Supporting Information for

The role of clade competition in the diversification of North American canids

Daniele Silvestro[1,2,3], **Alexandre Antonelli**[1,4], **Nicolas Salamin**[2,3], **Tiago B. Quental**[5]

[1] *Department of Biological and Environmental Sciences, University of Gothenburg, Carl Skottsbergs gata 22B, 413 19 Gothenburg, Sweden;*

[2] *Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland;*

[3] *Swiss Institute of Bioinformatics, Quartier Sorge, 1015 Lausanne, Switzerland;*

[4] *Gothenburg Botanical Garden, Carl Skottsbergs gata 22A, 413 19 Gothenburg, Sweden;*

[5] *University of São Paulo (USP), Department of Ecology, São Paulo, SP, Brazil.*

Corresponding authors: silvestro.daniele@gmail.com (D.S.); tbquental@usp.br (T.B.Q.)

This PDF file includes:

1 Fossil data sets

We included in the final data sets only fossil occurrences found in North America and identified to a species level. Because we focused on a single continent, in some cases a species appearance might be the result of a migration event (i.e., the species originated elsewhere and subsequently dispersed into North America) and a species disappearance can be a local extinction (i.e. the species survived, but elsewhere). Although this is a general issue in diversification analyses based on fossil data, in this case it likely has a small effect since canids originated and diversified in the largest part in North America (1, 2)

The ages of most occurrences were provided with a temporal range, which reflects the uncertainty associated with the age of the fossil records and generally relies on the estimated boundaries of stratigraphic units (3). As an additional quality filter on the raw data, we excluded from the analyses the occurrences with temporal range larger than 15 Myr, which correspond to only about 5% of the data (Fig. S9).

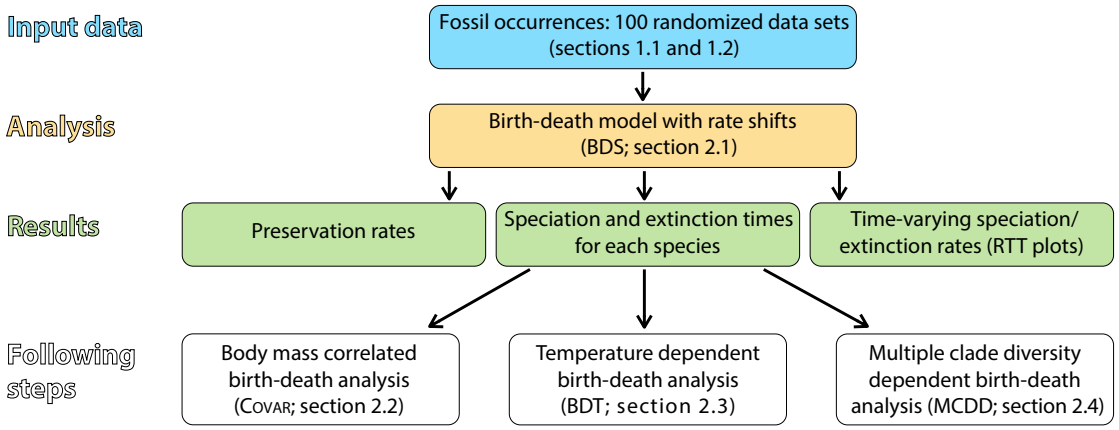
The final Canidae data sets included: 29 Hesperocyoninae species (269 occurrences), 68 Borophaginae species (829 occurrences), and 23 Caninae species (388 occurrences). The final data sets for putative competitor clades included (in addition to the Canidae subfamilies): 34 Amphicyonidae species (129 occurrences), 27 Ursidae species (205 occurrences), 13 Nimvravidae species (55 occurrences), 5 Barbourofelidae species (17 occurrences), and 36 Felidae species (338 occurrences). We treated the temporal ranges of the fossil occurrences as uncertainties around their estimated age. **In order to account for age uncertainties in our subsequent analyses, we generated 100 randomized data sets for each clade, by resampling the age of each occurrence uniformly within the respective temporal range**, following the procedure described by Silvestro *et al.* (4, 5). All the data are accessible from the Paleobiology database (http://paleobiodb.org) and the input files used for diversification rate analyses (see Supplementary Methods) are available from the authors (and will be provided at http://sourceforge.net/projects/pyrate/ upon acceptance of the paper).

2 Detailed methods and analysis settings

We carried out the analyses of the fossil data sets based on the Bayesian framework implemented in the program PyRate (4, 6). We first analyzed the fossil data sets under a birth-death model with time-varying rates and estimated for each clade: 1) the parameters of the preservation process, 2) the times of speciation and extinction of each species, and 3) the speciation and extinction rates and their variation through time (section 2.1). We then used the estimated times of speciation and extinction of all species to carry out three additional sets of analyses to test whether speciation and extinction rate dynamics correlate with 1) changes in body mass (section 2.2), 2) global temperature as a proxy for environmental changes (section 2.3), 3) competition and/or positive interaction between species through diversity dependence (section 2.4).

In the following sections we present the models used and provide flowcharts outlining the workflow followed in each analytical step to highlight the most relevant input, methodological details, and output. All methods shown here are implemented within the open source program PyRate (6), which is available at https://github.com/dsilvestro/PyRate.

2.1 Birth-death model with shifts (BDS)



We analyzed the fossil occurrence data sets using PyRate (4, 6) to **jointly estimate the preservation rate, the times of speciation and extinction for all species**, and the speciation and extinction rates through a Markov Chain Monte Carlo (MCMC) algorithm. In all analyses, we used the birth-death MCMC algorithm (BDMCMC; 4, 7) to simultaneously estimate the number and magnitude of shifts in speciation and extinction rates. Thus, the times of speciation and extinction used in all the subsequent analyses (sections 2.2–2.4) were obtained while accounting for the heterogeneity of preservation, speciation, and extinction rates.

Preservation process and the times of speciation and extinction

The preservation process was modeled as a non-homogeneous Poisson process (NHPP), in which the rate parameter is a function of time and the mean rate equals q and quantifies the expected mean number of fossil occurrences per lineage per million years (Myr). To allow for preservation heterogeneity among lineages, we used the Gamma model (4), which sets the rate of preservation to vary according to a gamma distribution with mean q and shape parameter α. Decreasing values of the shape parameter indicate increasing rate heterogeneity, e.g. with $\alpha = 20$ the gamma distribution has a variance of 0.05 whereas with $\alpha = 0.5$ the variance is 2.

Based on the NHPP model of preservation, we estimated, along with the parameters q and α, the times of speciation and extinction ($\mathbf{s}$ and $\mathbf{e}$) of N species using the following likelihood function (4):

$$P(X|\mathbf{s},\mathbf{e},q,\alpha) = \prod_{n=1}^{N}\sum_{k=1}^{K}\frac{P_{NHPP}(x_n|s_n,e_n,q,\alpha)}{K} \tag{1}$$

where $X = \{x_1, x_2, ..., x_N\}$ is a set of fossil occurrences for the N species, x_n a vector of fossil occurrences assigned to species n, and K is the number of discrete categories used to approximate the gamma distribution (in our analyses we used PyRate's default value $K = 4$ (4, 8)).

Speciation and extinction rates

We used a birth-death process with time-varying speciation and extinction rates as a prior on the times of speciation and extinction $\mathbf{s},\mathbf{e}$ (4). The parameters of a constant rate birth-death process are the speciation and extinction rates (λ and μ, respectively) and can be estimated through the following likelihood function:

$$P(\mathbf{s},\mathbf{e}|\lambda,\mu) = \lambda^B \mu^D e^{-(\lambda+\mu)S} \tag{2}$$

where B and D are the numbers of speciation and extinction events and S represents the total time lived by the N species (9), i.e. the sum of species lifespans:

$$S = \sum_{n=1}^{N} s_n - e_n. \tag{3}$$

We incorporated temporal variation of speciation and extinction rates by using a birth-death model with shifts (BDS), in which rates change across different time frames. The parameters estimated under this model are the vectors of speciation and extinction rates Λ, M and the temporal boundaries defining a vector of timeframes T. The number of time

frames is also jointly estimated by the BDMCMC algorithm. For a given set of T time frames the likelihood of a birth-death process is

$$P(\mathbf{s}, \mathbf{e}|\Lambda, M, T) = \prod_{\tau}^{T} \left[\lambda_{\tau}^{B_\tau} \exp\left(-\lambda_\tau S_\tau\right)\right] \prod_{\tau}^{T} \left[\mu_{\tau}^{D_\tau} \exp\left(-\mu_\tau S_\tau\right)\right] \quad (4)$$

where, for each time frame τ, B_τ and D_τ are the respective numbers of speciation and extinction events, λ_τ and μ_τ are the speciation and extinction rates, and S_τ is the fraction of the total time lived (Equation 3) observed within τ (4).

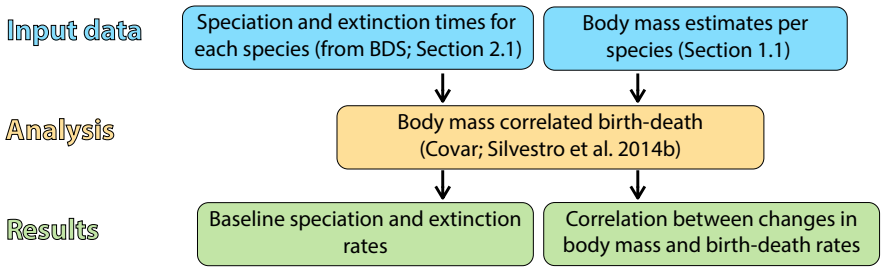
Analysis settings

We analyzed each data set (3 canid clades and 5 additional carnivore clades) independently under a birth-death model with rate shifts (BDS). We ran 1,250,000 BDMCMC iterations, discarded the first 250,000 as burnin, and sampled every $1,000^{th}$ to obtain posterior estimates of the parameters. We monitored chain mixing and effective sample sizes by examining the log files in Tracer (10). Proposal distributions and priors followed the default settings implemented in PyRate (6, 8). We replicated the analyses on the 100 randomized data sets of each clade (see section 1) and calculated estimates of $\mathbf{s}$ and $\mathbf{e}$ as the mean of the posterior samples from each replicate. Thus, **we obtained 100 posterior estimates of the times of speciation and extinction for all species and used them as input data in all the subsequent analyses** (sections 2.2–2.4), which, therefore, focused exclusively on the estimation of the birth-death parameters (i.e. without re-modeling preservation and re-estimating times of speciation and extinction). This procedure reduced drastically the computational burden, while allowing us to account for the preservation process and the uncertainties associated with the fossil ages.

The analyses performed under the BDS model also served to infer the temporal dynamics of the speciation and extinction rates, without any specific *a priori* hypothesis. We calculated the marginal speciation and extinction rates within 1 Myr time bins, while incorporating

model uncertainty, and summarized the results in rates-through-time (RTT) plots (Fig. 2; S10–S12) (4, 6).

2.2 Body mass correlated diversification (Covar model)



In the COVAR model, changes in speciation and extinction rates correlate with changes in body mass through the correlation parameters $(\alpha_\lambda, \alpha_\mu)$, which are estimated from the data (6). The birth-death rates are transformed on a lineage-specific . Thus, given a vector $\mathcal{B}$ of estimated body masses, we obtained the speciation and extinction rates of the i^{th} species with the following transformations:

$$\lambda(\mathcal{B}_i) = \exp\big(\log(\lambda_0) + \alpha_\lambda \log(\mathcal{B}_i)\big), \quad \mu(\mathcal{B}_i) = \exp\big(\log(\mu_0) + \alpha_\mu \log(\mathcal{B}_i)\big) \tag{5}$$

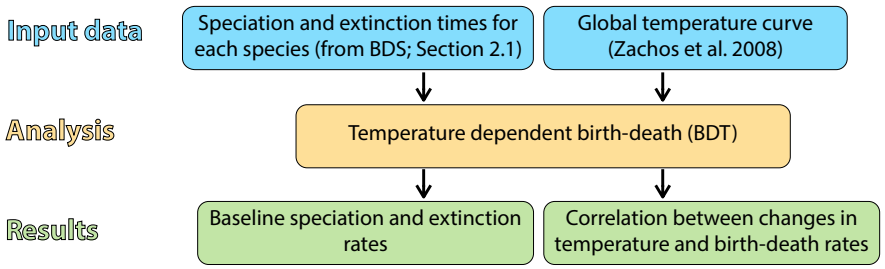
where λ_0, μ_0 are estimated baseline speciation and extinction rates, and the parameters $\alpha_\lambda, \alpha_\mu \in \mathbb{R}$ represent the coefficients of correlation between the rates and the log-transformed trait. We rescaled the body mass values by dividing them by their mean prior to the analyses, to obtain baseline speciation and extinction rates that represent the rates for a species of body mass equal to the mean size of all species within the clade (6). The likelihood of the COVAR birth-death process (Eqn. 2) is:

$$P(\mathbf{s}, \mathbf{e}, \mathcal{B} | \lambda_0, \mu_0, \alpha_\lambda, \alpha_\mu) = \prod_i^N \left[\lambda(\mathcal{B}_i)\mu(\mathcal{B}_i) e^{-\left[\lambda(\mathcal{B}_i)+\mu(\mathcal{B}_i)\right](s_i - e_i)} \right]. \tag{6}$$

Analysis settings and parameter interpretation

We analyzed each of the three canid data sets independently under the COVAR model and ran 100 analyses fixing the times of speciation and extinction to the ages estimated from the 100 replicated data sets under the BDS model (Section 2.1). We used the default gamma priors on the baseline speciation and extinction rates and normal priors $\mathcal{N}(0, 2)$ on the correlation parameters α_λ and α_μ (8). We ran 1,250,000 MCMC iterations with sampling frequency of 1,000 and combined the posterior samples of the parameters from the 100 replicates after excluding the first 250 samples as burnin. Posterior samples of the parameters were summarized over all replicates as mean values and 95% HPD. We considered the correlation to be statistically significant when 0 was not included within 95% of the highest posterior density (HPD) of α (e.g. 11).

2.3 Temperature dependent BD model (BDT)



We developed upon the birth-death model with time-varying rates described in Equation (4) to implement a correlation between speciation and extinction rates in canids and changes in global temperature through time, derived from stable isotope proxies (12). In this temperature-dependent birth-death model (hereafter referred to as BDT model) speciation and extinction rates for a given time frame τ are calculated based on the following transformations:

$$\lambda_\tau = \lambda_0 \exp(\gamma_\lambda \theta_\tau) \text{ and } \mu_\tau = \mu_0 \exp(\gamma_\mu \theta_\tau) \quad (7)$$

where θ_τ is the temperature value at time τ, λ_0, μ_0 are estimated baseline speciation and

extinction rates, and $\gamma_\lambda, \gamma_\mu$ are the parameters quantifying the correlation between changes in birth-death rates and temperature changes (13). We rescaled the temperature values by dividing them by their mean over the past 50 Ma (θ_{mean}), so that $\lambda(\theta_{mean}) = \lambda_0 \exp(\gamma_\lambda)$ and $\mu(\theta_{mean}) = \mu_0 \exp(\gamma_\mu)$.

In the BDT model, the time frames T of Equation (4) are not estimated from the data, but dictated by the time resolution of the temperature curve, in this case 0.1 Myr. The likelihood of the BDT model is therefore:

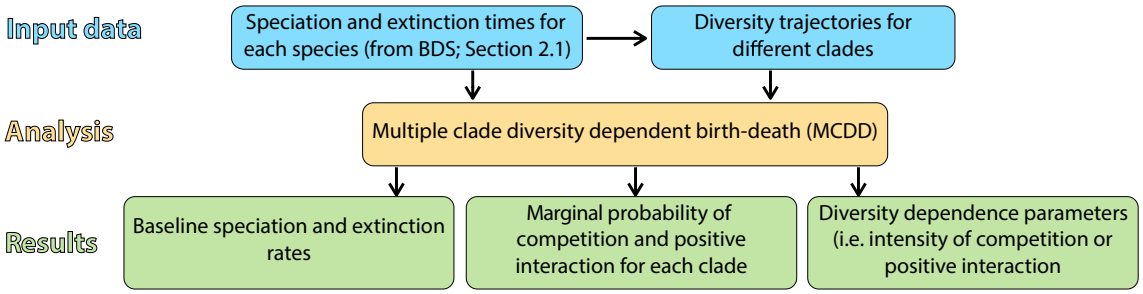
$$P(\mathbf{s}, \mathbf{e}, \Theta | \lambda_0, \mu_0, \gamma_\lambda, \gamma_\mu) = \prod_{\tau}^{T} \left[\lambda_\tau^{B_\tau} \exp\left(-\lambda_\tau S_\tau\right)\right] \prod_{\tau}^{T} \left[\mu_\tau^{D_\tau} \exp\left(-\mu_\tau S_\tau\right)\right] \quad (8)$$

where $\mathbf{s}$ and $\mathbf{e}$ are the speciation and extinction times for each species (as estimated from the BDS analyses; Section 2.1) and Θ is the vector of rescaled temperature values.

Analysis settings and parameter interpretation

We analyzed each of the three canid data sets independently under the BDT model and ran 100 analyses fixing the times of speciation and extinction to the ages estimated from the 100 replicated data sets under the BDS model (Section 2.1). We used gamma priors on the baseline speciation and extinction rates and normal priors $\mathcal{N}(0, 2)$ on the correlation parameters γ_λ and γ_μ (8). We ran 1,250,000 MCMC iterations with sampling frequency of 1,000 and combined the posterior samples of the parameters from the 100 replicates after excluding the first 250 samples as burnin. Posterior samples of the parameters were summarized over all replicates as mean values and 95% HPD. As for the parameters of the COVAR model, we considered the correlation to be statistically significant when 0 was not included within the 95% HPD of γ_λ and γ_μ.

