| **Model Structure** | **overdispersed** | **AIC** | **BIC** |
| --- | --- | --- | --- |
| ~ Warming | no \*\*\* | 16576.2 | 16600.2 |
| ~ Warming + Pal. Int. | no \*\*\* | 16521.9 | 16553.8 |
| ~ Cooling + Lag | no \*\*\* | 19523.6 | 19547.9 |
| ~ Cooling + Pal. Int. | no \*\*\* | 18883.5 | 18915.9 |
| ~ Warming + Lag | no \*\*\* | 15762.3 | 15794.1 |
| ~ Warming + Lag + Pal. Int. | no \*\*\* | 15696.8 | 15736.6 |
| ~ Cooling + Lag | no \*\*\* | 15756.3 | 15788.2 |
| ~ Cooling + Lag + Pal. Int. | no \*\*\* | 15730.3 | 15770.1 |
| ~ Warming + Lag + Warming:Lag | no \*\*\* | 15751.0 | 15790.7 |
| ~ Warming + Lag + Warming:Lag + Pal. Int. | no \*\*\* | 15674.4 | 15722.1 |
| ~ Cooling + Lag + Cooling:Lag | no \*\*\* | 15660.2 | 15700.1 |
| ~ Cooling + Lag + Cooling:Lag + Pal. Int. | no \*\*\* | 15637.8 | 15685.6 |

Suppl. Table 1 | 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, based on adaptive Gauss-Hermite quadrature approximation). Model comparison was based on Akaikes information criterion (AIC) and is consistent with Bayesian information criterion (BIC).

| **Palaeoclimate Interaction** | **Lower Quartile** | **Median** | **Upper Quartile** |
| --- | --- | --- | --- |
| Cooling-Cooling | 0.126 | 0.141 | 0.160 |
| Cooling-Warming | 0.095 | 0.114 | 0.148 |
| Warming-Cooling | 0.101 | 0.123 | 0.138 |
| Warming-Warming | 0.082 | 0.104 | 0.136 |

Suppl. Table 2 | Distribution of origination probability after each palaeoclimate interaction for all studied fossil groups based on predictions of final GLMMs.

| **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 |

Suppl. Table 3 | 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.

| **Group** | **Lower CI** | **Log Odds ratio** | **Upper CI** |
| --- | --- | --- | --- |
| Total | 1.58 | 1.64 | 1.70 |
| Annelida | 0.76 | 1.89 | 3.01 |
| Arthropoda | 2.42 | 2.67 | 2.92 |
| Brachiopoda | 1.88 | 2.02 | 2.16 |
| Bryozoa | 1.88 | 2.17 | 2.46 |
| Chordata | 0.95 | 1.20 | 1.46 |
| Cnidaria | 1.41 | 1.60 | 1.80 |
| Echinodermata | 0.94 | 1.23 | 1.52 |
| Hemichordata | 2.31 | 4.47 | 6.63 |
| Mollusca | 1.32 | 1.43 | 1.52 |
| Porifera | 1.33 | 1.63 | 1.93 |
| Stage 14:29 | 3.25 | 3.43 | 3.60 |
| Stage 30:45 | 0.33 | 0.49 | 0.66 |
| Stage 46:61 | 0.64 | 0.78 | 0.91 |
| Stage 62:77 | 1.74 | 1.95 | 2.17 |
| Stage 78:94 | 1.10 | 1.22 | 1.34 |

