

Supplementary Material

Searching for function: Reconstructing adaptive niche changes using geochemical and morphological data in planktonic Foraminifera

L.E. Kearns*, S.M. Bohaty, K.M. Edgar, S. Nogué and T.H.G. Ezard

*Correspondence:

Lorna Kearns

L.Kearns@soton.ac.uk

Contents

Introduction to supplementary material	2
Introduction to data	3
Importing of data	4
Multi-specimen analysis	5
Multiple Linear Regression	5
Individual analysis	5
Clustering Analysis	5
Analysis of variance	6
Stable Isotopes using subset of <i>Subbotina</i> geochemically analysed (Supplementary table 4) . .	6
Morphology using supplementary table 3 with all 300 <i>Subbotina</i> measurements	6
Multiple linear regression	7
With size	7
With weight	12
Including climate	14
Bibliography	16

Introduction to supplementary material

This supplementary material is intended to provide reproducible code for the results presented in the paper. All analyses were performed in the freely distributed R environment (R Core Team 2020). This supporting information was written using knitr (Xie 2021). The code uses the tidyverse (Wickham 2019) and packages within, mclust (Fraley, Raftery, and Scrucca 2020), kableExtra (Zhu 2021), scales (Wickham and Seidel 2020) and broom(Robinson, Hayes, and Couch 2021) packages as well as dependencies:

```
library(tidyverse)
library(mclust)
library(knitr)
library(kableExtra)
library(png)
library(scales)
library(broom)
library(bookdown)
print(sessionInfo(), local = FALSE)

## R version 4.0.3 (2020-10-10)
## Platform: x86_64-apple-darwin17.0 (64-bit)
## Running under: macOS Mojave 10.14.6
##
## Matrix products: default
## BLAS: /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/libRblas.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/libRlapack.dylib
##
## attached base packages:
## [1] stats      graphics  grDevices  utils      datasets  methods   base
##
## other attached packages:
## [1] bookdown_0.22    broom_0.7.4      scales_1.1.1     png_0.1-7
## [5] kableExtra_1.3.2 knitr_1.31       mclust_5.4.7     forcats_0.5.1
## [9] stringr_1.4.0    dplyr_1.0.6      purrr_0.3.4      readr_1.4.0
## [13] tidyr_1.1.3      tibble_3.1.2     ggplot2_3.3.3    tidyverse_1.3.0
##
## loaded via a namespace (and not attached):
## [1] tidyselect_1.1.1 xfun_0.22        haven_2.3.1      colorspace_2.0-1
## [5] vctrs_0.3.8      generics_0.1.0   viridisLite_0.4.0 htmltools_0.5.1.1
## [9] yaml_2.2.1       utf8_1.2.1       rlang_0.4.11     pillar_1.6.1
## [13] glue_1.4.2       withr_2.4.2      DBI_1.1.1        dbplyr_2.1.0
## [17] modelr_0.1.8     readxl_1.3.1     lifecycle_1.0.0  munsell_0.5.0
## [21] gtable_0.3.0     cellranger_1.1.0 rvest_0.3.6      evaluate_0.14
## [25] fansi_0.5.0      Rcpp_1.0.6       backports_1.2.1  webshot_0.5.2
## [29] jsonlite_1.7.2   fs_1.5.0         hms_1.0.0        digest_0.6.27
## [33] stringi_1.5.3    grid_4.0.3       cli_2.5.0        tools_4.0.3
## [37] magrittr_2.0.1   crayon_1.4.1     pkgconfig_2.0.3  ellipsis_0.3.2
## [41] xml2_1.3.2       reprex_1.0.0     lubridate_1.7.9.2 assertthat_0.2.1
## [45] rmarkdown_2.6    httr_1.4.2       rstudioapi_0.13  R6_2.5.0
## [49] compiler_4.0.3
```

Introduction to data

Data for all analyses comes from samples collected during IODP Expedition 342 (Norris et al. 2014) with sample information presented in Table S1. All data needed for this supplement is provided as supplementary .csv files and the captions, including units, are outlined below.

Supplementary Table 1 - IODP Expedition 342 samples used in this study. ID= Study specific sample ID, Exp. = IODP expedition number, Top Int = Top interval (m), Bottom Int = Bottom interval (m), Sample ages were calculated based on an age–depth model constructed using available biostratigraphic and magnetostratigraphic data for Sites U1408 and U1410 ((Norris et al. 2014)(Yamamoto et al. 2018)(Cappelli et al. 2019)). Age calibrations from the 2012 geologic timescale were used for middle Eocene geomagnetic polarity reversals (GTS2012; (Gradstein et al. 2012)(Ogg, Hinnov, and Huang 2012)).

Supplementary Table 2 - Planktic foraminifera stable carbon and oxygen isotope values based on multi-specimen samples used for palaeoceanographic analysis. Sample ID is the study specific sample ID. Age is presented in millions of years ago (Ma), Depth habitat refers to the inferred depth habitat of the foraminifera in the water column based on stable isotope studies as summarized in the main text, Oxygen = $\delta^{18}\text{O}$ VPDB (‰) and Carbon = $\delta^{13}\text{C}$ VPDB (‰) .

