**Title**

Marine origination rates linked to interactions of past temperature trends with short-term climate change

**Abstract**

Biodiversity dynamics are shaped by a complex interplay between current conditions and historic legacy. While a simple relationship is often used to link evolution with temperature, short-term climate change likely interacts with previous temperature trends when influencing the pace of origination. Such palaeoclimate interactions have been demonstrated for extinction risk, but the effect on evolutionary dynamics is untested. Here we show that origination probability in marine fossil groups is largely affected by palaeoclimate interactions. Short-term cooling adding to a long-term cooling trend increases the origination probability by 24%. This large effect is consistent through time and all studied groups. Our results demonstrate that biodiversity is controlled by a complex array of ecological and evolutionary factors, with mutual interactions. The mechanisms of the detected effect might be manifold but are likely connected to eustatic sea level drop caused by cumulative global cooling. The complex and cascading nature of palaeoclimate interactions might explain the ambiguous conclusions of previous studies on the relationship between temperature and diversity. Accounting for these complex interactions using a dynamic modelling approach improves existing models and enables a better mechanistic discernment of processes that generate and maintain biodiversity through time.

**Introduction**

Biodiversity responses to modern climate change are dependent on past climate (Antão et al., 2020; Svenning et al., 2015). Palaeoclimate has driven both extinction and diversification dynamics in the Quaternary (Lister, 2004; Maldonado-Coelho, 2012; Postigo-Mijarra et al., 2010; Svenning et al., 2015). The effect of past climate on evolutionary dynamics in deep time, however, is unknown.

Palaeontological models generally test first- or second-order relationships between Phanerozoic marine diversity and climate (Cárdenas & Harries, 2010; Douglas, 2009; Krug et al., 2009; Mayhew et al., 2008). Current analysis, per contra, shows that fossil biodiversity interacts with temperature by complex cascading effects (Mathes et al., 2020), which could blur general relationships. While these interactions of past temperature trends with short-term climate change have shown to be major drivers of extinction events in Earth’s history (Mathes et al., 2020), their effect on evolutionary dynamics is untested.

These palaeoclimate interactions theoretically initiate new evolutionary dynamics and hence affect origination rates. If short-term climate change adds to a long-term temperature trend in the same direction, species are more likely to loose adaptations to the climatic situation due to niche conservatism (Hopkins et al., 2014; Stigall, 2014; Wiens & Graham, 2005). This maladjustment might result in bottleneck and subsequent founder effects, as well as ecological releases (Button, 2017; Gilman et al., 2010; Raup, 1979). These ecological effects have been shown to influence rates of evolution and speciation (Templeton, 2008; Wahl et al., 2002). We hence expect that palaeoclimate interactions not only affect extinction dynamics, but also origination processes. Explicitely, we hypothesize that origination processes are stronger influenced by temperature change if the change adds to a previous temperature trend in the same direction (e.g. warming following long-term warming) rather than if the focal change withdraws previous trends (e.g. warming following long-term cooling).

Here we analyse how global trajectories of palaeoclimate can affect origination probabilities of twelve marine fossil phyla in the last 485 million years (Annelida, Arthropoda, Brachiopoda, Bryozoa, Chordata, Cnidaria, Echinodermata, Foraminifera, Hemichordata, Mollusca, Nematoda, Porifera). We use palaeoclimate interactions as explanatory variables, which are defined as the interaction of short-term climate change with long-term temperature trends. We first apply regression models using traditional palaeo-temperature variables to sample-standardized fossil data. We then compare performances of these traditional models to our models including palaeoclimate interactions within a dynamic modelling framework. Considering the general complexity of evolutionary processes, we hypothesize that incorporating dynamic interactions of palaeo-temperature estimates can lead to an improved mechanistic understanding of evolutionary patterns. Second, we test whether cumulative palaeoclimate interactions increase the origination probability of fossil taxa, which might explain the perceived ambiguity of existing studies on the relationship of temperature and biodiversity in Earth’s history.

**Results**

*Model comparison*

Models including dynamic palaeoclimate interactions performed better than their traditional counterparts in 6 out of 6 cases (Fig. 1, Table S 2), based on Akaike’s information criterion (AIC) and Bayesian information criterion (BIC). Models covering short-term cooling improved on average by 688 AIC values (665 BIC values), and models covering short-term warming by 196 AIC values (172 BIC values). Both AIC and BIC values hence indicate that all models including palaeoclimate interactions are more parsimonious than their counterparts, despite their increased complexity.

*Origination probability*

Cooling-cooling palaeoclimate interactions resulted in an increased origination probability (Fig. 2, Table S 7). We compared the origination probability for marine groups per palaeoclimate interaction, based on the output of our regression models. The overall origination response was on average lower after cooling-warming, warming-cooling, and warming-warming palaeoclimate interactions compared to the average origination probability of 13.2% (p < 0.001, Wilcoxon rank sum test). On the contrary, cooling-cooling palaeoclimate interactions resulted in an increased mean origination response of 14.8%.

*Effect size*

The effect of cooling-cooling palaeoclimate interactions on origination probability is large (Fig. 3, Table S 8). The absolute difference between the mean origination probability of cooling-cooling compared to all other palaeoclimate interactions was 3.2 (Bayesian estimation, 95% CI [3.1, 3.2]). This difference represents a 23.7% increase compared to all other interactions (Bayesian estimation, 95% CI [23.2%, 24.2%]). We evaluated the effect size of cooling-cooling on origination probability using Cohen’s d statistic, a frequently applied measure used to indicate the standardised difference between two means. Cohen’s d for cooling-cooling was greater than 1, indicating a large effect size (Bayesian estimation, 95% CI [1.0, 1.1]). All results for effect size are consistent with bootstrapping resampling and raw output from generalised linear mixed effect models.

*Group differences*

We additionally tested whether cooling-cooling palaeoclimate interactions raised the origination probability of all studied marine groups and if the signal remains robust through tiime (Fig. 4, Table S 9). This was the case for all 10 major groups and throughout all 94 geologic stages. Within the groups, hemichordates, bryozoans, brachiopods, and arthropods showed a higher than average response to cooling-cooling palaeoclimate interactions. Through time, the youngest and oldest periods (stage 78 to 94, and 14 to 29) showed the highest likelihoods. We hence found a large effect size of cooling-cooling palaeoclimate on origination probability, which remained consistent throughout all studied groups and throughout the whole Phanerozoic.

**Discussion**

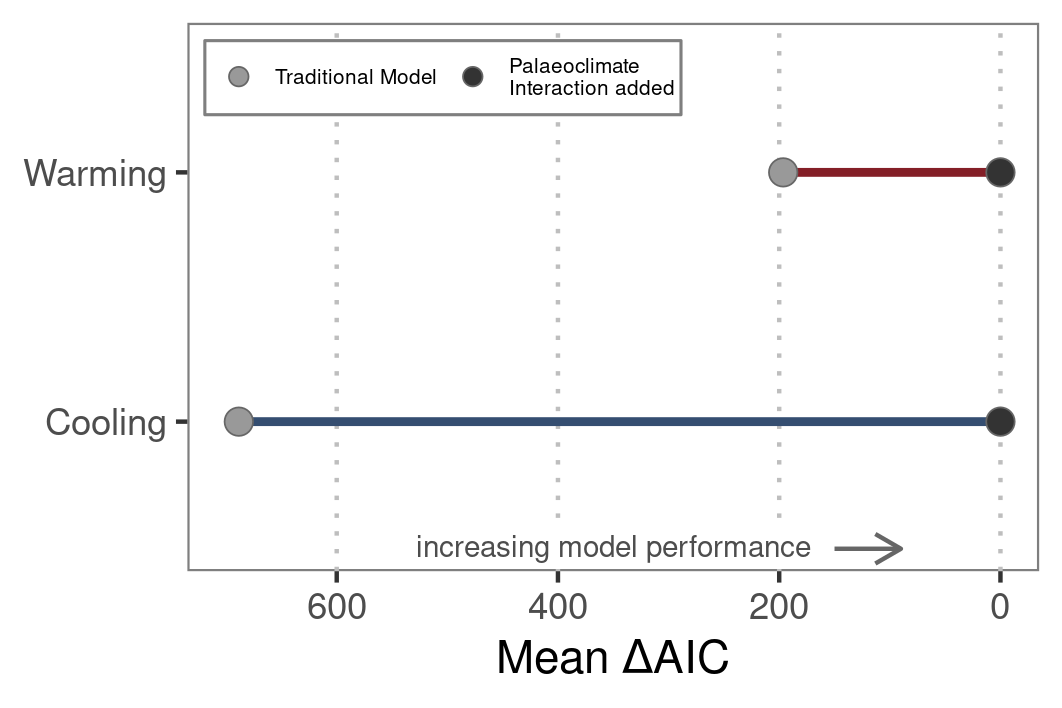
The interplay of short-term climate cooling adding to a long-term average global cooling resulted in a sharp increase of origination probability. We found this effect to be consistent through time and across all studied groups. Previous studies on the relationship between temperature and diversity came to ambiguous conclusions, ranging from a negative relationship (Mayhew et al., 2008), no relationship (John Alroy et al., 2000; Prothero, 1999), to a positive one (Mayhew et al., 2012). Our results show that complex interactions of climate and evolutionary dynamics exist, rendering the relationship between temperature and diversity non-linear. The large effect size and the cascading nature of palaeoclimate interactions might have obscured or even inverted the apparent relationship between temperature and origination, as it has been shown for other interactions in the fossil record (Ritterbush & Foote, 2017; Stigall, 2013). Explicitly accounting for dynamic interactions within a flexible modelling framework will hence provide a more robust foundation to assess the relationship between Phanerozoic marine diversity and climate. This has been recently demonstrated for modern terrestrial and marine assemblages, where biodiversity responses to recent climate change were conditional on the baseline climate (Antão et al., 2020).

