Occupancy history influences extinction risk of fossil marine microplankton groups

Isaiah E. Smith*, Ádám T. Kocsis, Wolfgang Kiessling

Department of Geography and Geosciences, GeoZentrum Nordbayern, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany

* Corresponding author

Correspondence email: isaiah.em.smith@gmail.com

Formatted: Font: (Default) Times New Roman, Complex Script Font: Times New Roman

Formatted: Index, Indent: Before: 0 cm, First line: 0 cm, Line spacing: single

Abstract

Geographic range has long been acknowledged as an important determinant of extinction risk. The trajectory of geographic range through time, however, has not received as much scientific attention. Here, we test the role of change in geographic range – assessed by a measure of proportional occupancy of grid cells – in determining the extinction risk in four major microplankton groups over the last 66 million years: foraminifera, calcareous nannofossils, radiolarians, and diatoms. Logistic regression was used to assess the importance of standing occupancy and~~,~~ occupancy change ~~, and sampling probability~~ in the extinction risk of species. We find that while standing occupancy is a major determinant of extinction risk in all microplankton groups, the change in occupancy accounts for an average of ~~52%~~41% ~~of~~ of the explanatory power ~~of~~ shared by the ~~three~~ two analyzed variables, with a maximum value of 77% ~~92%~~. ~~Sampling probability was also found to be consistently informative, with an average of 6% and a maximum value of 22%.~~ We also find that as temporal resolution decreases, the predictive ability of these variables increases. Our results highlight the importance of incorporating both geographic range and its change through time~~, as well as sampling probability,~~ into extinction models. The ability of occupancy trajectory to help predict extinction risk underlines the necessity of paleontological data in modern conservation efforts.

Formatted: Space Before: 0 pt, After: 0 pt

1. Introduction

There is a rich literature documenting the effect of smaller geographic range size~~s~~ increasing risk in contemporary and ancient extinctions (e.g. Foote et al., 2016; Foote et al., 2007; McKinney, 1997; Payne and Finnegan, 2007; Purvis et al., 2000; Staude et al., 2020). The International Union for the Conservation of Nature (IUCN) uses geographic range size as one of the five key criteria by which the risk status of a species is assessed in the "Red List of

Threatened Species" ~~risk status of a species is assessed~~ (Mace et al., 2008). The temporal trajectory of geographic range as a predictor of global extinction has been previously ~~sparingly~~ explored in the paleontological literature (Liow et al., 2010; Foote et al., 2007, Tietje and Kiessling, 2013; Kiessling and Kocsis, 2016; Saulsbury et al., 2023). Here, -we further build upon this topic and explore how it applies to marine microplankton. ~~but many studies looking at historic extinctions exclude this factor altogether. Increasing anthropogenic impact on biotic systems has been suggested as a contributing factor to species' extinction risks (Ceballos et al., 2015; De Vos et al., 2015; Payne et al., 2016; Wagler, 2011). Understanding factors that increase a species' risk of disappearance, and how they can be mitigated, is thus more important now than ever before.~~

Based on a data set of Cenozoic marine invertebrates from the Paleobiology Database (https://paleobiodb.org/), Kiessling and Kocsis (2016), suggested that the trajectory of geographic range has the potential to inform extinction risk. However, the coarse stratigraphic resolution of the macroinvertebrate record (geological stages, about 5 million years in duration) puts constraints on the fidelity of any approach that depends on the spatiotemporal distribution of species. Due to their sheer abundance, unicellular groups are less affected by such issues and can be used for finely resolved studies of assemblage changes (e.g., Strack et al. 2024) and biogeography (e.g., Swain et al., 2024). ~~Variations in the material that make up their tests (calcareous or siliceous) are expected to manifest as either differential fossil sampling or as evolutionary patterns, which presents an opportunity to explore the sensitivity of extinction models under a wider range of scenarios. While accounting for sampling,~~ W~~w~~e assess here the importance of geographic range (expressed as proportional grid occupancy) and its temporal trajectory on extinction risk of marine planktonic organisms. By using a temporally finely-resolved dataset of fossil plankton, ~~as well as with the explicit incorporation of fossil sampling~~

~~into our models,~~ we can better assess ~~whether~~ the degree to which the trajectory of geographic occupancy ~~actually~~ influences extinction risk in marine life.

Formatted: Left

2. Data and Methods

2.1 Sourcing and cleaning of raw data

We downloaded occurrence records from the Neptune Sandbox Berlin ("NSB"; Lazarus, 1994; Renaudie et al., 2020; data downloaded 30 August 2023) using the R package “NSBcompanion” (Renaudie, 2019; version 2.2) as well as the Triton database (Fenton et al. 2021, version 2). Four taxonomic groups were downloaded: planktonic foraminifera, calcareous nannofossils, radiolarians, and diatoms. ~~We downloaded occurrence records of four major marine microplankton groups from the Neptune Sandbox Berlin ("NSB"; Lazarus, 1994; Renaudie et al., 2020; data downloaded 30 August 2023) using the R package “NSBcompanion” (Renaudie 2019). These groups were planktonic foraminifera, calcareous nannofossils, radiolarians, and diatoms.~~ Each of t~~T~~he four data sets were downloaded with the taxonomy resolved using the IODP Taxonomic Name List Project (Renaudie et al., 2020), a built-in option that we specified prior to downloading. Additionally, questionably-identified taxa were excluded from the download. Open-nomenclature taxa and possibly problematic or reworked occurrences were also excluded using the built-in NSB download options. The NSB holds taxon occurrences stretching back to the late Jurassic, but we limit our analysis to the Cenozoic record (i.e. the last 66 million years) to ensure a consistent age range for all for taxonomic groups, since both the diatom and the radiolarian NSB records only exist for the Cenozoic.

The NSB database includes the estimated age as well as the modern and estimated paleo-coordinates (longitude and latitude) of each fossil occurrence calculated internally based on the plate tectonic reconstruction by Boyden et al. (2011, Fig. S1). Each dataset (foraminifera, calcareous nannofossils, radiolarians, diatoms) was cleaned to remove any occurrences that were missing age, paleo-coordinate, and/or relevant taxonomic information. The counts of unique species and the number of occurrence records in each data set (post-cleaning) are provided in Table 1. All cleaning and subsequent analyses were carried out in R 4.4~~2~~.3~~2~~ (R Core Team, 2022).

Data Set Summaries				
	Foraminifera	Calcareous Nannofossils	Radiolarians	Diatoms
Number of unique species	552	737	840	966
Number of occurrences	~~96 410~~	~~154 052~~ 150346	~~91 240~~	~~91 092~~

	96426		90201	91194
~~Proportion~~ Percent of species that are extant and ~~species with usable records~~ have records in more than ~~(>~~1 time bin~~)~~	~~0.051~~ 6.1	~~0.015~~ 1.5	~~0.093~~ 11.3	~~0.119~~ 14.6
Number of unique species-bin pairings with “gap-omission” approach, excluding ~~first occurrences (occupancy change = NA)~~NA occurrences for occupancy and occupancy change. Bin size = 1 million years.	~~3 471~~ 2951	~~4 814~~ 3817	~~5 626~~ 4512	~~4 938~~ 4225

Table 1. The number of unique species, the number of total occurrence records, the proportion of species that are both extant and occur in more than one bin, and the number of species-bin pairings (post-cleaning).

We assigned occurrences from each data set to time bins of either 0.1, 0.2, 0.5, or 1.0 million years, noting that 0.1 million years is currently the lower limit for global correlation. For each time bin size, the first bin stretched from the present (0 Ma) to either 0.1, 0.2, 0.5, or 1.0 million years into the past. Each subsequent bin encompassed the following increment stretching progressively further into the past. ~~Ages are reported in millions of years before present (Ma).~~

We assessed stratigraphic ranges as defined by the oldest and youngest fossil occurrences. Due to reworking and other processes, the documented raw ranges may not reflect the true durations of species. Therefore, we also applied the recommended “Pacman profiling” (Lazarus et al., 2012), a stratigraphic outlier correction, to reduce the impact of outliers and reworking on the data. The degree of Pacman trimming on the NSB data was determined via a calibration process that used speciation and extinction ages of a given subset of each taxonomic group. Based on this subset, the degree of trimming necessary to restore the “true” temporal ranges of species could be estimated. Calibration ages were sourced from the Triton database (Fenton et al. 2021) for foraminifera, Nigrini et al. (2006, obtained from Lazarus et al., 2012) for radiolarians, the "Barron Diatom Catalog" (Lazarus et al., 2014) for diatoms, and a custom species list constructed from Mikrotax (www.mikrotax.org; Huber et al., 2017) for calcareous nannofossils. Potential trim values ranging from 0% to 16% of the raw ranges, at 1% intervals, were analyzed. Pacman calibration was carried out on datasets after they had been trimmed to the last 66 million years. Trim values were selected such that they minimized the average absolute difference between the actual and the represented speciation or extinction ages of the species present in the calibration set. The best-performing trim values were implemented in this study, although the key results presented here do not change in the absence of Pacman profiling. ~~Those trim values were as follows: foraminifera (top: 15%, bottom 3%), calcareous nannofossils (top: 14%, bottom: 4%), diatoms (top: 11%, bottom: 4%), radiolarians (top: 10%, bottom: 6%). Per capita extinction rates were calculated using the formula from Foote (1999).~~ Those trim values were as follows: foraminifera (top: 15%, bottom 3%), calcareous nannofossils (top: 15%, bottom: 5%), diatoms (top: 11%, bottom: 4%), radiolarians (top: 10%, bottom: 7%). Per capita extinction rates were calculated using the formula from Foote (1999), without normalizing for interval length.

