The rise of grasslands is linked to atmospheric CO_2 decline in the late Palaeogene

Luis Palazzesi[1,2,*], Oriane Hidalgo[2,3], Viviana D. Barreda[1], Félix Forest[2,†], and Sebastian Höhna[4,5,,†]**

[1]Museo Argentino de Ciencias Naturales & Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires C1405DJR, Argentina
[2]Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, United Kingdom
[3]Institut Botànic de Barcelona (IBB, CSIC-Ajuntament de Barcelona), Catalonia, Spain
[4]GeoBio-Center, Ludwig-Maximilians-Universität München, Richard-Wagner-Str. 10, 80333 Munich, Germany
[5]Department of Earth and Environmental Sciences, Paleontology & Geobiology, Ludwig-Maximilians-Universität München, Richard-Wagner-Str. 10, 80333 Munich, Germany
[*]lpalazzesi@macn.gov.ar
[**]hoehna@lmu.de
[†]Equal contribution

ABSTRACT

Grasslands are predicted to experience a major biodiversity change by the year 2100 in part due to recent and projected increases in atmospheric CO_2 concentration. A better understanding of how grasslands have responded to past environmental changes will help predict the outcome of current and future environmental changes. Here, we explore the relationship between past atmospheric CO_2 and temperature fluctuations and the shifts in diversification rate of grasses (Poaceae) and daisies (Asteraceae), two exceptionally species-rich grassland families (~11,000 and ~23,000 species, respectively). To this end, we developed a novel Bayesian approach that simultaneously estimates diversification-rates through time from time-calibrated phylogenies and correlations between environmental variables and diversification rates. Additionally, we developed a new statistical approach that incorporates the information of the distribution of missing species in the phylogeny. We found strong evidence supporting a simultaneous increase in diversification rates for daisies and grasses after the most significant reduction of atmospheric CO_2 in the Cenozoic (~34 Mya). The fluctuations of paleo-temperatures, however, appear not to have had a significant relationship with the diversification of these grassland families. Overall, our results shed new light on our understanding of the origin of grasslands in the context of past environmental changes.

Introduction

The grassland biome (steppes, savannas and prairies) covers vast areas of the Earth's surface and today accounts for as much as one-third of the net primary production on land[1,2]. Although grasses (Poaceae) comprise the bulk of the biomass and plant population in grasslands, other plant families—in particular the daisies (Asteraceae)—are usually as much as (or even more) diverse than grasses (Fig. S1). The evolution of grasslands marked the emergence of a new landscape and provided the substrate for the adaptive radiation of other life forms that coevolved along with this biome, including grazing mammals[3] such as horses, wombats, and capybaras.

The age of a given biome is often estimated by detecting when particular representative taxonomic groups first appear in the fossil record. For example, the early evolution of the grassland biome—and open-habitat biomes in general—has been estimated from the fossil record of grass phytoliths (plant silica)[4] or from the record of fossil pollen of daisies, grasses and amaranths[5,6]. Phylogenetic trees based on DNA sequence data calibrated with fossils provide a powerful new perspective on the history of biomes[7]. This approach has been used to estimate the timing of tropical-rainforest evolution based on phylogenetic trees of plant groups that are characteristic of this biome (e.g., Malpighiales[8], Arecaceae[9], and the legume genus *Inga*[10]). Nevertheless, phylogenetic approaches have barely been used to study the evolutionary history of grassy biomes; most previous studies of grassland evolution have focused on the origins of C_4 grasslands[11]. Here we estimate when grasslands first expanded using phylogenetic trees of its two primary plant families, Asteraceae and Poaceae. We assembled a large calibrated phylogenetic tree for daisies and used the largest tree yet inferred for grasses[11] to explore temporal shifts in rates of lineage diversification, and to test correlations between diversification-rate shifts and past climatic fluctuations.

A major limitation when analyzing hyper-diverse groups—in our case Asteraceae with ~23,000 species and Poaceae with ~11,000 species—is the inevitable sparse species sampling (Figs. 1, 2). Although existing approaches for inferring rates

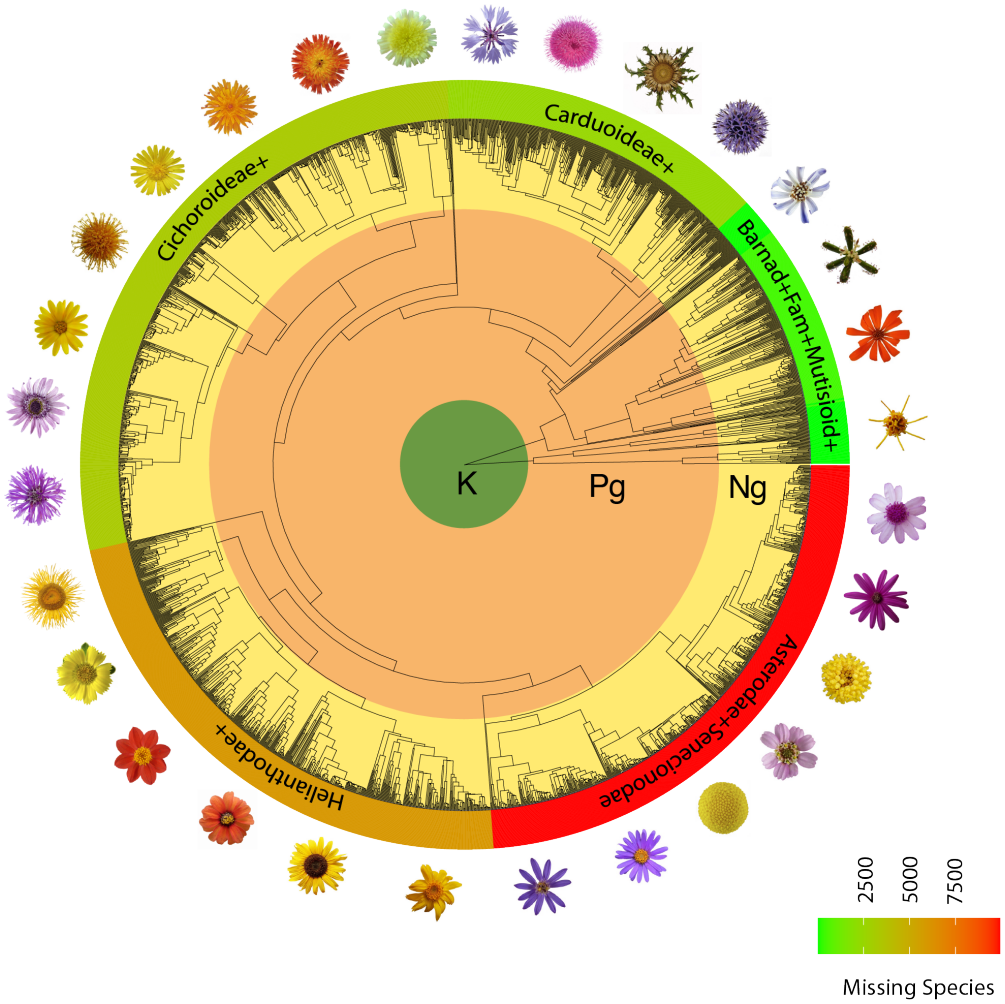

Figure 1. Phylogenetic tree scaled to geological time of Asteraceae with 2723 sampled tips. Asteraceae is one of the most species-rich families of flowering plants with more than 23,000 species. The number of non-sampled (missing) species increased enormously towards the more derived and specious lineages. For this reason, the sampling among clades is severely biased. Note that these rich and derived lineages evolved during the late Paleogene or early Neogene. K=Cretaceous, Pg=Paleogene, Ng=Neogene.

of lineage diversification (speciation and extinction) can accommodate incomplete species sampling[12,13], the distribution of missing species on the tree in these approaches is modeled in a simplistic and somewhat unrealistic manner. Previous work has shown that biased species sampling has a strong impact on diversification-rate estimates[14–16]. We develop a novel Bayesian approach for detecting diversification-rate shifts that incorporates a more realistic (non-uniform) model of species sampling and implemented it in the open-source software `RevBayes`[17]. Our model builds on the episodic birth-death process, where speciation and extinction rates are constant within an interval but may shift instantly to new rates at a rate-shift episode[18–21]. Furthermore, we tested for a correlation between diversification rate and two environmental variables —atmospheric CO_2 concentration and average global paleo-temperature— using one existing[22–27] and three new environmentally-dependent diversification models. We used an empirically informed and biologically realistic model to accommodate missing species that assigns unsampled species to their corresponding clades using taxonomic information.