Our results are remarkable as we expected to detect a weakened origination signal after cooling-cooling palaeoclimate interactions due to the “common cause” hypothesis (Shanan E. Peters, 2005). Sampling biases caused by low sea levels arise from the removal of originating taxa from the fossil record or a shift of their detection to a later origination datum, leading to an artificially reduced origination probability. These extrinsic biases of fossil data cannot be addressed by sampling standardisation (Shanan E. Peters, 2006). The apparently large effect size of cooling-cooling palaeoclimate interactions might hence even be an underestimation of the true magnitude. However, we only tested the effect of temperature on diversity and did not include other environmental parameters potentially affecting origination rates (e.g., sea-level, nutrient inputs, continental dispersion). The large effect size found in this study is thus related to temperature estimates only and could potentially be less relevant when compared to other environmental parameters. Nevertheless, current research shows that temperature remains a significant predictor of origination rates after accounting for additional environmental parameters (Mayhew et al., 2012). Further, temperature can act as a top-down effect, driving the change in other environmental parameters such as sea-level and shelf area weathering. Directly correlating temperature to origination rates, instead of fitting additional mediatory variables, might hence give a more precise estimate of the relationship between climate and evolutionary dynamics.

We emphasize that our results can only provide statistical conclusions, but no causal inferences. Causes for the detected effect might be manifold but are likely connected to eustatic sea level subsidence caused by global cooling. A drop in sea level due to glaciation results in reduced continental shelf area, which is the main habitat for the majority of the studied fossil groups. Increased habitat fragmentation and loss is correlated to the rate of ecological interactions (Schuler et al., 2017; Tilman, 1994; Valentine, 1968) and speciation rates (John Alroy, 2008; Mayr & O’Hara, 1986). Low sea levels lead to increased nutrient availability in the marine realm through exposure of continental shelf area to weathering (Broecker & Peng, 1982), a process known to drive genus origination rates (Cárdenas & Harries, 2010). The combination of short-term cooling on top of a long-term cooling trend exacerbates eustatic sea level drop, thus potentially reinforcing habitat fragmentation and loss as well as nutrient availability, leading to increased origination rates.

Cooling-cooling palaeoclimate interactions have been shown to be major driver of temperature related extinctions in the fossil record, with warming-warming interactions showing the same but less severe signal (Mathes et al., 2020). The synergetic combination of same-directional temperature changes is thought to move taxa out of their adaptation space, which is phylogenetically retained (Hopkins et al., 2014; Stigall, 2014; Wiens & Graham, 2005) When the environment changes so drastically that taxa are no longer adapted to it, they will either migrate, adapt through evolutionary change, or go extinct (Nogués-Bravo et al., 2018). During warming-warming palaeoclimate interactions, marine taxa could potentially escape adverse environment through range shifts. During cooling-cooling, however, geographic barriers resulting from sea level drop might impede migration, resulting in the observed increased extinction risk (Mathes et al., 2020) and origination probability (this study).

Our study shows that evolutionary processes interact with climate by complex cascading effects. Explicitly integrating these effects within a dynamic modelling framework leads to an improved discernment of origination patterns in the fossil record. Previous studies on the relationship of temperature and biodiversity in Earth’s history came to ambiguous conclusions, a non-conformance likely to be fixed in the future by accounting for palaeoclimate interactions. However, the mechanisms underlying the large association between evolution and the observed effect are still unclear. The interpretation offered here, a combination of niche conservatism and ecological effects arising from sea level change, could be at least partially tested by comparing responses of terrestrial and marine fossil groups to palaeoclimate interactions. If sea level subsidence plays a key role, terrestrial groups should show a significantly lower response to cooling-cooling than marine groups. Given the large effect of palaeoclimate interactions on both extinction and origination, as well as the consistency of this effect through time and studied groups, determining the underlying causes of palaeoclimate interactions will improve our mechanistic comprehension of evolutionary dynamics in Earths history (Dave Jablonski, 2008).

Figure 1 | Model comparison. Comparison of model performance for traditional models covering first- and second-order relationships of origination and temperature versus models allowing for palaeoclimate interactions. The latter are based on the same model structure, but explicitly allow for interactions of long-term temperature trends and short-term climate changes within a dynamic modeling framework.

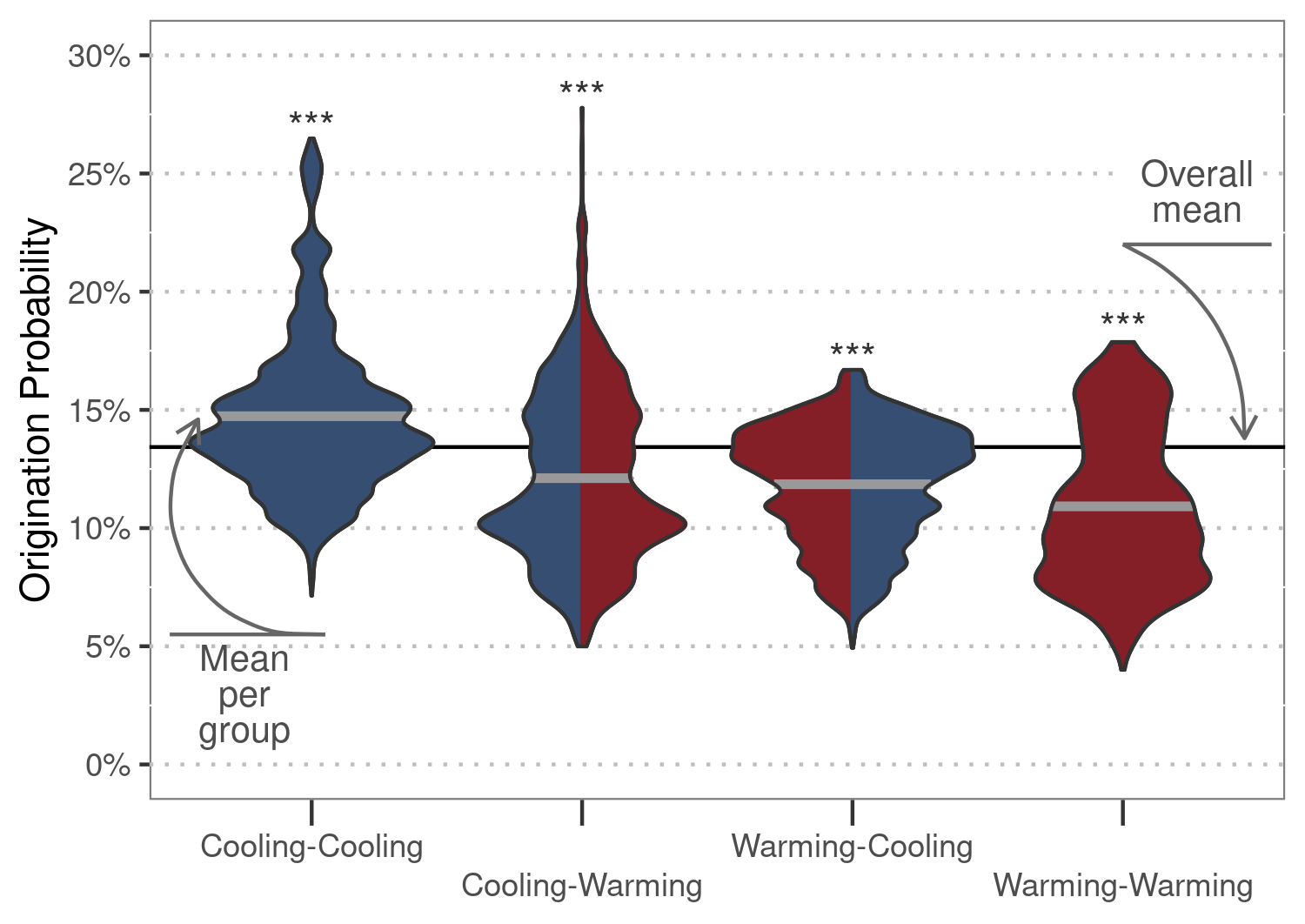
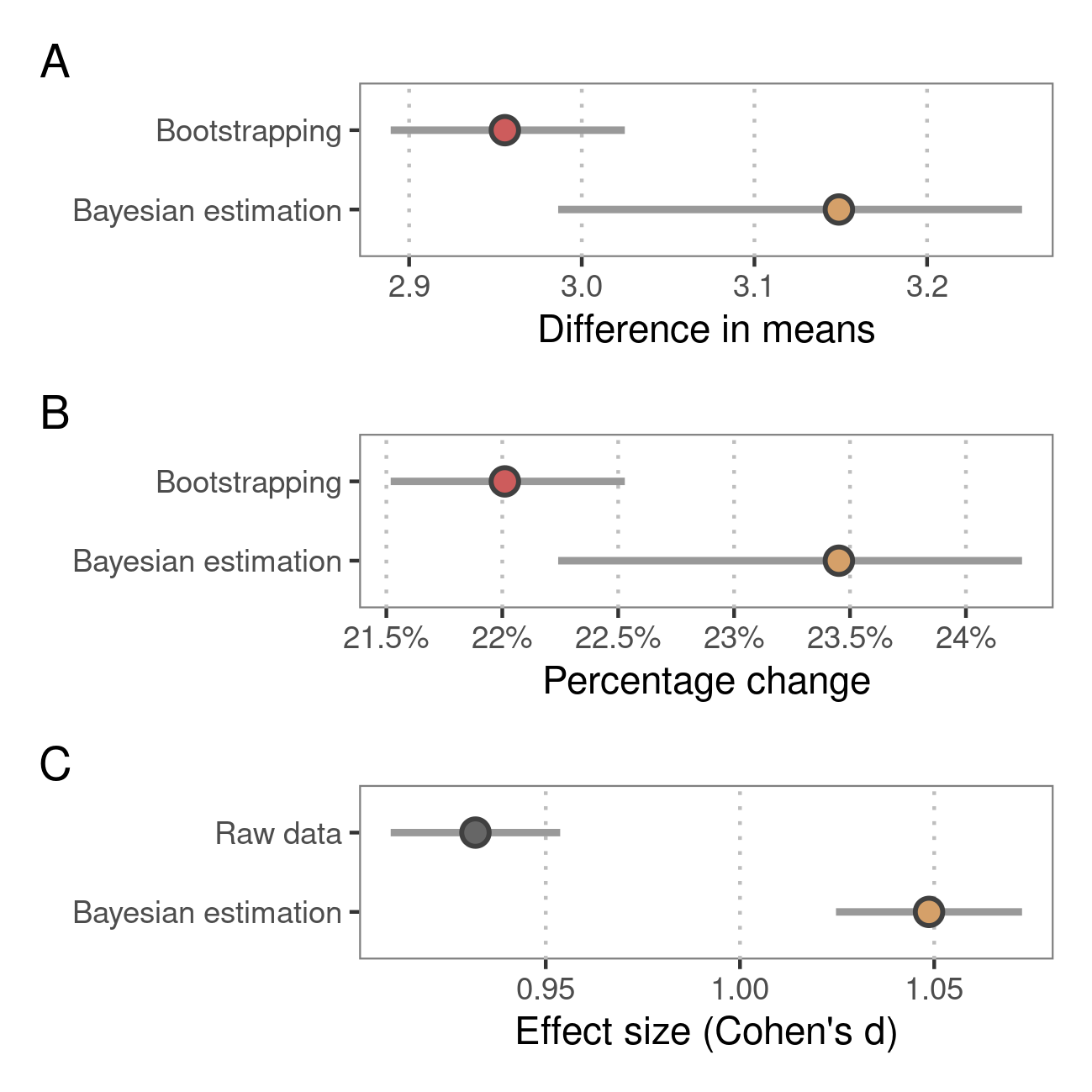
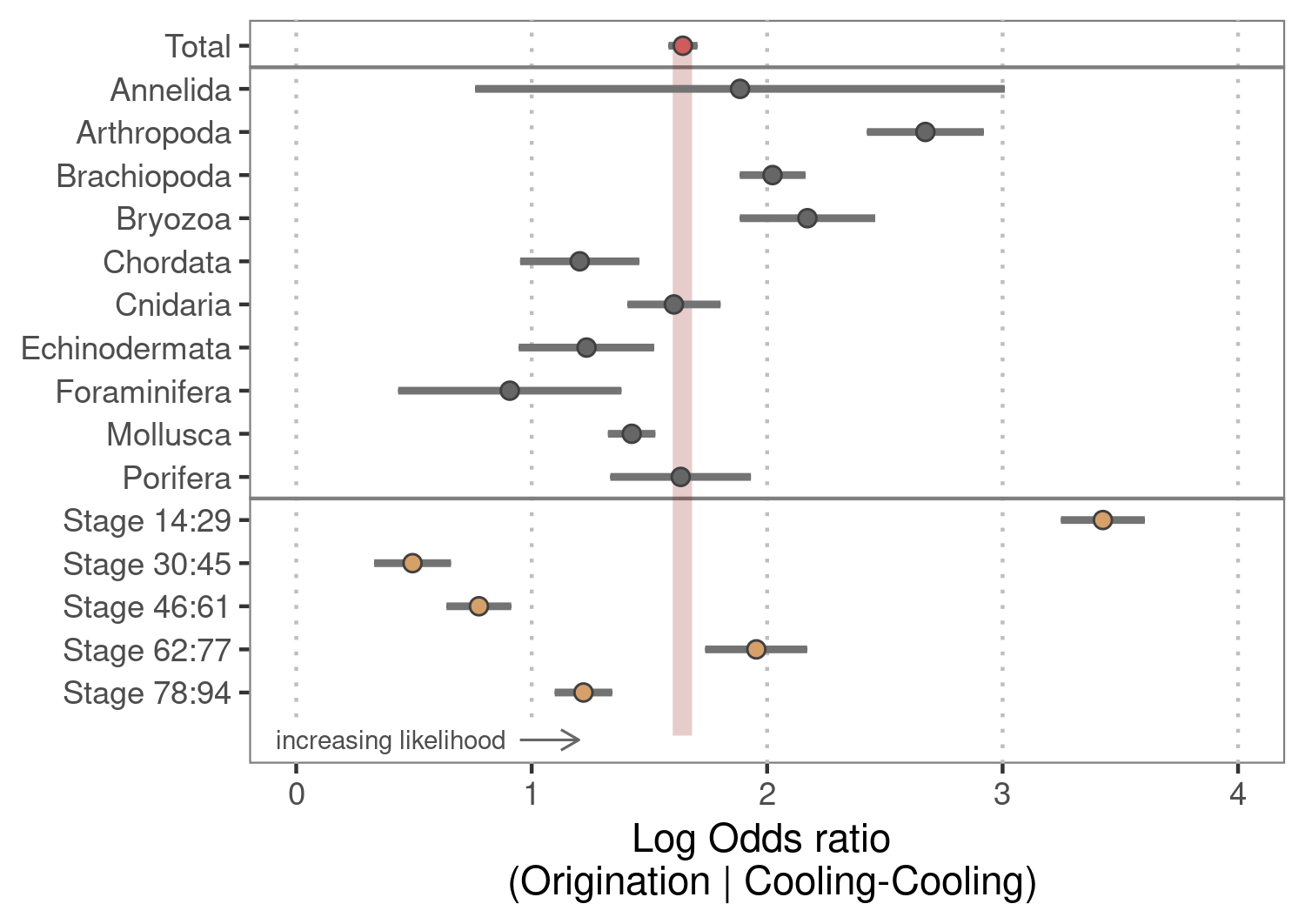


