

# Supplementary Figures for: Quantifying Age-dependent Extinction from Species Phylogenies

HELEN K. ALEXANDER<sup>1</sup>, AMAURY LAMBERT<sup>2,3</sup>, AND TANJA STADLER<sup>4</sup>

<sup>1</sup>*Institute for Integrative Biology, ETH Zürich, 8092 Zürich, Switzerland;*

<sup>2</sup>*Laboratoire de Probabilités et Modèles Aléatoires CNRS UMR 7599, UPMC Univ Paris 06,  
Paris, France;*

<sup>3</sup> *Center for Interdisciplinary Research in Biology CNRS UMR 7241, Collège de France, Paris,  
France;*

<sup>4</sup>*Department of Biosystems Science and Engineering, ETH Zürich, 4058 Basel, Switzerland*

**Corresponding author:** Helen K. Alexander, Institute for Integrative Biology, ETH  
Zürich, CHN H.74, Universitätsstrasse 16, 8092 Zürich, Switzerland; E-mail:  
helen.alexander@env.ethz.ch

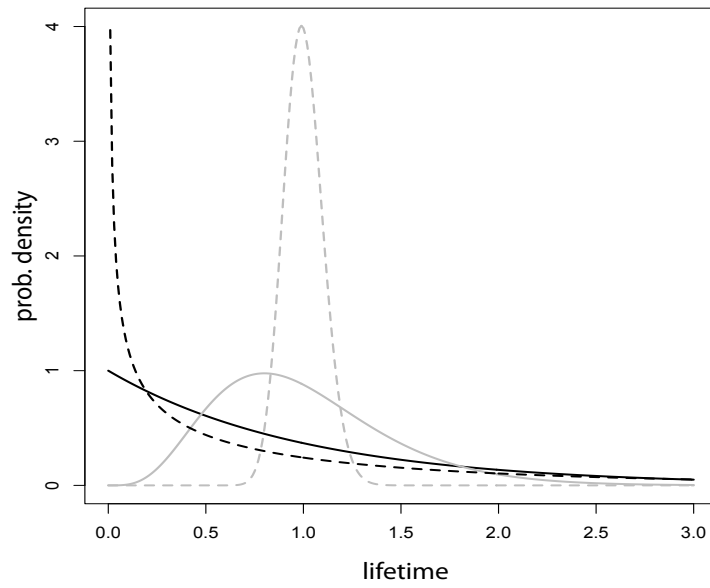


Figure S1: Lifetime distributions used in the simulation study: gamma distributions with mean  $\ell = 1$  and shape parameter  $k = 0.5$  (black dashed), 1 (black solid), 5 (grey solid), or 100 (grey dashed).  $k = 1$  corresponds to exponentially distributed species lifetimes (constant extinction rate),  $k < 1$  corresponds to more variable lifetimes (extinction rate decreases with age), and  $k > 1$  corresponds to less variable, bell-shaped lifetime distributions (extinction rate increases with age).

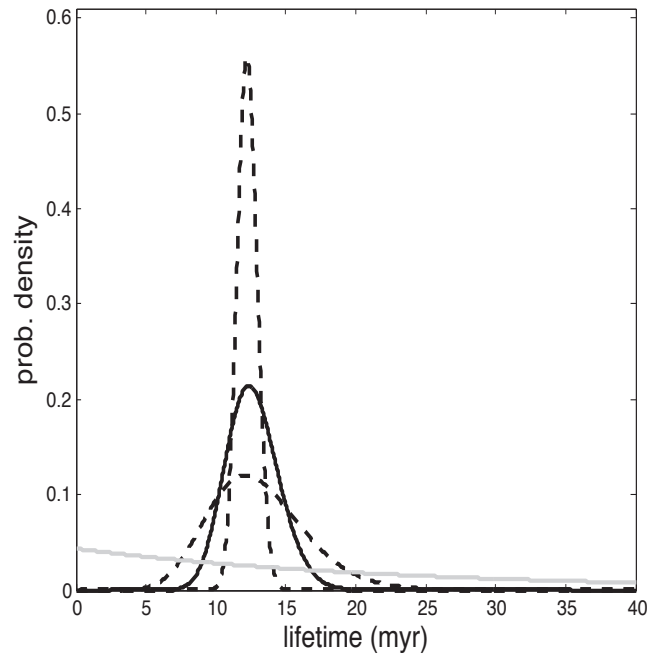


Figure S2: Estimated lifetime distribution, i.e. species age at extinction (in millions of years, myr), for one “typical” reconstructed Aves phylogeny (see main text Methods for description). Black lines: estimates under the gamma model – solid: maximum likelihood estimate (MLE); dashed: limits of 95% profile likelihood confidence interval. Grey line: MLE under the exponential model.

