Reconciling molecular phylogenies with the fossil record

Hélène Morlon[a,b,1], Todd L. Parsons[b], and Joshua B. Plotkin[b]

[a]Center for Applied Mathematics, Ecole Polytechnique, 91128 Palaiseau, France; and [b]Biology Department, University of Pennsylvania, Philadelphia, PA 19104

Edited* by Robert E. Ricklefs, University of Missouri, St. Louis, MO, and approved August 1, 2011 (received for review February 14, 2011)

Historical patterns of species diversity inferred from phylogenies typically contradict the direct evidence found in the fossil record. According to the fossil record, species frequently go extinct, and many clades experience periods of dramatic diversity loss. However, most analyses of molecular phylogenies fail to identify any periods of declining diversity, and they typically infer low levels of extinction. This striking inconsistency between phylogenies and fossils limits our understanding of macroevolution, and it undermines our confidence in phylogenetic inference. Here, we show that realistic extinction rates and diversity trajectories can be inferred from molecular phylogenies. To make this inference, we derive an analytic expression for the likelihood of a phylogeny that accommodates scenarios of declining diversity, time-variable rates, and incomplete sampling; we show that this likelihood expression reliably detects periods of diversity loss using simulation. We then study the cetaceans (whales, dolphins, and porpoises), a group for which standard phylogenetic inferences are strikingly inconsistent with fossil data. When the cetacean phylogeny is considered as a whole, recently radiating clades, such as the Balaneopteridae, Delphinidae, Phocoenidae, and Ziphiidae, mask the signal of extinctions. However, when isolating these groups, we infer diversity dynamics that are consistent with the fossil record. These results reconcile molecular phylogenies with fossil data, and they suggest that most extant cetaceans arose from four recent radiations, with a few additional species arising from clades that have been in decline over the last ∼10 Myr.

Inferring rates of speciation and extinction and the resulting pattern of diversity over geological time scales is one of the most fundamental but challenging questions in biodiversity studies (1–4). Traditionally, biologists have relied on the fossil record to study long-term diversity dynamics (4–7). However, many groups, including terrestrial insects, birds, and plants, lack an adequate fossil record. Methods have therefore been developed to estimate speciation and extinction rates and test hypotheses about the mechanisms governing diversification using phylogenies of extant species reconstructed from molecular data (1, 8–14). These methods have raised the possibility of inferring diversity dynamics for groups that lack a detailed fossil record.

Given the large number of taxonomic groups that lack fossil data, approaches that rely on extant taxa alone are critically important. However, recent studies have highlighted major inconsistencies between the diversity dynamics inferred from phylogenies and those dynamics inferred from the fossil record (4, 11, 15). For example, the fossil record clearly shows that the diversity of cetaceans has declined over the last 10 Myr (4, 16), whereas two recent phylogeny-based maximum likelihood analyses of this group would suggest that cetacean diversity has been expanding (17, 18). More generally, phylogeny-based maximum likelihood estimates of extinction rates are often close to zero, which is not realistic given that extinctions do, in fact, occur and can be frequent in many groups (1, 9, 11, 15, 19).

The current inconsistency between phylogenies and fossils is puzzling, and it casts serious doubt on phylogenetic techniques for inferring the history of species diversity (4). These concerns are especially problematic for groups that lack sufficient fossil data. There are several possible reasons for this inconsistency. On the one hand, because phylogenies of extant taxa lack direct information about extinct lineages, they simply may lack sufficient information to accurately estimate extinction rates or infer diversity dynamics (15, 20, 21). If this is the case, there is little hope for us to ever understand the history of diversification in groups or places lacking fossil data. On the other hand, there is also a possibility that the apparent inconsistency between phylogenies and the fossil record is a methodological artifact, which could be overcome if we develop the appropriate tools.

There is no doubt that the information provided by a reconstructed phylogeny is limited. It is well-recognized that alternative diversification scenarios can produce phylogenies with similar shapes, such that phylogenies may have little discriminatory power (4, 13, 14, 22, 23). It is, thus, understandable that phylogenetic inferences based on a single summary statistic, such as the widely used γ-statistic describing the temporal distribution of nodes in a phylogeny (24–27), fail to properly infer past diversity dynamics (4, 20, 21). It is more worrisome, however, that likelihood-based inferences, which use most of the information contained in phylogenies, also yield unrealistically low extinction rate estimates (9, 11, 19). However, this inconsistency likely arises from limitations in the current methods of phylogenetic inference, which typically assume that speciation rates are constant through time or across lineages and that speciation rates exceed extinction rates (1, 4).

Here, we begin by deriving an exact analytic expression for the likelihood of observing a given phylogeny that simultaneously accommodates undersampling of extant taxa, rate variation over time, and potential periods of declining diversity. Our derivation is based on the birth–death framework introduced in the work by Nee et al. (8) and further developed by Rabosky and Lovette (11) and by Maddison et al. (10, 28). Using Monte Carlo simulations, we quantify the ability of our likelihood-based inferences to detect clades in decline. We then apply this phylogenetic method to the cetaceans, and we compare the diversity dynamics inferred from their phylogeny with the dynamics inferred from the fossil record.

Results

We developed a method to infer diversity dynamics from phylogenies using a birth–death model of cladogenesis (1, 8, 9, 11). We assume that a clade has evolved according to a birth–death process, with per-lineage speciation and extinction rates, $\lambda(t)$ and $\mu(t)$, respectively, that can vary over time. We consider the phylogeny of n species sampled at present from this clade. We allow for the possibility that some extant species are not included in the sample by assuming that each extant species was sampled with probability $f \leq 1$. We measure time from the present to the past. Thus, $t = 0$ denotes the present, and t increases into the past. t_1 denotes the first time at which the ancestral species came into existence, and $\{t_2, t_3, \ldots, t_n\}$ denote the times of branching

Author contributions: H.M. and J.B.P. designed research; H.M., T.L.P., and J.B.P. performed research; H.M. and T.L.P. contributed new reagents/analytic tools; H.M. analyzed data; and H.M., T.L.P., and J.B.P. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

See Commentary on page 16145.

[1]To whom correspondence should be addressed. E-mail: helene.morlon@polytechnique.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1102543108/-/DCSupplemental.

events in the phylogeny, with $t_1 > t_2 > \ldots > t_n$. In particular, t_2 is the time of the most recent common ancestor of the sampled species (Fig. 1 has a schematic illustration of notations).

The probability density of observing such a phylogeny, conditioned on the presence of at least one descendant in the sample, is proportional to (*Materials and Methods* has details)

$$\mathcal{L}(t_1,\ldots,t_n) = \frac{f^n \Psi(t_2,t_1)\prod_{i=2}^{n}\lambda(t_i)\Psi(s_{i,1},t_i)\Psi(s_{i,2},t_i)}{1-\Phi(t_1)}, \quad [1]$$

where $\Psi(s, t)$ denotes the probability that a lineage alive at time t leaves exactly one descendant lineage at time $s < t$ in the reconstructed phylogeny, and $\Phi(t)$ denotes the probability that a lineage alive at time t has no descendant in the sample. $s_{i,1}$ and $s_{i,2}$ denote the times at which the daughter lineages introduced at time t_i themselves branch (or zero if the daughter lineage survives to the present without branching) (Fig. 1).

Adapting the approaches by Maddison et al. (10) and FitzJohn et al. (28), we derived the following exact likelihood expressions:

$$\Phi(t) = 1 - \frac{e^{\int_0^t \lambda(u)-\mu(u)du}}{\frac{1}{f} + \int_0^t e^{\int_0^s \lambda(u)-\mu(u)du}\lambda(s)ds} \quad [2]$$

and

$$\Psi(s,t) = e^{\int_s^t \lambda(u)-\mu(u)du}\left[1 + \frac{\int_s^t e^{\int_0^\tau \lambda(\sigma)-\mu(\sigma)d\sigma}\lambda(\tau)d\tau}{\frac{1}{f} + \int_0^s e^{\int_0^\tau \lambda(\sigma)-\mu(\sigma)d\sigma}\lambda(\tau)d\tau}\right]^{-2}. \quad [3]$$

Substituting Eqs. **2** and **3** into Eq. **1** gives the likelihood of the phylogeny in terms of the sampling fraction f and the speciation and extinction rates $\lambda(t)$ and $\mu(t)$, respectively. The net diversification rate at any given time, $\lambda(t) - \mu(t)$, can be positive, corresponding to a period of expanding diversity, or negative, corresponding to a period of declining diversity. The likelihood expression given by Eq. **1** is directly comparable with those expressions derived in the seminal work by Nee et al. (8) and in more recent works by Rabosky and Lovette (11), Maddison et al. (10), and FitzJohn et al. (28). When phylogenies are fully sampled and diversification rates are assumed constant over time and independent of a specific character, these likelihood expressions differ from one another only by the conditioning on the birth–death process (*Materials and Methods* and *SI Results*). Eq. **1** is, however, an exact analytical expression that simultaneously accounts for rate variation through time and undersampling, although FitzJohn (29) and Rabosky and Glor (30) proposed numerical procedures associated with this scenario, and Stadler (31) derived the likelihood analytically in the case of discrete rate shifts.

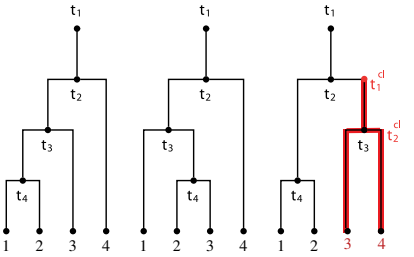

Fig. 1. Schematic figure illustrating the notations. We characterize a phylogeny by its branch intervals, which are denoted $\{(t_2, t_1), \{(s_{i,1}, t_i), (s_{i,2}, t_i)\}_{i=2,\ldots,n}\}$. The figure illustrates three example phylogenies that all have the same branching times, which are indicated as labels. The leftmost and middle trees have the same topology, whereas the rightmost tree has a distinct topology. *Left* and *Center* have corresponding tree branch intervals of $\{(t_2, t_1), (t_3, t_2), (0, t_2), (t_4, t_3), (0, t_3), (0, t_4), (0, t_4)\}$, whereas *Right* has branch intervals of $\{(t_2, t_1), (t_3, t_2), (t_4, t_2), (0, t_3), (0, t_3), (0, t_4), (0, t_4)\}$. The use of branch intervals, thus, encodes the topology of the tree up to the labeling of nodes. Note that, although the topology of the tree in *Right* is distinct from the topology of the trees in *Left* and *Center*, their likelihoods are identical when diversification rates are assumed homogeneous across lineages. When we instead assume that diversification rates vary at a known set of branching points (for example, at t_2 on the tree in *Right*), different topologies are no longer equally probable. In this case, the likelihood of the tree may be computed as a product of the likelihood of the subclade (red) given by Eq. **1** and the likelihood of the rest of the tree.

If we no longer assume that diversification rates are constant across lineages but, instead, assume that they change at a fixed set of branching points, Eq. **1** can readily be modified to take into account this rate heterogeneity. Under the assumption that rates vary only at a fixed set of observed branching points, our expression for $\Phi(t)$ and $\Psi(s, t)$ holds (*SI Results*). If a clade first appears at the branching point t_1^{cl} (Fig. 1 clarifies the notation), the likelihood function of the subtree corresponding to this clade is given by Eq. **1** with the clade-specific rates of speciation and extinction. The likelihood function of the remaining pruned parent tree (i.e., the whole tree minus the clade) is also given by Eq. **1** if we replace all terms corresponding to the subclade, such as

$$f^{n_{cl}}\Psi(t_1^{cl},t_2^{cl})\prod_{j=2}^{n_{cl}}\lambda(t_i^{cl})\Psi(s_{i,1}^{cl},t_i^{cl})\Psi(s_{i,2}^{cl},t_i^{cl}), \quad [4]$$

with the probability $1-\Phi(t_1^{cl})$ that the subclade did not go extinct (*SI Results*). The likelihood of the whole tree is then the product of the likelihood of the subtree and the pruned tree. More generally, the likelihood of a phylogeny in which various subclades have different diversification rates is obtained by multiplying together the likelihoods of each subclade and the likelihood of the remaining pruned parent tree. Hence, given a phylogeny, Eq. **1** can be used to estimate rates in multiple subclades as well as compare the performance of various parametric models for how these rates vary over time and across clades.

Applying the likelihood expression in Eq. **1** to phylogenies simulated with time-variable speciation and extinction rates, we found that diversity dynamics can be accurately inferred across a wide range of parameter values, including scenarios that feature periods of declining diversity (*Materials and Methods*, *SI Results*, and Figs. S1–S3). For example, one such simulated parameter set featured a constant extinction rate ($\mu_0 = 0.5$ events per arbitrary time unit) and a speciation rate that decayed exponentially over time (from $\lambda = 3$ at 10 time units in the past to $\lambda_0 = 0.25$ at present), and therefore, the net diversification rate switched from positive to negative over the clade's history. Even in this scenario of diversity expansion followed by diversity collapse, which is similar to the scenario of waxing and waning observed in the fossil record, our method produced unbiased parameter estimates (Figs. S1, rightmost data point, and S2), and we correctly inferred a negative diversification rate at present for 70% of such simulated phylogenies. This percentage increases to 81% and 92% when we consider only the subset of phylogenies with at least 10 or 20 species at present, respectively. We found similar results when the speciation rate was held constant and the extinction rate increased over time (Fig. S3).