2.4 Multiple Clade Diversity Dependence model (MCDD)



We implemented a birth-death model with diversity dependence (14) by combining the time varying birth-death model (Equations 4,8) with the diversity trajectories of sampled species (see below). We allowed for competition to take place not only among the species of a given clade, but also among species which are not closely related but share similar ecology. Therefore, **we extended the diversity-dependent birth-death model to assess the effects of competition within and between clades** and named this model Multiple Clade Diversity Dependence (MCDD). Estimating diversity dependence under the MCDD model involved the joint analysis of the three canid clades along with the five additional carnivore clades (Section 1). We obtained the diversity trajectories of sampled species for each clade by calculating at every point in time the number of living species (Fig. 4A; S13–S15) based on the times of speciation and extinction ($\mathbf{s}$, $\mathbf{e}$) estimated under the BDS model (Section 2.1). We calculated 100 diversity trajectories from the 100 replicated analyses under the BDS model (Section 2.1) The estimation of past species diversity might be biased by low preservation rates, taxonomic uncertainties, instances of anagenetic speciation (15–19). However, such trajectory curves are likely to provide a reasonably accurate representation of the true diversity changes in carnivore clades (20), in particular in North America, where the fossil record of mammals has been intensively sampled, studied, and revised (e.g. 2, 21–24).

Birth-death with diversity dependence

The MCDD model assumes that the birth-death rates of a clade varies through time as a function of its own diversity and/or the diversity of other competing clades. We indicate a trajectory curve of a clade i by δ_i, so that $\delta_i(t)$ represents the number of species from clade i that are living at time t. The speciation and extinction rates of a clade are linearly transformed as described by Etienne *et al.* (14). Thus, if we consider a set of C clades with trajectory curves $\mathcal{D} = \{\delta_1, \delta_2, ..., \delta_C\}$, the speciation and extinction rates for a given clade i at time t are obtained by the following transformations:

$$\lambda_i(t) = \max\left\{0,\ \lambda_i - \sum_{j=1}^{C} \lambda_i\Big[\delta_j(t) g_{ij}^{\lambda}\Big]\right\} \tag{9}$$

and

$$\mu_i(t) = \max\left\{0,\ \mu_i + \sum_{j=1}^{C} \mu_i\Big[\delta_j(t) g_{ij}^{\mu}\Big]\right\},$$

where λ_i, μ_i are baseline speciation and extinction rates of clade i, and $g_{ij}^{\lambda}, g_{ij}^{\mu}$ are the diversity dependence parameters transforming speciation and extinction rates, respectively. In the case of $i = j$, the diversity dependence parameter measures the effect of the diversity of a clade on its own speciation and extinction rates (self diversity dependence). **Each g parameter expresses a diversity dependence relationship between the diversity of a clade and the speciation/extinction rates of another clade.** Thus, the model is able to infer directionality of the reciprocal interactions between two clades, since they are modeled by two independent parameters (e.g. g_{ij}, g_{ji}).

Competition under the MCDD model results in a reduction of a clade's speciation rate and an increase of its extinction rate (14). This translates into $g_{ij}^{\lambda} > 0$ and $g_{ij}^{\mu} > 0$ indicating that the diversity of clade j negatively correlates with the speciation rate of clade i and causes the opposite (i.e., positive) effect on its extinction rate. On the contrary, $g_{ij}^{\lambda} < 0$ and $g_{ij}^{\mu} < 0$ indicate a **positive interaction** between clades, so that increasing diversity of a clade j

correlates with higher speciation rates and lower extinction rates in clade i. Finally, $g_{ij}^{\lambda} = 0$ and $g_{ij}^{\mu} = 0$ imply that no diversity dependent effects are detected and the diversification dynamics of clade i are independent of the diversity of clade j. In the case of clades with only partial temporal overlap (e.g. Borophaginae and Felidae; Fig. 4A), competitive effects or positive interactions can only occur when the diversity of both clades is greater than 0 and $g \neq 0$. The rates outside of the overlap time window or with $g = 0$ equal the baseline rates, based on Equation (9). Thus, **the baseline rates represent the speciation and extinction rates for a clade in the absence of diversity dependence (competition or positive interaction)**.

Based on the definitions given by the Equation (9), the intensity of diversity dependence (g) quantifies the proportion of rate change associated with the addition of one species in the competing clade. For instance, the competing effects given by $g_{ij}^{\lambda} = 0.1$ and $g_{ij}^{\mu} = 0.2$ imply that the addition of one species in clade j will decrease the speciation rate in clade i by 10% of the baseline rate (λ_i) and increase its extinction rate by 20% of the baseline rate (μ_i). Conversely, the extinction of a species in clade j will increase clade's i speciation rate and decrease its extinction rate by 10 and 20%, respectively. Opposite effects result from positive interaction, that is with $g < 0$.

In the MCDD model the birth-death likelihood function is based on the time varying birth-death model (Equation 4), where the time frames T are not estimated from the data, but determined by changes in the diversity trajectories. The likelihood of the MCDD model for a clade i is therefore:

$$P(\mathbf{s}, \mathbf{e}, \mathcal{D} | \lambda_i, \mu_i, g^{\lambda}, g^{\mu}) = \prod_{\tau}^{T} \left[\lambda_{\tau}^{B_{\tau}} \exp\left(-\lambda_{\tau} S_{\tau}\right) \right] \prod_{\tau}^{T} \left[\mu_{\tau}^{D_{\tau}} \exp\left(-\mu_{\tau} S_{\tau}\right) \right] \tag{10}$$

where $\mathbf{s}$ and $\mathbf{e}$ are the speciation and extinction times for each species (as estimated from the BDS analyses; section 2.1), $\mathcal{D} = \{\delta_1, \delta_2, ..., \delta_C\}$ is the set of diversity trajectories of all clades, λ_{τ} and μ_{τ} are the speciation and extinction rates at time τ transformed according to Equation

(9), and g^{λ}, g^{μ} are vectors of diversity dependence parameters, e.g. $g^{\lambda} = \{g^{\lambda}_{i1}, g^{\lambda}_{i2}, ..., g^{\lambda}_{iC}\}$, for a set of C clades.

A hierarchical model with Bayesian variable selection

We designed an MCMC algorithm that samples the parameters of the MCDD model from their posterior distributions based on the following considerations:

1. The number of diversity dependence parameters (p_C) increases more than linearly with the number of clades ($p_C = 2C^2$, for $C \geq 2$) and for a large number of clades there is a risk of over-parameterization (e.g. with the 8 clades analyzed here: $C = 8, p_C = 128$). Furthermore, the definition of appropriate priors for the parameters of the MCDD model can greatly affect the dimension of the sampled parameter space. **The use of a hierarchical Bayesian model with hyperprior distributions on the priors is an appropriate way to potentially reduce the effective parameter space** (5, 25).

2. Competition or positive interaction might come from several clades simultaneously. Additionally, different clades can have similar diversity trajectories and potentially result in similar effects on diversification rates. Thus, a joint analysis of all clades is preferable over pairwise analyses. This allows us to assess if competitive effects or positive interactions have an overall significant impact on the diversification dynamics of each clade, while accounting for the combined effects of multiple clades. **Joint and hierarchical parameter estimation is also an efficient way to avoid the risks of false positives that could result from multiple comparisons (26)**.

3. The complexity of the model and a high number of potentially interacting clades make it difficult to assess whether there is significant evidence for competition and render explicit model selection impractical for the purpose, since too many alternative models would have to be tested. The question of finding which and how many clades affect

the dynamics of speciation and extinction rates of each clade can be seen as a question of selecting the variables in a regression problem. Hence, **the use of Bayesian variable selection provides a powerful tool to explore potentially all models and jointly estimate the probability and the intensity of each diversity dependence parameter** (27).

Based on these considerations, we implemented the MCDD model using the Bayesian variable selection algorithm originally developed by Kuo and Mallick (28) within the PyRate analytical framework. We transformed the parametrization of the MCDD model by decomposing the diversity dependence parameters (g) into two auxiliary variables, the indicator variable I and the effect size of diversity dependence k (27, 28), so that

$$g_{ij} = I_{ij}k_{ij} \text{ for } i, j \in C. \tag{11}$$

The indicator variable (I_{ij}) can only take values equal to 0 or 1, thus denoting whether or not the birth-death rates of clade i are affected by the diversity of clade j. The effect size (k_{ij}) represents the intensity of the effect of clade j's diversity on speciation/extinction rates of clade i when $I = 1$. Thus, there is a competitive effect when $I = 1$ and $k > 0$, whereas positive interaction results from $I = 1$ and $k < 0$.

We treated the two auxiliary variables as independent *a priori* and estimated them from the data (28). Since the indicators can only take two values, we assigned the following prior probabilities

$$P(I_{ij} = 0) = \eta_i, \tag{12}$$

$$P(I_{ij} = 1) = 1 - \eta_i, \text{ for } j \in C.$$

Thus, for a given clade i, each of the indicators $I_{i1}, ..., I_{iC}$ equals 0 with prior probability η_i and it equals 1 with prior probability $1 - \eta_i$. We considered the prior probability η_i as an

unknown hyperparameter and estimated it from the data. The hyperparameter η measures the prior probability that the diversification process of a clade is not affected by diversity dependence from any of the analyzed clades.

The novel parameterization outlined above presents several desirable features. First, it allows the exploration of many potential models ranging from the simplest case in which a clade diversifies under constant birth-death rates in the absence of diversity dependence (e.g. with $\eta_i = 1,\ g_{i1}, ..., g_{iC} = 0$) to the most complex scenario in which each clade is subject to either competition or positive interactions from all clades (i.e. $g_{i1}, ..., g_{iC} \neq 0$). While all these models have identical parameterization, the effective parameter space can differ greatly depending on the values of the indicator variables I that can potentially remove entirely the effect of diversity dependence on speciation and extinction rates. Thus, **the use of Bayesian variable selection reduces the risk of overparameterization**, since both I and η are estimated from the data (27) (see also simulations below).

Second, this method allows us to **readily assess whether there is significant evidence for diversity dependence (competition or positive interactions)** based on the posterior samples of the g parameters, thus avoiding expensive calculations of the fit of alternative models, e.g. by Bayes factors (29). Indeed, the posterior probabilities that a clade i is subject to diversity dependent birth-death dynamics (by competitive or positive effects) can be calculated directly from the sampling frequencies of $g > 0$ and $g < 0$ (see below).

Finally, the decomposition of the diversity dependence parameters (Equation 11) into the auxiliary variables effect size (k) and indicator (I) helps to **avoid the excessive shrinkage towards small values of the competitive/positive effects that are supported by the data**, while removing negligible effects (background noise) (27, 30, 31). This feature is desirable because a clade might be subject, for instance, to strong competition (i.e. large k) from some of the clades considered in the analysis, whereas several other clades do not play any role in the clade's diversification.

Implementation and (hyper)priors

We implemented the MCDD model in a hierarchical MCMC algorithm that jointly estimates, for each clade i, the baseline speciation and extinction rates (λ_i, μ_i), the diversity dependence parameters $(g_{i1}, ...g_{iC})$ through their effect size $(k_{i1}, ...k_{iC})$ and indicator components $(I_{i1}, ...I_{iC})$, and the hyperparameter η_i. For the baseline speciation and extinction rates of all clades we used an exponential prior with scale parameter estimated in the MCMC as a hyperparameter (with exponential hyperprior with mean equal to 10). The prior probabilities for the indicators are given in Equations (12) (see above) based on the hyperparameters η, which were assigned uniform hyperprior distributions $P(\eta_i) = \mathcal{U}[0, 1]$. Finally, we used a uniform prior distributions for the effect size parameters $P(k_{ij}) = \mathcal{U}[-0.3, 0.3]$. Hence, the addition of one species in the diversity trajectory of clade j can decrease or increase the speciation and extinction rates of clade i by up to 30% of its baseline rates (based on Equation 9). We selected the boundaries of this prior after empirical testing with the data set to allow all plausible values, while maintaining good mixing of the MCMC.

We used sliding window proposals with reflection at the boundaries to update the effect sizes, and the hyperparameters and multiplier proposals for the baseline speciation and extinction rates (32). Proposals for the discrete indicator variables (I) were generated by random updates where 0 and 1 had equal probability. The effect size parameters k were updated even when the respective indicator was set to 0, in which case the acceptance was only determined by their uniform prior probability (see above) (28).

MCDD output

The most important results generated by the MCDD model are: 1) the baseline speciation and extinction rates for each clade, 2) the marginal probability of competition or positive interaction for each clade, and 3) the intensity of the diversity dependence between each pair of clades. The baseline rates and intensity of the diversity dependence are directly sampled in the MCMC, based on the MCDD birth-

death likelihood (Equation 10).

The marginal probability of competition or positive interaction is a measure of statistical support for these effects and can be calculated from the posterior samples of the g parameters. In order to assess whether a clade's diversification history is significantly affected by competition or by positive interactions, we computed, for each clade i, the frequency at which positive or negative diversity dependence parameters from at least one of the C clades was sampled across the MCMC iterations. The marginal probability of competition is

$$P_{comp}(i) = \frac{\sum_{m=1}^{M} Z_m^+}{M} \tag{13}$$

where M is the number of posterior samples, and $Z_m^+ = 1$ when at least one of the diversity dependence parameters for clade i indicates competition (i.e. $\max\{g_{i1}, ..., g_{iC}\} > 0$) and $Z_m^+ = 0$ when all diversity dependence parameters equal 0 or indicate positive interaction (i.e. $\max\{g_{i1}, ...g_{iC}\} \leq 0$). Note that such marginal probability of competition includes the effects that a clade's own diversity can have on its speciation/extinction rate (g_{ii}). Conversely, the probability of positive interaction is

$$P_{posi}(i) = \frac{\sum_{m=1}^{M} Z_m^-}{M} \tag{14}$$

where $Z_m^- = 1$ when at least one of the diversity dependence parameters is negative (i.e. $\min\{g_{i1}, ..., g_{iC}\} < 0$) and $Z_m^- = 0$ when all the indicators diversity dependence parameters equal to 0 or positive (i.e. $\min\{g_{i1}, ..., g_{iC}\} \geq 0$). Such frequencies thus represent the marginal probabilities that the birth-death process correlates (by competition or positive interactions) with any of the trajectory curves $\mathcal{D}$, regardless of which or how many clades are involved in the diversity dependence (cf. consideration 2 above). These marginal probabilities are calculated independently for speciation and for extinction (based on g^λ and g^μ, respectively), thus allowing us to pinpoint which aspects of the diversification process are more likely to undergo diversity dependence.

Analysis of empirical data sets

After some test runs aimed at assessing the extent of burnin and the sampling efficiency, we ran 30,000,000 MCMC iterations with sampling frequency of 10,000 to obtain posterior parameter estimates. We repeated the analyses on the 100 replicates, using the times of speciation and extinction estimated under the BDS model (see section 2.1). For each of the 8 carnivore clades we computed mean and 95% HPD of the baseline speciation and extinction rates (λ_i, μ_i), the hyperparameter η_i and the diversity dependence parameters g_i^{λ}, g_i^{μ}. We used the mean of the sampled diversity dependence parameters (e.g. g_{ij}) as a measure of intensity of competition (if positive) or positive interaction (if negative) between each pair of clades. The respective probabilities of competition or positive interaction between pairs of clades were calculated as the sampling frequency of positive or negative values. Finally, we obtained the marginal probabilities of competitive effects and positive interaction (P_{comp}, P_{posi}) for both speciation and extinction using Equations (13,14). **We considered competition or positive interaction to be significant when they obtained marginal probabilities above 0.95** (Fig. 4).

2.5 Robustness of the MCDD model

We assessed the robustness of our MCMC algorithm with Bayesian variable selection by three sets of 1000 simulations. The complexity of diversity dependence among multiple clades renders the simulation of data the under MCDD model unfeasible. Instead, we used standard birth-death simulations (4) to quantify the frequency of finding significant competition or positive interactions from randomly diversifying clades (false positive rate).

In the first set of analyses, we simulated clades under a constant rate birth-death model, which we considered as a null model, since the MCDD reduces to constant rate birth-death when removing the effect of competition and positive interactions (Equation 9). We simulated 1,000 independent clades aimed to match the size, rates, and lifespan of the carnivore

clades in our empirical data set. For each clade we drew the root age randomly from a uniform distribution $\mathcal{U}[15, 40]$ Ma (reflecting the temporal range of clade origins in our data; Figs. S13–S15) and sampled speciation and extinction rates from uniform distributions $\mathcal{U}[0.01, 0.4]$ (similarly to the range of rates estimated from our empirical data set under BDS; Figs. S10–S12). We also constrained the size of the simulated clades to range between 5 and 70 species to resemble the clade sizes in our empirical data set. We analyzed the clades under the MCDD model after grouping them in random groups of 8 clades. For each analysis we ran 15,000,000 MCMC iterations with sampling frequency of 5,000 and then calculated the marginal probabilities of competition and positive interaction (P_{comp}, P_{posi}) for both speciation and extinction in each clade (Equations 13,14). Based on a threshold of 0.95 we calculated the frequency of finding significant diversity dependence overall clades (F_{comp}, F_{posi}), which we considered as a 'false positive'. **The 'false positive' frequencies were below 2% for both competition and positive interaction, indicating that the MCDD model is robust against type I error.** The frequencies of significant diversity dependence were $F_{comp}(\lambda) = 0.018$ and $F_{comp}(\mu) = 0.008$ for competition, and $F_{posi}(\lambda) = 0.003$ and $F_{posi}(\mu) = 0.003$ for positive interaction.

We performed a second set of simulations (1,000 independent clades), in which we accounted for an amount of variation in speciation and extinction rates, reflecting the amount of rate variation reconstructed from the 8 carnivore clades (Figs. S10–S12). We introduced rate variation through rate shifts, i.e. using the BDS model (section 2.1). Since in most of the empirical clades analyzed here we found evidence for either constant rates or for one shift in speciation and/or extinction (Table S2), we randomly drew the number of rate shifts for each simulated clade from a Poisson distribution with rate parameter equal to 1. This assigns the highest probabilities to zero shifts (constant rates) and one shift, but allows for higher numbers of rate shifts. We sampled the temporal placement of the shifts (if any) from random uniform distributions between the root and the present and the speciation and extinction rates between shifts from random uniform distributions $\mathcal{U}[0.01, 0.4]$. The other

settings (root ages, clade sizes) were maintained as in the first set of simulations. The random birth-death settings used in these simulations can generate casual diversity dependence, since the diversity trajectory of a clade can randomly correlate with rate changes in another clade. The frequencies of significant diversity dependence from this set of simulations were nevertheless below 5% for both competition and positive interaction: $F_{comp}(\lambda) = 0.041$ and $F_{comp}(\mu) = 0.037$ for competition, and $F_{posi}(\lambda) = 0.016$ and $F_{posi}(\mu) = 0.024$.

