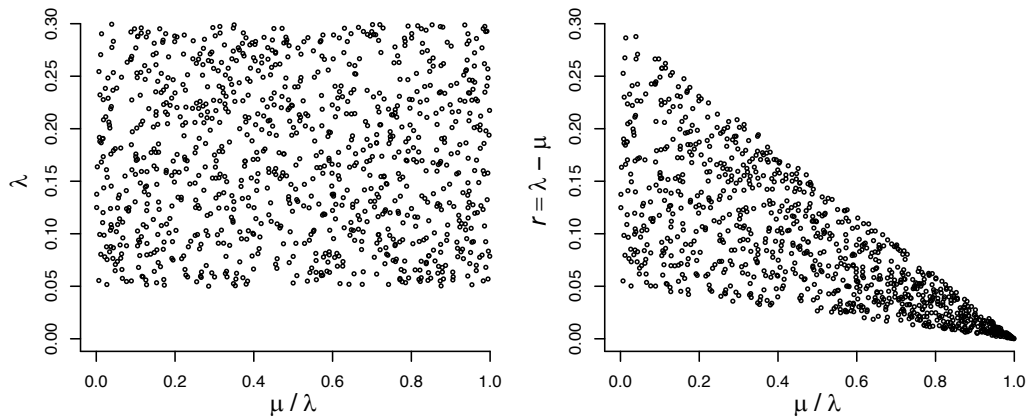


Tip rates, phylogenies, and diversification:
what are we estimating, and how good are the estimates?

Supplementary Information

simulations for evaluating speciation rate



simulations for evaluating net diversification rate

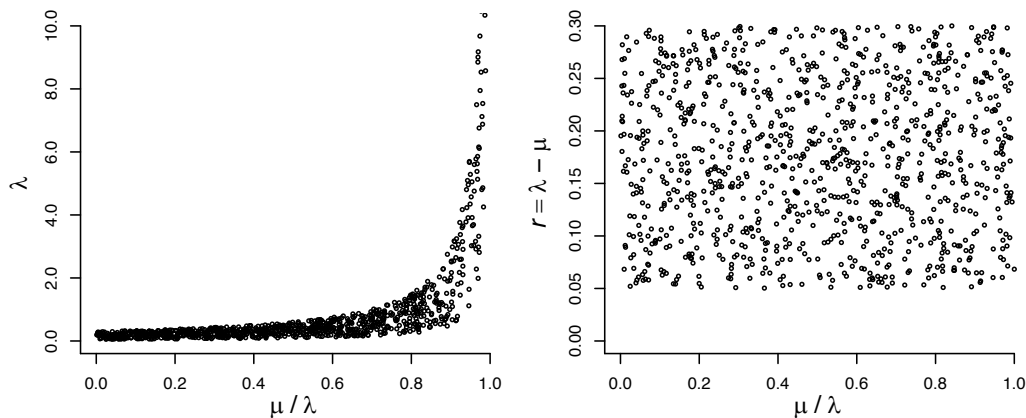


Figure S1. Details of the phylogeny simulations designed to evaluate the performance of the four tip metrics in terms of speciation rate and net diversification rate. From the top row, it is clear that when λ is sampled uniformly with respect to ε , the distribution of r is not uniform: the mean, range and variance in r decrease dramatically as ε increases. The reverse is true for the distribution of λ when r is sampled uniformly with respect to ε (bottom row). Our simulation design ensures that λ and r are sampled from identical uniform distributions with respect to ε and ensures comparability of the resulting error estimates.

