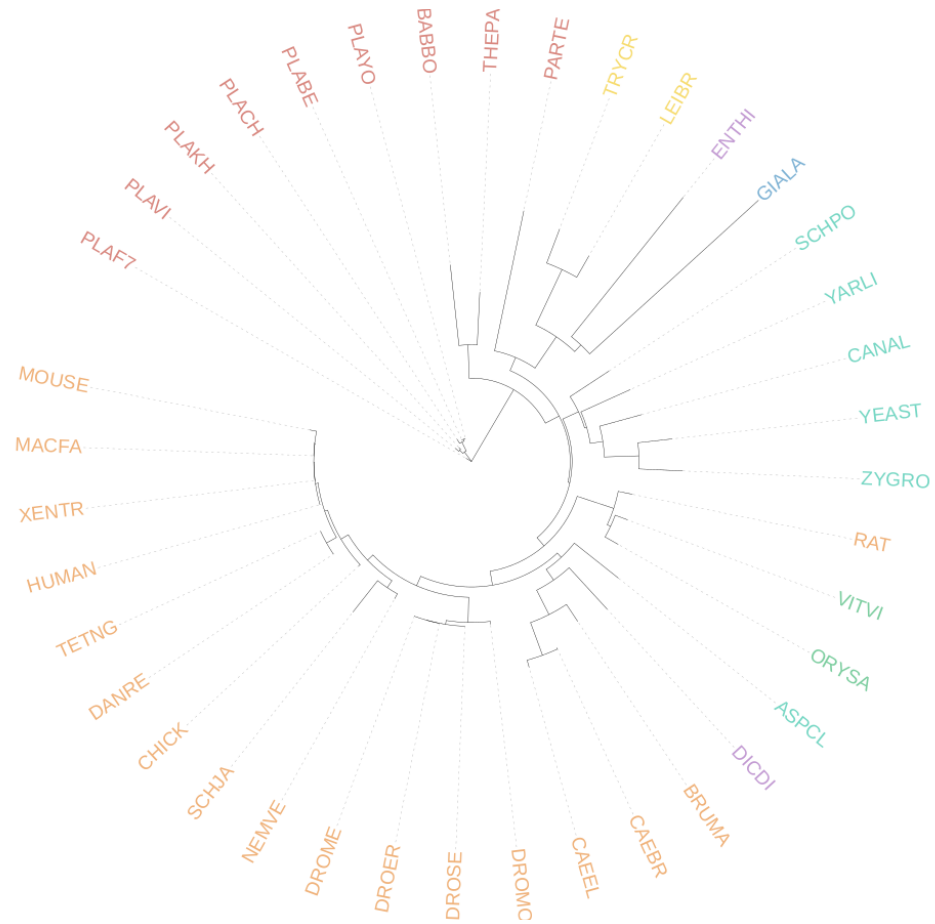
content

1.2 Question 2

content

2 Exercise 2

Figure 1: The tree inferred with the Neighbor-Joining algoritm on one gene familly
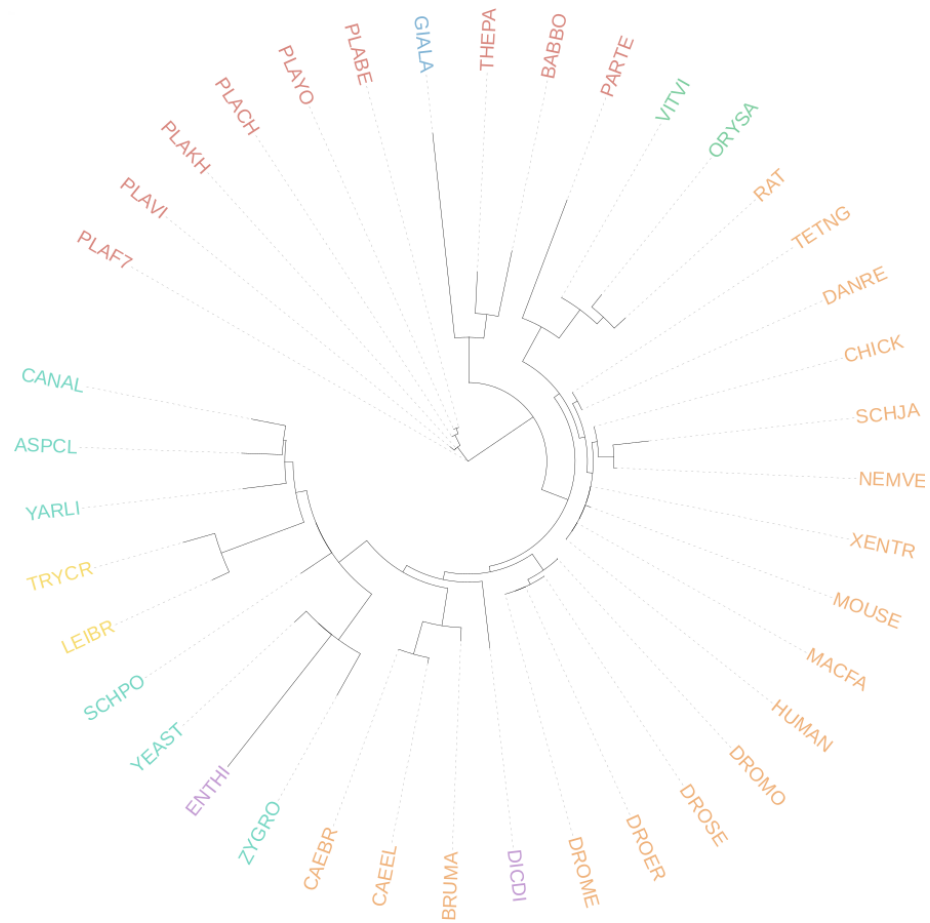
Figure 2: The tree inferred with the Most-Likelihood algoritm on one gene familly