Finally, we tested the performance of the MCDD model in a scenario in which speciation and extinction rates undergo very frequent variations through time. We simulated under BDS model 1,000 independent clades following the settings described above, but drawing the number of rate shifts from a Poisson distribution with rate parameter equal to 10. This yields on average 10 rate shifts in speciation and 10 in extinction, which is about 20 times more rate variation than estimated from the 8 carnivore clades analyzed here. Since diversity dependence implies rate changes correlated with diversity changes of a clade, we expect this simulation scenario to increase even further the chances of random correlations between rates and diversity trajectories of clades. The frequencies of significant diversity dependence were in this case higher than in the other simulations, although they remained below 10%: $F_{comp}(\lambda) = 0.084$ and $F_{comp}(\mu) = 0.064$ for competition, and $F_{posi}(\lambda) = 0.024$ and $F_{posi}(\mu) = 0.044$. These frequencies are likely to capture casual diversity dependence correlations, showing that, when clades diversify under highly variable (but random) birth-death rates, the chances of fortuitously reproducing diversity dependence increase slightly. However, given the limited amount of rate variation found across the clades in our data set (at most one rate shift), we consider this possibility as unlikely to affect our results.

3 Supplementary results and discussion

3.1 Preservation, speciation, and extinction rates (BDS model)

The mean preservation rates, averaged over 100 replicates, ranged across clades between 0.67 and 8.04 expected fossil occurrences per lineage per Myr, and the estimated heterogeneity parameters indicated considerable rate variation among the species of most clades (Table S1). **Under this range of preservation rates, we can expect speciation and extinction rates to be reliably inferred in the birth-death analyses**, as shown through extensive simulations by Silvestro *et al.* (4).

The BDS analyses identified rate shifts of both speciation and extinction rates in the two extinct clades, Hesperocyoninae and Borophaginae. Indeed, both a decrease in speciation rates and an increase in extinction rates were inferred throughout their evolutionary history (Figs. 1, S10). We did not find evidence of significant rate variation through time in the subfamily Caninae, despite large uncertainties on the rate estimates during the initial phase of their diversification (until around 18 Ma), and a weak tendency towards increased extinction rates toward the present (Figs. 2, S10). Among the other carnivore lineages included in the analyses, we found one shift in the speciation rate of Amphicyonidae (rate decrease at ca. 19 Ma) and one shift in the extinction rate of Felidae (increased extinction in the past 4 Myr). Rates were roughly constant in all other cases (Figs. S11, S12). The sampling frequencies of birth-death models with estimated number of rate shifts are provided in Table S2.

3.2 Diversification and body mass evolution (Covar model)

Body mass appears to have increased through time in each of the three canid subfamilies (Fig. 3A), although a relatively large range of body sizes was maintained over time. Large species, however, seem to have appeared at different times in the three clades. Some Hesperocyoninae species exceeded 10 Kg around 35 Ma, whereas similar sizes were reached by Borophaginae

only around 25 Ma. The latter reached sizes of 40 Kg between 20 and 8 Ma. Moderately large species of Canidae (> 10 Kg) appeared around 13–10 Ma.

The birth-death analyses under the Covar model did not recover evidences of correlations between speciation/extinction rates and changes in the body size (Fig. 3B; Table S3). This was evident from the correlation parameters (α_λ and α_μ), which we found to be approximately centered on zero in all clades, with HPDs intervals spanning both positive and negative values.

3.3 Correlations with climate change (BDT model)

The BDT model based on a global temperature curve revealed a trend towards positive correlation with speciation rates (i.e. $\gamma_\lambda > 0$) and negative correlation with extinction rates (i.e. $\gamma_\mu < 0$) in all three canid subfamilies (Fig. 3D). However, **the correlation was not significant in most of the cases**, based on the 95% HPD intervals (Table S4). The only significant correlation was found with the extinction rate of Borophaginae, suggesting that its increase through time (Fig. S10) might be associated with the global cooling that started after the Middle Miocene Climatic Optimum (ca. 18–16 Ma).

3.4 Competition and positive interactions (MCDD model)

The MCDD model applied to the 8 North American carnivore clades involved the estimation of the diversity dependence parameters (g) correlating the speciation and extinction rates of each lineage with the different diversity trajectories. The reconstructed diversity trajectories for each clade, obtained from the BDS analyses, are shown in Figs. 4A; S13–S15. The marginal probabilities of positive interaction calculated for speciation and for extinction rates were lower than the 0.95 significance threshold in all clades (Table S5). Thus, **we did not find evidence suggesting that the diversification of a clade profited from the diversity of any of the 8 clades**. This result conforms to our expectations, since all the clades considered here belong to the broad category of large mammalian carnivores, and

are thus expected to share similar ecology and potentially compete among each others. In contrast, **the MCDD model detected strong competitive effects on both speciation and extinction rates in canid subfamily Hesperocyoninae and on extinction in Borophaginae**, with marginal probabilities greater than 0.95. We found much lower marginal probabilities of competition in Caninae and no evidence of significant competition among the other clades (Table S5). We emphasize that positive interaction and competition are not mutually exclusive in the joint analysis of multiple clades. For instance, the diversification dynamics of a given clade could be positively correlated to the diversity of another clade, while being subject to significant competition from a third clade.

The evidence that 2 of the 3 canid subfamilies have their dynamics affected by competition, as obtained from the MCDD model, implies that the respective speciation and/or extinction rates significantly correlate with one or more diversity trajectories (diversity dependence). We summarized the entire set of parameters estimated under the MCDD model for each clade in the Tables S6–S13. The intensity of diversity dependence among clades (i.e. competition when $g > 0$, positive interaction when $g < 0$) shows that there is a large heterogeneity in the strength of these interactions and that different relationships link the dynamics of speciation and extinction rates, respectively (Figs. 1, S16). Additionally, the probabilities of the individual competitive effects between clades (independently of their intensity), are usually moderate or low (Tables S6–S13, columns $P(g < 0), P(g > 0)$) even in clades where significant competition is supported by marginal probabilities. This uncertainty might reflect the difficulty to tease apart the effect of individual clades with similar diversity curves and can be the result of weak and sparse signal of diversity dependence from different clades. On the other hand, the use of Bayesian variable selection and the hierarchical structure of the MCDD model ensure that such uncertainty be fully maintained in the credible intervals of the parameters (26–28). The effect of 'self-diversity dependence', here intended as within-clade competition (14), appears to be overall quite weak, except for speciation rates in Barbourofelidae (where, however, competition is not overall significant;

Figs. S16, Table S12), and is negligible in all clades for extinction rates.

Rate changes induced by competition

In order to visually assess the amount of rate variation induced by competition, we plotted the effects of 1 to 10 species of each clade on the speciation and extinction rates of Hesperocyoninae, Borophaginae, and Canine, using the estimated baseline rates and the intensities of diversity dependence. We calculated both the absolute rate variation (Fig. S17) and the relative variation, expressing the change as a percentage of the baseline rate (Fig. S18). We emphasize that these plots only represent theoretical rate changes considering only the effects of each individual clade, i.e. ignoring the combined effect of all clades that is instead accounted for in the MCDD model (see Equation 9). **These plots help visualizing which clades have the most important competitive effects and how large are the effects.** For instance, the addition of one species of Borophaginae correlates with an 4% decrease in the speciation rate of Heperocyoninae (Table S6) and addition of one species of Felidae correlates with a 13% increase in extinction rate of Borophaginae (Table S7).

3.5 Passive replacement and active displacement

The sequential succession of competing clades in life history has been traditionally explained by two main biological mechanisms: passive replacement and active displacement (2, 33–35). **In the case of passive replacement, an incumbent clade initially prevents a competing clade from radiating.** The latter clade can diversify at high net rates only after the incumbent clade declines, for instance due to extrinsic factors such as climate change, thus freeing ecological niche space. **In contrast, active displacement occurs when the rise in diversity of a clade drives the decline of another clade by outcompeting it on limited resources.** The hypotheses of passive replacement and active displacement between co-occurring clades have been discussed in terms of how much temporal overlap there is between potential competitors, suggesting that highly overlapping clades should indicate

an active competition process, while passive replacement might be characterized by a lag between the decline of the incumbent clade and the rise of the more recent clade (33). However, the two processes should also leave different signatures on the speciation and extinction rates. Under passive replacement, we predict that the decline of the incumbent clade correlates with an increase of net diversification rate (by increased speciation and/or decreased extinction; Fig. S19) in the competing clade, which has the ecological and evolutionary opportunity to expand and diversify. Conversely, active replacement involves a decrease of net diversification rates in a clade (by decreased speciation and/or increased extinction) linked to the raise of another clade, from which it is actively outcompeted. Thus, the diversification of the latter clade might remain unchanged (Fig. S19). Distinguishing between these two alternative processes is difficult by only looking at the diversity trajectories of clades and their overlap. For instance, if we considered only the diversity trajectories of Borophaginae and Caninae (Fig. S13), we could imagine that, as Borophaginae declined, they left empty niche space that could be filled by Caninae , leading to a recent radiation. Alternatively the decline of Borophaginae could have been caused by the increasing diversity of Caninae, which in turn might show no simultaneous increase in their net diversification rates.

A process-based approach that models the correlation between diversity and speciation and extinction rates, such as the MCDD model, can help us to better understand the evolutionary replacement of clades. **Although the MCDD model does not explicitly model the two scenarios outlined above as different processes, it still allows us to distinguish between passive replacement and active displacement.** This is possible by examining both the estimated correlation parameters and the diversity trajectory of the clades. If we consider two temporally sequential but overlapping clades, e.g. clades 1 and 2 in Fig. S19, a passive replacement should result in a diversity dependent relation between the diversity trajectory of the declining incumbent clade 1 and the net diversification rate of the new clade ($g_{21} > 0$). In this scenario the negative net diversification rate driving the decline of the incumbent clade does not correlate with the diversity trajectory of clade 2

($g_{12} \approx 0$). On the contrary, in the case of active displacement, the net diversification rates of the older clade should decrease with the rising diversity of competing clade 2 ($g_{12} > 0$). In this case, the increasing diversity of clade 2 does not need to be a consequence of diversity dependence mechanisms ($g_{21} \approx 0$).

Our results strongly support the role of active displacement within the Canidae family, where older clades undergo a decrease in speciation rates and increase in extinction rates both correlating to the diversity trajectories of more recent clades. This is observed for Hesperocyoninae, which appear to be actively displaced by Borophaginae and Felidae, and for Borophaginae, which itself experienced active displacement especially from Caninae and Felidae (Figs. S10, S16). In contrast, diversification rates of Caninae did not appear to profit from the decline and extinction of the older clades Hesperocyoninae and Borophaginae, as they had essentially constant speciation and extinction rates through time and we detected no competition from older clades (Figs. S10, S16).

4 Supplementary Figures (S1–S19)

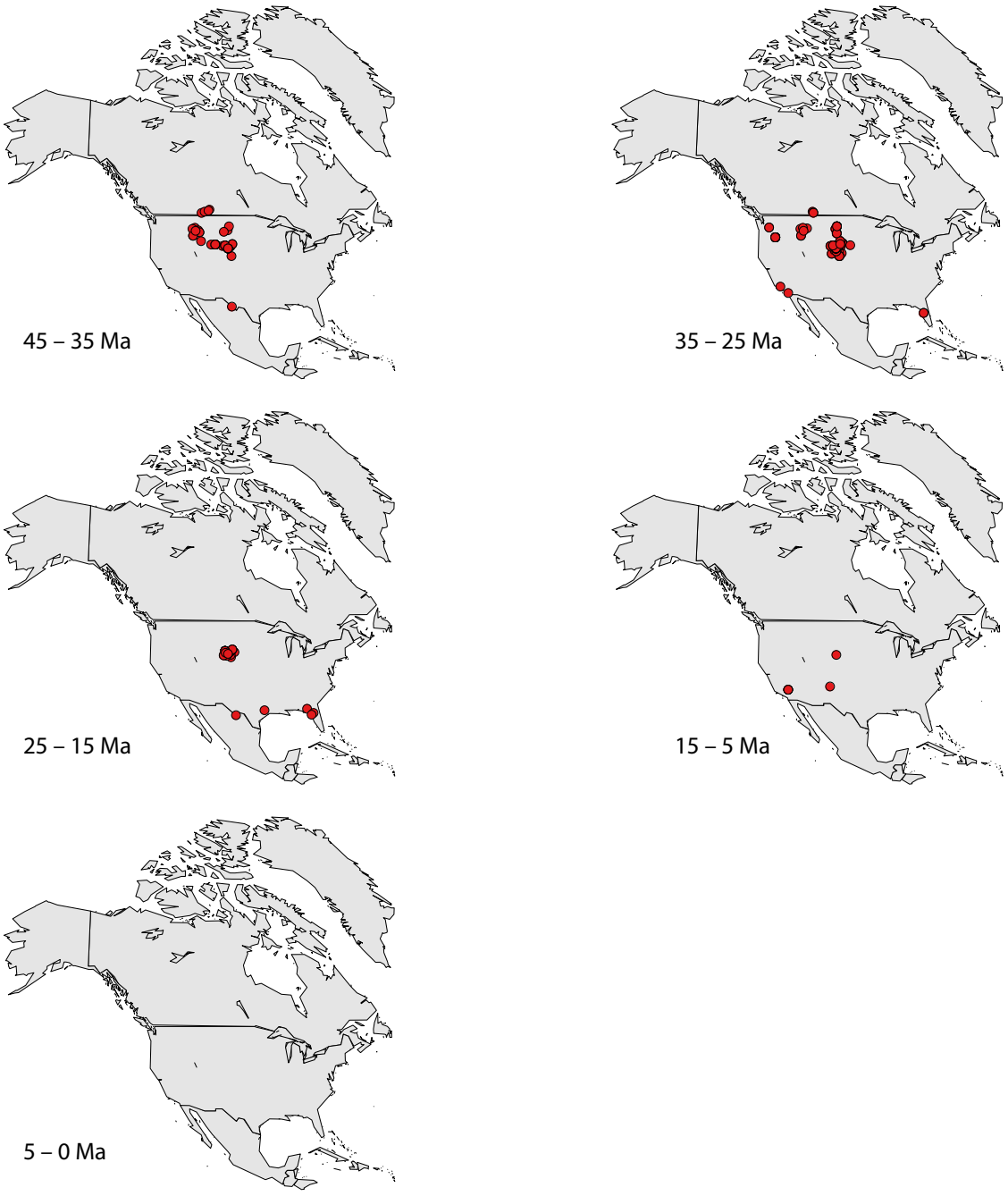


Figure S1: Fossil occurrences of Hesperocyoninae plotted in five time slices, based on the present geographic coordinates of the sampling localities and on their age estimates.

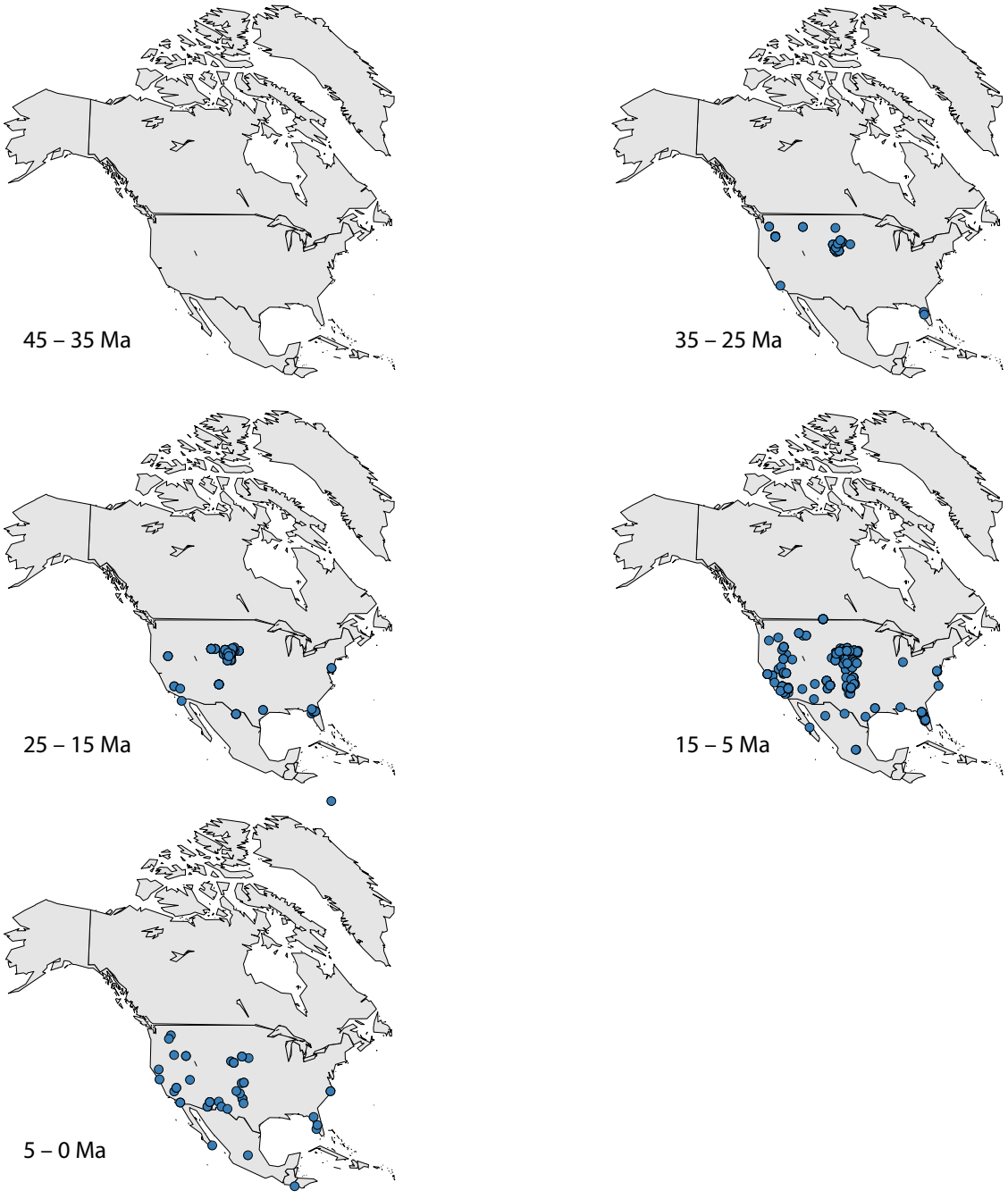


Figure S2: Fossil occurrences of Borophaginae plotted in five time slices, based on the present geographic coordinates of the sampling localities and on their age estimates.