Although our method produces unbiased parameter estimates, the confidence intervals around these estimates can be broad, especially when a phylogeny is small (*Materials and Methods* and Figs. S4 and S5). Figs. S4 and S5 show how tree size influences the confidence interval for estimates of the net diversification rate at present for an example set of parameters. These figures reflect the fact that even the most powerful asymptotically unbiased procedure (i.e., maximum likelihood) may require relatively large tree sizes to reject one model of diversification in favor of another model. Nonetheless, we will show below that our method allows us to confidently reject positive net diversification rates for some important empirical phylogenies, notably the cetaceans.

Confident that our likelihood expression (Eq. **1**) produces unbiased estimates and can accurately detect periods of declining diversity under our model assumptions, we used this expression to estimate diversification rates from a recently published phylogeny of the cetaceans (17) (*Materials and Methods*). This dated molecular phylogeny includes 87 of 89 extant cetacean species, missing only two species in the Delphinidae family. According to the fossil record, the global diversity of the cetaceans increased steadily during the late Oligocene to mid Miocene and subsequently declined monotonically over the past ~10 Myr (Fig. 2*D*) (4).

When we analyzed the cetacean phylogeny as a whole, a pure birth model was selected over all other models, including models that allowed speciation and extinction rates to vary exponentially or linearly through time (*Materials and Methods* and Table S1). In particular, the most likely, pure-birth model suggests that cetacean diversity has increased over the last 30 Myr—in direct contradiction to the fossil record. Similar results were found in previous phylogenetic studies of this group (17, 18). The striking discrepancy with the fossil record is not caused by a large number of extinct taxa in the stem group, because most Oligocene and all Miocene and younger taxa belong to the crown group (4). Thus, even when we allow speciation and extinction rates to vary over time, allow the extinction rate to exceed the speciation rate, and account for undersampling of extant taxa, the levels of extinction estimated from the phylogeny are unrealistically low (Fig. S6*A*), and we fail to identify known periods of diversity loss.

By analyzing the cetaceans as a whole, we have implicitly assumed that the pattern of diversification, $\lambda(t)$ and $\mu(t)$, is ho-

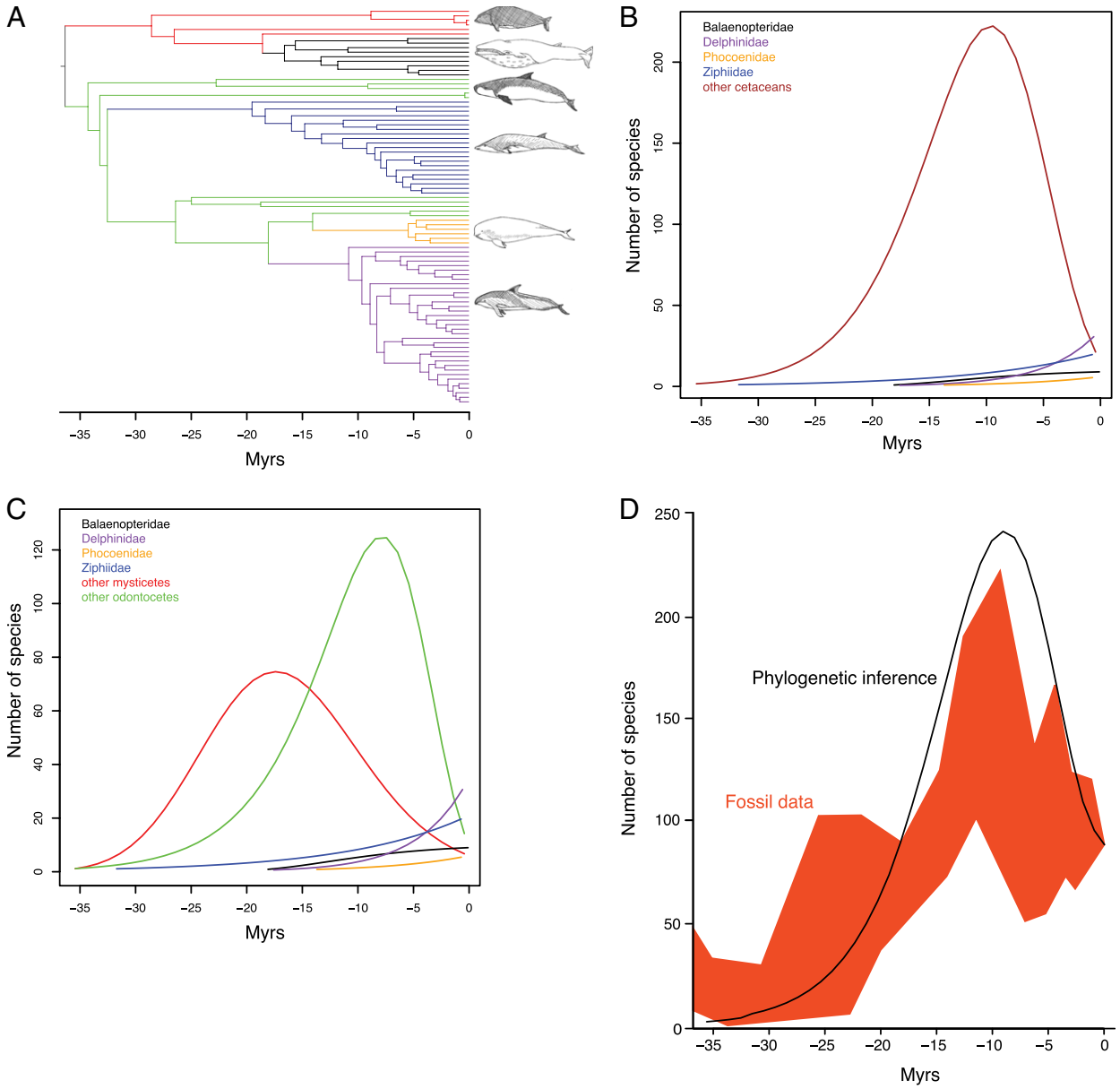

Fig. 2. Phylogenetic inferences of diversity are consistent with the fossil record. (*A*) The cetacean phylogeny. (*B*) The diversity trajectories inferred for each of the five primary cetacean groups; the best fit model is the B-variable model for the Balaenopteridae, the B-constant model for the Delphinidae, Phocoenideae, and Ziphiidae, and the B-constant, D-variable model for the rest of the cetaceans. (*C*) The diversity trajectories inferred for each of the six primary cetacean groups when the mysticetes and odontocetes are analyzed separately; the best fit model is the B-variable, D-constant model for the mysticetes and the B-constant, D-variable model for the odontocetes. (*D*) The total diversity curve inferred for the cetaceans obtained by summing the five individual diversity trajectories (black line in *B*) compared with lower and upper estimates of diversity derived from the fossil record [red, adapted from Quental and Marshall (4)].

mogeneous across clades. This assumption is likely violated, as suggested by the different temporal distributions of nodes in species-rich families compared with the rest of the tree (Fig. 2*A*). Whereas the four most speciose families (the Delphinidae, Balaenopteridae, Phocoenidae, and Ziphiidae) exhibit many recent nodes, the rest of the tree, comprising 10 other families, exhibits relatively few recent speciation events. This pattern suggests that the four species-rich families may have diversified faster than the smaller families, at least recently. To test this hypothesis, we compared the likelihoods of models that allow for different patterns of rate variation in different clades. In particular, we allowed for rate shifts at some or all of the nodes corresponding to the four largest families (*SI Results*); we found that the model allowing distinct patterns of rate variation in each of these families was strongly supported over alternative models (Table S2). When we isolated the four largest families using the rate heterogeneous approach outlined above, we found that the phylogeny of the remaining 16 cetacean species is consistent with a decline in diversity over the past ∼10 Myr (Figs. 2*B* and S6*F*). In particular, the most likely model for these 16 species featured a constant speciation rate through time and an exponentially increasing extinction rate, such that the net diversification rate switched from positive to negative over time (Fig. S6*F*); this model was strongly supported over all alternative patterns of diversification (Table 1 and Table S3). Using the most likely parameter values, we inferred that the diversity of the cetaceans, excluding the four largest families at present, peaked at more than 200 species about 10 Mya, and it subsequently crashed to its present value of 16 extant species (Fig. 2*B*).

The boom-then-bust pattern of diversity that we inferred for the cetaceans, excluding the four largest extant families, is especially notable given the well-known difficulty of inferring nonzero extinction rates from molecular phylogenies. Not only did we infer a positive extinction rate for these groups, we inferred an extinction rate significantly higher than the speciation rate over the past ∼10 Myr [inferred net diversification rate at present: $\lambda(0) - \mu(0) = -0.69$]. We performed a series of tests to determine the robustness of these inferences. First, the 95% χ^2 confidence interval around the estimated net diversification rate at present is (−1.54; −0.28), which allows us to confidently reject the hypothesis that diversity is increasing at present ($P < 0.05$). Second, assuming linear rather than exponential variation in diversification rates through time produced a similar boom-then-bust pattern of diversity (Fig. S7). Third, the diversity trajectory that we inferred was not qualitatively affected if we considered models with two or three rate shifts instead of the model with four shifts (Fig. S8). Finally, if we chose to analyze the mysticetes (excluding the Balaenopteridae) and odontocetes (excluding the Delphinidae, Phocoenidae, and Ziphiidae) separately rather than as a whole (i.e., if we allow for different rates in each of these two clades), then the inferred diversity trajectories are hump-shaped for both groups (Fig. 2*C*), and the present net diversification rates are both significantly negative [−0.26 (−0.49; −0.06) and −0.88 (−1.59; −0.18), respectively]. Under the most likely parameter values, the mysticete group peaked at about 80 species ∼9 Mya, and it then crashed to its present value of 6 extant species. Similarly, the odontocetes peaked at more than 120 species around 9 Mya, and this group retains only 10 species today. Summing these two diversity trajectories, we obtain a diversity curve qualitatively similar to the one obtained when treating the mysticetes and odontocetes as a single clade (Fig. S9). These results all suggest that the boom-then-bust pattern of diversity that we have inferred is not an artifact of the various choices that we made in our analysis.

The history of species diversity in the cetaceans as a whole has been extensively studied using the fossil record (4, 16). To compare our phylogenetic inferences with the fossil record, we summed the diversity trajectories of each individual group (the four largest extant families plus the remaining species). For each of the Balaneopteridae, Delphinidae, Phocoenidae, and Ziphiidae families, we used their best fit models (specified in Fig. 2), which all feature expanding diversity at present (Figs. 2 *B* and *C* and Fig. S6). However, the diversity expansion of these four families does not compensate for the diversity loss in the remaining cetaceans. As a result, the trajectory that we inferred for the cetaceans as a whole features a maximum diversity of almost 250 species about 9 Mya, with only 89 species surviving today (Fig. 2*D*). This trajectory is consistent with the history of diversity inferred from the fossil record (4), which shows a long period of steady species accumulation followed by a sharp decline in diversity starting ∼10 Mya (Fig. 2*D*).

Aside from analyzing historical patterns of net diversity, our phylogenetic inference technique allows us to study how speciation and extinction rates themselves have varied over time (11). Different groups feature different patterns of temporal variation in speciation and extinction rates (Fig. S6 *B–F*). The phylogenies of the Delphinidae, Phocoenidae, and Ziphiidae exhibit relatively constant speciation and suggest that diversity is expanding at present (Fig. S6 *C–F*). The phylogeny of the Balaenopteridae, by contrast, indicates a decay in the net diversification rate caused by a decay in the speciation rate (Fig. S6*B*), so that the Balaenopteridae are currently reaching a point of equilibrium diversity (zero net diversification rate) (14, 30). For the remaining cetaceans, the extinction rate has increased over time, whereas the speciation rate has remained relatively constant, resulting in a negative net diversification rate at present (Fig. S6*F*); the same pattern holds for the odontocetes analyzed separately (Fig. S10). Finally, the mysticete phylogeny indicates a constant extinction rate with a decaying speciation rate (Fig. S10). Thus, a variety of different scenarios operating in different taxonomic groups combine to produce the net diversity trajectories that we observed.