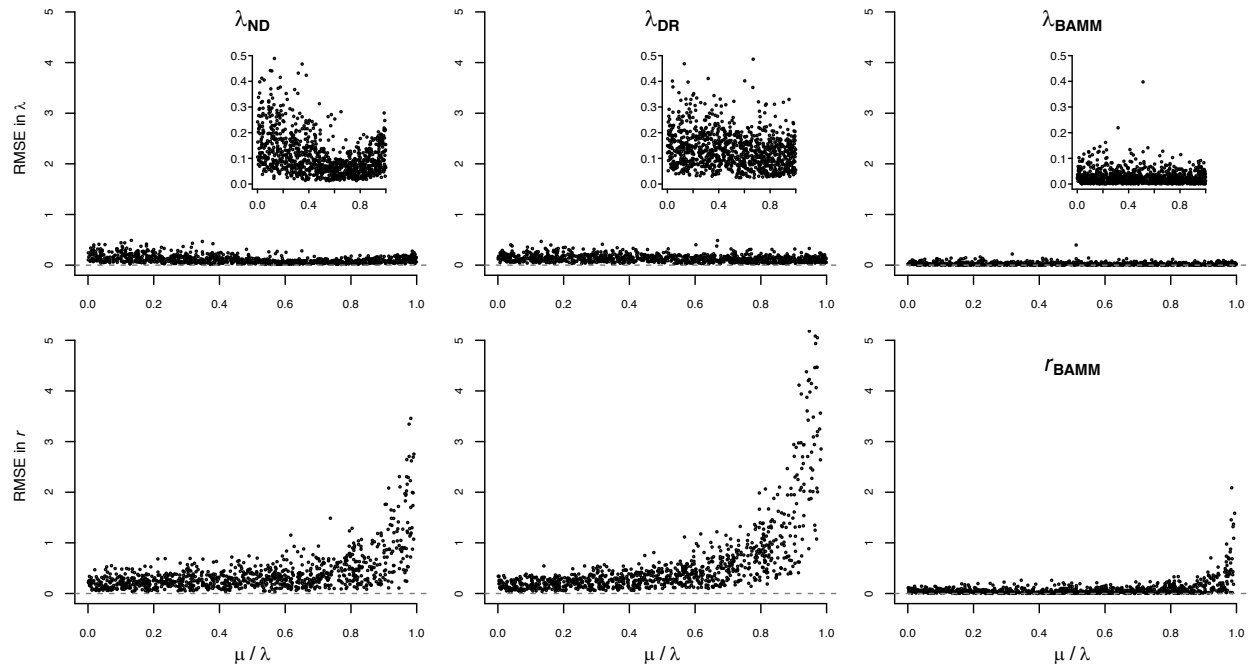


Figure S2. Root-mean-square error (RMSE) in λ (top) and r (bottom) for three different tip rate metrics, across a range of relative extinction rates. For BAMM, the estimated speciation and net diversification rates are presented in the top and bottom panels, respectively. Error of zero implies perfect accuracy. Inset plots show error in λ with truncated y-axis scale to facilitate comparison among metrics. All tip rate metrics track λ more accurately than they track r .

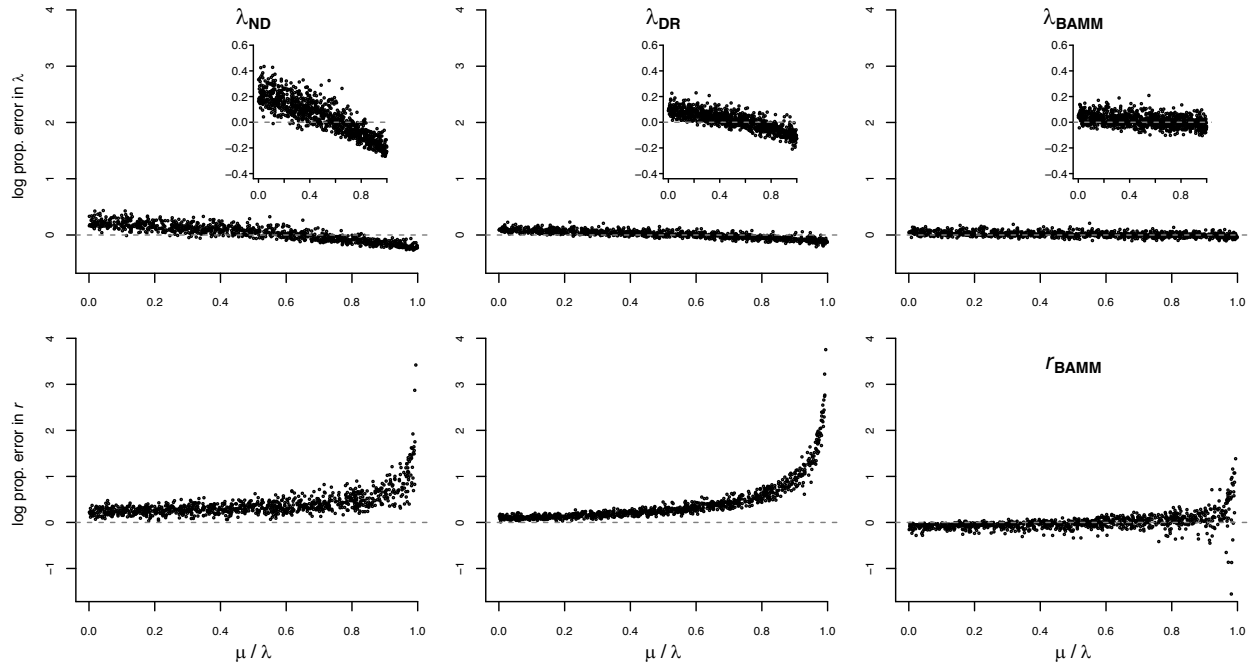


Figure S3. Log proportional accuracy in λ (top) and r (bottom) for different tip rate metrics, across a range of relative extinction rates. For BAMM, the estimated net diversification rate is presented. Proportional error of 0 implies perfect accuracy. Inset plots reveal greater detail in error for λ to ease metric comparison. All tip metrics track λ much more accurately than they track r , and λ_{BAMM} does so with the least amount of error. See Figure S4 for λ_{TB} .