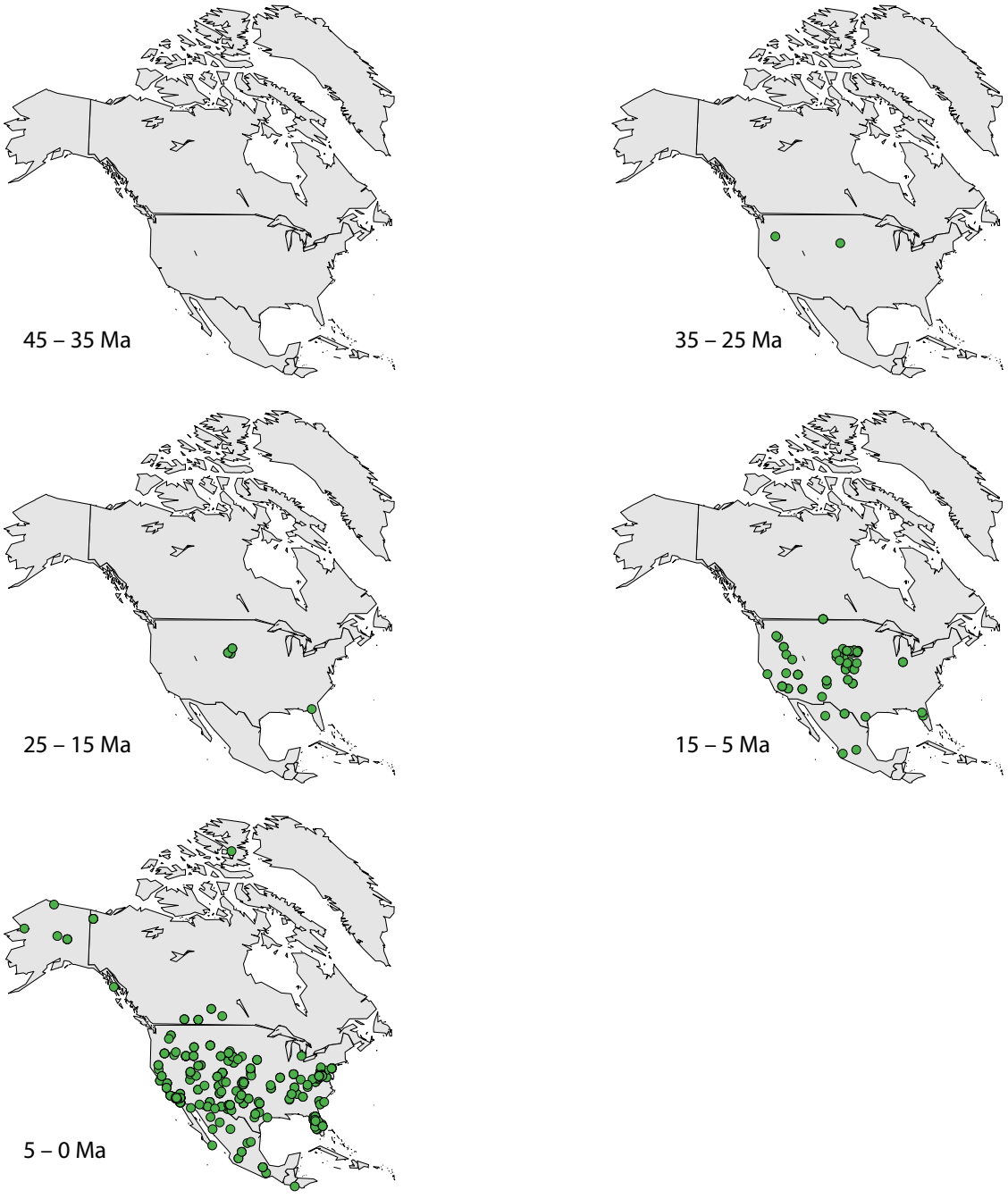


Figure S3: Fossil occurrences of Caninae plotted in five time slices, based on the present geographic coordinates of the sampling localities and on their age estimates.

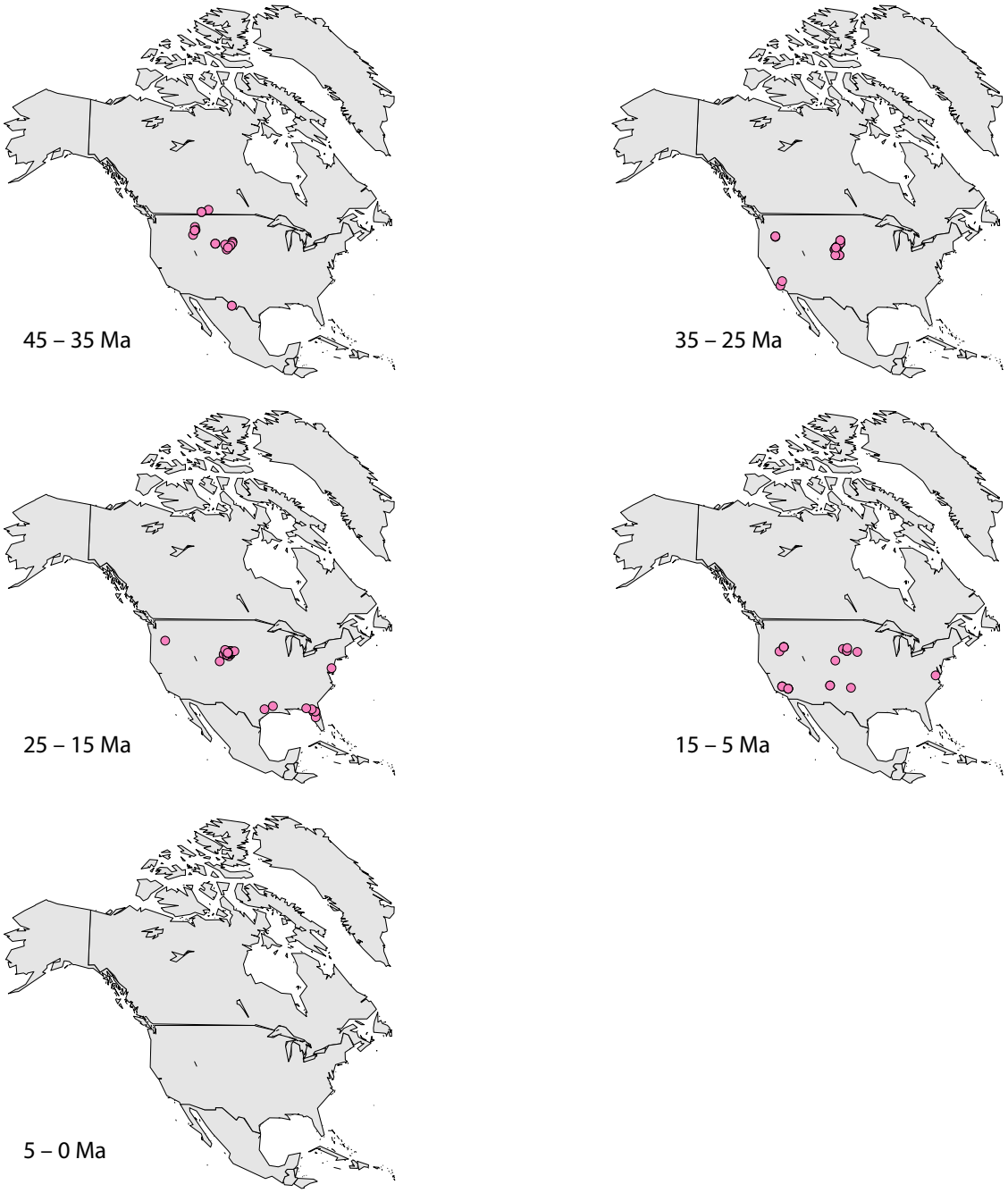


Figure S4: Fossil occurrences of Amphicyonidae plotted in five time slices, based on the present geographic coordinates of the sampling localities and on their age estimates.

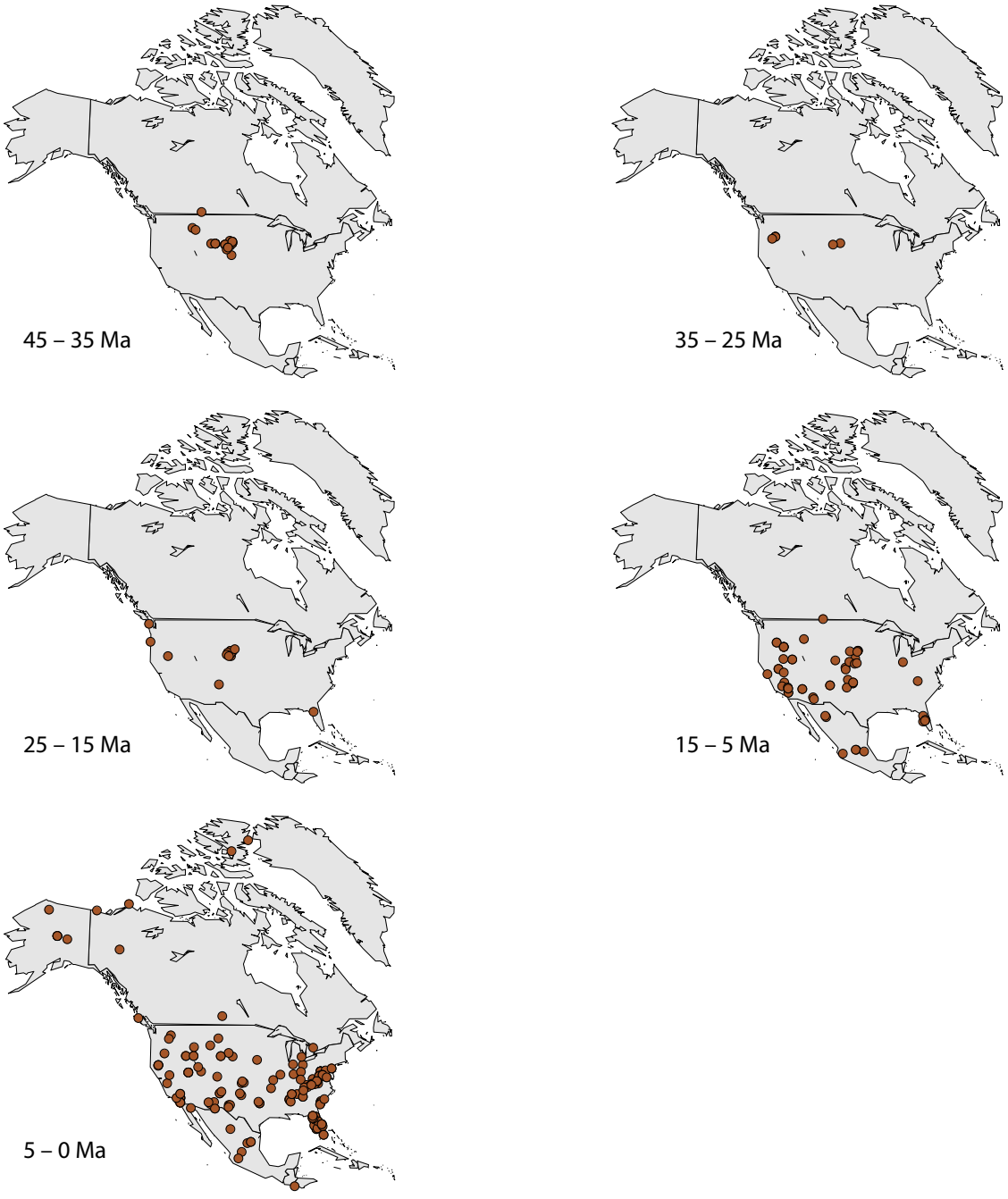


Figure S5: Fossil occurrences of Ursidae plotted in five time slices, based on the present geographic coordinates of the sampling localities and on their age estimates.

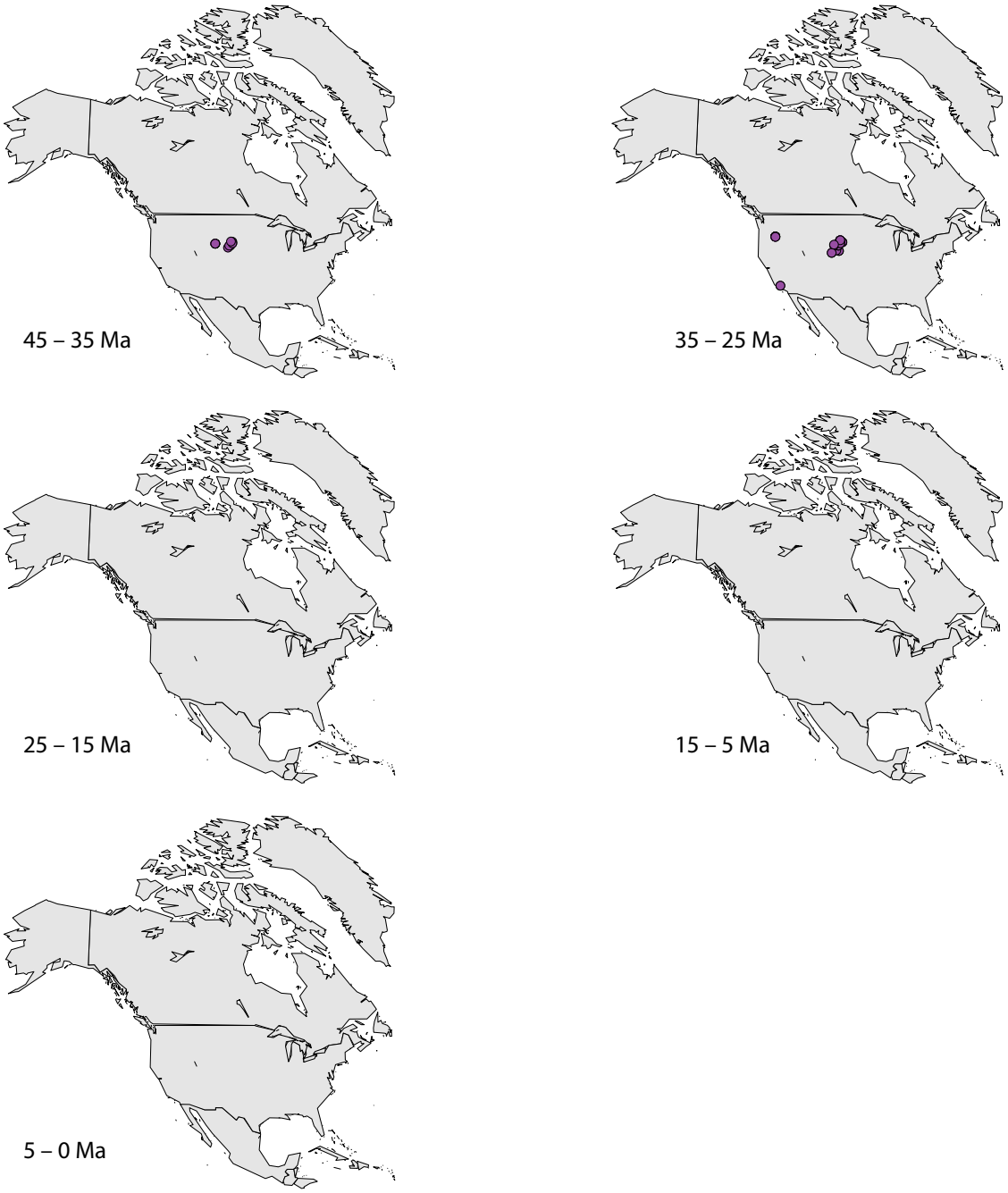


Figure S6: Fossil occurrences of Nimravidae plotted in five time slices, based on the present geographic coordinates of the sampling localities and on their age estimates.

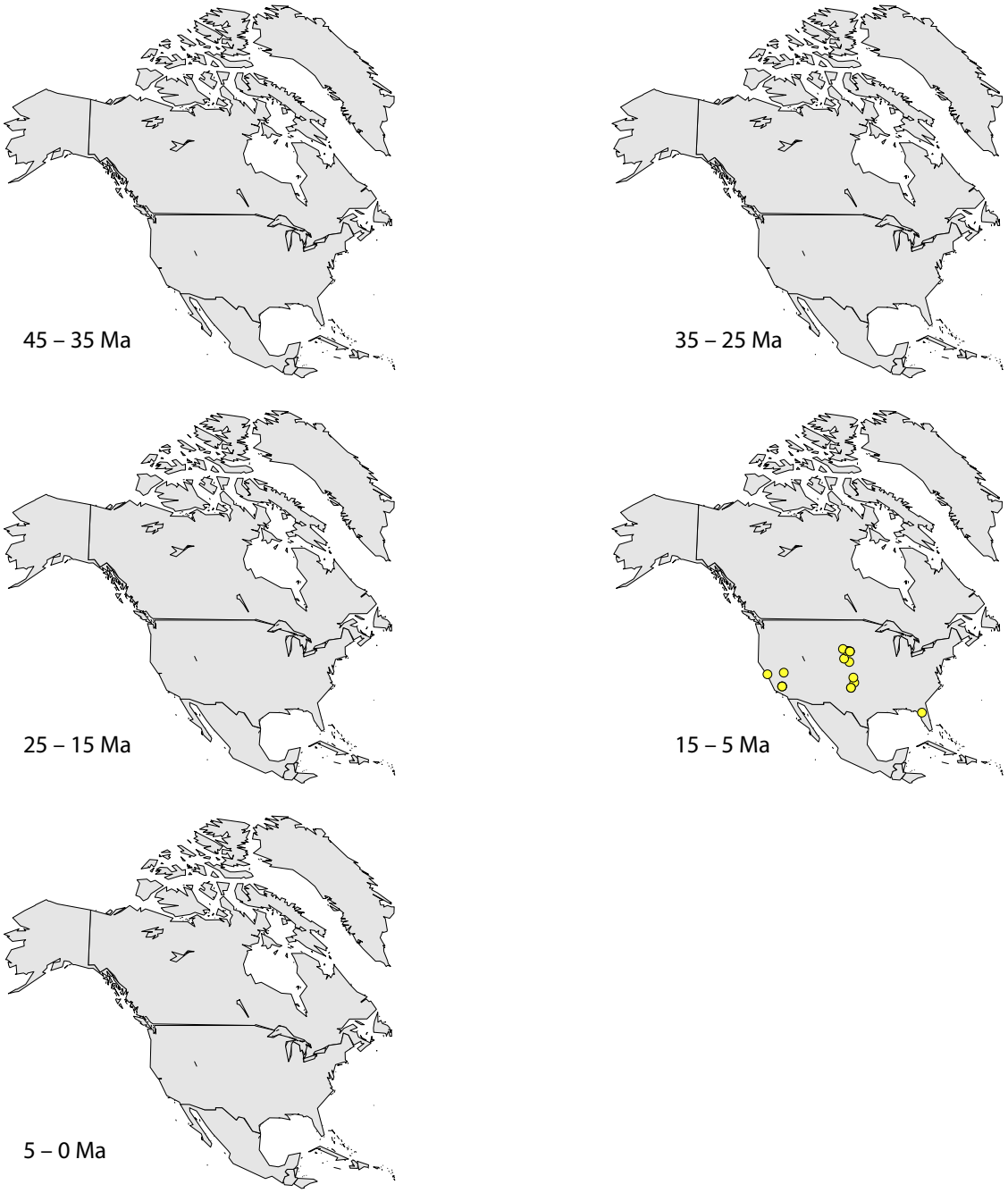


Figure S7: Fossil occurrences of Barbourofelidae plotted in five time slices, based on the present geographic coordinates of the sampling localities and on their age estimates.