2.2 Analysis of completeness

In order to quantify the degree to which sampling completeness affected downstream analyses, we employed two separate completeness metrics: the three-timer completeness metric (Alroy, 2008) and the simple completeness metric ("SCM"; Benton, 1985). The three-timer completeness metric is the ratio of "three-timer" taxa (those which occur in bin *i*-1, bin *i*, and bin *i*+1) to all taxa which occur in both bin *i*-1 and bin *i*+1 (irrespective of their presence in bin *i*, "part-timer" taxa). The three-timer metric ~~: the simple completeness metric ("SCM"; Benton, 1985) and the three-timer completeness metric (Alroy, 2008). The three-timer completeness metric~~ was calculated from the three-timer and part-timer counts returned by the "divDyn" R extension package (Kocsis et al., 2019; version 0.8.3). The simple completeness metric is the ratio of time bins with confirmed taxon occurrences to the inferred (by recorded observations before and after a focal time interval) number of time bins occupied by that taxon.

2.3 Calculating occupancy

For each data set, paleo-coordinates of samples were assigned to equal-area geographic cells using the R package "icosa" (Kocsis, 2020; version 0.11.1) for the calculation of proportional grid occupancy. Proportional grid occupancy is a recognized metric for assessing geographic range in the fossil record, where contemporaneous sampling is impossible and incomplete preservation is common (Foote et al., 2007; Darroch et al., 2022). Several cell-sizes were analyzed ranging in edge length between 3.33° to 2°. There was little variation in results within this range, so the highest resolution (4002 cells with 2° edge length, mean area of 1.3 x 10^5 km^2) was selected for this study. The present-day distribution of samples can be seen in supplementary Fig. S1.

As counts of occupied cells tend to be biased by sampling (Kiessling, 2005), we calculated proportional occupancy of each species in every time bin. Proportional occupancy is ~~simply~~ the number of geographic cells occupied by the species divided by the total number of sampled cells in a given time bin. For ~~clarity~~simplicity, we refer to what is actually proportional occupancy as ~~simply~~ *occupancy* from here forward. Furthermore, the number of unique Longhurst (2007) biogeographic planktonic provinces that were occupied by each species in each time bin was calculated, and the Pearson correlation of this value with the raw number of occupied geographic cells was calculated. Autocorrelation was accounted for by differencing temporally consecutive values prior to calculating correlation values~~s~~.

2.4 Change in occupancy

In addition to standing *occupancy*, the change in *occupancy* between consecutive time bins was calculated by taking the natural log of the ratio of *occupancy* in time bin *i* to *occupancy* in time bin *i*-1. The log transformation serves to standardize the magnitude of change and produces positive values for increases in *occupancy* (range expansions) and negative values for decreases in *occupancy* (range contractions). The correlation of the first differences between occupancy and occupancy change was computed to determine if the data was affected by multicollinearity.

Initially, instances where *occupancy* values in bin *i* or *i*-1 were 0 (no occurrences) were coded as missing data for *occupancy change* and removed from the final data set. While removing these records prevents the inclusion of undefined *occupancy change* values in the final data set, it greatly reduces the number of occurrences for a given taxon, especially for species whose sampling is fragmentary. This effect is magnified by the fact that for each time bin with zero occurrences of a given taxon (a “gap” in that taxon’s fossil record), two data points are removed from the final data set for that taxon. This overall loss of data becomes more pronounced with smaller bin sizes.

To combat this effect, we employed a “gap-omission” approach, whereby the change in *occupancy* was calculated based on the previous occurrence of the taxon (regardless of when that

Formatted: Font: Italic, Complex Script Font: Italic

was) rather than the previous time bin, *per se*. Thus, *i* and *i*-1 do not necessarily correspond to sequential time bins in this approach, but rather consecutive positive sampling intervals for each given taxon. With this approach, consecutive taxon occurrences are included even when separated by "gaps", thus retaining more data to the final data set. Although both approaches yield the same basic results (see Tables S1 and S2), we used the "gap-omission" approach for the sake of retaining a larger data set.

2.5 Binomial logistic modeling

For every species, a record of each time bin in which that species occurred was included in the final ~~table~~ dataset as a single row. Each unique species-bin pairing (row) is characterized with the *occupancy* and a binary extinction indicator in the focal time bin as well as the change in *occupancy* from the previous time bin. An extinction indicator value of 1 was assigned if an occurrence was the last time bin in which a species occurred for the entire data set (the species went extinct or permanently disappeared from the fossil record during this interval). An extinction value of 0 was assigned for all other occurrence records (the species did not go extinct during this interval). Species that are still extant, or those which only went extinct during the most recent time bin (which spans up to the present), would by default be assigned an extinction value of 1 in the most recent time bin. To avoid this edge effect, all occurrences from the most recent time bin were removed prior to model fitting.

~~Species that are still extant, or those which only went extinct during the most recent time bin (which spans up to the present), would by default be assigned an extinction value of 1 in the most recent time bin (because the most recent bin is the last occurrence of those species in the data set). To avoid this edge effect, all occurrences from the most recent time bin were removed prior to model fitting.~~

Binomial logistic models were constructed to examine the dependency of extinction on *occupancy* and *occupancy change*. Both ~~the additive and multiplicative effects of *occupancy*,~~variables ~~*occupancy change*, and *sampling* (represented by the three-timer value calculated for each bin), as well as the impact of each variable independently,~~ were examined with respect to the per-interval probability of extinction. The ~~multiplicative~~ saturated generalized linear model structure of "glm(*extinction* ~ *occupancy* * *occupancy_change* ~~* *sampling*~~, family = binomial(link = 'logit'))," was used. The stepAIC() function in the R package "MASS" (Ripley et al., 2013; version 7.3) was used to select t~~T~~he best-fitting model containing some combination of these variables and their interaction term~~was selected based on its AIC value~~.

2.6 Model performance and predictor importance

We calculated the adjusted amount of deviance (D^2 of Guisan and Zimmermann, 2000) accounted for by each computed logistic model. Deviance in a generalized linear model is analogous to variance of ordinary linear regression. In each of the 16 data sets (four groups with four time resolutions each), the Lindeman, Merenda, and Gold (1980, "LMG") indices of correlated input relative importance (henceforth referred to ~~simply~~ as "relative importance") were calculated for the *occupancy* and~~,~~ *occupancy change* ~~, and *sampling*~~ terms with respect to predicting the *extinction* term. ~~The relative importance of the *occupancy change* term was standardized as a proportion of "total importance".~~ This statistical approach was used to represent the explanatory power of each model term with respect to one another, an insight that is not directly apparent with simple model coefficients. ~~Additionally, for each of the additive model outputs, the ratio of the standardized *occupancy change* coefficient to the standardized~~

~~*occupancy* coefficient was calculated to estimate the relative explanatory power of the *occupancy change* term between data sets.~~

~~For each of these metrics, the standard error was estimated by bootstrapping the species that make up each set of occurrence records. For each of the 1000 bootstrapping iterations, the set of species whose occurrence records were retained in the data set was determined by resampling with replacement. After each resampled set of species and their occurrence records was determined, every model was re-fit and every metric was recalculated. The standard deviation of all bootstrapped values for a single metric approximates the standard error.~~

2.7 Extinction probabilities of extant species

The World Register of Marine Species (https://www.marinespecies.org/), with the assistance of the R package "taxize" (Chamberlain and Szoecs, 2013; version 0.9.100), was used to identify extant species. These data on extant taxa were downloaded on 25 September 2023.

In order to predict the extinction probabilities of extant species, the data sets were reanalyzed and re-fit to models using only the extinct species. Although this technique reduced the overall amount of data used to fit the model, it allowed for the prediction of extinction probabilities of extant species without circularity. Other than removing extant species, all other processes were carried out in the same way as described above.

After selecting the best model for each plankton group, that model was used to predict the extinction probability of extant species. Using the fitted models along with the *occupancy* and *occupancy change* values for each extant species in the present bin (that which ends at the present, 0 Ma), a probability of the binary response variable occurring as a 1 (extinction) can be calculated. This represents the probability that the species will *not* appear again during the next time bin of the same length (that which begins at the present, 0 Ma) or in other future time bins.

~~Because three-timer sampling calculations require bins both before and after the target bin, the present bin lacked a three-timer sampling value. Models were thus fitted without the sampling term for future extinction probability predictions.~~ Extinction predictions were made on extant species subsets without upper Pacman trimming, and the average probability of extinction for all extant species ~~were~~ was calculated in each data set.

2.8 Robustness testing

Further analyses tested the robustness of our results, specifically for the proportional occupancy's utility as a metric of geographic range ~~and for different datasets~~. The same analyses at a bin size of 1 million years were carried out using latitudinal range and change of latitudinal range instead of proportional occupancy and its change. Additionally, the same analyses were carried out using proportional occupancy of Longhurst (2007) provinces and the change in proportional occupancy of Longhurst provinces for data sorted into 1 million-year bins. Mixed-effect models, in which each taxon was considered as a random-effect, were also constructed to check if species identity substantially impacted the basic model results.