Results and Discussions

Our analyses demonstrate that the most dramatic increase in diversification rates in both Asteraceae and Poaceae (calibration scenario #1, see Methods) occurred from the late Oligocene (~28 Mya) to the early Miocene (~20 Mya) (Fig. 3 and Fig. S5). This diversification rate shift are robust to several model assumptions. We recovered the same diversification rate shifts regardless of the assumed number of time intervals (Fig. S6). Both autocorrelated diversification rate prior models qualitatively agree on the overall pattern of diversification rates (Gaussian Markov random field (GMRF) or Horseshoe Markov random field (HSRMF), Fig. S5 and S6). Only the uncorrelated diversification rate prior model differed in the inferred pattern (UCLN,

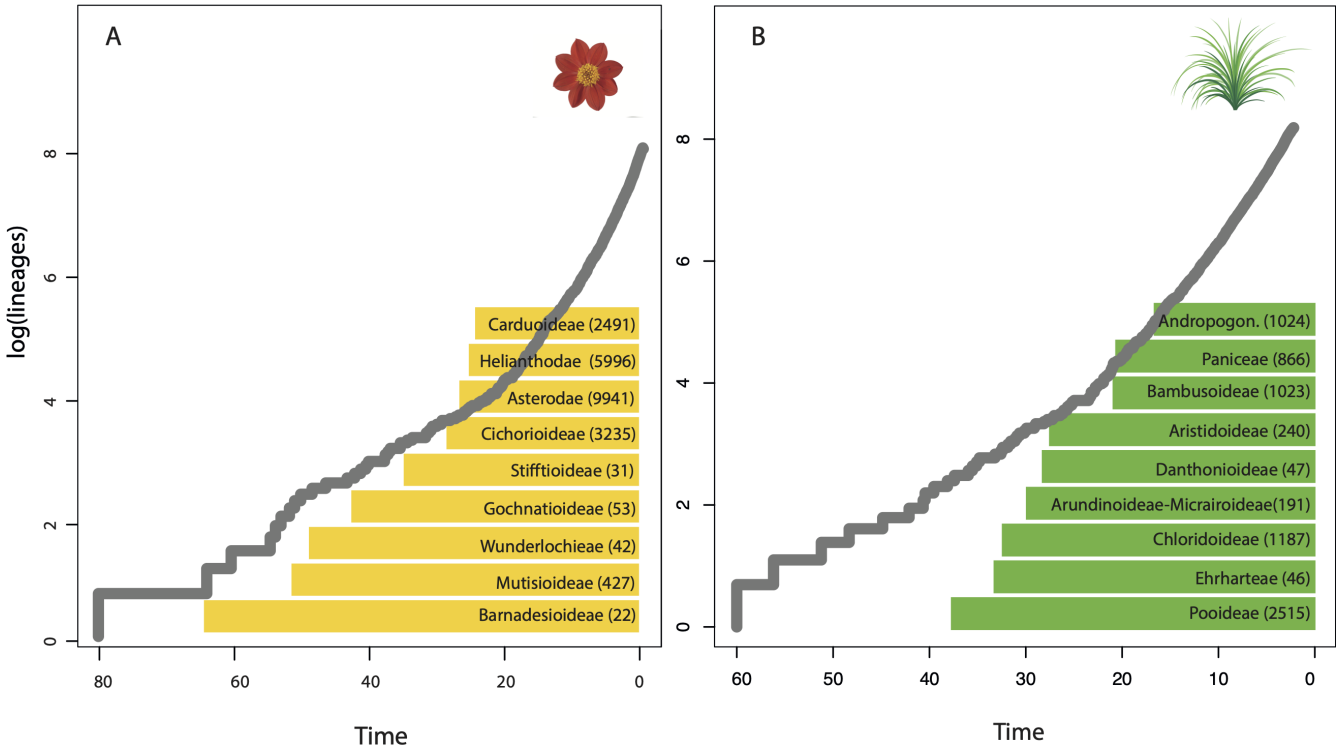

Figure 2. Lineage-Through Time (LTT) plots of Asteraceae (A) and Poaceae (B). Solid grey lines represent LTT curves derived from time-calibrated phylogenetic trees of Asteraceae and Poaceae (calibration scenario #1). Colored boxes depict the name and number of non-sampled (missing) species per clade that we integrated in our novel empirical taxon sampling. The shape of LTT curves have demonstrated to be a convenient summary metric for diversification diagnostics, particularly when diversification deviates from the expectation of constant rates[28,29]. However, the distribution of missing species might not be uniform —as it is the case of these angiosperm families— and can severely impact on diversification-rate estimates. Our work shows that the most important increase in diversification rate for both Asteraceae and Poaceae is completely unnoticeable using the LTT analysis, even when calibrated phylogenetic trees include a large number of species.

Fig. S5 and S6). However, the autocorrelated diversification rate prior models were significantly favored according to our Bayes factor analyses (GMRF for the daisy phylogenetic tree and HSMRF for the grass phylogenetic tree, Fig. S7). The diversification rate patterns were strongly influenced by the assumed incomplete taxon sampling (Fig. S8). In our simulation study we show that incorrectly assuming uniform taxon sampling and thus disregarding taxonomic information about the distribution of missing species strongly biases diversification rates (Fig. S22). Conversely, our empirical taxon sampling informed by a more accurate distribution of missing species has good power to detect the correct time-varying diversification rates and low false-positive rate when diversification rates are in reality constant (Fig. S22). Thus, we recommend to include as much information as possible regarding the distribution of missing species.

The respective diversification rates of Asteraceae and Poaceae (calibration scenario #1, see Methods) peak between 20 Mya and 15 Mya, and subsequently decreases for a brief period of time before increasing again from the late Miocene (~10 Mya, Fig. 3 and Fig. S5). Our second analysis using the Poaceae phylogeny calibrated with a Cretaceous phytolith (calibration scenario #2) detects an earlier peak for Poaceae at about 30 Mya (Fig. S4). The phylogenetic placement of this fossil phytolith has been debated[30], thus this last result should be considered with caution. Our estimates of low diversification rates prior to the Oligocene is consistent with the scarcity of fossil forms assigned to both daisies (Table S1) and grasses[4,31] known from this period. Similarly, our estimates of a high diversification rate in the late Oligocene and early Miocene is in line with the high diversity of fossil remains assigned to these groups[4,32]. The Cenozoic 'temporal hotspot' of grassland diversification (~30 Mya to ~15 Mya) –based on daisies and grasses (calibration #1 and #2) phylogenetic trees– coincides with one of the most fundamental changes in global climate in the geologic record; a marked decline of atmospheric CO_2 occurred during the Oligocene (~34 Mya), reaching modern levels by the latest Oligocene[33,34]. This scenario marks the onset of a cooler and more modern world (Coolhouse state), identified by the earliest Cenozoic glaciations in Antarctica, and the consequent drop in global paleotemperatures[35].

In line with the reconstructed climatic scenario, our analyses of correlation between diversification rates and CO_2 or paleo-temperature show very interesting results (Fig. 3 and S8). Diversification rates inferred from both the daisy and grasses phylogeny support correlation to CO_2 over paleo-temperature (Fig. S10). Surprisingly, the best fitting environmentally-