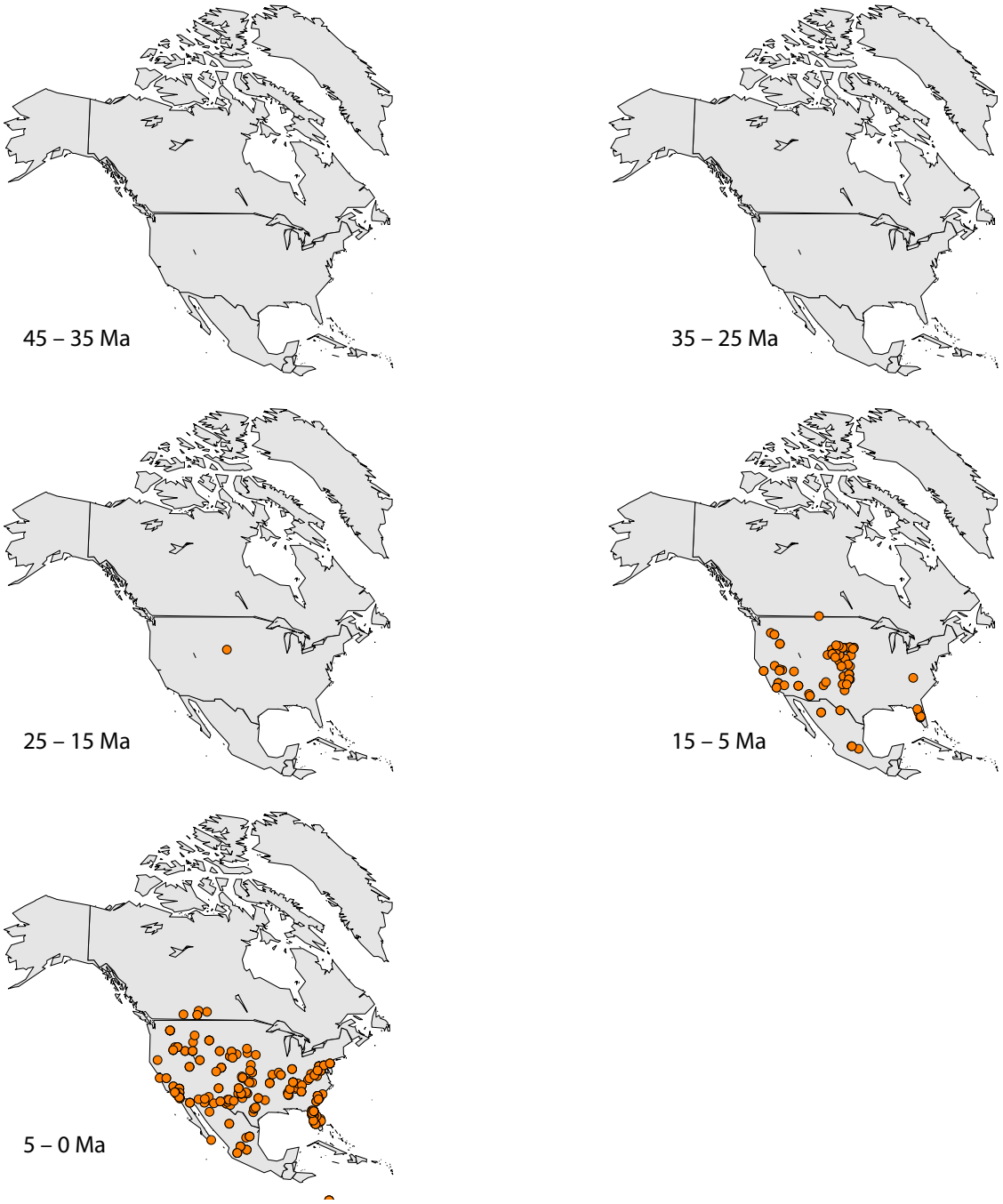


Figure S8: Fossil occurrences of Felidae plotted in five time slices, based on the present geographic coordinates of the sampling localities and on their age estimates.

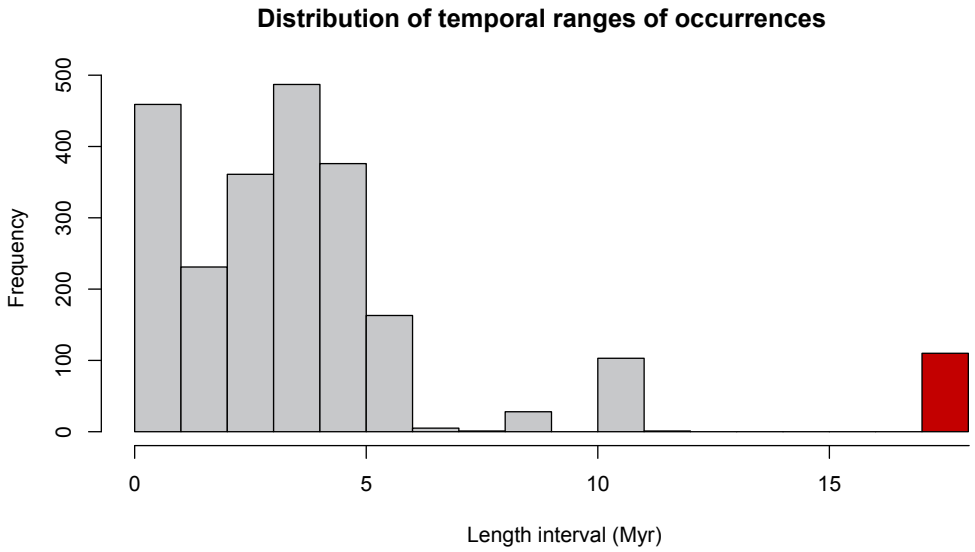


Figure S9: Temporal ranges of the fossil occurrences of all clades, quantifying the degree of uncertainty around the respective age estimates. Highlighted in red are the occurrences that were excluded from the analyses as low quality data.

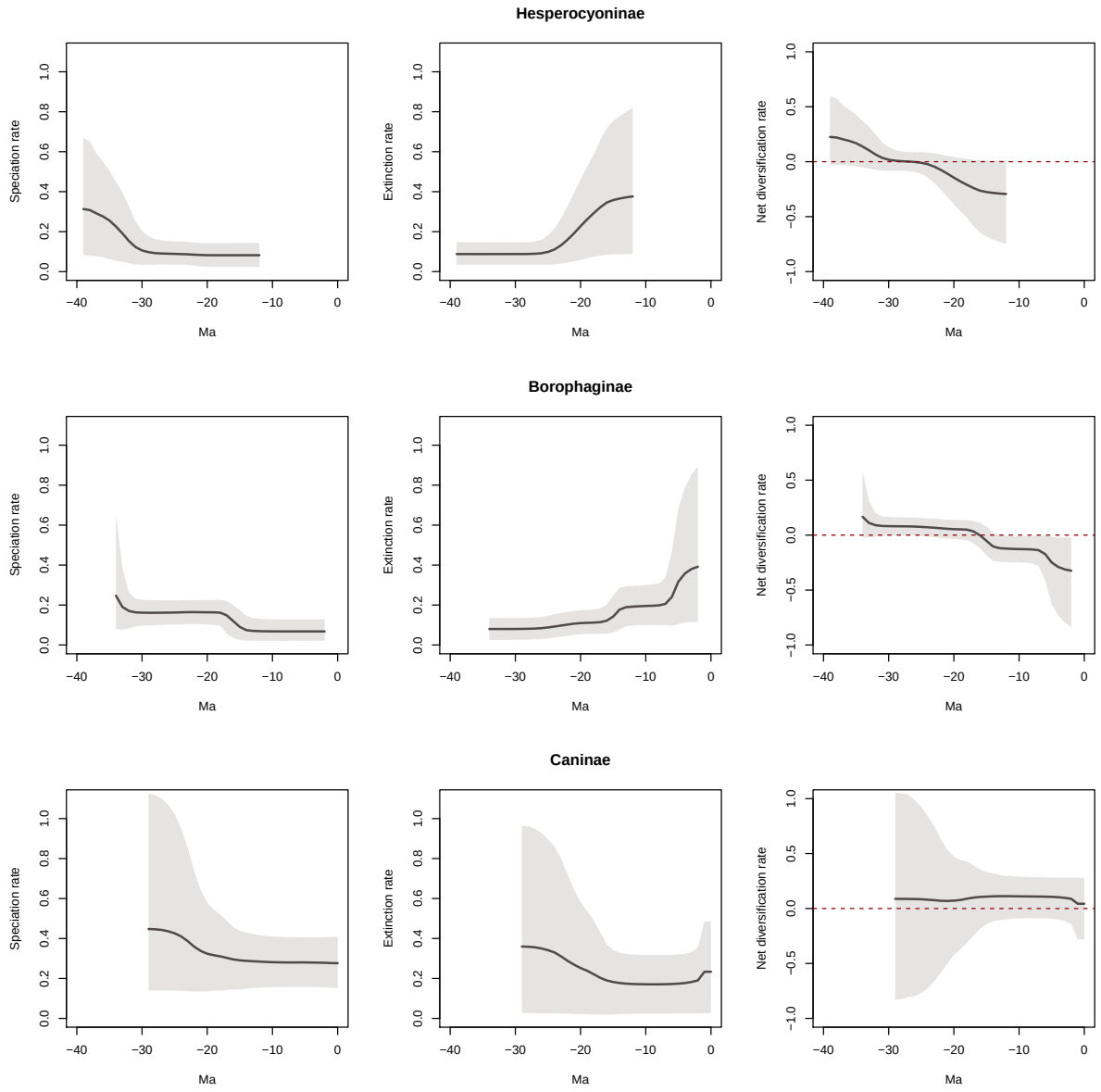


Figure S10: Speciation, extinction, and net diversification rates through time of the three Canidae subfamilies (only North American species) estimated by BDS analysis (4) under the birth-death shift model of diversification. Posterior estimates are averaged over 100 replicates to account for the uncertainties associated with the ages of the fossil occurrences.

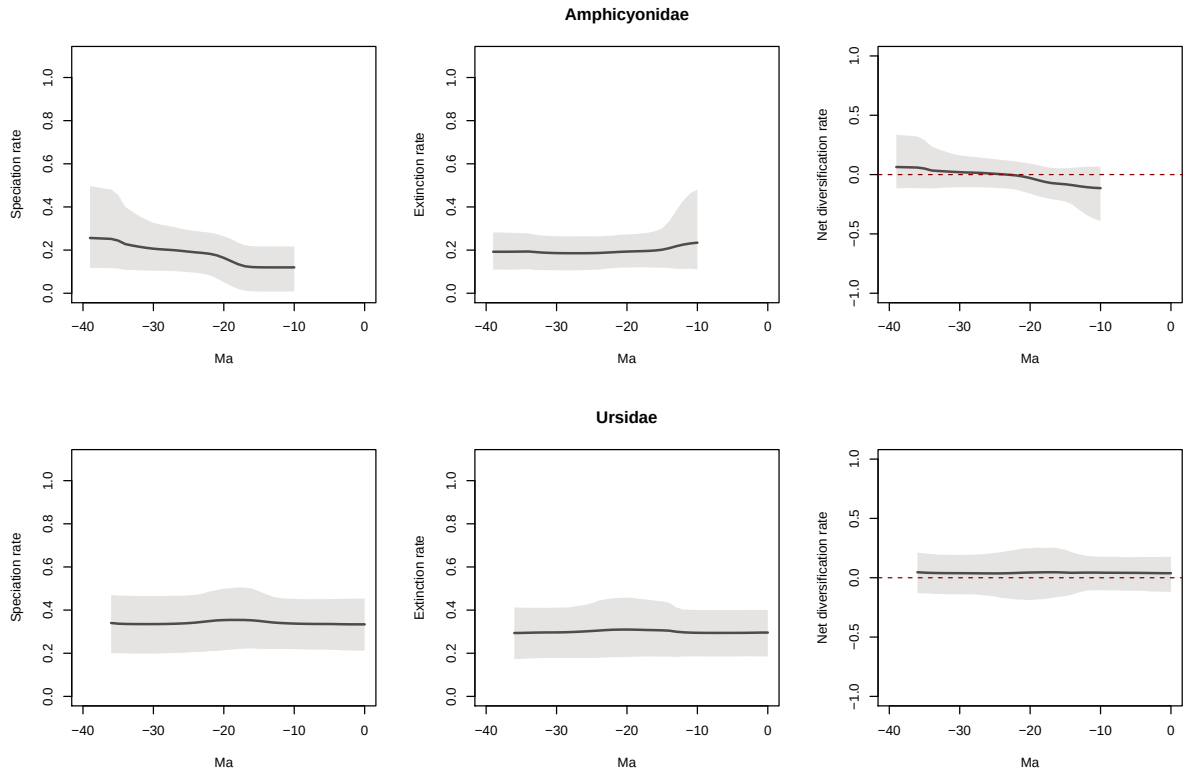


Figure S11: Speciation, extinction, and net diversification rates through time of Caniformia families Amphicyonidae and Ursidae (only North American species) estimated by BDS analysis (4) under the birth-death shift model of diversification. Posterior estimates are averaged over 100 replicates to account for the uncertainties associated with the ages of the fossil occurrences.

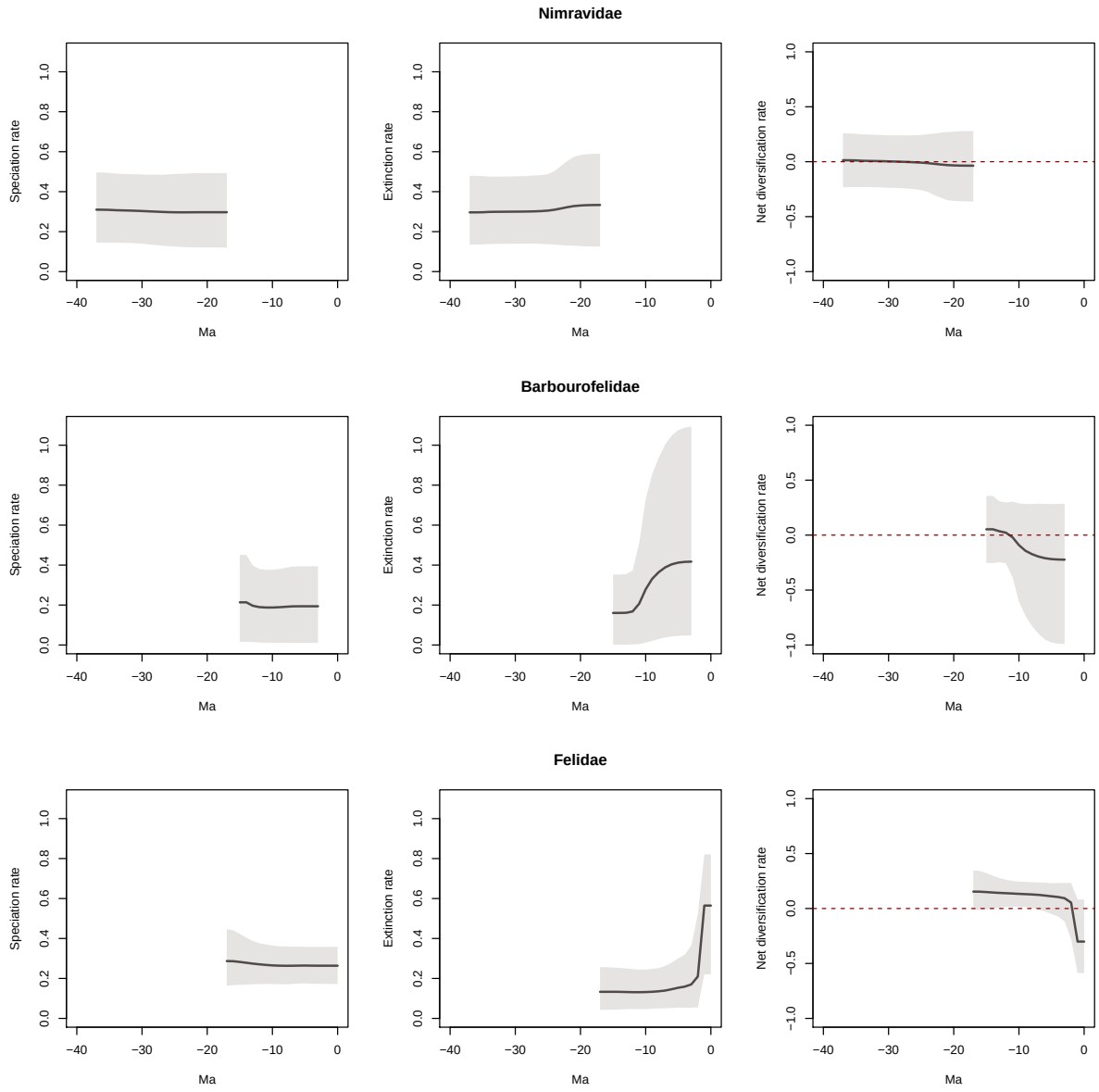


Figure S12: Speciation, extinction, and net diversification rates through time of Feliformia families Nimvravidae, Barbourofelidae, Felidae (only North American species) estimated by BDS analysis (4) under the birth-death shift model of diversification. Posterior estimates are averaged over 100 replicates to account for the uncertainties associated with the ages of the fossil occurrences.