Discussion

Phylogenies of extant taxa are increasingly used to infer macroevolutionary patterns. However, few studies have directly com-

Table 1. Diversification models fitted to the cetacean phylogeny after isolating the Balaneopteridae, Delphinidae, Phocoenidae, and Ziphiidae families

Model	nb	Description	LogL	AICc
B constant	1	No extinction and constant speciation rate	−63.17	128.48
BD constant	2	Constant speciation and extinction rates	−63.17	130.77
B variable	2	No extinction and exponential variation in speciation rate through time	−61.49	127.41
B variable, D constant	3	Exponential variation in speciation rate and constant extinction rate	−56.76	120.40
B constant, D variable	**3**	**Constant speciation rate and exponential variation in extinction rate**	**−56.33**	**119.54**
BD variable	4	Exponential variation in speciation and extinction rates	−56.22	121.99

nb denotes the number of parameters of each model. LogL stands for the maximum log likelihood; AICc stands for the second-order Akaike's information criterion (42). In the most likely model (bold), the speciation rate is constant through time, and the extinction rate increases exponentially through time. This model is supported by the AICc criterion against all other models and the likelihood ratio test against the two models nested within it (B constant and BD constant are both rejected; $P < 0.01$). The best fit model, which specifies an exponentially varying extinction rate, is also supported against models with linear variation in rates over time (Table S3).

pared diversity patterns inferred from phylogenies with those patterns estimated from the fossil record. The few studies attempting this comparison have uncovered major inconsistencies, suggesting that phylogenetic inferences are not reliable on their own (4). In this paper, we have shown that diversity dynamics inferred from phylogenies can be consistent with the fossil record if rate variation through time and among major taxonomic groups is taken into account.

The correspondence that we found between the diversity dynamics inferred from molecular phylogenies and the fossil record is remarkable given that we analyzed the cetaceans, a group that has been used specifically to illustrate major inconsistencies between phylogenies and fossil data (4). Our analysis indicates an important role for species turnover in shaping biodiversity, which is generally found in the fossil record (6) but has rarely been evidenced in molecular phylogenies (12, 32, 33). In addition, our analysis suggests that the net diversification rate has decreased over time in several taxonomic groups, which is often interpreted as a feature of evolutionary radiations (11, 25, 27). Our modeling approach has allowed us to unravel complex historical patterns, such as boom-then-bust patterns of species diversity. These patterns of diversity would have been difficult to discern simply by inspecting the phylogenies without the use of a quantitative cladogenesis model and a corresponding inference procedure.

When the cetacean phylogeny was considered as a whole, with the implicit assumption that diversification rates are homogeneous across lineages, we did not detect any extinction. However, after isolating recently radiating clades from the phylogeny, we recovered realistic extinction rates and diversity trajectories. This finding suggests an important general principle—recently radiating clades mask the signal of extinctions in other clades, but extinctions can be detected from a phylogeny after accounting for rate heterogeneity. These results support the view in Rabosky (15) that different tempos of diversification across lineages are responsible for the current inconsistency between phylogenies and the fossil record, and they suggest that this issue can be overcome (12). When accounting for rate heterogeneity, it is inherently difficult to identify which clades should be analyzed separately. Here, realistic diversity trajectories were obtained by separating the largest extant families. In other situations, more systematic ways to detect rate shifts may be needed (12).

Whereas the historical trajectory of species diversity that we have inferred from the cetacean phylogeny matches the fossil record, our analyses also make more specific inferences than this pattern alone. In particular, we have inferred that the vast majority of cetacean species present about 10 Mya were not within the Balaenopteridae, Delphinidae, Phocoenidae, or Ziphiidae families. In the future, this phylogenetic inference could be tested by detailed examination of the fossil record, with historical specimens identified to the family level.

If we are to extract meaningful information from phylogenies, it is crucial that we understand the strengths and limitations of various analytical approaches. There has been a recent focus on using the γ-statistic to detect declines in speciation rates (20, 21, 24–27). Although the γ-statistic has the advantage of simplicity, it was originally designed only to test deviations from the pure birth Yule process (24). As a result, it is typically not powerful enough to analyze complex diversity trajectories (4, 20, 21). The distribution of phylogenetic branch lengths (13), by contrast, allows us to test whether diversification rates depend on species' age, which is not accommodated by other methods. However, this approach is not powerful for testing more traditional hypotheses, such as whether diversification rates vary over absolute time (14). A recent method based on coalescent theory allows us to compare a variety of scenarios with constant vs. expanding diversity (14). However, this approach does not yet accommodate scenarios of declining diversity, and it relies on an approximate likelihood expression. The approach described here, which is closely related to the approaches used in refs. 8, 10, 11, and 28, rectifies several of these issues simultaneously.

Inferring long-term diversity dynamics without fossils is challenging. Obviously, any incorporation of fossil data to phylogenetic inference will improve our ability to understand diversity dynamics (20, 21, 34–36), and our likelihood expressions can be modified to incorporate some types of fossil information (*SI Results*). However, we have shown here that molecular phylogenies alone can recover diversity dynamics that are consistent with the fossil record. Thus, there is hope for us to reconstruct the history of species diversity in groups or regions that lack a reliable fossil record.

Materials and Methods

Likelihood of Observing a Given Phylogeny. To obtain our likelihood expression (Eq. **1**), we conditioned the cladogenesis process on having at least one lineage surviving to the present and being sampled. The denominator in the likelihood function accounts for this conditioning. Conditioning the process on survival is critical to obtain unbiased parameter estimates, particularly when the probability of survival is low (i.e., when extinction rates exceed speciation rates) (35). Our conditioning is different from the conditioning used by Nee et al. (8) and implemented in both the laser (37) and diversitree (28, 29) packages. Nee et al. (8) conditioned the process on the existence of a root node (i.e., a speciation event occurring at the time of the most recent common ancestor and the two descendant lineages surviving to the present). Modifying Eq. **1** to obtain the likelihood conditioned on the existence of a root node as in Nee et al. (8) is straightforward. Our conditioning allows for taking into account information on the root length (i.e., $t_1 - t_2$) when available, which is the case for subclades within the cetaceans. This form of conditioning also allows us to relax the assumption (8) that all lineages trace back to a single common ancestor at a given time T in the past. In future studies, this flexibility may allow for the combination of phylogenetic and fossil data to gain a more precise understanding of diversity dynamics (*SI Results*).

In Eq. **1**, the factor f^n accounts for the fact that each extant species was sampled with probability f. The factor $\Psi(t_2, t_1)$ corresponds to the probability of observing the given root. The $n - 1$ other factors correspond to the probabilities of observing a speciation event and the two descendant branch lengths at each of the $n - 1$ nodes. To obtain maximum likelihood estimates for a given model, we used the Nelder–Mead simplex algorithm implemented in R (38). R codes computing the likelihood and estimating the maximum likelihood parameters are provided in Dataset S1.

To determine an analytic expression for $\Psi(s, t)$, we first find

$$\Phi(t) = \mathbb{P}\{\text{a lineage is } \textit{not} \text{ in the sample} | \text{it was alive at the time } t\} \quad [5]$$

which following ref. 10, can be obtained through an ordinary differential equation. Notice that

$$\begin{aligned}\Phi(t+\Delta t) &= \mathbb{P}\left\{\begin{array}{c}\text{lineage goes extinct}\\ \text{in } (t, t+\Delta t)\end{array}\right\}\\ &\quad + \mathbb{P}\left\{\begin{array}{c}\text{no extinction and speciation,}\\ \text{but neither lineage is observed at present}\end{array}\right\}\\ &\quad + \mathbb{P}\left\{\begin{array}{c}\text{no extinction or and speciation in} (t, t+\Delta t),\\ \text{but lineage is not observed at present}\end{array}\right\}\\ &= \mu(t)\Delta t + (1-\mu(t)\Delta t)\lambda(t)\Phi^2(t)\\ &\quad + (1-\mu(t)\Delta t)(1-\lambda(t)\Delta t)\Phi(t) + o(\Delta t).\end{aligned} \quad [6]$$

Subtracting $\Phi(t)$, dividing by Δt, and taking $\Delta t \to 0$ yields

$$\frac{d\Phi}{dt} = \mu(t) - (\lambda(t) + \mu(t))\Phi(t) + \lambda(t)\Phi^2(t), \quad [7]$$

whereas

$$\Phi(0) = \mathbb{P}\{\text{a lineage is } \textit{not} \text{ in the sample} | \text{it was alive at time } 0\} = 1 - f. \quad [8]$$

Set $F(t) = 1 - \Phi(t)$. Then, $F(t)$ satisfies the Bernoulli equation

$$\frac{dF}{dt} = (\lambda(t) - \mu(t))F(t) - \lambda(t)F^2(t). \quad [9]$$

Letting $G(t) = 1/F(t)$, we have

$$\frac{dG}{dt} = -(\lambda(t) - \mu(t))G(t) + \lambda(t), \quad [10]$$

which is readily solved as

$$G(t) = e^{-\int_0^t \lambda(u)-\mu(u)du}\left(G(0) - \int_0^t e^{\int_0^s \lambda(u)-\mu(u)du}\lambda(s)ds\right), \quad [11]$$

where $G(0) = 1/f$. Solving for $\Phi(t)$ produces Eq. **2**. To determine $\Psi(s, t)$ itself, following ref. 10, we note that

$$\begin{aligned}\Psi(s, t+\Delta t) &= \mathbb{P}\left\{\begin{array}{l}\text{no extinction}\\ \text{in } (t, t+\Delta t)\end{array}\right\} \times \left(\mathbb{P}\left\{\begin{array}{l}\text{no speciation}\\ \text{in } (t, t+\Delta t)\end{array}\right\}\right.\\ &\quad \left. + \mathbb{P}\left\{\begin{array}{l}\text{speciation in } (t, t+\Delta t), \text{ but one of}\\ \text{the two lineages is not in the sample}\end{array}\right\}\right) \quad [12]\\ &\quad \times \mathbb{P}\left\{\begin{array}{l}\text{the lineage survives from } t \text{ to } s\\ \text{without any } \textit{observed} \text{ daughters}\end{array}\right\}\\ &= (1-\mu(t)\Delta t)\,((1-\lambda(t)\Delta t) + 2\lambda(t)\Delta t\Phi(t)\Psi(s, t).\end{aligned}$$

Subtracting $\Psi(s, t)$, dividing by Δt, and taking $\Delta t \to 0$ yields

$$\frac{d\Psi(s, t)}{dt} = ((2\Phi(t)-1)\lambda(t) - \mu(t))\Psi(s, t), \quad [13]$$

whereas $\Psi(s, s) = 1$, because the lineage can neither disappear nor give birth at a single instant. Solving this ordinary differential equation yields

$$\Psi(s, t) = e^{\int_s^t (2\Phi(u)-1)\lambda(u)-\mu(u)du}. \quad [14]$$

Substituting Eq. **2** into Eq. **14**, we get Eq. **3** (*SI Results*).

Models of Diversification. We considered models with time-constant diversification rates and models with time-variable diversification rates. When rates varied through time, we assumed one of two variations: either exponential variation, such that $\lambda(t) = \lambda_0 e^{\alpha t}$ and $\mu(t) = \mu_0 e^{\beta t}$, where λ_0 and μ_0 are the speciation and extinction rates at present, respectively, and α and β are the rates of change, or linear variation, such that $\lambda(t) = \max(0, \lambda_0 + \alpha t)$ and $\mu(t) = \max(0, \mu_0 + \beta t)$.

Computing Confidence Intervals. Given a phylogeny (simulated or empirical), we computed the 95% confidence interval around the maximum likelihood estimate of the net diversification rate according to the χ^2 distribution. To perform this computation, we changed variables to parameterize the likelihood function in terms of the net diversification rate and speciation rate. We then computed the confidence interval corresponding to a likelihood ratio test, with 1 degree of freedom and $P = 0.05$, by finding the minimum and maximal values of the net diversification rate within 3.84/2 log-likelihood units of the maximal log-likelihood value (39).

Cetacean Phylogeny. We analyzed the dated cetacean phylogeny constructed by Steeman et al. (17), which consists of 87 of 89 extant cetacean species. This phylogeny was derived from six mitochondrial and nine nuclear genes using the Bayesian phylogenetic inference implemented in MrBayes (40). It was calibrated using seven paleontological age constraints and the relaxed molecular clock approach implemented in r8S (41).

ACKNOWLEDGMENTS. We thank L. Harmon, C. Marshall, T. Quental, D. Rabosky, and M. Sanderson for discussions and/or comments on our manuscript, and we thank A. Chen-Plotkin for the scientific illustrations in Fig. 2. We also thank the editor and two anonymous reviewers for very constructive comments. H.M. acknowledges support from the Centre National de la Recherche Scientifique, and J.B.P. acknowledges funding from the Burroughs Wellcome Fund, the David and Lucile Packard Foundation, the Alfred P. Sloan Foundation, and the James S. McDonnell Foundation.