Supplementary Table 3 - Morphological measurements for all 300 *Subbotina* individuals as presented in Figure S1. Sample ID is the study specific sample ID, Age is presented in million of years ago (Ma), Area is presented as μm^2 .

Supplementary Table 4 - Stable isotope and morphological measurements for a subset of 112 *Subbotina* individuals selected from the 300 measured morphologically (Table S4). Sample ID is the study specific sample ID, Age is presented in million of years ago (Ma), oxygen = $\delta^{18}\text{O}$ VPDB (‰), carbon = $\delta^{13}\text{C}$ VPDB (‰) and Area is presented as μm^2 . Note this table is a subset of those presented in Table S4 and includes species identification

Supplementary Table 5 - Stable carbon and oxygen isotope differences between *Globigerinatheka* and both *Subbotina* and *Catapsydrax*. Age is presented as millions of years ago (Ma), Difference is in per mille (‰), MECO indicates whether that sample is in the MECO (Y) or not (N)

Importing of data

Firstly we need to import Table S2-5 to set variables for easy recall. For the following set of analyses the age of the sample (Age) needs to be converted to a factor. Morphological traits shown in Figure S1 need also need to be renamed for easier recall in supplementary table 3-5. Size (Area) is measured on a different scale to stable isotopes therefore needs to be transformed on the log scale to aid model fit and centered in both Table 4 and 5.

```
bulk <- read_csv("Supplementary_Table_S2.csv") #Multispecimen Stable Isotopes
morph <- read_csv("Supplementary_Table_S3.csv") #Subbotina Morphometrics
morph_iso <- read_csv("Supplementary_Table_S4.csv") #Subbotina Stable Isotopes
diff <- read_csv("Supplementary_Table_S5.csv") #Multispecimen Isotope Gradients
```

```
# Supplementary table 5 renaming and factoring
MECO <- as.factor(diff$MECO)
isotope <- as.factor(diff$`Stable Isotope`)
```

```
# Supplementary table 4 renaming and factoring
morph$Age <- as.factor(morph$Age)
morph <- morph %>% rename(size = Area)
morph <- morph %>% rename(shape = Aspect_Ratio)
morph$size <- log(morph$size) - 10
```

```
# Supplementary table 5 renaming and factoring
morph_iso$Age <- as.factor(morph_iso$Age)
morph_iso <- morph_iso %>% rename(size = Area)
morph_iso <- morph_iso %>% rename(shape = `Aspect Ratio`)
morph_iso$size <- log(morph_iso$size) - 10
```

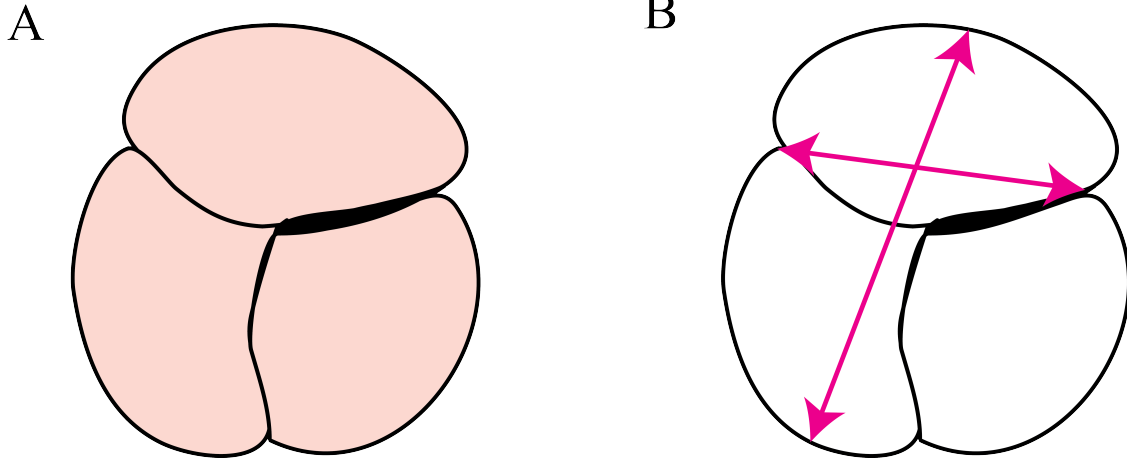


Figure S1: Schematic of the morphological traits used in this study. A) Test area, B) Test aspect (maximum feret diameter/minimum feret diameter).

Multi-specimen analysis

Multiple Linear Regression

To test if the MECO samples are different from the rest of the time slices. Y = MECO (40.1 Ma sample) , N = Not MECO (other 5 samples in analysis).

```
diff_lm <- lm(Difference ~ isotope + MECO, data=diff)
```

Table S6: Coefficient-Level estimates for a model fitted to estimate variation in isotopic differences, model formula: `lm(Difference ~ isotope + MECO, data=diff)`

Predictor	Coefficient	SE	t	<i>p</i>
(Intercept)	1.42	0.08	16.96	<0.001
isotopeoxygen	-0.71	0.11	-6.28	<0.001
MECOY	-0.56	0.15	-3.71	0.001

Individual analysis

Clustering Analysis

Individual *Subbotina* stable isotope measurements used to see if the samples can be separated in stable isotope space.

```
for_clust <- morph_iso %>% dplyr::select(Age, Oxygen, Carbon)
class <- for_clust$Age
X <- for_clust[, -1]
BIC <- mclustBIC(X)
mod1 <- Mclust(X, x=BIC)
```

Table S7: Assignment of individuals to either cluster 1 or 2 (columns) in each sample. Rows correspond to the inferred sample ages (Ma). Note there is no strong correlation between sample age and cluster assignment, implying no continuous trend, but that the two warmest samples (the oldest 43.5 Ma and the MECO at 40.14 Ma) both show very different relative frequencies to the cooler samples.