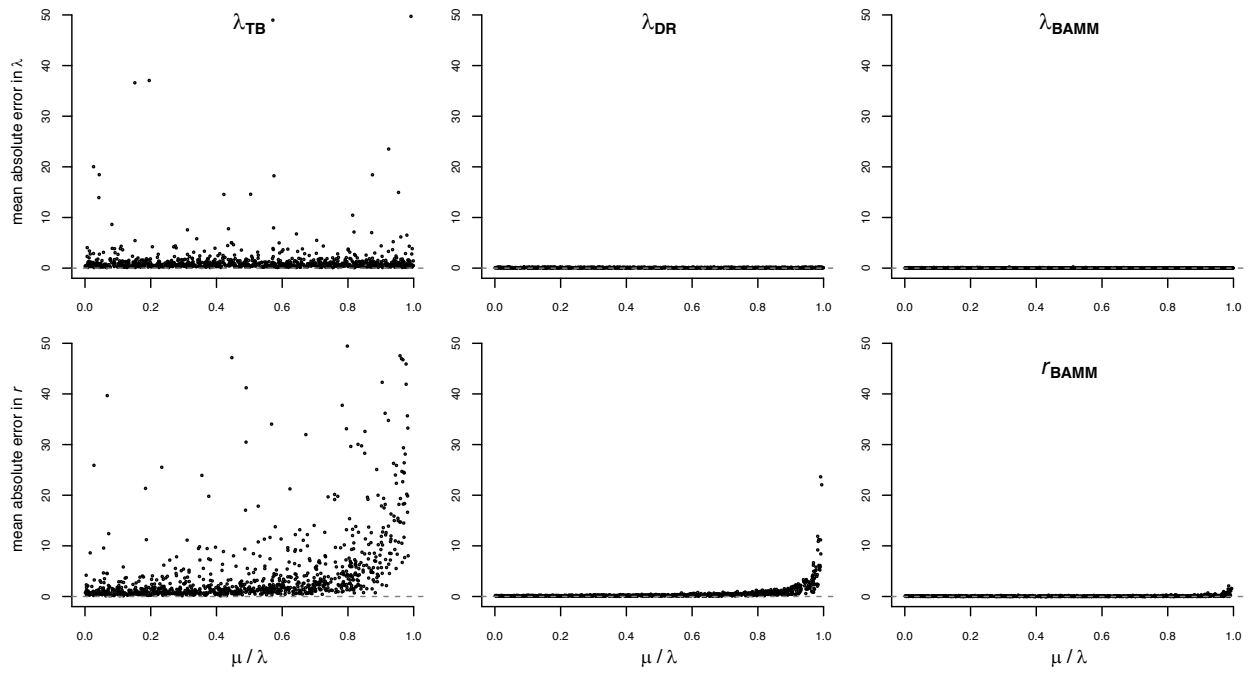


Figure S4. Mean absolute error in λ (top) and r (bottom) for λ_{TB} , with λ_{DR} and λ_{BAMM} on the same scale for comparison. For BMM, the estimated net diversification rate is presented. λ_{TB} more accurately tracks λ than r , but the amount of error is an order of magnitude greater than for other metrics.