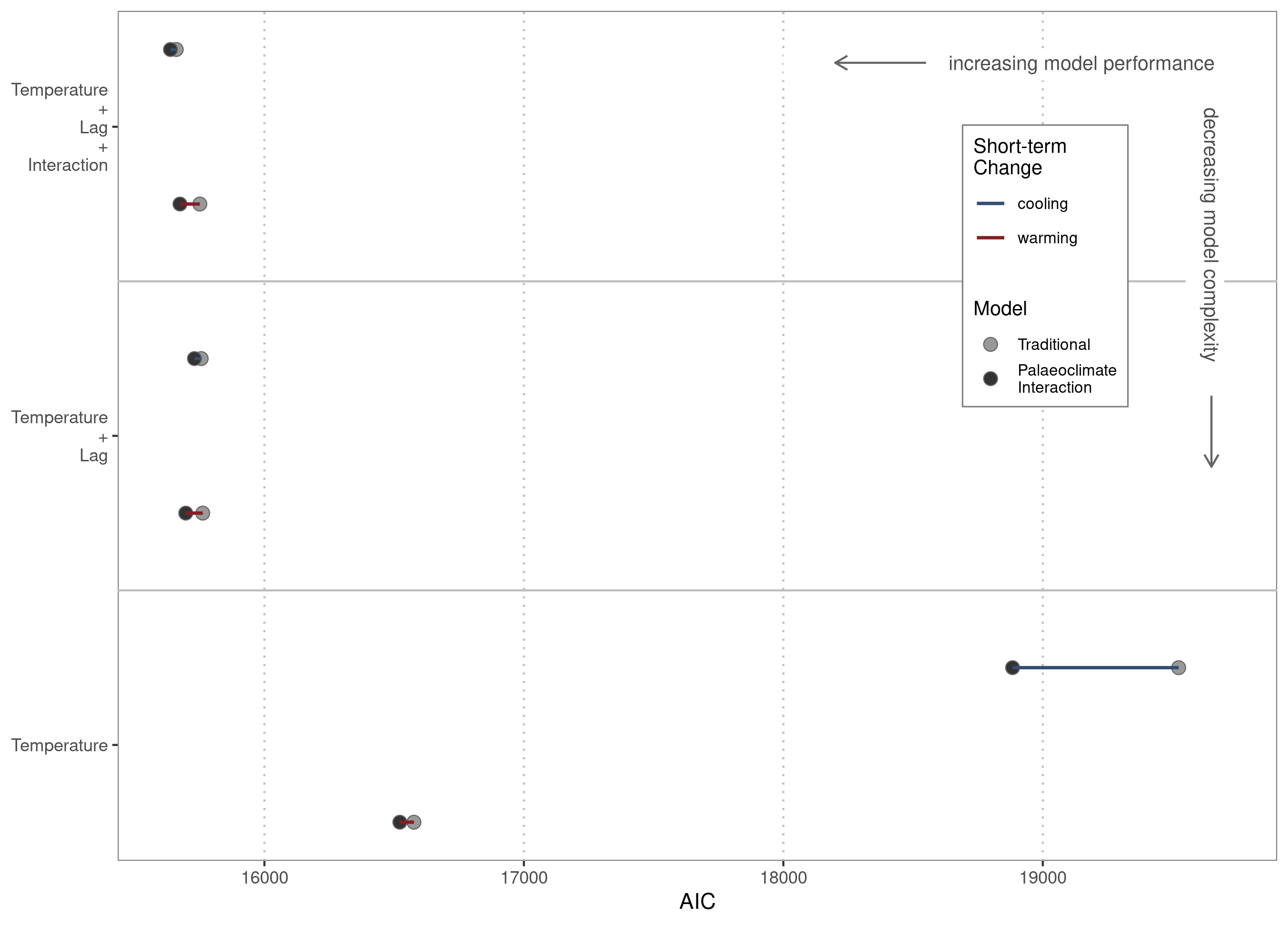
Suppl. Table 4 | Origination probability for all major phyla and throughout time after cooling-cooling compared to all other palaeoclimate interactions calculated as log odds ratio, including 95% confidence intervals. Stage 14 is the oldest stage included in analysis, and stage 94 the youngest.