Figure 2 | Origination probability. Total response distributions to palaeoclimate interactions as predicted by generalised linear mixed effect models. Thick grey lines show the mean origination probability per palaeoclimate interaction, and the black line depicts the mean probability for all interactions combined. Asterisks indicate significance for differences of means compared to the overall mean, based on Wilcoxon rank sum tests.

Figure 3 | Effect size. A) Difference in means between origination probability after cooling-cooling compared to all other palaeoclimate interactions. B) Overall increase of origination probability after cooling-cooling palaeoclimate interactions compared to baseline probability. C) Cohen’s d effect size of the observed change in probability after cooling-cooling palaeoclimate interactions. Red points show estimates given by bootstrapping simulation, brown points estimates given by Bayesian regression, and grey points estimates based on raw output of generalised linear mixed effect models. Grey lines delineate 95% confidence intervals.

Figure 4 | Group differences. Origination probability for all major phyla and across time after cooling-cooling compared to all other palaeoclimate interactions. The red point shows the overall response, brown points response per time slice, and grey points per phyla. Grey lines show 95% confidence intervals. Stage 14 is the oldest stage included in analysis, and stage 94 the youngest.

**Methodology**

Fossil data

We accessed data from the Paleobiology Database (PaleoDB, paleobiodb.org) using the chronosphere package interface (À. T. Kocsis & Raja, 2020) on 03 December 2020, including all occurrences from the Ediacaran to the Holocene. All analyses were conducted at the genus level. This taxonomic level is a compromise between uncertainty in the species level taxonomy of fossils and data loss at coarser taxonomic resolutions (Valentine, 1974). We filtered the data to only comprise marine animal taxa and heterotrophic protists, i.e. the same taxonomic groups listed in Sepkoski’s (2002) compendium, following common processing recommendations (À. T. Kocsis et al., 2019). Previous studies on the relationship between temperature and biodiversity were mainly based on Sepkoski’s compendium, rendering this data optimal for comparison purposes of our novel model structures. We binned the data to one of 80 geological stages (Ogg et al., 2016), ranging from the Ordovician to the Pleistocene. The Holocene as well as stages older than the Ordovician were excluded from the analysis. Ordovician stages were resolved using biozone and formaFair sampling of taxonomic richness and unbiased estimation of origination and extinction ratestion entries due to potential stratigraphic errors (À. T. Kocsis et al., 2019). Further data cleaning included removal of uncertain taxonomical ranks, duplicates in bins, single-collection, and single-reference taxa as well as missing higher-level taxonomy. Taxa confined to a single stage (singletons) were excluded as they tend to produce undesirable distortions of the fossil record (Foote, 2000). Collections from unlithified sediments were omitted to reduce sampling bias (John Alroy et al., 2008). We then applied shareholder quorum subsampling (SQS) for sampling standardisation (John Alroy, 2010). SQS is based on frequency distribution coverage of taxa, drawing collections until estimated coverage reaches a fixed target (i.e., until a shareholder quorum is attained). For this, we used a shareholder quorum of 0.8. SQS was applied excluding dominant taxa from all calculations involving frequencies and with a separate treatment of the largest collection in each time slice. We then transformed occurrence data into ranges congruent to a time series with one single origination and extinction event for each genus. The final data set contained ten major marine fossil phyla

Climate proxy data

To reconstruct temperature change over time, we used the tropical whole surface water (mixed layers < 300 m deep) oxygen isotope dataset from a compilation of marine carbonate isotopes (Veizer & Prokoph, 2015). The δ18O values of well-preserved calcareous shells are often considered as the best available temperature proxy for the fossil record (Song et al. 2019). To reduce bias while calculating palaeotemperature from the raw δ18O values, we followed the data processing of Reddin et al. (2018), including adjustments for the long-term trend in oxygen isotopic composition of seawater as well as averaging of tropical and subtropical records. We subsequently binned temperature data based on isotope values to geological stages to provide global mean temperature for each of the 80 stages used in our analysis.

Statistical analysis

*Generalised linear mixed effect models*

All analyses were carried out in R (R Core Team 2020). We used the lme4 package (Bates et al. 2015) to perform the analysis, and the tidyverse package (Wickham et al. 2019) for data processing and visualisation. We quantified the effect of temperature change interacting with past temperature trends on origination probability using generalised linear mixed effect models with a binomial family error (GLMMs) (Bolker et al. 2009). The additive and multiplicative effects of cumulative temperature change were tested against the probability of origination in each time interval in the form:   
glm (extinct ~ ∆Ttrend : ∆Tchange + (1 | Genus), family = binomial).