	1	2
38.5	13	5
39.56	13	7
40.14	3	15
41.31	14	3
42.1	15	5
43.5	1	18

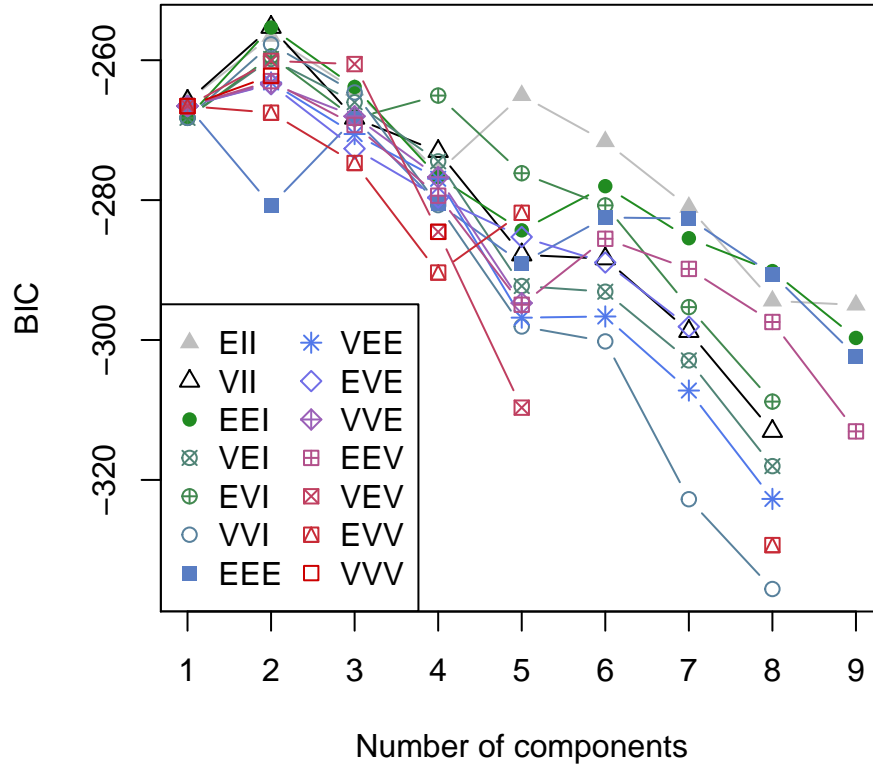


Figure S2: BIC plot from clustering analysis showing the best fitting model is a spherical, variable volume model (VII) with two clusters. Best fitting model in mclust is chosen by the model with the highest BIC, best fitting model in this analysis: BIC = -255

Analysis of variance

Stable Isotopes using subset of *Subbotina* geochemically analysed (Supplementary table 4)

These analyses also included species to understand whether it has any influence on stable isotopes before carrying species forward into further models. Species was found not to have any detectable impact ($F(6, 100) = 1.62$, $p > 0.05$) and is therefore not carried forward.

```
ox1<- aov(Oxygen ~ Age, data = morph_iso)
ox2<- aov(Oxygen ~ Age + Species, data = morph_iso)
#Species has no detectable impact so ox1 is carried forward to post-hoc test.
TUK_o <- TukeyHSD(ox1)

carb1 <- aov(Carbon ~ Age, data = morph_iso)
carb2 <- aov(Carbon ~ Age + Species, data = morph_iso)
#Species has no detectable impact so ox1 is carried forward to post-hoc test.
TUK_c <- TukeyHSD(carb1)
```

Morphology using supplementary table 3 with all 300 *Subbotina* measurements

```
aov_shape <- aov(shape ~ Age, data = morph)
aov_size <- aov(size ~ Age, data = morph)
```

Table S8: Summary table for one-way ANOVA: Oxygen~Age + Species

Term	df	Sum of squares	Mean sum of squares	F statistic	<i>p-value</i>
Age	5	8.27	1.65	12.26	<0.001
Species	6	1.31	0.22	1.62	0.149
Residuals	100	13.49	0.13		

Table S9: Summary table for one-way ANOVA: Carbon~Age + Species

Term	df	Sum of squares	Mean sum of squares	F statistic	<i>p-value</i>
Age	5	5.48	1.10	9.87	<0.001
Species	6	0.70	0.12	1.05	0.396
Residuals	100	11.10	0.11		

Table S10: Summary table for one-way ANOVA: Oxygen~Age showing significant differences ($p < 0.001$) between MECO samples (40.14 Ma) and those pre-MECO (bold) except at the beginning of the record 43.5 Ma

