Available online at www.sciencedirect.com

ScienceDirect

Geochimica et Cosmochimica Acta 270 (2020) 160–183

Geochimica et Cosmochimica Acta

www.elsevier.com/locate/gca

A robust calibration of the clumped isotopes to temperature relationship for foraminifers

N. Meinicke [a,*], S.L. Ho [a,b], B. Hannisdal [a], D. Nürnberg [c], A. Tripati [d,e], R. Schiebel [f], A.N. Meckler [a]

[a] *Bjerknes Centre for Climate Research and Department of Earth Science, University of Bergen, Norway*
[b] *Institute of Oceanography, National Taiwan University, Taipei City, Taiwan*
[c] *GEOMAR Helmholtz Centre for Ocean Research Kiel, Germany*
[d] *Department of Earth, Planetary, and Space Sciences, Department of Atmospheric and Oceanic Sciences, Institute for the Environment and Sustainability, American Indian Studies Center, Center for Diverse Leadership in Science, UCLA, USA*
[e] *Institut Universitaire Européen de la Mer (IUEM), France*
[f] *Max Planck Institute for Chemistry, Mainz, Germany*

Received 13 May 2019; accepted in revised form 18 November 2019; Available online 26 November 2019

Abstract

The clumped isotope (Δ_{47}) proxy is a promising geochemical tool to reconstruct past ocean temperatures far back in time and in unknown settings, due to its unique thermodynamic basis that renders it independent from other environmental factors like seawater composition. Although previously hampered by large sample-size requirements, recent methodological advances have made the paleoceanographic application of Δ_{47} on small (<5 mg) foraminifer samples possible.

Previous studies show a reasonable match between Δ_{47} calibrations based on synthetic carbonate and various species of planktonic foraminifers. However, studies performed before recent methodological advances were based on relatively few species and data treatment that is now outdated. To overcome these limitations and elucidate species-specific effects, we analyzed 14 species of planktonic foraminifers in sediment surface samples from 13 sites, covering a growth temperature range of ~0–28 °C. We selected mixed layer-dwelling and deep-dwelling species from a wide range of ocean settings to evaluate the feasibility of temperature reconstructions for different water depths. Various techniques to estimate foraminifer calcification temperatures were tested in order to assess their effects on the calibration and to find the most suitable approach.

Results from this study generally confirm previous findings that there are no species-specific effects on the Δ_{47}-temperature relationship in planktonic foraminifers, with one possible exception. Various morphotypes of *Globigerinoides ruber* were found to often deviate from the general trend determined for planktonic foraminifers.

Our data are in excellent agreement with a recent foraminifer calibration study that was performed with a different analytical setup, as well as with a calibration based exclusively on benthic foraminifers. A combined, methodologically homogenized dataset also reveals very good agreement with an inorganic calibration based on travertines. Our findings highlight the potential of the Δ_{47} paleothermometer to be applied to recent and extinct species alike to study surface ocean temperatures as well as thermocline variability for a multitude of settings and time scales.

Keywords: Carbonate clumped isotopes; Foraminifera; Paleothermometry

1. INTRODUCTION

The investigation of past climate change relies to a large degree on reconstructions of ocean conditions. Not only is

* Corresponding author.
E-mail address: niklas.meinicke@uib.no (N. Meinicke).

https://doi.org/10.1016/j.gca.2019.11.022
0016-7037/

the ocean a major player in the global climate system responsible for storage and redistribution of heat; ocean sediments also constitute vital archives that can be used for climate reconstructions on various time scales (e.g. Zachos et al., 2001). Several geochemical proxies have been used to reconstruct ocean conditions, such as sea surface temperature, prior to the instrumental era. These include paleothermometers, for instance stable oxygen isotopes ($\delta^{18}O$) measured in calcareous tests of foraminifers (Urey, 1947; Epstein et al., 1951; Emiliani, 1966; Pearson, 2012), Mg/Ca ratios of foraminifers (Nürnberg et al., 1996; Lea et al., 1999), the unsaturation of organic ketone molecules (U^{K}_{37}) produced by marine nannoplankton (Brassell et al., 1986; Prahl et al., 1988) and the TEX_{86} proxy based on membrane lipids of archaea (Schouten et al., 2002).

Although each of these proxies is characterized by its individual strengths and weaknesses, two sources of uncertainty are particularly problematic for the application on longer time scales: First, temperature proxies such as $\delta^{18}O$ and Mg/Ca in foraminifers that depend on the fluid composition the signal is formed from require precise knowledge of the seawater composition ($\delta^{18}O_{seawater}$ and Mg/Ca, respectively) at the time of formation. Second, the recorded signal in most proxies is to some extent influenced by biological processes that need to be accounted for. These so-called vital effects can be species-specific (e.g. Bemis et al., 1998; Turich et al., 2007; Regenberg et al., 2009; Ho et al., 2014; Ezard et al., 2015; Jentzen et al., 2018; Polik et al., 2018), thus increasing the uncertainty of environmental reconstructions, particularly for data that is derived from extinct species. Clumped isotopes have the potential to circumvent both of these problems.

The carbonate clumped isotope method is based upon the fact that the abundance of doubly substituted carbonate ions containing both rare isotopes ^{18}O and ^{13}C increases with colder temperature. While carbonate formed under equilibrium conditions generally contains more bonds between two heavy isotopes than expected for a random (stochastic) distribution, the amount of this excess is temperature dependent (e.g. Bigeleisen and Mayer, 1947; Urey, 1947; Eiler and Schauble, 2004; Schauble et al., 2006). Temperature-dependent equilibrium constants determine the relative abundance of the $^{13}C^{18}O^{16}O^{16}O^{2-}$ isotopologue in isotope exchange reactions (e.g. Wang et al., 2004; Schauble et al., 2006).

The relative abundance of these multiply substituted (clumped) isotopologues in a carbonate can therefore be used as a measure for its formation temperature (e.g. Eiler and Schauble, 2004; Ghosh et al., 2006; Schauble et al., 2006; Eiler, 2007). An important aspect distinguishing this paleothermometer from other approaches is its independence from the isotopic composition of the aqueous solution it precipitated from (Ghosh et al., 2006; Eiler, 2007). The Δ_{47} value measured in acid-liberated CO_2 reflects the excess abundance (in ‰) of doubly substituted molecules relative to a random distribution that is calculated for each sample (Ghosh et al., 2006). The relationship between the normalized Δ_{47} value and carbonate formation temperature has been defined by theoretical, experimental, and empirical calibrations (e.g. Ghosh et al., 2006; Ghosh et al., 2007; Tripati et al., 2010; Grauel et al., 2013; Henkes et al., 2013; Zaarur et al., 2013; Wacker et al., 2014; Kele et al., 2015; Bonifacie et al., 2017; Kelson et al., 2017; Breitenbach et al., 2018; Peral et al., 2018; Petersen et al., 2019).

The first two studies of foraminifers (Grauel et al., 2013; Tripati et al., 2010) indicated that within their sample sets there is evidence that foraminifers follow a single Δ_{47}-temperature relationship. Hence these studies paved the way for the application of clumped isotope thermometry on foraminifers from sedimentary archives (e.g. Tripati et al., 2014). However, due to a relative lack of data for the cold temperature end of these foraminifer calibrations, coupled with recent developments in data processing and correction methods, additional studies have been underway: The more recent works (Breitenbach et al., 2018; Peral et al., 2018; Piasecki et al., 2019) have utilized progress made in community-wide efforts to facilitate inter-laboratory data comparison using the recalculation of absolute isotope ratios (Daëron et al., 2016; Schauer et al., 2016), and further redefinition of carbonate standards (Bernasconi et al., 2018), building on the definition of an "absolute" reference frame by Dennis et al. (2011), as well as newly developed analytical approaches (Hu et al., 2014).

Peral et al. (2018) and Piasecki et al. (2019) focused mostly on planktonic and benthic foraminifers, respectively, and both concluded that foraminifer-based Δ_{47}-T calibrations agree with inorganic calibrations. Although the two equations are statistically indistinguishable from each other temperatures calculated with these calibrations diverge towards the cold end of ocean temperatures (~0 °C) by more than 2.5 °C. A challenge for surface sediment-based calibrations using foraminifers is the difficulty in determining the actual calcification temperature, particularly for planktonic foraminifers.

Additionally, the small temperature range recorded in foraminifers poses a persisting problem for foraminifer Δ_{47}-T calibrations, because large datasets are required to extract an accurate linear relationship from the relatively large uncertainty of individual measurements (Fernandez et al., 2017). The relatively low signal to noise ratio compared to other geochemical proxies such as $\delta^{18}O$ might mask smaller, potentially species-specific, secondary effects. These potential secondary effects include pH or kinetic effects suggested for other types of marine biogenic carbonates/marine invertebrate organisms (e.g. Bajnai et al., 2018; Daeron et al., 2019; Davies and John, 2019). Divergences among older foraminifer-based calibrations can partly be explained by methodological or inter-laboratory differences such as the ^{17}O correction (Schauer et al., 2016; Bernasconi et al., 2018; Petersen et al., 2019), the choice of standards (Bernasconi et al., 2018), the acid digestion temperature (Defliese et al., 2015) and the common acid bath vs. the micro-volume approach (reviewed in Spencer and Kim, 2015).

These uncertainties underline the importance of further studies investigating method- and laboratory-specific differences as well as potential species effects. Ultimately the aim is to determine a common foraminifer calibration to enable a widespread application of clumped isotope analysis in for-

aminifers. At the same time, using clumped isotopes on foraminifers yields enormous potential for paleoceanographic reconstructions when coupled with other proxies: As highlighted by Breitenbach et al. (2018) and Evans et al. (2018), clumped isotope measurements can be combined with other foraminifer-based proxies to disentangle ocean temperature from other influences, such as changing seawater composition (e.g. past Mg/Ca changes, Evans et al. (2018)).

Similarly to other foraminifer-based proxies such as $\delta^{18}O$ (e.g. Mulitza et al., 1997), the Δ_{47} signal in foraminifers could be used to reconstruct temperature gradients in the water column by comparing species from different depth habitats. In the case of Δ_{47} using a single calibration is advantageous as it allows for a direct comparison of data from various species without any additional uncertainty introduced by individual, species-specific calibrations.

Here, we present new foraminifer-based Δ_{47} data analyzed on 14 species of planktonic foraminifers from surface sediments from 13 sites, covering a calcification temperature range of ~0–28 °C. We study potential species-specific effects on the clumped isotope measurements and compare our results to recent Δ_{47}-T calibrations. Data from our study are combined with data from Peral et al. (2018) and Piasecki et al. (2019) to determine a common foraminifer-based calibration and compare it to inorganic Δ_{47}-T calibrations. Finally, we evaluate whether temperature reconstructions for different depth levels of the water column are feasible with the reduced sample requirements of our analytical approach.

2. MATERIAL AND METHODS

2.1. Sites and samples

Surface sediment samples (mostly 0–1 cm, see Table 1) from 13 sites in the Nordic Seas, the North Atlantic, Indian Ocean, and Pacific Ocean were used in this calibration study (Fig. 1 and Table 1). The sites were selected to cover a wide range of oceanographic conditions and species of foraminifers.

Monospecific samples of surface- as well as deep-dwelling foraminifers covering a growth temperature range of ~0–28 °C (see Section 2.4) were measured. A total number of 43 samples from 14 species of planktonic foraminifers was selected (Tables 1 and 2), including *Globigerina bulloides*, *Globigerinoides conglobatus*, *Globigerinoides ruber* pink, *Globigerinoides ruber* white sensu lato (s.l.) and sensu stricto (s.s.), *Globorotalia hirsuta*, *Globorotalia inflata*, *Globorotalia menardii*, *Globorotalia truncatulinoides*, *Globorotalia tumida*, *Neogloboquadrina dutertrei*, *Neogloboquadrina pachyderma*, *Orbulina universa*, *Pulleniatina obliquiloculata*, *Trilobatus sacculifer*, *Trilobatus trilobus*.

Species characteristics and assumptions regarding their ecology are crucial to the interpretation of the Δ_{47} data, in particular when various species are compared to each other. Table 2 provides a summary of the species-specific characteristics considered, such as the presence of photobiotic symbionts, spatial and seasonal distribution, preferred habitat depth, the tendency to form gametogenetic calcite prior to reproduction and the accumulation of thick calcite crusts.

2.2. Sample preparation

All samples were wet-sieved over a 63 μm sieve and dried at ≤50 °C. The coarse fraction was then dry-sieved into size fractions of < 150 μm, 150–250 μm, 250–315 μm, 315–355 μm and 355–400 μm, 400–500 μm and >500 μm. For each sample, at least 2 mg of foraminifer tests of each species were collected under the microscope. The preservation of all individual specimens was assessed under the microscope and translucent specimens were preferred for analysis where available. Only fully intact pristine-looking tests were selected for analysis. Broken specimens as well as specimens containing substantial infillings, secondary calcite overgrowth or oxide coatings were excluded from analysis. Additionally, SEM images were used for selected samples to confirm that the foraminifers were well preserved. The size fractions used for the analysis were individually selected for each sample (Table 3). We attempted to obtain enough adult specimen of each species to allow an accurate isotope analysis while keeping the size range as narrow as possible in order to limit ontogenetic effects. Therefore, the size fraction in which most of the adult specimens at a given site were found was selected for analysis. Size fractions with a small number of very large individuals were excluded as well as smaller size fractions potentially containing juvenile specimens.

A modified version of the cleaning protocol for foraminiferal Mg/Ca analysis published by Barker et al. (2003) was used to remove contaminants. Batches of 200 to 1300 μg of foraminifer tests were cleaned at a time with each sample being represented by at least three individually cleaned sub-samples. The foraminifers were placed between two glass plates and carefully crushed in order to crack open all chambers and allow for subsequent cleaning. The crushed tests were sonicated three times for 30 s with DI water and rinsed with DI water after each sonication step. Samples were then sonicated once for 15 s with methanol and subsequently rinsed three times with DI water. After removing excess DI water, the cleaned samples were dried in an oven at ≤50 °C. The comparison of several cleaning steps and intensities (Piasecki et al., 2019; see also Grauel et al., 2013, Peral et al., 2018) led to the decision to leave out the H_2O_2 treatment suggested by Barker et al. (2003) to remove organic material for Mg/Ca analysis.

2.3. Measurement procedure

All measurements took place between November 2016 and March 2018 with replicate measurements of individual samples spread over several weeks to months. Measurements were performed using a Thermo Scientific MAT 253Plus mass spectrometer coupled to a KIEL IV carbonate device (Thermo Fisher Scientific, Bremen, Germany) equipped with a Porapak trap to capture organic contaminants (Schmid and Bernasconi, 2010). The Porapak trap was operated at −20 °C during the measurement. Between runs, the trap was heated to 120 °C for at least one hour for cleaning. In the Kiel device, each aliquot is reacted individually with phosphoric acid at 70 °C.

We measured 15 to 30 (average n = 19) aliquots (100–130 μg each) for every sample. Average values for stable

Table 1
Sites from which planktonic foraminiferal specimens used in this study were selected.