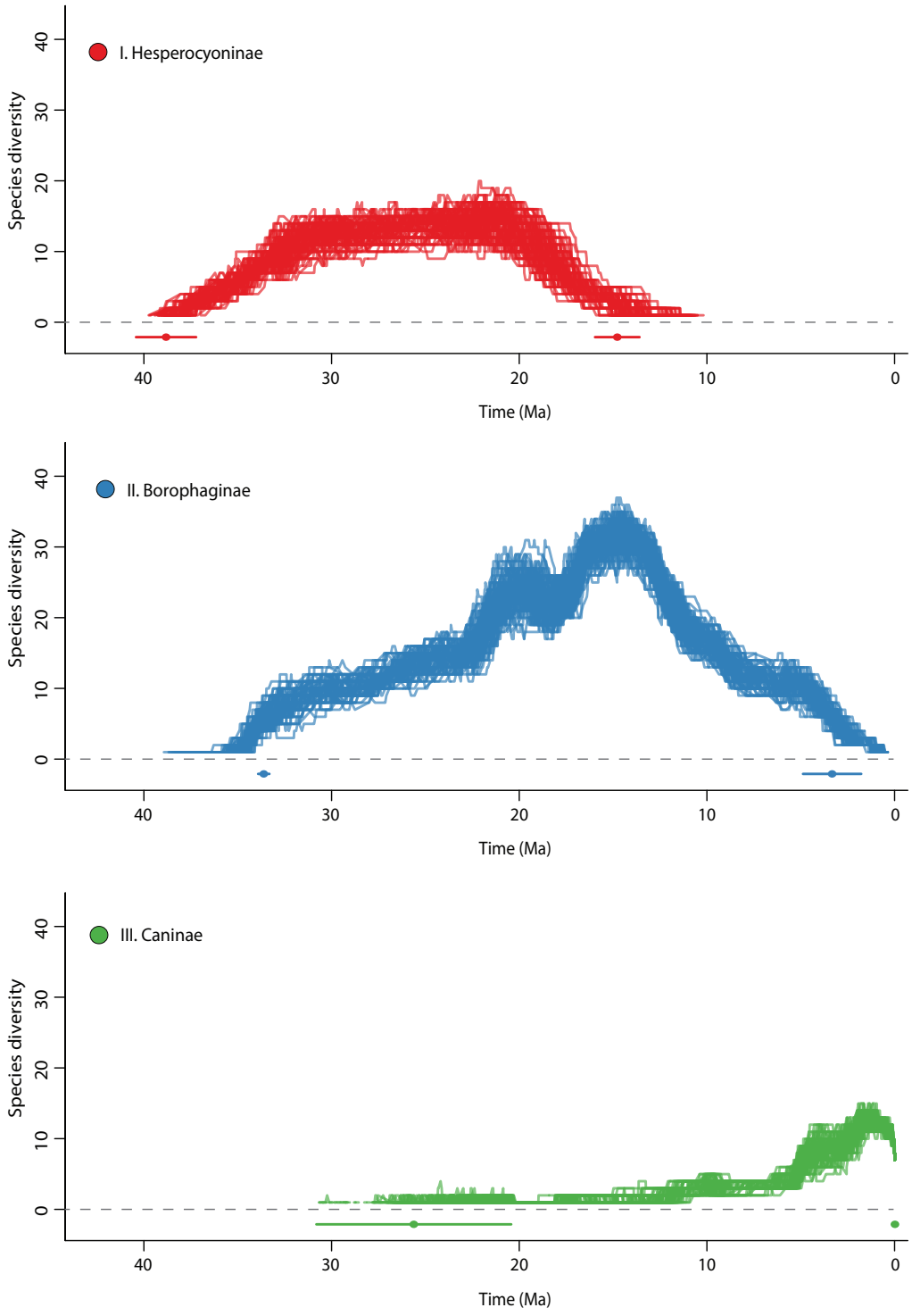


Figure S13: Diversity trajectories of the three Canidae subfamilies (only North American species) reconstructed under the BDS model. The curves are obtained from the times of speciation and extinction of each species estimated by modeling the preservation process (section 2.1). The trajectories derive from 100 data sets that incorporate the uncertainties associated with the fossil ages (section 1). The mid-points and temporal ranges of the first and last appearances of each clade are displayed below the dashed line.

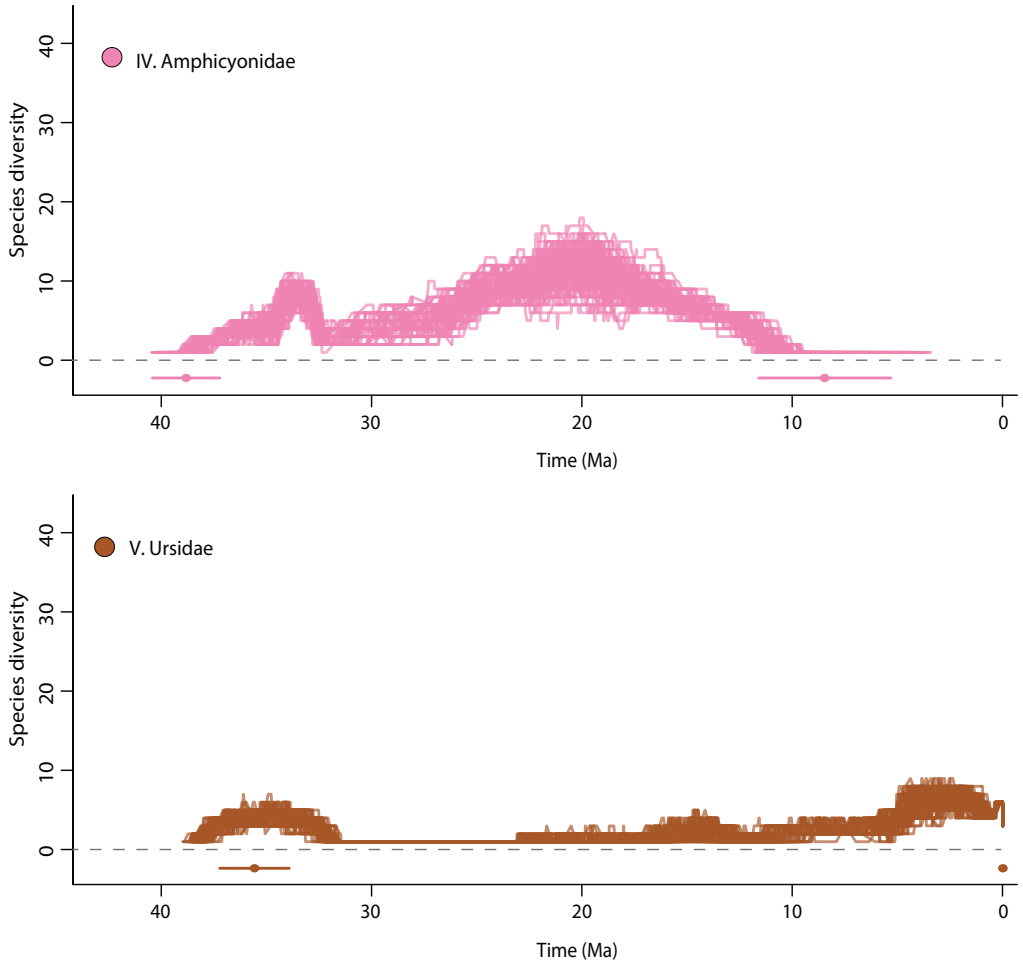


Figure S14: Diversity trajectories of Caniformia families Amphicyonidae and Ursidae (only North American species) reconstructed under the BDS model. The curves are obtained from the times of speciation and extinction of each species estimated by modeling the preservation process (section 2.1). The trajectories derive from 100 data sets that incorporate the uncertainties associated with the fossil ages (section 1). The mid-points and temporal ranges of the first and last appearances of each clade are displayed below the dashed line.

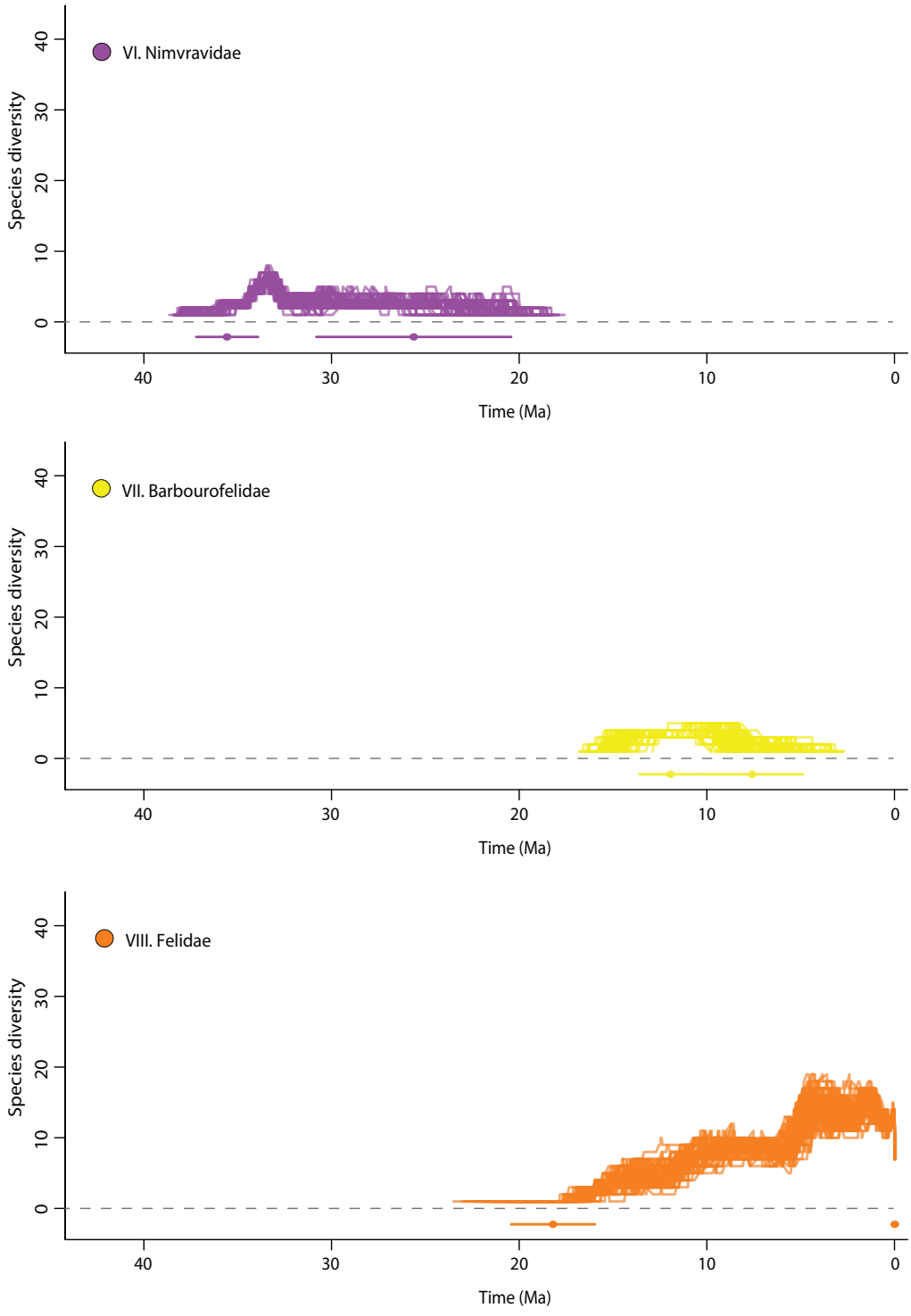


Figure S15: Diversity trajectories of Feliformia families Nimvravidae, Barbourofelidae, Felidae (only North American species) reconstructed under the BDS model. The curves are obtained from the times of speciation and extinction of each species estimated by modeling the preservation process (section 2.1). The trajectories derive from 100 data sets that incorporate the uncertainties associated with the fossil ages (section 1). The mid-points and temporal ranges of the first and last appearances of each clade are displayed below the dashed line.

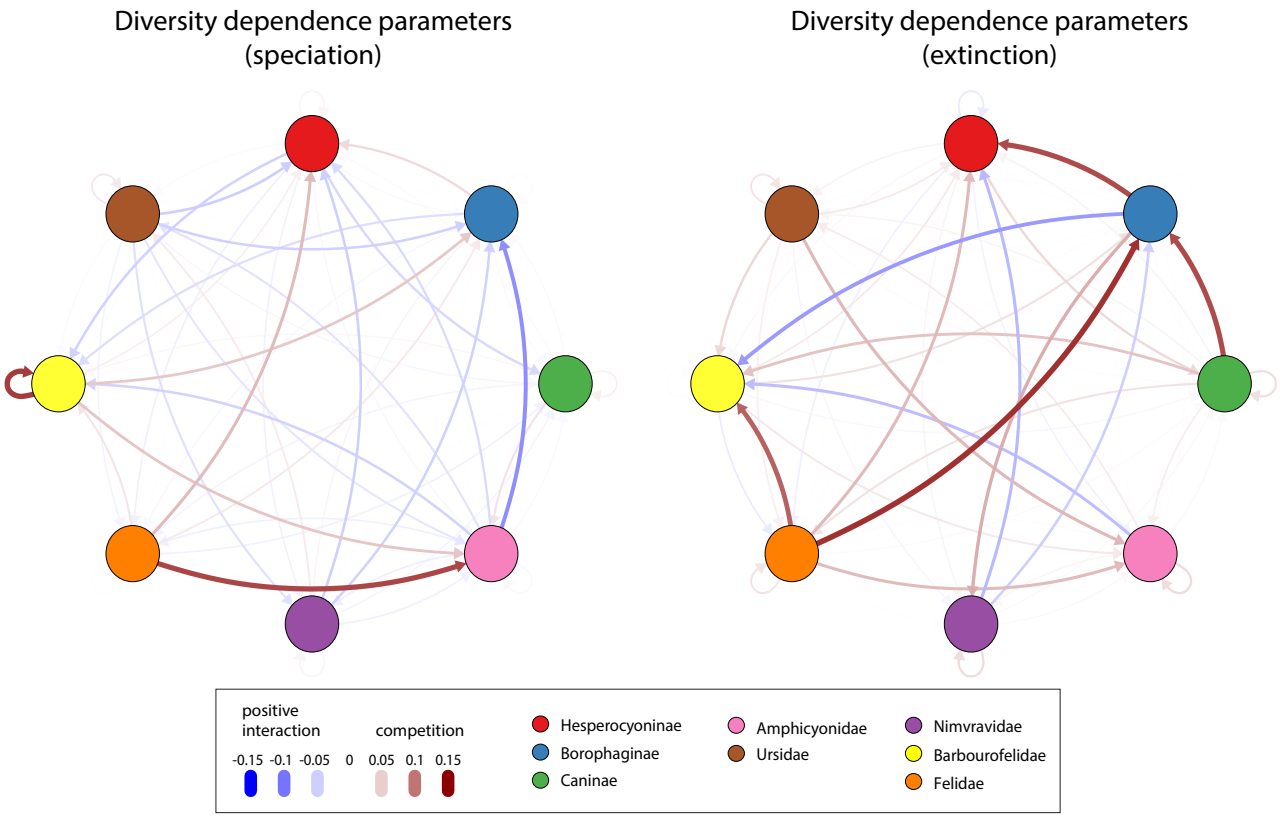


Figure S16: Posterior estimates of the diversity dependence parameters g (Equations 9, 11) between the eight North American carnivore clades included in the Multiple Clade Diversity Dependence analysis (MCDD; sections 2.4, 3.4). Among the competitive effects, only those affecting Hesperocyoninae (speciation and extinction) and Borophaginae (extinction) were found to ave an overall significant probability (Fig. 4; Table S5). None of the positive effects received significant support (Table S5).

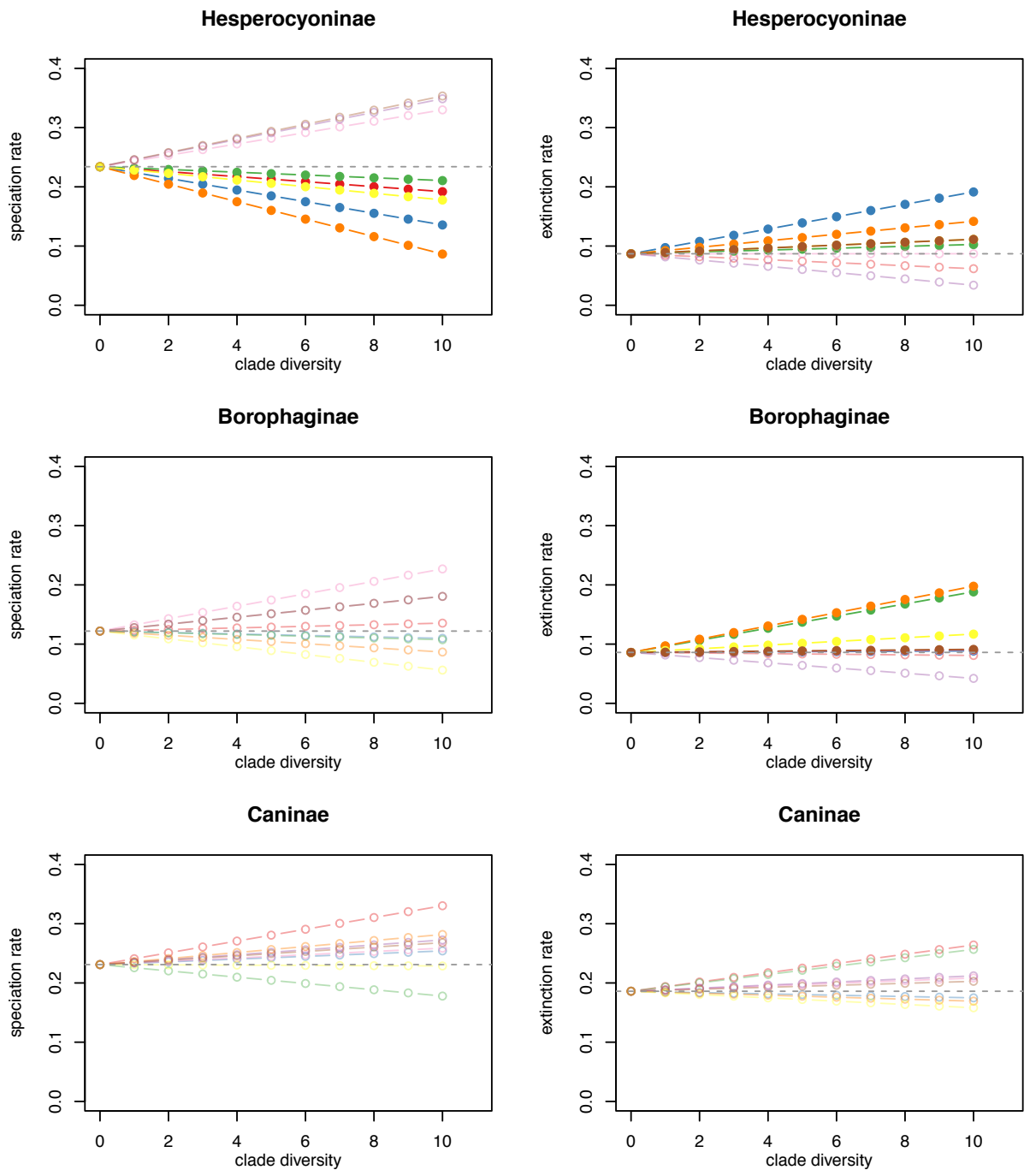


Figure S17: Rate changes induced by the diversity dependence correlations estimated in the three canid subfamilies under the MCDD model. The rates are obtained from the transformation of baseline rates based on the estimated competition parameters under the competitive pressure of 0 to 10 competing species for each of the 8 carnivore clades considered. Filled circles highlight the effects that were found to be overall significant based on their marginal probabilities (Fig. 4B; Table S5), i.e. competitive effects suppressing speciation rates in Hesperocyoninae and increasing extinction in Hesperocyoninae and Borophaginae. Open circles indicate positive or competitive effects which did not receive an overall significant support (Table S5).