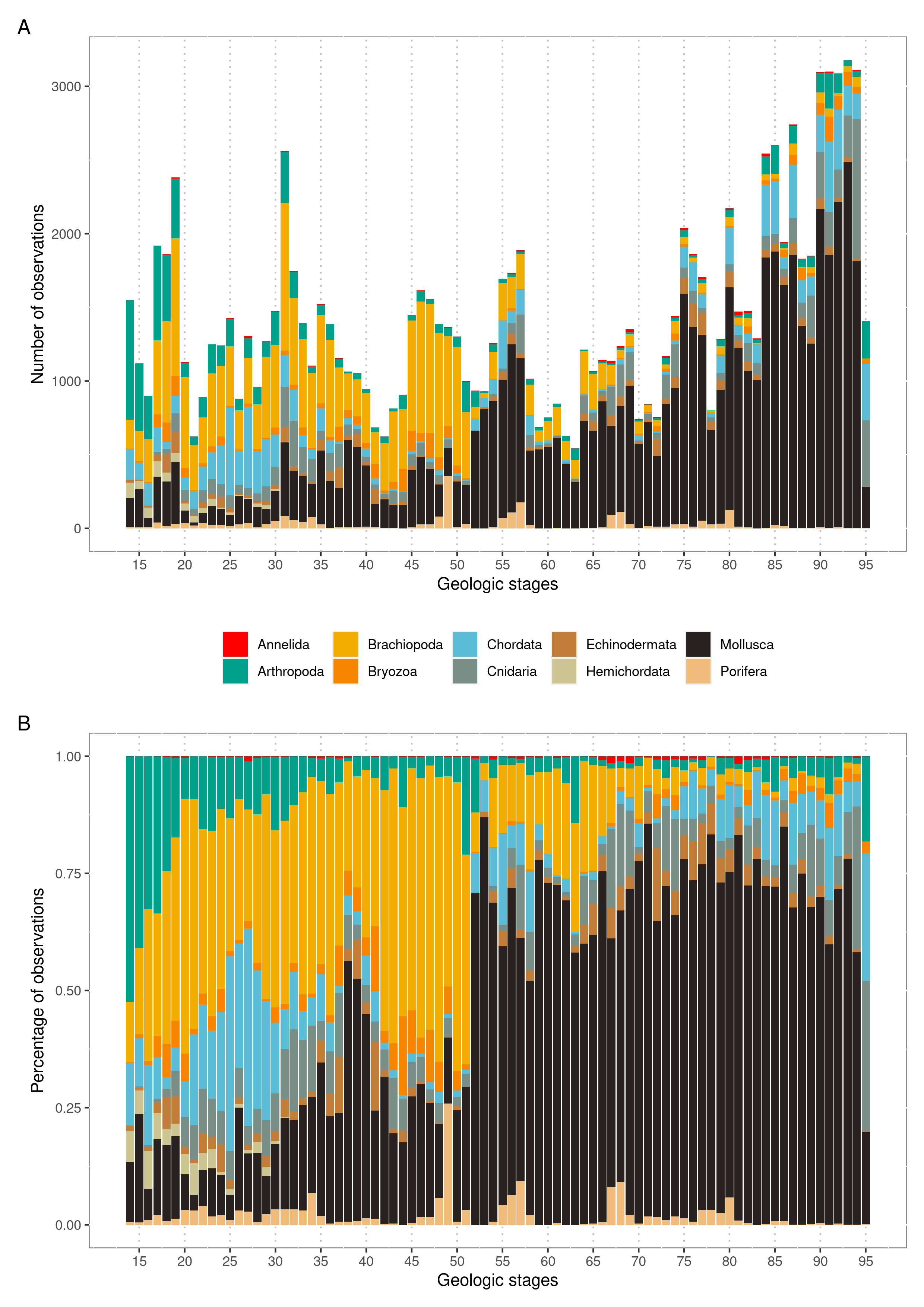
| **Phylum** | **Class** | **Order** | **Family** | **Genus** |
| --- | --- | --- | --- | --- |
| Annelida | 1 | 2 | 4 | 14 |
| Arthropoda | 5 | 23 | 164 | 615 |
| Brachiopoda | 5 | 21 | 293 | 1286 |
| Bryozoa | 2 | 15 | 127 | 279 |
| Chordata | 8 | 101 | 253 | 495 |
| Cnidaria | 3 | 15 | 185 | 571 |
| Echinodermata | 11 | 52 | 143 | 262 |
| Foraminifera | 3 | 6 | 23 | 81 |
| Hemichordata | 1 | 3 | 6 | 16 |
| Mollusca | 8 | 82 | 629 | 2463 |
| Nematoda | 1 | 1 | 1 | 1 |
| Porifera | 5 | 33 | 109 | 242 |