Station	Latitude °N	Longitude °E	Region	Depth [m]	Depth in the core [cm]	Age [ka BP]	Species
GS15-198-63MC	70.5	−2.8	Nordic Seas	2995	0–1	<625 ± 20 conventional ^{14}C age	*N. pachyderma*
GS15-198-38MC	70.1	−17.7	Denmark Strait	1610	0–1	<410 ± 15 conventional ^{14}C age	*N. pachyderma*
GS15-198-62MC	70.0	−13.6	Iceland Plateau	1423	0–1	<2995 ± 15 conventional ^{14}C age	*N. pachyderma*
GS06-144-19MC	63.8	5.2	Nordic Seas	922	0–7	recent (Yu et al., 2013)	*G. bulloides, G. inflata*
CD107 A ML 5A	52.9	−16.9	North Atlantic	3569	surface sediment		*G. bulloides, G. hirsuta, G. truncatulinoides, O. universa*
CD94 17B (OMEX)	48.9	−11.8	North Atlantic	1484	surface sediment		*G. bulloides, G. inflata, G. truncatulinoides, O. universa*
KL88	34.8	−27.7	North Atlantic	2060	surface sediment	subrecent	*G. bulloides, G. inflata, G. ruber* white s.l., *G. truncatulinoides*
CD145 A150	23.3	66.7	Arabian Sea	151	0–1		*N. dutertrei*
SO164-25–3	14.7	−59.7	Caribbean /North Atlantic	2720	0–1	1915 ± 30 (Regenberg et al., 2006)	*G. conglobatus, G. ruber* pink, *G. ruber* white s.s., *G. ruber* white s.l., *P. obliquiloculata, T. trilobus*
OJP2016 MW0691 1.5BC11	−1.0	157.8	Ontong Java Plateau	2016	0–5		*T. trilobus*
WIND 33B	−11.2	58.8	Indian Ocean	2871	0–2		*G. menardii, G. ruber* white s.s., *G. tumida, N. dutertrei, O. universa, P. obliquiloculata, T. sacculifer, T. trilobus*
SO225-53–1	−13.5	−162.1	Manihiki Plateau	3154	0–1	6230 ± 50 (Raddatz et al., 2017)	*G. conglobatus, G. ruber* white s.s., *G. tumida, O. universa, P. obliquiloculata, T. sacculifer, T. trilobus*
SO213-84–2	−45.1	174.6	South Pacific	992	0–1	4952 ± 238 (Molina-Kescher et al., 2014)	*G. bulloides, G. inflata, G. truncatulinoides*

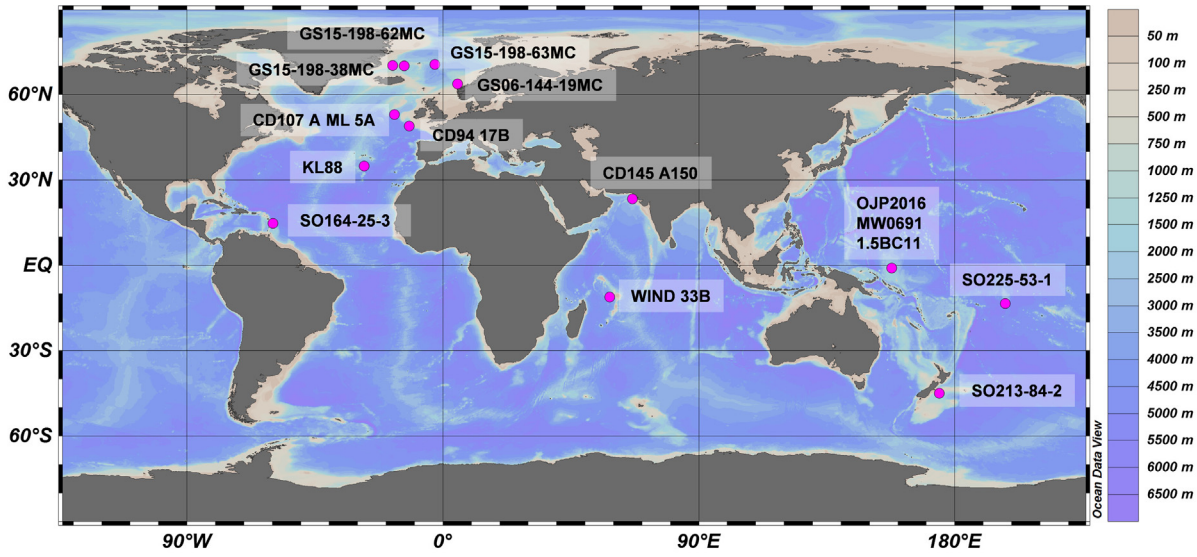

Fig. 1. Bathymetric chart generated using Ocean Data View (ODV, Schlitzer, 2018) showing surface sediment locations (pink filled circles), from which foraminifer specimens were selected. Bathymetric data from GlobHR (reference available in Ocean Data View).

carbon ($\delta^{13}C$) and oxygen isotopes ($\delta^{18}O$) as well as Δ_{47} were then calculated and used for the calibration. Samples were measured using the long-integration dual-inlet (LIDI) method described by Hu et al. (2014). This method measures the sample and reference gas separately with decreasing pressure from a micro-volume. Samples were measured first for 400 seconds with signals typically decreasing from ~16 V to ~10 V (m/z = 44). Afterwards the reference gas was adjusted to the same initial pressure and measured accordingly. The shot noise limit for these intensities and integration times is 0.03‰ when applying a typical scale decompression factor for this system.

Peak scans (varying high voltage between 9.4 and 9.6 kV) at m/z 44 intensities of 5, 10, 15, 20 and 25 V were performed once a day for the pressure baseline correction following Bernasconi et al. (2013) and Meckler et al. (2014). This and subsequent corrections were applied using the Easotope software package (John and Bowen, 2016). The "Brand parameters" suggested by Daëron et al. (2016) and Schauer et al. (2016) were used for the ^{17}O correction. In every run (maximum 46 aliquots), the sample measurements were bracketed by five blocks consisting of the four ETH carbonate standards ETH1 to ETH4 using the values reported in Bernasconi et al. (2018). Three of these standards were used to transfer the results into the absolute reference frame (Dennis et al., 2011), which corrects the measurements for offsets and scale compression, while the fourth standard was treated like a sample to monitor the corrections applied to the data. In addition, the long-term averages of ETH 1 and 2 were used to monitor the pressure baseline correction which should result in the same Δ_{47} values (Bernasconi et al., 2018). Baseline-corrected Δ_{48} values were used as a contamination monitor. No contamination was detected in any of the samples.

The average long-term reproducibility (1SD) of Δ_{47} measured in the carbonate standards after correction varies from 0.031‰ to 0.038‰ (see Appendix Table A1). Each replicate measurement was corrected using a total number of 60–80 standard measurements from the same and adjacent days. The exact number was chosen according to the instrument stability (see Piasecki et al., 2019 for more information). In addition to correcting for instrumental drift using carbonate standards, we distributed replicate measurements of all samples over long time intervals of up to several months to ensure that aliquots from as many samples as possible were measured in parallel.

The ETH standard values adopted in this study were reported by Bernasconi et al. (2018), who used a +0.062‰ correction (Defliese et al., 2015) for differences in acid fractionation between digestion at 70 °C and the classical 25 °C digestion temperature. Applying the recently updated acid fractionation correction of 0.066‰ for this temperature difference (Petersen et al., 2019) would increase all of our Δ_{47} values by 0.004‰. Should the ETH standard values be updated in the future it is possible to recalculate the values from this study using the replicate level raw data that is provided in the EarthChem database (https://doi.org/doi:10.1594/IEDA/111435).

2.4. Foraminifer calcification temperature estimates

In order to establish a calibration relating the Δ_{47} signal in planktonic foraminifers to water temperature, the calcification temperature for each species at each site needs to be estimated. Since our sample set comprises a large number of different species from a wide range of geographical regions, the estimation of calcification temperatures is subject to a number of uncertainties. Calcification temperatures were hence calculated using different approaches (Method 1 to 3) in order to find the optimal solution.

Method 1: If calcification depths and possible seasonality effects are known for the species and geographical regions, the water temperature can be taken from reanalysis data presented in the World Ocean Atlas (WOA; Locarnini

Table 2
Summary of species-specific characteristics for the planktonic foraminifers analyzed in this study (Schiebel and Hemleben, 2017, and references therein).

Species	Spinose	Symbiont-bearing	Typical habitat	Spacial distribution	Seasonality	Gametogenic calcite reported	Other secondary calcite reported	Remarks
Globigerina bulloides	yes	yes	mixed layer	temperate to sub-polar waters and upwelling regions	found year round with lower abundances in summer, typical of the spring bloom	yes	no	
Globigerinoides conglobatus	yes	yes	mixed layer	tropical and subtropical waters	more abundant in fall	yes	no	
Globigerinoides ruber pink	yes	yes	mixed layer	tropical to subtropical waters in the Atlantic Ocean	most abundant at highest T	no	no	
Globigerinoides ruber white s.l.	yes	yes	mixed layer	tropical to subtropical waters	low in stratified tropical waters	no	no	
Globigerinoides ruber white s.s.	yes	yes	upper mixed layer	tropical to subtropical waters	low in stratified tropical waters	no	no	
Globorotalia hirsuta	no	no	subsurface	temperate to subtropical waters	unclear	no	yes	
Globorotalia inflata	no	no	mixed layer to subsurface	subtropical to subpolar waters, hydrologic fronts	depending on front dynamics	no	yes	only specimens without calcite veneer selected for analysis
Globorotalia menardii	no	yes	thermocline	tropical to subtropical waters	low in stratified tropical waters	no	yes	
Globorotalia truncatulinoides	no	no	subsurface	tropical to temperate waters	more abundant in winter	no	yes	
Globorotalia tumida	no	no	thermocline	tropical to subtropical waters	low in stratified tropical waters	no	yes	
Neogloboquadrina dutertrei	no	yes	mixed layer to thermocline	tropical to temperate waters	low in stratified tropical waters	no	yes	
Neogloboquadrina pachyderma	no	no	mixed layer	polar waters	polar summer	no	yes	only left-coiling specimens selected for analysis
Orbulina universa	yes	yes	mixed layer to thermocline	tropical to temperate waters	summer in temperate waters	yes	no	
Pulleniatina obliquiloculata	no	no	lower mixed layer to thermocline	tropical to subtropical waters	more abundant in winter	yes	no	only smooth tests without gametogenic calcite selected for analysis
Trilobatus sacculifer	yes	yes	mixed layer	tropical to subtropical waters	low in stratified tropical waters	yes	no	
Trilobatus trilobus	yes	yes	mixed layer	tropical to subtropical waters	low in stratified tropical waters	yes	no	

Table 3
The specifications (site, species, size fraction, number of replicates) and average isotopic compositions (Δ_{47}, $\delta^{13}C_{calcite}$ and $\delta^{18}O_{calcite}$) of individual samples with corresponding $\delta^{18}O_{seawater}$ data (LeGrande and Schmidt, 2006), assumed calcification depth (Rippert et al., 2016, Schiebel and Hemleben, 2017, and references therein) and four different estimates of calcification temperature based on the World Ocean Atlas 2009 (Method 1), Shackleton (1974) (Methods 2 and 3) and Kim and O'Neil (1997) (Method 2). The calculated apparent calcification depth (ACD) is given for the Method 3.

Site and species	Size fraction [μm]	N*	Δ_{47} [‰]	SE	$\delta^{13}C_c$ [‰]	SE	$\delta^{18}O_c$ [‰]	SE	Assumed CD [m]*	$\delta^{18}O_{sw}$ [‰]	SE	T M1 [°C]*	SD	T S74 M2 [°C]*	SE	T S74 M3 [°C]*	SE	ACD M3 [m]*	T K97 M2 [°C]*	SE
CD94 17B (OMEX)																				
G. bulloides	250–315	16	0.706	0.006	−0.440	0.030	0.72	0.02	0–100	0.57	0.20	12.9	1.9	15.5	1.0	13.5	0.2	15	14.1	1.2
G. truncat.	355–500	15	0.713	0.008	1.050	0.040	1.48	0.03	200–500	0.55	0.20	11.0	0.3	12.3	1.2	11.6	0.8	204	10.6	1.4
G. inflata	355–400	15	0.710	0.008	0.960	0.010	1.35	0.03	0–500	0.56	0.20	12.0	1.6	13.0	1.2	12.3	1.0	115	11.3	1.4
O. universa	355–400	15	0.707	0.008	2.080	0.040	1.18	0.03	20–150	0.57	0.20	12.2	1.5	13.6	1.3	13.1	0.8	46	12.1	1.5
CD107 A ML 5A																				
G. bulloides	250–355	30	0.714	0.006	−0.410	0.020	1.38	0.04	0–100	0.52	0.20	11.9	1.5	12.6	1.6	11.6	0.8	99	10.9	1.8
*G. truncat.**	355–500	15	0.713	0.008	1.180	0.020	1.58	0.02	200–500	0.48	0.20	10.3	0.3	11.7	1.1	11.1	0.9	186	9.9	1.2
O. universa	355–500	15	0.704	0.005	2.030	0.020	1.05	0.02	20–150	0.52	0.20	11.5	1.2	14.0	1.2	12.1	0.5	32	12.4	1.3
G. hirsuta	>355	19	0.723	0.007	1.170	0.020	2.08	0.02	500–700	0.44	0.20	9.4	0.4	9.5	1.2	9.6	1.1	497	7.5	1.2
CD145 A150																				
N. dutertrei	355–500	20	0.679	0.008	1.420	0.040	−1.12	0.02	25–125	0.72	0.21	23.4	2.1	23.5	1.2	23.3	0.7	62	23.6	1.4
GS06-144-19MC																				
G. bulloides	250–315	18	0.713	0.006	−0.410	0.030	1.61	0.02	0–100	0.29	0.22	8.6	1.6	10.8	1.2	9.0	0.2	12	8.9	1.3
G. inflata	250–315	18	0.719	0.007	0.920	0.010	1.95	0.01	0–500	0.37	0.23	7.6	1.8	9.7	1.1	8.7	0.4	39	7.7	1.2
GS15-198-38MC																				
N. pachyderma	150–250	20	0.755	0.008	0.540	0.010	3.60	0.01	0–50	−0.17	0.23	0.2	1.5	1.0	1.2	0.4	0.3	111	−1.2	1.2
GS15-198-62MC																				
N. pachyderma	150–250	28	0.756	0.006	0.770	0.010	3.66	0.01	0–50	0.01	0.21	0.8	1.5	1.5	1.1	0.8	0.5	65	−0.7	1.0
GS15-198-63MC																				
N. pachyderma	250–315	21	0.750	0.006	0.390	0.010	2.78	0.01	0–50	0.28	0.20	4.2	1.7	6.1	1.1	4.3	0.3	18	3.9	1.1
KL88																				
G. ruber w. s.l.	315–355	18	0.715	0.008	1.050	0.030	0.61	0.02	30–50	1.03	0.20	19.4	1.9	17.7	1.2	17.0	1.1	120	16.7	1.3
G. bulloides	250–315	15	0.708	0.006	−1.180	0.040	1.41	0.03	0–100	1.01	0.21	19.3	2.3	14.4	1.4	13.0	1.1	414	13.0	1.5
*G. truncat.**	355–400	16	0.723	0.008	0.930	0.030	1.64	0.04	200–500	0.71	0.22	13.9	1.2	12.3	1.5	11.8	1.1	537	10.6	1.6
G. inflata	355–400	16	0.721	0.007	0.800	0.030	1.74	0.03	0–500	0.88	0.25	16.9	3.0	12.6	1.6	11.3	1.0	595	10.9	1.7
OJP2016 MW0691 1.5BC11																				
T. trilobus	355–400	16	0.669	0.005	2.220	0.030	−2.05	0.02	75–150	0.36	0.21	26.4	1.9	25.7	1.2	26.0	0.8	123	26.4	1.5