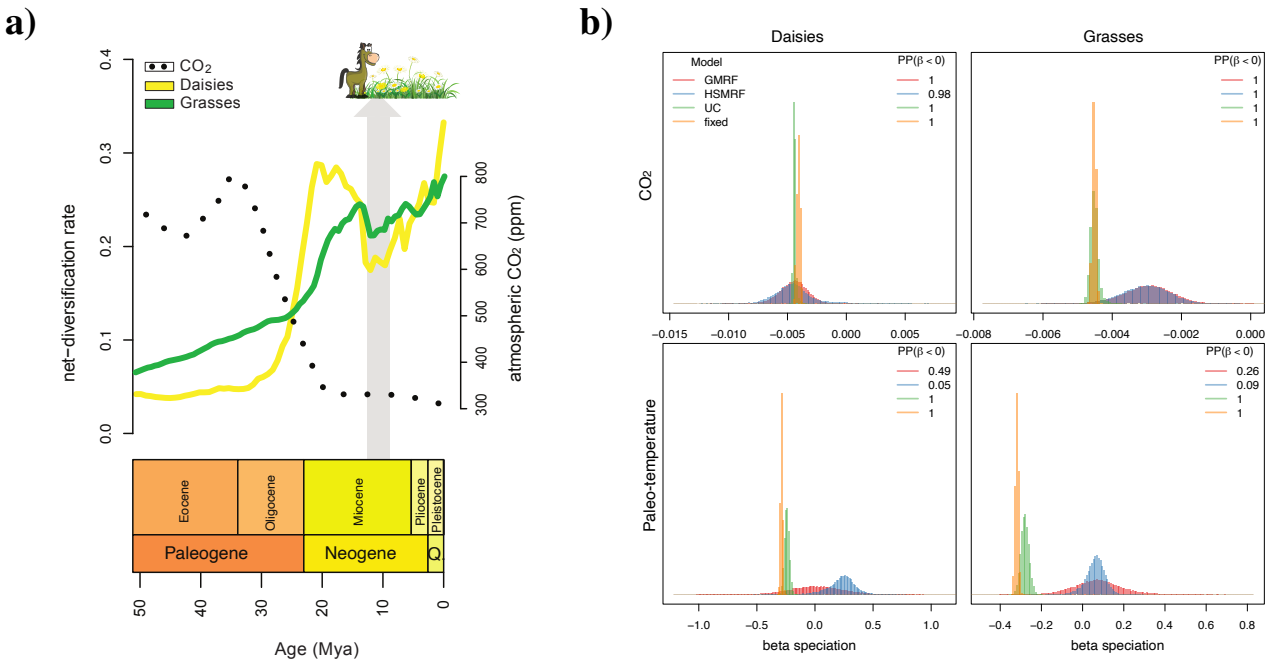

Figure 3. Estimating diversification rates and correlation to CO_2 and paleo-temperature. **a)** Diversification rates through time of daisies (yellow) and grasses (green) –calibration scenario #1– for the last 50 Mya. Dotted line represents atmospheric CO_2 fluctuations (22); note that the Oligocene steep decrease mirrors the onset of the increase in diversification of daisies and grasses. Grey arrow indicates a period (13-10 Mya) of lower diversification rates probably linked either with a brief increase in CO_2 (not represented in the dotted smoothed curve) or the explosive radiation of hypsodont grazers (*e.g.,* horses) and other mixed feeder grazers[3] who may have had a tremendous impact on grasslands through their effects on plant populations and community composition. **b)** The correlation coefficient (β) between diversification rates and CO_2 concentrations for daisies and grasses is significantly negative (posterior probability of 1.0 for all models except the HSRMF, which has a posterior probability of 0.98; Bayes factors of 37,501 and 49 respectively). The support for a negative correlation between paleo-temperature and diversification rates is ambiguous; the UC and *fixed* models show significant support (posterior probability of 1.0, Bayes factor of 37,501) while the autocorrelated models show no support (posterior probability between 0.05 and 0.095, Bayes factors supporting a *positive* correlation of 1.04 and 17.86 for the daisy dataset and 2.85 and 9.75 for the grasses dataset for the GMRF and HSMRF models respectively).

dependent diversification model for the daisy phylogeny was the uncorrelated lognormal (UC) variation model and for the grasses phylogeny the *fixed* rate model without additional variation. The support of the uncorrelated model over the two autocorrelated models (GMRF and HSRMF), although the autocorrelated models were favored when using time-varying diversification rates without environmental variables (Fig. S7), could stem from the use of vague prior distribution which allows for more rate variation in autocorrelated models[21]. However, regardless of the specific environmentally-dependent diversification model, we inferred a negative correlation between diversification rates and environmental CO_2 (Fig. S9). The resulting Bayes factors for a negative correlation were decisive with values of 37,501 for the *fixed*, UC and GMRF models and 49 for the HSMRF model (Fig. 3). We also see the same agreement between the four environmentally-dependent diversification models in our simulation study (Fig. S19 and S20). Thus, if there is a clear signal of correlation between the environmental variable and diversification rates, then our analyses appear robust to modeling of the additional component of time-varying diversification rates. This agreement can also be seen when all four environmentally-dependent diversification models show the same estimated diversification rates (Fig. S13-14). When the signal is less clear, as for the paleo-temperature analyses, the four models disagree and range from significant positive to significant negative correlation and the estimated diversification rates of the environmentally-dependent diversification models also differ (Fig. S13-14). Finally, our results of correlation between environmental CO_2 and diversification rates are also robust to the chosen epoch size (Fig. S11).

The negative correlation between diversification rates of these selected grassland families and atmospheric CO_2 might not be surprising; atmospheric CO_2—the main source of carbon for photosynthesis— serves as a fundamental substrate for plant growth. The available experimental evidence shows that low atmospheric CO_2 limits plant performance[36], although responses vary significantly between species. At a landscape scale, carbon limitation and water stress due to lower atmospheric CO_2

concentrations ('ecophysiological drought'), rather than water stress due to lower precipitation ('climatic drought'), cause changes in vegetation structure[37]. During the Last Glacial Maximum (LGM; ~21,000 years ago), for example, atmospheric CO_2 was at its lowest concentration in the history of land plants (~180–200 ppm)[38]. Models have predicted that the direct physiological impact of the of low CO_2 concentrations during the LGM drove the expansion of grasslands and dry shrublands at the expense of forest[39] (Fig. S2). Other modeling experiments indicate that low atmospheric CO_2, in combination with increased aridity and decreased temperatures, causes new xeric biomes to develop[38].

Although our primary hypothesis is that a CO_2-depleted atmosphere played a role in the geographic expansion and diversification of grassland families since the Oligocene, other environmental and biological variables could have also been involved. In particular, the decreasing temperatures, increasing aridity, and increasing seasonality of temperature and/or precipitation of the late Cenozoic have been traditionally linked to the early radiation of grasslands[40,41]. The role of cooling in the emergence of open-habitat grasses has been debated as the adaptation to low temperatures became prominent in the more derived groups of grasses[4,42]. Grazing mammals have been also important components in the evolution of grasslands; grazers and grassland ecosystems probably coevolved over millions of years[43]. Grazing increased species diversity according to experimental studies, as grazers prevent dominant plant species from monopolizing resources. Without grazing, tall, vegetatively reproducing plant species increase in cover and shade out short and sexually reproducing species[44]. Grazing also affects the flux of nutrients by accelerating the conversion of plant nutrients from forms that are unavailable for plant uptake to forms that can be readily used. Overall, grazing mammals have an important role in the diversity of present-day natural grasslands and we assume they might have done so during their early radiation. However, the explosive radiation of true hypsodonts may have negatively impacted grasslands' distribution and diversity (see below). Sorting out the relative importance of all these environmental and biological competing forces from the hypothesized CO_2-induced shift is challenging.

We detected a short decrease in diversification rates for daisy and grass plant groups during the mid-Miocene, about 13-10 Mya (Fig. 3). The causal mechanism underlying remains to be elucidated. However, we suspect that the dramatic radiation of hypsodont grazers –such as horses– and other mixed feeder grazers may have had an impact on grasslands[3,45]. Since the late Miocene (~10 Mya), however, the more recent expansion of C_4 grass lineages[11] may have contributed to the increased diversification rates in these groups. Plants using the C_4 photosynthetic pathway have anatomical and biochemical adaptations for concentrating CO_2 within leaf cells prior to photosynthesis, which may lead to a selective advantage over C_3 plants under conditions of low atmospheric CO_2. Although the evolutionary origin of C_4 photosynthesis in grasses most likely occurred early in the Cenozoic[30], their expansion and ecological dominance may have taken place during the last 10 Mya, by the late Miocene in warmer and fire-prone landscapes of the world[46]. Likewise, the evolution of hyper-diverse Asteraceae lineages (*e.g. Senecio*)[47] have also contributed to the increasing rates of diversification since the last 10 Mya. Our evidence also supports the notion that the ongoing rise of atmospheric CO_2 will likely altered vegetation distributions through differential effects on C_3 and C_4 plant types. In fact, modelling future distributions predicts the near-complete eradication of C_4 species across the globe for the next 50 year[48]; this implies that about half of the species in the grass family will be extinct. In summary, our study reveals episodic shifts in diversification rates of grasses and daises which are correlated with changes in atmospheric CO_2 (Fig. 3); these insights are made possible by the development of our new Bayesian phylogenetic approach which combines the episodic birth-death process[18–21] with environmentally-dependent diversification rates[22–24,26] and empirical taxon sampling[15,16,49]. Our environmentally-dependent and episodic birth-death diversification model provides a novel approach for exploring the evolution of hyper-diverse groups of plants and animals in the context of historical environmental changes.

Material and Methods

Grassland diversity. To quantify the taxonomic representativeness of vascular-plant families found in open-habitat landscapes (Fig. S1), we selected seven distantly distributed eco-regions dominated by grasslands from the World Wide Fund for Nature (WWF)[50]. Using the coordinate boundaries of each of the selected eco-regions, we extracted the vascular plant taxa (=Tracheophyta) from the Global Biodiversity Information Facility (GBIF), using the R 'RGBIF' package[51] with the option "hasGeospatialIssue=FALSE", that includes only records without spatial issues. Plant families were sorted according to the number of species, removing duplicated species.

