Supplementary Material for

Uneven spatial sampling distorts reconstructions of Phanerozoic seawater temperature

Lewis A. Jones[1] and Kilian Eichenseer[2]

[1]*Centro de Investigación Mariña, Grupo de Ecoloxía Animal, Universidade de Vigo, Spain*
LewisA.Jones@outlook.com; LewisAlan.Jones@uvigo.es
[2]*GeoZentrum Nordbayern, Friedrich-Alexander Universität Erlangen-Nürnberg, Germany*
Kilian.Eichenseer@fau.de

SUPPLEMENTARY METHODS

Data preparation

The oxygen isotope compilation of Veizer and Prokoph (2015) contains 22,280 δ^{18}O measurements from fossil calcareous shells within shallow marine depths (< 300 m). The original dataset does not provide the geographic coordinates of where fossils, and respective measurements, originate. However, using the supplementary information of δ^{18}O measurements (location, literature references, *etc*.) within the dataset, we assigned geographic coordinates to the data. The majority of the data—12,688 measurements (57%)—could be ascribed to a total of 97 ocean drilling program (DSDP, ODP and IODP) sites. A further 7,396 measurements (33%) were assigned to terrestrial localities, of which: (1) 1,181 measurements had unambiguous location names, for which coordinates could be assigned via geocoding in Google Earth; (2) 2,968 measurements were looked up manually, using the location names; (3) 655 measurements were looked up in their original publications; (4) 4,105 measurements were from publications which included multiple localities in a geographically restricted region—the midpoints of these regions were assigned to the corresponding measurements; and (5) 1,455 measurements had only country names as localities and were assigned coordinates equivalent to the centroid of the corresponding country (Fig. S2). Lastly, 2,196 measurements (10%) were discarded because neither their location information, nor the original publication enabled us to assign a locality. The final dataset contains 20,093 δ^{18}O samples.

Latitudinal temperature gradient models

We generated two latitudinal temperature gradient models for our analyses to represent different climate states throughout Earth history: (1) a 'Modern-type' steep latitudinal temperature gradient and (2) an 'Eocene-type' flattened latitudinal temperature gradient (Fig. S3). To generate these models, we used generalized additive models to estimate the relationship between temperature (y) and absolute (non-negative) latitude (palaeolatitude for the Eocene; x). These models were produced via the R package 'ggplot2' (Wickham, 2016), using the function stat_smooth(), and took the form:

$$(1) \qquad y = a + f(x) + \varepsilon,$$

with a denoting the intercept, f denoting the smooth function for x, which is estimated using restricted maximum likelihood, and ε the error term. To construct the Modern-type latitudinal temperature gradient model, we used a global grid of present-day mean annual sea surface temperature, downloaded from Bio-ORACLE (Tyberghein et al., 2012). This global grid is provided at a horizontal resolution of ~0.08° x 0.08° (approximately 9.2 km x 9.2 km at the equator). For the purposes of this study, we resampled this grid to a resolution of 1° x 1° and calculated the mean sea surface temperature values for each 1° latitudinal band. For the Eocene-type latitudinal temperature gradient model, a compilation of Late Palaeocene–Early Eocene temperature proxy measurements from the published literature was utilised (Zhang et al., 2019).

Sensitivity analyses

Exclusion of poorly constrained data
As 1,455 of our samples (~7%) were assigned centroid coordinates of their corresponding country (due to lack of more specific information), we evaluated the impact of excluding these samples. In general, we found that results were largely confirmatory (Fig. S4–S5). The most notable difference is the incompleteness of the time series relative to the one constructed with the full dataset, particularly in older intervals. While these results provide a more constrained view than those presented in the main text, the exclusion of data in this specific study would lead to a potentially false exacerbation of the issue addressed, resulting in overly negative results.

Alternative binning scheme
In our primary analysis, we followed a stratigraphic stage binning protocol. However, as stages are temporally uneven, we also repeated our analyses using equal-length time bins of 10 million years (myr) in duration. We assigned the $\delta^{18}O$ samples to these bins using the age estimates (GTS2012) provided in the Veizer and Prokoph (2015) data set. Confirming expectation, long-term trends in the results generally match those obtained at stage-level resolution (Figs. S6–S7). Due to the coarser resolution and the binning across major stratigraphic boundaries, short-term excursions in sampling and temperature are dampened with the 10-myr binning scheme. As most fossils are lacking precise, absolute age estimates, and instead are assigned to stratigraphic intervals with a range of possible absolute ages, we opted to focus our analyses on a binning scheme based on stratigraphic intervals (i.e. stages).