Although containing only records of planktonic foraminifers (many of which were sourced from the NSB), the Triton database includes information on the original purpose of each study from which records were sourced, as well as the age of speciation and extinction for each species. With this additional information, the Triton dataset can be used to confirm the suitability of methods used with the Neptune dataset with a different collection of fossil occurrences. Given that some studies may not record every present taxon if it is not a zonal marker or thought to be particularly informative, the Triton dataset was subset to include only studies whose purpose was noted as "community analysis" (Fenton et al. 2021). Because studies whose purpose was to analyze community structure would likely document all present species, by using this subset, studies that potentially excluded some species were removed from the final data set. Additionally, each included species history was subset to exclude any occurrences that occurred before or after the speciation and extinction ages noted in the Triton dataset~~ (Fenton et al. 2021)~~, respectively, reducing the potential impact of reworked fossils in the analysis. Because each species in Triton was trimmed in this manner, these data did not undergo Pacman profiling as the NSB data did. After these additional data-cleaning actions were taken, the Triton dataset had 197,832~~71~~ usable occurrence records and was analyzed in the same way as the NSB data.

3. Results

The utilized datasets show all four groups with elevated extinction rates coming out of the end-Cretaceous mass extinction ~~K-Pg ~~and returning to relative stasis approximately 5-15 Ma after the ~~K-Pg boundary~~event (Fig. 1). ~~Shortly thereafter, a~~All groups~~, but especially the siliceous groups,~~ underwent~~go~~ decreases in diversity, corresponding with the Paleocene-Eocene Thermal Maximum (PETM, Fig. S2~~4~~). ~~Both foraminifera and calcareous nannofossils retain relatively low and stable extinction rates (<0.2) for the rest of the Cenozoic (Fig. 1). ~~Radiolarians and diatoms show spikes in extinction (> 0.4) during the Eocene, including at the Eocene-Oligocene transition as well as at approximately 44 and 41 Ma. The extinction rates (Fig. 1) and diversity patterns (Fig. S2) of each plankton group match those of previous analyses of Neptune data (Jamson et al., 2022), and various biotic events in the Cenozoic, including the Eocene-Oligocene and the Oligocene-Miocene transitions, can be detected.

Formatted: Font: (Default) Times New Roman, Complex Script Font: Times New Roman

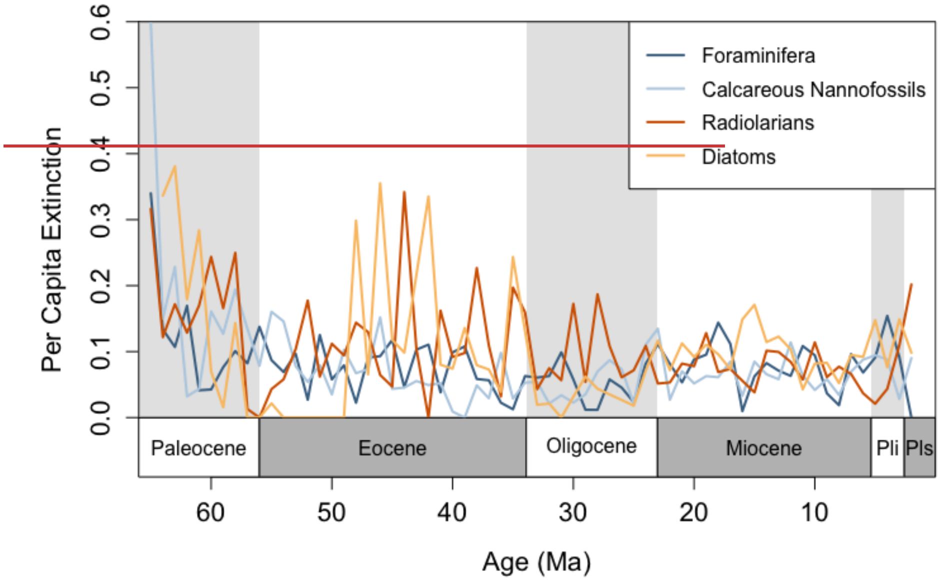
Formatted: Font: (Default) Times New Roman, Complex Script Font: Times New Roman

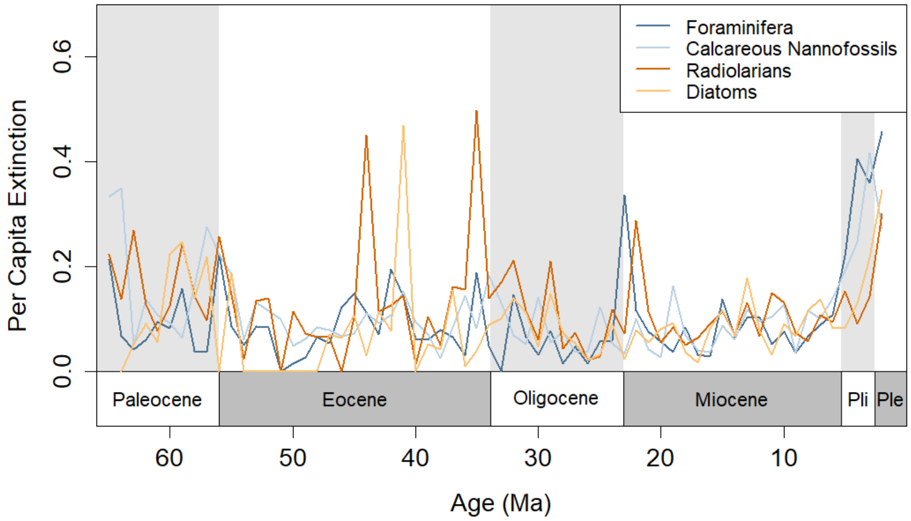
Formatted: Font: (Default) Times New Roman, Complex Script Font: Times New Roman

Formatted: Font: (Default) Times New Roman, Complex Script Font: Times New Roman

Formatted: Font: (Default) Times New Roman, Complex Script Font: Times New Roman

Formatted: Indent: First line: 0.63 cm, Line spacing: Double





Formatted: Centered

Figure 1. Per capita extinction rates calculated using the formula in Foote (1999) for each of the four taxonomic groups during the Cenozoic, calculated ~~for~~ with 1-million-year bins. The timescale of Gradstein et al. (2012) was used here to match age assignments in the NSB.

Formatted: Indent: Before: 0.5 cm, After: 0.51 cm

~~The extinction rates (Fig. 1) and diversity patterns (Fig. S4) of each plankton group match those of previous analyses of Neptune data (Jamson et al., 2022), and various biotic events in the Cenozoic can be detected.~~ *Occupancy* is statistically significant (i.e., statistically non-zero, $p = 0.05$) ~~I~~in all of the analyzed combinations of taxonomic group and bin size. *Occupancy change* is significant in all models except for calcareous nannofossils with a bin size of 0.2 Ma~~, standing *occupancy*, *occupancy change*, and *sampling* were all retained in the AIC-selected models as informative~~ (Table 2 ~~2~~). ~~When considering strictly the simpler additive models, both *occupancy* and *occupancy change* were determined significant in all models ($p = 0.05$). *Sampling* was deemed significant across all bin sizes for calcareous nannofossils, at 0.2 and 1.0 Ma for foraminifera, at 0.2 and 0.5 Ma for diatoms, and never for radiolarians.~~ The term for the interaction between occupancy and occupancy change is significant in all but three models: foraminifera with bin size of 0.2 Ma, foraminifera with a bin size of 0.5 Ma, and diatoms with a bin size of 1.0 Ma (Table 2).

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic

~~**Bin size (Ma):**~~	~~**0.1**~~	~~**0.2**~~	~~**0.5**~~	~~**1.0**~~
~~**Foraminifera**~~	~~*multiplicative*~~	~~*additive*~~	~~*multiplicative*~~	~~*multiplicative*~~
~~**Calcareous Nannofossils**~~	~~*multiplicative*~~	~~*multiplicative*~~	~~*multiplicative*~~	~~*multiplicative*~~
~~**Radiolarians**~~	~~*multiplicative*~~	~~*multiplicative*~~	~~*multiplicative*~~	~~*multiplicative*~~
~~**Diatoms**~~	~~*multiplicative*~~	~~*multiplicative*~~	~~*multiplicative*~~	~~*multiplicative*~~

~~**Table 2.** The AIC-selected model formula for each group and bin size combination, on data that was processed with the "gap-omission" approach. "*multiplicative*" refers to the model with the formula: $ex \sim oc * ch * p$,and "*additive*" refers to the model with the formula: $ex \sim oc + ch + p$. "*ex*" represents the binary response variable *extinction*, "*oc*" represents the *occupancy* term, "*ch*" represents the *occupancy change* term, and "*p*" represents three-timer sampling probability.~~

Occupancy Importance	Occupancy Change Importance	Interaction Importance	D^2	Bin Size (Ma)	Group
0.062 ± 0.010	0.052 ± 0.008	0.040 ± 0.007	0.155	1	F
0.050 ± 0.008	0.028 ± 0.006	--	0.093	0.5	F
0.049 ± 0.007	0.014 ± 0.004	--	0.062	0.2	F
0.036 ± 0.006	0.003 ± 0.001	0.002 ± 0.001	0.041	0.1	F
0.037 ± 0.006	0.040 ± 0.006	0.049 ± 0.007	0.126	1	N
0.045 ± 0.006	0.029 ± 0.005	0.029 ± 0.004	0.103	0.5	N
0.052 ± 0.005	--	0.009 ± 0.002	0.072	0.2	N
0.050 ± 0.004	0.006 ± 0.002	0.002 ± 0.001	0.058	0.1	N
0.015 ± 0.003	0.032 ± 0.005	0.026 ± 0.004	0.073	1	R
0.019 ± 0.003	0.018 ± 0.003	0.013 ± 0.003	0.050	0.5	R
0.016 ± 0.003	0.014 ± 0.003	0.003 ± 0.001	0.033	0.2	R
0.013 ± 0.003	0.004 ± 0.002	0.001 ± 0.0004	0.019	0.1	R
0.017 ± 0.004	0.020 ± 0.005	--	0.037	1	D
0.012 ± 0.003	0.011 ± 0.003	0.002 ± 0.001	0.025	0.5	D
0.007 ± 0.002	0.010 ± 0.003	0.003 ± 0.001	0.020	0.2	D
0.002 ± 0.001	0.007 ± 0.002	0.004 ± 0.001	0.014	0.1	D

Formatted Table

Formatted Table

Formatted Table

Table 2. Relative importance values shown with standard error for the *occupancy*, *occupancy change*, and interaction term for each group at each bin size. The D^2 values for each model are also reported. Darker blue corresponds to higher values. F = foraminifera, N = calcareous nannofossils, R = radiolarians, D = diatoms.