Palaeobotanical analysis. Asteraceae and Poaceae have a fairly similar fossil record; their oldest findings are known from the Late Cretaceous —which mainly comprise microscopic remains (that is, phytoliths[52] or pollen grains[53])— whereas the first indisputable macroscopic Asteraceae and Poaceae fossils are first known from the Eocene[54,55], with a substantial increase of diversity since since the Oligocene/Miocene. While the fossil record of Poaceae has been fully revised[4,56,57], the fossil record of Asteraceae has not been as carefully reviewed. We compiled published pollen and macroscopic fossil data for Asteraceae including all fossil species assigned to Asteraceae (Table S1). The earliest record of the Asteroideae (the clade that includes the most common open-habitat daisy tribes) occurs since the Late Oligocene of New Zealand but in very low frequencies. Fossils refer to this subfamily increased in abundance and diversity during the Miocene and Pliocene. Pollen referred to *Artemisia*, in

particular, did not become abundant until the Middle-Late Miocene with several reports from central Europe, Asia and North America. Pre-Miocene findings need further verification. Overall, the Late Oligocene and in particular the Miocene witnessed the major step in the diversification of Asteraceae; ca. 80% of the fossil species recorded have been assigned to this time interval.

Divergence-Time estimation. To construct the Asteraceae supertree (2,723 tips), we first inferred a backbone chronogram using 14 plastid DNA regions from 54 species, including representatives of all 13 subfamilies, with an additional four species of Calyceraceae used as outgroup taxa (Table S2). Sequences were compiled from GenBank and each region was aligned separately using MAFFT[58] with the options maxiterate 1000 and localpair. Two fossil constraints were applied: (i) a macrofossil (capitulum) and associated pollen (*Raiguenrayun cura* + *Mutisiapollis telleriae*) from the Eocene (45.6 Mya) to calibrate the non-Barnadesioideae Asteraceae clade[55] and; (ii) the fossil-pollen species *Tubulifloridites lilliei* type A from the late Cretaceous (72.1 Mya)[53] to calibrate the crown Asteraceae (considering *T. lilliei* as a stem group, see Huang *et al.*[59] for further discussion). Divergence-time estimates and phylogenetic relationships were inferred using `RevBayes`[17]. For the aligned molecular sequences we assume a general-time reversible substitution model with gamma-distributed rate variation among sites (GTR+Γ), an uncorrelated log-normal prior on substitution-rate variation across branches (UCLN relaxed clock), and a birth–death prior model on the distribution on node ages/tree topologies. A densely sampled phylogeny is crucial to identify shifts in diversification rates. Therefore, we constructed a supertree by inserting eleven individual sub-trees —representing all subfamilies of the Asteraceae except those less diverse or monotypic clades (that is, Gymnarrhenoideae, Corymbioideae, Hecastocleidoideae, Pertyoideae)— into the calibrated backbone chronogram. This method follows a previous study that constructed a supertree of grasses using the same approach[11]. Each of the eleven clades of Asteraceae was built using their own set of markers and the same phylogenetic approach as the one used to infer the backbone tree (Table S2). Sequence data for each of the eleven trees and their respective outgroup taxa were collected from Genbank using the NCBIminer tool[60]. The estimated ages of the nodes given by the backbone analysis were used to constrain the age of each of the eleven sub-trees (Table S2). Divergence-time estimates and phylogenetic relationships for each of the eleven sub-clades were estimated using `RevBayes` as described above. The eleven trees were grafted onto the backbone tree using the function 'paste.tree' from the `phytools` R package[61]. We used `GGTREE` R package[62] to plot the circle phylogenetic tree of Figure 1 and `phytools`[61] to include the concentric geological scale. The supertree of the grass family (3,595 taxa) was obtained from Spriggs *et al.*[11] (Table S3). They inferred two chronograms using two different calibration scenarios, that is, a younger scenario (#1) calibrated using an Eocene megafossil[54] and an older scenario (#2) calibrated using Cretaceous phytoliths[52]. We run our diversification analyses using these two chronograms.

Inferring Changes in Diversification Rate Through Time. Our species-diversification model is based on the *reconstructed evolutionary process* described by Nee *et al*[12] and more specifically on the episodic birth-death process[18–21]. We assume that each lineage gives birth to another species with rate λ (cladogenetic speciation events) and dies with rate μ (extinction event; see Figure 4). We model diversification rates (i.e., speciation and extinction rates) as constant within an interval but independent between intervals, where intervals are demarcated by instantaneous rate-shift events. We denote the vector of speciation rates $\Lambda = \{\lambda_1, \ldots, \lambda_k\}$ and extinction rates $\mathbf{M} = \{\mu_1, \ldots, \mu_k\}$ where λ_i and μ_i are the (constant) speciation and extinction rates in interval i. Additionally, we use the taxon sampling fraction at the present denoted by ρ[15,16]. Following the notation of May *et al.*[20], we construct a unique vector, $\mathbb{X}$, that contains all divergence times and rate-shift event times sorted in increasing order. It is convenient for notation to expand the vectors for all the other parameters so that they have the same number of elements $k = |\mathbb{X}|$. Let Ψ denote an inferred tree relating n species, comprising a tree topology, τ, and the set of branching times, $\mathbb{T}$. We use the notation $S(2, t_1 = 0, T)$ to represent the survival of two lineages in the interval $[t_1, T]$, which is the condition we enforce on the reconstructed evolutionary process. Transforming Equation (A4) in May *et al*[20] to our model yields the probability density of a reconstructed tree as:

$$
\begin{aligned}
&f(\Psi|N(t_1=0)=2,S(2,t_1=0,T)) \\
&= \frac{2^{n-1}}{n!} \\
&\times \left(1+\sum_{i=0}^{k}\left(\frac{\mu_i}{\mu_i-\lambda_i}\times e^{\sum_{j=0}^{i-1}(\mu_j-\lambda_j)(x_{j+1}-x_j)}\times\left(e^{(\mu_i-\lambda_i)(x_{i+1}-x_i)}-1\right)\right)-\frac{\rho-1}{\rho}\times e^{\sum_{i=0}^{k}(\mu_i-\lambda_i)(x_{i+1}-x_i)}\right)^{-2}\times\left(e^{-\log(\rho)\sum_{j=0}^{k}(\mu_j-\lambda_j)(x_{j+1}-x_j)}\right)^2 \\
&\times\prod_{i\in\mathbb{I}_T}\left[\lambda_i\times\left(1+\sum_{l=i}^{k}\left(\frac{\mu_l}{\mu_l-\lambda_l}\times e^{\sum_{j=0}^{l-1}(\mu_j-\lambda_j)(x_{j+1}-x_j)}\times\left(e^{(\mu_l-\lambda_l)(x_{l+1}-x_l)}-1\right)\right)-\frac{\rho-1}{\rho}\times e^{\sum_{i=i}^{k}(\mu_l-\lambda_l)(x_{l+1}-x_l)}\right)^{-2}\right. \\
&\left.\times e^{-\log(\rho)\sum_{j=i}^{k}(\mu_j-\lambda_j)(x_{j+1}-x_j)}\right]. \qquad (1)
\end{aligned}
$$

The first term, $\frac{2^{n-1}}{n!}$ corresponds to the combinatorial constant for the number of labelled histories[18], the second term corresponds to the condition of two initial lineage at the root of the phylogeny surviving until the present, and the third term corresponds to the product of all speciation events and the new lineages surviving until the present.

Empirical taxon-sampling model. Here we develop an *empirical* taxon sampling model that uses taxonomic information on the membership of unsampled species to clades and speciation times of unsampled species, which is an extension to the work by Höhna *et al*[15,16] and similar to the approach used by Stadler and Bokma[49]. The main difference of our approach and the approach by Stadler and Bokma[49] is that their model uses a constant-rate birth-death process (compared to our episodic birth-death process). Additionally, Stadler and Bokma[49] derive the density of the missing species using a random probability s of an edge being sampled, which differs from our approach where we integrate over the time of the missing speciation event. Nevertheless, at least for the constant-rate birth-death process, both approaches arrive at the same final likelihood function.