We calculated the change in temperature for each geologic change compared to the previous stage (∆Tchange). We computed long-term temperature trends (∆Ttrend) from ordinary least squares. Each of the trends covered a successively growing time of temperature history: Trend.st1 ranged one stage back, trend.st2 two stages, …, trend.st10 ten stages. To exclude ∆Tchange from the long-term trend, each ∆Ttrend started one stage prior to the focal stage. We controlled for the fact that observations on the same genus are non-independent by including genus identity as a random effect. We used a dynamic model framework to select the best performing ∆Ttrend. This implies calculating 10 GLMMs for each ∆Ttrend while keeping ∆Tchange fixed. Instead of calculating one model for all ∆Tchange, we distinguished short-term warming from short-term cooling. This enabled us to quantify the differences in origination likelihood per palaeoclimate interaction (cooling-cooling, warming-cooling, cooling-warming, warming-warming). For comparison of model performance, Akaike’s information criterion (AIC) was used (Burnham und Anderson 2002). For likelihood approximation, we used adaptive Gauss-Hermite quadrature (Pinheiro und Chao 2006), which reduces bias and is more robust than other approximation methods (Bolker et al. 2009). We used 25 quadrature points per scalar integral, which is the maximum number of nodes in the quadrature formula of the lme4 package.

*Model comparison*

Previous studies used different estimates of temperature to deduce an association between origination and temperature (Cárdenas und Harries 2010; Erwin 2009; Krug et al. 2009; Mayhew et al. 2008; Mayhew et al. 2012). Simple models correlate static temperature estimates at each interval to biodiversity at the same interval. More complex models include temporal lags of temperature, allowing for a suspended response of biota to temperature change. Most recent models are based on both estimates but additionally allow for interactions of static estimates and temporal lags. We hence calculated a GLMM for each model structure from simple to complex. We compared the performance of each model to a model allowing for palaeoclimate interactions. The latter were based on the same model structure, but explicitly allowed for interactions of long-term temperature trends and short-term climate changes within a dynamic modelling framework. Dynamic implies that we let the long-term trend vary for each palaeoclimate interaction, and subsequently selected the best performing model. Comparison of model performance was based on AIC values with results being consistent with Bayesian information criterion (Neath und Cavanaugh 2012).

*Origination probability*

We separated the predicted origination response for each type of palaeoclimate interaction from our GLMMs. To test whether the response after a specific palaeoclimate interaction significantly differs from the mean origination probability, we used Wilcoxon rank sum tests.

*Effect size*

We applied bootstrapping (Mooney und Duval 1993) and Bayesian estimation (Kruschke 2013) to the predicted origination response of our GLMMS to calculate the difference in means between origination probability after cooling-cooling compared to all other palaeoclimate interactions. Bootstrapping was based on 2000 repetitions. For Bayesian estimation, we used a markov chain monte carlo method with four chains, 1000 iterations, and 500 warm-up iterations. We used an informed prior for the group mean based on the origination response of the whole sample, and a default exponential prior with a rate of 1/29 for the group variances. Overall increase of origination probability after cooling-cooling palaeoclimate interactions was calculated by dividing the estimated difference of means by the baseline probability of origination. Cohen’s d effect size was calculated based on the raw output of our GLMMs as well as posterior distributions after Bayesian estimation using the formula:

*Group differences*

Based on the results from our GLMMS, we calculated the log-odds ratio for each major fossil group and for five periods spanning all studied stages. The oldest period ranged from the Tremadocian to Lochkovian, the second from the Pragian to Artinskian, the third from the Kungurian to Pliensbachian, the fourth from the Toarcian to Turonian, and the youngest period from the Coniacian to Pleistocene.

References

Alroy, John (2008): Dynamics of origination and extinction in the marine fossil record. In: *Proceedings of the National Academy of Sciences* 105 (Supplement 1), S. 11536–11542.

Alroy, John (2010a): Fair sampling of taxonomic richness and unbiased estimation of origination and extinction rates. In: *The Paleontological Society Papers* 16, S. 55–80.

Alroy, John (2010b): Geographical, environmental and intrinsic biotic controls on Phanerozoic marine diversification. In: *Palaeontology* 53 (6), S. 1211–1235.

Alroy, John; Aberhan, Martin; Bottjer, David J.; Foote, Michael; Fürsich, Franz T.; Harries, Peter J. et al. (2008): Phanerozoic trends in the global diversity of marine invertebrates. In: *Science* 321 (5885), S. 97–100.

Alroy, John; Koch, Paul L.; Zachos, James C. (2000): Global climate change and North American mammalian evolution. In: *Paleobiology* 26 (S4), S. 259–288.

Antão, Laura H.; Bates, Amanda E.; Blowes, Shane A.; Waldock, Conor; Supp, Sarah R.; Magurran, Anne E. et al. (2020): Temperature-related biodiversity change across temperate marine and terrestrial systems. In: *Nature ecology & evolution*.

Bates, Douglas; Mächler, Martin; Bolker, Ben; Walker, Steve (2015): Fitting Linear Mixed-Effects Models Using lme4. In: *J. Stat. Soft.* 67 (1). DOI: 10.18637/jss.v067.i01.

Bolker, Benjamin M.; Brooks, Mollie E.; Clark, Connie J.; Geange, Shane W.; Poulsen, John R.; Stevens, M. Henry H.; White, Jada-Simone S. (2009): Generalized linear mixed models: a practical guide for ecology and evolution. In: *Trends in ecology & evolution* 24 (3), S. 127–135.

Broecker, Wallace S.; Peng, Tsung-Hung (1982): Tracers in the Sea. In: *Lamont-Doherty Geol. Observ. Publ.*

Burnham, Kenneth P.; Anderson, David R. (2002): A practical information-theoretic approach. In: *Model selection and multimodel inference, 2nd ed. Springer, New York*.

Button, David J.; Lloyd, Graeme T.; Ezcurra, Martín D.; Butler, Richard J. (2017): Mass extinctions drove increased global faunal cosmopolitanism on the supercontinent Pangaea. In: *Nature communications* 8 (1), S. 1–8.

Cárdenas, Andrés L.; Harries, Peter J. (2010): Effect of nutrient availability on marine origination rates throughout the Phanerozoic eon. In: *Nature Geoscience* 3 (6), S. 430–434.

Chen, I-Ching; Hill, Jane K.; Ohlemüller, Ralf; Roy, David B.; Thomas, Chris D. (2011): Rapid range shifts of species associated with high levels of climate warming. In: *Science* 333 (6045), S. 1024–1026.

Dietl, Gregory P.; Kidwell, Susan M.; Brenner, Mark; Burney, David A.; Flessa, Karl W.; Jackson, Stephen T.; Koch, Paul L. (2015): Conservation Paleobiology: Leveraging Knowledge of the Past to Inform Conservation and Restoration. In: *Annu. Rev. Earth Planet. Sci.* 43 (1), S. 79–103.

Erwin, Douglas H. (2009): Climate as a driver of evolutionary change. In: *Current Biology* 19 (14), R575-R583.

Foote, Mike (2000): Origination and extinction components of taxonomic diversity: general problems. In: *Paleobiology* 26 (S4), S. 74–102.

Gilman, Sarah E.; Urban, Mark C.; Tewksbury, Joshua; Gilchrist, George W.; Holt, Robert D. (2010): A framework for community interactions under climate change. In: *Trends in ecology & evolution* 25 (6), S. 325–331.

Hopkins, Melanie J.; Simpson, Carl; Kiessling, Wolfgang (2014): Differential niche dynamics among major marine invertebrate clades. In: *Ecology letters* 17 (3), S. 314–323.

Humboldt, Alexander von (1808): Ansichten der Natur mit Wissenschaftlichen Erlåuterungen, Vol. 1: Tübingen.

Jablonski, David (2008): Extinction and the spatial dynamics of biodiversity. In: *Proceedings of the National Academy of Sciences* 105 (Supplement 1), S. 11528–11535.

Kocsis, Adam T.; Reddin, Carl J.; Alroy, John; Kiessling, Wolfgang (2019): The R package divDyn for quantifying diversity dynamics using fossil sampling data. In: *Methods in Ecology and Evolution* 10 (5), S. 735–743.

Krug, Andrew Z.; Jablonski, David; Valentine, James W.; Roy, Kaustuv (2009): Generation of Earth's first-order biodiversity pattern. In: *Astrobiology* 9 (1), S. 113–124.

Kruschke, John K. (2013): Bayesian estimation supersedes the t test. In: *Journal of Experimental Psychology: General* 142 (2), S. 573.

Mayhew, Peter J.; Bell, Mark A.; Benton, Timothy G.; McGowan, Alistair J. (2012): Biodiversity tracks temperature over time. In: *Proceedings of the National Academy of Sciences of the United States of America* 109 (38), S. 15141–15145.

Mayhew, Peter J.; Jenkins, Gareth B.; Benton, Timothy G. (2008): A long-term association between global temperature and biodiversity, origination and extinction in the fossil record. In: *Proceedings. Biological sciences* 275 (1630), S. 47–53.

Mayr, Ernst; O'Hara, Robert J. (1986): The biogeographic evidence supporting the Pleistocene forest refuge hypothesis. In: *Evolution* 40 (1), S. 55–67.

Mooney, Christopher Z.; Duval, Robert D. (1993): Bootstrapping: A nonparametric approach to statistical inference: sage (95).

Neath, Andrew A.; Cavanaugh, Joseph E. (2012): The Bayesian information criterion: background, derivation, and applications. In: *Wiley Interdisciplinary Reviews: Computational Statistics* 4 (2), S. 199–203.

Nogués-Bravo, David; Rodríguez-Sánchez, Francisco; Orsini, Luisa; Boer, Erik de; Jansson, Roland; Morlon, Helene et al. (2018): Cracking the code of biodiversity responses to past climate change. In: *Trends in ecology & evolution* 33 (10), S. 765–776.

Ogg, James George; Ogg, Gabi M.; Gradstein, Felix M. (2016): A concise geologic time scale: 2016: Elsevier.

Pecl, Gretta T.; Araújo, Miguel B.; Bell, Johann D.; Blanchard, Julia; Bonebrake, Timothy C.; Chen, I-Ching et al. (2017): Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. In: *Science* 355 (6332).

Peters, Shanan E. (2005): Geologic constraints on the macroevolutionary history of marine animals. In: *Proceedings of the National Academy of Sciences* 102 (35), S. 12326–12331.