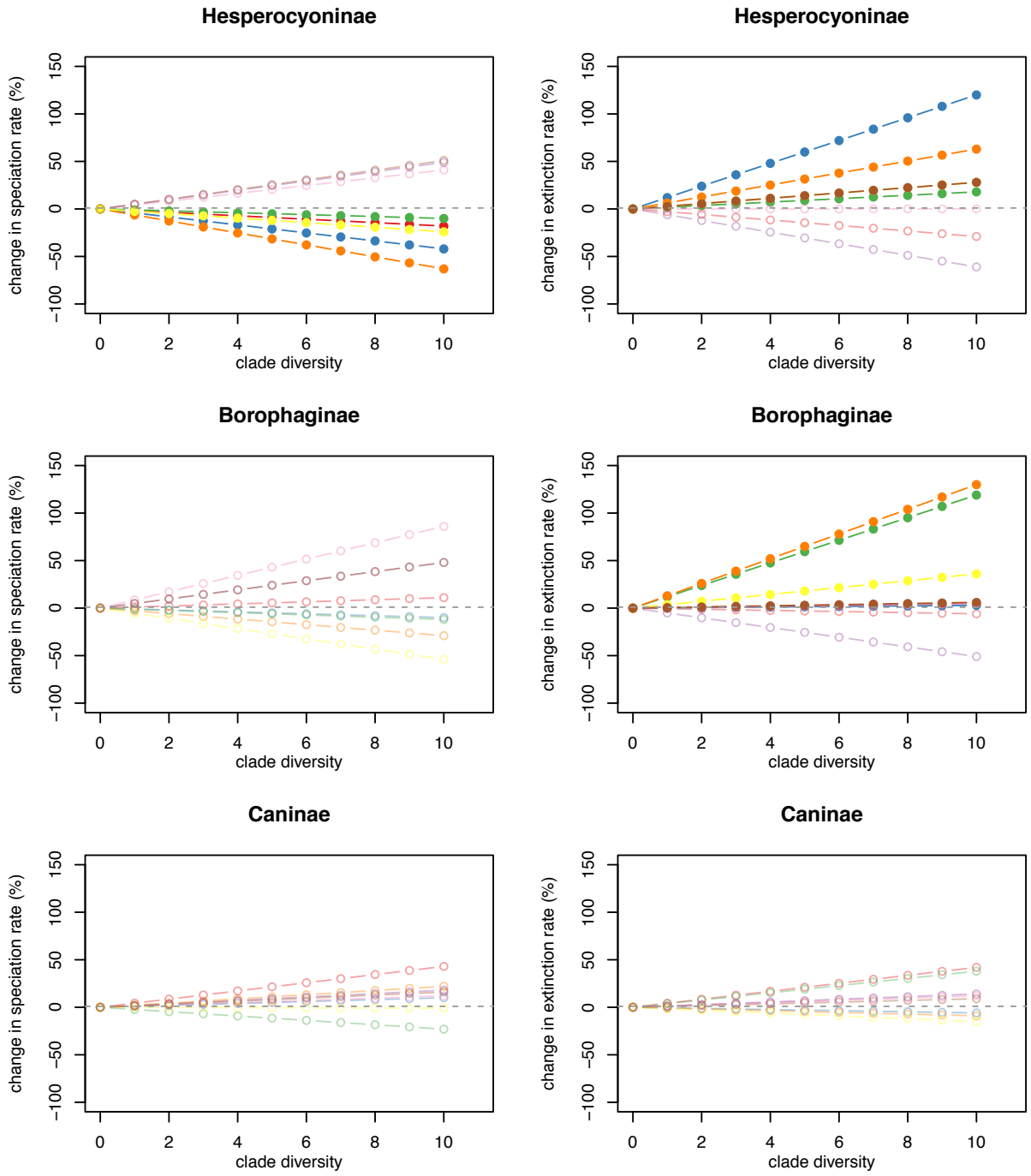


Figure S18: Reconstructed proportion of rate variation in the three canid subfamilies under the MCDD model. The rate changes caused by competition from different clades (see Fig. S17) are given as percentage of the baseline speciation and extinction rates.Filled circles highlight the effects that were found to be overall significant based on their marginal probabilities (Fig. 4B; Table S5), i.e. competitive effects suppressing speciation rates in Hesperocyoninae and increasing extinction in Hesperocyoninae and Borophaginae. Open circles indicate positive or competitive effects which did not receive an overall significant support (Table S5).

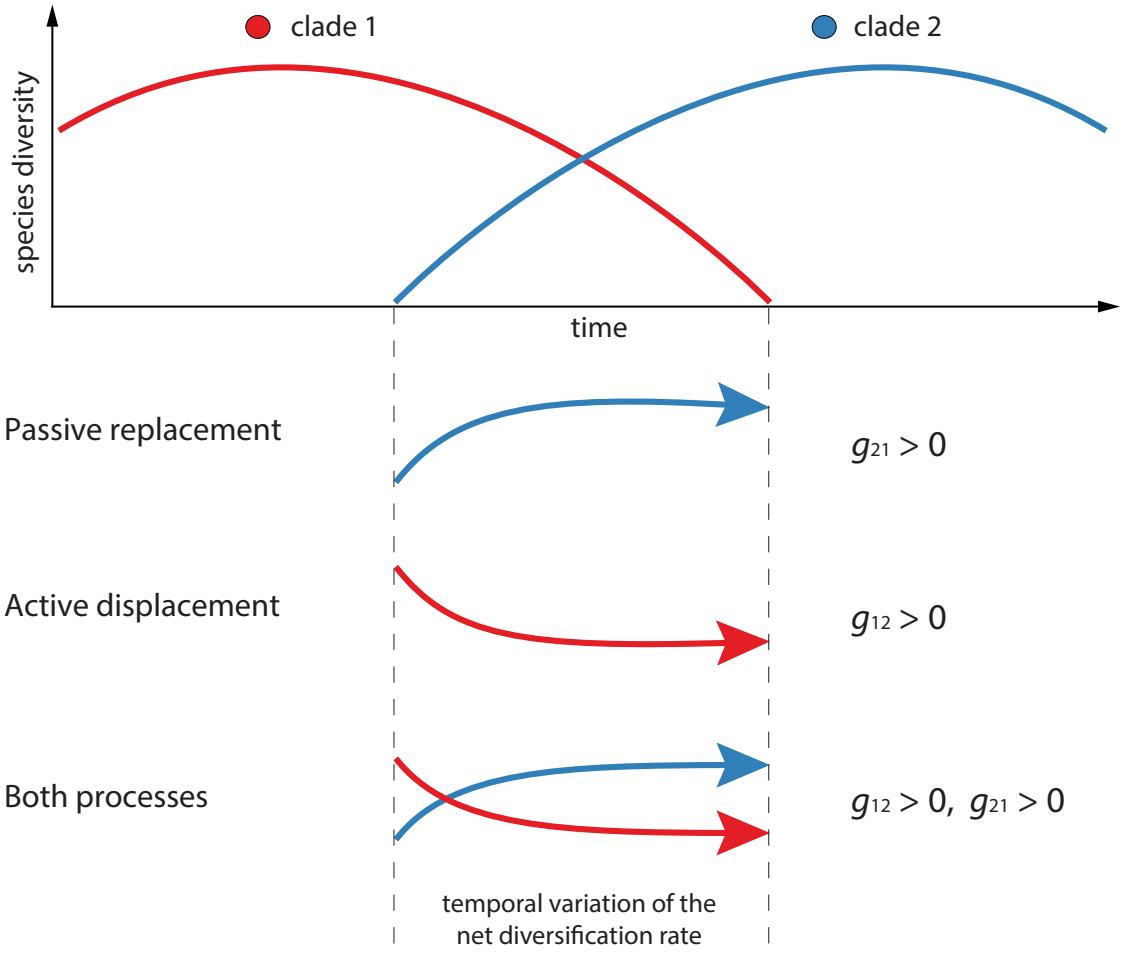


Figure S19: Predictions regarding the trends in net diversification rates under passive replacement and active displacement of clades. Under passive replacement, the decline of the incumbent clade correlates with an increase of net diversification rate (by increased speciation and/or decreased extinction) in the competing clade, which has the ecological and evolutionary opportunity to expand and diversify. Under active replacement the diversification of a new competing clade drives the decrease of net diversification rates in an older clade (by decreased speciation and/or increased extinction).

5 Supplementary Tables (S1–S13)

Table S1: Posterior estimates of the preservation rate (expected number of occurrences per lineage per Myr) and heterogeneity parameter (shape parameter of the Gamma distribution) estimated under the BDS model. Lower and upper bounds of the 95% HPDs are provided in parentheses.

Clade	Preservation rate		Heterogeneity parameter	
Hesperocyoninae	4.279	(3.502 – 5.07)	0.42	(0.314 – 0.545)
Borophaginae	1.896	(1.59 – 2.216)	1.318	(0.977 – 1.683)
Caninae	6.952	(5.929 – 8.009)	0.571	(0.304 – 0.888)
Amphicyonidae	0.807	(0.548 – 1.095)	2.44	(0.68 – 5.749)
Ursidae	8.044	(5.8 – 10.266)	0.539	(0.338 – 0.762)
Nimvravidae	1.55	(0.756 – 2.555)	3.756	(0.244 – 11.326)
Barbourofelidae	0.674	(0.273 – 1.179)	10.319	(1.779 – 19.524)
Felidae	5.702	(4.669 – 6.795)	0.823	(0.647 – 1.012)

Table S2: Sampling frequencies of the number of rates (k_λ and k_μ) estimated by BDMCMC under BDS model for the three canid subfamilies and averaged over 100 replicates (standard deviation across replicates is given in parentheses). Numbers in bold highlight the number of rates with the highest probability.

Clade	Parameter	Sampling frequency (number of rates)			
		1	2	3	4
Hesperocyaninae	λ	0.184 (0.134)	**0.764** (0.127)	0.050 (0.019)	0.001 (0.001)
	μ	0.175 (0.158)	**0.744** (0.147)	0.078 (0.051)	0.003 (0.003)
Borophaginae	λ	0.058 (0.048)	**0.589** (0.094)	0.301 (0.086)	0.049 (0.039)
	μ	0.036 (0.033)	**0.635** (0.120)	0.298 (0.121)	0.028 (0.013)
Caninae	λ	**0.704** (0.126)	0.258 (0.109)	0.036 (0.016)	0.003 (0.002)
	μ	**0.630** (0.152)	0.218 (0.068)	0.133 (0.102)	0.018 (0.015)
Amphicyonidae	λ	0.376 (0.155)	**0.561** (0.137)	0.06 (0.025)	0.002 (0.002)
	μ	**0.792** (0.079)	0.170 (0.057)	0.033 (0.029)	0.004 (0.006)
Ursidae	λ	**0.853** (0.038)	0.107 (0.013)	0.032 (0.024)	0.007 (0.01)
	μ	**0.877** (0.032)	0.098 (0.013)	0.022 (0.022)	0.003 (0.006)
Nimvravidae	λ	**0.924** (0.023)	0.075 (0.022)	0.002 (0.002)	0
	μ	**0.897** (0.027)	0.101 (0.027)	0.002 (0.002)	0
Barbourofelidae	λ	**0.666** (0.118)	0.322 (0.116)	0.012 (0.012)	0
	μ	**0.558** (0.145)	0.432 (0.147)	0.010 (0.005)	0
Felidae	λ	**0.850** (0.036)	0.131 (0.028)	0.018 (0.014)	0.002 (0.002)
	μ	0.139 (0.122)	**0.743** (0.108)	0.110 (0.024)	0.008 (0.004)

Table S3: Posterior estimates of the parameters $\alpha_\lambda, \alpha_\mu$ quantifying the correlation between species-specific body mass and speciation and extinction rates (COVAR model). 95% HPDs are given in parentheses.

Clade	Correlation parameters (95% HPD)			
	α_λ		α_μ	
Hesperocyaninae	-0.140	(-0.952 – 0.664)	-0.140	(-0.942 – 0.669)
Borophaginae	-0.072	(-0.257 – 0.109)	-0.073	(-0.255 – 0.107)
Caninae	-0.136	(-0.703 – 0.415)	-0.042	(-0.692 – 0.597)

Table S4: Posterior estimates of the parameters $\gamma_\lambda, \gamma_\mu$ quantifying the correlation between global temperature dynamics and speciation and extinction rates (BDT model). 95% HPDs are given in parentheses. The value in bold highlights the parameter displaying a significant correlation, i.e. γ significantly different from 0.

Clade	Correlation parameters (95% HPD)			
	γ_λ		γ_μ	
Hesperocyaninae	0.584	(-2.105 – 3.353)	-1.701	(-3.839 – 0.572)
Borophaginae	1.523	(-0.069 – 3.194)	**-2.695**	(-3.813 – -1.539)
Caninae	0.446	(-1.015 – 1.897)	-1.458	(-3.207 – 0.238)

Table S5: Marginal probabilities of positive interaction and competition indicated as P_{posi} and P_{comp}, respectively (MCDD model). The marginal probabilities are given for each clade and calculated separately for speciation and extinction based on Equations (13,14). Values in bold highlight significant probabilities (at a 0.95 threshold).

	P_{posi}		P_{comp}	
Clade	Speciation	Extinction	Speciation	Extinction
Hesperocyoninae	0.794	0.868	**0.962**	**0.950**
Borophaginae	0.811	0.622	0.885	**0.990**
Caninae	0.515	0.437	0.349	0.509
Amphicyonidae	0.509	0.463	0.911	0.779
Nimravidae	0.607	0.561	0.589	0.739
Felidae	0.520	0.565	0.368	0.593
Barburofelidae	0.870	0.932	0.911	0.888
Ursidae	0.351	0.266	0.455	0.454

Table S6: Posterior parameter estimates for Hesperocyoninae under the MCDD model. The clades are here numbered as follows: 1) Hesperocyoninae, 2) Borophaginae, 3) Caninae, 4) Amphicyonidae, 5) Nimvravidae, 6) Felidae, 7) Barbourofelidae, 8) Ursidae.

Parameter		Mean	95%HPD		Pos. interaction P(g<0)	Competition P(g>0)
Baseline rates	λ_1	0.234	0.052	0.502		
	μ_1	0.087	0.013	0.227		
Hyperprior	η_1	0.469	0.003	0.845		
Competition	g_{11}^{λ}	0.018	-0.056	0.122	0.088	0.395
parameters	g_{12}^{λ}	0.042	-0.001	0.166	0.016	0.556
	g_{13}^{λ}	0.01	-0.239	0.3	0.246	0.296
	g_{14}^{λ}	-0.041	-0.266	0.07	0.355	0.091
	g_{15}^{λ}	-0.049	-0.3	0.091	0.375	0.124
	g_{16}^{λ}	0.063	-0.184	0.3	0.16	0.454
	g_{17}^{λ}	0.024	-0.21	0.3	0.205	0.325
	g_{18}^{λ}	-0.051	-0.3	0.08	0.379	0.116
	g_{11}^{μ}	-0.029	-0.228	0.068	0.33	0.086
	g_{12}^{μ}	0.12	0.0	0.285	0.009	0.693
	g_{13}^{μ}	0.018	-0.222	0.3	0.223	0.315
	g_{14}^{μ}	0.0	-0.246	0.257	0.227	0.232
	g_{15}^{μ}	-0.061	-0.3	0.183	0.467	0.163
	g_{16}^{μ}	0.063	-0.161	0.3	0.145	0.455
	g_{17}^{μ}	0.028	-0.207	0.3	0.198	0.341
	g_{18}^{μ}	0.028	-0.159	0.3	0.172	0.32

Table S7: Posterior parameter estimates for Borophaginae under the MCDD model. The clades are here numbered as follows: 1) Hesperocyoninae, 2) Borophaginae, 3) Caninae, 4) Amphicyonidae, 5) Nimvravidae, 6) Felidae, 7) Barbourofelidae, 8) Ursidae.

