

Assignment 4.1

Mahdi

October 31, 2020

Which is the ancestral state of the OBP gene family size in the phylogeny?

```
((((((((((((dmel<5>_52:5,(dsec<1>_51:1,dsim<0>_53:1)<4>_52:4)<7>_52:8,(dyak<3>_55:10,
dere<2>_50:10)<6>_52:3)<11>_52:31,dana<10>_50:44)<15>_50:11,(dpse<9>_45:1,dper<8>_45:1)
<14>_45:54)<19>_49:7,dwil<18>\*_62:62)<21>_49:1,((dmoj<13>_43:31,dvir<12>_41:31)<17>_44:12,
dgri<16>_46:43)<20>_46:20)<23>_49:187,agam<22>\*_83:250)<25>_44:10,bmor<24>_46:260)
<27>_43:20,tcas<26>_49:280)<31>_40:20,amel<30>_21:300)<33>_36:50,(phum<29>\*_5:260,
apis<28>_18:260)<32>\*_21:90)<34>_30;
```

The most recent common ancestral state (34) has a size of 30.

Which is the general mode of evolution of the OBP gene family across Hexapoda? Has it undergone different dynamics across the surveyed species?

Overall, the size of gene families across Hexapoda is changing because the p value is significant so we can reject the null hypothesis that there is no change across the branches. 6 species out of 18 are expanding, 7 species are contracting and 5 species do not change size.

Are there differences in the OBP dynamics between Drosophila and the other Hexapoda species surveyed? Is the birth-and-death process constant across the phylogeny?

Yes, there is a difference since the 2 lambda model has a higher base family likelihood than the single lambda model. The birth-and-death process is different between the two groups since the 2 lambda model finds two different lambda values for them.

Which is the dN/dS ratio in the three duplicated OBP genes? Are there different functional constraints among copies? Which is the most likely reason of this observation?

The dN/dS ratios of the duplicated genes are 0.340, 0.182 and 0.543. Since the values are all different, there are different functional constraints among the copies. If the genes are under functional constraint, that must mean the original sequence must be preserved. As long as one of the gene duplicates is conserved, (i.e. has a high functional constraint), the other duplicates can have a lower constraints since the first duplicate is performing the required function.

Is there any specific site under positive selection?

Codon 72 is under positive selection