

# Does gene tree discordance explain the mismatch between macroevolutionary models and empirical patterns of tree shape and branching times?

## Supplementary Figures

Tanja Stadler<sup>1,2,\*</sup>, James H. Degnan<sup>3</sup>, Noah A. Rosenberg<sup>4</sup>

February 23, 2016

<sup>1</sup> ETH Zürich, Department of Biosystems Science and Engineering, Mattenstrasse 26, 4058 Basel, Switzerland

<sup>2</sup> Swiss Institute of Bioinformatics (SIB), Switzerland

<sup>3</sup> Department of Mathematics and Statistics, University of New Mexico, 311 Terrace NE, Albuquerque, NM, 87131, USA

<sup>4</sup> Department of Biology, Stanford University, 371 Serra Mall, Stanford, CA 94305, USA

\* Corresponding author: [tanja.stadler@bsse.ethz.ch](mailto:tanja.stadler@bsse.ethz.ch)

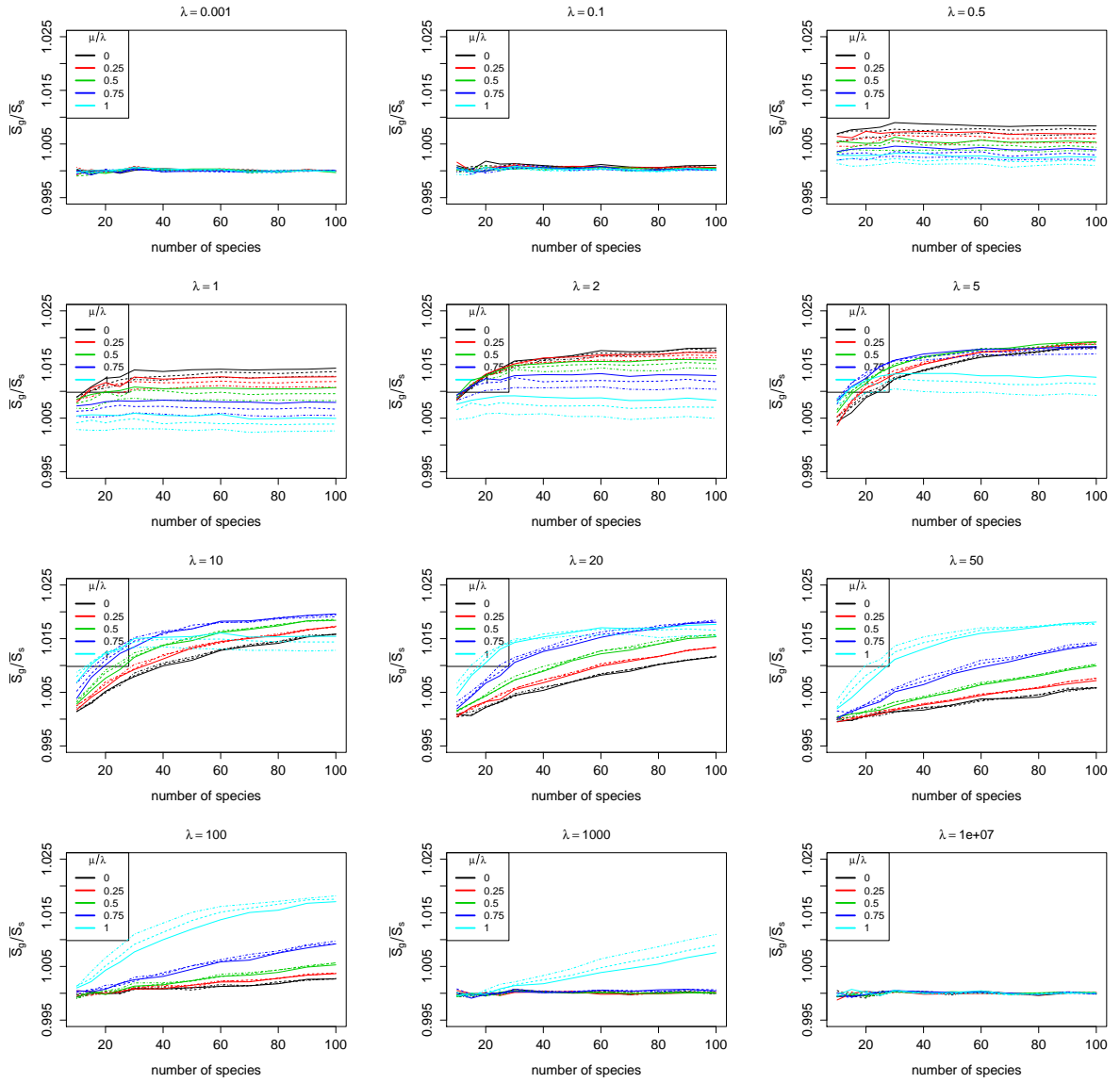


Figure 1: (Supplementary Figure). Mean Sackin statistic of gene trees divided by mean Sackin statistic of species trees ( $\bar{S}_g/\bar{S}_s$ ). Solid lines correspond to complete species sampling  $\rho = 1$ , dashed lines to sampling probability  $\rho = 0.75$ , and dot-dashed lines to sampling probability  $\rho = 0.5$ . Plots are obtained based on 100,000 simulated species tree–gene tree pairs at each choice of parameter values, taking means separately for the gene trees and the species trees.