Study limitations
Our study is not without limitations. Firstly, the results presented in this work are sensitive to the type of latitudinal temperature gradient used in analyses. While we used two different types of gradient to

constrain this sensitivity, the shape and strength of the latitudinal temperature gradient has likely varied considerably throughout Earth history (Zhang et al., 2019). The utilised gradients in this study represent two climatic extremes, a steep latitudinal temperature gradient (i.e. icehouse), and a flattened latitudinal temperature gradient (i.e. greenhouse). Given that our results agree for almost all stages (84 out of 88) regarding the direction of temperature bias, a high-level of confidence can be afforded to our findings. Although the actual magnitude of temperature offset is more difficult to constrain due to variation in the latitudinal temperature gradients through time, our analysis provides an estimate of the potential range of temperature offset. Secondly, our approach is simple in that it does not directly account for longitudinal differences in temperature during the extraction process. While disregarding this variation may impact upon point estimates, longitudinal temperature differences are generally small in comparison to latitudinal differences, and therefore the overall findings are likely to remain unchanged. Furthermore, due to unspecific geographic information provided for some data ($n = 1,455$) from the Phanerozoic $\delta^{18}O$ compilation (Veizer and Prokoph, 2015), centroid coordinates within the bounds of the available geographic information (i.e. country names) had to be generated. Whilst this may have influenced the results presented here, we opted for this approach as discarding this data would have led to an underrepresentation of sampling, particularly for older time intervals which lack data from ocean drilling programs. Finally, due to smaller latitudinal variation in temperature, palaeotemperature reconstructions based on deep-water samples (not included in this study) may be less significantly influenced by variation in spatial sampling. However, few deep-water samples are preserved prior to the Albian due to subduction processes (Veizer and Prokoph, 2015).

SUPPLEMENTARY REFERENCES

Gradstein, F.M., Ogg, J.G., Schmitz, M., and Ogg, G., 2012, The Geologic Time Scale 2012 - 1st Edition: United States of America, Elsevier, https://www.elsevier.com/books/the-geologic-time-scale-2012/gradstein/978-0-444-59425-9 (accessed January 2020).

Tyberghein, L., Verbruggen, H., Pauly, K., Troupin, C., Mineur, F., and Clerck, O.D., 2012, Bio-ORACLE: a global environmental dataset for marine species distribution modelling: Global Ecology and Biogeography, v. 21, p. 272–281, doi:10.1111/j.1466-8238.2011.00656.x.

Veizer, J., and Prokoph, A., 2015, Temperatures and oxygen isotopic composition of Phanerozoic oceans: Earth-Science Reviews, v. 146, p. 92–104, doi:10.1016/j.earscirev.2015.03.008.

Wickham, H., 2016, ggplot2: Elegant Graphics for Data Analysis: Houston, Texas, USA, Springer, Use R!

Zhang, L., Hay, W.W., Wang, C., and Gu, X., 2019, The evolution of latitudinal temperature gradients from the latest Cretaceous through the Present: Earth-Science Reviews, v. 189, p. 147–158, doi:10.1016/j.earscirev.2019.01.025.

SUPPLEMENTARY TABLES

Table S1. Temporal bins used in this study for binning $\delta^{18}O$ samples. Binning scheme is based on GTS2012 (Gradstein et al., 2012), with Pleistocene stages amalgamated.