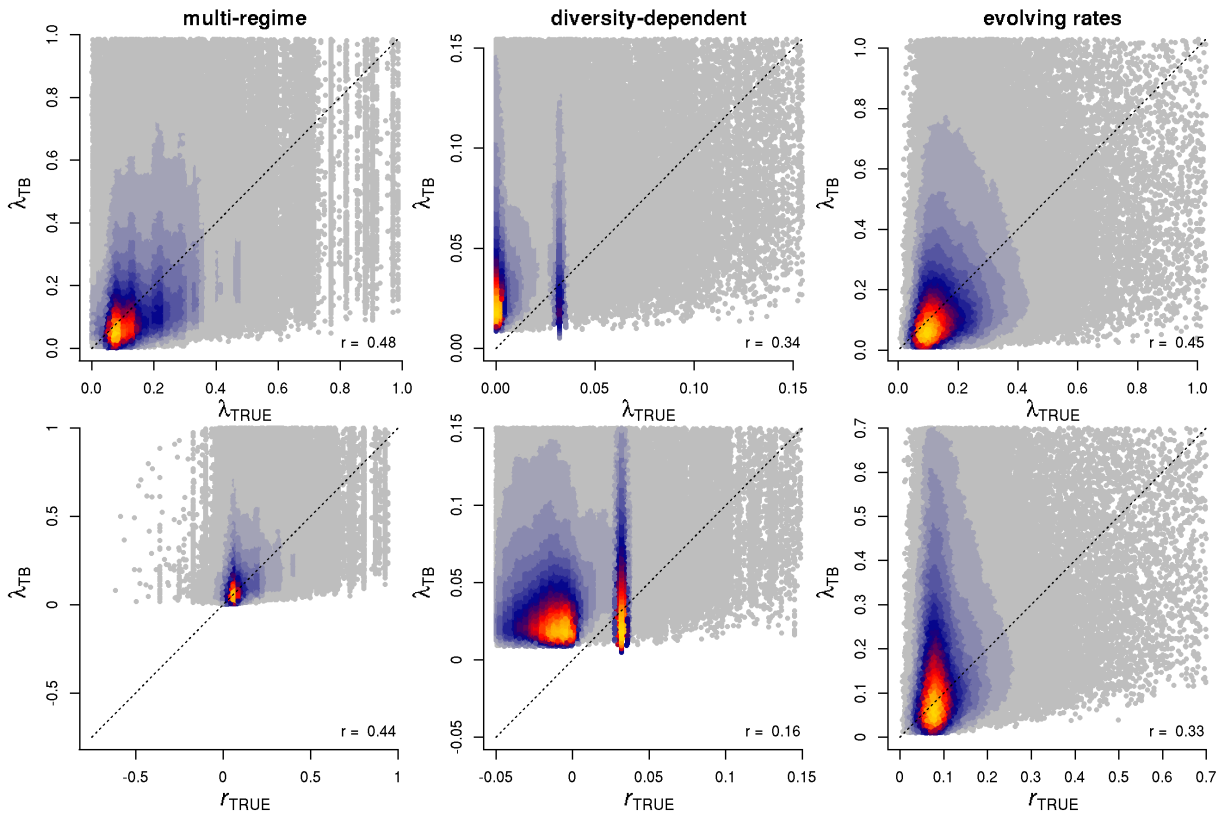


Figure S5. True tip rates (top row: λ_{TRUE} , bottom row: r_{TRUE}) in relation to λ_{TB} . Tip rates were compared separately for different major categories of phylogeny simulations (rows). Plotting region is restricted to the 99th percentile of true rates, but Spearman correlations between true and estimated rates (lower right of each figure panel) are based on the full range of the data. Colors indicate the density of points in the scatter plots. λ_{TB} is a largely unbiased but noisy measure of true speciation rate.

