Section 3: Impact of fossils under various MiSSE models

Under a single rate model of birth and death, the inclusion of fossil data reduces uncertainty, but not necessarily the bias, in the estimates of functions of speciation and extinction, namely, the turnover rate ($\tau = \lambda + \mu$; Beaulieu and O'Meara 2016) and the extinction fraction ($\varepsilon = \mu/\lambda$). We explored these results further by conducting additional single rate regime simulations where trees were generated using a speciation rate of λ = 0.4 and setting the extinction rate to take on three values, μ = 0.1, 0.2, 0.3, to ensure three different values of extinction fraction, ε = 0.25, 0.50, and 0.75, respectively. In each scenario, trees were grown until they reached 200 taxa. We then used a range of fossil sampling (ψ = 0.01, 0.05, and 0.10 fossils/Myr), as well as varying the number of k-type fossils, where either only half were randomly chosen to be included in the analysis (i.e., half k fossils missing), none were included at all (i.e., only m-type fossils included), or converted to stratigraphic ranges. When converting a fossil set into a set of stratigraphic ranges, we simply take the oldest and youngest fossil on each edge, removing all others, prior to pruning unsampled extinct lineages from the full tree.

In all cases, and across three metrics of diversification (turnover, extinction fraction, and net diversification), when comparing extant-only versus a full fossil sample under a given sampling scenario, there was minimal improvement in the rate estimates. Perhaps the most dramatic differences occurred when the extinction fraction was the smallest (i.e., ε = 0.25) and when the fossil sampling rates were highest (ψ = 0.10 fossils/Myr), which resulted in the full fossil set having both the bias and the variance being an order of magnitude smaller than estimates from extant-only. However, when extinction fraction was the highest, at ε = 0.75, the inclusion of fossils had very minor improvements in terms of bias and variance. When a portion of the k-type fossils were removed, there was a notable upward bias, but not the variance, in the rate estimates (see Table S1), which is consistent with the contour plot shown in Figure 2. We also note that when the fossil set is converted to stratigraphic ranges, there is rather severe downward bias in estimates of both turnover rate and extinction fraction as both extinction fraction and increases. Interestingly, bias in the estimates of net diversification remains nominal generally across every combination of generating parameters (see Table S1).

To test the behavior of our MiSSE model when fossils are included, we conducted a set of simulations under scenarios of moderate difficulty. As mentioned above, the underlying model is like BAMM in that it ignores character information altogether. Our likelihood-based, "hidden state only" model assumes that a "shift" in diversification denotes a lineage tracking some unobserved, hidden state. We evaluated the performance of our MiSSE model by simulating trees under three scenarios and then estimating the fit and bias of the inferred rates from these trees. In all cases, we assumed that all the models started with the true trees and that any fossils that were sampled were placed precisely, both in time and on the tree. These assumptions are quite optimistic; our goal was to find the impact of fossils on diversification rate estimates, not assess the impact of fossils on all aspects of tree inference (which would require simulation of morphology and other traits, state-based fossilization rates, tree inference, homology assessment, and more), nor make it difficult for fossils by adding realistic issues of difficulty with placement, taphonomic biases, and more. The first simulation scenario (scenario 1), we introduced two turnover rate regimes, where regime B (τ_B =0.50 events/Myr) had a turnover rate that was 2x the rate of regime A (τ_A=0.25 events/Myr). The extinction fraction was set at 0.75 for both regimes.

In the second scenario (scenario 2), the rates were set such that there were two unique turnover rate and extinction fraction regimes. Specifically, we assumed regime *B* (τ_B=0.50 events/Myr) had a turnover rate that was 2x the rate of regime *A* (τ_A=0.25 events/Myr), but where regime *A* had a lower extinction fraction (ε_A=0.55) than regime *B* (ε_B=0.75). With the turnover rates and extinction fractions structured in this way the net diversification rates between these regimes are nearly identical. Finally, we simulated a scenario (scenario 3) that assumed both regime *A* and *B* had identical turnover rates (τ_A=τ_B=0.375 events/Myr) and extinction fractions (ε_A=ε_B=0.75). This scenario was meant to examine the impact that including fossils has on the support for more complex models when the true model contains homogeneous diversification rates. In all three scenarios we simulated 100 trees and assumed that the transition between regimes *A* and *B* was *q*=0.005 transitions/Myr.