Stage	Age (Ma)		Stage	Age (Ma)	
	Maximum	Minimum		Maximum	Minimum
Holocene	0.0117	0	Changhsingian	254.17	252.17
Pleistocene	2.588	0.0117	Wuchiapingian	259.9	254.17
Piacenzian	3.6	2.588	Capitanian	265.1	259.9
Zanclean	5.333	3.6	Wordian	268.8	265.1
Messinian	7.246	5.333	Roadian	272.3	268.8
Tortonian	11.62	7.246	Kungurian	279.3	272.3
Serravallian	13.82	11.62	Artinskian	290.1	279.3
Langhian	15.97	13.82	Sakmarian	295.5	290.1
Burdigalian	20.44	15.97	Asselian	298.9	295.5
Aquitanian	23.03	20.44	Gzhelian	303.7	298.9
Chattian	28.1	23.03	Kasimovian	307	303.7
Rupelian	33.9	28.1	Moscovian	315.2	307
Priabonian	38	33.9	Bashkirian	323.2	315.2
Bartonian	41.3	38	Serpukhovian	330.9	323.2
Lutetian	47.8	41.3	Visean	346.7	330.9
Ypresian	56	47.8	Tournaisian	358.9	346.7
Thanetian	59.2	56	Famennian	372.2	358.9
Selandian	61.6	59.2	Frasnian	382.7	372.2
Danian	66	61.6	Givetian	387.7	382.7
Maastrichtian	72.1	66	Eifelian	393.3	387.7
Campanian	83.6	72.1	Emsian	407.6	393.3
Santonian	86.3	83.6	Pragian	410.8	407.6
Coniacian	89.8	86.3	Lochkovian	419.2	410.8
Turonian	93.9	89.8	Pridoli	423	419.2
Cenomanian	100.5	93.9	Ludfordian	425.6	423
Albian	113	100.5	Gorstian	427.4	425.6
Aptian	125	113	Homerian	430.5	427.4
Barremian	129.4	125	Sheinwoodian	433.4	430.5
Hauterivian	132.9	129.4	Telychian	438.5	433.4
Valanginian	139.8	132.9	Aeronian	440.8	438.5
Berriasian	145	139.8	Rhuddanian	443.4	440.8
Tithonian	152.1	145	Hirnantian	445.2	443.4
Kimmeridgian	157.3	152.1	Katian	453	445.2
Oxfordian	163.5	157.3	Sandbian	458.4	453
Callovian	166.1	163.5	Darriwilian	467.3	458.4
Bathonian	168.3	166.1	Dapingian	470	467.3
Bajocian	170.3	168.3	Floian	477.7	470
Aalenian	174.1	170.3	Tremadocian	485.4	477.7
Toarcian	182.7	174.1	Stage 10	489.5	485.4
Pliensbachian	190.8	182.7	Jiangshanian	494	489.5
Sinemurian	199.3	190.8	Paibian	497	494
Hettangian	201.3	199.3	Guzhangian	500.5	497
Rhaetian	208.5	201.3	Drumian	504.5	500.5
Norian	228	208.5	Stage 5	509	504.5
Carnian	237	228	Stage 4	514	509
Ladinian	242	237	Stage 3	521	514
Anisian	247.2	242	Stage 2	529	521
Olenekian	251.2	247.2	Fortunian	541	529
Induan	252.17	251.2			