We include information on the missing speciation events by integrating over the known interval when these speciation events must have occurred (that is, between the stem age t_c of the MRCA of the clade and the present). This integral of the probability density of a speciation event is exactly the same as one minus the cumulative distribution function of a speciation event[16],

$$F(t_c|N(t_1)=1,t_1\leq t\leq T)=1-\frac{1-P(N(T)>0|N(t_c)=1)\exp(r(t_c,T))}{1-P(N(T)>0|N(t_1)=1)\exp(r(t_1,T))}, \qquad (2)$$

where t_1 is the age of the root. The probability of survival is given by:

$$
\begin{aligned}
&P(N(T)>0|N(t_c)=1) \\
&= \left(1+\sum_{i=c}^{k}\left(\frac{\mu_i}{\mu_i-\lambda_i}\times e^{\sum_{j=c}^{i-1}(\mu_j-\lambda_j)(x_{j+1}-x_j)}\times\left(e^{(\mu_i-\lambda_i)(x_{i+1}-x_i)}-1\right)\right)-\frac{\rho-1}{\rho}\times e^{\sum_{i=c}^{k}(\mu_i-\lambda_i)(x_{i+1}-x_i)}\right)^{-1} \qquad (3)
\end{aligned}
$$

where $k=|\mathbb{X}|$. Let us define n as the number of sampled species, m as the total number of species in the study group, $\mathbb{K}$ as the set of missing species per clade and $|\mathbb{K}|$ the number of clades with missing species. Additionally, we define c_i as the time of most recent common ancestor of the i^{th} clade. Then, the joint probability density of the sampled reconstructed tree and the empirically informed missing speciation times is

$$
\begin{aligned}
f(\Psi,\mathbb{K}|N(t_1=0)=2,S(2,t_1=0,T)) &= f(\Psi|N(t_1=0)=2,S(2,t_1=0,T)) \\
&\times\frac{(m-1)!}{(n-1)!}\prod_{i=1}^{|\mathbb{K}|}\frac{1}{k_i!}\left(1-F(t|N(c_i)=1,c_i\leq t\leq T)\right)^{k_i} \qquad (4)
\end{aligned}
$$

Prior models on diversification rates. Our model assumes that speciation and extinction rates are piecewise constant but can be different for different time intervals (Figure 4). Thus, we divide time into equal-length intervals (e.g., Δt=1). Following Magee *et al*.[21], we specify prior distributions on the log-transformed speciation rates ($\ln(\lambda_i)$) and extinction rates ($\ln(\mu_i)$)

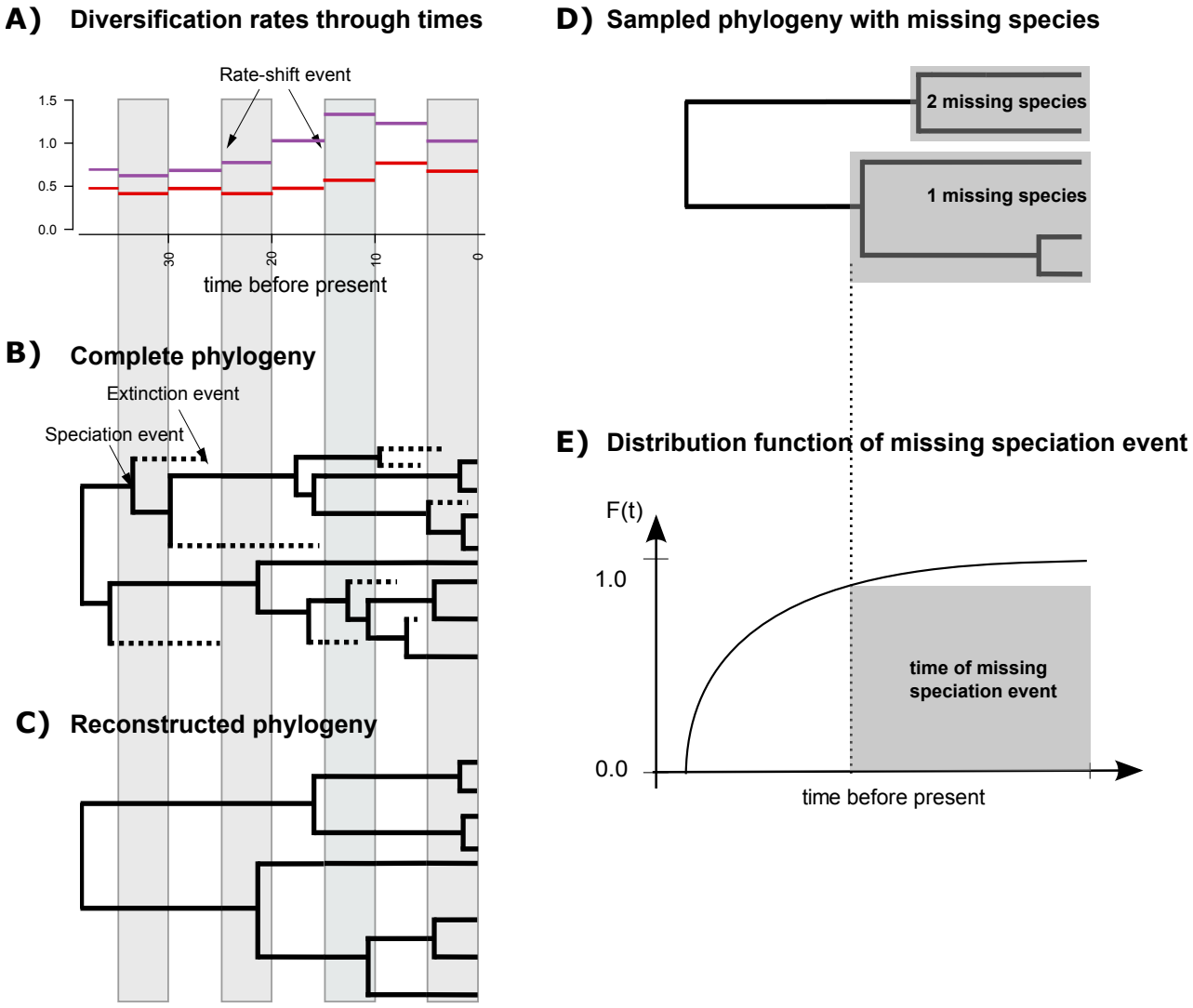

Figure 4. Cartoon of the birth-death process with rate-shift events and empirical taxon sampling. A) Depiction of the speciation (purple lines) and extinction (red lines) rates through time. Here we assume that speciation and extinction rates are episodically constant, that is, diversification rates shift instantly and only at the beginning of an episode. Each episode lasts 5 time units in this example. B) A realization (complete phylogeny) of the birth-death process. Lineages that have no extant or sampled descendant are shown as dashed lines and surviving lineages are shown as solid lines. C) Reconstructed phylogeny corresponding exactly to the one shown in B with the extinct lineages pruned away. Thus, plot C depicts the "observed" phylogeny from which the speciation times are retrieved. D) Sampled phylogeny with gray boxes depicting named clades with known number of missing species. The phylogenyis the same as in C with fewer taxa. E) Distribution function of the time of the missing speciation event. The missing speciation event could have occurred any time between the crown age of the named clade and the present time (gray box). The distribution function is integrated over and hence the uncertainty of the missing speciation event accounted for.

because the rates are only defined for positive numbers and our prior distributions are defined for all real numbers. We apply and compare three different prior models: (i) an uncorrelated log-normal (UCLN) prior distribution, (ii) a Gaussian Markov random field (GMRF) prior[21], and (iii) a Horseshoe Markov random field (HMRF) prior[21]. The first prior distribution specifies temporally uncorrelated speciation and extinction rates, whereas the second and third prior distributions are autocorrelated prior models. The assumption of autocorrelated rates might make more sense biologically (an interval of high speciation rates is likely to be followed by another interval with high speciation) but also improves our ability to estimate parameters[21]. Nevertheless, our inclusion of both uncorrelated and autocorrelated prior distributions allows for testing whether an uncorrelated or autocorrelated model is preferred.

The prior distribution on the speciation rates λ_i and extinction rates μ_i are set in exactly the same form in our models with their respective hyperprior parameters. Thus, for the sake of simplicity, we omit the prior distribution on the extinction rates here in the text. Our first prior distribution, the uncorrelated log-normal (UCLN) distributed prior, specifies the same prior probability for each speciation rate λ_i,

$$\ln(\lambda_i) \sim \text{Normal}(m, \sigma) \quad . \tag{5}$$

Thus, each speciation rate is independent and identically distributed.

Our second, prior distribution, the Gaussian Markov random field (GMRF) prior, models rates in an autocorrelated form analogous to a discretized Brownian motion. That is, we assume that diversification rates $\lambda(t)$ and $\mu(t)$ are autocorrelated and the rates in the next time interval will be centered at the rates in the current time interval,

$$\ln(\lambda_i) \sim \text{Normal}(\ln(\lambda_{i-1}), \sigma_\lambda) \quad . \tag{6}$$

The standard deviation σ regulates the amount of change between each time interval.