For each simulation replicate, within each scenario, we fit four MiSSE models that broadly capture the complexity of the scenarios. Specifically, we fit a single regime model, a model that only allows turnover rate to vary ($\tau_A \neq \tau_B$), a model that only allowed extinction fraction to vary, and a model that allowed both to vary ($\tau_A \neq \tau_B$ and $\varepsilon_A \neq \varepsilon_B$). Each model was evaluated using a two-step optimization routine. The first step consists of a bounded stochastic simulated annealing run for 5000 iterations, followed by a bounded subplex routine that searches parameter space until the maximum likelihood is found. Since the state space allowed each model to be nested within the most complex model where turnover rate and extinction fraction were allowed to vary, we model-averaged the rate estimates using the AIC weights. However, before doing so, we culled the model set to remove any resulting redundant model fits. If the maximum likelihood estimates for, say, τ_A and τ_B, in the turnover rate varying only model, take on the same value, it is essentially the same as including a single turnover rate model twice, and this would lower the weight of other models as a consequence. It is recommended in these situations to remove the more complex of the two from the set (see Burnham and Anderson 2003).

The main effect of including fossils is to reduce the variance of the estimates across the different rate regimes, calculated as the mean of the squared errors from the mean of the model averaged parameter estimate, $\hat{x}_{model-ave}$, not the bias, calculated as the mean of $\hat{x}_{model-ave}$ - x_{true}. In fact, scenario 2, which assumed differences in both turnover rate and extinction fraction, but not net diversification, proved difficult regardless of whether fossils were included into the analyses. There was a consistent downward bias in the rates associated with shifts to regime *B* even as increased, even when many fossils were included. With regards to converting the fossil set to stratigraphic ranges, there was, again, a general tendency for a downward bias in both turnover rate and extinction fraction. However, when compared to extant-only and a full fossil set assuming the canonical FBD, the estimates of net diversification for stratigraphic ranges generally performed better, albeit with higher variance associated with the estimates compared to the other two analysis types. We do note that with scenario 3, which assumed no rate differences among regimes *A* and *B*, the use of stratigraphic ranges consistently and erroneously inferred higher net diversification rates for regime *A*. Finally, as expected, removing *k* fossils completely from the set results in very high biases related to turnover rates specifically, as increases, and rather severe downward biases generally in estimates of net diversification rates.