1. Ricklefs RE (2007) Estimating diversification rates from phylogenetic information. *Trends Ecol Evol* 22:601–610.
2. Benton MJ (2009) The Red Queen and the Court Jester: Species diversity and the role of biotic and abiotic factors through time. *Science* 323:728–732.
3. Rabosky DL (2009) Ecological limits and diversification rate: Alternative paradigms to explain the variation in species richness among clades and regions. *Ecol Lett* 12: 735–743.
4. Quental TB, Marshall CR (2010) Diversity dynamics: Molecular phylogenies need the fossil record. *Trends Ecol Evol* 25:434–441.
5. Raup DM (1972) Taxonomic diversity during the Phanerozoic. *Science* 177:1065–1071.
6. Alroy J (2008) Colloquium paper: Dynamics of origination and extinction in the marine fossil record. *Proc Natl Acad Sci USA* 105(Suppl 1):11536–11542.
7. Alroy J (2010) The shifting balance of diversity among major marine animal groups. *Science* 329:1191–1194.
8. Nee S, May RM, Harvey PH (1994) The reconstructed evolutionary process. *Philos Trans R Soc Lond B Biol Sci* 344:305–311.
9. Nee S (2006) Birth-death models in macroevolution. *Annu Rev Ecol Evol Syst* 37:1–17.
10. Maddison WP, Midford PE, Otto SP (2007) Estimating a binary character's effect on speciation and extinction. *Syst Biol* 56:701–710.
11. Rabosky DL, Lovette IJ (2008) Explosive evolutionary radiations: Decreasing speciation or increasing extinction through time? *Evolution* 62:1866–1875.
12. Alfaro ME, et al. (2009) Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proc Natl Acad Sci USA* 106:13410–13414.
13. Venditti C, Meade A, Pagel M (2010) Phylogenies reveal new interpretation of speciation and the Red Queen. *Nature* 463:349–352.
14. Morlon H, Potts MD, Plotkin JB (2010) Inferring the dynamics of diversification: A coalescent approach. *PLoS Biol* 8:e1000493.
15. Rabosky DL (2010) Extinction rates should not be estimated from molecular phylogenies. *Evolution* 64:1816–1824.
16. Marx FG, Uhen MD (2010) Climate, critters, and cetaceans: Cenozoic drivers of the evolution of modern whales. *Science* 327:993–996.
17. Steeman ME, et al. (2009) Radiation of extant cetaceans driven by restructuring of the oceans. *Syst Biol* 58:573–585.
18. Slater GJ, Price SA, Santini F, Alfaro ME (2010) Diversity versus disparity and the radiation of modern cetaceans. *Proc Biol Sci* 277:3097–3104.
19. Purvis A (2008) Phylogenetic approaches to the study of extinction. *Annu Rev Ecol Evol Syst* 39:310–319.
20. Quental TB, Marshall CR (2009) Extinction during evolutionary radiations: Reconciling the fossil record with molecular phylogenies. *Evolution* 63:3158–3167.
21. Liow LH, Quental TB, Marshall CR (2010) When can decreasing diversification rates be detected with molecular phylogenies and the fossil record? *Syst Biol* 59:646–659.
22. Mooers AO, Heard SB (1997) Inferring evolutionary process from phylogenetic tree shape. *Q Rev Biol* 72:31–54.
23. Losos JB (2011) Seeing the forest for the trees: The limitations of phylogenies in comparative biology. *Am Nat* 177:709–727.
24. Pybus OG, Harvey PH (2000) Testing macro-evolutionary models using incomplete molecular phylogenies. *Proc Biol Sci* 267:2267–2272.
25. Harmon LJ, Schulte JA, 2nd, Larson A, Losos JB (2003) Tempo and mode of evolutionary radiation in iguanian lizards. *Science* 301:961–964.
26. McPeek MA (2008) The ecological dynamics of clade diversification and community assembly. *Am Nat* 172:E270–E284.
27. Phillimore AB, Price TD (2008) Density-dependent cladogenesis in birds. *PLoS Biol* 6: e71.
28. FitzJohn RG, Maddison WP, Otto SP (2009) Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. *Syst Biol* 58:595–611.
29. FitzJohn RG (2010) Quantitative traits and diversification. *Syst Biol* 59:619–633.
30. Rabosky DL, Glor RE (2010) Equilibrium speciation dynamics in a model adaptive radiation of island lizards. *Proc Natl Acad Sci USA* 107:22178–22183.
31. Stadler T (2011) Mammalian phylogeny reveals recent diversification rate shifts. *Proc Natl Acad Sci USA* 108:6187–6192.
32. Ricklefs RE (2003) Global diversification rates of passerine birds. *Proc Biol Sci* 270: 2285–2291.
33. Ricklefs RE (2006) Global variation in the diversification rate of passerine birds. *Ecology* 87:2468–2478.
34. Paradis E (2004) Can extinction rates be estimated without fossils? *J Theor Biol* 229: 19–30.
35. Etienne RS, Apol MEF (2009) Estimating speciation and extinction rates from diversity data and the fossil record. *Evolution* 63:244–255.
36. Alfaro M, Santini F (2010) Evolutionary biology: A flourishing of fish forms. *Nature* 464:840–842.
37. Rabosky DL (2006) LASER: A maximum likelihood toolkit for detecting temporal shifts in diversification rates from molecular phylogenies. *Evol Bioinform* 2:247–250.
38. Nelder JA, Mead R (1965) A simplex method for function minimization. *Comput J* 7: 308–313.
39. Bolker BM (2008) *Ecological Models and Data in R* (Princeton University Press, Princeton).
40. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.
41. Sanderson MJ (2002) Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. *Mol Biol Evol* 19:101–109.
42. Burnham KP, Anderson DR (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (Springer-Verlag, New York), 2nd Ed.

Supporting Information

Morlon et al. 10.1073/pnas.1102543108

SI Results

Simulations: Methods and Results. We used forward time Monte Carlo simulations (1) to quantify the accuracy of phylogenetic inferences based on Eq. **1**. We assumed a simple model of cladogenesis in which the extinction rate is constant and the speciation rate decays over time (Fig. S1), or conversely, a model with constant speciation and increasing extinction rate (Fig. S3). The simulations were started at time zero with a single lineage. As time moved forward, we simulated Poisson events with rate $\lambda + \mu$ (i.e., speciation plus extinction rates at the time of the previous event; for the first event, we used the rates at time zero). At each event, a lineage picked at random was replaced by two descendant lineages with probability $\frac{\lambda}{\lambda+\mu}$ and removed with probability $\frac{\mu}{\lambda+\mu}$, and the speciation (or extinction) rate was updated (the other rate remained constant). The process was simulated until time exceeded a predetermined value.

Fig. S1 shows maximum likelihood parameter estimates for phylogenies simulated under the model, with constant extinction rate and time-decaying speciation rate, for five different parameter sets. In practice, given a suite of phylogenies simulated under a given parameter set, we computed, for each of these phylogenies, the maximum likelihood estimates of the diversification parameters. We then computed the median and 5% and 95% quantiles of these estimates. Fig. S2 illustrates the distribution of parameter estimates across phylogenies. Fig. S3 illustrates that unbiased parameter estimates are also obtained for phylogenies simulated under the model with constant speciation rate and time-increasing extinction rate.

Figs. S4 and S5 illustrate how confidence limits of estimated parameters depend on tree size. We simulated 10,000 phylogenies with the parameters corresponding to the rightmost data point in Fig. S1. The simulated net diversification rate is −0.25. Of the 10,000 phylogenies, we considered (*i*) the 186 phylogenies with estimated net diversification rates at present within −0.26, −0.24 (Fig. S4) and (*ii*) 300 randomly sampled phylogenies (Fig. S5). For each of these phylogenies, we computed the 95% confidence interval using a χ^2 criterion as described in the text.

Diversity Dynamics of the Cetaceans. When the cetacean phylogeny is considered as a whole, we find no support for the presence of extinction (Table S1 and Fig. S6*A*). The constant rate pure birth model (B constant) is selected against all other models based on either likelihood ratio tests or the second-order Akaike's information criterion (AIC_c). More generally, pure birth models with constant (B constant), exponentially varying (B variable E), and linearly varying (B variable L) speciation rates are all selected against their analogs with extinction.

Besides models in which diversification rates are homogeneous across clades, we considered models in which rates change instantaneously at a fixed set of branching points. In particular, we tested for rate shifts at the bases of the four richest families (the Balaenopteridae, Delphinidae, Phocoenidae, and Ziphiidae). To do this testing, we performed a forward selection procedure similar to the procedure used in ref. 2. We first assumed that there is a single rate shift at one of these four possible nodes. For each of these possible breakpoints, we selected the best fit model for both the corresponding subclade and the remaining phylogeny among a set of 10 models with constant, exponentially varying, or linearly varying diversification rates (Table S1 has a list of models). We then computed the maximum likelihood of the corresponding combined phylogeny, as described in the text. Finally, we selected the breakpoint that maximized this likelihood and tested the statistical support of the breakpoint using a likelihood ratio test. We iterated the process and tested for the statistical support of each additional breakpoint. We found that the model allowing distinct patterns of rate variation in the four most speciose families (the Balaneopteridae, Delphinidae, Phocoenidae, and Ziphiidae families) was strongly supported over models allowing fewer distinct patterns (Table S2).

When isolating the Balaneopteridae, Delphinidae, Phocoenidae, and Ziphiidae families, the model best describing the phylogeny of the remaining 16 cetacean species is a model in which the speciation rate remains constant while the extinction rate varies exponentially through time (Table S3 and Fig. S6*F*). In particular, this model is more likely than models where the time variation is linear.

When we assume a linear time variation in diversification rates, the diversity trajectory of the cetaceans features a boom-then-bust pattern similar to the one obtained assuming an exponential time variation with a diversity peak ∼10 Mya (Fig. S7).

When we consider the most likely models with only three or two rate shifts (Fig. S8), the inferred trajectory is qualitatively similar to the one obtained when considering the four-shifts model (Fig. 2).

The model in which the same diversification rates are applied to the mysticetes (excluding the Balaenopteridae) and the odontocetes (excluding the Delphinidae, Phocoenidae, and Ziphiidae) is not rejected against a model in which different diversification rates are allowed for each group [number of parameters (nb) = 11, maximal log likelihood (LogL) = −247.64, AIC_c = 518.93 compared with values for the four-shifts model in Table S2; $P > 0.1$ according to a likelihood ratio test]. If we nevertheless analyze the odontocetes and mysticetes separately rather than as a whole, the inferred diversity dynamics of the cetaceans are qualitatively similar to the dynamics inferred when analyzing the phylogeny of the odontocetes and mysticetes combined (Fig. S9 compared with Fig. 2*B*, brown curve).

When analyzing the mysticetes and odontocetes separately, their respective phylogenies suggest that the speciation rate has decreased in the mysticetes, whereas the extinction rate has increased in the odontocetes (Fig. S10). In both cases, these dynamics produce a switch from positive to negative diversification rates over time (i.e., a diversity increase followed by a diversity decline).

Details of the Derivation of $\Psi(s,t)$. Substituting Eq. **2** into Eq. **14** (in the text), we get

$$\Psi(s,t) = e^{\int_s^t \left(1 - 2\frac{e^{\int_0^u \lambda(\sigma)-\mu(\sigma)d\sigma}}{\frac{1}{f} + \int_0^u e^{\int_0^\tau \lambda(\sigma)-\mu(\sigma)d\sigma}\lambda(\tau)d\tau}\right)\lambda(u)-\mu(u)du}$$

$$= e^{\int_s^t \lambda(u)-\mu(u) - 2\frac{e^{\int_0^u \lambda(\sigma)-\mu(\sigma)d\sigma}\lambda(u)}{\frac{1}{f} + \int_0^u e^{\int_0^\tau \lambda(\sigma)-\mu(\sigma)d\sigma}\lambda(\tau)d\tau}du}$$

$$= e^{\int_s^t \lambda(u)-\mu(u)du - 2\ln\left|\frac{\frac{1}{f} + \int_0^t e^{\int_0^\tau \lambda(\sigma)-\mu(\sigma)d\sigma}\lambda(\tau)d\tau}{\frac{1}{f} + \int_0^s e^{\int_0^\tau \lambda(\sigma)-\mu(\sigma)d\sigma}\lambda(\tau)d\tau}\right|}$$

$$= e^{\int_s^t \lambda(u)-\mu(u)du - 2\ln\left|1+\frac{\int_s^t e^{\int_0^\tau \lambda(\sigma)-\mu(\sigma)d\sigma}\lambda(\tau)d\tau}{\frac{1}{f}+\int_0^s e^{\int_0^\tau \lambda(\sigma)-\mu(\sigma)d\sigma}\lambda(\tau)d\tau}\right|}$$

$$= e^{\int_s^t \lambda(u)-\mu(u)du}\left[1+\frac{\int_0^t e^{\int_0^\tau \lambda(\sigma)-\mu(\sigma)d\sigma}\lambda(\tau)d\tau}{\frac{1}{f}+\int_0^s e^{\int_0^\tau \lambda(\sigma)-\mu(\sigma)d\sigma}\lambda(\tau)d\tau}\right]^{-2}. \quad \text{[S1]}$$

Likelihood for Specific Forms of the Time Dependence of the Speciation and Extinction Rates. The general likelihood expression given by Eq. **1** may be calculated for specific forms of the time dependency of the speciation and extinction rates [$\lambda(t)$ and $\mu(t)$] by calculating $\Phi(t)$ (Eq. **2**) and $\Psi(s,t)$ (Eq. **3**). For example, if $\lambda(t)$ and $\mu(t)$ are assumed constant over time, then

$$\lambda(t) = \lambda_0 \text{ and} \quad \text{[S2]}$$

$$\mu(t) = \mu_0. \quad \text{[S3]}$$

We find

$$\Phi(t) = 1 - \frac{e^{(\lambda_0-\mu_0)t}}{\frac{1}{f}+\frac{\lambda_0}{\lambda_0-\mu_0}\left(e^{(\lambda_0-\mu_0)t}-1\right)} \quad \text{[S4]}$$

and

$$\Psi(s,t) = e^{(\lambda_0-\mu_0)(t-s)}\left[1+\frac{\frac{\lambda_0}{\lambda_0-\mu_0}(e^{(\lambda_0-\mu_0)t}-e^{(\lambda_0-\mu_0)s})}{\frac{1}{f}+\frac{\lambda_0}{\lambda_0-\mu_0}(e^{(\lambda_0-\mu_0)s}-1)}\right]^{-2}. \quad \text{[S5]}$$

If $\lambda(t)$ and $\mu(t)$ are assumed to vary linearly with time,

$$\lambda(t) = \lambda_0 + \alpha t \text{ and} \quad \text{[S6]}$$

$$\mu(t) = \mu_0 + \beta t, \quad \text{[S7]}$$

where λ_0 (μ_0) is the speciation (extinction) rate at present and α (β) is the linear variation in speciation (extinction) rate over time.