SUPPLEMENTARY FIGURES
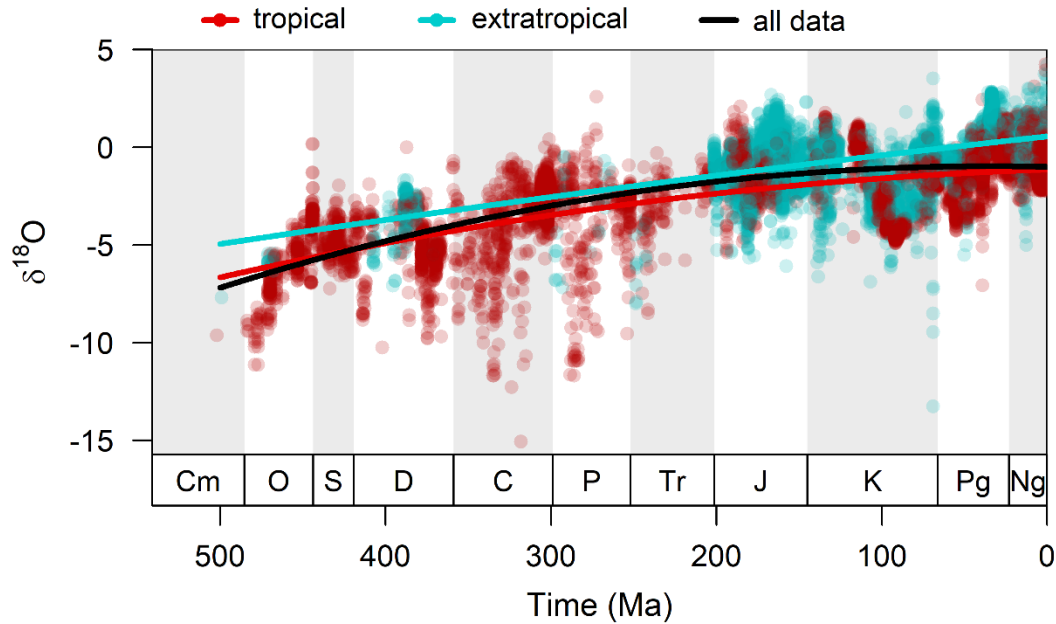

Figure S1. Phanerozoic secular trends in the $\delta^{18}O$ data. Quadratic polynomial regressions of $\delta^{18}O$ against time are shown for the tropical–subtropical data ($\leq 35°$ absolute palaeolatitude, red), for the extratropical data ($\geq 35°$ absolute palaeolatitude, turquoise), and across all data (black). The resulting regression lines are similar for the tropical–subtropical and the extratropical data, except that the predicted $\delta^{18}O$ values are higher (i.e. cooler) for the extratropical data, and have a less dominant quadratic component. The trend across all data has a more dominant quadratic component, and is less reliable due to being strongly influenced by changes in latitudinal sampling (see Fig. 3). For the conversion of $\delta^{18}O$ to temperature, we used the trend derived by Veizer and Prokoph (2015), based on tropical–subtropical data, due to the better sampling in the tropics in most of the record.
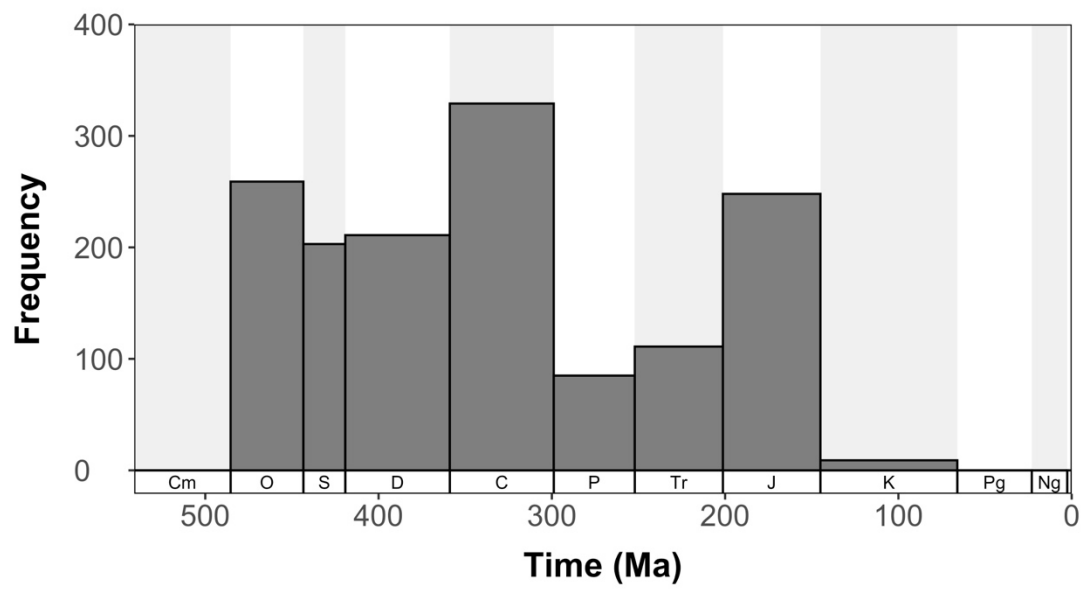
Figure S2. Temporal distribution of samples with only country names provided as locality information. These samples were assigned the centroid coordinates of their respective country of origin.
