

**Appendix S2.** Diagnostics of the Bayesian inference analysis of the 38 taxa *Petrocodon* dataset of combined ITS and *trnLF* sequence data.

Number of taxa = 38

Number of characters = 1562

Average standard deviation of split frequencies: 0.005094

Analysis completed in 15879 seconds

Analysis used 15879.58 seconds of CPU time

Likelihood of best state for "cold" chain of run 1 was -7093.85

Likelihood of best state for "cold" chain of run 2 was -7099.15

Acceptance rates for the moves in the "cold" chain of run 1:

With prob. Chain accepted changes to

55.33 % param. 1 (revmat) with Dirichlet proposal

26.46 % param. 2 (revmat) with Dirichlet proposal

66.69 % param. 3 (revmat) with Dirichlet proposal

19.39 % param. 4 (state frequencies) with Dirichlet proposal

18.53 % param. 5 (state frequencies) with Dirichlet proposal

91.78 % param. 7 (gamma shape) with multiplier

46.15 % param. 8 (prop. invar. sites) with sliding window

12.14 % param. 9 (topology and branch lengths) with extending TBR

21.31 % param. 9 (topology and branch lengths) with LOCAL

Acceptance rates for the moves in the "cold" chain of run 2:

With prob. Chain accepted changes to

55.66 % param. 1 (revmat) with Dirichlet proposal

26.08 % param. 2 (revmat) with Dirichlet proposal

66.65 % param. 3 (revmat) with Dirichlet proposal

19.26 % param. 4 (state frequencies) with Dirichlet proposal

18.74 % param. 5 (state frequencies) with Dirichlet proposal

91.80 % param. 7 (gamma shape) with multiplier

45.65 % param. 8 (prop. invar. sites) with sliding window

12.08 % param. 9 (topology and branch lengths) with extending TBR

21.44 % param. 9 (topology and branch lengths) with LOCAL

Chain swap information for run 1:

	1	2	3	4
1	0.30	0.05	0.01	
2	250701	0.37	0.09	
3	249593	250641	0.43	
4	250809	249372	248884	

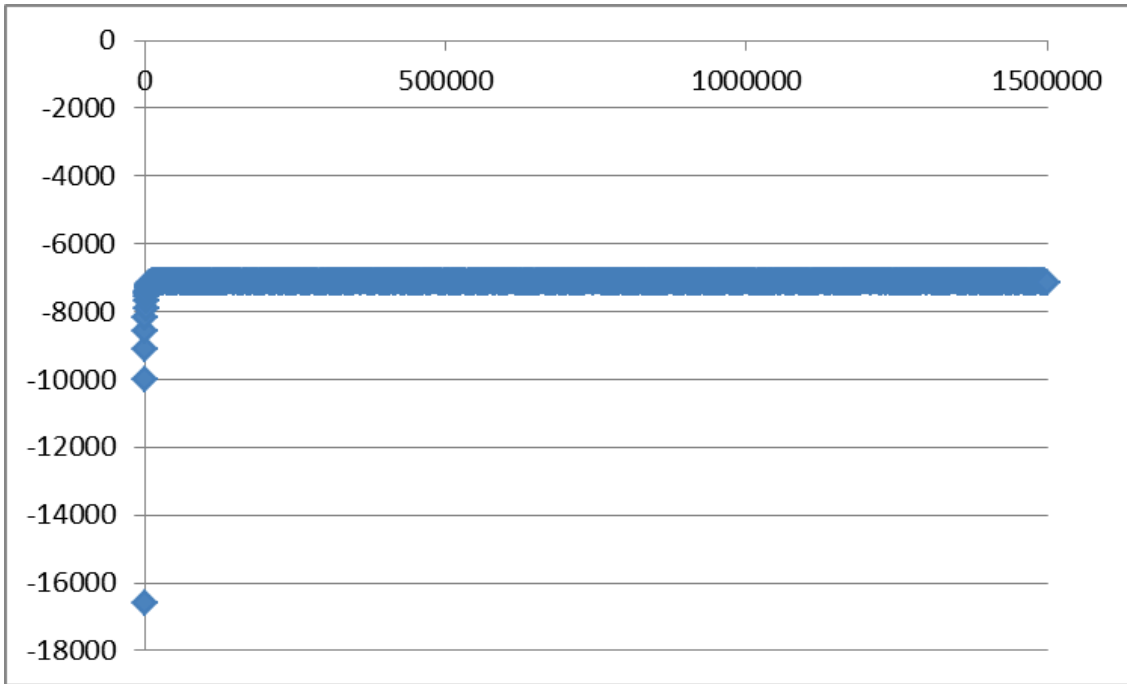
Chain swap information for run 2:

	1	2	3	4
1	0.30	0.05	0.01	
2	249123	0.36	0.09	
3	251143	249280	0.41	
4	249748	250178	250528	

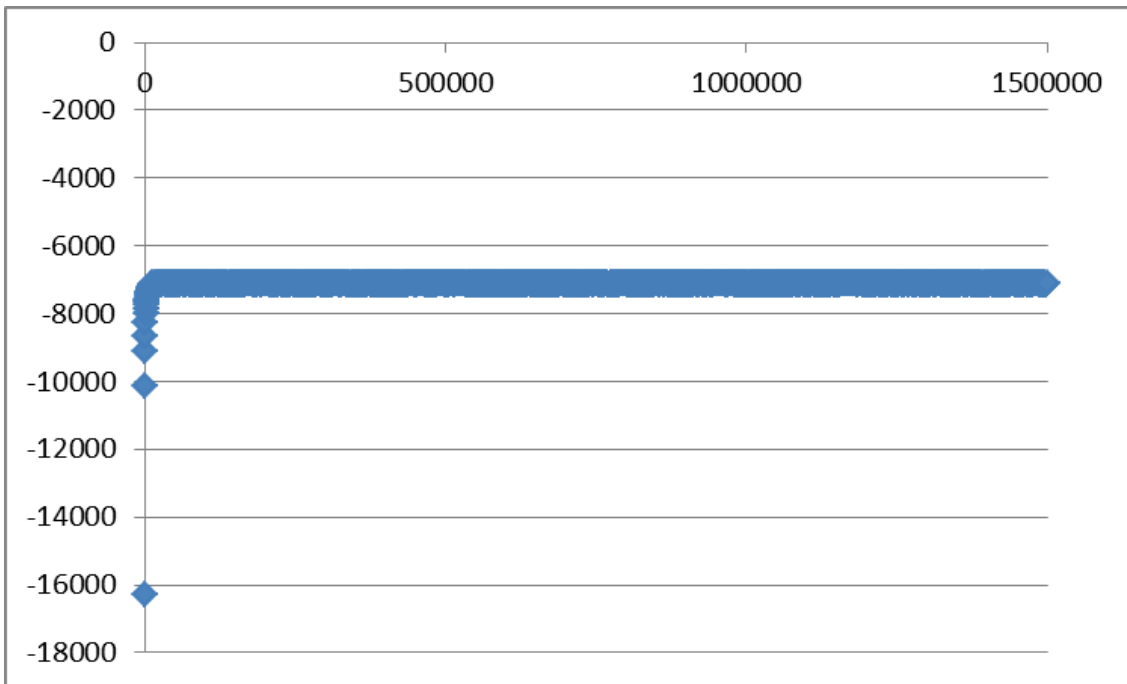
Upper diagonal: Proportion of successful state exchanges between chains