Term	df	Sum of squares	Mean sum of squares	F statistic	<i>p-value</i>
Age	5	8.27	1.65	11.84	<0.001
Residuals	106	14.80	0.14		

Table S11: Results from TUKEY-HD test following one-way ANOVA with oxygen

Term	Difference	Estimate (ppm)	Lower estimate (ppm)	Upper estimate (ppm)	<i>p-value</i>
Age	39.56-38.5	0.04	-0.32	0.39	>0.999
Age	40.14-38.5	-0.66	-1.02	-0.30	<0.001
Age	41.31-38.5	0.06	-0.31	0.43	0.997
Age	42.1-38.5	-0.15	-0.50	0.20	0.826
Age	43.5-38.5	-0.47	-0.83	-0.12	0.003
Age	40.14-39.56	-0.70	-1.05	-0.34	<0.001
Age	41.31-39.56	0.02	-0.33	0.38	>0.999
Age	42.1-39.56	-0.18	-0.53	0.16	0.631
Age	43.5-39.56	-0.51	-0.86	-0.16	<0.001
Age	41.31-40.14	0.72	0.35	1.09	<0.001
Age	42.1-40.14	0.51	0.16	0.86	<0.001
Age	43.5-40.14	0.19	-0.17	0.54	0.653
Age	42.1-41.31	-0.21	-0.57	0.15	0.544
Age	43.5-41.31	-0.53	-0.90	-0.17	<0.001
Age	43.5-42.1	-0.33	-0.67	0.02	0.080

Table S12: Summary table for one-way ANOVA: Carbon~Age showing significant differences ($p < 0.005$) between MECO samples (40.14 Ma) and pre-MECO samples (bold)

Term	df	Sum of squares	Mean sum of squares	F statistic	<i>p-value</i>
Age	5	5.48	1.10	9.84	<0.001
Residuals	106	11.80	0.11		

Multiple linear regression

With size

Table S13: Results from TUKEY-HD test following one-way ANOVA with carbon

Term	Difference	Estimate (ppm)	Lower estimate (ppm)	Upper estimate (ppm)	<i>p-value</i>
Age	39.56-38.5	0.46	0.15	0.78	<0.001
Age	40.14-38.5	0.69	0.36	1.01	<0.001
Age	41.31-38.5	0.26	-0.06	0.59	0.186
Age	42.1-38.5	0.13	-0.19	0.44	0.850
Age	43.5-38.5	0.26	-0.06	0.58	0.185
Age	40.14-39.56	0.23	-0.09	0.54	0.307
Age	41.31-39.56	-0.20	-0.52	0.12	0.480
Age	42.1-39.56	-0.33	-0.64	-0.03	0.024
Age	43.5-39.56	-0.20	-0.51	0.11	0.406
Age	41.31-40.14	-0.42	-0.75	-0.09	0.004
Age	42.1-40.14	-0.56	-0.87	-0.24	<0.001
Age	43.5-40.14	-0.43	-0.75	-0.11	0.002
Age	42.1-41.31	-0.14	-0.46	0.18	0.811
Age	43.5-41.31	-0.01	-0.33	0.32	>0.999
Age	43.5-42.1	0.13	-0.18	0.44	0.825

Table S14: Summary table for one-way ANOVA: Shape~Age

Term	df	Sum of squares	Mean sum of squares	F statistic	<i>p-value</i>
Age	5	0.10	0.02	3.25	0.007
Residuals	294	1.84	0.01		

Table S15: Summary table for one-way ANOVA: Size~Age

Term	df	Sum of squares	Mean sum of squares	F statistic	<i>p-value</i>
Age	5	1.02	0.20	3.18	0.008
Residuals	294	18.78	0.06		

```

c1.s <- lm(Carbon ~ Oxygen + size + Age, data= morph_iso)
c2.s <- lm(Carbon ~ (Oxygen + size + Age)^2, data=morph_iso)
anova(c1.s,c2.s)

```

Carbon with size

```

## Analysis of Variance Table
##
## Model 1: Carbon ~ Oxygen + size + Age
## Model 2: Carbon ~ (Oxygen + size + Age)^2
##   Res.Df    RSS Df Sum of Sq    F Pr(>F)
## 1     104 11.6425
## 2      93  9.1557 11    2.4868 2.2964 0.01554 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```
# Interactions with age are affecting so assume a constant size with age
c3.s <- lm(Carbon~ (Oxygen+size)^2 + Age, data = morph_iso)
anova(c2.s,c3.s)
```

```
## Analysis of Variance Table
##
## Model 1: Carbon ~ (Oxygen + size + Age)^2
## Model 2: Carbon ~ (Oxygen + size)^2 + Age
##   Res.Df    RSS   Df Sum of Sq    F Pr(>F)
## 1      93  9.1557
## 2     103 11.5736 -10    -2.418 2.4561 0.01189 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## c3 best fitting model that includes in interaction between carbon and size
```

```
## null device
##           1
```