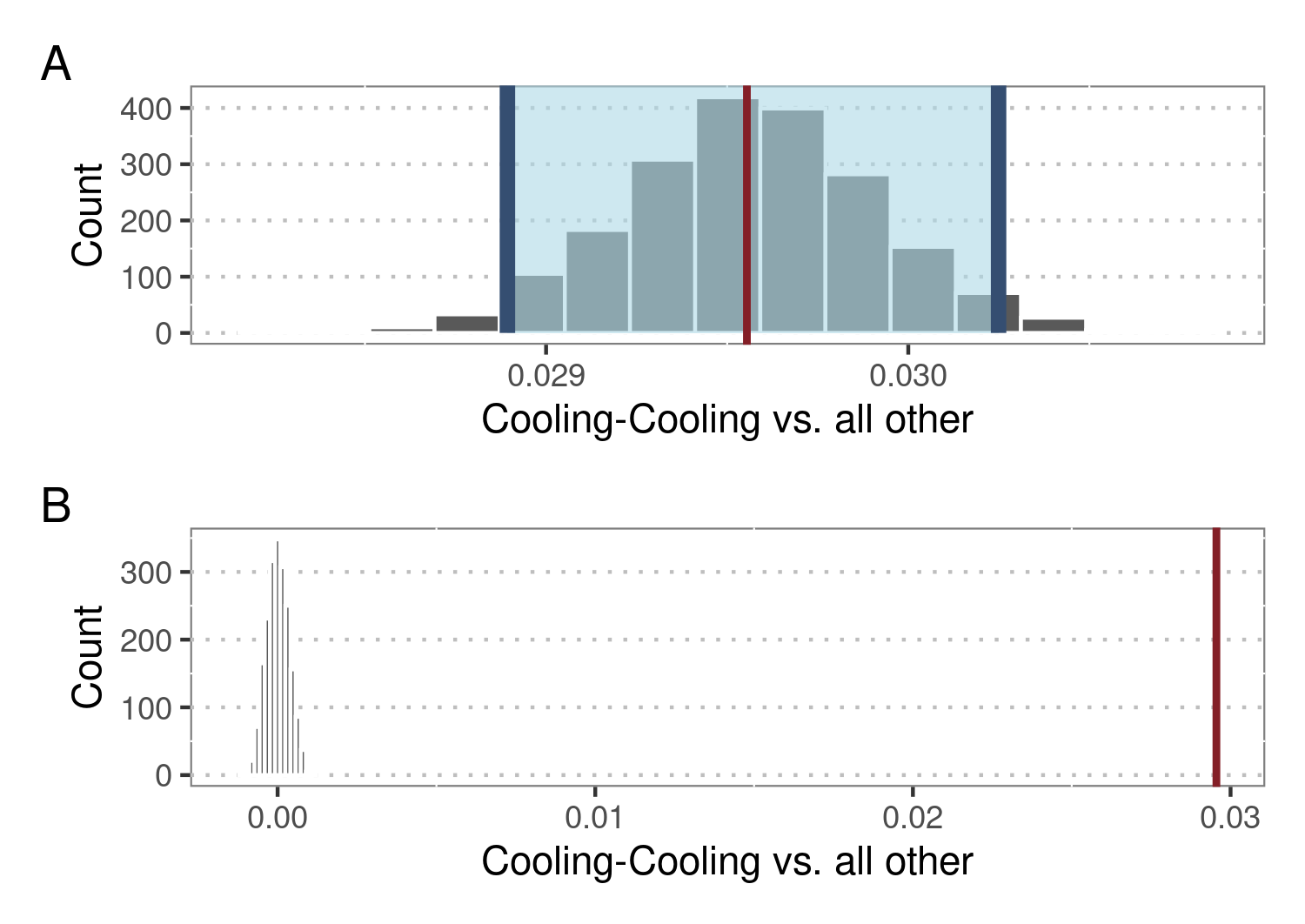
Suppl. Table 5 | Number of classes, order, families, and genera within every major phylum after data-cleaning and sampling-standardisation.