Lower diagonal: Number of attempted state exchanges between chains

A) BI run characteristics

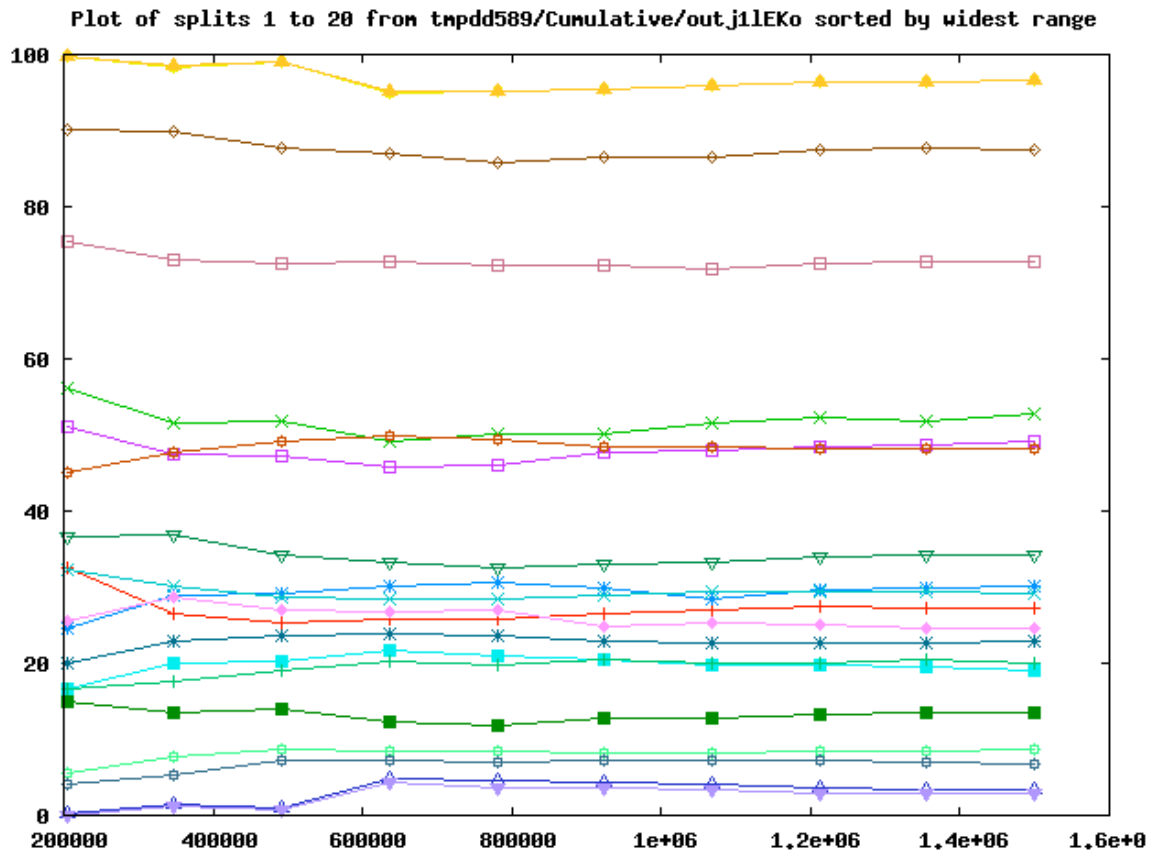


Run1

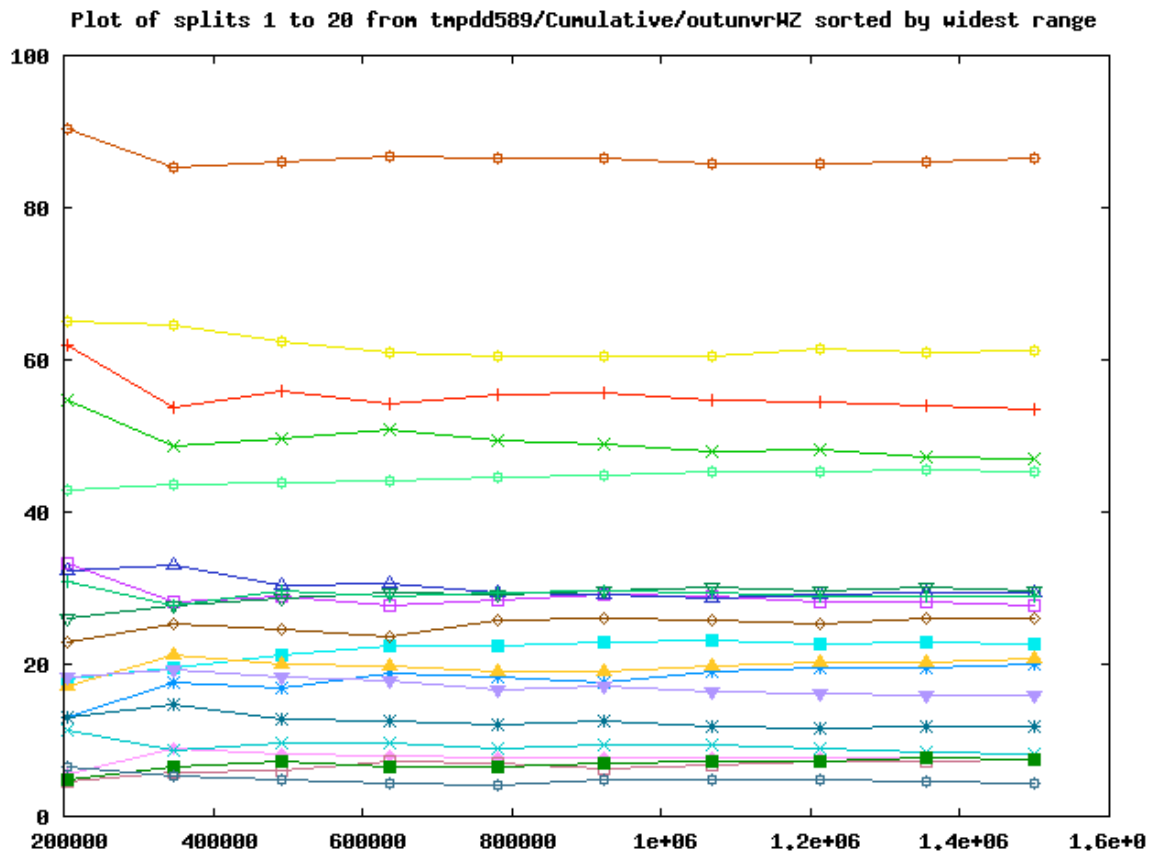


Run2

B) Scatterplots of  $-\ln$  Likelihood values  $v$  generation time to determine burn-in.