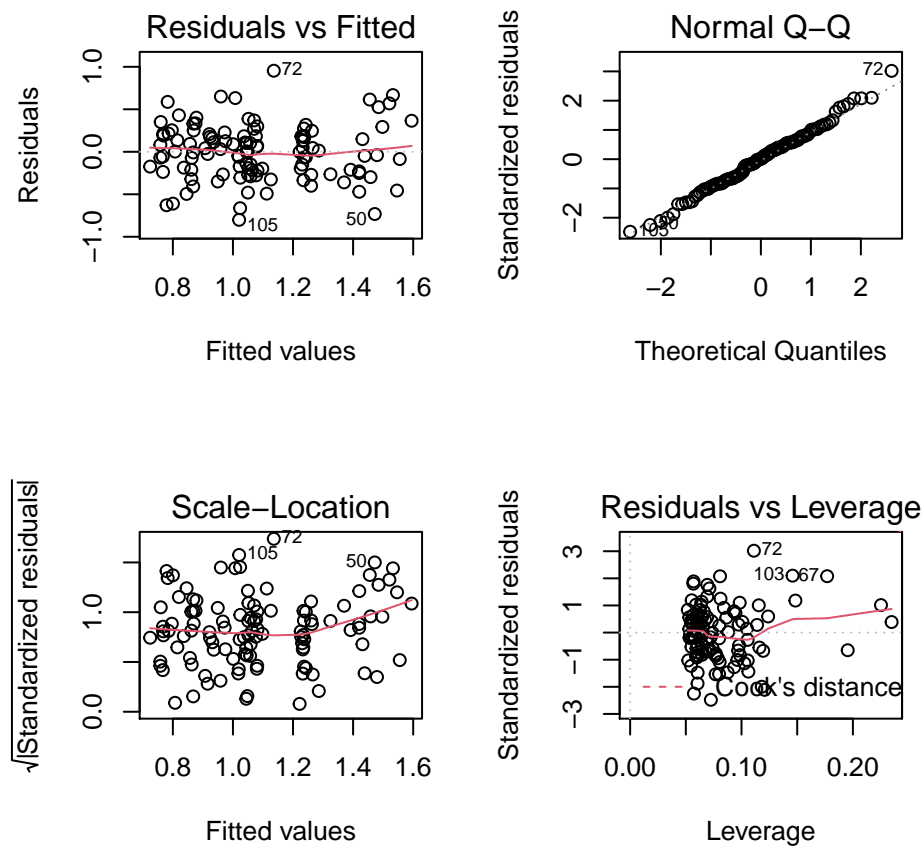


Figure S3: Diagnostic plot for model: $\text{Carbon} \sim (\text{Oxygen} + \text{size})^2 + \text{Age}$

Table S16: Model comparison for carbon linear regression models

	Model	Reisidual df	df	Residual sum of squares	Sum of squares	F statis- tic	p- value	AIC	Δ AIC
1	(Oxygen+size+Age) ²	93	9.16					77.38	0.00
2	Oxygen+size+Age	104	11.64	-11	-2.49	2.3	0.016	82.29	4.91
3	(Oxygen+size) ² +Age	103	11.57	1	0.07	0.7	0.405	83.63	6.25

Table S17: Model summary for best fitting model following mdoel reduction. Model: Carbon \sim (Oxygen+size)² + Age

Predictor	Coefficient	SE	t	p-value
(Intercept)	0.88	0.43	2.04	0.044
Oxygen	0.36	0.55	0.64	0.522
size	-0.10	0.30	-0.32	0.747
Age39.56	0.46	0.11	4.18	<0.001
Age40.14	0.60	0.13	4.59	<0.001
Age41.31	0.28	0.11	2.43	0.017
Age42.1	0.10	0.11	0.88	0.382
Age43.5	0.23	0.12	1.91	0.059
Oxygen:size	-0.29	0.37	-0.78	0.436

Remove all other variables to see if size without environmental variables has an impact on carbon

```
c1.s.1 <- lm(Carbon ~ size, data =morph_iso)
```

Table S18: Univariate model summary. Model: Carbon~size

Predictor	Coefficient	SE	t	p-value
(Intercept)	0.65	0.22	2.95	0.004
size	0.29	0.15	1.99	0.050

```
o1 <- lm(Oxygen ~ Carbon + size + Age, data=morph_iso)
o2 <- lm(Oxygen ~ (Carbon + size + Age)^2 , data= morph_iso)
anova(o1,o2)
```

Oxygen with size

```
## Analysis of Variance Table
##
## Model 1: Oxygen ~ Carbon + size + Age
## Model 2: Oxygen ~ (Carbon + size + Age)^2
##   Res.Df    RSS Df Sum of Sq    F Pr(>F)
## 1     104 14.705
## 2      93 11.582 11    3.1228 2.2795 0.01634 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
o3 <- lm(Oxygen ~ (Carbon*size)^2 + Age, data = morph_iso)
anova(o2,o3)
```

```
## Analysis of Variance Table
##
## Model 1: Oxygen ~ (Carbon + size + Age)^2
## Model 2: Oxygen ~ (Carbon * size)^2 + Age
##   Res.Df    RSS Df Sum of Sq    F Pr(>F)
## 1      93 11.582
## 2     103 14.643 -10   -3.0603 2.4573 0.01185 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## null device
##           1
```