(*continued on next page*)

Table 3 (*continued*)

Site and species	Size fraction [µm]	N*	Δ_{47} [‰]	SE	$\delta^{13}C_c$ [‰]	SE	$\delta^{18}O_c$ [‰]	SE	Assumed CD [m]*	$\delta^{18}O_{sw}$ [‰]	SE	T M1 [°C]*	SD	T S74 M2 [°C]*	SE	T S74 M3 [°C]*	SE	ACD M3 [m]*	T K97 M2 [°C]*	SE
SO164-25-3																				
G. ruber w. s.s.	315–355	20	0.653	0.006	1.190	0.050	−2.29	0.02	0–30	0.72	0.20	27.4	0.8	28.1	1.2	27.3	0.2	20	29.5	1.6
G. ruber w. s.l.	250–315	18	0.682	0.006	0.670	0.050	−2.21	0.03	30–50	0.81	0.20	27.1	0.7	28.1	1.3	27.3	0.2	22	29.5	1.7
G. ruber pink	315–355	18	0.678	0.008	1.450	0.020	−2.18	0.02	0–50	0.74	0.21	27.3	0.8	27.8	1.2	27.3	0.2	24	29.0	1.5
T. trilobus	315–355	17	0.665	0.009	1.380	0.050	−1.85	0.04	50–100	0.98	0.23	26.0	0.9	27.4	1.6	27.0	0.4	42	28.6	2.0
P. obliquiloculata	355–500	16	0.664	0.005	1.230	0.040	−0.91	0.02	100–125	1.13	0.20	24.2	0.8	24.2	1.2	24.2	0.8	113	24.5	1.4
G. conglobatus	>400	16	0.670	0.008	2.300	0.060	−1.30	0.03	75–125	1.09	0.21	24.8	1.1	25.6	1.4	25.5	0.7	87	26.3	1.7
SO213-84-2																				
G. bulloides	315–355	30	0.729	0.005	0.400	0.030	2.02	0.03	100–150	0.01	0.20	8.3	0.4	8.0	1.4	7.7	0.9	271	6.0	1.5
*G. truncat.**	315–355	16	0.731	0.007	1.010	0.030	2.17	0.03	200–500	−0.01	0.20	7.4	0.6	7.3	1.3	7.3	0.9	361	5.2	1.3
G. inflata	315–355	20	0.729	0.008	1.240	0.010	1.90	0.02	0–500	0.04	0.20	8.8	0.9	8.6	1.2	8.1	1.0	214	6.6	1.3
SO225-53-1																				
G. ruber w. s.s.	250–355	17	0.684	0.007	1.750	0.050	−1.59	0.02	50–150	0.68	0.21	26.6	1.4	25.1	1.2	25.4	0.9	132	25.7	1.5
T. trilobus	250–355	22	0.683	0.006	1.850	0.040	−1.40	0.03	75–150	0.71	0.21	26.2	1.2	24.6	1.5	24.6	0.9	147	25.0	1.8
T. sacculifer	>355	15	0.676	0.005	2.480	0.060	−1.37	0.04	75–150	0.71	0.21	26.2	1.2	24.4	1.4	24.5	0.9	150	24.8	1.8
P. obliquiloculata	355–500	21	0.678	0.006	1.500	0.030	−0.24	0.06	150–200	0.72	0.20	23.3	1.2	19.9	1.9	19.1	1.1	253	19.3	2.2
O. universa	>355	16	0.672	0.007	3.170	0.050	−1.26	0.06	50–200	0.68	0.21	25.8	2.1	23.8	1.8	24.0	0.9	160	24.1	2.2
G. conglobatus	355–500	20	0.690	0.007	2.360	0.020	−1.23	0.04	100–200	0.72	0.20	24.8	1.8	23.9	1.6	23.9	0.9	163	24.1	2.0
G. tumida	>355	16	0.688	0.006	2.280	0.040	0.10	0.06	125–300	0.61	0.25	21.6	3.5	18.1	2.1	17.3	1.1	282	17.2	2.4
WIND 33B																				
G. ruber w. s.s.	250–355	19	0.686	0.008	1.590	0.030	−1.74	0.02	0–30	0.40	0.20	26.9	1.2	24.6	1.1	24.6	1.0	58	25.0	1.4
T. trilobus	250–355	23	0.671	0.005	1.670	0.030	−1.51	0.02	50–100	0.32	0.21	23.0	2.1	23.4	1.2	23.5	1.0	71	23.6	1.4
T. sacculifer	>355	25	0.675	0.006	2.350	0.030	−1.55	0.02	50–100	0.32	0.21	23.0	2.1	23.5	1.2	23.7	1.0	69	23.7	1.5
N. dutertrei	355–400	20	0.678	0.005	1.720	0.020	−0.64	0.03	20–125	0.34	0.21	23.5	3.2	20.0	1.4	19.7	0.8	112	19.4	1.7
P. obliquiloculata	355–500	21	0.680	0.005	0.990	0.010	−0.80	0.03	100–125	0.31	0.21	19.6	2.5	20.5	1.4	20.4	0.8	104	20.0	1.7
O. universa	>400	19	0.679	0.006	2.370	0.030	−1.37	0.03	20–150	0.34	0.21	22.6	3.8	22.9	1.4	22.8	1.0	78	22.9	1.7
G. menardii	>400	20	0.702	0.006	1.520	0.020	0.40	0.06	20–100	0.35	0.21	24.5	2.5	15.9	1.8	15.6	0.9	182	14.6	2.1
G. tumida	>355	20	0.700	0.005	1.660	0.020	0.71	0.05	125–300	0.29	0.20	15.1	2.5	14.4	1.6	14.2	0.9	215	12.9	1.8

* N = number of replicate measurements, CD = calcification depth, S74 = Shackleton 1974, K97 = Kim and O'Neil, 1997, M1-3 = Methods 1–3, *G. truncat.* = *Globorotalia truncatulinoides*.

et al., 2010). Due to the fact that typical foraminifer water depth habitats vary over time, both dependent on the availability of prey and ontogeny (Schiebel and Hemleben, 2017), the environmental signal recorded by the bulk foraminifer tests is rather an average across the entire life cycle of individuals and assemblages (e.g. Deuser and Ross, 1989). We therefore used published apparent calcification depths from studies utilizing other temperature proxies such as oxygen isotopes and Mg/Ca on planktonic foraminifers (e.g. Schiebel and Hemleben, 2017 and references therein) as basis for atlas-based calcification temperatures (Table 3). This approach suffers from insufficient information regarding foraminiferal calcification depths for individual regions and species. Also, the temperature information derived from the World Ocean Atlas may not provide the same accuracy everywhere because the data is interpolated to all ocean regions and standard depth levels. Here we used the annual mean water temperature of the assumed calcification depth as basis for further calculations. Seasonal temperature variability was factored into the uncertainty calculations. For Method 1, the overall temperature uncertainty is given by the standard deviation of all monthly temperatures at the assumed calcification depth of each species at a given site.

Method 2: Calcification temperatures can be derived from oxygen isotope ($\delta^{18}O$) measurements of each foraminifer sample by applying empirical calibration equations, which relate the $\delta^{18}O$ of the calcite tests ($\delta^{18}O_{calcite}$) to water temperature (e.g. Shackleton et al., 1973). For this approach, the $\delta^{18}O$ of the seawater ($\delta^{18}O_{seawater}$) is needed, which we obtained for the assumed calcification depths from the database of LeGrande and Schmidt (2006). Due to species-specific disequilibrium effects (suggested by Urey, 1947; Shackleton et al., 1973), specific $\delta^{18}O$-T calibrations have been derived for certain species and ocean regions (reviewed in Pearson, 2012). However, as such calibrations are only available for some of the species studied here, we decided to apply the multi-species $\delta^{18}O$-temperature equations of Kim and O'Neil (1997) and Shackleton (1974) for the entire dataset, acknowledging that some of the reconstructed temperatures may be biased by species-specific effects. The extent of such effects, however, is still a matter of debate (Niebler et al., 1999; Schiebel and Hemleben, 2017). We tested the sensitivity of our results to corrections for species-specific differences using the available information (Appendix Table A2). Applying species-specific $\delta^{18}O$ corrections led to a calibration line within the error of the uncorrected $\delta^{18}O$ data (Table A3). Furthermore, using species-specific corrections hardly changes the influence of individual species on the slope of the calibration line (Fig. A1). Because of the uncertainty introduced by the large spread of published values for species-specific corrections and the lack of improvement to our fit when applying a correction, we decided against applying any correction to the $\delta^{18}O_{calcite}$ data used for the calibration. Temperature estimates from two commonly used calibrations (Shackleton, 1974, equation D; Kim and O'Neil, 1997 modified by Bemis et al., 1998 Table 1) were compared. Following the recommendation of Bemis et al. (1998) and Pearson (2012), factors of 0.20‰ (Epstein et al., 1953) and 0.27‰ (Hut, 1987) were used to convert from VSMOW to VPDB for the $\delta^{18}O$-T calibrations of Shackleton (1974) and Kim and O'Neil (1997), respectively. The uncertainty of each calcification temperature estimate was calculated as a combination of several individual uncertainties: We used the standard deviation of the measured $\delta^{18}O_{calcite}$ values to account for the variability of the sample material and the uncertainty of the isotope measurement. Mean $\delta^{18}O_{seawater}$ values were calculated for the depth intervals that were assumed to best represent the calcification depths (Table 3, same as in Method 1) of the samples. The standard deviation of $\delta^{18}O_{seawater}$ over this depth interval was taken as uncertainty. An additional 0.2‰ were added to account for the uncertainty introduced by the gridded dataset (following Peral et al. (2018)).

Method 3: In order to avoid relying on assumed depth habitats, calcification temperatures were also estimated from hypothetical $\delta^{18}O$ depth profiles of calcite formed in the water column ($\delta^{18}O_{calculated}$). For this approach, the WOA temperature data are used in combination with a published $\delta^{18}O$-T calibration (Shackleton, 1974) to produce vertical profiles of hypothetical $\delta^{18}O$ to which the measured foraminiferal $\delta^{18}O_{calcite}$ is compared, in order to determine the apparent calcification depth (ACD) and subsequently the corresponding WOA temperature. This approach has previously been applied in Mg/Ca-temperature calibration studies on foraminifers (e.g. Groeneveld and Chiessi, 2011). The $\delta^{18}O_{calculated}$ of calcite is calculated for the entire water column at each sample site, combining WOA water temperature data (Locarnini et al., 2010) and $\delta^{18}O_{seawater}$ values (LeGrande and Schmidt, 2006). Method 3 has the advantage that no assumptions regarding habitat depth are needed, neither for the atlas-derived water temperature nor for the $\delta^{18}O$ of the water. This way, seasonal or ontogenetic variations in the calcification depth are accounted for as well.

If the measured $\delta^{18}O_{calcite}$ of a sample was not found in the $\delta^{18}O_{calculated}$ values (calculated temperatures warmer/colder than the observed maximum/minimum annual mean water temperature), the annual mean water temperature at 0 m depth was used. Therefore, extreme temperature cases not represented in WOA are excluded with this method, eliminating temperatures warmer or colder than observed at these sites. Nonetheless, Method 3 is associated with several uncertainties stemming from both the analytical and the natural variability of foraminiferal $\delta^{18}O$, as well as from the atlas $\delta^{18}O_{seawater}$. Uncertainties were propagated using a Monte Carlo approach. First, assuming a conservative error of 0.2‰ for the atlas $\delta^{18}O_{seawater}$ (following Peral et al. (2018)), we generated 10,000 iterations of $\delta^{18}O_{calculated}$ (Step 1 in Fig. 2A), using the equation of choice. For all these iterations, we then performed Step 2 (Fig. 2B) considering the uncertainty in $\delta^{18}O_{calcite}$ measurements (estimated from the standard deviation of replicate measurements) to obtain the ACD and calcification temperatures. We then calculated average ACDs and calcification temperatures for each sample from the individual iterations. Temperature estimates using all three approaches are compared in Table 3.

2.5. Linear regression

In order to account for the uncertainty in both Δ_{47} and calcification temperature, we calculate regression slopes and

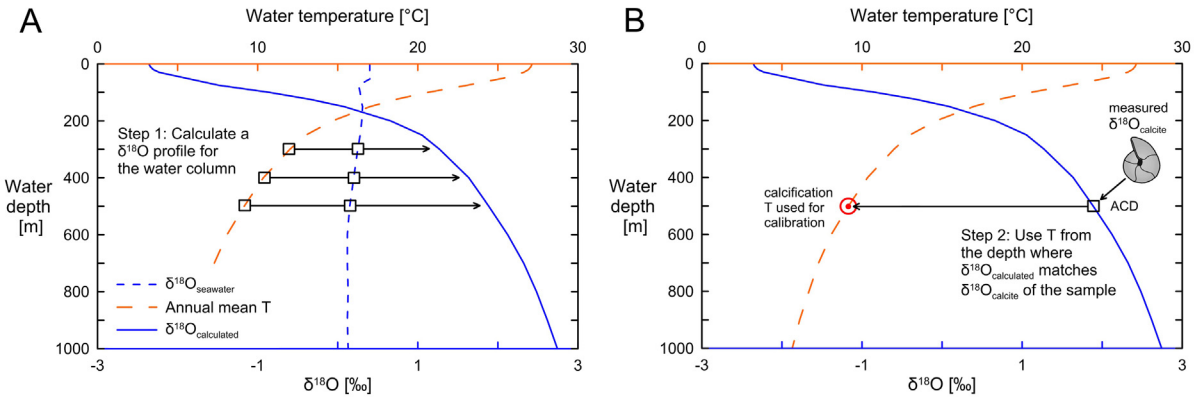

Fig. 2. Schematic drawing illustrating the two-step process (Method 3) used to assess apparent calcification temperatures from a combination of WOA-based temperature data and a δ^{18}O-T calibration. A: Annual mean temperature data (Locarnini et al., 2010) and $\delta^{18}O_{seawater}$ data (LeGrande and Schmidt, 2006) are used to generate a vertical $\delta^{18}O_{calculated}$ profile for calcite formed at any given location. B: The comparison between the measured $\delta^{18}O_{calcite}$ value of a foraminifer sample and the theoretical $\delta^{18}O_{calculated}$ profile is used to find the apparent calcification depth (ACD) of a foraminifer species. The annual mean water temperature at the ACD serves as best estimate for the calcification temperature of the foraminifer.

intercepts using the method of York et al. (2004). This method is commonly used in regression analysis of clumped isotope calibration data (e.g. Huntington et al., 2009; Grauel et al., 2013; Peral et al., 2018), thus helps facilitate the intercomparison of calibrations across studies. We estimated the uncertainty on the slope and intercept and 95% confidence envelopes on the regression lines using quantiles of 100,000 bootstrap samples. These were obtained by randomly resampling with replacement from the original data with its associated uncertainties, therefore maintaining the original sample size.