The signs of the *occupancy* and the *occupancy change* coefficients ~~were~~ are always negative ~~in single-term models~~ (Table 3), meaning that smaller instantaneous geographic range sizes and more-negative changes in geographic range size both correspond to larger extinction probabilities (Fig. 2). The D^2 values ~~and AIC values~~ of each analyzed model combination ~~for each 1-million-year binned data~~ set are ~~depicted~~ shown in Table 2~~Fig. 3~~. The maximum D^2 for any model ~~was~~ is 0.155~~29~~, occurring for the ~~multiplicative~~ saturated foraminifera model at a bin size of 1.0 Ma ~~(Table 3)~~. The relative importance of the *occupancy change* term averages~~d~~ ~~52.0%~~0.019, or about 41% of the total explanatory power shared by *occupancy* and *occupancy change*, with higher values being achieved with larger bin sizes.

Occupancy	Occupancy Change	Interaction	Bin Size (Ma)	Group
-4.975	-0.624	-7.860	1	F
-4.419	-0.700	--	0.5	F
-4.863	-0.463	--	0.2	F
-4.406	-0.389	2.299	0.1	F
-3.275	-0.393	-11.363	1	N
-4.080	-0.343	-9.593	0.5	N
-4.576	--	-4.313	0.2	N
-4.516	-0.522	1.970	0.1	N
-2.063	-0.711	-4.907	1	R
-2.484	-0.598	-2.858	0.5	R
-2.138	-0.922	2.021	0.2	R
-1.902	-0.513	1.443	0.1	R
-2.095	-0.627	--	1	D
-2.028	-0.734	2.096	0.5	D
-1.494	-0.830	2.139	0.2	D
-0.918	-0.799	2.036	0.1	D

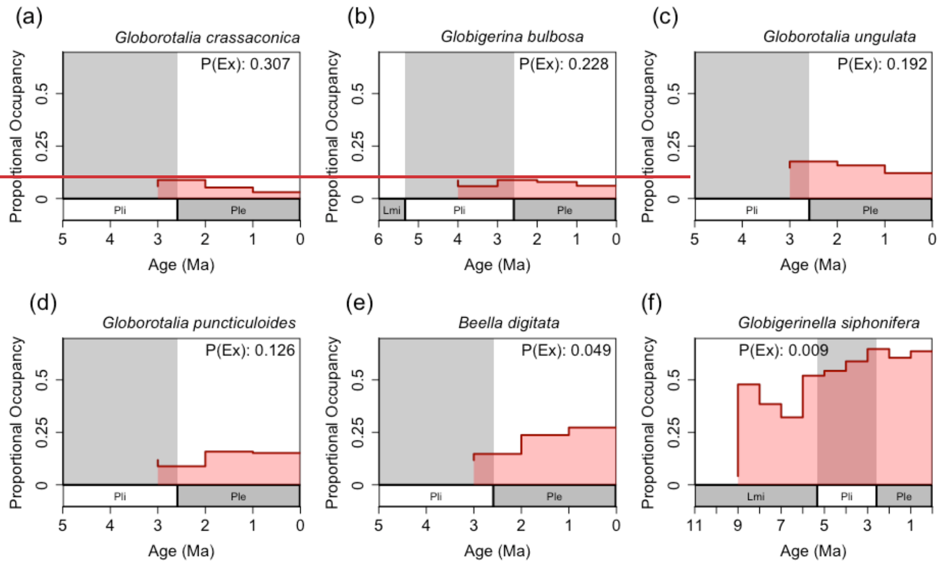
Table 3. Model coefficients for the occupancy, occupancy change, and interaction term for each group and bin size. Darker blue corresponds to more positive values, and darker red corresponds to more negative values. F = foraminifera, N = calcareous nannofossils, R = radiolarians, D = diatoms.

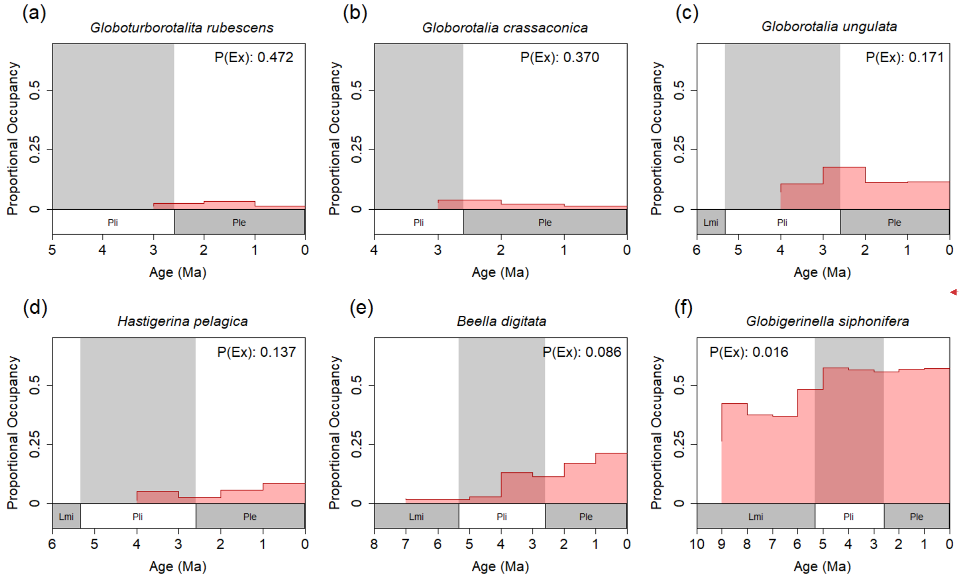
Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

Formatted Table

Formatted: Indent: Before: 2.25 cm, First line: 0 cm, After: 1.51 cm, Line spacing: single



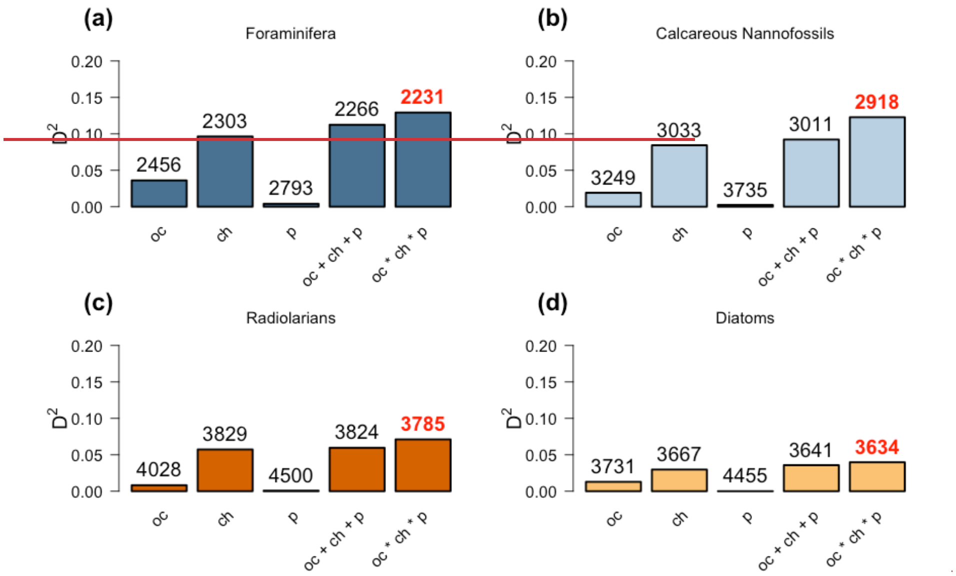


Formatted: Centered

Figure 2. A selection of proportional occupancy through time plots for 6 species of foraminifers sourced from the NSB within~~in~~ one-million-year bin size. In each panel, the current extinction probability of that species—predicted for the next 1 million years using that

species' historical geospatial records~~with the model that was fit using extinct species of the same taxonomic grouping~~ — is shown. Panels A through F are ordered according to decreasing current extinction probability. Notice the association of relatively small standing occupancy values and relatively large occupancy decreases with increased probability of extinction. "Lmi" = Late Miocene, "Pli" = Pliocene, "Ple" = Pleistocene.