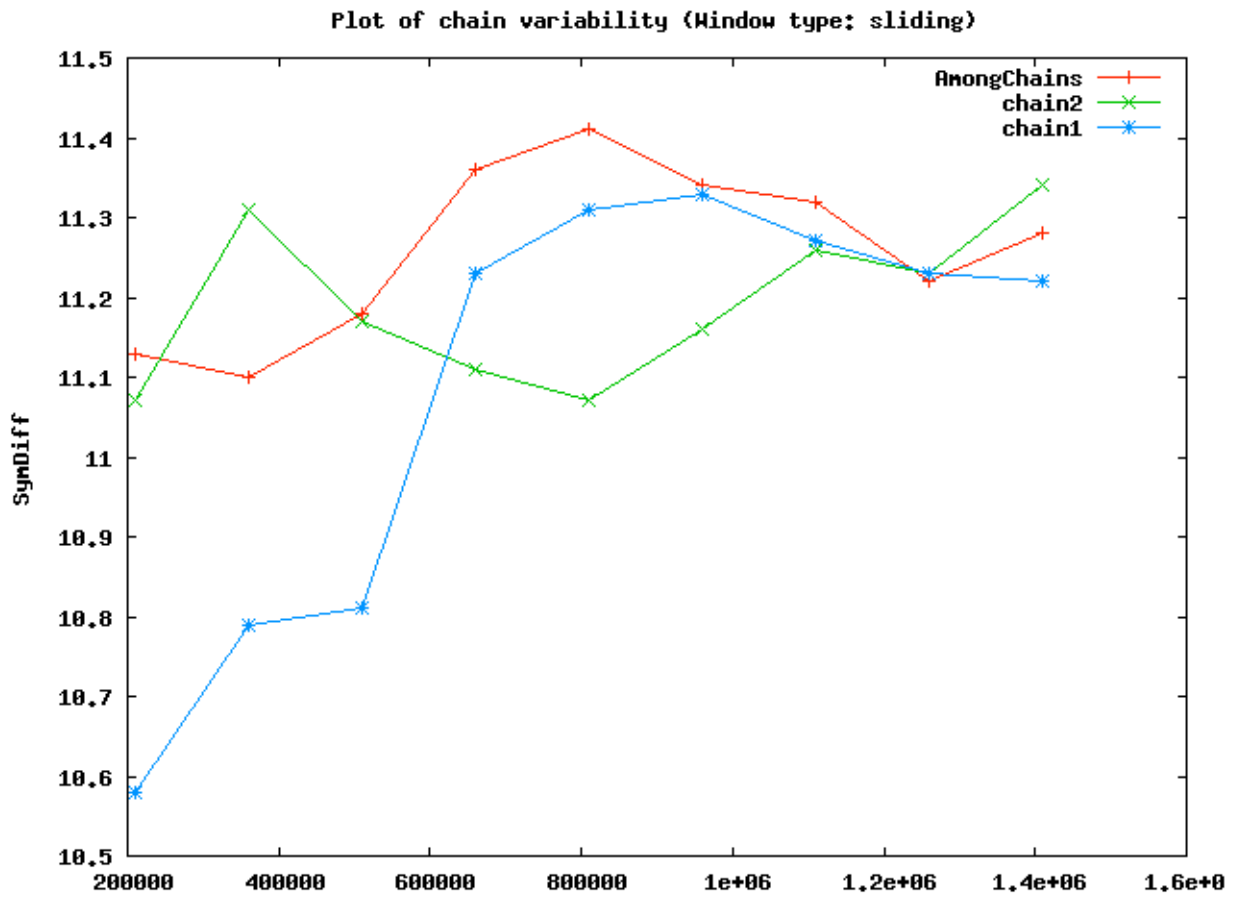


Run 1

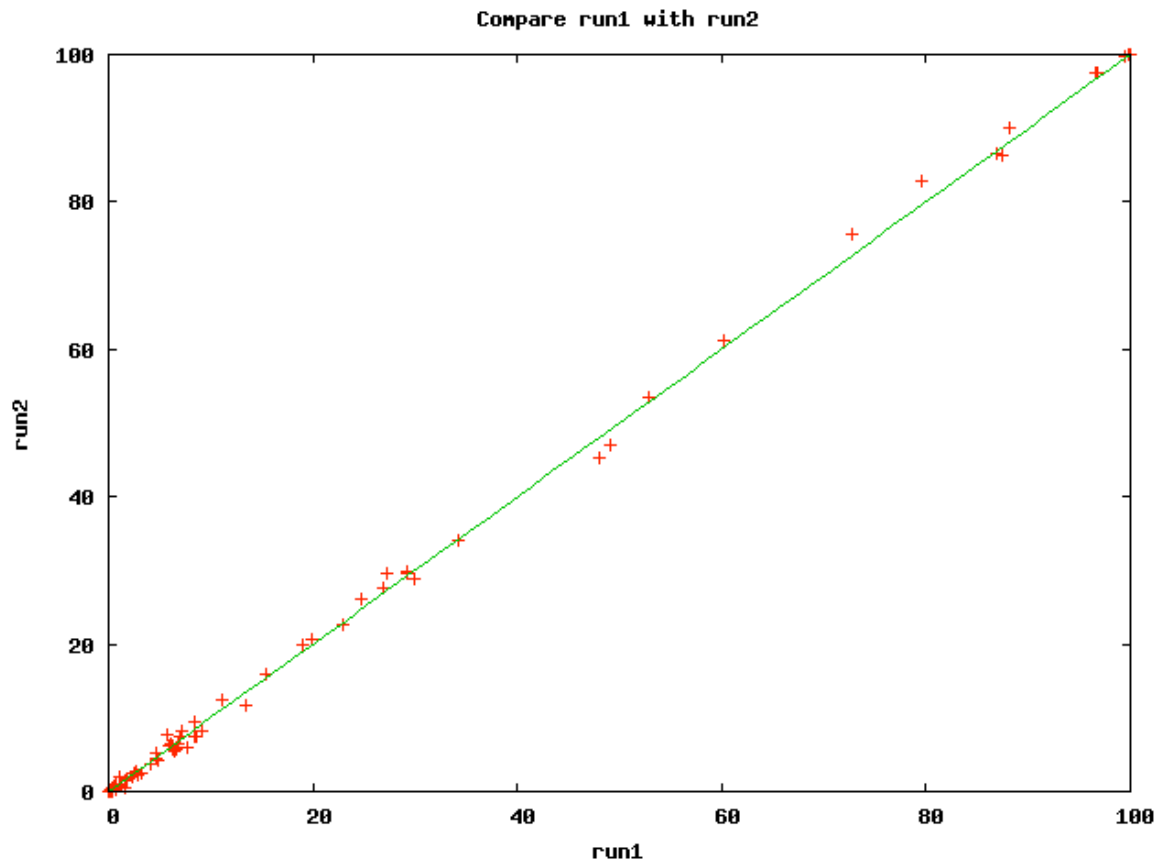


Run 2

C) Plots of splits sorted by widest range of run 1 and run 2.



D) Symmetric tree differences within and between run 1 v run 2 of combined data.



E). Scatterplot of posterior probabilities of run 1 v run 2 of combined data.