3. RESULTS

3.1. Δ_{47}, δ^{18}O and δ^{13}C data

Average data for each sample as well as environmental parameters such as the estimated $\delta^{18}O_{seawater}$ values and calcification temperatures reconstructed using various approaches (see Section 2.4) are summarized in Table 3. The average Δ_{47} data cover a range of 0.103‰ with a standard error of the mean for individual samples of 0.005–0.009. The lowest (0.653‰) and highest (0.756‰) Δ_{47} value correspond to the lowest (−2.29‰) and highest (3.66‰) $\delta^{18}O_{calcite}$ values, respectively (Fig. 3 A). Overall, there is a strong positive correlation (0.95 using Pearson's product-moment correlation) between both variables for the calculated averages. The standard deviation of replicate δ^{18}O measurements is 0.05–0.26‰ (standard error: 0.01–0.06‰). Mean δ^{13}C values for the samples measured in this study range between −1.2‰ and 3.2‰ with standard deviations between 0.04 and 0.24‰ (Fig. 3 B) and standard errors between 0.01 and 0.06‰. The δ^{13}C values and the Δ_{47} signal do not show a clear relationship. The standard errors of the mean and the standard deviations are not correlated with the mean Δ_{47}, δ^{18}O and δ^{13}C, respectively. Moreover, there is no systematic difference in the isotopic composition between species with and without photosymbionts.

3.2. Calcification temperatures

Because the calcification temperatures of planktonic foraminifers are challenging to estimate, we approximated them using three approaches (see Section 2.4). All three methods reveal strong correlations (correlation coefficient between −0.91 and −0.95 using Pearson's product momentum correlation) between estimated calcification temperature ($10^6/T^2$, T in K) and Δ_{47}. Detailed information on the regression lines derived from the different temperature estimates can be found in the Appendix (Table A3). Despite the strong correlations that were found for all the different methods, calcification temperature datasets differ from each other (Fig. 4A–D, Table 3).

The differences in estimated calcification temperature are largest between Method 1 using the World Ocean Atlas 2009 and the methods using δ^{18}O-T relationships (Fig. 4D). While the slopes of linear regression models for Methods 1 and 2 are similar (Fig. 4A), the dataset using WOA-based temperatures is characterized by larger variability (up to 13 °C temperature difference for similar Δ_{47} values, Fig. 4A). This is reflected in a lower correlation coefficient (−0.91 compared to −0.95 using Pearson's product-moment correlation). Hence, the uncertainty related to insufficient ecological information for certain species and/or regions seems to lead to a larger uncertainty for temperature estimates using Method 1 compared to δ^{18}O-based approaches (Methods 2 and 3).

When comparing calcification temperatures which were derived from two different δ^{18}O-T calibration equations (Shackleton, 1974; Kim and O'Neil, 1997) using Method 2 (Fig. 4D), temperature estimates for the tropical species largely agree (average temperature difference 0.7 °C), while towards the cold end of the calibration temperature estimates increasingly deviate from each other (maximum 2.2 °C). As a result, the regression calculated using the Kim and O'Neil (1997) δ^{18}O-T calibration reveals a flatter slope (Fig. 4B). The coldest temperature estimates

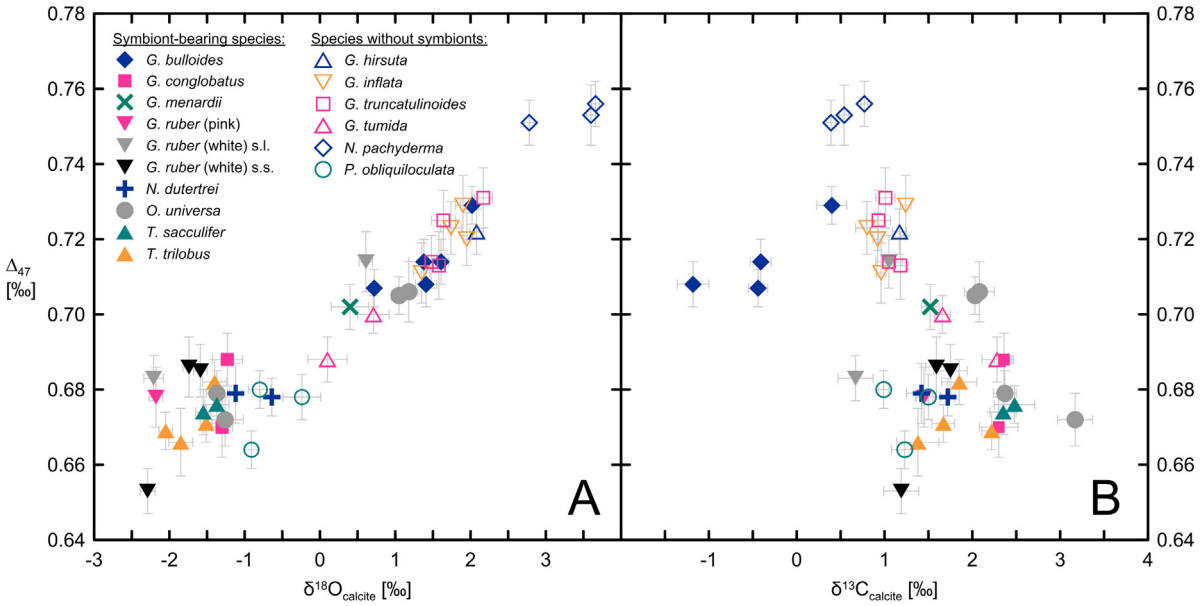

Fig. 3. Stable oxygen (A) and carbon isotopes (B) plotted against Δ_{47}. Symbols and colors represent different planktonic foraminifer species. Error bars illustrate one standard error for Δ_{47} and one standard deviation for $\delta^{18}O$ and $\delta^{13}C$. Note the positive correlation between $\delta^{18}O_{calcite}$ and Δ_{47}.

calculated using Kim and O'Neil (1997) are well below 0 °C. These temperature estimates do not agree with the available temperature data from the WOA, neither with the annual mean temperature nor seasonal extremes. In contrast, temperatures calculated using the $\delta^{18}O$-T calibration of Shackleton (1974) are in agreement with the temperature ranges reported in the WOA.

Temperatures estimated using Method 3 show a good agreement within the error estimates with those derived from Method 2 (Fig. 4C and D), and both datasets reveal similar correlation coefficients with Δ_{47}. The slopes of both linear regressions agree within error (Fig. 4C). Most samples reveal less than 1 °C temperature differences between both methods (Methods 2 and 3). In most cases with larger offsets the measured $\delta^{18}O$ value of the sample is lighter (on average 0.22‰) than the $\delta^{18}O_{calculated}$ value at the sea surface. As Method 3 uses the temperature at sea surface whenever the $\delta^{18}O_{calcite}$ value is lighter than the $\delta^{18}O_{calculated}$ value at 0 m, the resulting temperature estimate is lower than the one that is only based on the $\delta^{18}O$-T calibration.

4. DISCUSSION

4.1. Relationship between Δ_{47} and foraminifer calcification temperature

The method used to estimate calcification temperatures of planktonic foraminifers has an influence on the resulting Δ_{47} vs. temperature calibration. Fig. 4D demonstrates that the differences between Method 1 and Method 2 are pronounced for some of the samples, whereas others are less sensitive to the choice of method. We attribute this result to the varying accuracy of the ecological assumptions made for individual sites and species when using Method 1. Therefore, the WOA-based temperature estimates generated using Method 1 appear less applicable for a temperature calibration than temperature estimates based on the better established $\delta^{18}O$-T-relationship (Methods 2 and 3). Selecting an appropriate $\delta^{18}O$-T calibration for the reconstruction of calcification temperatures from the $\delta^{18}O$ values is however crucial, as there are systematic differences between temperature estimates generated with different calibrations (Fig. 4D).

It is unlikely that the coldest temperature estimates generated with the calibration of Kim and O'Neil (1997) are accurate, as they do not agree with the WOA temperature data at these sites. A possible cause for this discrepancy between the two calibrations is inherent in the way they were generated: Both calibrations use data from a temperature range of 0–500 °C by combining foraminifer data and inorganic calcite data from a study of O'Neil et al. (1969). The calibration of Shackleton (1974) combined this extensive inorganic dataset with foraminifer samples covering a temperature range of 0–7 °C and was specifically proposed to represent cold temperature carbonate samples (Shackleton, 1974; reviewed in Pearson, 2012). In contrast, the calibration of Kim and O'Neil (1997) used samples from a temperature range of 10–40 °C. Therefore, the latter equation may be less suitable to be applied to foraminifers calcifying at low water temperatures that are beyond the calibrated range. While acknowledging a remaining uncertainty related to the choice of $\delta^{18}O$-T calibration, we surmise that the calibration of Shackleton (1974) is the most reliable basis for temperature reconstructions from diverse settings and for a large number of different species based on the arguments outlined above.

Whether this equation is used directly (Method 2) or combined with available temperature data (Method 3) has only minor influence on the resulting calibration line

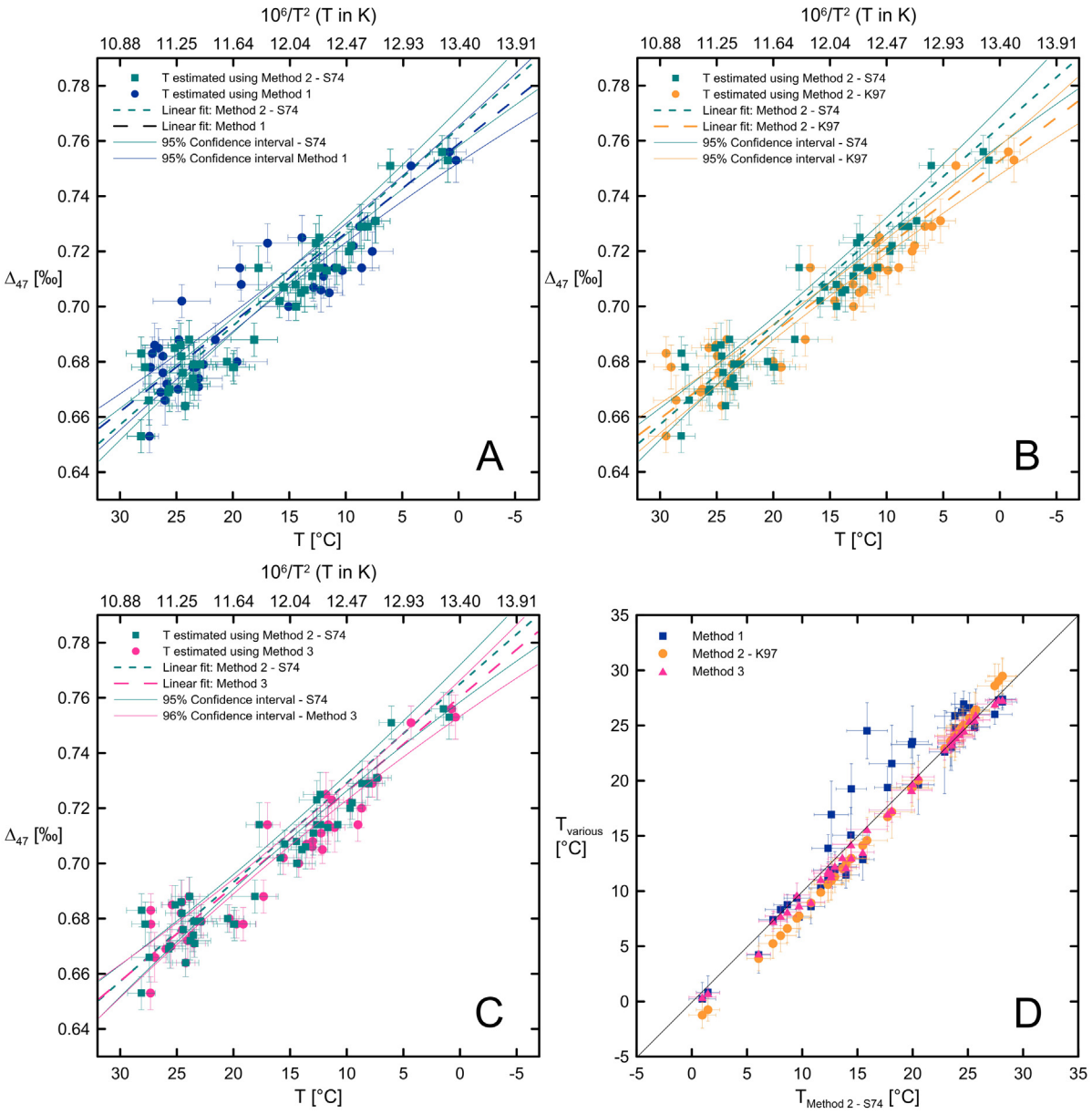

Fig. 4. Calcification temperatures estimated from three different methods versus Δ_{47} (A-C) and against each other (D). A: Method 1 using WOA-based temperature data (Locarnini et al., 2010) and a linear regression model (blue dots and blue dashed line) compared to Method 2 using the Shackleton (1974) equation (green filled squares and green dashed line). The 95% confidence intervals for Method 1 (blue) and Method 2 (green) are shown as solid lines. B: Method 2 using the Kim and O'Neil (1997) equation (orange dots and orange dashed line) and the Shackleton (1974) equation (green filled squares and green dashed line). Solid lines represent the 95% confidence intervals. C: Comparing Method 2 (green filled squares and green dashed line) to Method 3 (pink dots and pink dashed line), both approaches using the same $\delta^{18}O_{calcite}$ vs. temperature calibration (Shackleton, 1974). Solid lines represent the 95% confidence intervals. D: Comparison of temperatures estimated using Method 2 (Shackleton, 1974) with other temperature datasets used in A–C (blue dots: Method 1; orange dots: Method 2 using Kim and O'Neil (1997); pink dots: Method 3). Error bars represent the temperature uncertainty (A–D) in x direction and one standard error of the mean Δ_{47} (A–C) or the uncertainty of the temperature estimates (D) in y direction, as given in Table 3. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(Fig. 4C and D). One potential weakness of Method 3 compared to Method 2 is indicated by the observation that some sea surface sample values are not represented by the $\delta^{18}O_{calculated}$ profile. This could be explained by some unaccounted-for species-specific disequilibrium effects. Another reason for these $\delta^{18}O_{calcite}$ values that point to warmer temperatures than the annual mean SST could be seasonality effects in the life cycle of certain foraminifer species. These effects may bias the signal towards summer temperatures. This could for example affect samples of *G. ruber* pink as well as *N. pachyderma* since both species may calcify relatively close to the sea surface and are reported to

reach maximum abundances during summer season (Table 2).