We find

$$\Phi(t) = 1 - \frac{e^{(\lambda_0-\mu_0)t+\frac{\alpha-\beta}{2}t^2}}{\frac{1}{f}+\int_0^t e^{(\lambda_0-\mu_0)s+\frac{\alpha-\beta}{2}s^2}(\lambda_0+\alpha s)ds} \quad \text{[S8]}$$

and

$$\Psi(s,t) = e^{(\lambda_0-\mu_0)(t-s)+\frac{\alpha-\beta}{2}(t^2-s^2)} \times \left[1+\frac{\int_s^t e^{(\lambda_0-\mu_0)\tau+\frac{\alpha-\beta}{2}\tau^2}(\lambda_0+\alpha\tau)d\tau}{\frac{1}{f}+\int_0^s e^{(\lambda_0-\mu_0)\tau+\frac{\alpha-\beta}{2}\tau^2}(\lambda_0+\alpha\tau)d\tau}\right]^{-2}, \quad \text{[S9]}$$

which can be integrated numerically.

If $\lambda(t)$ and $\mu(t)$ are assumed to vary exponentially with time,

$$\lambda(t) = \lambda_0 e^{\alpha t} \text{ and} \quad \text{[S10]}$$

$$\mu(t) = \mu_0 e^{\beta t}, \quad \text{[S11]}$$

where λ_0 (μ_0) is the speciation (extinction) rate at present and α (β) is the exponential variation in speciation (extinction) rate over time.

We find

$$\Phi(t) = 1 - \frac{\frac{\lambda_0}{\alpha}(e^{\alpha t}-1) - \frac{\mu_0}{\beta}(e^{\beta t}-1)}{\frac{1}{f}+\int_0^t e^{\frac{\lambda_0}{\alpha}(e^{\alpha\tau}-1)-\frac{\mu_0}{\beta}(e^{\beta\tau}-1)+\alpha\tau}d\tau} \quad \text{[S12]}$$

and

$$\Psi(s,t) = \left[\frac{\lambda_0}{\alpha}(e^{\alpha t}-e^{\alpha s}) - \frac{\mu_0}{\beta}(e^{\beta t}-e^{\beta s})\right] \times \left[1+\frac{\lambda_0\int_s^t e^{\frac{\lambda_0}{\alpha}(e^{\alpha\tau}-1)-\frac{\mu_0}{\beta}(e^{\beta\tau}-1)+\alpha\tau}d\tau}{\frac{1}{f}+\lambda_0\int_0^s e^{\frac{\lambda_0}{\alpha}(e^{\alpha\tau}-1)-\frac{\mu_0}{\beta}(e^{\beta\tau}-1)+\alpha\tau}d\tau}\right]^{-2}, \quad \text{[S13]}$$

which can be integrated numerically.

Relationship to Ref. 3. In ref. 3, Nee et al. use forward time results from ref. 4 to derive a likelihood for $\{x_2,\ldots,x_n\}$, the times of the first, second, etc., branching events in a phylogeny, assuming that the phylogeny can be traced back to a single common ancestor at time $x = 0$ in the past (they take time $x = T$ to be the present). Letting $\tilde{\lambda}(x)$ and $\tilde{\mu}(x)$ be the speciation and extinction rates, respectively, at time x from the time of the most recent common ancestor, they introduce two functions:

$$P(x,T) = \mathbb{P}\{\text{a single lineage alive at time } x \text{ is not extinct at time } T\} = \frac{1}{1+\int_x^T \tilde{\mu}(s)e^{\int_x^s \tilde{\mu}(u)-\tilde{\lambda}(u)du}ds} \quad \text{[S14]}$$

and

$$u_{x,T} = 1 - \mathbb{P}\{\text{no offspring in } (x,T)\} = 1 - P(x,T)e^{\int_x^T \tilde{\mu}(u)-\tilde{\lambda}(u)du}. \quad \text{[S15]}$$

Using these expressions, one can readily write down a likelihood for the tree rooted at $x_2 \geq 0$, the time of the first known branching event:

$$(1-u_{x_2,T})^2\prod_{i=3}^n \tilde{\lambda}(x_i)\prod_{i=3}^n P(x_i,T)(1-u_{x_i,T}). \quad \text{[S16]}$$

This likelihood expression does not appear in ref. 3, but it can be readily obtained from results in sections 5 and 6 in ref. 3.

Although at first glance, it may seem that considering only the birth times may be less informative than considering the entire branch lengths, in reality, up to conditioning (and taking $f = 1$), our expression is equivalent to Eq. **S16**, which we show below (see also refs. 5 and 6). The equivalence between our

likelihood and the likelihood in Nee et al. (3), which depends only on branching times (not branching intervals), shows that trees with the same branching times but different topology have the same likelihood under our model assumptions (7).

To reconcile our backward time approach with the forward time approach used in ref. 3, we ignore the phylogenetic root, and we take $t = T - x$:

$$\mu(t) = \tilde{\mu}(T-t) = \tilde{\mu}(x) \text{ and} \quad \textbf{[S17]}$$

$$\lambda(t) = \tilde{\lambda}(T-t) = \tilde{\lambda}(x) \quad \textbf{[S18]}$$

(i.e., the rates at time t before the present). We begin by observing that, if a species that first appears at time $t_i = T - x_i$ before the present is in our sample, then for each species, the phylogeny must contain branches

$$\left(t'_{i,1}, t_i\right), \left(t'_{i,2}, t'_{i,1}\right), \ldots, \left(0, t'_{i,k}\right)$$

connecting the birth time to the present. Moreover, if the tree is rooted at time t_2, there are two species tracing back to t_2 and thus, two such collections of intervals. For example, for the tree in Fig. 1 *Left*, we get (ignoring the root)

$$\begin{array}{lll} (t_3, t_2), & (t_4, t_3), & (0, t_4), \\ & (0, t_2), & \\ & (0, t_3), & \\ & (0, t_4), & \end{array}$$

whereas for the tree in Fig. 1 *Right*, we have

$$\begin{array}{c} (t_3, t_2), (0, t_3), \\ (t_4, t_2), (0, t_4), \\ (0, t_3), \\ (0, t_4). \end{array}$$

When ignoring the phylogenetic root, the likelihood of observing a given phylogeny (without any conditioning) is

$$f^n \prod_{i=2}^{n} \lambda(t_i) \Psi\left(s_{i,1}, t_i\right) \Psi\left(s_{i,2}, t_i\right). \quad \textbf{[S19]}$$

Now, from the analytical expression of $\Psi(s,t)$ (Eq. **5**), we see that, if $s \le u \le t$, then

$$\Psi(s,u)\Psi(u,t) = \Psi(s,t). \quad \textbf{[S20]}$$

Thus, the likelihood may be written as

$$f^n \lambda(t_2) \Psi(0,t_2)^2 \prod_{i=3}^{n} \lambda(t_i) \Psi(0,t_i). \quad \textbf{[S21]}$$

Next, observe that, by definition (which may also be verified by direct calculation),

$$P(x,T) = 1 - \Phi(T-x) = 1 - \Phi(t) \quad \textbf{[S22]}$$

and

$$\begin{aligned} u(x,T) &= 1 - P(x,T) e^{\int_x^T \tilde{\mu}(u) - \tilde{\lambda}(u) du} \\ &= 1 - (1 - \Phi(T-x)) e^{\int_0^{T-x} \mu(u) - \lambda(u) du} \\ &= 1 - (1 - \Phi(t)) e^{\int_0^t \mu(u) - \lambda(u) du}. \end{aligned} \quad \textbf{[S23]}$$

Last, from Eq. **3**, we see that

$$\begin{aligned} \Psi(0,t) &= e^{\int_0^t \lambda(u) - \mu(u) du} \left[1 + \int_0^t e^{\int_0^\tau \lambda(\sigma) - \mu(\sigma) d\sigma} \lambda(\tau) d\tau\right]^{-2} \\ &= e^{-\int_0^t \lambda(u) - \mu(u) du} (1 - \Phi(t))^2 \\ &= P(x,T)(1 - u(x,T)). \end{aligned} \quad \textbf{[S24]}$$

Thus, our expression for the likelihood becomes

$$\begin{aligned} \lambda(t_2)\Psi(0,t_2)^2 \prod_{i=3}^{n} \lambda(t_i)\Psi(0,t_i) &= \tilde{\lambda}(x_2) P(x_2,T)^2 (1 - u(x_2,T))^2 \\ &\times \prod_{i=3}^{n} \tilde{\lambda}(x_i) P(x_i,T)(1 - u(x_i,T)), \end{aligned} \quad \textbf{[S25]}$$

which divided by $\tilde{\lambda}(x_2) P(x_2,T)^2$ (i.e., conditioning on a speciation event at time t_2/x_2 with both lines from that event in the sample), yields Eq. **S16**.

Incorporating Fossil Data. This form of conditioning also allows us to relax the assumption (3) that all lineages trace back to a single common ancestor at a given time T in the past. In future studies, this flexibility may allow for the combination of phylogenetic and fossil data to gain a more precise understanding of diversity dynamics.

If our data trace back to N_a ancestors at some time T in the past, then the data are the disjoint union of N_a trees:

$$\left\{\left(t_2^1, t_1^1\right), \left\{\left(s_{i,1}^1, t_i^1\right), \left(s_{i,2}^1, t_i^1\right)\right\}_{i=2,\ldots,n^1}\right\}, \ldots,$$

$$\left\{\left(t_2^{N_a}, t_1^{N_a}\right), \left\{\left(s_{i,1}^{N_a}, t_i^{N_a}\right), \left(s_{i,2}^{N_a}, t_i^{N_a}\right)\right\}_{i=2,\ldots,n^{N_a}}\right\}. \quad \textbf{[S26]}$$

The probability that each of the N_a ancestors has some descendant in the sample at $t = 0$ is $(1 - \Phi(T))^{N_a}$, and the conditional likelihood is

$$\frac{\prod_{j=1}^{N_a} f^{n^j} \Psi\left(t_2^j, t_1^j\right) \prod_{i=1}^{n^j} \lambda\left(t_i^j\right) \Psi\left(s_{i,1}^j, t_i^j\right) \Psi\left(s_{i,2}^j, t_i^j\right)}{(1 - \Phi(T))^{N_a}}. \quad \textbf{[S27]}$$

Suppose that, in addition to the N_a ancestors of the sample alive at time T, we also know from fossil data that there were at least N_e species alive at time T that have left no observed descendants. This event has probability $\Phi(T)^{N_e}$, which can be incorporated into a joint likelihood for the data by multiplication:

$$\frac{\prod_{j=1}^{N_a} f^{n^j} \Psi\left(t_2^j, t_1^j\right) \prod_{i=1}^{n^j} \lambda\left(t_i^j\right) \Psi\left(s_{i,1}^j, t_i^j\right) \Psi\left(s_{i,2}^j, t_i^j\right)}{(1 - \Phi(T))^{N_a}} \Phi(T)^{N_e}. \quad \textbf{[S28]}$$

Likelihood When Rates Vary Across Lineages. In this section, we consider the situation where different lineages are allowed to have different speciation and extinction rates. Over the lifetime of a species, these rates vary according to $\lambda(t)$ and $\mu(t)$, but when a new species appears, it has extinction and speciation rates $\tilde{\lambda}(t)$ and $\tilde{\mu}(t)$ that may be different from the parent lineage.