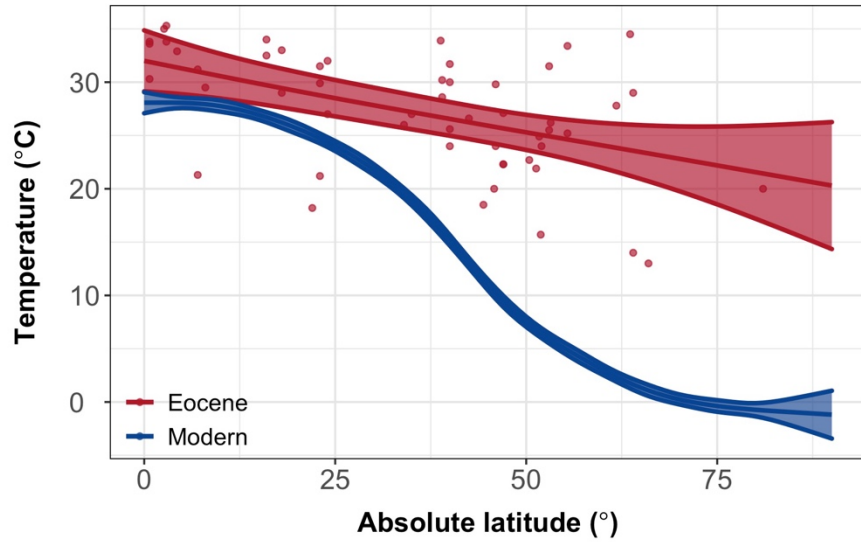
Figure S3. Modelled 'Modern-type' and 'Eocene-type' latitudinal temperature gradients. Models were produced using a generalized additive model, with temperature regressed against latitude. The Modern (blue line) latitudinal temperature gradient model was produced using a global grid of present-day mean annual sea surface temperature, downloaded from Bio-ORACLE (Tyberghein et al., 2012). The Eocene model (red line) was produced using a late Palaeocene–early Eocene compilation of temperature proxy measurements (red points) from the published literature (Zhang et al., 2019). Respective coloured shading depicts the 95% confidence interval of the models.
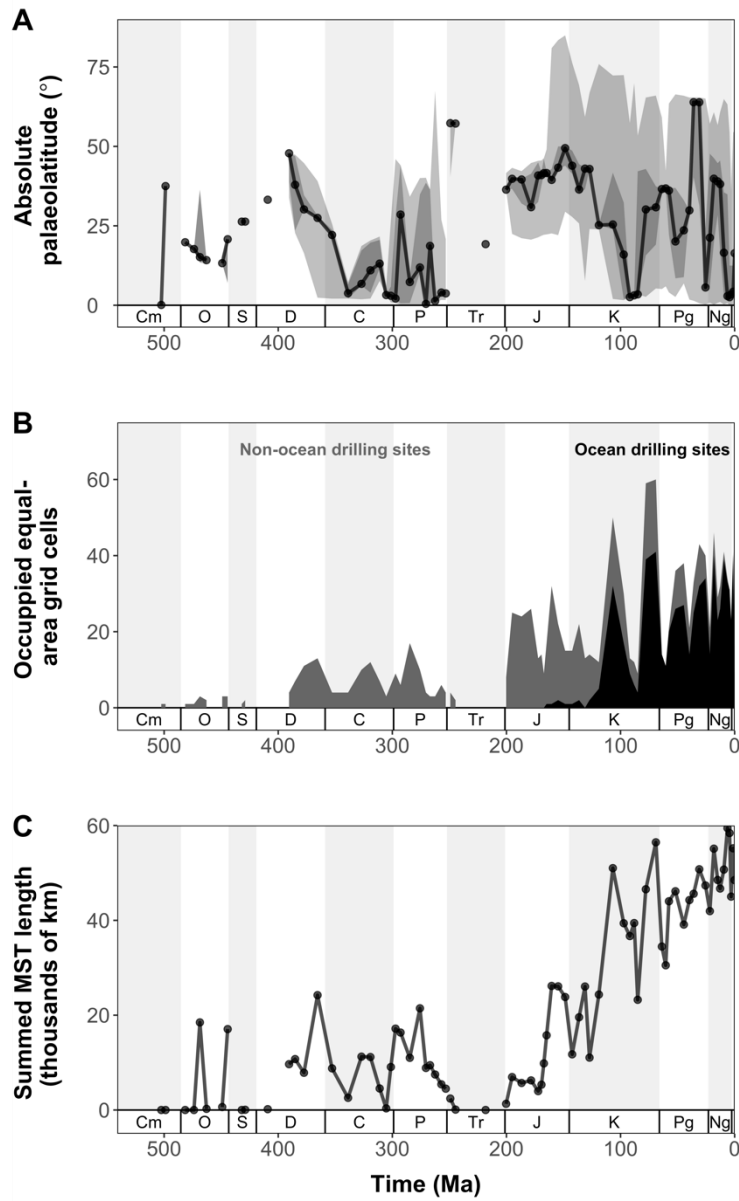
Figure S4. Stage-level spatial sampling metrics of Phanerozoic δ^{18}O samples (excluding samples with non-specific locality information). (A) Absolute palaeolatitudinal median (black line), interquartile range (dark grey ribbon) and maximum/minimum palaeolatitudes (light grey ribbon) of δ^{18}O samples. (B) The number of occupied equal-area grid cells of δ^{18}O samples. Cells