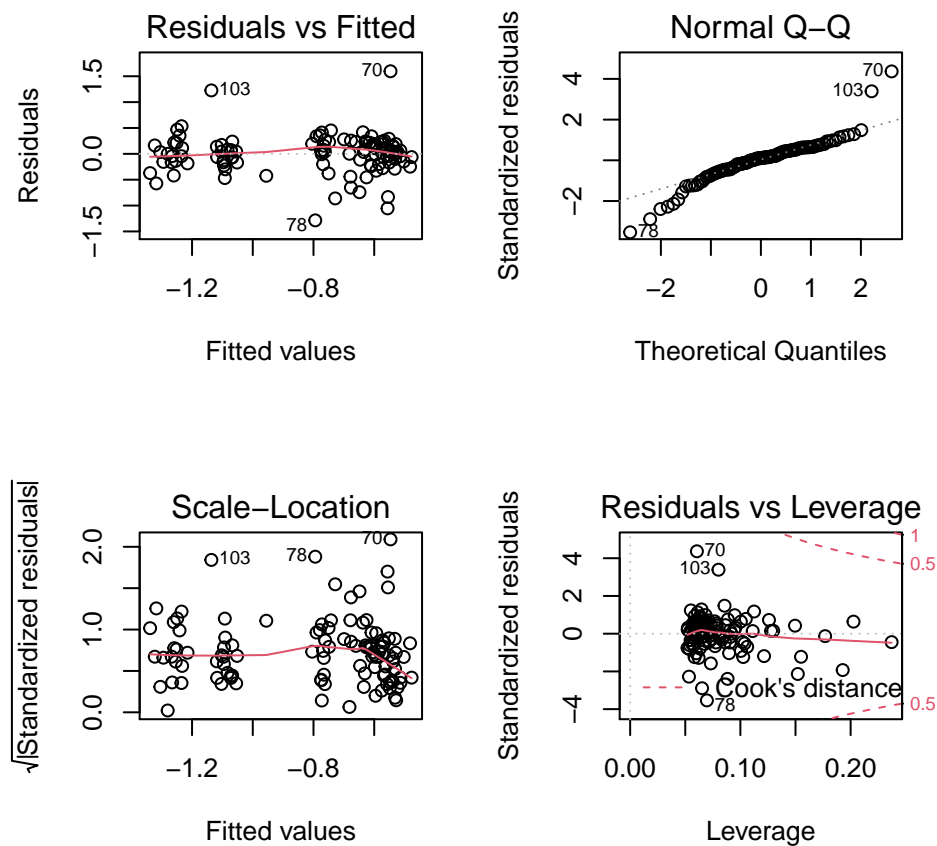


Figure S4: Diagnostic plot for model: $\text{Oxygen} \sim (\text{Carbon} + \text{size})^2 + \text{Age}$

Table S19: Model comparison for oxygen linear regression models

	Model	Residual df	df	Residual sum of squares	Sum of squares	F statis- tic	p- value	AIC	Δ AIC
1	(Carbon+size+Age) ²	93	11.58					103.71	0.00
2	Carbon+size+Age	104	14.71	-11	-3.12	2.28	0.016	108.45	4.74
3	(Carbon+size) ² +Age	103	14.64	1	0.06	0.50	0.480	109.97	6.26

Table S20: Model summary for best fitting model following mdoel reduction. Model: Oxygen \sim (Carbon+size)² + Age

Predictor	Coefficient	SE	t	p-value
(Intercept)	-0.15	0.64	-0.23	0.821
Carbon	-0.49	0.62	-0.80	0.426
size	-0.26	0.42	-0.61	0.546
Age39.56	0.08	0.13	0.61	0.540
Age40.14	-0.61	0.15	-4.05	<0.001
Age41.31	0.09	0.13	0.69	0.494
Age42.1	-0.14	0.13	-1.11	0.270
Age43.5	-0.44	0.13	-3.47	<0.001
Carbon:size	0.26	0.39	0.66	0.509

With weight

Now we replace size with weight to see if weight has any influence on the stable isotopes we measured. The best fitting models are similar for both size and weight. Where size was not detectable in either carbon (Table S17) or oxygen (Table S20) regressions, weight was detectable in carbon (Table S22). Furthermore, the model including weight for carbon explains more variation (lower residual sum of squares; 9.9264) than the model containing size (11.5736) . We include only size in the manuscript as this is a more commonly used measure.

```
c1.w <- lm(Carbon ~ Oxygen + Age, data= morph_iso)
c2.w <- lm(Carbon ~ (Oxygen + Weight + Age)^2, data=morph_iso)
anova(c1.w,c2.w)
```