Our third prior distribution, Horseshoe Markov random field (HSMRF) prior, is very similar to the GMRF but additionally allows for the variance to change between time intervals,

$$\gamma_i \sim \text{halfCauchy}(0, 1) \tag{7}$$

$$\ln(\lambda_i) \sim \text{Normal}(\ln(\lambda_{i-1}), \sigma\gamma_i) \quad . \tag{8}$$

The HSMRF prior model is more adaptive than the GMRF; it allows for more extreme jumps between intervals while favoring/smoothing more constant rate trajectories if there is no evidence for rate changes.

These three prior models of diversification rates provide the null models of our analyses as it does not assume any dependence to an environmental variable. We use this model first to estimate diversification rates through time before testing for a correlation of the speciation or extinction rate to an environmental variable (e.g., atmospheric CO_2 or paleo-temperature). Magee *et al.*[21] found that 100 epochs perform well for autocorrelated models. Since we do not know how many bins (i.e., epochs) should be used for the episodic birth-death process, we test various numbers of equal-sized epochs (4, 10, 20, 50, 100 and 200). We show both the mean posterior diversification rates (Fig. S5) as well as select the best fitting model based on the number of epochs (Fig. S7).

Correlation between speciation and extinction rate to CO_2. Previously, Condamine *et al.*[22] introduced an environmentally-dependent diversification model. In their model, diversification rates are correlated with an environmental variable[22–27]. For example, the speciation rate can be modeled as $\lambda(t) = \lambda_0 e^{\beta \times CO_2(t)}$ (see Box 1 in Condamine *et al.*[22]), which is equivalent to $\ln(\lambda(t)) = \ln(\lambda_0) + \beta \times CO_2(t)$. Since we are using the episodic birth-death process which has piecewise-constant diversification rates, we modify the original continuous-time environmentally-dependent diversification model to $\ln(\lambda_i) = \ln(\lambda_0) + \beta \times CO_{2,i}$, which is equivalent to and more conveniently written as $\ln(\lambda_i) = \ln(\lambda_{i-1}) + \beta \times \Delta CO_{2,i}$ where $\Delta CO_{2,i} = CO_{2,i} - CO_{2,i-1}$. Note that we only use the so-called exponential dependency and not the linear dependency[24] because the linear dependency can result in negative rates which are mathematically and biologically impossible[63].

We applied this original environmentally-dependent diversification model and three new environmentally-dependent diversification models. The original environmentally-dependent diversification model of Condamine *et al.*[22] does not accommodate diversification-rate variation that is independent of the environmental variable. Instead, our three new environmentally-dependent diversification models build on our diversification-rate prior models which allow for rate-variation through time (see above). Thus, our environmentally-dependent diversification models will collapse to the episodic birth-death model if rates of diversification and atmospheric CO_2 are uncorrelated and hence inherently allows for diversification rate variation. The linkage of environmental variable and diversification rates without allowing for independent diversification rate variation might provide spurious results, as has been noticed for trait evolution[64] and state-dependent diversification rates[65]. We explore this potential of misattribution of diversification rate variation to the environmental variable in our model selection procedure and simulation study (see below).

As before, we omit the description of the extinction rates in the text for the sake of notational simplicity. Both speciation and extinction rates are model exactly in the same way with their corresponding set of hyperparameters (e.g., see the Tables S4-S7). Our first environmentally-dependent diversification model has a *fixed* linkage between the diversification rate variation and variation in the environmental variable;

$$\begin{aligned} \lambda_0 &\sim \text{Uniform}(0, 100) && (9)\\ \ln(\lambda_i) &= \ln(\lambda_{i-1}) + \beta_\lambda \times \Delta\text{CO}_2 \quad . && (10) \end{aligned}$$

This model does not have a counterpart in the above diversification rate priors, but is included as a comparison to the work Condamine *et al*[22].

Our second environmentally-dependent diversification model adds uncorrelated lognormal variation on top of the variation in the environmental variable;

$$\begin{aligned} \lambda_0 &\sim \text{Uniform}(0, 100) && (11)\\ \ln(\hat{\lambda}_i) &= \ln(\hat{\lambda}_{i-1}) + \beta_\lambda \times \Delta\text{CO}_2 && (12)\\ \varepsilon_i &\sim \text{Normal}(0, \sigma) && (13)\\ \ln(\lambda_i) &= \ln(\hat{\lambda}_i) + \varepsilon_i \quad . && (14) \end{aligned}$$

Thus, this model collapses to the above UCLN model if there is no correlation between the environmental variable and diversification rates ($\beta = 0$). Importantly, the difference in the variation of the diversification rates and environmental variable is independent in each epoch, contributed by the variable ε_i. The environmental-dependent part of the diversification rates $\hat{\lambda}_i$ is equivalent to the *fixed* environmentally-dependent diversification model.

Our third environmentally-dependent diversification model adds correlated lognormal variation on top of the *fixed* environmentally-dependent diversification mode;

$$\begin{aligned} \lambda_0 &\sim \text{Uniform}(0, 100) && (15)\\ \ln(\lambda_i) &\sim \text{Normal}(\ln(\lambda_{i-1}) + \beta_\lambda \times \Delta\text{CO}_2), \sigma) \quad . && (16) \end{aligned}$$

This model an extension of the above GMRF model and collapses to it if there is no correlation between the environmental variable and diversification rates ($\beta = 0$). As the GMRF model is a discretized Brownian motion model, this environmentally-dependent extension can be considered as a Brownian motion with trend model, where the trend is predicted by the environmental variable. Instead of writing this model with a separate environmentally-dependent part $\hat{\lambda}_i$ and autocorrelated part ε_i, we directly use the combined environmentally-dependent and independent rate variation as the mean for the next time interval. Nevertheless, we want to emphasize this equivalence to bridge the connection to the UCLN model above.

Finally, our fourth environmentally-dependent diversification model extends the above HSRMF to allow for diversification rates predicted by the environmental variable;

$$\begin{aligned} \lambda_0 &\sim \text{Uniform}(0, 100) && (17)\\ \gamma_i &\sim \text{halfCauchy}(0, 1) && (18)\\ \ln(\lambda_i) &\sim \text{Normal}(\ln(\lambda_{i-1}) + \beta_\lambda \times \Delta\text{CO}_2), \sigma\gamma_i) \quad . && (19) \end{aligned}$$

This model follows the same extension as the environmentally-dependent GMRF model with local adaptability of the rate variation through the parameter γ_i, as before for the HSMRF.

In all our four models, we denote the correlation coefficient by β. If $\beta > 0$ then there is a positive correlation between the speciation rate and CO_2, that is, if the CO_2 increases then the speciation rate will also increases. By contrast, if $\beta < 0$ then there is a negative correlation between the speciation rate and CO_2, that is, if the CO_2 concentration increases then the speciation rate will decrease. Finally, if $\beta = 0$ then there is no correlation and our environmentally-dependent diversification model collapses to corresponding episodic birth-death model.

All four models have the same parameter for the initial speciation rate λ_0 with a uniform prior distribution between 0 and 100. The models are constructed in increasing complexity and all three new models can collapse either to the *fixed* environmentally-dependent diversification model or to their environmentally-independent episodic birth-death process.

Environmental Data. In our analyses we tested for correlation between two environmental factors: CO_2 and temperature. The concentration of atmospheric CO_2 throughout the Cenozoic were compiled by Beerling & Royer[33] using terrestrial and marine proxies. An updated dataset was provided by Dr. Dana Royer. Paleo-temperature fluctuations come from Zachos et al. (2001)[66]. Raw data were extracted from ftp://ftp.ncdc.noaa.gov/pub/data/paleo/.

Analogous to our tests about the number of epochs for the diversification rate analyses, we computed the arithmetic mean for the environmental variable for 1-, 2- and 5-million year intervals. We both estimated the correlation between the environmental variable and diversification rates for each interval size and performed model selection using Bayes factors.

Model Selection. We performed three sets of empirical diversification rate analyses for each dataset. We estimated the diversification rates over time using three different models, we estimated the environmentally-dependent diversification rates using four different models, and we applied two different taxon sampling schemes. For the first two sets of analyses we performed standard model selection in a Bayesian framework using Bayes factors[67]. Thus, we computed the marginal likelihood for each model using stepping-stone sampling[68] as implemented in `RevBayes`. We run 128 stepping stones with each stone comprising of its own MCMC run with 2,000 iteration and on average 1,374 moves per iteration (i.e., the runs being equivalent to standard single-move-per-iteration software with 2,748,000 iterations).