containing non-ocean drilling site δ^{18}O samples (marine fossils collected on land) are depicted in grey, while ocean drilling sites are depicted in black. (C) The summed minimum spanning tree (MST) length of δ^{18}O samples, i.e. the total, minimal length of segments connecting all their palaeocoordinates.
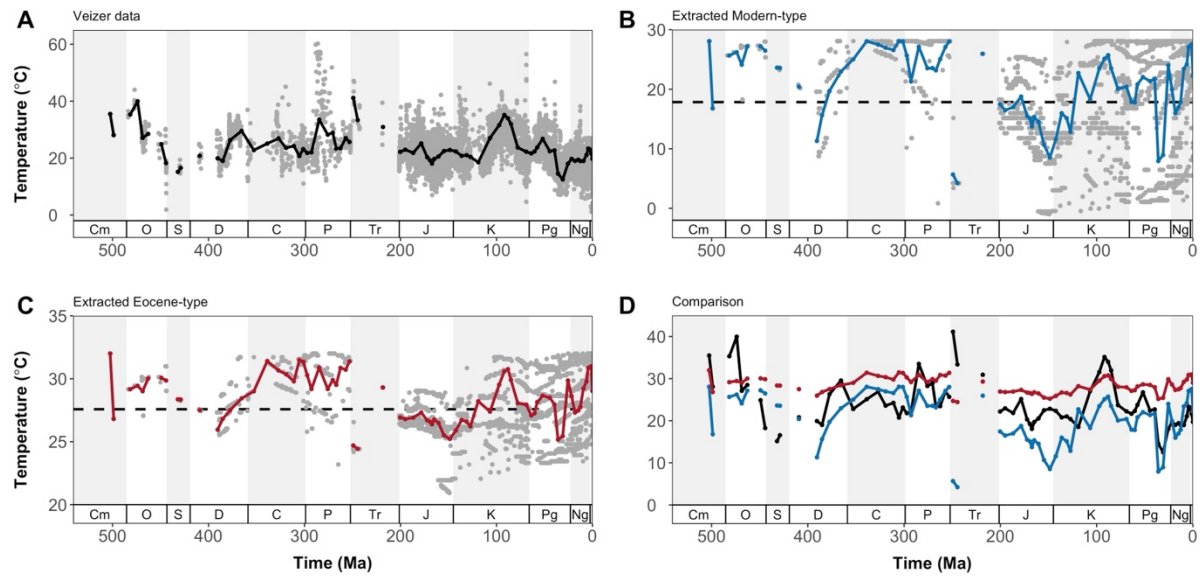
Figure S5. Stage-level Phanerozoic temperature reconstructions (excluding samples with non-specific locality information). (A) Mean stage-level temperatures (black line) calculated from the oxygen isotope compilation of Veizer and Prokoph (2015). (B) Mean stage-level extracted temperatures (blue line) from the Modern latitudinal temperature gradient model, and Modern global mean temperature (black dashed line). (C) Mean stage-level extracted temperatures (red line) from the Eocene latitudinal temperature gradient model, and Eocene global mean temperature (black dashed line). (D) Comparison plot of stage-level global temperature estimates for the Veizer dataset (black line), extracted Modern (blue line) and extracted Eocene (red line). (A–C) Individual datapoints are plotted in grey.
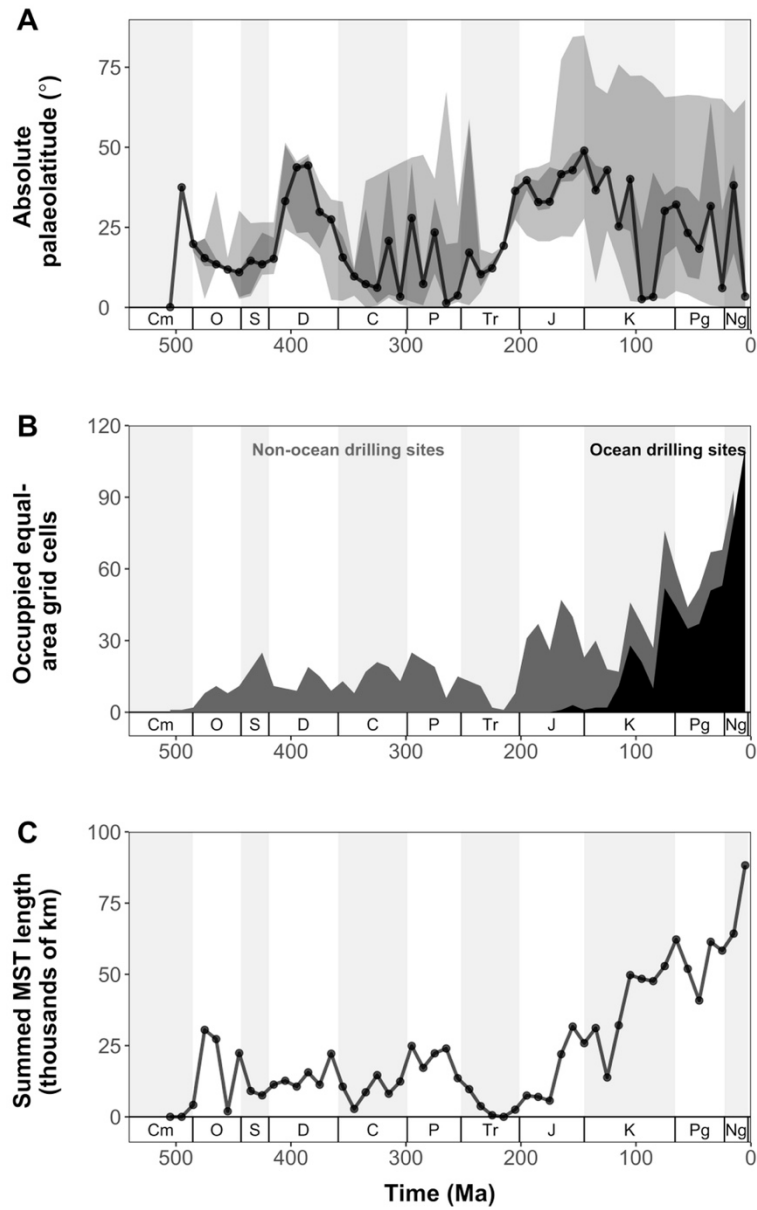
9