We, henceforth, refer to the pair of functions, $\boldsymbol{x} = (\lambda, \mu)$, as the species type and write $\mathfrak{X}$ for the set of all possible types. If $\lambda(t)$ and $\mu(t)$ are chosen from a parametric family [e.g., in the text, we use $\lambda(t) = \lambda_0 e^{\alpha t}$ and $\mu(t) = \mu_0 e^{\beta t}$], the type may be equivalently represented by those parameters (e.g., λ_0, α, μ_0, and β). Note that, after the rates are allowed to vary across the tree, all topologies are no longer equally probable.

Previously, we considered two functions: $\Phi(t)$, the probability that a lineage alive at time t has no descendant in the sample,

and $\Psi(s, t)$, the probability that a lineage alive at time t leaves exactly one descendant lineage at time $s < t$ in the reconstructed phylogeny. We will now have to consider typed functions $\Phi_x(t)$ and $\Psi_x(s, t)$.

Below, we begin by deriving a general expression for $\Phi_x(t)$ and $\Psi_x(s, t)$. We then show that, if we assume that the rates only change at a known set of branching points, these expressions reduce to the expressions found previously when rates were assumed to be homogeneous across lineages.

Deriving $\Phi_x(t)$ and $\Psi_x(s, t)$. We take an approach informed by the general branching process (8), which offers a particularly clear conceptual framework. Consider a lineage of type $x_p = (\lambda_p, \mu_p)$ that is known to have been alive at time t (the parent lineage). Let L_p be the remaining lifespan of the parent lineage (i.e., it dies at $t - L_p$). Owing to the memoryless property of the exponential distribution, we have

$$\mathbb{P}\{L_p > t - s\} = e^{-\int_s^t \mu_p(u)du}. \quad \text{[S29]}$$

Moreover, because we model the phylogeny using a Markov branching process, the birth times of descendant lineages born after time t will be the points of a time-inhomogeneous Poisson point process on $(t, t - L_p)$ with rate function $\lambda_p(t)$. Hence, the probability density of n descendant lineages born at time $s_1,\ldots,s_n$ in (s, t) [with $s < t + L$] is

$$\frac{e^{-\int_s^t \lambda_p(u)du}}{n!}\lambda_p(s_1)\ldots\lambda_p(s_n)ds_1\ldots ds_n. \quad \text{[S30]}$$

Suppose these descendant lineages have type $x_1, \ldots, x_n$. For the moment, we make no assumption about the types (e.g., types can be all identical, all different, change at known set of branching points, etc.). The probability that the i^{th} leaves no observed descendant is $\Phi_{\mathbf{x}_i}(s_i)$, and thus, the probability that no descendant of a lineage born in (s, t) is observed in the sample at time $t = 0$ is

$$\Pi_{\mathbf{x}_p}(s,t) = e^{-\int_s^t \lambda_p(u)du}\sum_{n=0}^{\infty}\frac{1}{n!}\prod_{i=1}^{n}\int_s^t \lambda_{\mathbf{x}_i}(s_i)\Phi_{\mathbf{x}_i}(s_i)ds_i. \quad \text{[S31]}$$

Notice that, if all types are identical, $\Phi_{\mathbf{x}_1} = \cdots = \Phi_{\mathbf{x}_n} = \Phi_{\mathbf{x}_p}$, then

$$\begin{aligned}\Pi_{\mathbf{x}_p}(s,t) &= e^{-\int_s^t \lambda_p(u)du}\sum_{n=0}^{\infty}\frac{1}{n!}\prod_{i=1}^{n}\int_s^t \lambda_p(s_i)\Phi_{\mathbf{x}_p}(s_i)ds_i \\ &= e^{-\int_s^t \lambda_p(u)du}\sum_{n=0}^{\infty}\frac{1}{n!}\left(\int_s^t \lambda_p(s)\Phi_{\mathbf{x}_p}(s)ds\right)^n \\ &= e^{\int_s^t \lambda_p(u)(\Phi_{\mathbf{x}_p}(u)-1)du}. \quad \text{[S32]}\end{aligned}$$

Finally, there are two (mutually exclusive) ways that the parent lineage may leave no observed descendants.

i) The parent lineage survives to the present but is not observed, and it leaves no descendant lineage. Thus, $L_p > t$, and the probability of this event is

$$(1-f_p)e^{-\int_0^t \mu_p(u)du}\Pi_{\mathbf{x}_p}(0,t). \quad \text{[S33]}$$

ii) The parent lineage dies at some $0 < s < t$. Thus, $L_p = t - s$, and the probability of this event is

$$\int_0^t e^{-\int_0^t \mu_p(u)du}\mu_p(s)\Pi_{\mathbf{x}_p}(s,t)ds. \quad \text{[S34]}$$

Thus,

$$\begin{aligned}\Phi_{\mathbf{x}_p}(t) = (1-f_p)e^{-\int_0^t \mu_p(u)du}&\Pi_{\mathbf{x}_p}(0,t) \\ &+ \int_0^t e^{-\int_s^t \mu_p(u)du}\mu_p(s)\Pi_{\mathbf{x}_p}(s,t)ds. \quad \text{[S35]}\end{aligned}$$

Notice $\Phi_{\mathbf{x}_p}(0) = 1 - f_p$. When all types are identical, $\Phi_{\mathbf{x}_1} = \cdots = \Phi_{\mathbf{x}_n} = \Phi_{\mathbf{x}_p}$, we may substitute the calculations above to obtain

$$\begin{aligned}\Phi_{\mathbf{x}_p}(t) = (1-f_p)e^{\int_0^t \lambda_p(u)(\Phi_{\mathbf{x}_p}(u)-1)-\mu_p(u)du}&\\ + \int_0^t e^{-\int_s^t \lambda_p(u)(\Phi_{\mathbf{x}_p}(u)-1)-\mu_p(u)du}&\mu_p(s)ds. \quad \text{[S36]}\end{aligned}$$

Differentiating left- and right-hand sides by t, we obtain the differential equation given in the text. Thus, when all types are identical, we recover our previous result.

Computing $\Psi_{\mathbf{x}_p}(s,t)$ is similar. Either the parent lineage survives to s, leaving no descendant lineage, which proceeding as above, has probability

$$e^{-\int_s^t \mu_p(u)du}\Pi_{\mathbf{x}_p}(s,t), \quad \text{[S37]}$$

or it dies at some $s < \sigma < t$, with exactly one of its descendant lineages having a unique descendant in the sample, which arguing as before, has probability

$$\begin{aligned}\int_s^t e^{-\int_\sigma^t \mu_p(u)du}\mu_p(\sigma)e^{-\int_s^t \lambda_p(u)du}\sum_{n=0}^{\infty}\frac{1}{n!}\sum_{i=1}^{n}&\left(\prod_{\substack{j=1\\ j\neq i}}^{n}\int_\sigma^t \lambda_{\mathbf{x}_i}(s_j)\Phi_{\mathbf{x}_j}(s_j)ds_j\right) \\ \times\int_\sigma^t \lambda_{\mathbf{x}_i}(s_i)\Psi_{\mathbf{x}_i}(s,s_i)ds_i.& \quad \text{[S38]}\end{aligned}$$

Again, it can be verified that this computation reduces to our previous expression when $\Phi_{\mathbf{x}_1} = \cdots = \Phi_{\mathbf{x}_n} = \Phi_{\mathbf{x}_p}$.

Likelihood When We Assume That the Rates only Change at a Known Set of Branching Points. As we saw above, without additional assumptions, such as assuming $\Phi_{\mathbf{x}_1} = \cdots = \Phi_{\mathbf{x}_n} = \Phi_{\mathbf{x}_p}$, no closed form is available for $\Phi_{\mathbf{x}_p}(t)$ or $\Psi_{\mathbf{x}_p}(s,t)$. However, if we assume that the rates only change at a known set of branching points, then all descendant lineages that appear in the expression for $\prod_{\mathbf{x}}(s,t)$, which leave no observed descendants, are assumed to be of the same form as the parent lineage. In this case, $\Phi_{\mathbf{x}_p}(t)$ and $\Psi_{\mathbf{x}_p}(s,t)$ reduce to the expressions found previously when rates were assumed to be homogeneous across lineages. We can then assign a pair $\boldsymbol{x} = (\lambda, \mu)$ for each subtree, forming the appropriate $\Psi_{\mathbf{x}}$ for all branches corresponding to species in that subtree and forming the likelihood as a product of all $\Psi_{\mathbf{x}}(s,t)$. By assumption, at each time t_i, a new species of type $\mathbf{x}_i$ appears $\boldsymbol{x}_i = (\lambda_i, \mu_i)$. We, thus, obtain the likelihood for the tree,

$$\frac{f_{\mathbf{x}_1}\Psi_{\mathbf{x}_1}(t_2,t_1)\prod_{i=1}^{n} f_{\mathbf{x}_i}\lambda_{p_i}(t_i)\Psi_{\mathbf{x}_{p_i}}(s_{i,1},t_i)\Psi_{\mathbf{x}_i}(s_{i,2},t_i)}{1-\Phi_{\mathbf{x}_1}(t_1)}, \quad \text{[S39]}$$

where the parent lineage p_i produces the descendant lineage i at time t_i, and $f_{\mathbf{x}_i}$ is the probability that species i is observed given that it survives to the present.

This approach is best illustrated by an example. In Fig. S11, we adopt the convention that the right branch is the new lineage, whereas the left branch is the parental line.

The corresponding likelihood function is

$$\frac{\Psi_{\mathbf{x}_1}(t_2,t_1)\lambda_{\mathbf{x}_1}(t_2)\Psi_{\mathbf{x}_1}(t_3,t_2)f_{\mathbf{x}_2}\Psi_{\mathbf{x}_2}(0,t_2)\lambda_{\mathbf{x}_1}(t_3)f_{\mathbf{x}_1}\Psi_{\mathbf{x}_1}(0,t_3)\Psi_{\mathbf{x}_3}(t_4,t_3)\lambda_{\mathbf{x}_3}(t_4)f_{\mathbf{x}_3}\Psi_{\mathbf{x}_3}(0,t_4)f_{\mathbf{x}_4}\Psi_{\mathbf{x}_4}(0,t_4)}{1-\Phi_{\mathbf{x}_1}(t_1)}. \quad \textbf{[S40]}$$

In computing the likelihood of a tree composed of species of different types, it is convenient to subdivide the tree. Suppose that the first lineage of type $\boldsymbol{x}_{\mathrm{cl}}$ appeared at time t_{cl} and that all descendants of this species are of type $\boldsymbol{x}_{\mathrm{cl}}$. We can form the likelihood function for the subtree using our expression for identical rates, conditioning on nonextinction by dividing by

$$1-\Phi_{\mathbf{x}_{\mathrm{cl}}}(t_{\mathrm{cl}}). \quad \textbf{[S41]}$$

After the best fit parameters are found for the subtree, the rates for the pruned parent lineage tree can be inferred by replacing terms corresponding to the subclade,

$$\lambda(T_{\mathrm{clade}})\Psi_{\mathbf{x}_{\mathrm{cl}}}(s_1,t_{\mathrm{cl}})\Psi_{\mathbf{x}_{\mathrm{cl}}}(s_2,T_{\mathrm{clade}})\ldots, \quad \textbf{[S42]}$$

with the probability that the clade leaves at least one descendant, $1-\Phi_{\mathrm{clade}}(T_{\mathrm{clade}})$, which may be calculated using the rates inferred from the subtree. The product of the likelihood of the subtree and the pruned tree is then exactly the likelihood of the whole tree. This procedure can then be applied recursively starting from terminal clades, which have no subclades, moving to the clades of which they are part, and moving back to the ancestor.

Stochastic Type Variation. We conclude by briefly considering a second approach that, although beyond the scope of our current study, has considerable promise. Suppose that we had a parametric model, $v_\theta(\boldsymbol{x}_p, \boldsymbol{x}')$, for the probability that a parent lineage of type $\boldsymbol{x}_p$ has a descendant lineage of type $\boldsymbol{x}'$. Then, the probability of having descendant lineages of type $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ at times $s_1, \ldots, s_n$ is

$$\frac{e^{-\int_s^t\lambda_p(u)du}}{n!}\lambda_p(s_1)\cdots\lambda_p(s_n)v_{\boldsymbol{\theta}}(\mathbf{x}_p,\mathbf{x}_1)\cdots v_{\boldsymbol{\theta}}(\mathbf{x}_p,\mathbf{x}_n)ds_1\cdots ds_n d\mathbf{x}_p\cdots d\mathbf{x}_n, \quad \textbf{[S43]}$$

and $\Pi_{\mathbf{x}_p}(s,t)$ reduces to

$$e^{\int_s^t\lambda_{\mathbf{x}_p}(s)\left(\int_{\mathbf{x}}\Phi_{\mathbf{x}_p}(s)v_{\boldsymbol{\theta}}(\mathbf{x}_p,\mathbf{x})d\mathbf{x}-1\right)-\mu(u)du}. \quad \textbf{[S44]}$$

Proceeding as before, we can obtain closed expressions for $\Phi_{\mathbf{x}_p}$, $\Psi_{\mathbf{x}_p}$, and the likelihood of observing a given tree as a function of the parameters θ. It is an interesting question for future study to determine an appropriate functional form of $v_\theta(x, x')$ and parameters θ.