We tested the support for the environmental correlation using Bayes factors computed from the posterior odds. Our prior probability for the correlation coefficient β was symmetric and centered at zero, that is, we specified exactly a probability of 0.5 that $\beta < 0$ and $\beta > 0$. Thus, the prior probability ratio of $\frac{P(\beta<0)}{P(\beta>0)} = 1.0$ Then, to compute the Bayes factor for in support of a negative correlation is simply the number of MCMC samples with $\beta < 0$ divided over the total number of MCMC samples.

We did not compute marginal likelihoods for the two different sampling schemes; the uniform taxon sampling and the empirical taxon sampling. Empirical taxon sampling uses additional data, the age ranges of the missing speciation events, and two analyses with different data cannot be compared using traditional model selection. Instead, we performed a simulation study to show the robustness of our parameter estimates under empirical taxon sampling and the resulting bias if wrongly uniform taxon sampling was assumed.

Simulation Study. We performed two sets of simulations; focusing (a) on the environmentally-correlated diversification model, and (b) the incomplete taxon same scheme. First, we simulated phylogenies under the UCLN and GMRF environmentally-correlated diversification model using the R package `TESS`[69,70]. We set the diversification rate variation to $\sigma = \{0, 0.02, 0.04\}$ and correlation coefficient to $\beta = \{0, -0.005, -0.01\}$. Thus, our simulations included the constant-rate birth-death process (when $\sigma = 0$ and $\beta = 0$), time-varying but environmentally independent diversification rates (when $\sigma > 0$ and $\beta = 0$), the *fixed* environmentally-dependent diversification model (when $\sigma = 0$ and $\beta \neq 0$), and the time-varying and environmentally-dependent diversification model (when $\sigma > 0$ and $\beta \neq 0$). For each setting, we simulated ten diversification rate trajectories (Figure S14 and S15) and trees (Figure S12 and S13). We analyzed each simulated tree under the same four environmentally-dependent diversification model as in our empirical analysis (see above).

Second, we simulated phylogenetic trees under empirical taxon sampling to validate the correctness of our model derivation. Unfortunately, simulation of empirical taxon sampling is not straight forward. We circumvented the problem by randomly adding the missing species to the daisy phylogenetic tree, then drawing new divergence times under (a) a constant-rate birth-death process, and (b) a time-varying episodic birth-death process with rates taken from the empirical estimates. Then, we pruned the additional species to mimic empirical taxon sampling. The simulations under the constant-rate birth-death process provide information about falsely inferring diversification rate variation (false positives) and the simulations under the time-varying episodic birth-death process provide information about the power to correctly inferring diversification rate variation (power analysis). We simulated 100 trees under each setting and analyzed each tree using the GMRF prior model with both empirical and uniform taxon sampling. The MCMC inference settings were identical to the empirical analyses.

Software Implementation and Availability. Both models, the episodic birth-death process and the environmentally-dependent diversification model, are implemented in the Bayesian phylogenetics software `RevBayes`[17]. Moreover, the implementation is not restricted to the models we introduce here because `RevBayes` is built on the principle of probabilistic graphical models[71]. The graphical model approach provides full flexibility to extend or modify the current analyses to other models and assumption, for example, testing for correlation to multiple environmental variables. `RevBayes` is open-source and freely available from https://github.com/revbayes/revbayes. The analysis from this paper are described in detail in several tutorials available at http://revbayes.github.io/tutorials.html.

References

1. Vitousek, P. M. Grassland ecology: Complexity of nutrient constraints. *Nat. Plants* **1**, 1–2 (2015).
2. Sala, O. E. *et al.* Global Biodiversity Scenarios for the Year 2100. *Sci.* **287**, 1770–1774 (2000).
3. Janis, C. M., Damuth, J. & Theodor, J. M. Miocene ungulates and terrestrial primary productivity: where have all the browsers gone? *Proc. Natl. Acad. Sci. United States Am.* **97**, 7899–7904 (2000).
4. Strömberg, C. A. Evolution of Grasses and Grassland Ecosystems. *Annu. Rev. Earth Planet. Sci.* **39**, 517–544 (2011).
5. Hill, R., Truswell, E., McLoughlin, S. & Dettmann, M. E. The evolution of the Australian flora: fossil evidence. In *Flora of Australia: Volume 1, Introduction*, vol. i, 251–320 (1999).
6. Palazzesi, L. & Barreda, V. Fossil pollen records reveal a late rise of open-habitat ecosystems in Patagonia. *Nat. Commun.* **3** (2012).

7. Pennington, R. T., Cronk, Q. C. B. & Richardson, J. A. Introduction and synthesis: plant phylogeny and the origin of major biomes. *Philos. Transactions Royal Soc. B: Biol. Sci.* **359**, 1455–1464 (2004).

8. Davis, C. C., Webb, C. O., Wurdack, K. J., Jaramillo, C. A. & Donoghue, M. J. Explosive radiation of Malpighiales supports a mid-cretaceous origin of modern tropical rain forests. *The Am. naturalist* **165**, E36–E65 (2005).

9. Couvreur, T., Forest, F. & Baker, W. J. Origin and global diversification patterns of tropical rain forests: inferences from a complete genus-level phylogeny of palms. *BMC Biol.* **9**, 44 (2011).

10. Richardson, J. E., Pennington, R. T., Pennington, T. D. & Hollingsworth, P. M. Rapid diversification of a species-rich genus of neotropical rain forest trees. *Sci.* **293** (2001).

11. Spriggs, E. L., Christin, P.-A. & Edwards, E. J. C4 photosynthesis promoted species diversification during the miocene grassland expansion. *PLoS ONE* **9**, e105923 (2014).

12. Nee, S., May, R. M., Harvey, P. H., Trans, P. & Lond, R. S. The reconstructed evolutionary process. *Philos. transactions Royal Soc. Lond. Ser. B, Biol. sciences* **344**, 305–11 (1994).

13. Yang, Z. & Rannala, B. Bayesian phylogenetic inference using DNA sequences: a Markov Chain Monte Carlo Method. *Mol. biology evolution* **14**, 717–24 (1997).

14. Cusimano, N. & Renner, S. S. Slowdowns in diversification rates from real phylogenies may not be real. *Syst. Biol.* **59**, 458–464 (2010).

15. Höhna, S., Stadler, T., Ronquist, F. & Britton, T. Inferring speciation and extinction rates under different sampling schemes. *Mol. Biol. Evol.* **28**, 2577–2589 (2011).

16. Höhna, S. Likelihood inference of non-constant diversification rates with incomplete taxon sampling. *PLoS ONE* **9**, e84184 (2014).

17. Höhna, S. *et al.* RevBayes: Bayesian phylogenetic inference using graphical models and an interactive model-specification language. *Syst. Biol.* **65**, 726–736 (2016).

18. Stadler, T. Mammalian phylogeny reveals recent diversification rate shifts. *Proc. Natl. Acad. Sci. USA* **108**, 6187–6192 (2011).

19. Höhna, S. The time-dependent reconstructed evolutionary process with a key-role for mass-extinction events. *J. Theor. Biol.* **380**, 321–331 (2015).

20. May, M. R., Höhna, S. & Moore, B. R. A Bayesian approach for detecting the impact of mass-extinction events on molecular phylogenies when rates of lineage diversification may vary. *Methods Ecol. Evol.* **7**, 947–959 (2016).

21. Magee, A. F., Höhna, S., Vasylyeva, T. I., Leaché, A. D. & Minin, V. N. Locally adaptive Bayesian birth-death model successfully detects slow and rapid rate shifts. *PLoS Comput. Biol.* **16**, e1007999 (2020).

22. Condamine, F. L., Rolland, J. & Morlon, H. Macroevolutionary perspectives to environmental change. *Ecol. Lett.* **16**, 72–85 (2013).

23. Morlon, H. *et al.* RPANDA: an R package for macroevolutionary analyses on phylogenetic trees. *Methods Ecol. Evol.* **7**, 589–597 (2016).

24. Lewitus, E. & Morlon, H. Detecting environment-dependent diversification from phylogenies: a simulation study and some empirical illustrations. *Syst. biology* **67**, 576–593 (2018).

25. Condamine, F. L., Rolland, J., Höhna, S., Sperling, F. A. H. & Sanmartín, I. Testing the role of the Red Queen and Court Jester as drivers of the macroevolution of Apollo butterflies. *Syst. Biol.* **67**, 940–964 (2018).

