Supporting Information

Huang et al. 10.1073/pnas.1409886112

SI Materials and Methods

Simulations. Phylogenies were simulated under a constant-rates birth–death process (1) using the program SimTreeSDD (2). The program was updated to return not only the resulting phylogeny of extant species but also the time of speciation for each species. This new version of SimTreeSDD is now available (3).

For each of the simulated trees, we quantified the change in EH for four extinction intensities (m) corresponding to the loss of 10, 30, 50, or 80% of the tips on the tree. Trees were subjected to five extinction regimes—old bias based on absolute or relative age, young bias based on absolute or relative age, or random with respect to age. For each extinction regime and rate, 100 replicates were performed. Before and after each extinction event, we computed the PD of each tree, given as the sum of all branch lengths of the phylogeny (4), and the corresponding SD, given as the sum of absolute ages of species on the tree. We then compared each of the old- or young-based extinction regimes against a null model represented by the random regime (5–8) using PDexcess and SDexcess (9).

To explore the effect of age on species survivorship, we fit a binomial logistic regression model with either absolute or relative age as the explanatory variable and survivorship during the simulated extinction event as the binary response variable. The regression coefficient is the natural logarithm (log-odds) of the odds ratio p/(1−p), where p is the probability of surviving through the event (10). We fit the model separately for each of the five extinction regimes and four extinction intensities described above.

We also repeated our analyses on trees generated by the R (11) package, paleotree (12). We simulated morphotaxa and their relationships in the fossil record under the budding cladogenesis model (13), based on the same speciation and extinction parameters and minimum number of extant taxa outlined above. Similarly, we generated 1,000 sets of absolute ages each for the high and low number of extinction intensities used in the previous analyses. These sets then were converted into time-scaled phylogenies for analyses. SimTreeSDD trees are used in the main text (Figs. 2 and 3), and results using paleotree trees are presented here as Figs. S2 and S3.

Plio–Pleistocene Extinctions in California. To reconstruct the phylogeny of extant and extinct species of California Pectinidae, we first assembled from GenBank a molecular dataset comprising six markers—the mitochondrial 12S rRNA, 16S rRNA, and cytochrome c oxidase subunit I, as well as nuclear 18S rRNA, 28S rRNA, and histone H3—covering 76 living pectinid species worldwide and three outgroup species from the Propeamussiidae (14). Extinct species were added to the dataset as topological constraints based on their genus identities. More specific constraints were implemented based on refs. 15–19.

Fossil node calibrations were carried out based on the first appearance of pectinid genera obtained from an existing global marine bivalve database (20, 21). For nodes for which fossil data were available, we used as prior a truncated normal distribution with the mean set to the earliest date (stage midpoint) among sister clades and an SD of 2.5 Ma. Calibration was carried out at the genus level and above for taxa with molecular data, so a species’ time of bifurcation with its sister on the phylogeny can occur after its first appearance in the fossil record. The root of the tree, representing the origin of Pectinidae was designated a truncated normal prior with a mean of 251 Mya and SD of 5 Ma. Based on stratigraphic ranges of pectinid species compiled from primary literature (18, 22–26), the origin of each extinct species was assigned a truncated normal prior with a mean based on first appearance and an SD of 1 Ma, with its termination represented by noncontemporaneous tip dates based on the species’ last appearance and a uniform distribution of ± 0.25 Ma (27).

Based on the molecular data, topological constraints, and fossil calibrations, we used an uncorrelated lognormal relaxed clock implemented in BEAST 1.8 (28–30) to infer the pectinid phylogeny comprising 79 living (global) and 75 extinct (California) species. Four separate Markov chain Monte Carlo runs of 60 million generations were carried out with a sampling interval of 1,000. The first 10 million of all posterior trees were discarded upon convergence that was checked in Tracer 1.5 (31), and the remaining trees were subsampled to every 200,000 iterations to generate 1,000 fully resolved trees.

Species not present in California as well as those that were extinct before the Plio–Pleistocene extinction event 2.5 Mya were trimmed from the trees, resulting in a 50-species phylogeny with 25 of the tips going extinct at this time. Because we were analyzing an extinction that happened around 2.5 Mya (Plio–Pleistocene), we trimmed back to this date the branch lengths of living species that survived the extinction event. We also carried out a separate reconstruction based on the 14 living and 75 extinct Californian species, excluding at the onset species not present in California, to determine if our results were sensitive to the amount of data used to reconstruct the phylogeny. All subsequent calculations were performed for each of the 1,000 posterior trees, separately for the global species and Californian species calibrations.

Following the procedure for the simulated trees, we computed PDexcess and performed the binomial logistic regression to obtain the log-odds of survivorship with respect to relative age for this particular extinction event. We also computed log-odds for the null model with the same extinction rate.


Fig. S1. Proportion of original EH lost or remaining after simulated species extinction independent of absolute or relative age. Shown are outcomes of PD based on bifurcating phylogenies as well as SD for extinction intensities (m) between 10% and 80%. Trees were simulated using SimTreeSDD under $\epsilon = \mu/\lambda = 0.8$. Patterns are similar for trees simulated under $\epsilon = 0.2$ and with paleotree (12).
Fig. S2. PD$_{\text{excess}}$ and SD$_{\text{excess}}$. (A–D) Figure components are the same as in Fig. 2. Trees used here were simulated with paleotree (12), whereas trees used in Fig. 2 were simulated with SimTreeSDD (3).

Fig. S3. Taxon age selectivity. (A–J) Figure components are the same as in Fig. 3. Trees used here were simulated with paleotree (12), whereas trees used in Fig. 3 were simulated with SimTreeSDD (3).
Fig. S4. \( PD_{excess} \) of California scallop (Pectinidae) species during the Plio–Pleistocene extinction event (~2.5 Mya) in which 25 of the 50 species present at that time went extinct. Results are similar for each posterior set of 1,000 trees calibrated with a global molecular dataset of 79 living and 75 extinct species (Upper) and a California dataset of 14 living and 75 extinct species (Lower).
Fig. 55. Age selectivity log-odds of survival during the Plio–Pleistocene extinction event (~2.5 Mya) in which 25 of the 50 species present at that time went extinct. Results are similar for each posterior set of 1,000 trees calibrated with a global molecular dataset of 79 living and 75 extinct species (Left) and a California dataset (Right) of 14 living and 75 extinct species (Right). None of the log-odds computed for trees of both calibrations show statistical significance based on a binomial logistic regression model for the actual event (red) and random, age-independent species loss (cyan).