On the other hand, assumptions regarding foraminifer ecology might not represent all of the data equally well and can potentially introduce errors, as evident from Fig. 4A. Consequently, the advantage of Method 3 over Method 2 is its independence from any assumptions concerning the calcification depth of the analyzed specimens. Moreover, since the calcification temperatures are derived from WOA data, extreme temperature values outside the observed annual mean temperature are excluded. We therefore use calcification temperatures derived with Method 3 for all subsequent calculations.

The relationship between Δ_{47} and calcification temperatures, which were derived from Method 3 and the application of the York regression, leads to the following regression equation:

$$\Delta_{47} = (0.0397\ 0.0021) * 10^6/T^2 + (0.2259\ 0.0255)\ \text{(T in K)} \quad (1)$$

The Δ_{47} data for all measured species follow a clear linear relationship (Fig. 5) albeit with noticeable scatter, which is most pronounced at the warm end of the calibration (>22 °C). Interestingly, earlier studies (Tripati et al., 2010, Grauel et al., 2013) have observed the opposite (i.e., larger variability at the cold temperature end, see Section 4.2). The observed variability is not related to a specific site being systematically offset from the general trend (Fig. 5A), but may stem from a variety of reasons as discussed in the following.

The small amounts (<5 mg) of sample material measured in this study may have led to a slightly larger scatter than observed in the recent foraminifer-based Δ_{47}-T calibration by Peral et al. (2018) where 16–20 mg were used. The samples in this study integrate over fewer individual specimens (minimum ~100) than studies measuring larger samples and could be affected by individual tests that deviate from the mean Δ_{47} values. Hence, the scatter of the Δ_{47} signal could likely be reduced further by measuring more replicates at the expense of slightly larger sample requirements.

Besides the measurement procedure, there are several potential reasons for individual samples to deviate from the described Δ_{47}-T relationship, related to either the calcification temperatures that were calculated from $\delta^{18}O_{calcite}$ or the Δ_{47} values. Because calcification temperatures can only be estimated, any divergence from the true calcification temperature can potentially cause affected samples to deviate from the general trend. Surface water conditions in particular can be highly variable (e.g. on a seasonal scale), potentially influencing the isotopic signal recorded by the foraminifers (e.g. Curry et al., 1983). However, potential seasonal temperature effects are largely accounted for by the combined approach of calculating $\delta^{18}O$ values and using WOA-based water temperatures (Method 3).

Despite the fact that planktonic foraminifers appear to calcify slightly offset from isotopic equilibrium with respect to stable oxygen isotopes of ambient seawater (Daeron et al., 2019), we do not see clear evidence that species-specific disequilibrium effects on $\delta^{18}O$ enhanced variability in our dataset. Species-specific correction for disequilibrium effects using published values would result in colder temperature estimates for several tropical surface species (Appendix Table A2). However, since there is a wide range of published disequilibrium correction factors (e.g. 0.0–1.0‰ for *G. ruber* (white), Niebler et al. (1999)), it remains difficult to assess its influence on the temperature estimates. Depending on the choice of correction factor for each species, the scatter of tropical surface-dwelling species in the Δ_{47} signal may in fact increase. In any case, a disequilibrium correction of the $\delta^{18}O_{calcite}$ signal moves the data of several warm-water surface species in the same direction and thus will not reduce the scatter of the data (Appendix Table A2 and Fig. A1).

Early diagenetic alterations such as secondary calcite precipitates grown at colder water temperatures may bias calcification temperature reconstructions from $\delta^{18}O$ measurements towards colder values (e.g. Pearson, 2012). However, such alteration should affect both the $\delta^{18}O$ and the Δ_{47} signal and bias all samples from the same site towards colder temperatures. Especially if two species from the same site are characterized by similar calcification depths, early diagenetic effects should be similar for both. Yet, the calculated calcification temperatures from $\delta^{18}O_{calcite}$ of different surface-dwelling species from the same sites generally agree well: For example, at site SO164-25-3 located in the Caribbean differences of up to 0.029‰ in the Δ_{47} signal were observed for surface species assumed to calcify at similar depth. For the same species the $\delta^{18}O$-based temperature estimates are characterized by a relatively small difference ≤1.8 °C (see Fig. 5A) and SEM images taken for several species at this site did not reveal any signs of secondary calcification. Moreover, we take the aforementioned good agreement between the temperatures calculated from $\delta^{18}O$ (Method 2) and temperature estimates based on WOA data in combination with published calcification depths (Method 1) as indication that the $\delta^{18}O$ signal is not altered (average temperature difference for tropical species: 0.7 °C). Due to the fact that clumped isotopes are only dependent on the mineral formation temperature, a stronger influence of early diagenesis on the Δ_{47} signal than on $\delta^{18}O$ is unlikely and has not been observed, even in samples as old as 44 Ma (Leutert et al., 2019).

Possible short-term variability of the surface water $\delta^{18}O$ due to salinity changes could introduce larger variability of the $\delta^{18}O_{calcite}$ signal of surface-dwelling species (reviewed in Pearson, 2012). For instance, site SO164-25-3 is located in an area that is influenced by the Amazon and Orinoco River plumes and may thus be experiencing considerable salinity changes (Schmuker and Schiebel, 2002). A surface-water salinity effect that influences the oxygen isotope signal of the upper water column could potentially bias the estimated calcification temperature of surface species at site SO164-25-3. For this site, however, we measured several surface-dwelling species including three morphotypes of the same species, *G. ruber* (*G. ruber* white s.s., s.l. and *G. ruber* pink). Although the $\delta^{18}O_{calcite}$ measurements from all three morphotypes agree well, not only among this species but also with other species from the same site, the Δ_{47} signal of these three samples reveals notable differences (~0.03‰). Moreover, none of the species measured for this site is characterized by a particularly variable $\delta^{18}O_{calcite}$

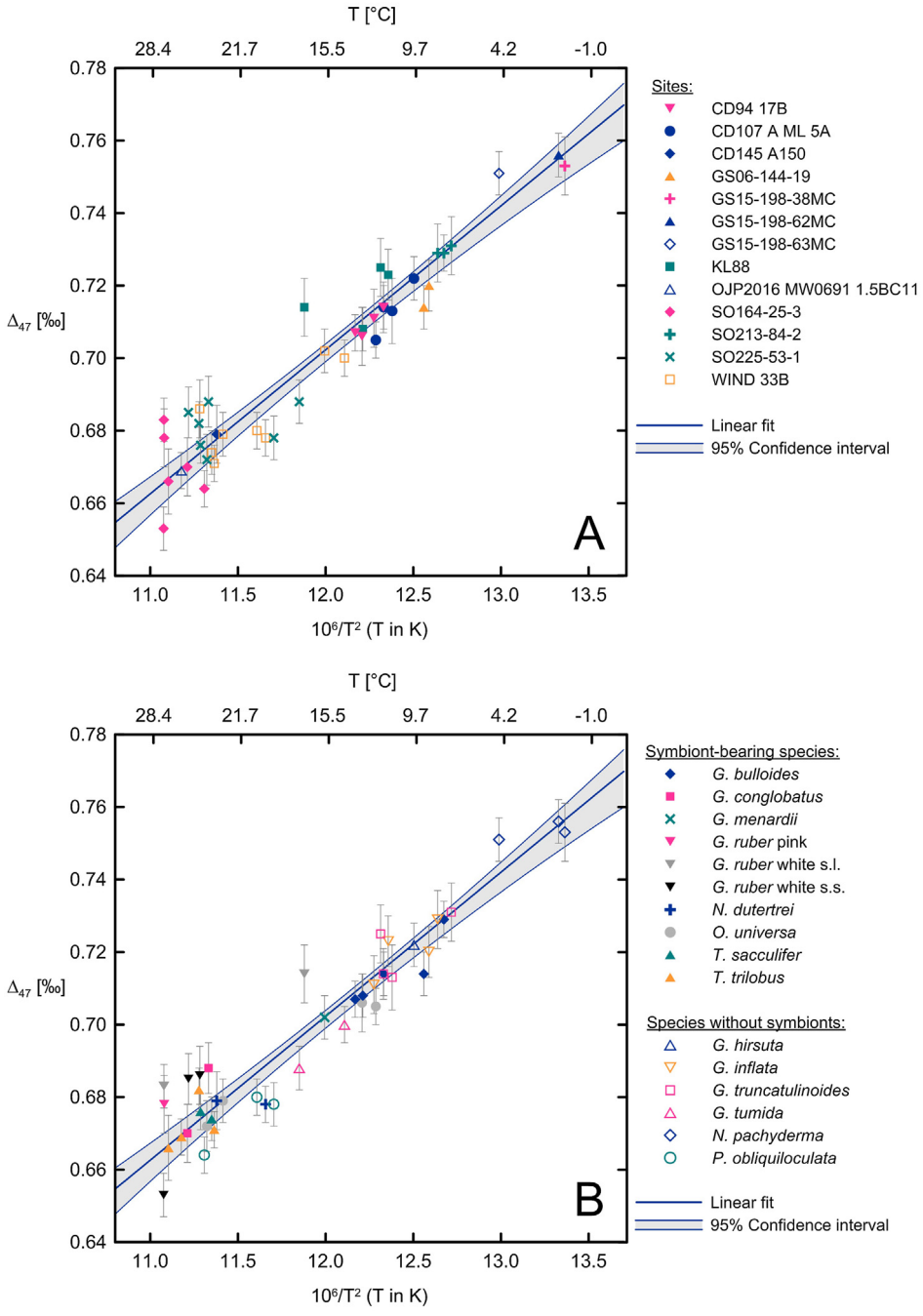

Fig. 5. The Δ_{47}-temperature relationship for planktonic foraminifers measured in this study by site (A) and species (B). Error bars represent the standard error of the mean Δ_{47} values (in per mil) (Table 3). Calcification temperatures are given in $10^6/T^2$ (T in K) and °C. The blue line and gray shaded area show the linear regression (Eq. (1)) and 95% confidence interval.

signal (SD of 0.09–0.16) and the calculated calcification temperatures are similar to the WOA-based estimates.

As suggested by various studies (e.g. Spero et al., 1997), pH affects the $\delta^{18}O$ signal of foraminiferal calcite towards more negative values with increasing pH. This effect was estimated by Zeebe (1999) to amount to −1.42‰ per unit of pH. The presence of photosymbionts is expected to increase the internal pH of foraminifers by up to 0.5 pH units (Rink et al., 1998) and thereby could bias temperatures calculated from $\delta^{18}O_{calcite}$ towards warmer values (reviewed in Pearson, 2012). For benthic foraminifers, on the other hand, Marchitto et al. (2014) found no clear pH effect. If pH had a strong influence on the calcification temperatures estimated from $\delta^{18}O_{calcite}$ in this study we would expect all symbiont-bearing species to reveal systematically warmer temperatures and disagree with atlas-based (Method 1) temperatures. Although this is not the case we cannot exclude that pH effects contributed some

additional scatter to the $\delta^{18}O_{calcite}$ signal of surface-dwelling species. Potential pH effects on the Δ_{47} signal will be discussed in Section 4.2. We suggest that the number of replicate measurements is the most important factor causing scatter in our foraminiferal Δ_{47}-T dataset. Nonetheless, we will investigate the data for species-specific effects possibly contributing to the scatter of the Δ_{47} signal in the following section.

4.2. Species-specific effects

The question whether the Δ_{47} in planktonic foraminifers is influenced by any species-specific effects is of vital importance to the application of this proxy for paleoceanography, since the absence of species effects would imply that Δ_{47} can be applied far back in time, despite evolutionary changes in species composition. Previous Δ_{47}-T calibration studies on foraminifers found large scatter and a potential discrepancy between foraminifers and inorganic calibrations at the cold end of the calibration (Tripati et al., 2010; Grauel et al., 2013), which were attributed to kinetic effects during the calcification process on foraminifers in cold-water conditions resulting in lower and more variable Δ_{47} values. In this study, in contrast, we observe neither increased scatter nor deviations towards low Δ_{47} in the cold-end foraminifer samples (Fig. 5). Cold-water species such as *N. pachyderma* from multiple sites do not reveal systematically negative, larger residuals. While most species are distributed relatively close to the calibration line ($\leq$0.01‰) and do not reveal any systematic offset, there are a few exceptions (Fig. 6A): One of two samples measured on *G. conglobatus* (site SO225-53-1 from the Manihiki Plateau) plots ~0.015‰ above the linear fit. Furthermore, all but one of the *G. ruber* samples from multiple sites are characterized by higher Δ_{47} values (residuals 0.01–0.02‰). This includes all three morphotypes of *G. ruber* measured in this study. Furthermore, all three samples of *P. obliquiloculata* show lower Δ_{47} values with two of them <-0.010‰.

Taking the uncertainty of the calcification temperature estimates and the clumped isotope measurements into account, the measured, species-specific Δ_{47} data from this study do not reveal any statistically significant deviation from the linear relationship determined for the entire dataset. Some species are only represented by a single sample, whereas up to six samples from different sites were included for species that are frequently used for paleoceanography (such as *G. bulloides*, *G. ruber and N. pachyderma*). The limited temperature ranges of individual species do not allow for the calculation of individual, species-specific regression lines.