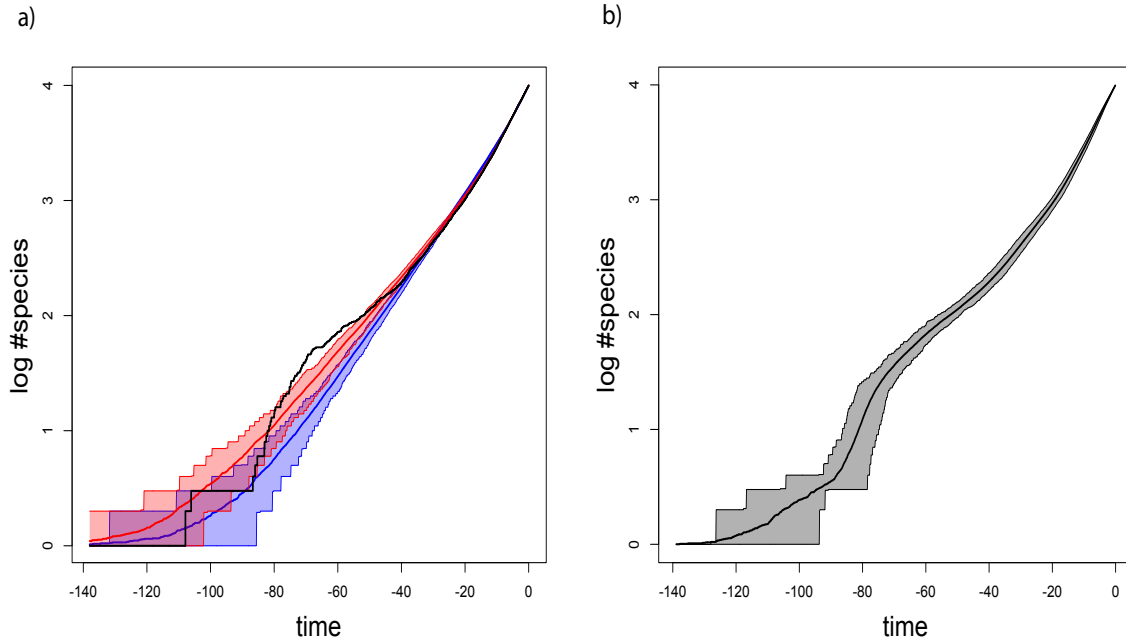


Figure S3: Lineage-through-time plots for complete Aves trees of 9993 species. Panel (a): Black curve: Number of lineages in one “typical” tree (see main text Methods) drawn from the posterior distribution of trees reconstructed using the Hackett backbone by Jetz et al. (*Nature* 491:444-448, 2012). Coloured: Mean (thick curve) and 5-95% quantile range (shaded area) of the number of lineages in 100 bootstrap simulated trees. Simulations were conducted under the maximum likelihood parameters inferred for the “typical” reconstructed tree under the gamma model (red) or the exponential model (blue). Panel (b): mean (thick curve) and 5-95% quantile range (shaded area) of the number of lineages in 100 trees drawn from the posterior distribution of trees reconstructed using the Hackett backbone by Jetz et al.

