



Inferring Phylogenies:

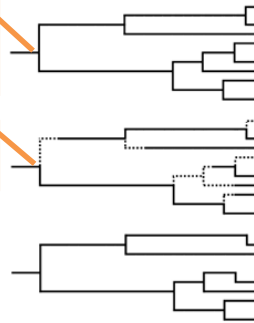
BEAST2 and the Protracted Birth-Death Model

In nature, speciation takes time. Despite this, many phylogenetic inference programs use models that assume speciation to happen instantly. One such model is the Birth – Death model (BD). One model which does take speciation time into account is Protracted Birth – Death (PBD).

A new species instantly becomes a good species

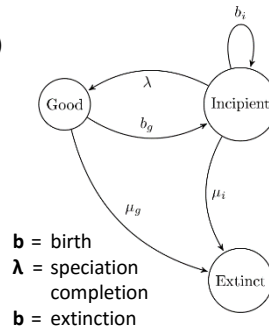
A new species is incipient first

Resulting phylogeny



PBD ($\lambda = \infty$)
a.k.a BD

PBD ($\lambda < \infty$)



PBD adds an incipient stage before a species becomes a good species. In this research, we test if BEAST2, a phylogenetic inference program based on BD, can accurately recover trees that were simulated using the assumptions of PBD.

Method

parameters	toy_example_2.RDa
slrg	0.50
siri	0.50
scr	0.10
erg	0.10
eri	0.10
age	5.00
sequence_length	1000.00

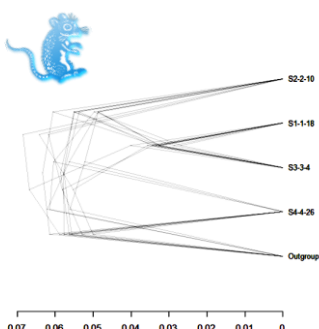
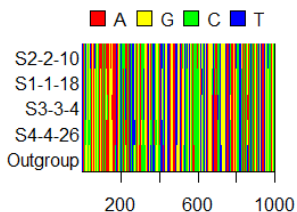
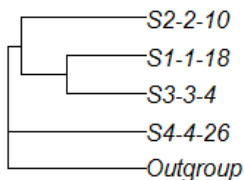
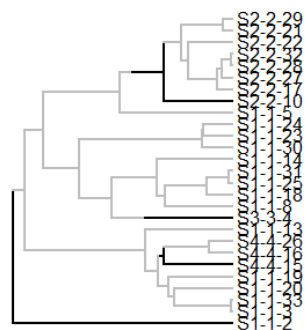
Step 1:
Simulate desired parameter files

Step 2:
Simulate one incipient species tree per parameter file

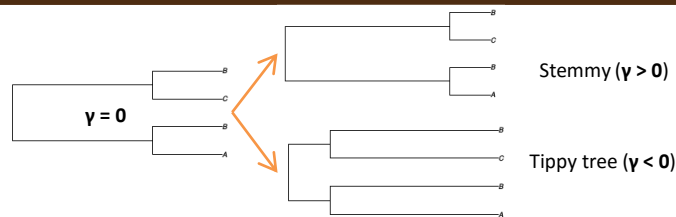
Step 3:
Sample two monophyletic species trees per incipient tree

Step 4: Simulate two DNA alignments per sampled species tree

Step 5: Run BEAST2 two times per alignment to get posteriors (eight in total per parameter file)

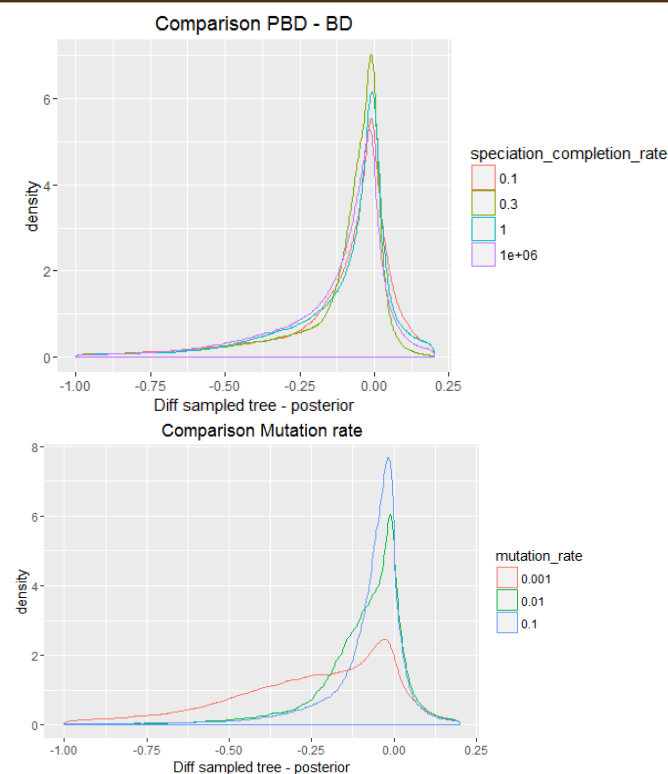


The γ statistic



The gamma statistic provides insight in the distribution of branching events within a tree. It shows how 'tippy' or 'stemmy' a tree is.

Results



Comparisons of the γ statistic of the sampled PBD tree and the posterior. 0 indicates no difference, <0 a higher posterior γ , and >0 a higher γ of the sampled PBD tree.

Different speciation completion rates (SCR) appear to have little impact on the error in the recovery of the gamma PBD trees by BEAST2.

Other parameters, like the mutation rate shown here, have a much greater impact on the difference in γ .

Discussion & Conclusion

Under all combinations of parameter settings, the gamma statistic was recovered quite well. We had expected lower SCR to increase deviations, but other parameters appear to have more impact on the correct recovery of a PBD tree than protractedness.

For this reason, with the results we have now, we think it is not necessary to include PBD in BEAST2. However, the gamma statistic is just one way to summarize a tree in one number – there could very well be other statistics that are less well recovered by BEAST2. This is why it would be prudent to look at other methods of comparing the sampled species trees and the posterior in further analysis.



References

Etienne, R. S., & Rosindell, J. (2012). Prolonging the past counteracts the pull of the present: protracted speciation can explain observed slowdowns in diversification. *Systematic Biology*, **61**(2), 204-213.