Formatted: Indent: Before: 0.5 cm



~~**Figure 3.** D^2 values (bars) and AIC values (numbers displayed above each bar) for each examined model formula. Shown for each of the NSB data sets with bin size = 1 million years. The lowest AIC value for each data set is shown in red.~~

~~**Bin size (Ma):**~~	~~**0.1**~~	~~**0.2**~~	~~**0.5**~~	~~**1.0**~~
~~**Foraminifera**~~	~~0.042~~	~~0.053~~	~~0.079~~	~~0.129~~
~~**Calcareous Nannofossils**~~	~~0.053~~	~~0.053~~	~~0.090~~	~~0.123~~
~~**Radiolarians**~~	~~0.015~~	~~0.025~~	~~0.060~~	~~0.071~~
~~**Diatoms**~~	~~0.021~~	~~0.030~~	~~0.027~~	~~0.040~~

~~**Table 3.** The D^2 of each best-fit model (see Table 2) for each data set at each bin size.~~

D^2 values and relative importance values of the *occupancy change term* increase systematically with bin size, and calcareous organisms tend to have higher ~~higher D^2~~ values than ~~values than~~ siliceous organisms for both metrics ~~for their best-fit model~~ (Table 2~~3~~), indicating better overall explanatory power in calcareous organisms. The~~A~~ maximum relative importance value of *occupancy* is 0.062, and the maximum relative importance value of *occupancy change* is 0.052, both of which were reported for the foraminifera data set with time bins of 1.0 Ma ~~91.9% was achieved for radiolarians binned to 1.0 million years~~ (the D^2 of this particular ~~multiplicative~~ saturated model ~~was~~is 0.155 ~~0.071~~). ~~In all but the smallest bin size, radiolarians consistently have the highest standardized coefficient ratios and relative importance values of the *occupancy change* term.~~ Not surprisingly, sampling completeness increases with larger temporal grain. Foraminifera and calcareous nannofossils ~~had~~ have consistently higher three-timer sampling completeness than the siliceous groups across all bin sizes (Table 4, see Table S3 for SCM completeness).

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

~~Three-Timer Completeness~~				
~~Bin Size (million years)~~	~~**0.1**~~	~~**0.2**~~	~~**0.5**~~	~~**1.0**~~
~~Foraminifera~~	~~0.821 ± 0.006~~	~~0.899 ± 0.006~~	~~0.928 ± 0.007~~	~~0.933 ± 0.009~~
~~Calcareous Nannofossils~~	~~0.874 ± 0.004~~	~~0.899 ± 0.005~~	~~0.925 ± 0.006~~	~~0.936 ± 0.008~~
~~Radiolarians~~	~~0.792 ± 0.005~~	~~0.825 ± 0.006~~	~~0.886 ± 0.007~~	~~0.926 ± 0.008~~
~~Diatoms~~	~~0.729 ± 0.006~~	~~0.781 ± 0.007~~	~~0.838 ± 0.008~~	~~0.861 ± 0.010~~

~~**Table 4.** Three-timer completeness scores (Alroy 2008) calculated for each full data set at each of the four examined bin sizes. Shown with 95% confidence intervals.~~

Three-Timer Completeness				
Bin Size (million years)	**0.1**	**0.2**	**0.5**	**1.0**
Foraminifera	0.869 ± 0. 005	0.921 ± 0. 005	0.957 ± 0. 006	0.949 ± 0.009
Calcareous Nannofossils	0.891 ± 0. 004	0.918 ± 0. 005	0.942 ± 0. 006	0.950 ± 0. 008
Radiolarians	0.857 ± 0. 005	0.879 ± 0. 006	0.922 ± 0. 007	0.947 ± 0. 008
Diatoms	0.830 ± 0. 006	0.867 ± 0. 007	0.898 ± 0. 008	0.909 ± 0. 010

Table 4. Three-timer completeness scores (Alroy 2008) calculated for each full data set at each of the four examined bin sizes. Shown with 95% confidence intervals.

When fit to extinct-only data subsets, AIC-selected models ~~were~~ still ~~either the additive or multiplicative three-term models~~retain both the *occupancy* and the *occupancy change* term for all groups at all bin sizes ~~(Table S4)~~. Extant diatoms show~~ed~~ a significantly higher probability of extinction than ~~all other groups, followed by extant radiolarians, and then by~~ the two calcareous groups, and e~~. E~~xtinct radiolarians ~~had~~ have a significantly lower ~~extinction~~ probability of extinction than calcareous nannofossils and diatoms (Table 5). ~~Extinction values through time for each group are visualized in Fig. 1.~~

	Species

Formatted: Indent: Before: 0 cm

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

	Extant	Extinct
Foraminifera	0.130 ± 0.049	0.235 ± 0.033
Calcareous Nannofossils	0.070 ± 0.044	0.262 ± 0.016
Radiolarians	0.177 ± 0.014	0.208 ± 0.012
Diatoms	0.191 ± 0.010	0.269 ± 0.020

Table 5 ~~5~~**.** Mean of all extant-organism extinction probabilities (left) and extinct-organism extinction probabilities (right) produced for each data set. Means are shown with 95% confidence interval.

The raw number of cells occupied by each species during each time bin ~~showed~~ shows a strong positive correlation with the number of occupied Longhurst provinces for each taxonomic group at each bin size, even when autocorrelation ~~was~~ is removed (Table S4~~5~~). The same analyses conducted on proportional occupancy of Longhurst provinces and the change in proportional occupancy of Longhurst provinces, as well as on the latitudinal range of species and the change in latitudinal range of species, yield~~ed~~ best-fit models that ~~related~~retain ~~all three model terms multiplicatively~~both variables and their interaction term (~~Figs. S2 and S3~~Tables S5 and S6). When the logistic modeling ~~was~~ is applied to the Triton data, ~~the three-term multiplicative or additive models always performed the best~~both the *occupancy* and the *occupancy change* terms are retained across the three largest bin sizes (~~Fig~~Table~~.~~ S9~~6~~). The maximum D^2 value of 0.072 is achieved with a bin size of 1.0 Ma, and the LMG values of *occupancy* and *occupancy change* are 0.050 and 0.016, respectively. Tests for multicollinearity show only minor correlation (mean = 0.23, maximum = 0.32) between the examined variables (Table S7). Although several datasets failed to reach convergence, the mixed-effect models show similar results to those reported as part of the main analysis (Table S8).

Formatted: Indent: Before: 2 cm, After: 2.01 cm

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

4. Discussion

4.1 Geographic range as a driver of microplankton diversity

~~Our observations agree with the findings of Jamson et al. (2022), which observed increased extinction rates for siliceous organisms during the PETM. Additionally, the authors found a net diversification rate of around 0, as well as a stable extinction rate, for calcareous groups between the PETM and the modern. Our results show diatoms and radiolarians having several spikes in extinction between 40-50 Ma, which may indicate the termination of the Early Eocene Climatic Optimum. Jamson et al. (2022) observed a brief increase in diatom extinction rate during this event.~~

~~There is another localized spike in extinction for the siliceous groups around 35Ma, perhaps corresponding with the end of the Middle Eocene Climatic Optimum. Jamson et al. (2022) observed a brief increase in diatom extinction rate during this event.~~

While absolute geographic range size is an informative predictor of extinction risk, various other factors relating to geographic range also play an important role in global biodiversity patterns. Powell and Glazier (2017) found that in the same four groups of microplankton analyzed in this study, latitudinal diversity gradients are produced by asymmetric shifts in geographic range, rather than variations in diversification rate with respect to latitude. On the contrary, Raja and Kiessling (2021) found that the extratropics had higher average origination rates than the tropics. Supporting Powell and Glazier’s (2017) finding that asymmetric shifts in geographic range are key drivers of latitudinal diversity gradients, Raja and Kiessling (2021) showed that dispersal was more likely to occur from the extratropics towards the tropics. Both studies suggest that latitudinal diversity gradients, and thus the geographic distribution of a species, are closely linked to paleoclimate regimes. Indeed, changes in global circulation patterns, ~~as well as~~ water column stratification, and temperature are all among the major ~~key~~ influences on global plankton diversity (Lowery et al., 2020)

Although the ranges of marine microplankton have been known to shift in response to climate (Ying et al., 2024, Chaabane et al., 2024), this is not always the case, as species ranges sometimes fail to keep up with shifting temperature zones ~~As paleoclimate regimes shift over geologic time, marine microplankton ranges often do not shift to follow changes in temperature; instead, populations may simply go extinct as local temperatures become inhospitable without a comparable range expansion into more hospitable zones~~ (Trubovitz et al., 2020). Trubovitz et al. (2023) found that radiolarian abundance is not a significant predictor of extinction risk and that external drivers (such as climate) are more likely to predict extinctions. Thus, while some species do migrate in response to climate change, larger geographic ranges may provide a geographic cushion to species that don’t~~with a geographic cushion~~: as local temperatures change, more widespread species undergo a more drawn-out series of local extirpations before global extinction occurs. This agrees with the well-established phenomenon, and that which we also report here, that larger instantaneous geographic occupancy reduces a species’ risk of extinction (Foote et al., 2016; Foote et al., 2007; McKinney, 1997; Payne and Finnegan, 2007; Purvis et al., 2000; Staude et al., 2020).~~.~~

Additionally, the trajectory of a species’ geographic range through time might indirectly reflect shifts in regional or global climate. As paleoclimate zones shift, geographic cells may become inhospitable to a species, and the species may undergo extirpation in that geographic cell. A more rapid change in a species’ occupancy through time may reflect a more rapid change in paleoclimate and hospitable regions of Earth. Continued reduction in occupancy over time can thus provide insight into the effects of long-term climatic, geographic, or biological trends on the extinction probability of marine microplankton.