Figure S6. Spatial sampling metrics (within 10-million-year time bins) of Phanerozoic $\delta^{18}O$ samples. (A) Absolute palaeolatitudinal centroid (black line), interquartile range (dark grey ribbon) and maximum/minimum palaeolatitudes (light grey ribbon) of $\delta^{18}O$ samples. (B) The number of occupied equal-area grid cells of $\delta^{18}O$ samples. Cells containing non-ocean drilling site $\delta^{18}O$ samples (marine fossils collected on land) are depicted in grey, while ocean drilling sites are depicted in black. (C) The summed minimum spanning tree (MST) length of $\delta^{18}O$ samples, i.e. the total, minimal length of segments connecting all their palaeocoordinates.
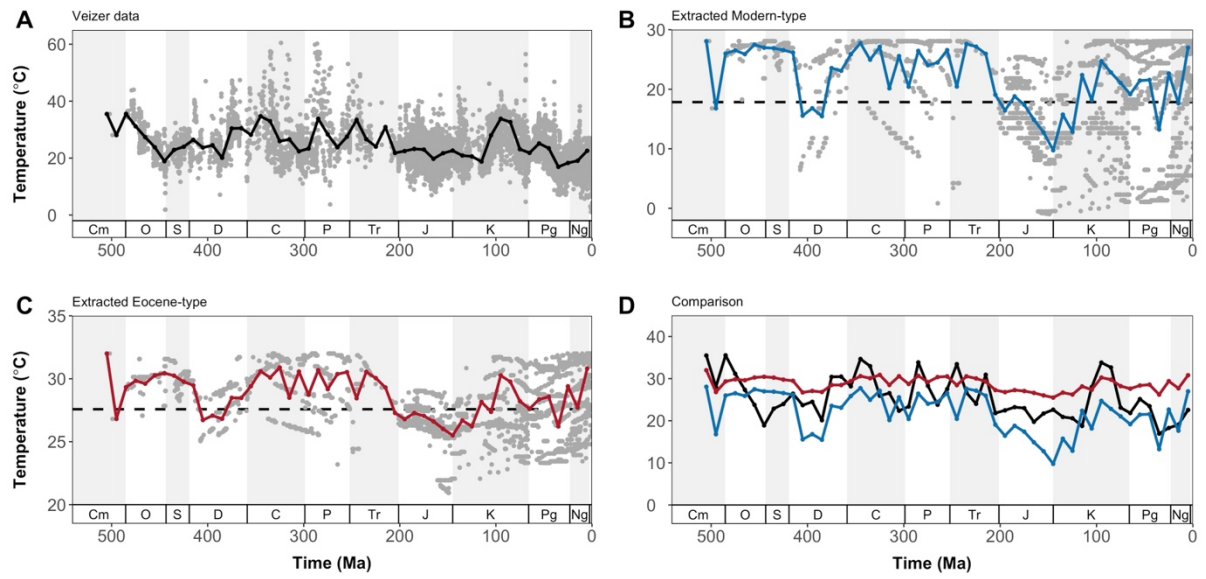
Figure S7. Phanerozoic temperature reconstructions (within 10-million-year time bins). (A) Mean temperatures (black line) calculated from the oxygen isotope compilation of Veizer and Prokoph (2015). (B) Mean extracted temperatures (blue line) from the Modern latitudinal temperature gradient model, and Modern global mean temperature (black dashed line). (C) Mean extracted temperatures (red line) from the Eocene latitudinal temperature gradient model, and Eocene global mean temperature (black dashed line). (D) Comparison plot of global temperature estimates for the Veizer dataset (black line), extracted Modern (blue line) and extracted Eocene (red line). (A–C) Individual datapoints are plotted in grey.