| **Type** | **Model** | **AIC** | **BIC** | **ΔAIC** | **ΔBIC** |
| --- | --- | --- | --- | --- | --- |
| Warming | trend.st1 | 16628.5 | 16652.5 | 42.7 | 42.7 |
| trend.st2 | 16621.0 | 16645.0 | 35.2 | 35.2 |
| trend.st3 | 16615.1 | 16639.1 | 29.3 | 29.3 |
| trend.st4 | 16620.0 | 16644.0 | 34.2 | 34.2 |
| trend.st5 | 16599.5 | 16623.4 | 13.7 | 13.6 |
| trend.st6 | 16585.8 | 16609.8 | 0.0 | 0.0 |
| trend.st7 | 16586.9 | 16610.9 | 1.1 | 1.1 |
| trend.st8 | 16594.9 | 16618.9 | 9.1 | 9.1 |
| trend.st9 | 16625.4 | 16649.3 | 39.6 | 39.5 |
| trend.st10 | 16628.9 | 16652.8 | 43.1 | 43.0 |
| Cooling | trend.st1 | 19159.7 | 19184.0 | 159.5 | 159.5 |
| trend.st2 | 19160.5 | 19184.7 | 160.3 | 160.2 |
| trend.st3 | 19120.0 | 19144.2 | 119.8 | 119.7 |
| trend.st4 | 19000.2 | 19024.5 | 0.0 | 0.0 |
| trend.st5 | 19001.9 | 19026.2 | 1.7 | 1.7 |
| trend.st6 | 19003.2 | 19027.5 | 3.0 | 3.0 |
| trend.st7 | 19037.4 | 19061.7 | 37.2 | 37.2 |
| trend.st8 | 19060.5 | 19084.8 | 60.3 | 60.3 |
| trend.st9 | 19072.3 | 19096.6 | 72.1 | 72.1 |
| trend.st10 | 19080.0 | 19104.3 | 79.8 | 79.8 |