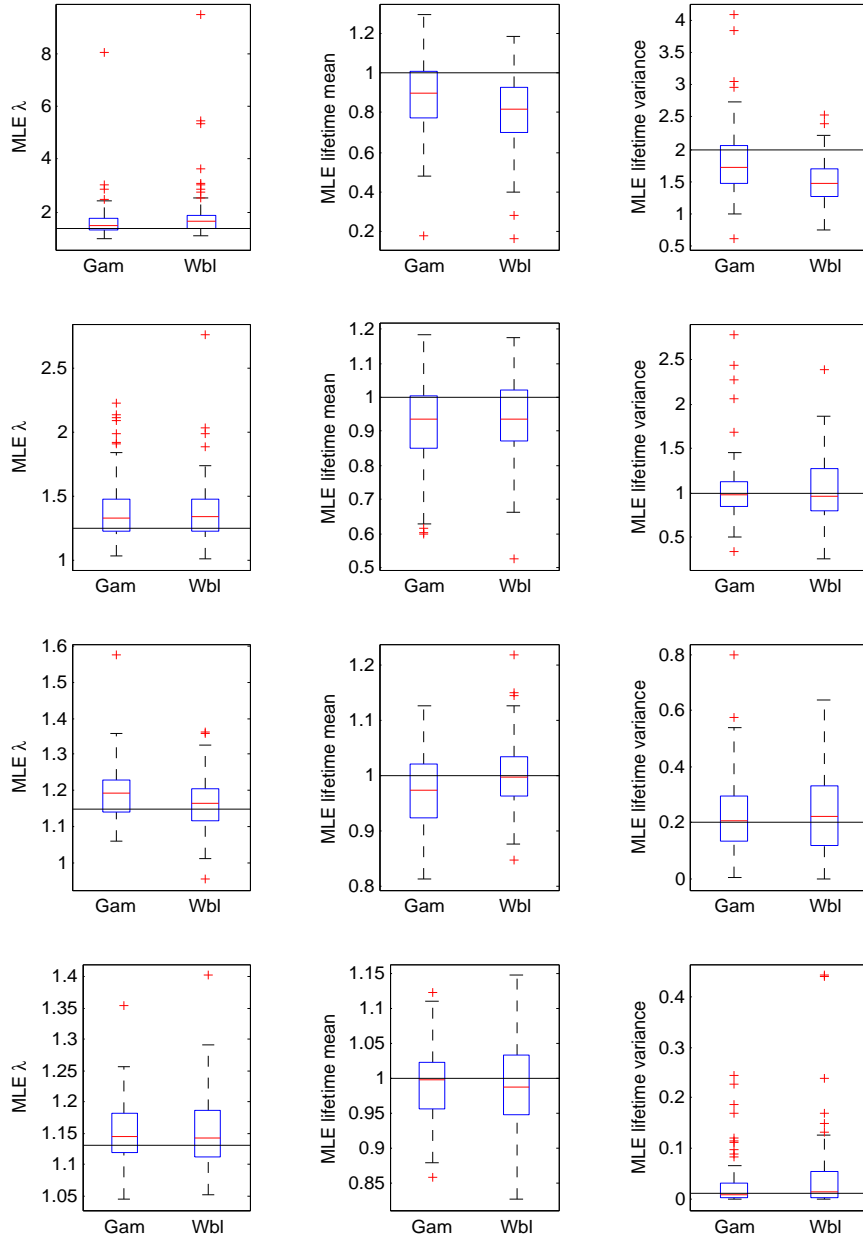


Figure S4: Comparison of maximum likelihood estimates inferred assuming the gamma lifetime distribution model, from sets of 100 trees simulated under either a gamma lifetime distribution (left box in each plot) or a Weibull lifetime distribution (right box in each plot). Each box contains the interquartile range with the median marked as a line. The trees simulated under the gamma distribution are those from the simulation study described in the main text. The parameters of the Weibull distribution used for simulation are chosen to match both the mean (always one) and variance of the gamma distribution in each case. Top row: variance = 2 (shape parameter of the gamma distribution is  $k_G = 0.5$ ); second row: variance = 1 ( $k_G = 1$ ); third row: variance = 0.2 ( $k_G = 5$ ); bottom row: variance = 0.01 ( $k_G = 100$ ). Note that when  $k_G = 1$ , both the gamma and the Weibull distribution are identical, namely collapsing to the exponential distribution; thus, this case illustrates the variability between two sets of simulations conducted under the same model.