Carbon with weight

```
## Analysis of Variance Table
##
## Model 1: Carbon ~ Oxygen + Age
## Model 2: Carbon ~ (Oxygen + Weight + Age)^2
##   Res.Df    RSS Df Sum of Sq    F    Pr(>F)
## 1     105 11.7216
## 2      93  8.0079 12    3.7137 3.594 0.0002031 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Interactions with age are affecting so assume a constant size with age
c3.w <- lm(Carbon~ (Oxygen+Weight)^2 + Age, data = morph_iso)
anova(c2.w,c3.w)
```

```
## Analysis of Variance Table
##
## Model 1: Carbon ~ (Oxygen + Weight + Age)^2
## Model 2: Carbon ~ (Oxygen + Weight)^2 + Age
##   Res.Df    RSS   Df Sum of Sq    F Pr(>F)
## 1      93 8.0079
## 2     103 9.9264 -10   -1.9184 2.228 0.02248 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

c3.w best fitting model that includes in interaction between carbon and weight

Table S21: Model comparison for carbon linear regression models

	Model	Reisidual df	df	Residual sum of squares	Sum of squares	F statis- tic	p- value	AIC	ΔAIC
1	(Oxygen+weight+Age) ²	93	8.01					62.38	0.00
2	Oxygen+weight+Age	105	11.72	-12	-3.71	3.59	0	81.05	18.67
3	(Oxygen+weight) ² +Age	103	9.93	2	1.80	10.42	0	66.43	4.05

```
o1.w <- lm(Oxygen ~ Carbon + Age , data=morph_iso)
o2.w <- lm(Oxygen ~ (Carbon + Weight + Age)^2 , data= morph_iso)
anova(o1.w,o2.w)
```

Oxygen with weight

```
## Analysis of Variance Table
```

Table S22: Model summary for best fitting model following mdoel reduction. Model: Carbon ~ (Oxygen+weight)² + Age

Predictor	Coefficient	SE	t	p-value
(Intercept)	0.23	0.19	1.21	0.229
Oxygen	-0.33	0.19	-1.71	0.089
Weight	34.07	12.84	2.65	0.009
Age39.56	0.46	0.10	4.48	<0.001
Age40.14	0.63	0.12	5.40	<0.001
Age41.31	0.31	0.11	2.95	0.004
Age42.1	0.09	0.10	0.90	0.368
Age43.5	0.16	0.12	1.34	0.184
Oxygen:Weight	14.84	15.46	0.96	0.339

```
##
## Model 1: Oxygen ~ Carbon + Age
## Model 2: Oxygen ~ (Carbon + Weight + Age)^2
##   Res.Df    RSS Df Sum of Sq      F   Pr(>F)
## 1     105 14.705
## 2      93 11.116 12    3.5891 2.5022 0.006829 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

o3.w <- lm(Oxygen ~ (Carbon*Weight)^2 + Age, data = morph_iso)
```

Table S23: Model comparison for oxygen linear regression models

	Model	Reisidual df	df	Residual sum of squares	Sum of squares	F statistic	p- value	AIC	ΔAIC
1	(CarbonWeight+Age) ²	103	13.27					98.97	0.00
2	Carbon+Age	93	11.12	10	2.16	1.8	0.070	99.11	0.14
3	(Carbon+Weight) ² +Age	105	14.71	-12	-3.59	2.5	0.007	106.45	7.48

Table S24: Model summary for best fitting model following model reduction. Model: Oxygen ~ (Carbon+Weight)² + Age

Predictor	Coefficient	SE	t	p-value
(Intercept)	-0.64	0.29	-2.24	0.027
Carbon	-0.31	0.27	-1.15	0.254
Weight	13.24	23.05	0.57	0.567
Age39.56	0.14	0.13	1.10	0.275
Age40.14	-0.47	0.15	-3.20	0.002
Age41.31	0.15	0.13	1.21	0.230
Age42.1	-0.12	0.12	-1.06	0.291
Age43.5	-0.46	0.12	-3.73	<0.001
Carbon:Weight	7.38	19.91	0.37	0.711

Including climate

With bulk oxygen isotopes to represent temperature changes at different water depths. First rename the variables for easier recall and select only oxygen isotopes before combining the datasets.

```
bulk$Age <- as.factor(bulk$Age)
bulk_oxygen <- bulk %>% dplyr::select(Genus, Oxygen, Age) %>% spread(Genus,Oxygen)
with_bulk <- merge(bulk_oxygen,morph_iso, by="Age")
subthermocline <- with_bulk$Catapsydrax
thermocline <- with_bulk$Subbotina
surface <- with_bulk$Globigerinatheka