26. Condamine, F. L., Rolland, J. & Morlon, H. Assessing the causes of diversification slowdowns: temperature-dependent and diversity-dependent models receive equivalent support. *Ecol. letters* **22**, 1900–1912 (2019).

27. Meseguer, A. S., Antoine, P.-O., Fouquet, A., Delsuc, F. & Condamine, F. L. The role of the Neotropics as a source of world tetrapod biodiversity. *Glob. Ecol. Biogeogr.* **29**, 1565–1578 (2020).

28. Pybus, O. G. & Harvey, P. H. Testing macro-evolutionary models using incomplete molecular phylogenies. *Proc R Soc B Biol Sci* **267** (2000).

29. Janzen, T., Höhna, S. & Etienne, R. S. Approximate bayesian computation of diversification rates from molecular phylogenies: introducing a new efficient summary statistic, the nltt. *Methods Ecol. Evol.* **6**, 566–575 (2015).

30. Christin, P.-A. *et al.* Molecular Dating, Evolutionary Rates, and the Age of the Grasses. *Syst. Biol.* **63**, 153–165 (2014).

31. Azevedo, J. A. *et al.* On the young savannas in the land of ancient forests. *Neotropical diversification: Patterns processes* 271–298 (2020).

32. Barreda, V., Palazzesi, L., Tellería, M. C., Katinas, L. & Crisci, J. V. Fossil pollen indicates an explosive radiation of basal Asteracean lineages and allied families during Oligocene and Miocene times in the Southern Hemisphere. *Rev. Palaeobot. Palynol.* **160**, 102–110 (2010).

33. Beerling, D. J. & Royer, D. L. Convergent Cenozoic CO2 history. *Nat. Geosci.* **4**, 418–420 (2011).

34. Rae, J. W. *et al.* Atmospheric co2 over the past 66 million years from marine archives. *Annu. Rev. Earth Planet. Sci.* **49** (2021).

35. Westerhold, T. *et al.* An astronomically dated record of earth's climate and its predictability over the last 66 million years. *Sci.* **369**, 1383–1387 (2020).

36. Robinson, J. M. Speculations on carbon dioxide starvation, Late Tertiary evolution of stomatal regulation and floristic modernization (1994).

37. Crucifix, M., Betts, R. A. & Hewitt, C. D. Pre-industrial-potential and Last Glacial Maximum global vegetation simulated with a coupled climate-biosphere model: Diagnosis of bioclimatic relationships. *Glob. Planet. Chang.* **45**, 295–312 (2005).

38. Cowling, S. A. & Sykes, M. T. Physiological Significance of Low Atmospheric CO2 for Plant–Climate Interactions. *Quat. Res.* **52**, 237–242 (1999).

39. Harrison, S. P. & Prentice, C. I. Climate and CO2 controls on global vegetation distribution at the last glacial maximum: Analysis based on palaeovegetation data, biome modelling and palaeoclimate simulations. *Glob. Chang. Biol.* **9**, 983–1004 (2003).

40. Webb, S. D. A history of savanna vertebrates in the new world. part i: North america. *Annu. Rev. Ecol. Syst.* **8**, 355–380 (1977).

41. Axelrod, D. I. Rise of the grassland biome, central north america. *The Bot. Rev.* **51**, 163–201 (1985).

42. Edwards, E. J. & Smith, S. A. Phylogenetic analyses reveal the shady history of c4 grasses. *Proc. Natl. Acad. Sci.* **107**, 2532–2537 (2010).

43. Blair, J., Nippert, J. & Briggs, J. Grassland ecology. In *Ecology and the Environment*, 389–423 (Springer, 2014).

44. Belsky, A. J. Effects of grazing, competition, disturbance and fire on species composition and diversity in grassland communities. *J. Veg. Sci.* **3**, 187–200 (1992).

45. Jardine, P. E., Janis, C. M., Sahney, S. & Benton, M. J. Grit not grass: concordant patterns of early origin of hypsodonty in great plains ungulates and glires. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **365**, 1–10 (2012).

46. Scheiter, S. *et al.* Fire and fire-adapted vegetation promoted c4 expansion in the late miocene. *New Phytol.* **195**, 653–666 (2012).

47. Kandziora, M., Kadereit, J. W. & Gehrke, B. Dual colonization of the palaearctic from different regions in the afrotropics by senecio. *J. Biogeogr.* **44**, 147–157 (2017).

48. Forrestel, E. J. & Edwards, E. J. The grasslands future biogeography of c3 and c4. *Grasslands Clim. Chang.* 234 (2019).

49. Stadler, T. & Bokma, F. Estimating speciation and extinction rates for phylogenies of higher taxa. *Syst. Biol.* **62**, 220–230 (2013).

50. Olson, D. M. *et al.* Terrestrial ecoregions of the world: a new map of life on Earth. *Biosci.* **51** (2001).

51. Chamberlain, S. rgbif: Interface to the Global 'Biodiversity' Information Facility 'API' (2016).

52. Prasad, V. *et al.* Late Cretaceous origin of the rice tribe provides evidence for early diversification in Poaceae. *Nat. Commun.* **2**, 480 (2011).

53. Barreda, V. D. *et al.* Early evolution of the angiosperm clade Asteraceae in the Cretaceous of Antarctica. *Proc. Natl. Acad. Sci.* **112**, 10989–10994 (2015).

54. Crepet, W. L. & Feldman, G. D. The earliest remains of grasses in the fossil record (1991).

55. Barreda, V. D. *et al.* An extinct Eocene taxon of the daisy family (Asteraceae): Evolutionary, ecological and biogeographical implications. *Annals Bot.* **109**, 127–134 (2012).

56. Jacobs, B. F., Kingston, J. D. & Jacobs, L. L. The origin of grass-dominated ecosystems. *Ann Mo. Bot Gard* **86** (1999).

57. Kellogg, E. A. Fossil Record and Dates of Diversification. In *Flowering Plants. Monocots*, 103–107 (Springer International Publishing, Cham, 2015).

58. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. biology evolution* **30**, 772–80 (2013).

59. Huang, C.-H. *et al.* Multiple Polyploidization Events across Asteraceae with Two Nested Events in the Early History Revealed by Nuclear Phylogenomics. *Mol. biology evolution* **33**, 2820–2835 (2016).

60. Xu, X., Dimitrov, D., Rahbek, C. & Wang, Z. NCBIminer: Sequences harvest from Genbank. *Ecography* **38**, 426–430 (2015).

61. Revell, L. J. phytools: An R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217–223 (2012).

62. Yu, G., Smith, D. K., Zhu, H., Guan, Y. & Lam, T. T.-Y. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* **8**, 28–36 (2017).

63. Gamisch, A. Technical comment on condamine et al.(2019): a cautionary note for users of linear diversification dependencies. *Ecol. letters* **23**, 1169 (2020).

64. May, M. R. & Moore, B. R. A bayesian approach for inferring the impact of a discrete character on rates of continuous-character evolution in the presence of background-rate variation. *Syst. Biol.* **69**, 530–544 (2020).

65. Rabosky, D. L. & Goldberg, E. E. Model inadequacy and mistaken inferences of trait-dependent speciation. *Syst. Biol.* **64**, 340–355 (2015).

66. Zachos, J., Pagani, M., Sloan, L., Thomas, E. & Billups, K. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Sci.* **292**, 686–693 (2001).

67. Kass, R. & Raftery, A. Bayes factors. *J. Acoust. Soc. Am.* **90**, 773–795 (1995).

68. Xie, W., Lewis, P., Fan, Y., Kuo, L. & Chen, M. Improving marginal likelihood estimation for bayesian phylogenetic model selection. *Syst. Biol.* **60**, 150–160 (2011).

69. Höhna, S. Fast simulation of reconstructed phylogenies under global time-dependent birth-death processes. *Bioinforma.* **29**, 1367–1374 (2013).

70. Höhna, S., May, M. R. & Moore, B. R. TESS: an R package for efficiently simulating phylogenetic trees and performing Bayesian inference of lineage diversification rates. *Bioinforma.* **32**, 789–791 (2016).

71. Höhna, S. *et al.* Probabilistic Graphical Model Representation in Phylogenetics. *Syst. Biol.* **63**, 753–771 (2014).

Acknowledgements

We thank Will Freyman and Dana Royer for their comments on earlier versions of this manuscript. Luis Palazzesi received a Marie Curie International Incoming Fellowships from the European Union for project GRASSLANDS (Proposal 329652 & 912652 — FP7-PEOPLE-2012-IIF). Additional financial support was provided by ANPCyT and CONICET from the Argentinian Government. Sebastian Höhna was supported by the Deutsche Forschungsgemeinschaft (DFG) Emmy Noether-Program HO 6201/1-1.