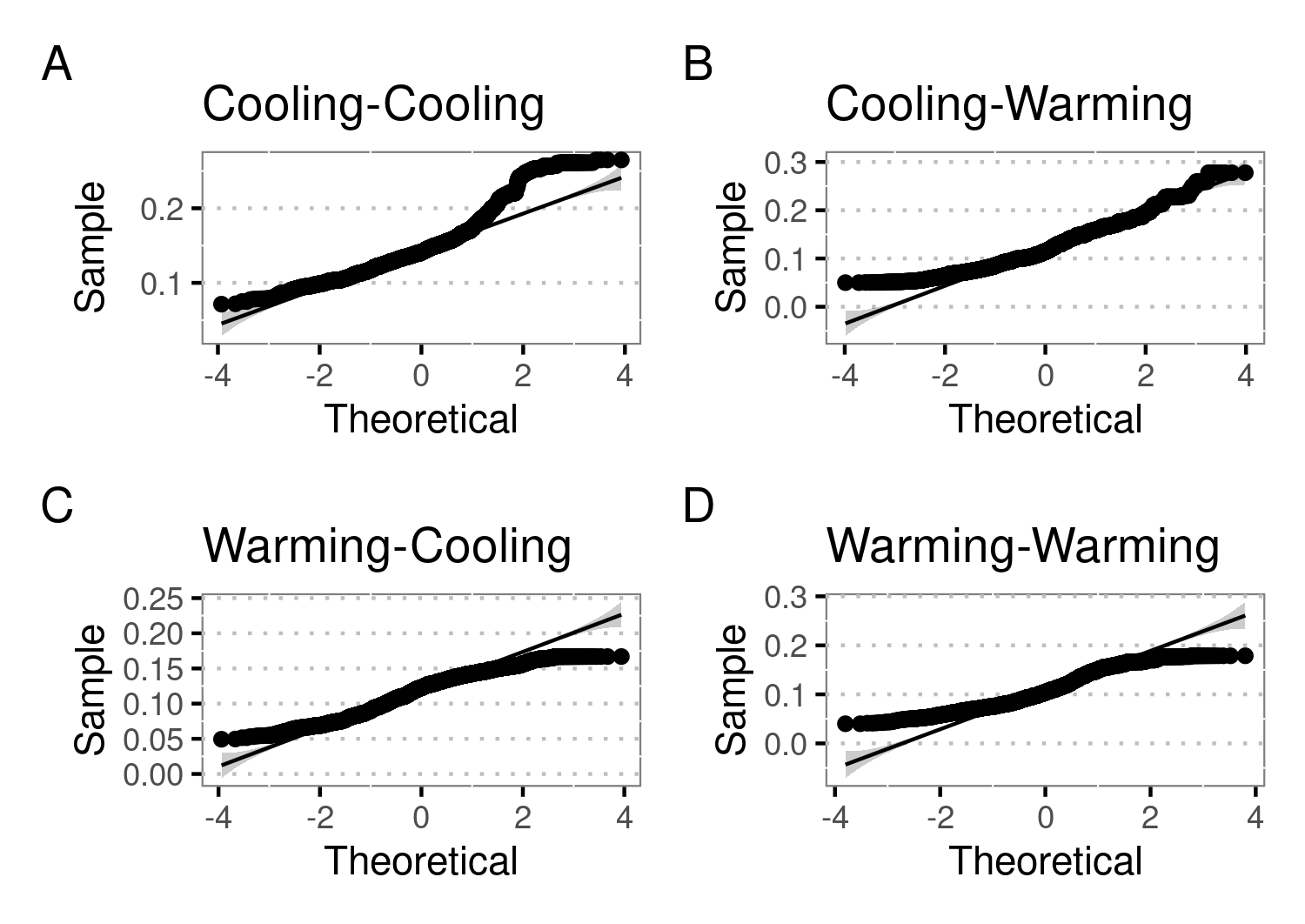
Suppl. Table 6 | Selection of the final model using a dynamic model framework. 10 GLMMs for each ∆Ttrend were calculated while keeping ∆Tchange fixed. The final model for both short-term warming and cooling was then selected using the AIC of each model.

Suppl. Figure 1 | Model comparison per group. 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.