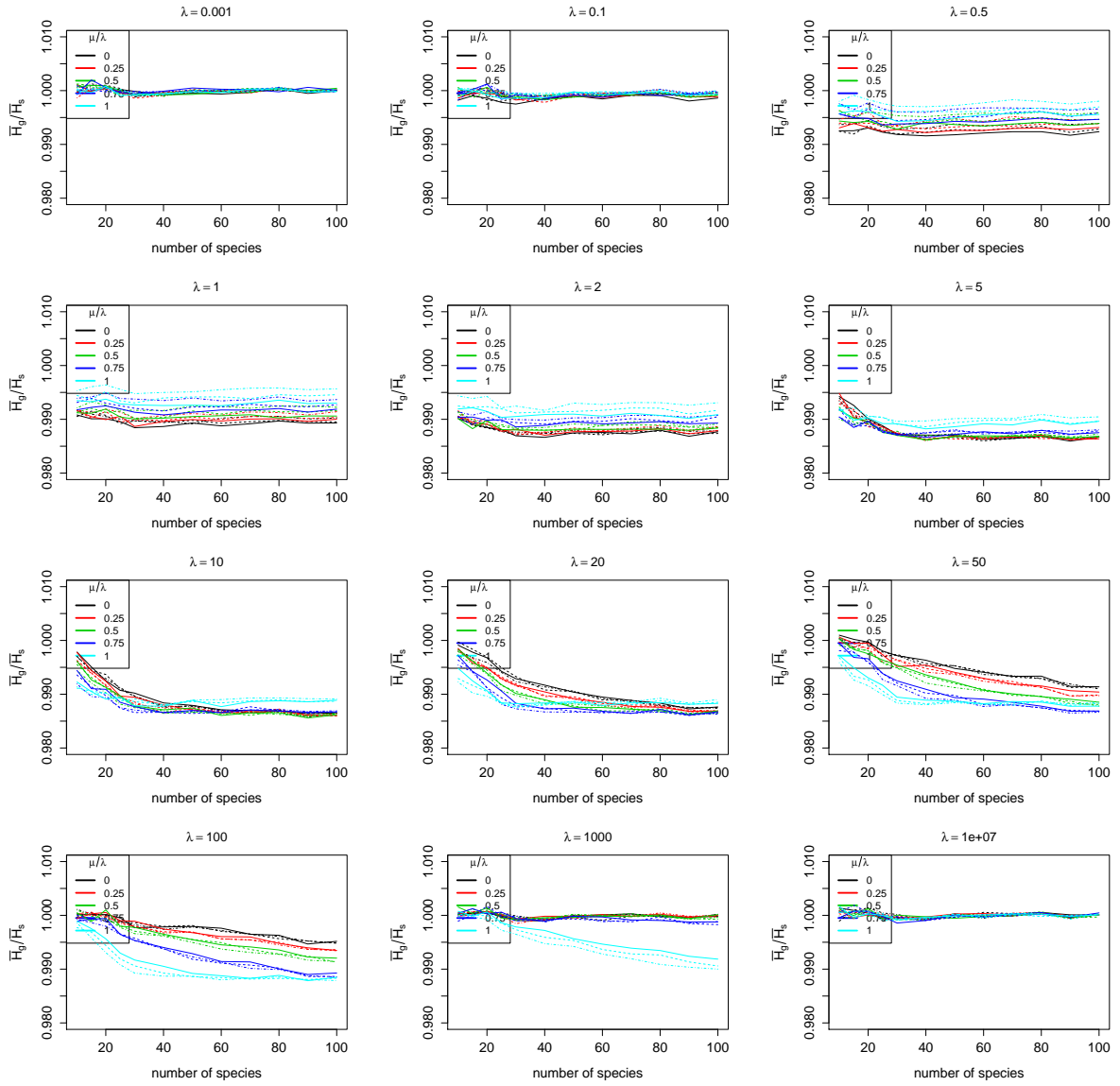


Figure 2: (Supplementary Figure). Mean number of cherries in gene trees divided by mean number of cherries in species trees ( $\overline{H}_g/\overline{H}_s$ ). Solid lines correspond to complete species sampling  $\rho = 1$ , dashed lines to sampling probability  $\rho = 0.75$ , and dot-dashed lines to sampling probability  $\rho = 0.5$ . Plots are obtained based on 100,000 simulated species tree–gene tree pairs at each choice of parameter values, taking means separately for the gene trees and the species trees.

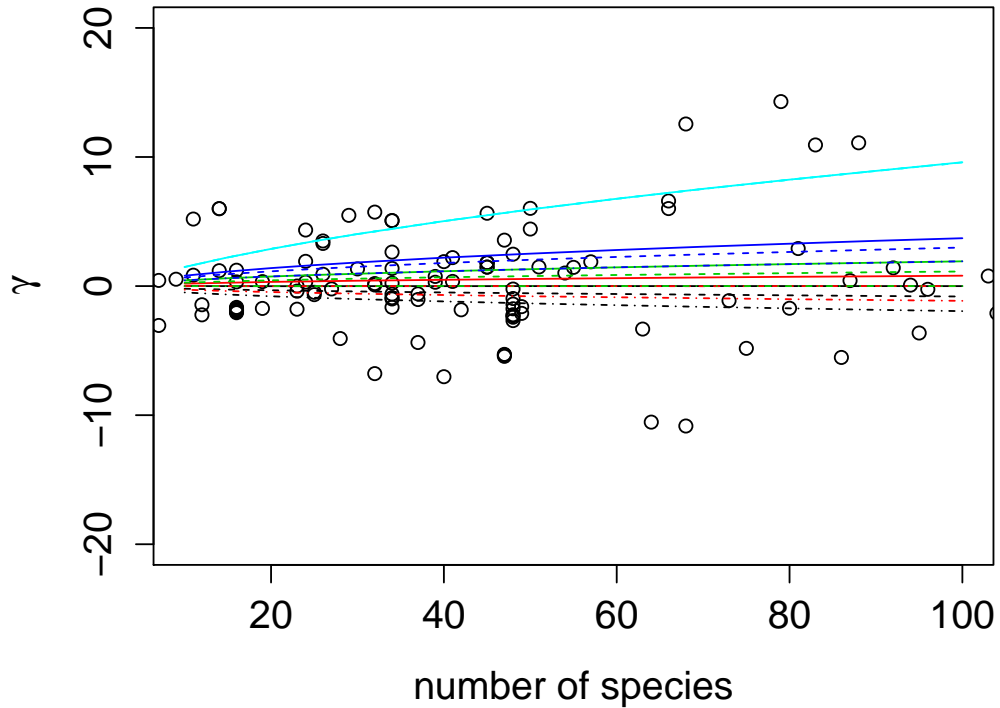


Figure 3: (Supplementary Figure). The  $\gamma$  statistic for empirical trees from TreeBASE. Each black dot represents a tree. The figure also shows simulated species tree  $\gamma_s$  values. The species tree  $\gamma_s$  value depends on turnover and species sampling; line colors indicate turnover, and line type indicates species sampling, as in Figure 2. Note that in the simulations, we subtracted mean species tree  $\gamma_s$  values from mean gene tree  $\gamma_g$  values to produce  $\bar{\gamma}_g - \bar{\gamma}_s$ . As the number of points is too low here to obtain mean  $\gamma$  values for empirical trees at each tree size, we only plot each empirically observed  $\gamma$  statistic.