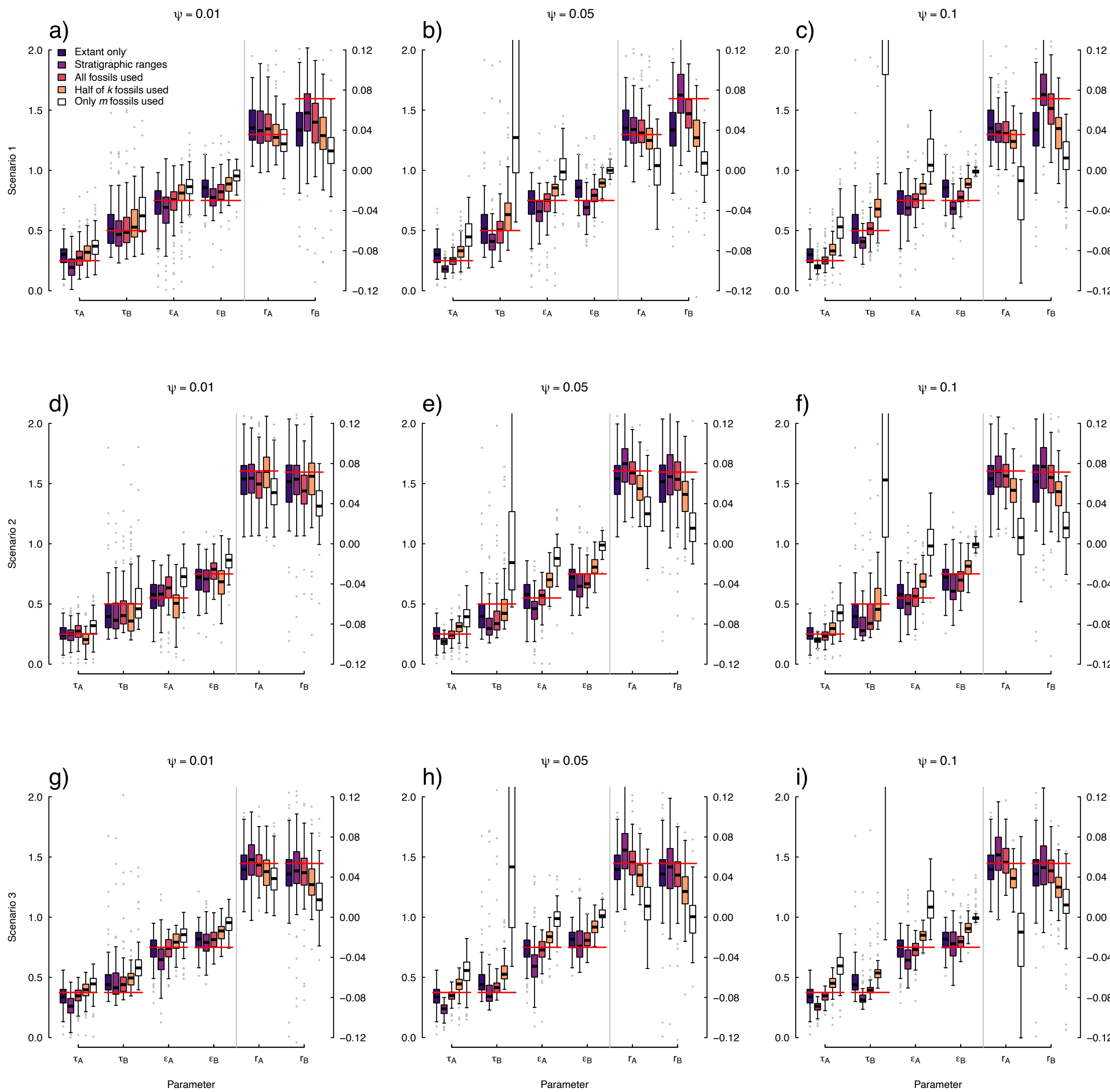

Figure S1. The uncertainty surrounding estimates of turnover rate (τ_A and τ_B), extinction fraction (ε_A and ε_B), and net diversification (r_A and r_B) when the generating model assumes: (a-c) regime *B* had a turnover rate that was 2x the rate of regime *A* (τ_A=0.25, τ_B=0.50) and extinction fraction was set at 0.75 for both regimes; (d-f) there were two unique turnover rate and extinction fraction regimes, specifically, regime *B* had a turnover rate that was 2x the rate of regime *A* (τ_A=0.25, τ_B=0.50), but where regime *A* had a lower extinction fraction than regime *B* (ε_A=0.55, ε_B=0.75); or (g-h) both regime *A* and *B* had identical turnover rates (τ_A=τ_B=0.375) and extinction fractions (ε_A=ε_B=0.75). In all cases, different assumptions about the fossils were fit to the same data set, and parameters were model averaged across a set of four models that variously allowed turnover rate and extinction fraction to vary (see text above). The solid red lines represent the generating value for each parameter in each regime; the right y-axis represents the net diversification scale.