4.2 History of occupancy / legacy effects

The trajectory of various ecological variables through time has been shown to impact the current and future direction of species diversity trends. These legacies may include past climatic events or geographic range shifts influencing modern distributions or extinctions of species (Svenning et al., 2015). The interaction of historic information with current information can provide insight about ecological processes that neither historic nor current information could provide on its own.

The historic trajectory of climate change impacts the probability of extinction occurring with a short-term change in climate. A warming event occurring after a long-term warming trend leads to greater extinction rates (Mathes et al., 2021~~a~~) ~~and lower origination rates (Mathes et al., 2021b)~~ than a warming event occurring after a long-term cooling trend. Understanding the historical conditions leading up to a study period of interest ~~thus~~ may thus be essential to understanding the key drivers as to what goes extinct versus what survives.

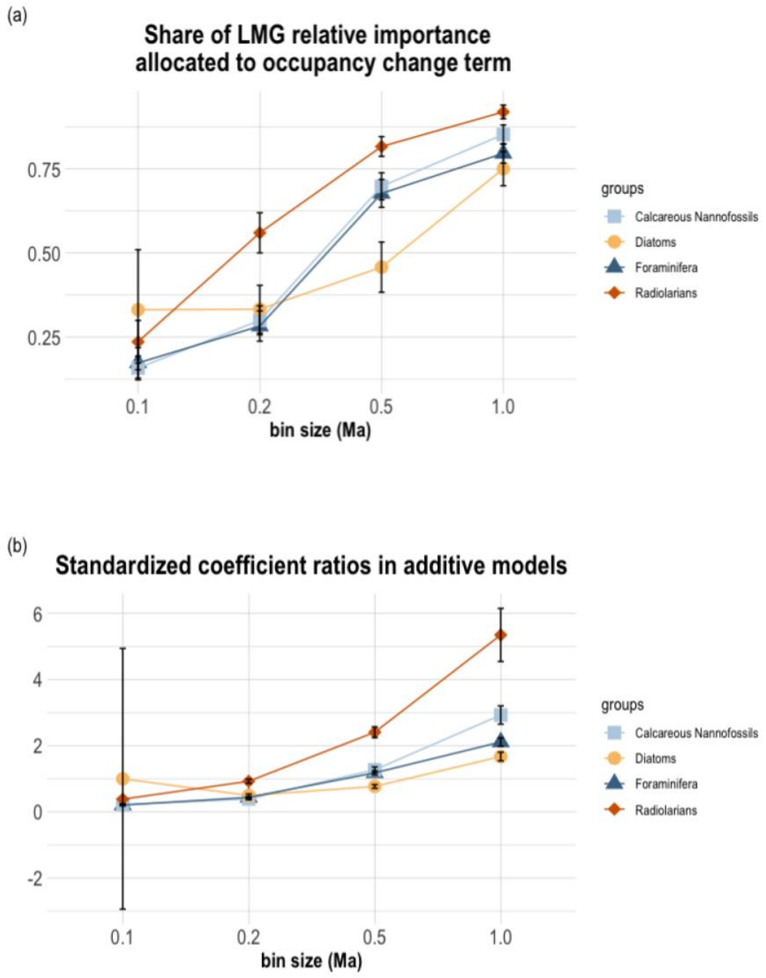
Although the effect of climate and geographic range legacies on instantaneous geographic range ~~itself~~ is well-studied (Svenning et al., 2015), the effect of geographic range legacies on instantaneous extinction probability has not received as much attention. Of course, populations of species cannot “look” backwards, but are instead influenced by the current conditions present in an environment. The predictive capability of the *occupancy change* term may thus be an indicator of continued unfavorable conditions (perhaps spanning millions of years) acting on a population at a given time. Kiessling and Kocsis (2016) found that the legacy of geographic range (represented as its change to the present from the previous bin) is an informative predictor for extinction risk in marine macroinvertebrates. Our results build upon those of Kiessling and Kocsis (2016), demonstrating both that these findings hold true for marine microplankton, as well as that temporal scale (bin size) is a key variable in detecting the importance of geographic legacy effects.

4.3 Scale dependency of extinction drivers:

Although previous studies have analyzed various drivers of extinction through geologic time, relatively little research has gone into understanding the scale-dependency of these extinction drivers. Scale-dependency in extinction studies manifests in various variables, such as area (Fagan et al., 2005; Guardiola ~~and Rodá~~et al., 2013) or taxon age (Henao Diaz et al., 2019). Analyzing data at different temporal scales is also imperative to detect true ecological signals (Hewitt et al., 2010). We find ~~here~~ that as temporal resolution decreases (bin size increases), the relative ~~explanatory power~~importance of both the *occupancy* and *occupancy change* variables increases ~~while the relative explanatory power of the other two variables decreases~~ (~~Fig~~Table~~.~~ 2~~4~~).

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic



~~**Figure 4. A.)** Share of the Lindemann, Merenda and Gold relative importance (1980) allocated to the *occupancy change* term, taken as a proportion of "total importance". Shown for each data set at each bin size. **B.)** Standardized coefficient ratios in AIC-selected models for each bin size. Error bars represent the bootstrap-estimated standard error. Note that as bin size increases (temporal resolution decreases), the relative importance of the occupancy~~

Formatted: Indent: First line: 1.27 cm, Space After: 0 pt, Widow/Orphan control

Formatted: Indent: First line: 1.27 cm, Space After: 0 pt, Line spacing: Double, Widow/Orphan control

~~change term as well as the standardized coefficient ratio (occupancy change : standing occupancy) both increase.~~

This could result from there being more records in a single temporal bin as bin size increases, thus ~~creating a more complete set of data in any given time bin~~increasing statistical power. ~~This may be especially true when working with fragmentary fossil data, such as that which we analyze here.~~ With larger bin sizes, it is easier to detect biological signals that may otherwise be lost in the noise of fragmentary data~~. The temporal scale by which we analyze these data can influence our understanding of extinction risk in marine microplankton~~. We show here that the seemingly-arbitrary selection of temporal bin size can have major impacts on conclusions drawn about microplankton diversification, and that coarser resolutions may more reliably indicate actual macroevolutionary trends.~~.~~

4.4 Calcareous vs siliceous microfossils

In general, the explanatory power of each of the model terms is smaller in the siliceous groups than in the calcareous groups. Although *occupancy* and *occupancy change* were found to be informative across all groups, the signals are weaker in diatoms and radiolarians (Table 2~~3~~, Table 3). ~~As both siliceous and calcareous organisms are sensitive to nutrient limitation (Kamikuri et al., 2005; Lazarus et al., 2009; Bristow et al., 2017),~~ T~~t~~his discrepancy likely results from minor variations in sampling, as evidenced by lower 3-timer completeness values for the two siliceous groups. ~~Furthermore, the *sampling* term was found to be significant in the three-term additive model across all bin sizes for calcareous nannofossils, in two bin sizes for both foraminifera and diatoms, and never for radiolarians. The greater importance of the *sampling* term in the calcareous groups suggests that fossil preservation and sampling rate has a greater influence on when a species disappears from the fossil record, a piece of information that is~~

~~crucial to distinguishing “true” from “apparent” extinction. This~~The ~~discrepancy~~ difference may also be a result of variable taphonomic pathways between the calcareous and siliceous organisms (Boltovskoy 1994). Nonetheless, both *occupancy* and *occupancy change* are important predictors of extinction regardless of the group, and these findings further underscore the importance of accounting for sampling when analyzing paleontological data.

4.5 Robustness testing

There is a strong correlation between the number of occupied Longhurst provinces and the number of individual occupied geographic cells for each species-bin pairing. This demonstrates that, although the locations of the various drilling expeditions that sourced much of the data in the Neptune database are not entirely random, when taken together they still account for a diverse spread of planktonic biogeographic regions around the globe. This supports the idea that the collection of data contained in the Neptune database~~, while only a snapshot of the global microplankton fossil record,~~ is comprehensive enough to study large-scale biogeographic trends. Additionally, AIC-selected models contained both the *occupancy* and *occupancy change* terms even when geographic range was measured as latitudinal expanse or as a proportion of occupied Longhurst provinces (Tables S5 and S6). This suggests that the significance of proportional occupancy change in predicting extinction is not merely an artifact of data processing.

The AIC-selected model for each bin size in the Triton data set ~~was either the three-term additive or multiplicative model~~always retains both *occupancy* and *occupancy change* as significant except with a bin size of 0.1 Ma (Table S9). Although the Triton dataset has substantially more occurrence records after preprocessing, it has consistently lower diversity compared to the other taxonomic groups from the NSB (Fig. S2~~4~~). This could indicate a greater propensity for “lumping” in the Triton dataset than in the NSB, which in turn could change how

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

spatio-temporal signals manifest. The similar results obtained from the Triton dataset further confirm~~s~~ the suitability of these methods with an alternative data set and reaffirm~~s~~ the importance of *occupancy* and *occupancy change* when modeling extinction.