To further test the influence of individual species or genera on the Δ_{47}-T calibration, we removed consecutively certain taxa from the dataset and compared the resulting slopes of the calibrations to the slope of the entire dataset (Fig. 6B). All of the slopes calculated for such data subsets fall within the 95% confidence interval of the slope calculated for the entire dataset. The two species that have the largest influence on the slope are *N. pachyderma* and *G. ruber*. This observation could be related to the position of the data from these two species at the cold (*N. pachyderma*) and warm end (*G. ruber*) of the dataset as the regression line is particularly sensitive to data at both ends of the temperature range. The dataset without *N. pachyderma* is characterized by a flatter slope (0.0380 compared to 0.0397 for the entire dataset). This deviation could be explained by the fact that the exclusion of this cool (<10 °C), high-latitude species from the dataset reduces the entire temperature-range by 7 °C and hence raises the uncertainty of the

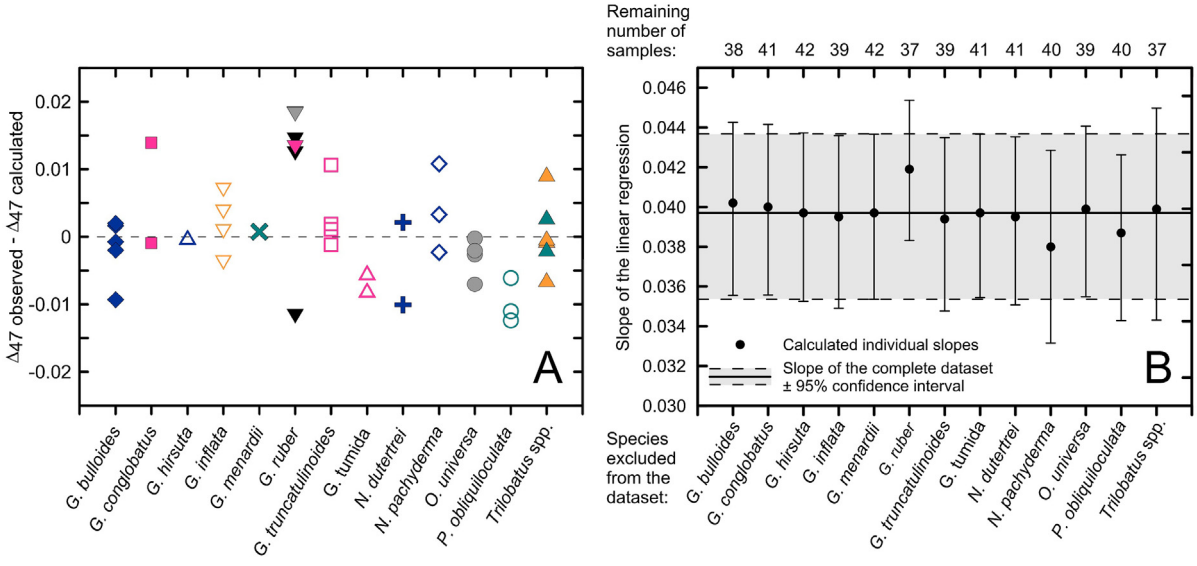

Fig. 6. Evaluation of possible species-specific effects on the calibration; A: Δ_{47} residuals of single foraminifer species, calculated as deviation from the linear regression presented in Fig. 5 (individual species/species groups are displayed by different symbols and colors similar to Fig. 5B). The residuals show no significant trend. B: Calculated slopes of the linear regression for subsets of the data (black dots) excluding one species at a time to test the influence of individual species on the calibration. The black line represents the slope of the complete dataset (gray area: 95% confidence interval of the mean slope). Error bars represent the 95% confidence intervals of the slopes.

Δ_{47}-T calibration (standard error raised from 0.0021‰ to 0.0025‰).

For *G. ruber*, all but one sample are offset by on average 0.015‰ to a higher Δ_{47} value (lower temperature) from the calculated Δ_{47}-T regression line of the entire dataset (Figs. 5B, 6A). The linear fit calculated for the dataset without *G. ruber* exhibits the steepest slope (0.0419) of any of the subsets presented in Fig. 6B, suggesting that these data exert the strongest influence on the regression, albeit not significant at the 95% confidence level. This relatively large effect on the slope despite the temperature range of the dataset being largely unaltered may imply that the *G. ruber* Δ_{47}-T dependency is different from other planktonic foraminifer species. However, the observation that not all *G. ruber* data show the offset to higher Δ_{47} values is inconsistent with this hypothesis. For Caribbean site SO164-25-3, for example, we analyzed three different morphotypes of *G. ruber* (*G. ruber* white s.s, s.l. and *G. ruber* pink). Two of these samples (*G. ruber* white s.l. and *G. ruber* pink) show the offset to higher Δ_{47}, whereas the third morphotype (*G. ruber* white s.s.) reveals a lower (by ~0.03‰) Δ_{47} value than the other two and plots below the regression line (Fig. 5B). Moreover, Δ_{47} in *G. ruber* was also analyzed in the recent study by Peral et al. (2018), which did not reveal systematic species-specific behavior of *G. ruber*, although one of three samples was characterized by slightly higher Δ_{47}. Overall, evidence regarding species-specific effects in *G. ruber* is inconclusive and we do not consider *G. ruber* to calcify systematically offset from other species of foraminifers with respect to Δ_{47}. Individual samples of other species (such as *G. conglobatus*) also deviate from the linear regression to a similar degree. A combination of the intrinsic uncertainty of the clumped isotope measurement together with natural variability of the sample material could explain the observed scatter of the Δ_{47}-T data, including the apparent deviation of *G. ruber* from the trend.

However, we cannot rule out the existence of relatively small and possibly variable secondary influences on the Δ_{47} signal during the calcification process in the surface water. The fact that the shallowest surface-dwelling species *G. ruber* shows the strongest deviation of individual samples from the general Δ_{47}-T relationship raises the question, whether there could be additional effects besides temperature on Δ_{47} in species living close to the sea surface. This has been described for other groups of marine calcifying organisms as well. Tropical shallow-water corals, for example, show increasing Δ_{47} with increasing calcification rates (Saenger et al., 2012). Kinetic effects on the Δ_{47} signal related to growth rates in brachiopods were shown by Bajnai et al. (2018). Moreover, Davies and John (2019) reported evidence for a constant offset of echinoid Δ_{47} values from inorganic calcite that might be related to internal pH of the calcifying fluid in echinoids being offset from seawater pH. If this was true for foraminifers as well, the effect would be expected to be larger for symbiont-bearing species such as *G. ruber* because the internal pH in these species was reported to be higher (Rink et al., 1998). In contrast to this hypothesis, however, Tang et al. (2014) observed in inorganic calcite precipitation experiments that the Δ_{47}-T relationship is largely insensitive to pH and growth rate at the external and internal conditions expected during foraminifer calcification. Tripati et al. (2015) also found that even major changes in ocean chemistry (pH and salinity) expected during the Cenozoic have only small to negligible effects on the Δ_{47} signal of marine carbonates.

One potential explanation for increased scatter of the Δ_{47} values in surface-dwelling planktonic foraminifers could be the additional influence of photosynthesis on the calcification process: Photosymbionts may potentially cause disequilibrium effects on the recorded Δ_{47} signal because they are not only altering the microenvironment from which calcite is excreted but also by affecting calcification rates (de Nooijer et al., 2014). An effect of strong photosynthesis in plants on Δ_{47} measured in residual CO_2 gas was demonstrated by Laskar and Liang (2016), who also reported plant photosynthesis to decouple the $\delta^{18}O$ and Δ_{47} values. However, studies investigating photosynthesis in foraminifers found that both the kind of symbionts and the concentration of chlorophyll as a measure for photosynthetic activity in *G. ruber* are similar to other species, such as *T. sacculifer* (Fujiki et al., 2014; Takagi et al., 2019). This would suggest

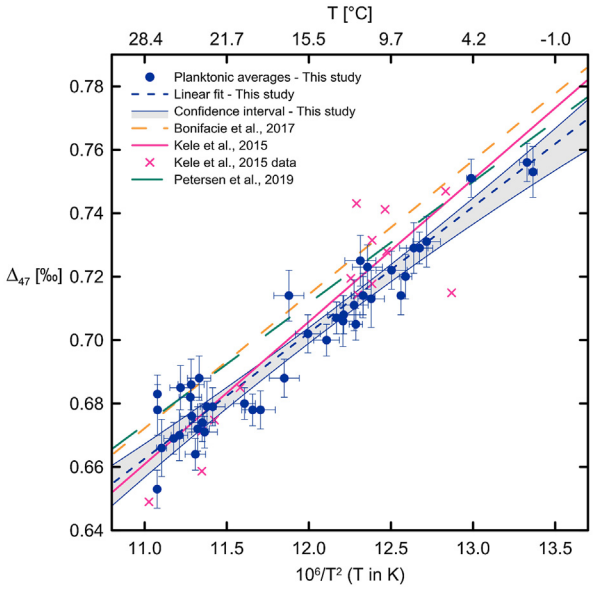

Fig. 7. Comparison of various Δ_{47}-T calibrations. Clumped isotope data from this study (blue symbols) and the calculated linear regression (dashed blue line) including confidence intervals (gray shaded area) is compared with published data and calibration lines including the compilation of Bonifacie et al. (2017; yellow dashed line), the travertine calibration of Kele et al. (2015; pink line and symbols) recalculated by Bernasconi et al. (2018), and the composite synthetic carbonate calibration of Petersen et al. (2019; green dashed line). The calibration of Bonifacie et al. (2017) and Petersen et al. (2019) contain data from several individual studies that were measured, standardized, and corrected in various ways. Therefore, the individual data of the Bonifacie et al. (2017) and Petersen et al. (2019) compilations are characterized by a larger variability and thus, are not presented here. Error bars in x and y-direction represent the temperature uncertainty and one standard error of the mean Δ_{47}, respectively, as presented in Table 3. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

that any effect photosynthesis may have on Δ_{47} in *G. ruber* should also occur in other symbiont-bearing species, whereas we do not observe any systematic difference between symbiont-bearing and non-symbiotic foraminifer species (Fig. 5B). Because *G. ruber* has a shallower habitat depth than most other planktonic foraminifers (e.g. Wang, 2000), the influence of photosynthesis might be larger on this species than on deeper-dwelling symbiotic species. A higher symbiont density under high-light conditions likely affects the calcification process of *G. ruber* as already suggested for boron isotopes (see Hönisch and Hemming, 2004). However, this effect cannot explain the differences of the Δ_{47} signal in three morphotypes of *G. ruber* from the same site (see Fig. 5). In particular the *G. ruber* (white) s.s. morphotype reported to live closest to the sea surface yields Δ_{47} values closest to the calibration line. More work is needed to test possible secondary influences on the incorporation of the Δ_{47} signal and evaluate whether there might be any significant species-specific effects in *G. ruber* and similar shallow-dwelling species.

4.3. Comparison with other clumped isotope calibrations

Our data is in overall good agreement with recent clumped isotope calibrations of Bonifacie et al. (2017), Kele et al. (2015) as recalculated by Bernasconi et al. (2018), and Petersen et al. (2019) (Fig. 7). Our linear regression has a slope between the flatter Petersen et al. (2019) calibration and steeper slopes published by Bonifacie et al. (2017) and Kele et al. (2015) (Table 4). Since the slope and intercept are negatively correlated for the regression analysis the intercept of our regression is higher than the ones published by Bonifacie et al. (2017) and Kele et al. (2015). The Petersen et al. (2019) calibration, which is based on a compilation of synthetic carbonate data shows a slight but apparently systematic offset towards higher Δ_{47} values compared to our data. Most of the datasets included in this compilation, as well as in the compilation by Bonifacie et al. (2017), however, used different analytical and data correction procedures compared to our study.

Bonifacie et al. (2017) combines data from various existing calibrations that were generated on a variety of analytical setups and using different standards. Also, several of the datasets included in the Bonifacie et al. (2017) calibration were calculated with the parameter set for ^{17}O correction by Gonfiantini et al. (1995), which were later shown to have caused discrepancies among samples with very different bulk compositions (Daëron et al., 2016; Schauer et al., 2016). For example, the intercept of the travertine calibration of Kele et al. (2015) decreased by 0.038‰ with the updated ^{17}O abundance correction of Brand et al. (2010) (Bernasconi et al., 2018). However, the original data are included in the compilation of Bonifacie et al. (2017). Recalculating the calibration of Bonifacie et al. (2017) using the updated Kele dataset might therefore lead to a lower intercept and move this calibration closer to our calibration (Eq. (1)).

The underlying data of the (fully recalculated, see Bernasconi et al., 2018) travertine calibration by Kele et al. (2015) agree well with our foraminifer data (Fig. 7). This is most likely due to similarity in analytical set-up, raw data treatment and correction using the same carbonate standard values. Differences in data correction and/or analytical procedures may thus explain some degree of systematic offset between calibrations. In addition, Fernandez et al. (2017) pointed out that a small temperature range of some Δ_{47}-T calibrations is among the important factors that can explain discrepancies between various calibration lines. Given the small temperature range biogenic samples like foraminifers cover, the slight discrepancies between calibrations are not surprising.

4.4. Comparison of foraminifer-based calibrations

Comparing our data to recent Δ_{47}-T calibrations of Peral et al. (2018), Piasecki et al. (2019) and Breitenbach et al. (2018) (Fig. 8) contributes to the ongoing debate regarding

Table 4
Our foraminifer-based Δ_{47}-T calibration compared to recent clumped isotope calibrations in the 25 °C reference frame. All calibrations except for Bonifacie et al. (2017) and Petersen et al. (2019) were calculated using ETH carbonate standards and the correction for temperature dependend acid fractionation that was published by Defliese et al. (2015). The equation published by Bonifacie et al (2017) was converted to the 25 °C reference frame using a correction for temperature-dependent acid fractionation of 0.082‰ (Defliese et al., 2015) on the intercept. The intercept of the Petersen et al. (2019) calibration was lowered by −0.004‰ (see Section 2.3) for comparability with equations based on the acid fractionation factors published by Defliese et al. (2015).

Regression	Slope * $10^6/T^2 \pm 1$ SE	Intercept ± 1 SE	Type of material
Bonifacie et al. (2017)	0.0422 ± 0.0019	0.208 ± 0.0207	Various – compilation of existing calibration data
Breitenbach et al. (2018) (foraminifer dataset)	0.0315 ± 0.008	0.313 ± 0.1	Foraminifers
Breitenbach et al. (2018) (Cambridge calibration)	0.0448 ± 0.007	0.154 ± 0.08	Natural cave carbonates (cave pearls)
Kele et al. (2015) (fully recalculated)	0.0449 ± 0.001	0.167 ± 0.01	Inorganic carbonates (travertines, tufas)
Peral et al. (2018)	0.04163 ± 0.00084	0.2056 ± 0.0011	Foraminifers
Petersen et al. (2019)	0.0383 ± 1.7E^{-6}	0.258 ± 1.7E^{-5}	Compilation of synthetic carbonate data
Piasecki et al. (2019)	0.0460 ± 0.005	0.159 ± 0.064	Benthic foraminifers
This study	0.0397 ± 0.0021	0.2259 ± 0.0255	Planktonic foraminifers
Combined calibration (Breitenbach et al., 2018, Peral et al., 2018, Piasecki et al., 2019, this study)	0.0418 ± 0.0016	0.2017 ± 0.0195	Foraminifers
Combined calibration (Peral et al., 2018, Piasecki et al., 2019, this study)	0.0431 ± 0.0016	0.1876 ± 0.0189	Foraminifers

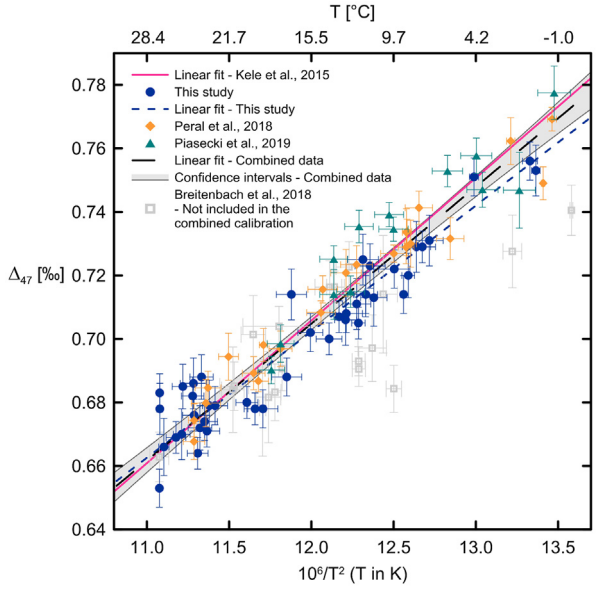

Fig. 8. Compilation of available foraminiferal Δ_{47}-T calibrations. The planktonic foraminifer data from this study (blue symbols and dashed line) are compared to foraminiferal data from Peral et al. (2018) with temperatures recalculated using Method 3 (orange diamonds), benthic foraminiferal data (site averages) from Piasecki et al. (2019) (green triangles) and planktonic foraminiferal data from Breitenbach et al. (2018) with recalculated temperatures using Method 3 (gray squares). Moreover, a combined foraminifer calibration (black dashed line) including 95% confidence intervals (gray shaded area) for the data from this study, Peral et al. (2018), and Piasecki et al. (2019) is presented with the recalculated calibration line of Kele et al. (2015; pink line) for comparison. Error bars in y-direction represent one standard error of the mean Δ_{47}. Error bars in x-direction show the calculated temperature uncertainty for Breitenbach et al. (2018), Peral et al. (2018) and this study and an assigned constant uncertainty of ±1 °C for Piasecki et al., 2019. Note that the foraminiferal data from Breitenbach et al. (2018) were not included in the combined calibration due to the larger variability of the dataset (also discussed in Breitenbach et al., 2018). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

inter-laboratory differences and improves foraminifer-based Δ_{47}-T calibration efforts. In order to treat the datasets consistently, we recalculated calcification temperatures for the datasets of Peral et al. (2018) and Breitenbach et al. (2018) using Method 3. The previously published bottom water temperatures for the benthic dataset of Piasecki et al. (2019) were kept. Bottom water conditions are assumed to be relatively constant over time, such that instrumental measurements can be regarded as reliable to calibrate Δ_{47}-data of benthic foraminifers. Such an approach independently verifies that the calculated planktonic calcification temperatures are realistic, provided that neither planktonic nor benthic foraminifers record Δ_{47} values significantly offset from the inorganic Δ_{47}-T relationship.