Peters, Shanan E. (2006): Genus extinction, origination, and the durations of sedimentary hiatuses. In: *Paleobiology* 32 (3), S. 387–407.

Pinheiro, José C.; Chao, Edward C. (2006): Efficient Laplacian and adaptive Gaussian quadrature algorithms for multilevel generalized linear mixed models. In: *Journal of Computational and Graphical Statistics* 15 (1), S. 58–81.

Prothero, Donald R. (1999): Does climatic change drive mammalian evolution. In: *GSA today* 9 (9), S. 1–7.

R Core Team (2020): R. A language and environment for statistical computing. Version 4.0.0 (2020-04-24). Vienna, Austria: R Foundation for Statistical Computing. Online verfügbar unter https://www.R-project.org/.

Raup, David M. (1979): Size of the Permo-Triassic bottleneck and its evolutionary implications. In: *Science* 206 (4415), S. 217–218.

Reddin, Carl J.; Kocsis, Ádám T.; Kiessling, Wolfgang (2018): Marine invertebrate migrations trace climate change over 450 million years. In: *Global Ecology and Biogeography* 27 (6), S. 704–713.

Ritterbush, Kathleen A.; Foote, Michael (2017): Association between geographic range and initial survival of Mesozoic marine animal genera: circumventing the confounding effects of temporal and taxonomic heterogeneity. In: *Paleobiology* 43 (2), S. 209–223.

Schuler, Matthew S.; Chase, Jonathan M.; Knight, Tiffany M. (2017): Habitat patch size alters the importance of dispersal for species diversity in an experimental freshwater community. In: *Ecology and evolution* 7 (15), S. 5774–5783.

Sepkoski, J. John (2002): A compendium of fossil marine animal genera. In: *Bulletins of American paleontology* 363, S. 1–560.

Song, Haijun; Wignall, Paul B.; Song, Huyue; Dai, Xu; Chu, Daoliang (2019): Seawater temperature and dissolved oxygen over the past 500 million years. In: *Journal of Earth Science* 30 (2), S. 236–243.

Stigall, Alycia L. (2013): Analysing links between biogeography, niche stability and speciation: the impact of complex feedbacks on macroevolutionary patterns. In: *Palaeontology* 56 (6), S. 1225–1238.

Stigall, Alycia L. (2014): When and how do species achieve niche stability over long time scales? In: *Ecography* 37 (11), S. 1123–1132.

Svenning, Jens-Christian; Eiserhardt, Wolf L.; Normand, Signe; Ordonez, Alejandro; Sandel, Brody (2015): The Influence of Paleoclimate on Present-Day Patterns in Biodiversity and Ecosystems. In: *Annu. Rev. Ecol. Evol. Syst.* 46 (1), S. 551–572.

Templeton, Alan R. (2008): The reality and importance of founder speciation in evolution. In: *BioEssays* 30 (5), S. 470–479.

Tilman, David (1994): Competition and biodiversity in spatially structured habitats. In: *Ecology* 75 (1), S. 2–16.

Valentine, James W. (1968): Climatic regulation of species diversification and extinction. In: *Geological Society of America Bulletin* 79 (2), S. 273–276.

Valentine, James W. (1974): Temporal bias in extinctions among taxonomic categories. In: *Journal of Paleontology*, S. 549–552.

Veizer, Jan; Prokoph, Andreas (2015): Temperatures and oxygen isotopic composition of Phanerozoic oceans. In: *Earth-Science Reviews* 146, S. 92–104.

Wahl, Lindi M.; Gerrish, Philip J.; Saika-Voivod, Ivan (2002): Evaluating the impact of population bottlenecks in experimental evolution. In: *Genetics* 162 (2), S. 961–971.

Wickham, Hadley; Averick, Mara; Bryan, Jennifer; Chang, Winston; McGowan, Lucy; François, Romain et al. (2019): Welcome to the Tidyverse. In: *Journal of Open Source Software* 4 (43), S. 1686.

Wiens, John J.; Graham, Catherine H. (2005): Niche conservatism: integrating evolution, ecology, and conservation biology. In: *Annual review of ecology, evolution, and systematics* 36, S. 519–539.

**Supplemental Information**

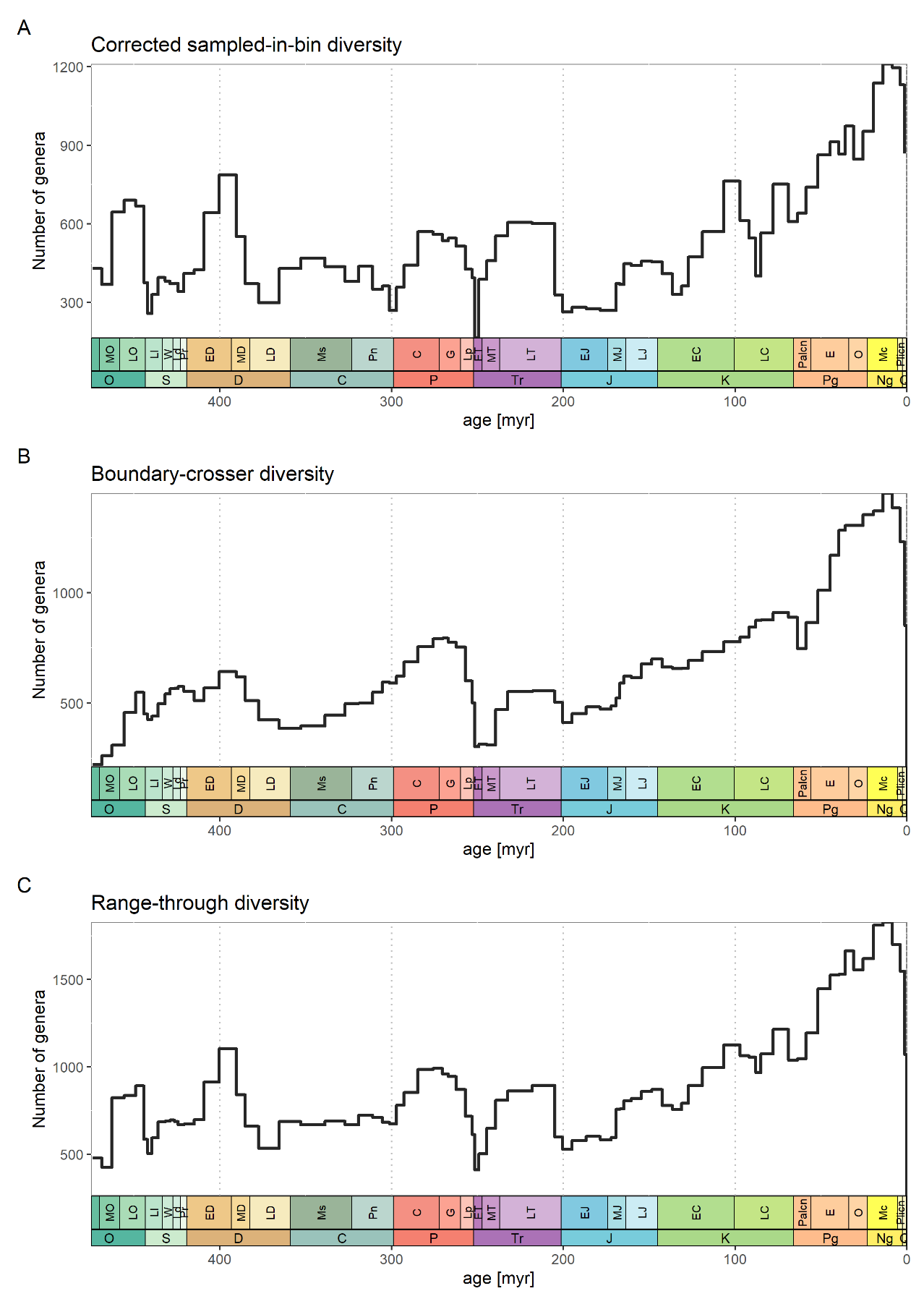


Fig. S 1 | Number of fossil genera within data sets. Diversity metrics shown here are based on filtered and sampling-standardised data. A) Sampled-in-bin diversity corrected for three-timer sampling completeness. B) Boundary-crosser diversity, which is the number of taxa with ranges crossing the boundaries of the interval. C) Range-through diversity, which is based on all taxa in the interval.