1. Harmon LJ, Weir JT, Brock CD, Glor RE, Challenger W (2008) GEIGER: Investigating evolutionary radiations. *Bioinformatics* 24:129–131.
2. Alfaro ME, et al. (2009) Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proc Natl Acad Sci USA* 106:13410–13414.
3. Nee S, May RM, Harvey PH (1994) The reconstructed evolutionary process. *Philos Trans R Soc Lond B Biol Sci* 344:305–311.
4. Kendall DG (1948) On the generalized "birth-and-death" process. *Ann Math Stat* 19:1–15.
5. Maddison WP, Midford PE, Otto SP (2007) Estimating a binary character's effect on speciation and extinction. *Syst Biol* 56:701–710.
6. FitzJohn RG, Maddison WP, Otto SP (2009) Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. *Syst Biol* 58:595–611.
7. Ford D, Matsen FA, Stadler T (2009) A method for investigating relative timing information on phylogenetic trees. *Syst Biol* 58:167–183.
8. Jagers P (1975) *Branching Processes with Biological Applications* (Wiley, New York).

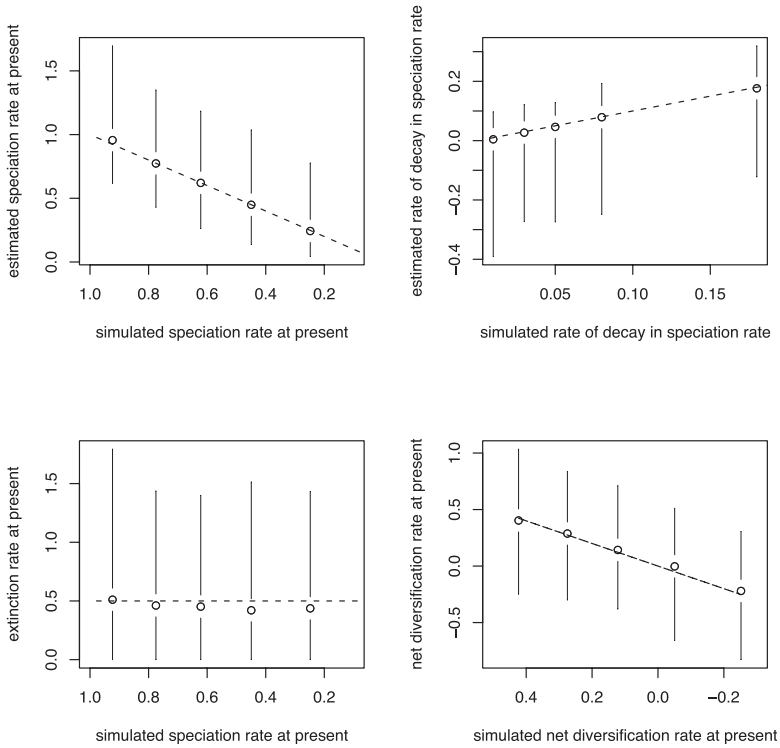

Fig. S1. Unbiased parameter estimates including negative net diversification. We simulated 1,000 phylogenies for each parameter set. The extinction rate was held constant in all five parameter sets (0.5 per time unit), whereas the speciation rate at present, rate of decay in speciation rate, and total simulation time varied across parameter sets. The true simulated parameters of diversification are indicated by dashed lines (expressed as number of events per time unit). The total simulation durations for each of the five parameter sets were (from left to right) 8, 8.5, 9.5, 10, and 10 (arbitrary time units; these durations were chosen to reduce variability in tree sizes across parameter sets). The corresponding mean tree sizes were (from left to right) 77, 57, 38, 18, and 14. Points and error bars indicate the median and 5% and 95% quantiles of the maximum likelihood parameter estimates (*Materials and Methods*). Unbiased estimates were obtained whether the net diversification rate at present was positive (corresponding to expanding clades; *Left*) or negative (corresponding to declining clades; *Right*).

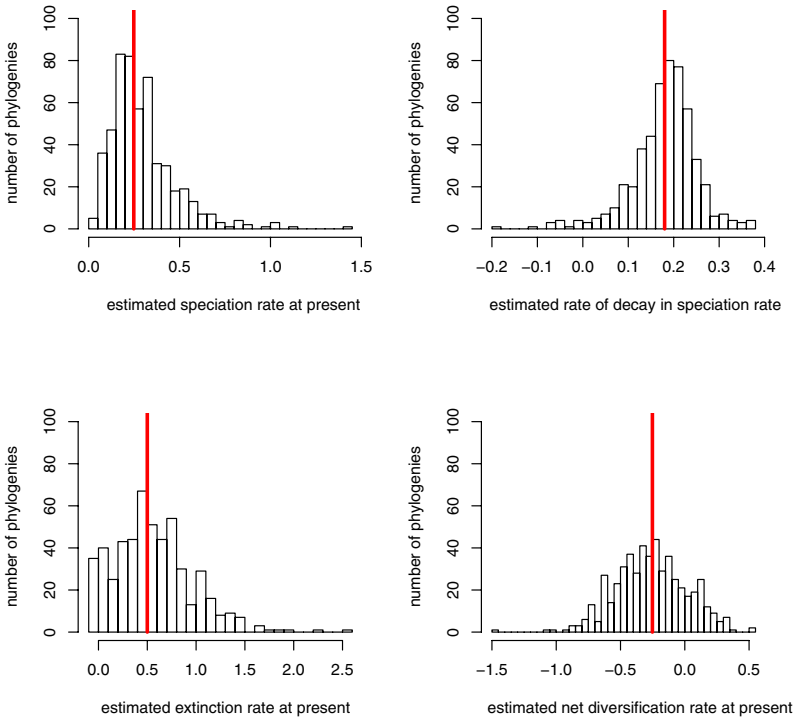

Fig. S2. The histograms represent the distribution of parameter estimates for the 889 phylogenies with at least 10 tips simulated with the parameters corresponding to the rightmost data point in Fig. S1. The red line indicates the true simulated parameters of diversification. The net diversification rate at present is correctly inferred to be negative in 81% of these phylogenies.

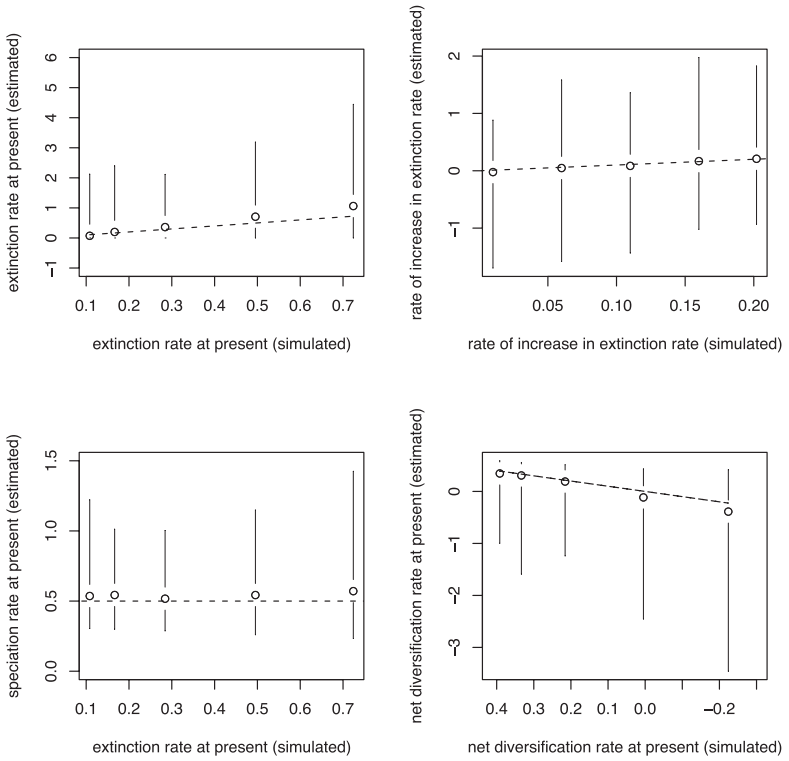

Fig. S3. Unbiased parameter estimates including negative net diversification. We simulated 1,000 phylogenies for each parameter set. The speciation rate was held constant in all five parameter sets (0.5 per time unit), whereas the extinction rate at present, rate of increase in extinction rate, and total simulation time varied across parameter sets. The total simulation durations for each of the five parameter sets were (from left to right) 8, 8.5, 9.5, 10, and 10 (arbitrary time units). The corresponding mean tree sizes were (from left to right) 29, 34, 45, 43, and 14. The true simulated parameters of diversification are indicated by dashed lines (expressed in number of events per time unit). Points and error bars indicate the median and 95% quantile range of the maximum likelihood parameter estimates. Unbiased estimates were obtained whether the net diversification rate at present was positive (corresponding to expanding clades; *Left*) or negative (corresponding to declining clades; *Right*).

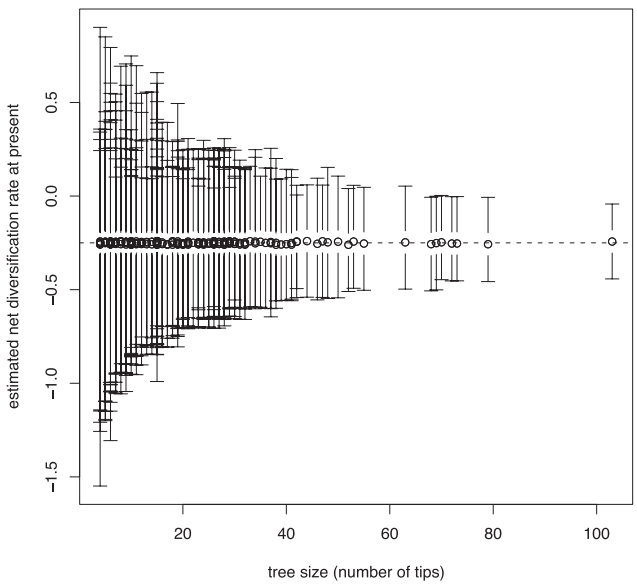

Fig. S4. Confidence intervals around the estimated net diversification rate at present as a function of tree size (number of tips). Each data point and error bar corresponds for a single phylogeny to the estimated net diversification rate at present and 95% confidence interval around it, respectively. Results for the 186 phylogenies with estimated net diversification rates at present falling within −0.26 and −0.24.

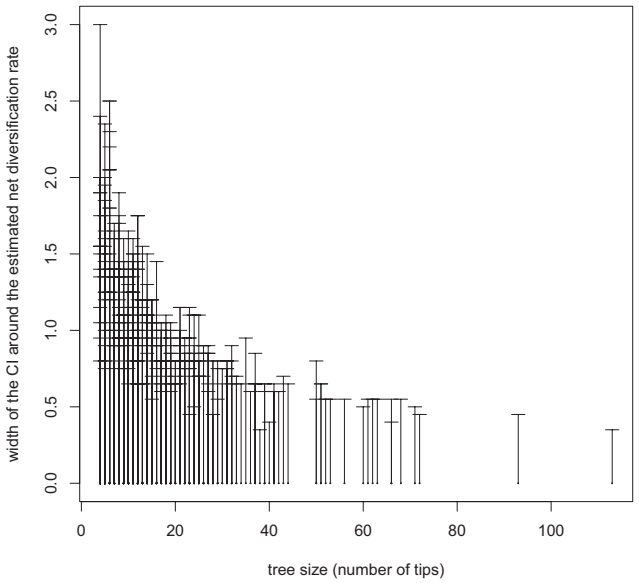

Fig. S5. Confidence intervals around the estimated net diversification rate at present as a function of tree size (number of tips). Each bar corresponds for a single phylogeny to the width of the 95% confidence interval around the maximum likelihood estimate. Results are for the 300 randomly sampled phylogenies.