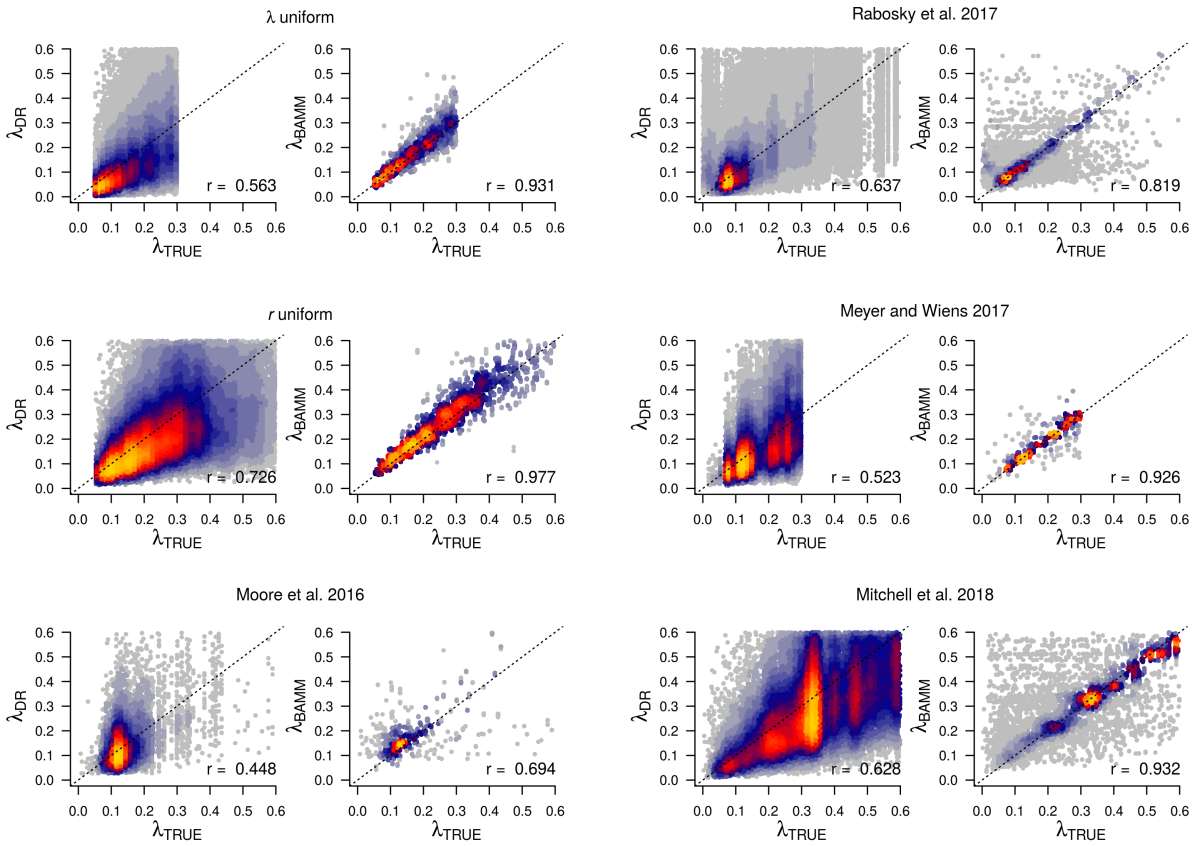


Figure S6. True tip rates (λ_{TRUE}) in relation to estimated tip rates from multi-regime phylogenies, as inferred from the best two tip rate metrics, λ_{DR} and λ_{BAMM} . Data are separated by source, to confirm that patterns described in the main text are not driven by any one simulation study. Spearman's correlation is presented in the bottom right corner. Colors indicate the density of points in the scatter plots. Regardless of the dataset, λ_{BAMM} performs noticeably better than λ_{DR} .