In general, the datasets are in good agreement across the entire temperature range from −1 to 28 °C. The benthic foraminiferal Δ_{47} data of Piasecki et al. (2019) was generated in the same laboratory as the data from this study and seem to indicate a slight deviation from our planktonic Δ_{47}-T calibration towards higher Δ_{47} values for temperatures below 15 °C. At the same time, the variability of the benthic Δ_{47} data at the cold end of the calibration is relatively large. This can be explained by individual data points that contain less replicate measurements due to sample limitations.

The planktonic foraminiferal Δ_{47}-T data from Breitenbach et al. (2018) are characterized by a larger scatter. While all four datasets presented in Fig. 8 overlap for the warm end of the Δ_{47}-T calibration, Δ_{47} values at the cold end (<13 °C) tend to be lower by ~0.02–0.03‰ in the Breitenbach dataset. Breitenbach et al. (2018) acknowledge the small number of samples and replicates as well as the relatively large scatter of the dataset, which was generated with the primary purpose of comparing Δ_{47} and Mg/Ca.

The data from Peral et al. (2018) show excellent agreement with our measurements, with confidence intervals overlapping for the whole temperature range. This is particularly noteworthy as the two datasets were derived with completely different analytical setups: In this study, samples were digested at 70 °C in a Kiel IV carbonate preparation device with a short Porapak column and subsequently measured on a Thermo Fisher Scientific MAT 253 Plus in microvolume mode with the LIDI approach. In contrast, Peral et al. (2018) used a common acid bath operated at 90 °C, a GC column for contaminant removal, and carried out the isotope measurements on a VG Isoprime mass spectrometer under constant gas pressure. The good agreement of the data provides further evidence that different measurement techniques provide comparable Δ_{47} data as long as they are corrected using the same carbonate standards (in this case ETH 1-4; Bernasconi et al. 2018) and the "Brand parameters" for the ^{17}O abundance correction (Daëron et al., 2016; Schauer et al., 2016).

Based on these considerations, the various recent datasets containing foraminiferal Δ_{47} data with comparable data treatment were combined to enhance the accuracy of an overarching Δ_{47}-T calibration valid for all foraminifer species. We excluded the Δ_{47} data of Breitenbach et al. (2018) due to the small number of replicate measurements (although we report a version of a combined foraminifer-based calibration including this dataset in Table 4). The resulting Δ_{47}-T calibration encompassing the data of Peral et al. (2018), Piasecki et al. (2019) and this study (Eq. (2), Fig. 8) falls within the error of the regression exclusively derived from our data (which is characterized by a slightly flatter slope) and emphasizes the conformity and compatibility of the three Δ_{47} datasets:

$$\Delta_{47} = (0.0431\ 0.0016) * 10^6/T^2 + (0.1876\ 0.0189)\ (T \text{ in K}) \quad (2)$$

The recalculated version of the Kele et al. (2015) calibration (see Bernasconi et al., 2018) lies within the confidence interval of the combined foraminiferal Δ_{47}-T calibration presented here. For the calibrated temperature range our combined calibration yields temperature estimates within 1 °C of the Kele et al. (2015) calibration with the largest difference at the cold end of the calibrated temperature range. Given the fact that the Kele et al. (2015) calibration used the same carbonate standards for the corrections, this

agreement suggests that foraminifers follow the same Δ_{47}-T relationship as the travertines. We note that this agreement between our combined foraminifer-based calibration and the Kele travertine calibration does not exclude the possibility that all of these carbonates are influenced by some degree of disequilibrium fractionation (Watkins and Hunt, 2015; Daeron et al., 2019).

For future studies using foraminifer samples we recommend applying our combined foraminifer-based calibration (Eq. (2)) rather than the recalculated travertine calibration of Kele et al. (2015). Although the Kele et al. (2015) calibration has the advantage of covering a much wider temperature range our combined calibration is based on a large number of foraminifer samples from different studies and laboratories. Hence it is characterized by a smaller uncertainty within the normal ocean temperature range compared to the Kele et al. (2015) calibration. However, the reconstructed temperatures applying either of the two calibrations fall within less than 1 °C of each other. Using the long-integration dual-inlet (LIDI) method a sample size of 2–5 mg of foraminifers is enough for 20 to 40 replicate measurements, which is the equivalent of a temperature uncertainty of 1.5 °C or less on the measurement.

4.5. Water column temperature gradients

A widely accepted approach to gain a deeper understanding on past oceanographic changes is to combine geochemical information (e.g. combined analyses of foraminiferal Mg/Ca and $\delta^{18}O$) from calcitic tests of shallow and deep-dwelling foraminifer species, allowing reconstruction of water column stratification. Based on the notion that there are no discernible species effects on the Δ_{47}-T calibration presented above (Eq. (2)) and the close agreement with the travertine-based Kele calibration (Kele et al., 2015), we test how reliably vertical temperature gradients can be reconstructed from foraminiferal Δ_{47} data. We compare Δ_{47}–derived temperatures from various species from two Pacific and Indian Ocean sites to annual mean water temperatures and seasonal extremes at these locations (Locarnini et al., 2010) (Fig. 9). To avoid circular reasoning, the Δ_{47}–temperature estimates were derived from the Kele calibration and plotted at the respective assumed calcification depths of the species, based on the available ecological information (c.f. Table 3). This exercise should be seen as a feasibility study.

Within error, the Δ_{47}–temperature estimates from almost all foraminifer species compare to the annual mean temperature at the respective sample locations. The absolute temperature difference of ~15 °C between the two sites is well reflected in the Δ_{47}–temperature signal. On the vertical scale, shallow-dwelling species commonly show higher Δ_{47}–temperatures than deep-dwelling species, and the reconstructed temperature differences reflect the different gradients at the two sites very well. Only *G. menardii* and *G. ruber* from site WIND 33B in the Indian Ocean yield Δ_{47}–temperatures that appear too cold for their assumed calcification depths (by 5 and 6 °C, respectively, Fig. 9). The calcification temperature reconstructed for *G. menardii* suggests a habitat in the lower thermocline at this site, lower than commonly assumed, which is also seen in the deeper $\delta^{18}O$-based apparent calcification depth (Table 3). The apparent cold bias of *G. ruber* is a recurrent feature

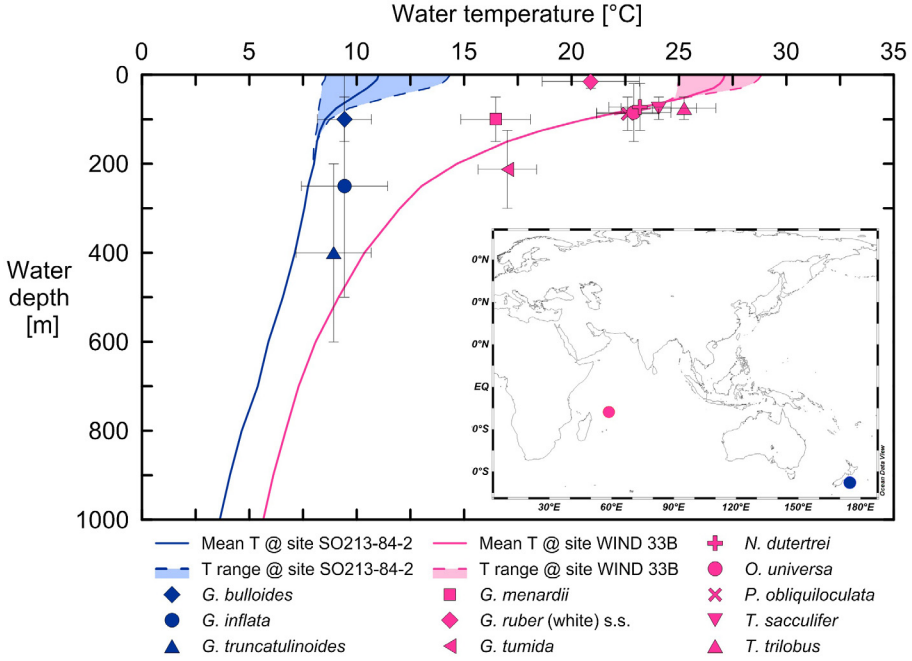

Fig. 9. Atlas-based mean annual water temperature (solid lines) and seasonal temperature range (dashed lines) (Locarnini et al., 2010) for sites SO213-84-2 (blue) and WIND 33-B (pink) and clumped isotope temperatures (using the calibration published by Kele et al., 2015, recalculated by Bernasconi et al. (2018)) plotted against water depth/assumed calcification depth (Table 3). Error bars in x-direction represent the temperature uncertainty due to the standard mean error of the measurement while the error bars in y-direction show the uncertainty of the calcification depth based on the available information presented in Table 3.

observed at various sites and ocean settings in this study as discussed above. Overall, this exercise demonstrates that Δ_{47} can be used to reconstruct vertical temperature gradients within the water column while avoiding the uncertainty introduced by the use of individual, species-specific calibrations for other foraminifer-based geochemical proxies.

5. CONCLUSION

By analyzing Δ_{47} in 14 species from 13 globally distributed core-top samples, this study confirms findings from previous studies (Tripati et al., 2010; Grauel et al., 2013; Peral et al., 2018) that found foraminifers to follow the same relationship between Δ_{47} and the carbonate formation temperature as inorganic calcite. The substantial number of different foraminifer species analysed here, as well as the large number of samples from different sites for some species greatly increases confidence in this finding. Although small species-specific effects within the analytical uncertainty cannot be completely ruled out, no significant systematic effect could be identified in this study. The only possible deviation from the Δ_{47}-T relationship that cannot be explained by the uncertainty associated with foraminiferal ecology is the mixed-layer species *G. ruber*, showing apparent cold biases in some samples. However, the results for this species remain inconclusive, warranting a more detailed study on the clumped isotope signal in *G. ruber*.

We demonstrate that results from different laboratories and various measurement setups are in good agreement when the Δ_{47} data are corrected using the same carbonate standards and the latest ^{17}O abundance correction parameters. The combination of natural variability, relatively large uncertainties of the estimated calcification temperatures and the comparatively small natural temperature range affect the precision of any Δ_{47}-T calibration based only on foraminifers. We minimize this problem by combining several available foraminifer-based calibrations and calculating a common Δ_{47}-T calibration. Within the error, this combined calibration is identical to the recalculated travertine calibration of Kele et al. (2015) (see Bernasconi et al., 2018). Temperatures reconstructed using either of the two calibrations fall within less than 1 °C of each other. Because of the smaller uncertainty within the ocean temperature range, we recommend using our combined calibration (Eq. (2)) for foraminifer samples. Finally, we show that the reconstruction of temperature profiles through the water column from clumped isotope measurements is feasible using micro-volume measurements on different species within the same sample.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ACKNOWLEDGEMENTS

We thank Eystein Jansen at the University of Bergen and NORCE Norwegian Research Centre AS and Trond Dokken and Bjørg Risebrobakken at NORCE for the provision of sample material. We would also like to thank Henrik Sadatzki at the University of Bergen for helping with the sample preparation. Radiocarbon dates for the GS15 samples were kindly provided by the ice2ice project funded by the European Research Council under European Community´s Seventh Framework Programme (FP7/2007-2013) ERC grant agreement 610055. We thank Eivind W. N. Støren for access to the facilities of EARTHLAB and Enver Alagoz, Pål Tore Mørkved and the clumped isotope group at the University of Bergen for laboratory assistance and technical support during the analysis. We are grateful to Alison Piasecki for discussions and input. We also thank the reviewers for the constructive comments that helped to improve the manuscript. This work was funded by the Trond Mohn Foundation and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 638467). The UCLA contribution was supported by the Department of Energy through BES grant DE-FG02-13ER16402.

RESEARCH DATA

The raw replicate level data of this study are available at EarthChem (URL: https://doi.org/doi:10.1594/IEDA/111435).

APPENDIX A

See Figs. A1 and A2 and Table A1–A3.

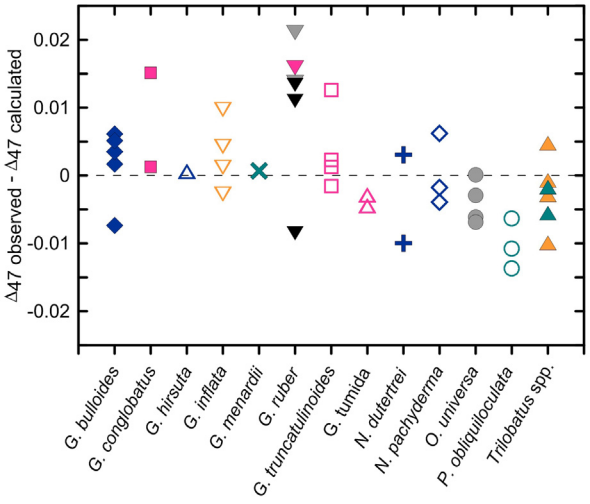

Fig. A1. Residual values for the resulting linear regression after species-specific disequilibrium correction factors (Appendix Table A2) were applied. The data are differentiated for each species (displayed by different symbols and colors as in Fig. 5B). Applying species-specific disequilibrium corrections to the $\delta^{18}O_{calcite}$ data prior to calculating the calcification temperature dos not result in smaller residual values compared to Fig. 6A.