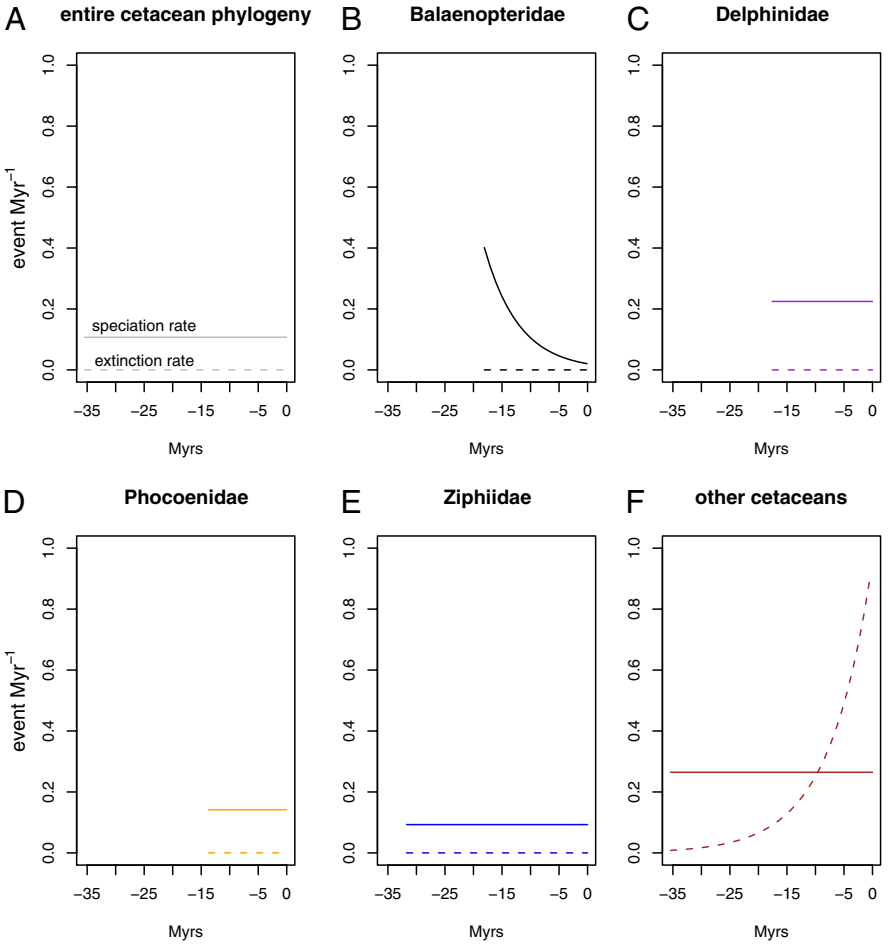

Fig. S6. Temporal variation in speciation and extinction rates in cetaceans. When the cetacean phylogeny is considered as a whole (*A*), the inferred net diversification rate is positive throughout the duration of the clade's history. However, when recently radiating clades (*B–E*; the Balaenopteridae, Delphinidae, Phocoenidae, and Ziphiidae) are isolated from the cetacean phylogeny, the phylogeny of the remaining cetaceans indicates an extinction rate that has increased over time, producing negative net diversification rates over the past ~10 Myr (*F*).

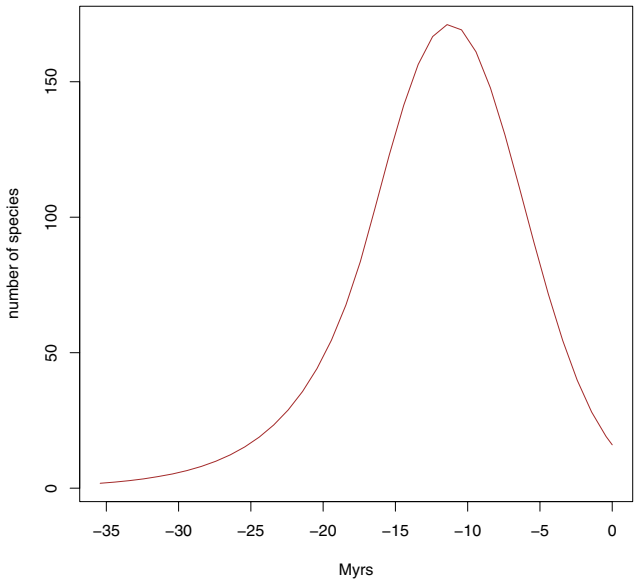

Fig. S7. The diversity trajectory of the cetaceans (excluding species from the Balaenopteridae, Delphinidae, Phocoenidae, and Ziphiidae families) inferred under the assumption that speciation and extinction rates vary linearly through time.

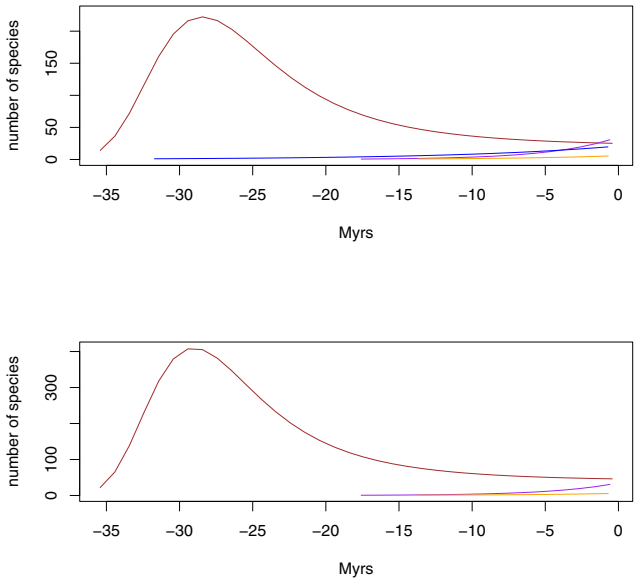

Fig. S8. The temporal variation in diversity obtained when considering the three-shift and two-shift models, respectively.

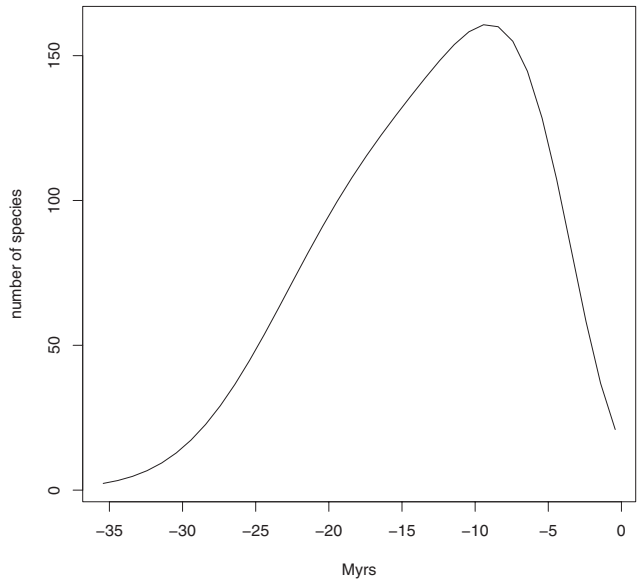

Fig. S9. The diversity trajectory of the cetaceans (excluding species from the Balaenopteridae, Delphinidae, Phocoenidae, and Ziphiidae families) inferred when analyzing the phylogenies of the odontocetes and mysticetes separately (obtained by summing these two individual trajectories). This diversity trajectory is qualitatively similar to the one inferred when analyzing the phylogeny of the odontocetes and mysticetes combined.

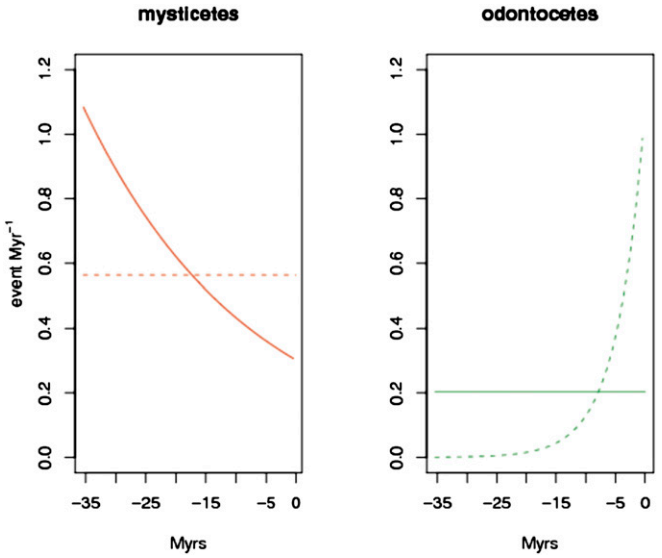

Fig. S10. Analysis of the mysticete phylogeny (excluding the Balaenopteridae) indicates that the speciation rate has decreased over time. Analysis of the odontocete phylogeny (excluding the Delphinidae, Phocoenidae, and Ziphiidae) indicates that the extinction rate has increased over time.

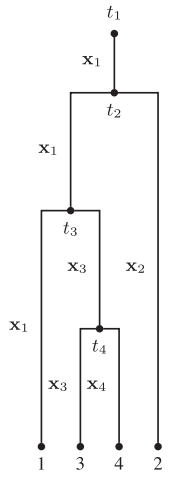

Fig. S11. Tree with types and speciation times. The i^{th} species, with type x_i, arises at time t_i, The types are indicated to the left of each branch, with the ancestral type on the left and the newly appearing species on the right. The leaves are labeled with the species number. Note that, although each descendant line has a new label, it may be that $x_2 = x_1$, etc.

Table S1. Various diversification models fitted to the cetacean phylogeny as a whole

Model	nb	Description	LogL	AICc
B constant	**1**	**No extinction and constant speciation rate**	**−279.0280**	**560.0796**
BD constant	2	Constant speciation and extinction rates	−279.0280	562.1270
B variable E	2	No extinction and exponential variation in speciation rate	−278.9887	562.0485
B variable L	2	No extinction and linear variation in speciation rate	−278.9896	562.0502
B variable E, D constant	3	Exponential variation in speciation rate and constant extinction rate	−278.9887	564.1204
B variable L, D constant	3	Linear variation in speciation rate and constant extinction rate	−278.9896	564.1221
B constant, D variable E	3	Constant speciation rate and exponential variation in extinction rate	−279.0280	564.1989
B constant, D variable L	3	Constant speciation rate and linear variation in extinction rate	−279.0280	564.1989

nb denotes the number of parameters of each model. LogL stands for the maximal log likelihood; AIC_c stands for the second-order Akaike's information criterion.

Table S2. Statistical support for rate shifts in the cetacean phylogeny

Model	nb	Description	LogL	AIC_c
No shift	1	Best fit model	−279.03	560.08
One shift	5	Best fit model: shift in the Delphinidae	−262.93*	536.22
Two shifts	6	Best fit model: shifts in the Delphinidae and Phocoenidae	−260.17†	532.85
Three shifts	7	Best fit model: shifts in the Delphinidae, Phocoenidae and Ziphiidae	−256.13‡	526.94
Four shifts	**8**	**Best fit model: shifts in the Delphinidae, Phocoenidae, Ziphiidae, and Balaenopteridae**	**−250.13**	**517.14**

nb denotes the number of parameters of each model; this number depends on the assumed number of shifts as well as the number of parameters in the best fit models for the corresponding subclades and pruned phylogeny. The best fit model for the pruned phylogeny was the BD variable model across the one-, two-, and three-shift models and the B-constant D-variable model for the four-shift model. LogL stands for the maximum log likelihood of the entire phylogeny; AIC_c stands for the second-order Akaike's information criterion (1). Footnotes indicate the statistical support of the corresponding model against the model with exactly one less rate shift based on the likelihood ratio test. Each additional shift is supported by both the AIC_c criterion and the likelihood ratio test. The four-shift model could not be compared with the three-shift model using the likelihood ratio test, because these models were not nested. In the most likely model (bold), each of the four largest families has different rates than the remainder of the tree.

*$P < 0.001$.

†$P < 0.05$.

‡$P < 0.01$.

1. Burnham KP, Anderson DR (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (Springer-Verlag, New York), 2nd Ed.

Table S3. Diversification models fitted to the cetacean phylogeny after isolating the Balaneopteridae, Delphinidae, Phocoenidae, and Ziphiidae families

Model	nb	Description	LogL	AIC_c
B constant	1	No extinction and constant speciation rate	−63.17	128.48
BD constant	2	Constant speciation and extinction rates	−63.17	130.77
B variable	2	No extinction and exponential variation in speciation rate through time	−61.49	127.41
B variable L	2	No extinction and linear variation in speciation rate	−58.57	121.56
B variable, D constant	3	Exponential variation in speciation rate and constant extinction rate	−56.76	120.40
B variable L, D constant	3	Linear variation in speciation rate and constant extinction rate	−58.45	123.79
B constant, D variable	**3**	**Constant speciation rate and exponential variation in extinction rate**	**−56.33**	**119.54**
B constant, D variable L	3	Constant speciation rate and linear variation in extinction rate	−56.60	120.10
BD variable	4	Exponential variation in speciation and extinction rates	−56.22	121.99
BD variable L	4	Linear variation in speciation and extinction rates	−56.39	122.31

nb denotes the number of parameters of each model. LogL stands for the maximum log likelihood; AIC_c stands for the second-order Akaike's information criterion. L stands for linear variation through time.

Other Supporting Information Files

Dataset S1 (TXT)