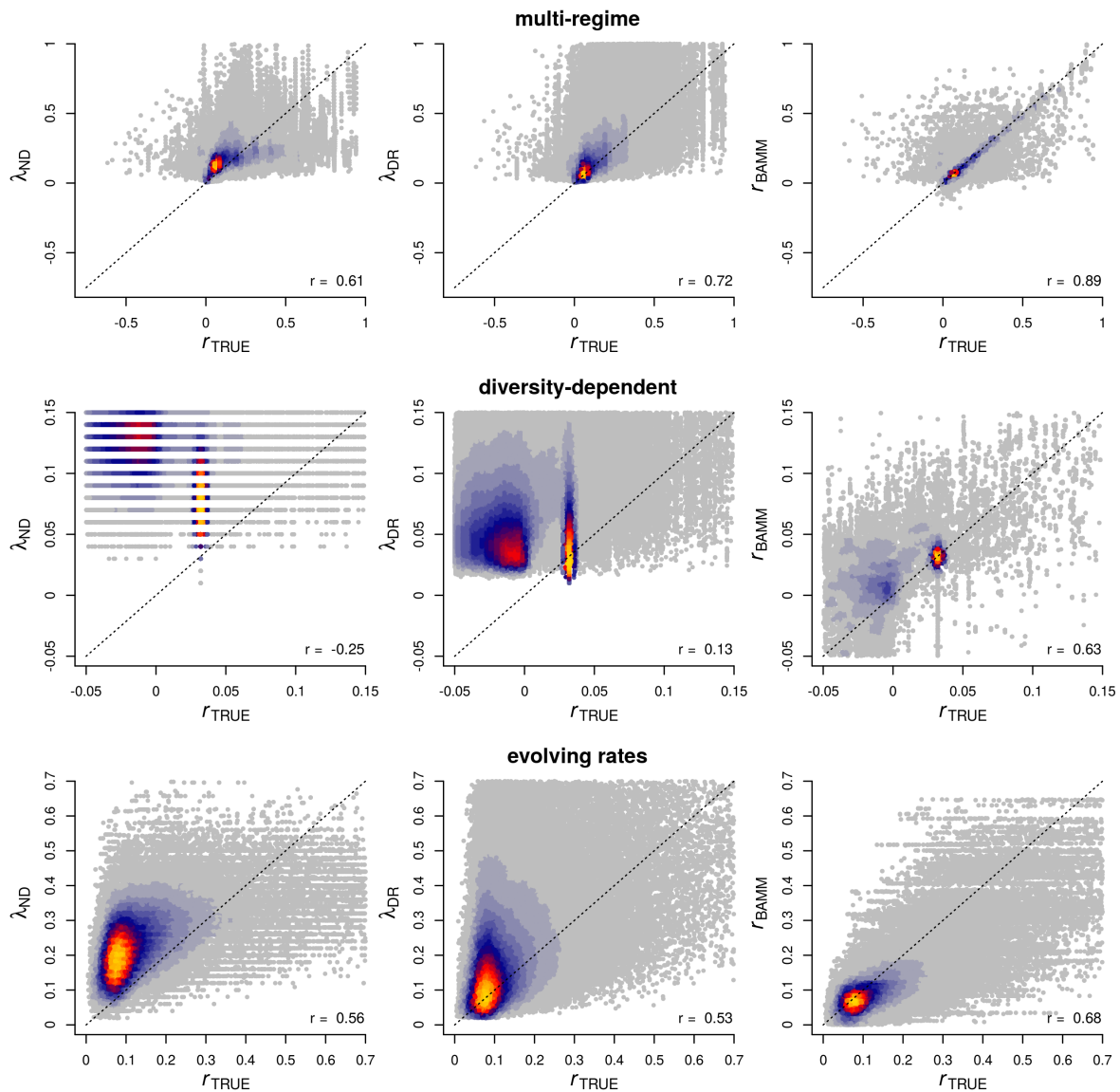


Figure S7. True net diversification tip rates (r_{TRUE}) in relation to estimated tip rates. Tip rates were compared separately for different major categories of phylogeny simulations (rows). Plotting region is restricted to the 99th percentile of true rates, but Spearman correlations between true and estimated rates (lower right of each figure panel) are based on the full range of the data. Colors indicate the density of points in the scatter plots. The horizontal gaps in λ_{ND} for diversity-dependent trees are an artefact of all trees having the same crown age. Relative performance comparison aside, correlations with r_{TRUE} are lower than with λ_{TRUE} (Figure 2).