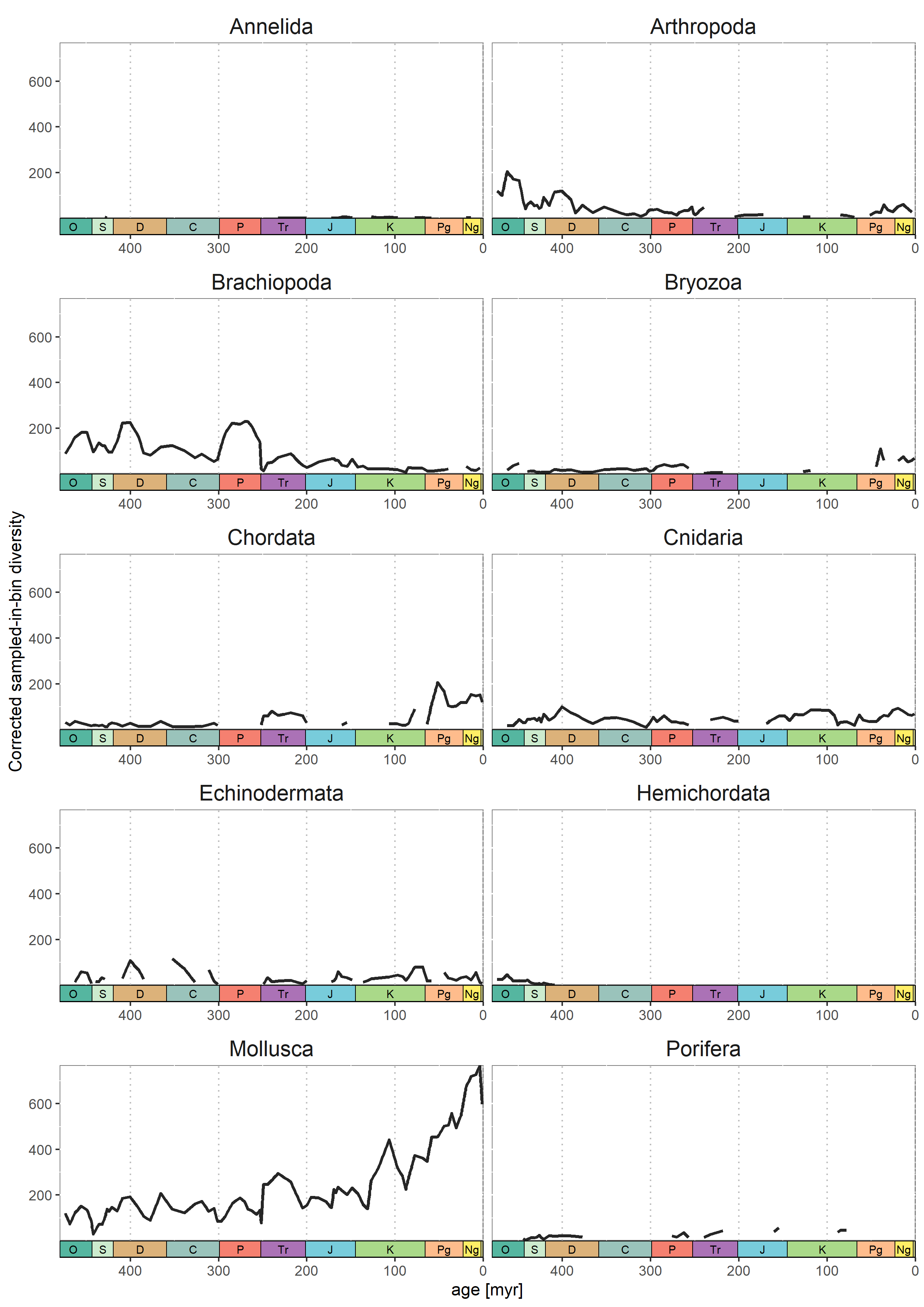


Fig. S 2 | Number of fossil genera for every major phylum, using the corrected sampled-in-bin diversity metric.

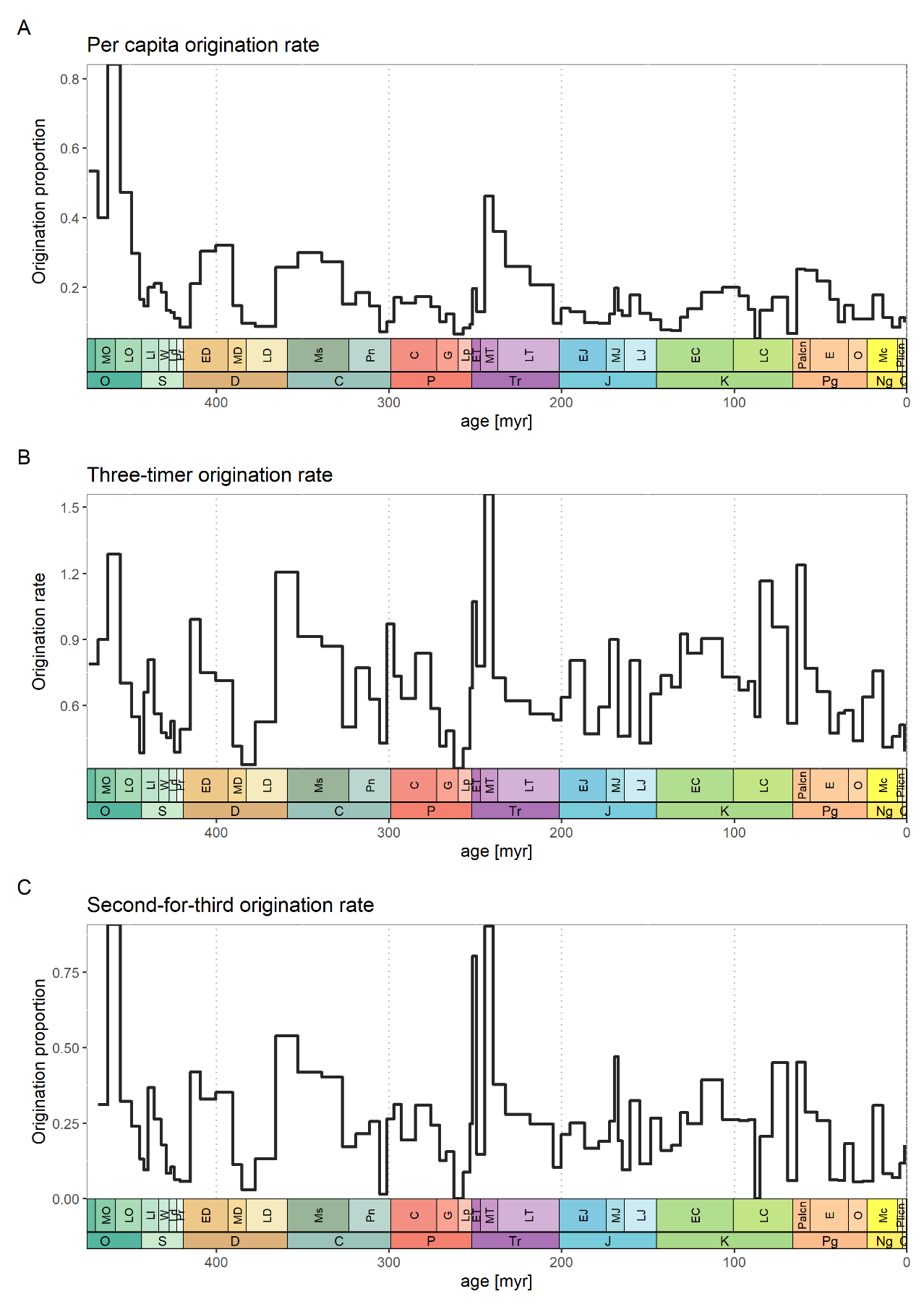


Fig. S 3 | Origination rates for all studied phyla based on filtered and sampling-standardised data. A) Per capita origination with values not normalised with bin lengths. B) Three-timer origination rates with values normalised with bin lengths. C) Second-for-third extinction propotions based on second-for-third substitution of taxa categories.

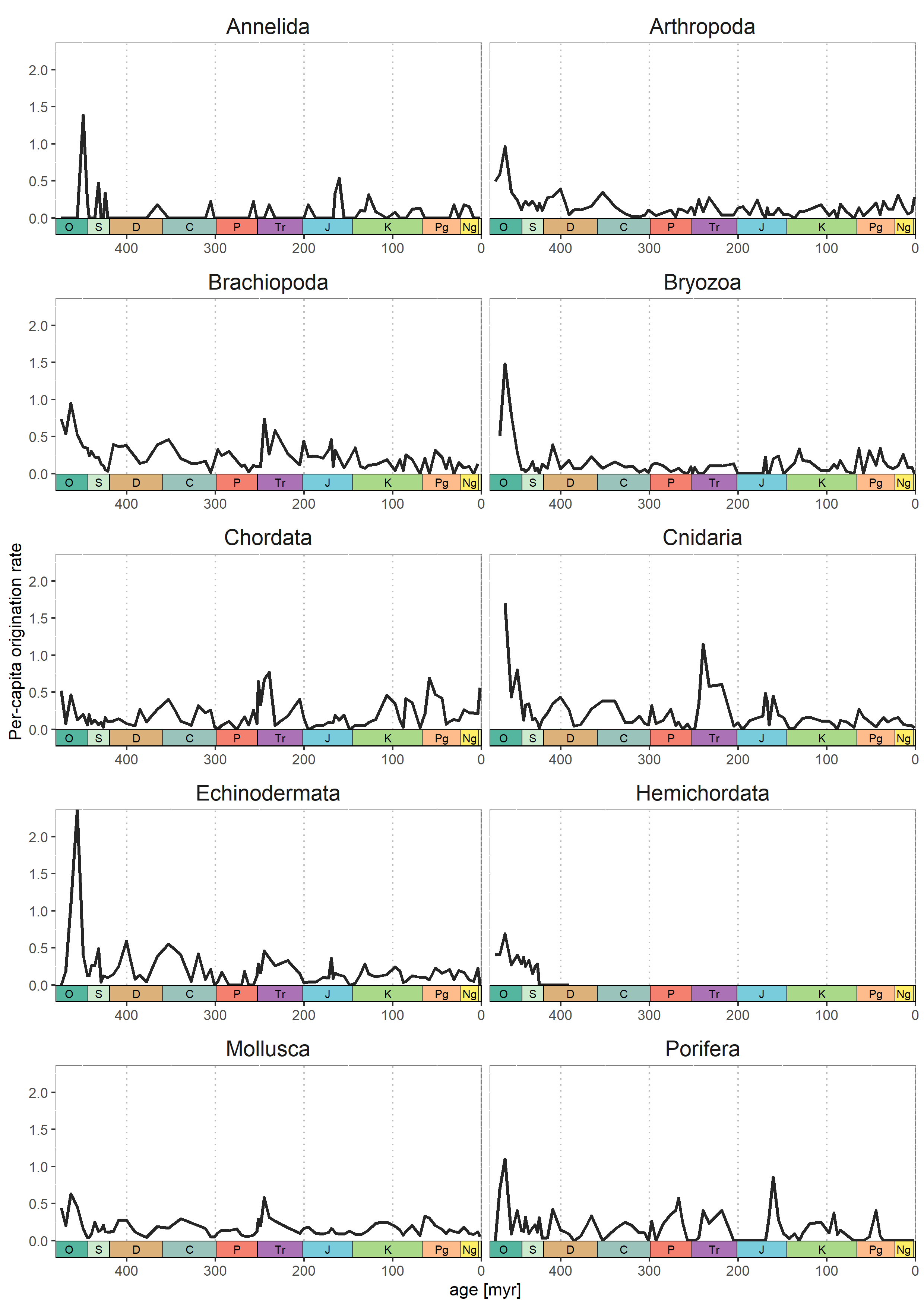


Fig. S 4 | Per-capita origination rates for every major phylum.