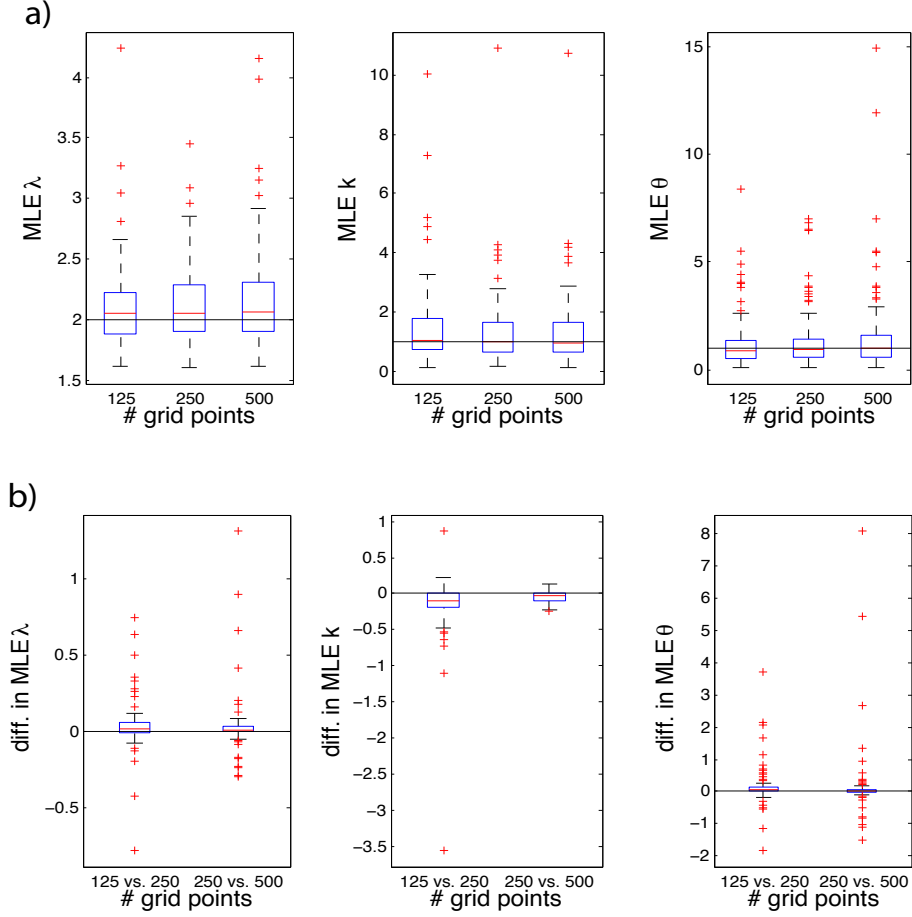


Figure S5: Comparison of parameter estimates in simulated trees, across choices of number of grid points used for numerical evaluation of the likelihood. We use the set of 100 trees from the simulation study with complete sampling,  $n = 1000$  extant species, and true values of  $\eta = 1$  and  $k = 1$ . Panel (a): Distribution of maximum likelihood estimates (MLEs) across trees. The box contains the interquartile range with the median marked as a line. The horizontal line indicates the true parameter value. Panel (b): Distribution of tree-by-tree difference in MLE as number of grid points increases from 125 to 250 or 250 to 500 (i.e. the MLE obtained for a given tree using the larger number of points, minus that obtained using the smaller number of points). The box contains the interquartile range with the median marked as a line. The horizontal line is at zero for reference.