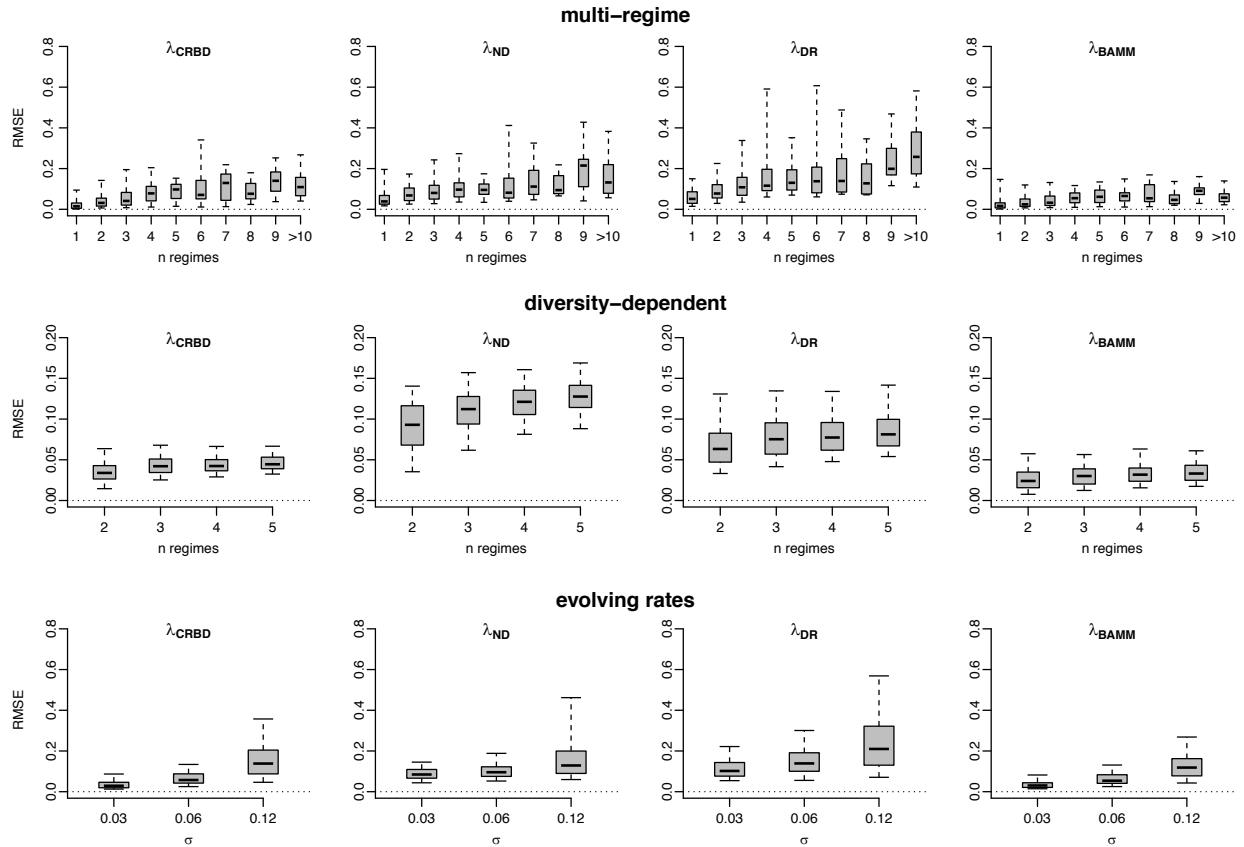


Figure S8. Root-mean-square error (RMSE) in speciation rates as a function of the magnitude of rate heterogeneity in each simulated phylogeny. Results are presented separately for different categories of rate variation (Table 1); left column shows estimates from a constant-rate birth-death model for reference. The boxes and whiskers represent the 0.25 – 0.75, and the 0.05 – 0.95 quantile ranges, respectively. In some cases, λ_{ND} and λ_{DR} had more error than a simple CRBD model with no variation in tip rates. λ_{BAMM} had the least amount of error across all amounts of rate heterogeneity.