Taken together, our findings suggest that the change in geographic occupancy is an important metric for predicting extinction across ~~the tree of~~marine life. Kiessling and Kocsis (2016) looked exclusively at skeletal macroinvertebrates, whereas we here ~~we~~ analyze~~d~~ several protist lineages of marine plankton. The broad taxonomic scope of these findings emphasizes the fundamental importance of the trajectory in geographic range as a biological metric, which can be a key aspect of taxon dynamics through time. Although the explanatory power of the model may seem low (up to 15.5~~12.9~~%), it is an important factor given the many other variables that influence extinction risk (McKinney 1997)

4.6 Future Perspectives

Formatted: Font: Bold, Complex Script Font: Bold

Formatted: Indent: First line: 0 cm

Although modern studies can track geographic occupancy change over the course of decades (if there is a history of consistent data collection), estimates of marine species durations average between 5-10 million years (Foote and Raup, 1996; Raup, 1991), much longer than human-collected records can encompass. To fully understand the change of occupancy through a species' duration, records extending beyond those which could have been manually recorded by conservation biologists are needed. Although some modern conservation practitioners have been hesitant to fully embrace long-term paleontological data, this study provides yet another argument for the incorporation of historical perspectives and fossil evidence in conservation efforts (Dietl et al., 2019; Kiessling et al., 2019; Smith et al., 2018).

While for simplicity's sake, this study only looked at the interaction of *occupancy* and the first degree of *occupancy change* (bin number *i* to *i*-1), future iterations could incorporate entire occupancy histories into model fitting using even more advanced techniques. This may help the model overcome variations in sampling intensity or localized paleoenvironmental events, and let the models provide information not only on decline, but information on *continued* decline — another hallmark of increased extinction risk.

5. Conclusion

In providing evidence that the geological history of species distributions plays a significant role in species extinction risk, our study demonstrates the importance of paleontological data for assessing modern species extinction risk. ~~Incorporating paleontological range data may allow for more-informed decisions about how to spend limited conservation funding in the future.~~These findings provide empirical support for the connection between continued range reduction and ultimate global extinction in marine microplankton. ~~-~~We also demonstrate here the importance of temporal grain in detecting biological signal in fragmentary fossil data.

Data availability statement

All data and code is currently accessible at the reserved stable repository DOI: 10.5281/zenodo.15174296 ~~10.5281/zenodo.7745607~~ (Smith, 2025~~4~~).

Author contributions

WK and AK developed the conceptual framework. IS constructed the analytical pipeline and carried out analyses and drafted manuscript. All authors contributed to the development of the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Formatted: Font: 12 pt, Complex Script Font: 12 pt

Formatted: HTML Preformatted, Indent: Before: 0 cm, First line: 0 cm, Pattern: Clear (White)

Formatted: Font: (Default) Times New Roman, 12 pt, Not Bold, Complex Script Font: Times New Roman, 12 pt, Not Bold

Formatted: Font: 12 pt, Font color: Black, Complex Script Font: 12 pt

Acknowledgements

The study was supported by the Deutsche Forschungsgemeinschaft (Ko 5382/2-1) and was embedded in the Research Unit TERSANE (FOR 2332). The authors would like to thank Johan Renaudie for ongoing assistance with the Neptune Sandbox Berlin, as well as for providing insightful comments and suggestions.

References

Alroy, J. (2008). Dynamics of origination and extinction in the marine fossil record. *Proceedings of the National Academy of Sciences*, *105*(supplement_1), 11536-11542.

Benton, M. J. (1985). Mass extinction among non-marine tetrapods. *Nature*, *316*(6031), 811-814.

Boltovskoy, D. (1994). The sedimentary record of pelagic biogeography. *Progress in Oceanography*, *34*(2-3), 135-160.

Boyden, J. A., Müller, R. D., Gurnis, M., Torsvik, T. H., Clark, J. A., Turner, M., Ivey-Law, H., Watson, R. J., & Cannon, J. S. (2011). Next-generation plate-tectonic reconstructions using GPlates, *in* Baru, C., and Keller, G. R., eds., Geoinformatics: Cyberinfrastructure for the Solid Earth Sciences: Cambridge, Cambridge University Press, p. 95-114.

~~Bristow, L. A., Mohr, W., Ahmerkamp, S., & Kuypers, M. M. (2017). Nutrients that limit growth in the ocean. *Current Biology*, *27*(11), R474-R478.~~

~~Burnham K.P., Anderson D.R. (2002). Model selection and multimodel inference: a practical information-theoretic approach. Springer, New York.~~

~~Ceballos, G., Ehrlich, P. R., Barnosky, A. D., García, A., Pringle, R. M., & Palmer, T. M. (2015). Accelerated modern human‐induced species losses: Entering the sixth mass extinction. *Science advances*, *1*(5), e1400253.~~ Chaabane, S., de Garidel-Thoron, T., Meilland, J., Sulpis, O., Chalk, T. B., Brummer, G. J. A., Mortyn, P.G., Giraud, X., Howa, H., Casajus, N., Kuroyanagi, A., Beaugrand, G., & Schiebel, R. (2024). Migrating is not enough for modern planktonic foraminifera in a changing ocean. *Nature*, 1-7.

Chamberlain, S., Szocs, E. (2013). "taxize - taxonomic search and retrieval in R." *F1000Research*. https://f1000research.com/articles/2-191/v2.

Darroch, S. A., Saupe, E. E., Casey, M. M., & Jorge, M. L. (2022). Integrating geographic ranges across temporal scales. *Trends in Ecology & Evolution*.

~~de Lafontaine, G., Napier, J. D., Petit, R. J., & Hu, F. S. (2018). Invoking adaptation to decipher the genetic legacy of past climate change. *Ecology*, *99*(7), 1530-1546.~~

~~De Vos, J. M., Joppa, L. N., Gittleman, J. L., Stephens, P. R., & Pimm, S. L. (2015). Estimating the normal background rate of species extinction. *Conservation biology*, *29*(2), 452-462.~~

Dietl, G. P., Smith, J. A., & Durham, S. R. (2019). Discounting the past: the undervaluing of paleontological data in conservation science. *Frontiers in Ecology and Evolution*, 108.

Fagan, W. F., Aumann, C., Kennedy, C. M., & Unmack, P. J. (2005). Rarity, fragmentation, and the scale dependence of extinction risk in desert fishes. *Ecology*, *86*(1), 34-41.

Fenton, I. S., Woodhouse, A., Aze, T., Lazarus, D., Renaudie, J., Dunhill, A. M., Young, J. R., & Saupe, E. E. (2021). Triton, a new species-level database of Cenozoic planktonic foraminiferal occurrences. *Scientific Data*, *8*(1), 160.

Formatted: English (United States)

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

Foote, M., Crampton, J. S., Beu, A. G., Marshall, B. A., Cooper, R. A., Maxwell, P. A., & Matcham, I. (2007). Rise and fall of species occupancy in Cenozoic fossil mollusks. *Science*, *318*(5853), 1131-1134.

Foote, M., Ritterbush, K. A., & Miller, A. I. (2016). Geographic ranges of genera and their constituent species: structure, evolutionary dynamics, and extinction resistance. *Paleobiology*, *42*(2), 269-288.

Foote, M. (1999). Morphological diversity in the evolutionary radiation of Paleozoic and post-Paleozoic crinoids. *Paleobiology*, *25*(S2), 1-115.

Gradstein, F.M., Ogg, J.G., Schmitz, M.D., and Ogg, G.M. (2012). The Geologic Time Scale. Elsevier, Amsterdam.

Guardiola, M., Pino, J., & Rodà, F. (2013). Patch history and spatial scale modulate local plant extinction and extinction debt in habitat patches. *Diversity and Distributions*, *19*(7), 825-833.

Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological modelling*, *135*(2-3), 147-186.

Henao Diaz, L. F., Harmon, L. J., Sugawara, M. T., Miller, E. T., & Pennell, M. W. (2019). Macroevolutionary diversification rates show time dependency. *Proceedings of the National Academy of Sciences*, *116*(15), 7403-7408.

Hewitt, J. E., Thrush, S. F., & Lundquist, C. (2010). Scale-dependence in ecological systems. *Encyclopedia of Life Sciences (ELS)*, 1-7.

Huber, B. T., Petrizzo, M. R., Young, J. R., Falzoni, F., Gilardoni, S. E., Bown, P. R., & Wade, B. S. (2016). Pforams@ microtax. *Micropaleontology*, *62*(6), 429-438. Jamson, K. M., Moon, B. C., & Fraass, A. J. (2022). Diversity dynamics of microfossils from the Cretaceous to the Neogene show mixed responses to events. *Palaeontology*, *65*(4), e12615.

~~Huber, B. T. et al. (2017b). Pforams@microtax: A new online taxonomic database for planktonic foraminifera. *Micropaleontology*. **62**(6): 429-438.~~

Jamson, K. M., Moon, B. C., & Fraass, A. J. (2022). Diversity dynamics of microfossils from the Cretaceous to the Neogene show mixed responses to events. *Palaeontology*, *65*(4), e12615.

~~Kamikuri, S. I., Nishi, H., Moore, T. C., Nigrini, C. A., & Motoyama, I. (2005). Radiolarian faunal turnover across the Oligocene/Miocene boundary in the equatorial Pacific Ocean. *Marine Micropaleontology*, *57*(3-4), 74-96.~~

Kiessling, W., & Kocsis, Á. T. (2016). Adding fossil occupancy trajectories to the assessment of modern extinction risk. *Biology letters*, *12*(10), 20150813.