Parameter		Mean	95%HPD		Pos. interaction P(g<0)	Competition P(g>0)
Baseline rates	λ_2	0.122	0.033	0.215		
	μ_2	0.086	0.036	0.143		
Hyperprior	η_2	0.616	0.245	0.944		
MCDD	g_{21}^{λ}	-0.011	-0.169	0.098	0.144	0.092
	g_{22}^{λ}	0.01	0.0	0.069	0.019	0.22
	g_{23}^{λ}	0.012	-0.163	0.257	0.122	0.224
	g_{24}^{λ}	-0.086	-0.281	0.0	0.5	0.012
	g_{25}^{λ}	-0.048	-0.3	0.05	0.328	0.077
	g_{26}^{λ}	0.029	-0.029	0.201	0.052	0.324
	g_{27}^{λ}	0.054	-0.039	0.3	0.069	0.42
	g_{28}^{λ}	-0.048	-0.3	0.036	0.32	0.07
	g_{21}^{μ}	-0.006	-0.077	0.023	0.13	0.041
	g_{22}^{μ}	0.003	-0.021	0.032	0.034	0.067
	g_{23}^{μ}	0.119	0.0	0.296	0.031	0.592
	g_{24}^{μ}	0.005	-0.061	0.117	0.068	0.107
	g_{25}^{μ}	-0.051	-0.3	0.057	0.354	0.079
	g_{26}^{μ}	0.13	0.0	0.293	0.022	0.665
	g_{27}^{μ}	0.036	-0.055	0.3	0.083	0.279
	g_{28}^{μ}	0.006	-0.185	0.3	0.167	0.192

Table S8: Posterior parameter estimates for Caninae under the MCDD model. The clades are here numbered as follows: 1) Hesperocyoninae, 2) Borophaginae, 3) Caninae, 4) Amphicyonidae, 5) Nimvravidae, 6) Felidae, 7) Barbourofelidae, 8) Ursidae.

Parameter		Mean	95%HPD		Pos. interaction P(g<0)	Competition P(g>0)
Baseline rates	λ_3	0.231	0.043	0.38		
	μ_3	0.186	0.038	0.331		
Hyperprior	η_3	0.746	0.204	1.0		
MCDD	g^{λ}_{31}	-0.043	-0.268	-0.0	0.265	0.017
	g^{λ}_{32}	-0.01	-0.14	0.046	0.105	0.037
	g^{λ}_{33}	0.023	-0.0	0.28	0.038	0.176
	g^{λ}_{34}	-0.012	-0.296	0.057	0.139	0.072
	g^{λ}_{35}	-0.018	-0.3	0.096	0.178	0.086
	g^{λ}_{36}	-0.022	-0.244	0.0	0.159	0.029
	g^{λ}_{37}	0.001	-0.197	0.217	0.104	0.115
	g^{λ}_{38}	-0.016	-0.3	0.042	0.149	0.068
	g^{μ}_{31}	0.042	0.0	0.27	0.018	0.254
	g^{μ}_{32}	-0.006	-0.086	0.0	0.124	0.024
	g^{μ}_{33}	0.038	0.0	0.273	0.023	0.23
	g^{μ}_{34}	0.012	-0.059	0.299	0.075	0.139
	g^{μ}_{35}	0.014	-0.115	0.3	0.093	0.164
	g^{μ}_{36}	-0.009	-0.22	0.071	0.126	0.066
	g^{μ}_{37}	-0.015	-0.3	0.125	0.18	0.096
	g^{μ}_{38}	0.009	-0.087	0.3	0.086	0.131

Table S9: Posterior parameter estimates for Amphicyonidae under the MCDD model. The clades are here numbered as follows: 1) Hesperocyoninae, 2) Borophaginae, 3) Caninae, 4) Amphicyonidae, 5) Nimvravidae, 6) Felidae, 7) Barbourofelidae, 8) Ursidae.

Parameter		Mean	95%HPD		Pos. interaction P(g<0)	Competition P(g>0)
Baseline rates	λ_4	0.228	0.067	0.413		
	μ_4	0.154	0.038	0.265		
Hyperprior	η_4	0.661	0.231	1.0		
MCDD	g^{λ}_{41}	-0.014	-0.147	0.04	0.143	0.037
	g^{λ}_{42}	0.003	-0.0	0.038	0.021	0.109
	g^{λ}_{43}	0.034	-0.121	0.3	0.106	0.274
	g^{λ}_{44}	0.014	-0.028	0.161	0.037	0.202
	g^{λ}_{45}	-0.024	-0.3	0.065	0.212	0.086
	g^{λ}_{46}	0.122	0.0	0.295	0.029	0.62
	g^{λ}_{47}	0.055	-0.045	0.3	0.071	0.345
	g^{λ}_{48}	-0.033	-0.3	0.006	0.24	0.054
	g^{μ}_{41}	-0.02	-0.212	0.018	0.219	0.043
	g^{μ}_{42}	0.002	-0.085	0.053	0.074	0.075
	g^{μ}_{43}	0.025	-0.116	0.3	0.104	0.235
	g^{μ}_{44}	0.041	-0.0	0.274	0.026	0.267
	g^{μ}_{45}	0.014	-0.145	0.3	0.124	0.192
	g^{μ}_{46}	0.062	0.0	0.293	0.039	0.368
	g^{μ}_{47}	0.03	-0.088	0.3	0.091	0.25
	g^{μ}_{48}	0.064	-0.0	0.296	0.042	0.368

Table S10: Posterior parameter estimates for Nimvravidae under the MCDD model. The clades are here numbered as follows: 1) Hesperocyoninae, 2) Borophaginae, 3) Caninae, 4) Amphicyonidae, 5) Nimvravidae, 6) Felidae, 7) Barbourofelidae, 8) Ursidae.

Parameter		Mean	95%HPD		Pos. interaction P(g<0)	Competition P(g>0)
Baseline rates	λ_5	0.275	0.052	0.523		
	μ_5	0.197	0.031	0.412		
Hyperprior	η_5	0.639	0.098	1.0		
MCDD	g^{λ}_{51}	-0.021	-0.24	0.062	0.182	0.07
	g^{λ}_{52}	0.018	-0.053	0.255	0.064	0.188
	g^{λ}_{53}	0.005	-0.205	0.3	0.168	0.194
	g^{λ}_{54}	-0.025	-0.292	0.05	0.212	0.079
	g^{λ}_{55}	0.018	-0.14	0.3	0.119	0.223
	g^{λ}_{56}	0.002	-0.215	0.298	0.176	0.186
	g^{λ}_{57}	0.0	-0.217	0.297	0.18	0.181
	g^{λ}_{58}	-0.038	-0.3	0.059	0.278	0.081
	g^{μ}_{51}	0.017	-0.139	0.3	0.123	0.216
	g^{μ}_{52}	0.07	-0.0	0.286	0.025	0.405
	g^{μ}_{53}	0.0	-0.268	0.248	0.178	0.181
	g^{μ}_{54}	0.017	-0.143	0.3	0.125	0.217
	g^{μ}_{55}	0.04	-0.07	0.3	0.083	0.289
	g^{μ}_{56}	0.004	-0.209	0.3	0.171	0.189
	g^{μ}_{57}	0.0	-0.217	0.299	0.18	0.18
	g^{μ}_{58}	-0.017	-0.3	0.166	0.223	0.134

Table S11: Posterior parameter estimates for Felidae under the MCDD model. The clades are here numbered as follows: 1) Hesperocyoninae, 2) Borophaginae, 3) Caninae, 4) Amphicyonidae, 5) Nimvravidae, 6) Felidae, 7) Barbourofelidae, 8) Ursidae.

Parameter		Mean	95%HPD		Pos. interaction P(g<0)	Competition P(g>0)
Baseline rates	λ_6	0.227	0.031	0.375		
	μ_6	0.179	0.034	0.337		
Hyperprior	η_6	0.738	0.232	1.0		
MCDD	g^{λ}_{61}	-0.023	-0.3	0.005	0.183	0.052
	g^{λ}_{62}	-0.015	-0.144	0.008	0.136	0.012
	g^{λ}_{63}	-0.018	-0.231	0.034	0.141	0.037
	g^{λ}_{64}	-0.023	-0.297	0.001	0.189	0.049
	g^{λ}_{65}	-0.003	-0.3	0.172	0.136	0.123
	g^{λ}_{66}	0.007	-0.051	0.194	0.057	0.127
	g^{λ}_{67}	0.002	-0.186	0.2	0.097	0.111
	g^{λ}_{68}	-0.02	-0.3	0.026	0.16	0.062
	g^{μ}_{61}	-0.023	-0.3	0.062	0.202	0.074
	g^{μ}_{62}	-0.005	-0.074	0.0	0.11	0.019
	g^{μ}_{63}	0.034	-0.037	0.3	0.062	0.25
	g^{μ}_{64}	0.01	-0.123	0.254	0.091	0.123
	g^{μ}_{65}	-0.002	-0.3	0.178	0.136	0.126
	g^{μ}_{66}	0.034	0.0	0.273	0.03	0.222
	g^{μ}_{67}	-0.027	-0.3	0.086	0.242	0.079
	g^{μ}_{68}	0.024	-0.069	0.3	0.075	0.201

Table S12: Posterior parameter estimates for Barbourofelidae under the MCDD model. The clades are here numbered as follows: 1) Hesperocyoninae, 2) Borophaginae, 3) Caninae, 4) Amphicyonidae, 5) Nimvravidae, 6) Felidae, 7) Barbourofelidae, 8) Ursidae.

Parameter		Mean	95%HPD		Pos. interaction P(g<0)	Competition P(g>0)
Baseline rates	λ_7	0.373	0.007	1.2		
	μ_7	0.385	0.018	1.015		
Hyperprior	η_7	0.418	0.0	0.871		
MCDD	g^{λ}_{71}	-0.048	-0.3	0.189	0.426	0.178
	g^{λ}_{72}	-0.04	-0.253	0.05	0.345	0.078
	g^{λ}_{73}	0.015	-0.227	0.3	0.24	0.329
	g^{λ}_{74}	-0.047	-0.3	0.144	0.41	0.144
	g^{λ}_{75}	0.001	-0.248	0.299	0.288	0.293
	g^{λ}_{76}	0.033	-0.124	0.3	0.146	0.361
	g^{λ}_{77}	0.125	-0.105	0.3	0.097	0.66
	g^{λ}_{78}	-0.01	-0.3	0.231	0.305	0.261
	g^{μ}_{71}	-0.014	-0.3	0.236	0.326	0.259
	g^{μ}_{72}	-0.082	-0.268	0.0	0.626	0.03
	g^{μ}_{73}	0.058	-0.155	0.3	0.153	0.447
	g^{μ}_{74}	-0.061	-0.3	0.2	0.481	0.189
	g^{μ}_{75}	-0.001	-0.275	0.273	0.293	0.289
	g^{μ}_{76}	0.109	0.0	0.3	0.049	0.601
	g^{μ}_{77}	0.009	-0.237	0.3	0.27	0.31
	g^{μ}_{78}	0.044	-0.188	0.3	0.185	0.406

Table S13: Posterior parameter estimates for Ursidae under the MCDD model. The clades are here numbered as follows: 1) Hesperocyoninae, 2) Borophaginae, 3) Caninae, 4) Amphicyonidae, 5) Nimvravidae, 6) Felidae, 7) Barbourofelidae, 8) Ursidae.

Parameter		Mean	95%HPD		Pos. interaction P(g<0)	Competition P(g>0)
Baseline rates	λ_8	0.348	0.081	0.691		
	μ_8	0.239	0.067	0.384		
Hyperprior	η_8	0.794	0.339	1.0		
MCDD	g^{λ}_{81}	0.01	-0.005	0.199	0.038	0.124
	g^{λ}_{82}	-0.005	-0.053	0.004	0.069	0.015
	g^{λ}_{83}	-0.003	-0.077	0.056	0.069	0.042
	g^{λ}_{84}	-0.035	-0.254	-0.0	0.214	0.011
	g^{λ}_{85}	0.027	-0.0	0.29	0.044	0.193
	g^{λ}_{86}	-0.003	-0.064	0.029	0.062	0.034
	g^{λ}_{87}	0.008	-0.072	0.27	0.065	0.114
	g^{λ}_{88}	0.03	-0.0	0.219	0.024	0.248
	g^{μ}_{81}	0.02	-0.0	0.239	0.025	0.144
	g^{μ}_{82}	0.007	-0.006	0.079	0.014	0.082
	g^{μ}_{83}	0.018	-0.0	0.246	0.027	0.126
	g^{μ}_{84}	0.027	0.0	0.255	0.022	0.182
	g^{μ}_{85}	-0.001	-0.221	0.176	0.093	0.085
	g^{μ}_{86}	-0.005	-0.181	0.047	0.08	0.049
	g^{μ}_{87}	-0.005	-0.286	0.088	0.107	0.075
	g^{μ}_{88}	0.033	-0.0	0.271	0.023	0.197

References

1. R H Tedford, X Wang, and B E Taylor. Phylogenetic systematics of the North American fossil Caninae (Carnivora: Canidae). *Bull Am Mus Nat Hist*, (325):1–218, 2009.

2. B Van Valkenburgh, X Wang, and J Damuth. Cope's rule, hypercarnivory, and extinction in north american canids. *Science*, 306(5693):101–104, 2004.

3. F M Gradstein, J G Ogg, M Schmitz, and G Ogg, editors. *The Geologic Time Scale*. Elsevier, 2012.

4. D Silvestro, J Schnitzler, L H Liow, A Antonelli, and N Salamin. Bayesian estimation of speciation and extinction from incomplete fossil occurrence data. *Syst Biol*, 63(3): 349–367, 2014.

5. D Silvestro, B Cascales-Miñana, C D Bacon, and A Antonelli. Revisiting the origin and diversification of vascular plants through a comprehensive Bayesian analysis of the fossil record. *New Phytol*, doi:10.1111/nph.13247, 2015.

6. D Silvestro, N Salamin, and J Schnitzler. PyRate: A new program to estimate speciation and extinction rates from incomplete fossil record. *Methods Ecol Evol*, 5:1126–1131, 2014.

7. M Stephens. Bayesian analysis of mixture models with an unknown number of components – an alternative to reversible jump methods. *Ann Stat*, 28(1):40–74, 2000.

8. D Silvestro and J Schnitzler. PyRate manual: available from http://sourceforge.net/projects/pyrate/files/manual.pdf, 2014.

9. N Keiding. Maximum likelihood estimation in the birth-death process. *The Annals of Statistics*, 3(2):363–372, 1975.

10. A Rambaut and A J Drummond. Tracer: Available from http://beast.bio.ed.ac.uk/tracer., 2007.

11. N Lartillot and R Poujol. A phylogenetic model for investigating correlated evolution of substitution rates and continuous phenotypic characters. *Mol Biol Evol*, 28(1):729–744, 2011.

12. J C Zachos, G R Dickens, and R E Zeebe. An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature*, 451(7176):279–283, 2008.

13. F L Condamine, J Rolland, and H Morlon. Macroevolutionary perspectives to environmental change. *Ecol Lett*, 16 Suppl 1:72–85, 2013.

14. R S Etienne, B Haegeman, T Stadler, T Aze, P N Pearson, A Purvis, and A B Phillimore. Diversity-dependence brings molecular phylogenies closer to agreement with the fossil record. *Proc R Soc Lond B*, 279(1732):1300–1309, 2012.

15. N D Newell. Adequacy of the fossil record. *Journal of Paleontology*, 33(3):488–499, 1959.

16. J Alroy, C R Marshall, R K Bambach, K Bezusko, M Foote, F T Fursich, T A Hansen, S M Holland, L C Ivany, D Jablonski, D K Jacobs, D C Jones, M A Kosnik, S Lidgard, S Low, A I Miller, P M Novack-Gottshall, T D Olszewski, M E Patzkowsky, D M Raup, K Roy, J J Sepkoski, M G Sommers, P J Wagner, and A Webber. Effects of sampling standardization on estimates of phanerozoic marine diversification. *Proc Natl Acad Sci USA*, 98(11):6261–6266, 2001.

17. M Foote and D M Raup. Fossil preservation and the stratigraphic ranges of taxa. *Paleobiology*, 22(2):121–140, 1996.

18. J Alroy. *Fair sampling of taxonomic richness and unbiased estimation of origination and extinction rates.*, pages 55–80. Number 16. Paleontological Society Papers, 2010.

19. L H Liow and J D Nichols. *Estimating rates and probabilities of origination and extinction using taxonomic occurrence data: Capture-recapture approaches.*, pages 81–94. University of California Press, 2010.

20. L H Liow and J A Finarelli. A dynamic global equilibrium in carnivoran diversification over 20 million years. *Proc R Soc Lond B*, 281(1778):20132312, 2014.

21. J Alroy. Cope's rule and the dynamics of body mass evolution in North American fossil mammals. *Science*, 280(5364):731–734, 1998.

22. J Alroy, P L Koch, and J C Zachos. Global climate change and North American mammalian evolution. *Paleobiology*, 26(4):259–288, 2000.

23. J Alroy. *Speciation and extinction in the fossil record of North American mammals*, pages 301–323. Cambridge University Press, 2009.

24. X Wang and R H Tedford. *The Dog Family, Canidae, and Their Evolutionary History.* Columbia University Press, New York, 2008.

25. A Gelman, J B Carlin, H S Stern, and D B. Rubin. *Bayesian Data Analysis, Second Edition (Chapman & Hall/CRC Texts in Statistical Science).* 2004.

26. A Gelman, J Hill, and M Yajima. Why we (usually) don't have to worry about multiple comparisons. *Journal of Research on Educational Effectiveness*, 5(2):189–211, 2012.

27. R B O'Hara and M. J Sillanpää. A review of Bayesian variable selection methods: what, how and which. *Bayesian Anal*, 4(1):85–117, 2009.

28. L Kuo and B Mallick. Variable selection for regression models. *Sankhya Ser. B*, 60(1): 65–81, 1998.

29. R E Kass and A E Raftery. Bayes factors. *J Amer Stat Assoc*, 90(430):773–795, 1995.

30. A Miller. *Subset Selection in Regression.* Chapman and Hall/CRC, Boca Raton, Florida, USA, 2002.

31. C M Carvalho, N G Polson, and J G Scott. The horseshoe estimator for sparse signals. *Biometrika*, 97(2):465–480, 2010.

32. F Ronquist, J P Huelsenbeck, P Van Der Mark, and P Lemey. Bayesian phylogenetic analysis using MrBayes. *The Phylogenetic Handbook*, Chapter 7:1–63, 2007.

33. B Van Valkenburgh. Major patterns in the history of carnivorous mammals. *Annu. Rev. Earth Planet. Sci.*, 27:463–493, 1999.

34. J J Sepkoski, F K McKinney, and S Lidgard. Competitive displacement among post-Paleozoic cyclostome and cheilostome bryozoans. *Paleobiology*, 26(1):7–18, 2000.

35. B Van Valkenburgh. Déjà vu: the evolution of feeding morphologies in the Carnivora. *Integr Comp Biol*, 47(1):147–163, 2007.