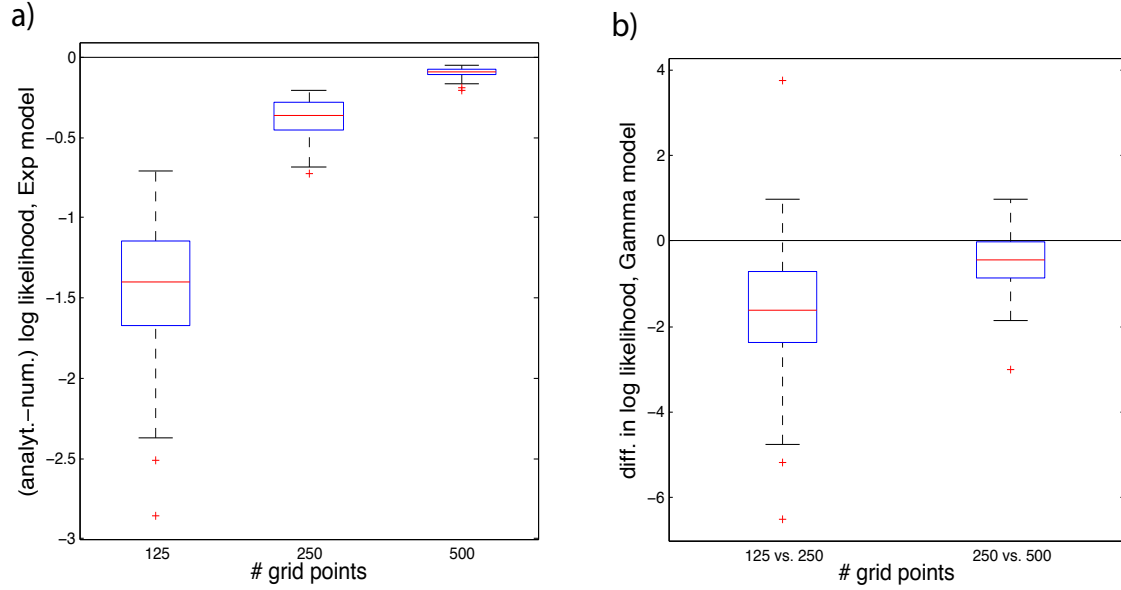


Figure S6: Comparison of maximum log likelihood values, across choices of number of grid points used for numerical evaluation of the likelihood, in the same set of simulated trees used for Figure S5. Panel (a): Distribution across trees of the difference between analytically and numerically evaluated maximum log likelihood under the exponential model, for each given number of grid points. Panel (b): Distribution across trees of the difference between maximum log likelihood evaluated numerically under the gamma model with successive numbers of grid points (i.e. the maximum log likelihood obtained using the larger number of points minus that obtained using the smaller number of points). In both panels, the box contains the interquartile range with the median marked as a line, and a horizontal line is at zero for reference.

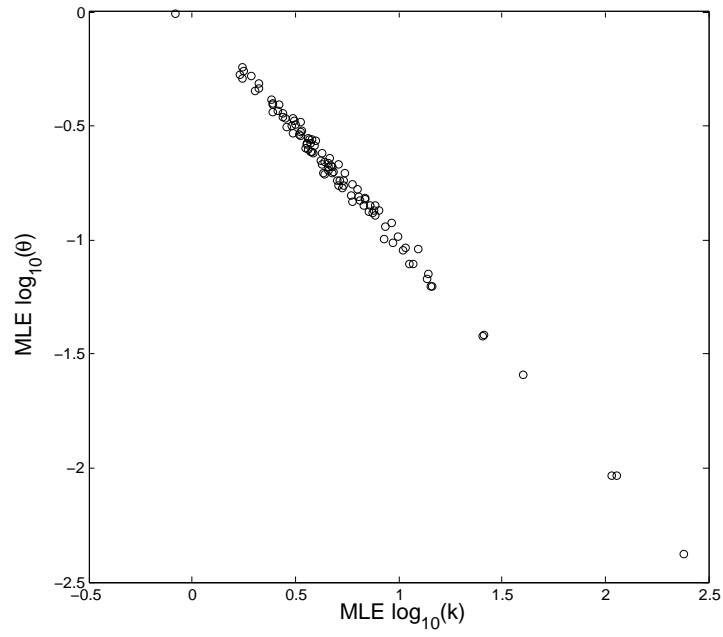


Figure S7: Correlation in estimated lifetime distribution parameters  $k$  (shape) and  $\theta$  (scale), illustrated by a scatter plot of maximum likelihood estimates of  $k$  and  $\theta$  (on base-10 log scales) inferred under the gamma lifetime distribution model in each tree. The set of 100 trees is from the simulation study with complete sampling,  $n = 1000$  extant species, and true parameter values of  $\eta = 0.25$  and  $k = 5$ .