Kiessling, W., Raja, N. B., Roden, V. J., Turvey, S. T., & Saupe, E. E. (2019). Addressing priority questions of conservation science with palaeontological data. *Philosophical Transactions of the Royal Society B*, *374*(1788), 20190222.

Kiessling, W. (2005). Habitat effects and sampling bias on Phanerozoic reef distribution. *Facies*, *51*(1-4), 24-32.

Kocsis, Á. (2020). icosa: global triangular and penta-hexagonal grids based on tessellated icosahedra, R package version 0.10. 0. In.

Kocsis, Á. T., Reddin, C. J., Alroy, J., & Kiessling, W. (2019). The R package divDyn for quantifying diversity dynamics using fossil sampling data. *Methods in Ecology and Evolution*, *10*(5), 735-743.

Formatted: German (Germany)

Lazarus, D., Barron, J., Renaudie, J., Diver, P., & Türke, A. (2014). Cenozoic planktonic marine diatom diversity and correlation to climate change. *PLoS One*, *9*(1), e84857.

Lazarus, D., Weinkauf, M., & Diver, P. (2012). Pacman profiling: a simple procedure to identify stratigraphic outliers in high-density deep-sea microfossil data. *Paleobiology*, *38*(1), 144-161.

~~Lazarus, D. B., Kotrc, B., Wulf, G., & Schmidt, D. N. (2009). Radiolarians decreased silicification as an evolutionary response to reduced Cenozoic ocean silica availability. *Proceedings of the National Academy of Sciences*, *106*(23), 9333-9338.~~

Lazarus, D. (1994). Neptune: a marine micropaleontology database. *Mathematical Geology*, *26*, 817-832.

Lindeman, R. H., Merenda, P. F., & Gold, R. Z. (1980). Introduction to bivariate and multivariate analysis. *Scott, Foresman & Company Glenview, IL*.

Liow, L. H., Skaug, H. J., Ergon, T., & Schweder, T. (2010). Global occurrence trajectories of microfossils: environmental volatility and the rise and fall of individual species. *Paleobiology*, *36*(2), 224-252.

Longhurst, A. R. (2007). *Ecological geography of the sea*. Elsevier.

Lowery, C. M., Bown, P. R., Fraass, A. J., & Hull, P. M. (2020). Ecological response of plankton to environmental change: thresholds for extinction. *Annual Review of Earth and Planetary Sciences*, *48*(1), 403-429.

Mace, G. M., Collar, N. J., Gaston, K. J., Hilton-Taylor, C., Akçakaya, H. R., Leader-Williams, N., Milner-Gulland, E. J., & Stuart, S. N. (2008). Quantification of extinction risk: IUCN's system for classifying threatened species. *Conservation biology*, *22*(6), 1424-1442.

~~Mathes, G. H., Kiessling, W., & Steinbauer, M. J. (2021b). Deep-time climate legacies affect origination rates of marine genera. *Proceedings of the National Academy of Sciences*, *118*(36), e2105769118.~~

Mathes, G. H., van Dijk, J., Kiessling, W., & Steinbauer, M. J. (2021~~a~~). Extinction risk controlled by interaction of long-term and short-term climate change. *Nature Ecology & Evolution*, *5*(3), 304-310.

McKinney, M. L. (1997). Extinction vulnerability and selectivity: combining ecological and paleontological views. *Annual review of ecology and systematics*, *28*(1), 495-516.

Nigrini, C., Sanfilippo, A., & Moore Jr., T. C. (2006). Cenozoic radiolarian biostratigraphy: a magnetobiostratigraphic chronology of Cenozoic sequences from ODP Sites 1218, 1219, and 1220, Equatorial Pacific. Proceedings of the Ocean Drilling Program, Scientific Results 199:1–76.

Payne, J. L., & Finnegan, S. (2007). The effect of geographic range on extinction risk during background and mass extinction. *Proceedings of the National Academy of Sciences*, *104*(25), 10506-10511.

~~Payne, J. L., Bush, A. M., Heim, N. A., Knope, M. L., & McCauley, D. J. (2016). Ecological selectivity of the emerging mass extinction in the oceans. *Science*, *353*(6305), 1284-1286.~~

~~Peters, S. E., & McClennen, M. (2016). The Paleobiology Database application programming interface. *Paleobiology*, *42*(1), 1-7.~~

Powell, M. G., & Glazier, D. S. (2017). Asymmetric geographic range expansion explains the latitudinal diversity gradients of four major taxa of marine plankton. *Paleobiology*, *43*(2), 196-208.

Formatted: German (Germany)

Purvis, A., Gittleman, J. L., Cowlishaw, G., & Mace, G. M. (2000). Predicting extinction risk in declining species. *Proceedings of the royal society of London. Series B: Biological Sciences*, *267*(1456), 1947-1952.

Raja, N. B., & Kiessling, W. (2021). Out of the extratropics: The evolution of the latitudinal diversity gradient of Cenozoic marine plankton. *Proceedings of the Royal Society B*, *288*(1950), 20210545.

Raup, D. M. (1991). A kill curve for Phanerozoic marine species. *Paleobiology*, *17*(1), 37-48.

R Core Team (2022). A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria https://www.R-project.org/.

Renaudie, J., Lazarus, D. B., & Diver, P. (2020). NSB (Neptune Sandbox Berlin): An expanded and improved database of marine planktonic microfossil data and deep-sea stratigraphy. *Palaeontologia Electronica*, *23*(1), 1-28.

Renaudie, J. (2019). plannapus/NSBcompanion: NSBcompanion 2.1 (v2.1.1). Zenodo. https://doi.org/10.5281/zenodo.3408190.

Ripley, B., Venables, B., Bates, D. M., Hornik, K., Gebhardt, A., Firth, D., & Ripley, M. B. (2013). Package ‘mass’. *Cran r*, *538*(113-120), 822.

Saulsbury, J. G., Parins-Fukuchi, C. T., Wilson, C. J., Reitan, T., & Liow, L. H. (2023). Age-dependent extinction and the neutral theory of biodiversity. *Proceedings of the National Academy of Sciences*, *121*(1), e2307629121.

~~Schlegel, M., & Hülsmann, N. (2007). Protists A textbook example for a paraphyletic taxon. *Organisms Diversity & Evolution*, *7*(2), 166-172.~~

Smith, I. (2025~~4~~). Preprint -- Occupancy Trajectory Microplankton. DOI: 10.5281/zenodo.15174296~~10.5281/zenodo.7745607~~. Zenodo.

Smith, J. A., Durham, S. R., & Dietl, G. P. (2018). Conceptions of long-term data among marine conservation biologists and what conservation paleobiologists need to know. *Marine conservation paleobiology*, 23-54.

Staude, I. R., Navarro, L. M., & Pereira, H. M. (2020). Range size predicts the risk of local extinction from habitat loss. *Global Ecology and Biogeography*, *29*(1), 16-25.

Strack, T., Jonkers, L., C. Rillo, M., Baumann, K. H., Hillebrand, H., & Kucera, M. (2024). Coherent response of zoo-and phytoplankton assemblages to global warming since the Last Glacial Maximum. *Global Ecology and Biogeography*, *33*(6), e13841.

Svenning, J. C., Eiserhardt, W. L., Normand, S., Ordonez, A., & Sandel, B. (2015). The influence of paleoclimate on present-day patterns in biodiversity and ecosystems. *Annual Review of Ecology, Evolution, and Systematics*, *46*(1), 551-572.

Swain, A., Woodhouse, A., Fagan, W. F., Fraass, A. J., & Lowery, C. M. (2024). Biogeographic response of marine plankton to Cenozoic environmental changes. *Nature*, 1-8.

Tietje, M., & Kiessling, W. (2013). Predicting extinction from fossil trajectories of geographical ranges in benthic marine molluscs. *Journal of Biogeography*, *40*(4), 790-799.

Trubovitz, S., Lazarus, D., Renaudie, J., & Noble, P. J. (2020). Marine plankton show threshold extinction response to Neogene climate change. *Nature communications*, *11*(1), 5069.

Trubovitz, S., Renaudie, J., Lazarus, D., & Noble, P. J. (2023). Abundance does not predict extinction risk in the fossil record of marine plankton. *Communications Biology*, *6*(1), 554.

~~Vilhena, D. A., & Smith, A. B. (2013). Spatial bias in the marine fossil record. *PLoS One*, *8*(10), e74470.~~

Formatted: Default Paragraph Font, Font: (Default) Calibri, (Asian) Calibri, Complex Script Font: Calibri

Formatted: Font: (Default) Times New Roman, 12 pt, Complex Script Font: Times New Roman, 12 pt

Formatted: Font: (Default) Times New Roman, 12 pt, Complex Script Font: Times New Roman, 12 pt

Formatted: Font: (Default) Times New Roman, 12 pt, Complex Script Font: Times New Roman, 12 pt

~~Wagler, R. (2011). The anthropocene mass extinction: An emerging curriculum theme for science educators. *The American Biology Teacher*, *73*(2), 78-83.~~Ying, R., Monteiro, F. M., Wilson, J. D., Ödalen, M., & Schmidt, D. N. (2024). Past foraminiferal acclimatization capacity is limited during future warming. *Nature*, 1-5.

Formatted: German (Germany)

Formatted: Font: (Default) Times New Roman, 12 pt, Complex Script Font: Times New Roman, 12 pt, German (Germany)

Formatted: Font: (Default) Times New Roman, 12 pt, Complex Script Font: Times New Roman, 12 pt