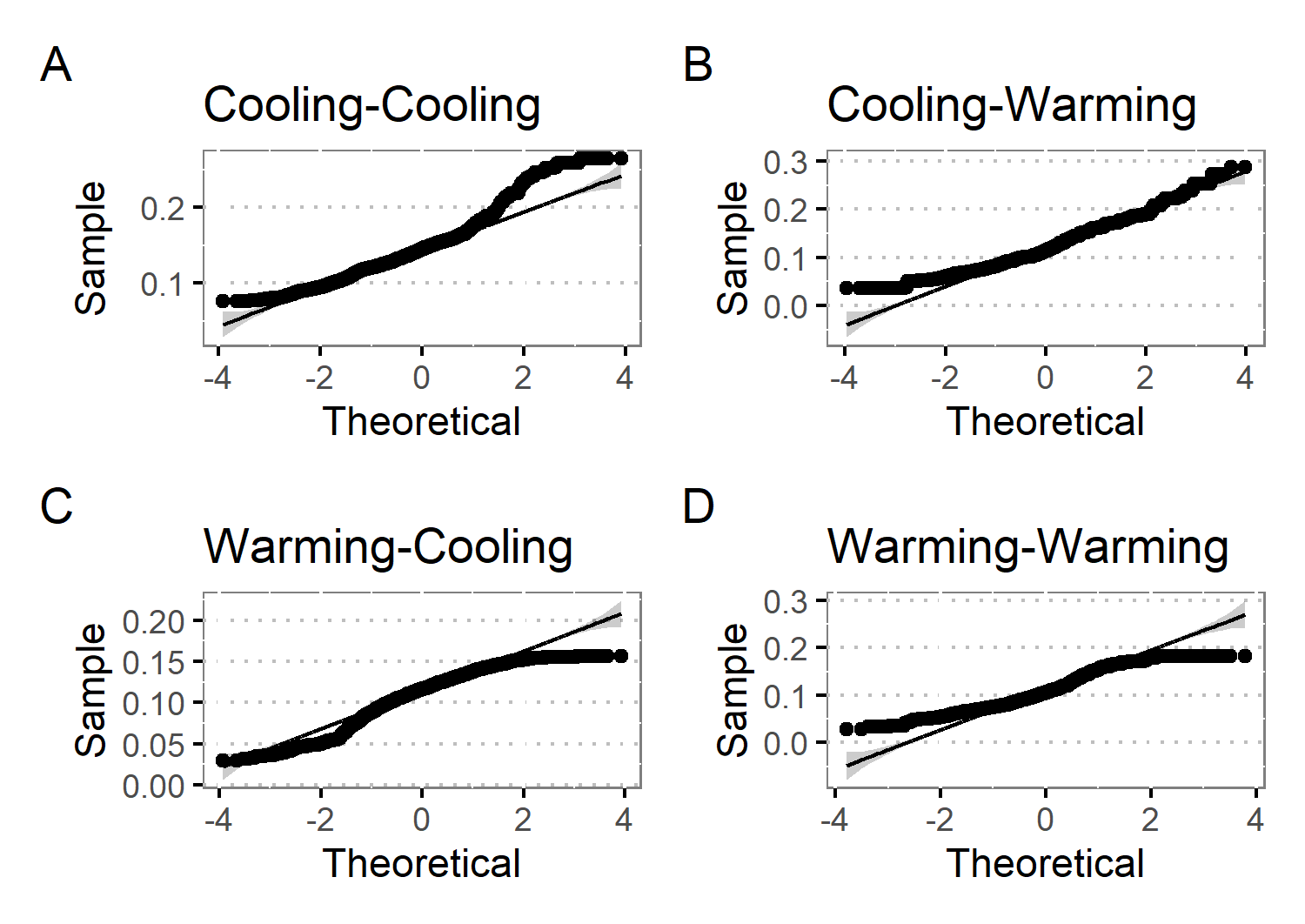


Fig. S 5 | Quantile-quantile plots for predicted responses of marine fossil groups after palaeoclimate interactions. As these plots indicate deviations from normality, further estimates (difference in means, percentage change, effect change) were calculated using non-parametric methods.

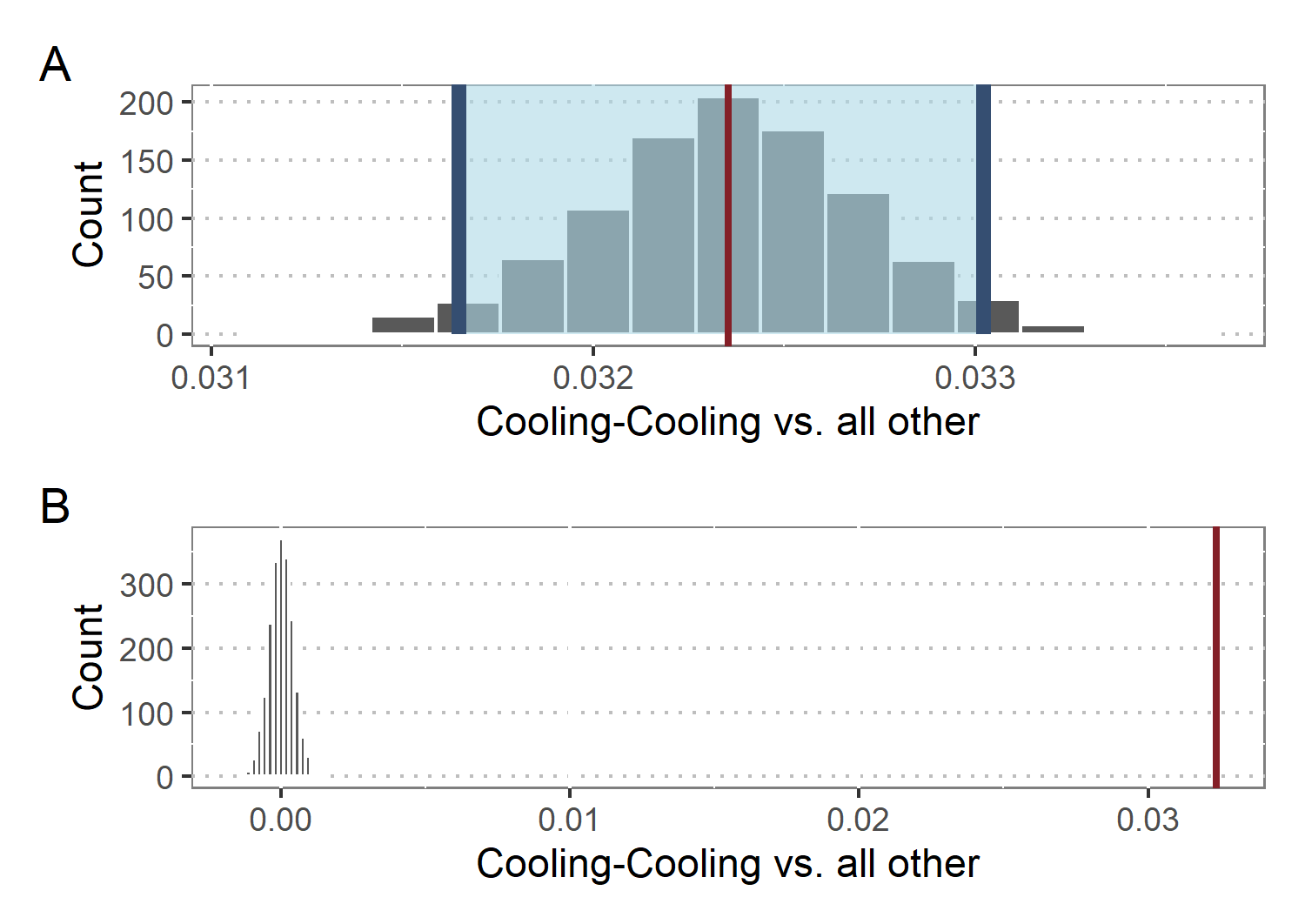


Fig. S 6 | Difference in means of origionation responses after cooling-cooling compared to all other palaeoclimate interactions, based on bootstrapping resampling. A) Bootstrapped distribution of differences in means. Red line shows observed difference and shaded area shows 95% confidence interval. B) Simulation-based null distribution of difference in means compared to observed difference (red line).

Table S 1 | Number of classes, order, families, and genera within every major phylum after data-cleaning and sampling-standardisation.

| **Phylum** | **Class** | **Order** | **Family** | **Genus** |
| --- | --- | --- | --- | --- |
| Annelida | 1 | 2 | 4 | 11 |
| Arthropoda | 6 | 25 | 114 | 519 |
| Brachiopoda | 6 | 22 | 296 | 1292 |
| Bryozoa | 2 | 15 | 114 | 257 |
| Chordata | 9 | 95 | 227 | 412 |
| Cnidaria | 3 | 15 | 177 | 575 |
| Echinodermata | 9 | 49 | 125 | 237 |
| Hemichordata | 1 | 3 | 6 | 20 |
| Hyolitha | 1 | 1 | 1 | 1 |
| Mollusca | 9 | 85 | 621 | 2449 |
| Nematoda | 1 | 1 | 1 | 1 |
| Porifera | 5 | 33 | 115 | 253 |

Table S 2 | Comparison of model performance for traditional models covering first- and second-order relationships of origination and climate versus models allowing for palaeoclimate interactions. The latter are based on the same model structure, but explicitly allow for interactions of long-term temperature trends and short-term climate changes within a dynamic modeling framework (+ Pal. Int). Each model was tested for overdispersion including significance (indicated by stars). Model comparison was based on Akaikes information criterion (AIC) and is consistent with Bayesian information criterion (BIC).