bulk_o <- lm(Oxygen ~ (Carbon*size)^2 + subthermocline + thermocline + surface, data=with_bulk)
```

Table S25: Model summary for linear regression including multi-specimen analysis. Model: Oxygen \sim (Carbon*size)² + subthermocline + thermocline + surface

Predictor	Coefficient	SE	t	p-value
(Intercept)	-0.46	0.74	-0.63	0.529
Carbon	-0.49	0.61	-0.81	0.420
size	-0.25	0.42	-0.60	0.547
subthermocline	0.35	0.16	2.21	0.029
thermocline	0.44	0.19	2.30	0.024
surface	-0.51	0.47	-1.08	0.281
Carbon:size	0.26	0.39	0.66	0.509

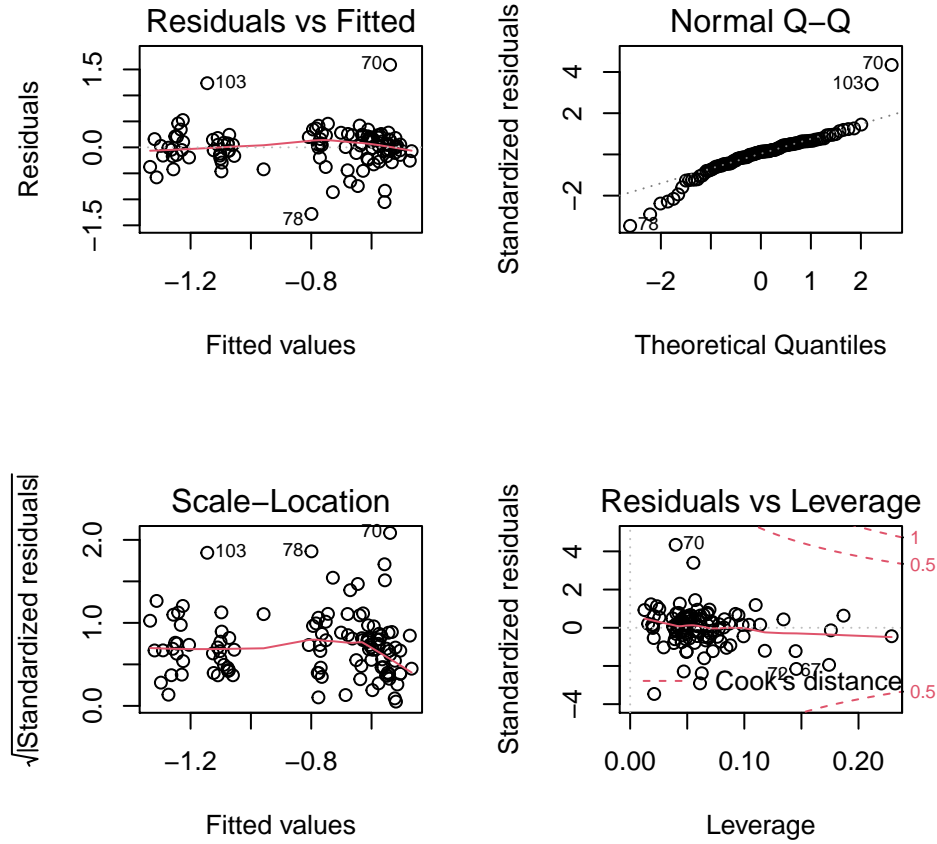


Figure S5: Diagnostic plot for model: Oxygen \sim (Carbon*size)² + subthermocline + thermocline + surface

Bibliography

- Cappelli, C., P. R. Bown, Thomas Westerhold, S. M. Bohaty, M. de Riu, V. Lobba, Y. Yamamoto, and C. Agnini. 2019. “The Early to Middle Eocene Transition: An Integrated Calcareous Nannofossil and Stable Isotope Record From the Northwest Atlantic Ocean (Integrated Ocean Drilling Program Site U1410).” *Paleoceanography and Paleoclimatology* 34 (12): 1913–30. <https://doi.org/10.1029/2019PA003686>.
- Fraley, Chris, Adrian E. Raftery, and Luca Scrucca. 2020. *Mclust: Gaussian Mixture Modelling for Model-Based Clustering, Classification, and Density Estimation*. <https://mclust-org.github.io/mclust/>.
- Gradstein, Felix M, James George Ogg, Mark D Schmitz, and Gabi M Ogg. 2012. *The Geologic Time Scale 2012*. elsevier.
- Norris, Richard D., P. A. Wilson, P. Blum, A. Fehr, C. Agnini, A. Bornemann, S. Boulila, et al. 2014. “Proceedings of the Integrated Ocean Drilling Program; Paleogene Newfoundland Sediment Drifts and MDHDS Test; Expedition 342 of the Riserless Drilling Platform, St. George, Bermuda, to St. John’s, Newfoundland (Canada); Sites U1402-U1411, 2 June-1 August 2.”
- Ogg, James G., Linda A. Hinnov, and Chunju Huang. 2012. “Cretaceous.” In, 1-2:793–853. Elsevier B.V. <https://doi.org/10.1016/B978-0-444-59425-9.00027-5>.
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Robinson, David, Alex Hayes, and Simon Couch. 2021. *Broom: Convert Statistical Objects into Tidy Tibbles*. <https://CRAN.R-project.org/package=broom>.
- Wickham, Hadley. 2019. *Tidyverse: Easily Install and Load the Tidyverse*. <https://CRAN.R-project.org/package=tidyverse>.
- Wickham, Hadley, and Dana Seidel. 2020. *Scales: Scale Functions for Visualization*. <https://CRAN.R-project.org/package=scales>.
- Xie, Yihui. 2021. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. <https://yihui.org/knitr/>.
- Yamamoto, Y., H. Fukami, W. Taniguchi, and P.C. Lippert. 2018. “Data Report: Updated Magnetostratigraphy for IODP Sites U1403, U1408, U1409, and U1410.” *Proceedings of the Integrated Ocean Drilling Program,. 342: College Station, TX (Integrated Ocean Drilling Program)*. <https://doi.org/10.2204/iodp.proc.342.2014>.
- Zhu, Hao. 2021. *kableExtra: Construct Complex Table with Kable and Pipe Syntax*. <https://CRAN.R-project.org/package=kableExtra>.