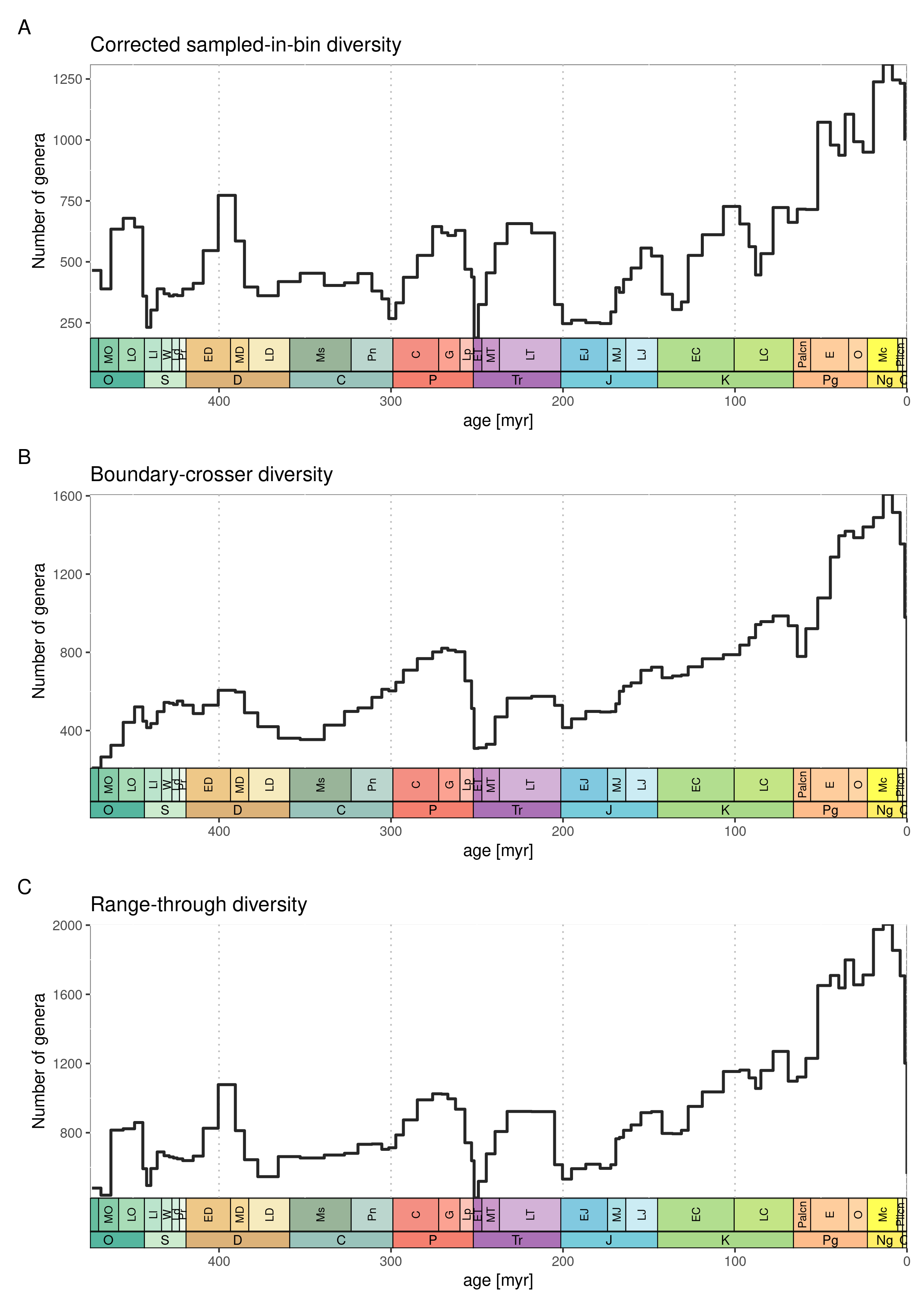
Suppl. Figure 2 | A) Total number of observations per phyla and stage. B) Percentage of total number of observations per phyla and stage. Stage 14 is the oldest, and stage 95 the youngest period.