These two trees are quite coherent with the knowed clades witch are mostly put together. There is some mistakes we *TRYCR* and *LEIBR*. So we see that the tree contain a lot of useful informations but need to be compared with other gene famillies to get rid of wrong branches.

The results of NJ and ML algoritm are equivalent, except for the *RAT* witch is wrongly positionned with NJ.
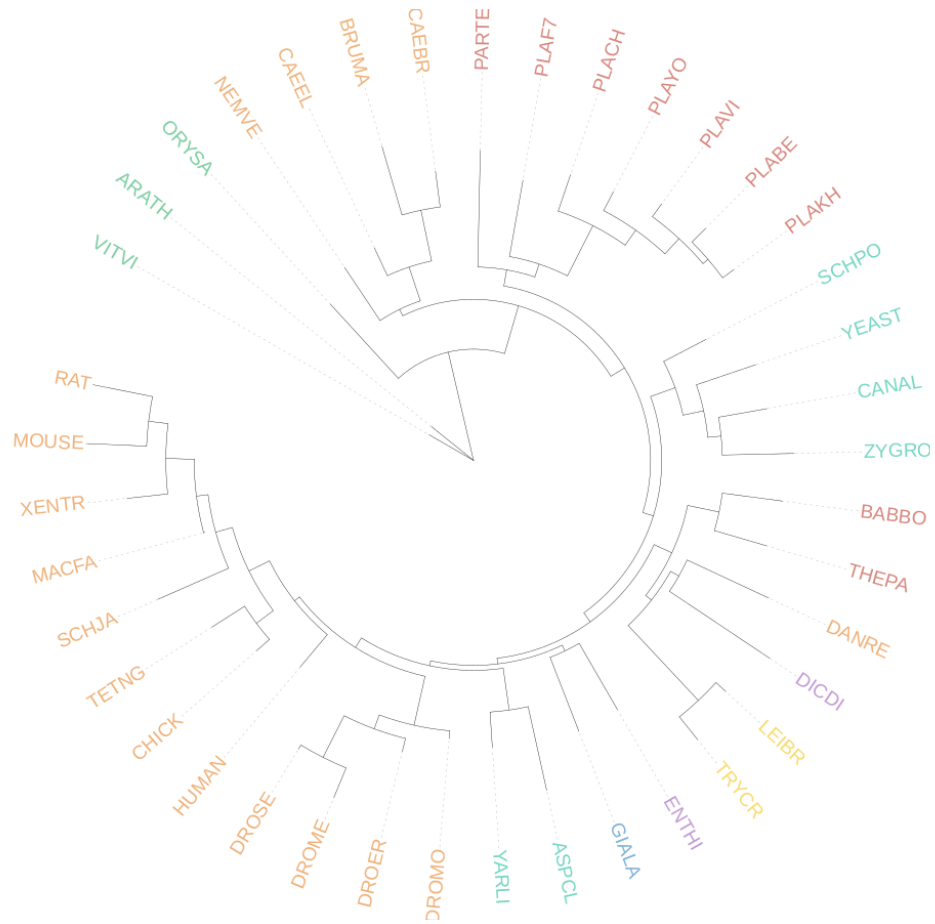
3 Exercise 3