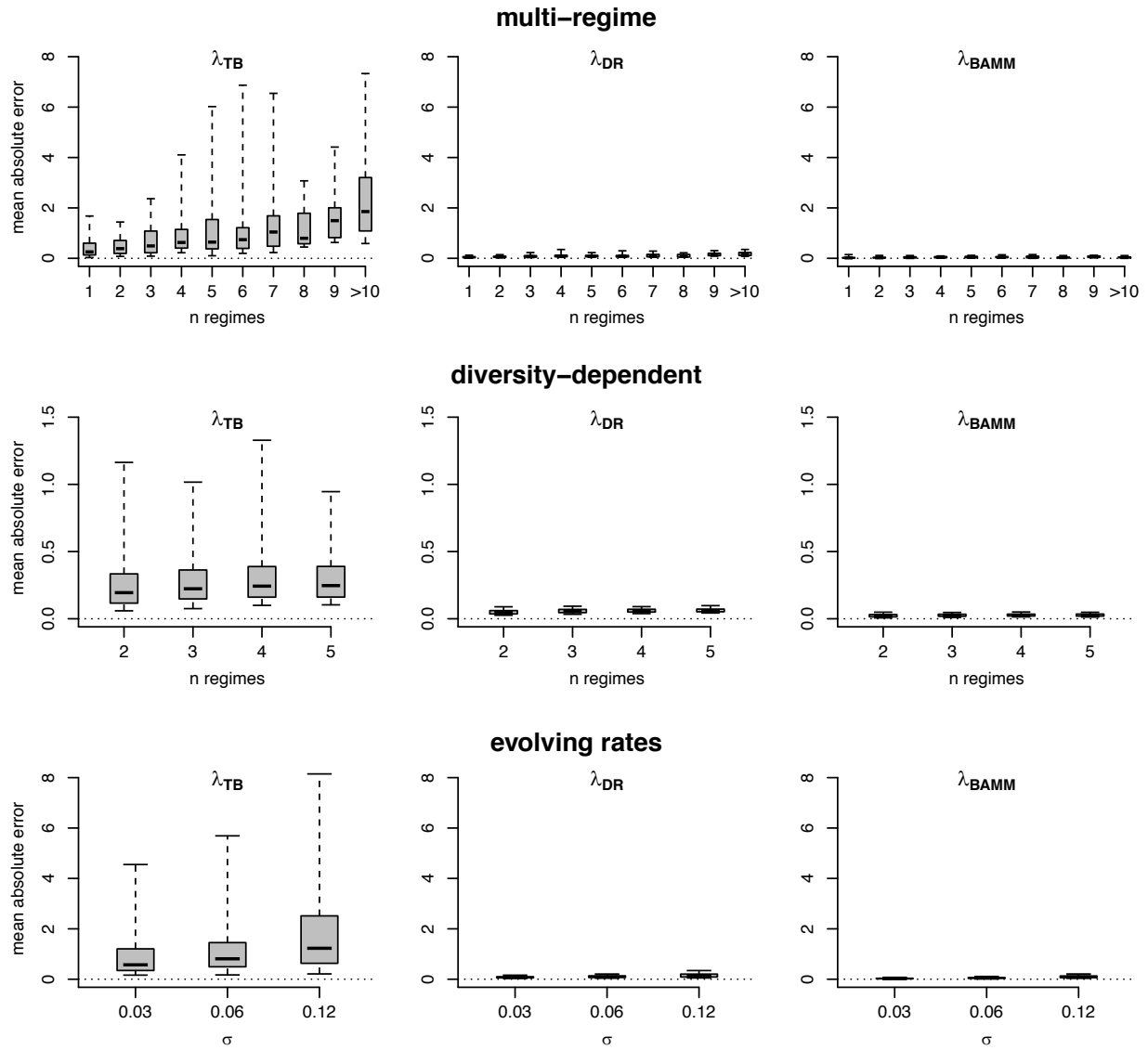


Figure S9. Mean per-tip absolute error in λ_{TB} as a function of the magnitude of rate heterogeneity in each simulated phylogeny. λ_{DR} and λ_{BAMM} are included on the same scale for comparison. Results are presented separately for different categories of rate variation (Table 1). The boxes and whiskers represent the 0.25 – 0.75, and the 0.05 – 0.95 quantile ranges, respectively. Error in λ_{TB} generally increases with increasing rate heterogeneity, and this error is substantially greater than error in other tip rate metrics.

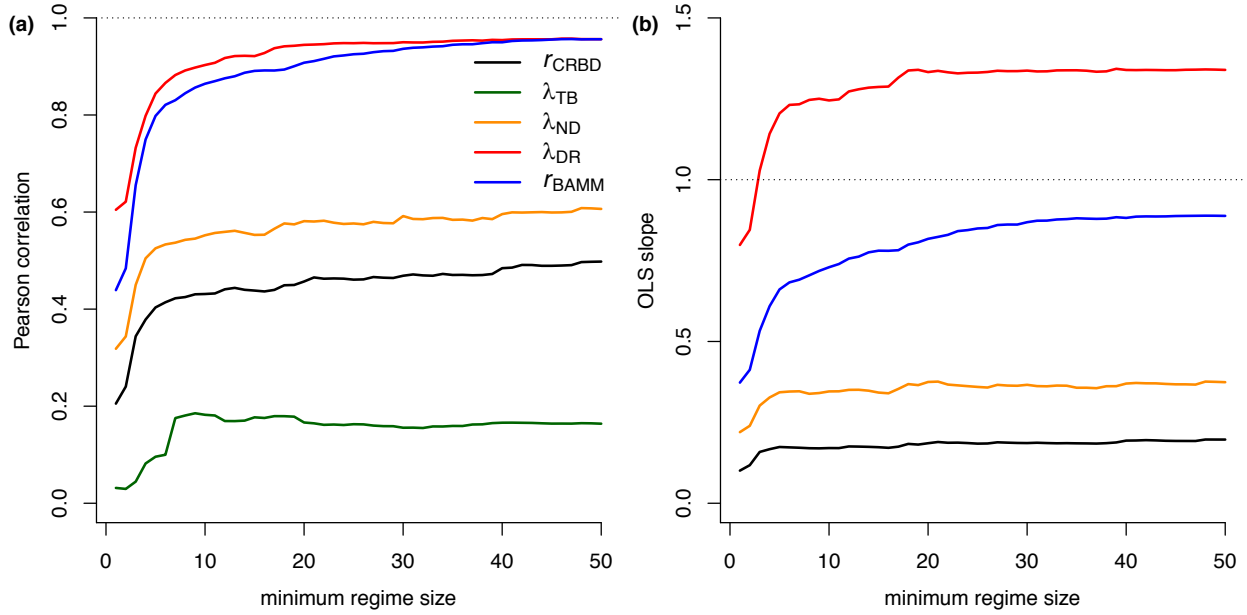


Figure S10. Performance of tip rate metrics as a function of regime size, including Pearson correlation (a) and OLS regression slope (b) for mean rates with respect to r_{TRUE} . λ_{DR} and r_{BAMM} outperform the other metrics when summarized in this fashion, although λ_{DR} overestimates the rate of net diversification (more so than it overestimated λ_{TRUE} , Figure 4). The x-axis denotes the minimum regime size across which performance was summarized. For example, $x = 20$ corresponds to the correlations and slopes computed for all regimes with 20 or more tips; a value of $x = 1$ is the corresponding results for all regimes. The OLS slope for λ_{TB} is not visible as it ranges between 7 and 9.