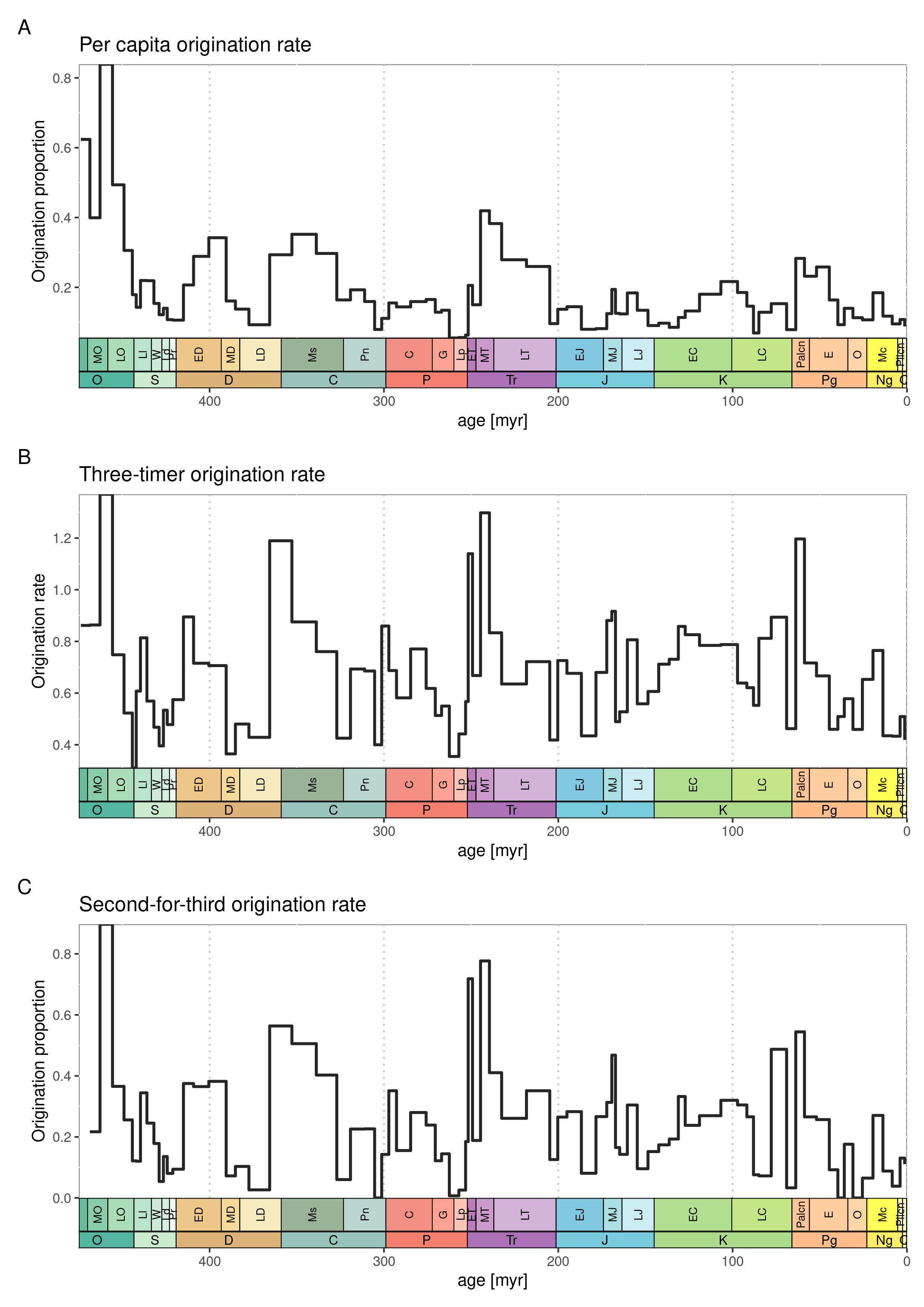


Suppl. Figure 3 | Difference in means of origination responses after cooling-cooling compared to all other palaeoclimate interactions, based on bootstrapping resampling. A) Bootstrapped distribution of differences in means. Red line shows observed difference and shaded area shows 95% confidence interval. B) Simulation-based null distribution of difference in means compared to observed difference (red line).



Suppl. Figure 4 | 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.

Suppl. Figure 5 | 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.

Suppl. Figure 6 | 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 proportions based on second-for-third substitution of taxa categories.