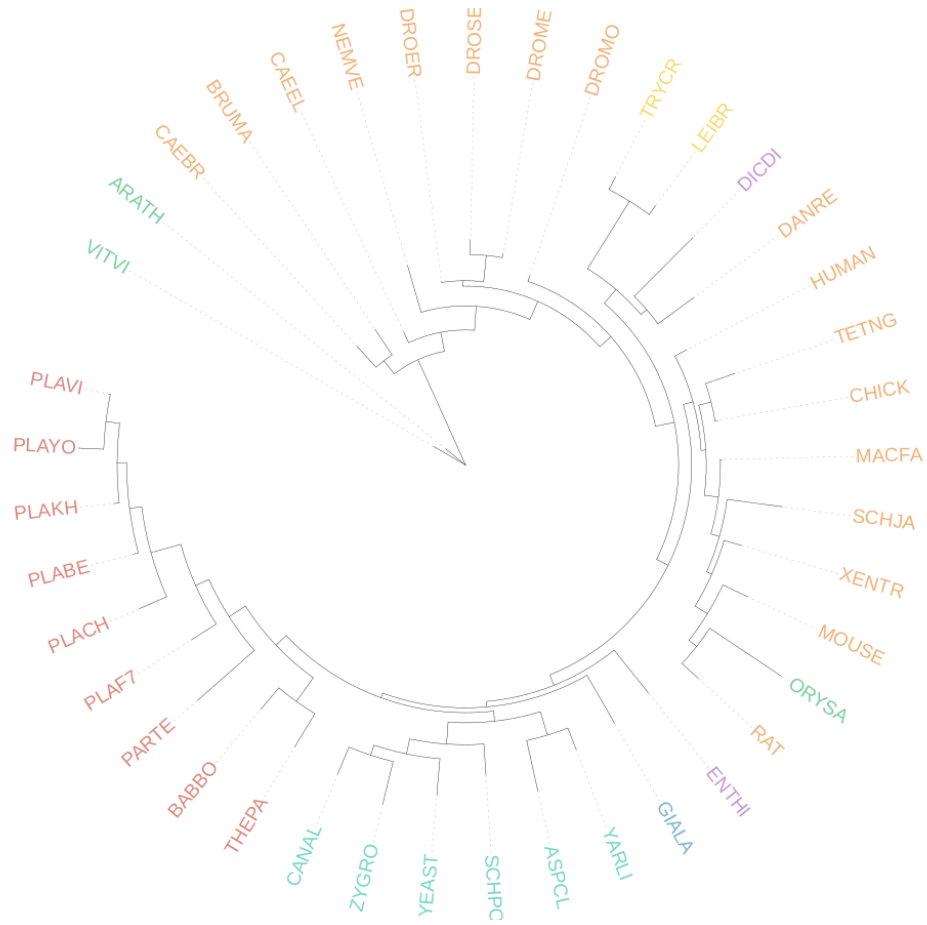Figure 3: The tree inferred with the Neighbor-Joining algoritm on all gene famillies

Figure 4: The tree inferred with the Most-Likelihood algoritm on all gene families

We can observe that the results are even worst than the last result with one familly. This can be explained by the fact that we do not manage the batch effect.

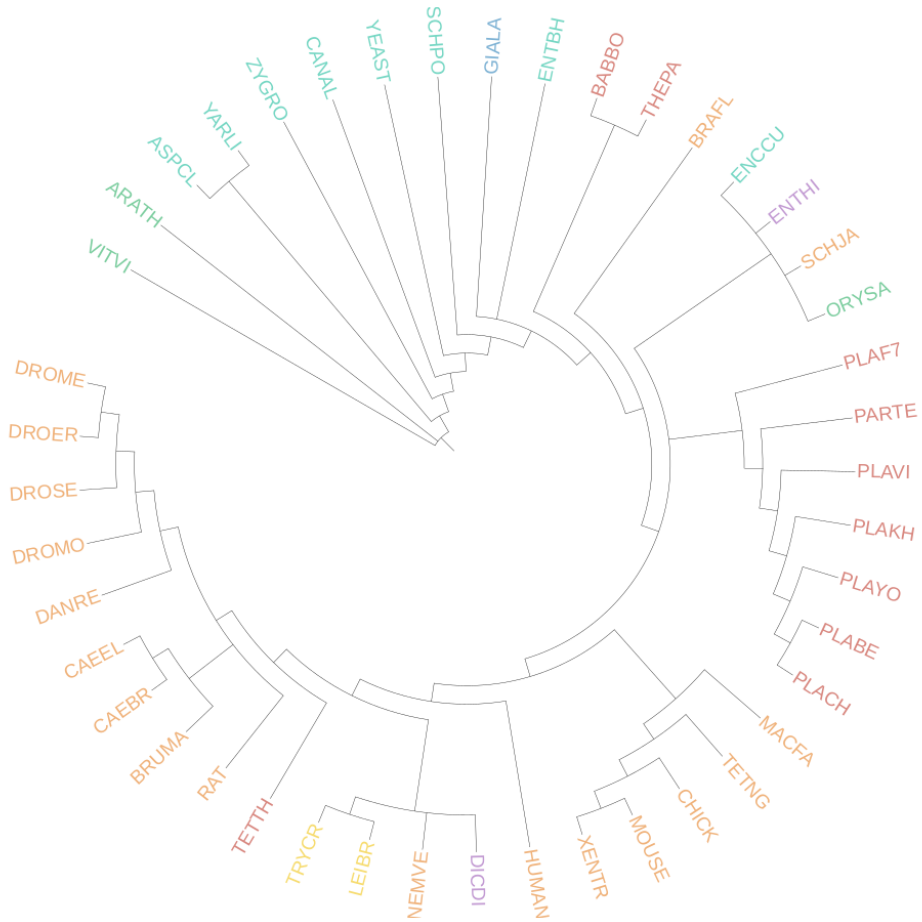The results are again similar between the two different algoritms.