| **Model Structure** | **Overdispersed** | **AIC** | **BIC** |
| --- | --- | --- | --- |
| ~ Warming | no \*\*\* | 15762.2 | 15786.0 |
| ~ Warming + Pal. Int. | no \*\*\* | 15702.5 | 15734.3 |
| ~ Cooling + Lag | no \*\*\* | 18427.6 | 18451.8 |
| ~ Cooling + Pal. Int. | no \*\*\* | 17869.4 | 17901.5 |
| ~ Warming + Lag | no \*\*\* | 14976.0 | 15007.6 |
| ~ Warming + Lag + Pal. Int. | no \*\*\* | 14901.3 | 14940.8 |
| ~ Cooling + Lag | no \*\*\* | 14763.2 | 14794.8 |
| ~ Cooling + Lag + Pal. Int. | no \*\*\* | 14748.5 | 14788.1 |
| ~ Warming + Lag + Warming:Lag | no \*\*\* | 14969.9 | 15009.4 |
| ~ Warming + Lag + Warming:Lag + Pal. Int. | no \*\*\* | 14886.2 | 14933.6 |
| ~ Cooling + Lag + Cooling:Lag | no \*\*\* | 14680.0 | 14719.6 |
| ~ Cooling + Lag + Cooling:Lag + Pal. Int. | no \*\*\* | 14667.4 | 14714.9 |

Table S 3 | Model summaries for all palaeclimate interactions based on short-term warming and short-term cooling. Ten different models with each having a different long-term trend, ranging from one stage (trend.st1) to ten stages (trend.st10), were calculated while keeping the short-term change fixed. Subsequently, one final model was selected exposing the best performance. This dynamic approach to defining the length of the long-term trend guarantees the selection of a model explaining the highest variance, without conduction multiple hypothesis testing. Shown are the coefficients for the intercept and interaction terms, including stars indicating significance based on adaptive Gauss-Hermite quadrature approximation.

| **Type** | **Model** | **Intercept** | **Interaction** |
| --- | --- | --- | --- |
| Warming | trend.st1 | -1.89 +- 0.03 \*\*\* | 0.05 +- 0.05 |
| trend.st2 | -1.91 +- 0.03 \*\*\* | -0.23 +- 0.07 \*\*\* |
| trend.st3 | -1.91 +- 0.03 \*\*\* | -0.4 +- 0.1 \*\*\* |
| trend.st4 | -1.91 +- 0.03 \*\*\* | -0.2 +- 0.11 |
| trend.st5 | -1.92 +- 0.03 \*\*\* | -0.43 +- 0.1 \*\*\* |
| trend.st6 | -1.93 +- 0.03 \*\*\* | -0.58 +- 0.1 \*\*\* |
| trend.st7 | -1.93 +- 0.03 \*\*\* | -0.6 +- 0.11 \*\*\* |
| trend.st8 | -1.93 +- 0.03 \*\*\* | -0.61 +- 0.12 \*\*\* |
| trend.st9 | -1.9 +- 0.03 \*\*\* | -0.21 +- 0.13 |
| trend.st10 | -1.89 +- 0.03 \*\*\* | 0.05 +- 0.13 |
| Cooling | trend.st1 | -1.81 +- 0.02 \*\*\* | 0.17 +- 0.04 \*\*\* |
| trend.st2 | -1.83 +- 0.02 \*\*\* | 0.26 +- 0.07 \*\*\* |
| trend.st3 | -1.81 +- 0.02 \*\*\* | 0.76 +- 0.11 \*\*\* |
| trend.st4 | -1.79 +- 0.02 \*\*\* | 1.31 +- 0.11 \*\*\* |
| trend.st5 | -1.81 +- 0.02 \*\*\* | 1.27 +- 0.11 \*\*\* |
| trend.st6 | -1.82 +- 0.02 \*\*\* | 1.32 +- 0.11 \*\*\* |
| trend.st7 | -1.85 +- 0.02 \*\*\* | 1.27 +- 0.12 \*\*\* |
| trend.st8 | -1.87 +- 0.02 \*\*\* | 1.26 +- 0.12 \*\*\* |
| trend.st9 | -1.88 +- 0.02 \*\*\* | 1.31 +- 0.13 \*\*\* |
| trend.st10 | -1.89 +- 0.02 \*\*\* | 1.34 +- 0.14 \*\*\* |

Table S 4 | Continuation of Table S 3, showing values for Akaikes information criterion (AIC), Bayesiane information criterion (BIC), and their respective delta values. The model with the lowest Aic value (∆Aic == 0.0) was selected for further analysis and selection was consistent with BIC values.

| **Type** | **Model** | **AIC** | **BIC** | **ΔAIC** | **ΔBIC** |
| --- | --- | --- | --- | --- | --- |
| Warming | trend.st1 | 15908.6 | 15932.5 | 28.8 | 28.9 |
| trend.st2 | 15897.7 | 15921.5 | 17.9 | 17.9 |
| trend.st3 | 15893.4 | 15917.3 | 13.6 | 13.7 |
| trend.st4 | 15906.8 | 15930.6 | 27.0 | 27.0 |
| trend.st5 | 15889.7 | 15913.5 | 9.9 | 9.9 |
| trend.st6 | 15880.1 | 15903.9 | 0.3 | 0.3 |
| trend.st7 | 15879.8 | 15903.6 | 0.0 | 0.0 |
| trend.st8 | 15883.4 | 15907.2 | 3.6 | 3.6 |
| trend.st9 | 15907.1 | 15931.0 | 27.3 | 27.4 |
| trend.st10 | 15909.6 | 15933.4 | 29.8 | 29.8 |
| Cooling | trend.st1 | 18159.7 | 18183.8 | 140.3 | 140.3 |
| trend.st2 | 18159.1 | 18183.2 | 139.7 | 139.7 |
| trend.st3 | 18123.4 | 18147.6 | 104.0 | 104.1 |
| trend.st4 | 18024.9 | 18049.0 | 5.5 | 5.5 |
| trend.st5 | 18028.7 | 18052.8 | 9.3 | 9.3 |
| trend.st6 | 18019.4 | 18043.5 | 0.0 | 0.0 |
| trend.st7 | 18048.1 | 18072.2 | 28.7 | 28.7 |
| trend.st8 | 18067.8 | 18091.9 | 48.4 | 48.4 |
| trend.st9 | 18076.8 | 18100.9 | 57.4 | 57.4 |
| trend.st10 | 18082.5 | 18106.7 | 63.1 | 63.2 |

Table S 5 | Generalised mixed effect model summaries for the final palaeoclimate interaction model based on short-term warming and cooling. Both models were selected from a pool of models with differing long-term temperature trends based on model performance.

| **Model** | **Term** | **Estimate** | **Std.error** | **Z value** | **P value** | **Group** |
| --- | --- | --- | --- | --- | --- | --- |
| Warming | Intercept | -1.930 | 0.026 | -74.396 | \*\*\* | fixed |
| Warming:Trend.st7 | -0.604 | 0.108 | -5.592 | \*\*\* | fixed |
| Random effect | 0.729 | NA | NA |  | genus |
| Cooling | Intercept | -1.824 | 0.021 | -86.255 | \*\*\* | fixed |
| Cooling:Trend.st6 | 1.321 | 0.109 | 12.072 | \*\*\* | fixed |
| Random effect | 0.423 | NA | NA |  | genus |

Table S 6 | Continuation of Table S 5, showing model performance measures as well as overdisperison tests for each final GLMM. Significance estimates for beeing not overdispersed are indicated by stars.

| **Model** | **Sigma** | **LogLik** | **AIC** | **BIC** | **Deviance** | **DF residual** | **Overdispersion** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Warming | 1 | -7936.88 | 15879.77 | 15903.62 | 14052.15 | 20919 | no \*\*\* |
| Cooling | 1 | -9006.70 | 18019.40 | 18043.53 | 17190.21 | 22981 | no \*\*\* |

Table S 7 | Distribution of origination probability after each palaeoclimate interaction for all studied fossil groups based on predictions of final GLMMs.

| **Palaeoclimate Interaction** | **Lower Quartile** | **Median** | **Upper Quartile** |
| --- | --- | --- | --- |
| Cooling-Cooling | 0.127 | 0.145 | 0.161 |
| Cooling-Warming | 0.093 | 0.114 | 0.147 |
| Warming-Cooling | 0.099 | 0.116 | 0.131 |
| Warming-Warming | 0.082 | 0.106 | 0.139 |

Table S 8 | Effect size estimates for the difference in means of origination probability, overall increase of origination probability and Cohen’s d effect size including 95% confidence intervals (CI) and the method to calculate each estimate.

| **Parameter** | **Lower CI** | **Estimate** | **Upper CI** | **Method** |
| --- | --- | --- | --- | --- |
| Difference in means | 3.17 | 3.24 | 3.30 | Bootstrapping |
| 3.09 | 3.15 | 3.22 | Bayesian Estimate |
| Percentage change | 23.82 | 24.36 | 24.86 | Bootstrapping |
| 23.24 | 23.70 | 24.23 | Bayesian Estimate |
| Cohen's d | 0.98 | 1.00 | 1.02 | Raw Data |
| 1.02 | 1.05 | 1.07 | Bayesian Estimate |

Table S 9 | Origination probability for all major phyla and throughout time after cooling-cooling compared to all other palaeoclimate interactions calculated as log odds ratio, including 95% confidence intervals. Stage 14 is the oldest stage included in analysis, and stage 94 the youngest.

| **Group** | **Lower CI** | **Log Odds ratio** | **Upper CI** |
| --- | --- | --- | --- |
| Total | 2.01 | 2.08 | 2.14 |
| Annelida | 0.90 | 2.95 | 5.01 |
| Arthropoda | 2.53 | 2.77 | 3.01 |
| Brachiopoda | 2.17 | 2.31 | 2.45 |
| Bryozoa | 2.19 | 2.49 | 2.79 |
| Chordata | 1.75 | 2.01 | 2.27 |
| Cnidaria | 1.71 | 1.91 | 2.12 |
| Echinodermata | 1.54 | 1.85 | 2.16 |
| Hemichordata | 2.54 | 4.79 | 7.04 |
| Mollusca | 1.87 | 1.97 | 2.07 |
| Porifera | 1.47 | 1.76 | 2.05 |
| Stage 14:29 | 3.32 | 3.48 | 3.65 |
| Stage 30:45 | 1.03 | 1.22 | 1.41 |
| Stage 46:61 | 0.88 | 1.02 | 1.15 |
| Stage 62:77 | 1.60 | 1.84 | 2.07 |
| Stage 78:94 | 2.42 | 2.56 | 2.69 |