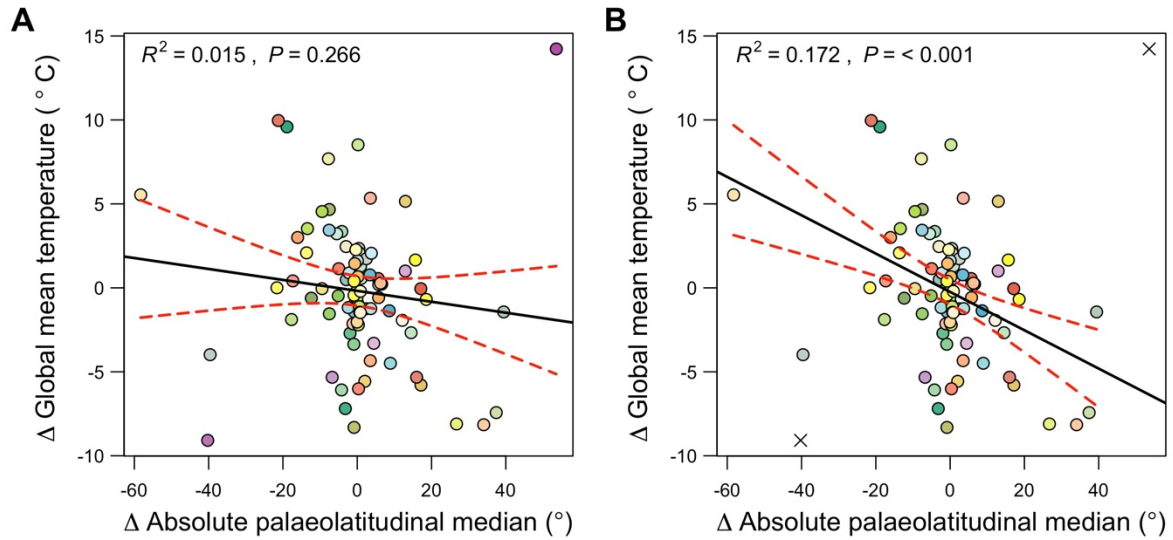
Figure S8. First differences of stage-level, absolute (non-negative) palaeolatitudinal median versus first differences of global mean δ^{18}O temperatures. A) Linear regression including all stages does not suggest a relationship between the two variables ($R^2 = 0.015$; $P = 0.266$). B) Linear regression, with the Olenekian outlier excluded, results in a significant negative relationship ($R^2 = 0.172$; $P < 0.001$). The Olenekian data was removed after differencing, hence both differences involving the Olenekian (crosses) do not appear in analysis. Data points are coloured by stage in accordance with the geological timescale of Gradstein et al (2012).
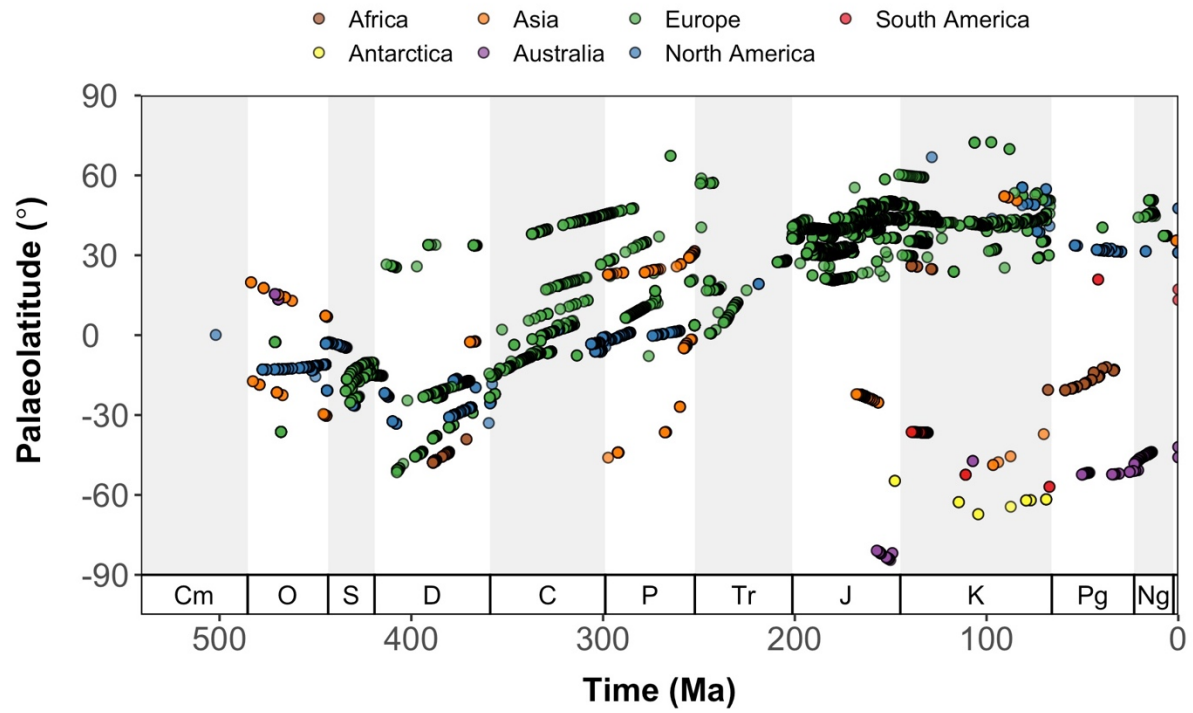
Figure S9. Palaeolatitudinal plot of Phanerozoic non-ocean drilling (marine fossils collected on land) $\delta^{18}O$ samples. Each point is coloured by its respective continent of origin (based on present-day coordinates). Note the preferential sampling of Europe across the entire time series, as well as the palaeolatitudinal shift of sampling.