4 Exercise 4

Figure 5: The tree inferred with the Neighbor-Joining algoritm and a supertree combination algoritm on all gene famillies
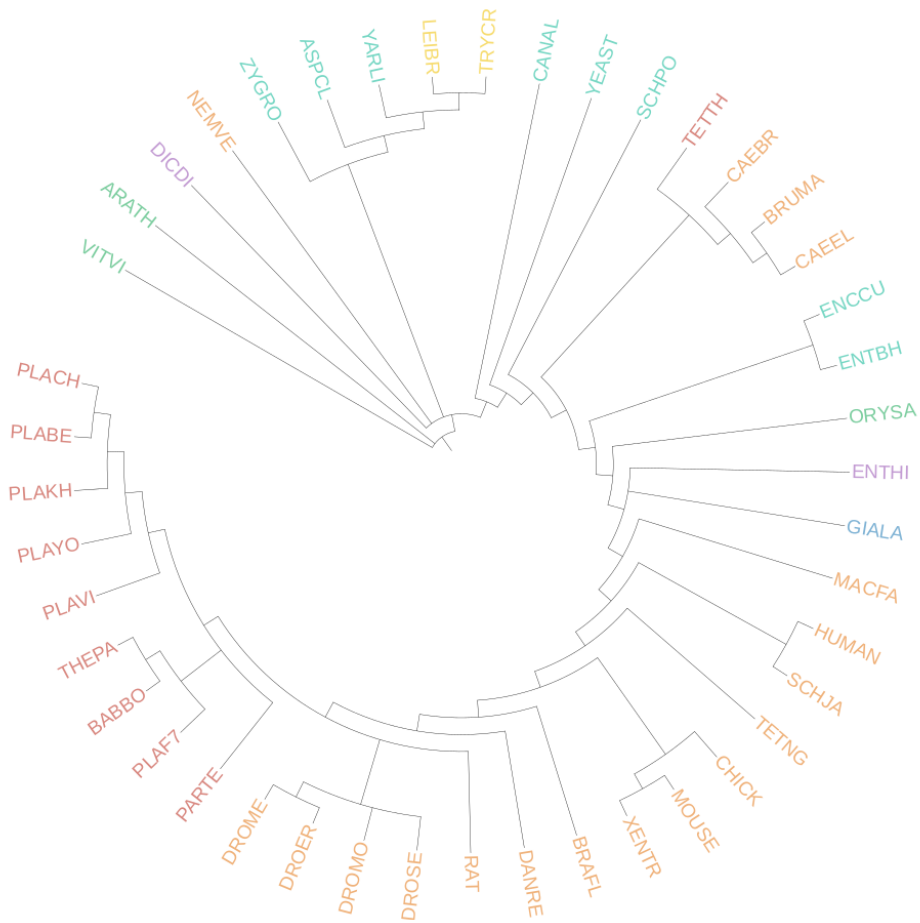
Figure 6: The tree inferred with the Most-Likelihood algoritm and a supertree combination algoritm on all gene famillies

These two last trees are well ordered and regular in terms of mutatiobal distances. Even if the trees are not the same, their quality seem to be good in anycase.

5   Exercise 5

With the last three exercises the bests trees are without supris the supertrees of exercise 4. In fact its combine all the informations contained in each gene families available and sumerize it in the consesus tree.