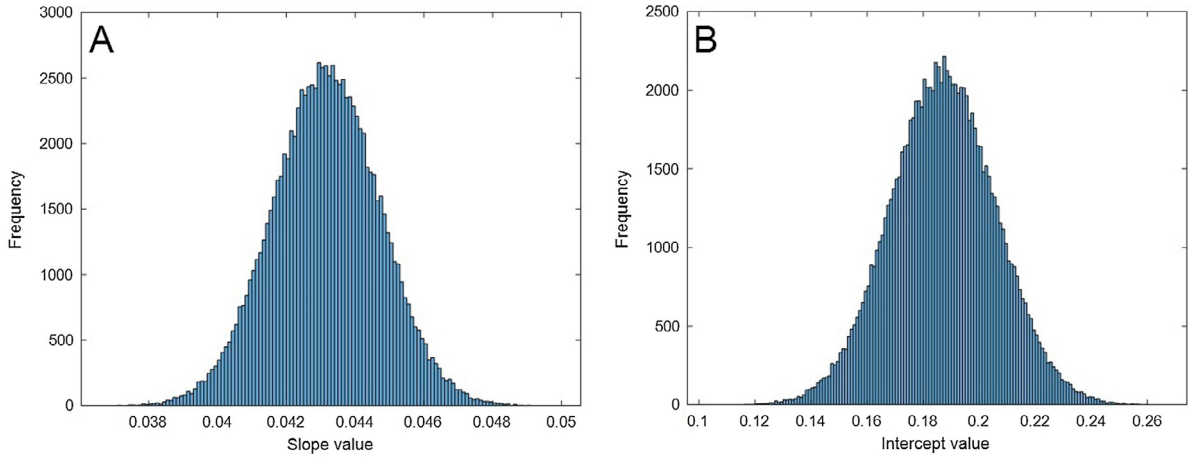

Fig. A2. Histograms showing the distribution of slopes (A) and intercepts (B) of the combined calibration across the 100,000 iterations calculated.

Table A1
External reproducibility of standard measurements. For the measuring interval from October 2016 to October 2017, the carbonate standards ETH 1, 3 and 4 were used for correction and the carbonate standard ETH 2 for monitoring. From November 2017 to March 2018, ETH 1, 2 and 3 were utilized for correction and ETH 4 for monitoring.

Standard	$\delta^{13}C$ VPDB SD external	$\delta^{18}O$ VPDB SD external	Δ_{47} CDES SD external
Measurement interval from 2016/10/21 13:14 to 2017/10/14 17:06			
ETH-1 (n = 957)	0.021	0.042	0.036
ETH-2 (n = 967), monitoring	0.022	0.045	0.035
ETH-3 (n = 982)	0.027	0.048	0.033
ETH-4 (n = 967)	0.020	0.042	0.031
Measurement interval from 2017/11/30 16:17 to 2018/03/14 07:29			
ETH-1 (n = 458)	0.045	0.087	0.037
ETH-2 (n = 402)	0.061	0.097	0.034
ETH-3 (n = 460)	0.063	0.106	0.036
ETH-4 (n = 456), monitoring	0.059	0.112	0.038

Table A2
Disequilibrium offset correction factors tested for the $\delta^{18}O_{calcite}$ measurements that were used to estimate calcification temperatures.

Species	Disequilibrium offset correction	Reference
Globigerina bulloides	0.25	Niebler et al. (1999)
Globigerinoides conglobatus	−0.2	Niebler et al. (1999)
Globigerinoides ruber pink	−0.4	Steph et al. (2009)
Globigerinoides ruber white s.l.	−0.4	Steph et al. (2009)
Globigerinoides ruber white s.s.	−0.4	Steph et al. (2009), Rippert et al. (2016)
Globorotalia hirsuta	−0.15	Niebler et al. (1999)
Globorotalia inflata	0	Niebler et al. (1999), Cléroux et al. (2013)
Globorotalia menardii	−0.2	Niebler et al. (1999), Steph et al. (2009)
Globorotalia truncatulinoides	−0.05	Niebler et al. (1999), Cléroux et al. (2013)
Globorotalia tumida	0	Niebler et al. (1999), Steph et al. (2009), Cléroux et al. (2013)
Neogloboquadrina dutertrei	−0.27	Niebler et al. (1999), Cléroux et al. (2013)
Neogloboquadrina pachyderma	−0.75	Niebler et al. (1999)
Orbulina universa	−0.3	Niebler et al. (1999)
Pulleniatina obliquiloculata	−0.3	Niebler et al. (1999)
Trilobatus sacculifer	−0.6	Steph et al. (2009), Rippert et al. (2016)
Trilobatus trilobus	−0.6	Steph et al. (2009), Rippert et al. (2016)

Table A3
Correlation coefficients (Pearson's product-moment correlation), slopes and intercepts calculated using various calcification temperatures for the following linear regression: $\Delta_{47} = (m \pm SE) * 10^6/T^2 + (b \pm SE)$ (T in K).

Method used to estimate calcification temperature	Correlation coefficient	Slope (m)	SE	Intercept (b)	SE
Method 1 – World Ocean Atlas 2009	−0.9120348	0.0375	0.0022	0.2543	0.0269
Method 2 – Kim and O'Neil, 1997	−0.9494287	0.0360	0.0018	0.2677	0.0219
Method 2 – Shackleton, 1974	−0.9493258	0.0415	0.0021	0.2054	0.0253
Method 3 – Shackleton, 1974	−0.947849	0.0397	0.0021	0.2259	0.0255
Method 3 – Shackleton, 1974 – disequilibrium-corrected	−0.9487466	0.0411	0.0021	0.2056	0.0263

REFERENCES

Bajnai D., Fiebig J., Tomasovych A., Milner Garcia S., Rollion-Bard C., Raddatz J., Loffler N., Primo-Ramos C. and Brand U. (2018) Assessing kinetic fractionation in brachiopod calcite using clumped isotopes. *Sci. Rep.* **8**, 533.

Barker S., Greaves M. and Elderfield H. (2003) A study of cleaning procedures used for foraminiferal Mg/Ca paleothermometry. *Geochem. Geophys. Geosyst.* **4**, n/a-n/a.

Bemis B. E., Spero H. J., Bijma J. and Lea D. W. (1998) Reevaluation of the oxygen isotopic composition of planktonic foraminifera: experimental results and revised paleotemperature equations. *Paleoceanography* **13**, 150–160.

Bernasconi S. M., Hu B., Wacker U., Fiebig J., Breitenbach S. F. and Rutz T. (2013) Background effects on Faraday collectors in gas-source mass spectrometry and implications for clumped isotope measurements. *Rapid Commun. Mass Spectrom.* **27**, 603–612.

Bernasconi S. M., Müller I. A., Bergmann K. D., Breitenbach S. F. M., Fernandez A., Hodell D. A., Jaggi M., Meckler A. N., Millan I. and Ziegler M. (2018) Reducing uncertainties in carbonate clumped isotope analysis through consistent carbonate-based standardization. *Geochem. Geophys. Geosyst.*.

Bigeleisen J. and Mayer M. G. (1947) Calculation of equilibrium constants for isotopic exchange reactions. *J. Chem. Phys.* **15**, 261.

Bonifacie M., Calmels D., Eiler J. M., Horita J., Chaduteau C., Vasconcelos C., Agrinier P., Katz A., Passey B. H., Ferry J. M. and Bourrand J.-J. (2017) Calibration of the dolomite clumped isotope thermometer from 25 to 350°C, and implications for a universal calibration for all (Ca, Mg, Fe)CO3 carbonates. *Geochim. Cosmochim. Acta* **200**, 255–279.

Brand W. A., Assonov S. S. and Coplen T. B. (2010) Correction for the 17O interference in δ (13C) measurements when analyzing CO2 with stable isotope mass spectrometry (IUPAC Technical Report). *Pure Appl. Chem.* **82**, 1719–1733.

Brassell S., Eglinton G., Marlowe I., Pflaumann U. and Sarnthein M. (1986) Molecular stratigraphy: a new tool for climatic assessment. *Nature* **320**, 129–133.

Breitenbach S. F. M., Mleneck-Vautravers M. J., Grauel A.-L., Lo L., Bernasconi S. M., Müller I. A., Rolfe J., Gázquez F., Greaves M. and Hodell D. A. (2018) Coupled Mg/Ca and clumped isotope analyses of foraminifera provide consistent water temperatures. *Geochim. Cosmochim. Acta*.

Cléroux C., deMenocal P., Arbuszewski J. and Linsley B. (2013) Reconstructing the upper water column thermal structure in the Atlantic Ocean. *Paleoceanography* **28**, 503–516.

Curry W. B., Thunell R. C. and Honjo S. (1983) Seasonal changes in the isotopic composition of planktonic foraminifera collected in Panama Basin sediment traps. *Earth Planet. Sci. Lett.* **64**, 33–43.

Daëron M., Blamart D., Peral M. and Affek H. (2016) Absolute isotopic abundance ratios and the accuracy of Δ47 measurements. *Chem. Geol.* **442**, 83–96.

Daeron M., Drysdale R. N., Peral M., Huyghe D., Blamart D., Coplen T. B., Lartaud F. and Zanchetta G. (2019) Most Earth-surface calcites precipitate out of isotopic equilibrium. *Nat. Commun.* **10**, 429.

Davies A. J. and John C. M. (2019) The clumped (13C18O) isotope composition of echinoid calcite: further evidence for "vital effects" in the clumped isotope proxy. *Geochim. Cosmochim. Acta* **245**, 172–189.

de Nooijer L. J., Spero H. J., Erez J., Bijma J. and Reichart G. J. (2014) Biomineralization in perforate foraminifera. *Earth Sci. Rev.* **135**, 48–58.

Defliese W. F., Hren M. T. and Lohmann K. C. (2015) Compositional and temperature effects of phosphoric acid fractionation on Δ47 analysis and implications for discrepant calibrations. *Chem. Geol.* **396**, 51–60.

Dennis K. J., Affek H. P., Passey B. H., Schrag D. P. and Eiler J. M. (2011) Defining an absolute reference frame for 'clumped' isotope studies of CO2. *Geochim. Cosmochim. Acta* **75**, 7117–7131.

Deuser W. and Ross E. (1989) Seasonally abundant planktonic foraminifera of the Sargasso Sea; succession, deep-water fluxes, isotopic compositions, and paleoceanographic implications. *J. Foraminiferal Res.* **19**, 268–293.

Eiler J. M. (2007) "Clumped-isotope" geochemistry—the study of naturally-occurring, multiply-substituted isotopologues. *Earth Planet. Sci. Lett.* **262**, 309–327.

Eiler J. M. and Schauble E. (2004) 18O13C16O in Earth's atmosphere. *Geochim. Cosmochim. Acta* **68**, 4767–4777.

Emiliani C. (1966) Isotopic paleotemperatures. *Science* **154**, 851–857.

Epstein S., Buchsbaum R., Lowenstam H. and Urey H. C. (1951) Carbonate-water isotopic temperature scale. *GSA Bull.* **62**, 417–426.

Epstein S., Buchsbaum R., Lowenstam H. A. and Urey H. C. (1953) Revised carbonate-water isotopic temperature scale. *Geol. Soc. Am. Bull.* **64**, 1315–1326.

Evans D., Sagoo N., Renema W., Cotton L. J., Müller W., Todd J. A., Saraswati P. K., Stassen P., Ziegler M., Pearson P. N., Valdes P. J. and Affek H. P. (2018) Eocene greenhouse climate revealed by coupled clumped isotope-Mg/Ca thermometry. *Proceedings of the National Academy of Sciences*.

Ezard T. H. G., Edgar K. M. and Hull P. M. (2015) Environmental and biological controls on size-specific δ13C and δ18O in recent planktonic foraminifera. *Paleoceanography* **30**, 151–173.

Fernandez A., Müller I. A., Rodríguez-Sanz L., van Dijk J., Looser N. and Bernasconi S. M. (2017) A reassessment of the precision of carbonate clumped isotope measurements: implications for calibrations and paleoclimate reconstructions. *Geochem. Geophys. Geosyst.* **18**, 4375–4386.

Fujiki T., Takagi H., Kimoto K., Kurasawa A., Yuasa T. and Mino Y. (2014) Assessment of algal photosynthesis in planktic foraminifers by fast repetition rate fluorometry. *J. Plankton Res.* **36**, 1403–1407.

Ghosh P., Adkins J., Affek H., Balta B., Guo W., Schauble E. A., Schrag D. and Eiler J. M. (2006) 13C–18O bonds in carbonate minerals: a new kind of paleothermometer. *Geochim. Cosmochim. Acta* **70**, 1439–1456.

Ghosh P., Eiler J., Campana S. E. and Feeney R. F. (2007) Calibration of the carbonate 'clumped isotope' paleothermometer for otoliths. *Geochim. Cosmochim. Acta* **71**, 2736–2744.

Gonfiantini R., Stichler W. and Rozanski K. (1995) Reference and intercomparison materials for stable isotopes of light elements. *IAEA TECDOC* **825**, 67–74.

Grauel A.-L., Schmid T. W., Hu B., Bergami C., Capotondi L., Zhou L. and Bernasconi S. M. (2013) Calibration and application of the 'clumped isotope' thermometer to foraminifera for high-resolution climate reconstructions. *Geochim. Cosmochim. Acta* **108**, 125–140.

Groeneveld J. and Chiessi C. M. (2011) Mg/Ca ofGloborotalia inflata as a recorder of permanent thermocline temperatures in the South Atlantic. *Paleoceanography* **26**, n/a-n/a.

Henkes G. A., Passey B. H., Wanamaker A. D., Grossman E. L., Ambrose W. G. and Carroll M. L. (2013) Carbonate clumped isotope compositions of modern marine mollusk and brachiopod shells. *Geochim. Cosmochim. Acta* **106**, 307–325.

Ho S. L., Mollenhauer G., Fietz S., Martínez-Garcia A., Lamy F., Rueda G., Schipper K., Méheust M., Rosell-Melé A., Stein R. and Tiedemann R. (2014) Appraisal of TEX86 and TEX86L thermometries in subpolar and polar regions. *Geochim. Cosmochim. Acta* **131**, 213–226.

Hu B., Radke J., Schluter H. J., Heine F. T., Zhou L. and Bernasconi S. M. (2014) A modified procedure for gas-source isotope ratio mass spectrometry: the long-integration dual-inlet (LIDI) methodology and implications for clumped isotope measurements. *Rapid Commun. Mass Spectrom.* **28**, 1413–1425.

Huntington K. W., Eiler J. M., Affek H. P., Guo W., Bonifacie M., Yeung L. Y., Thiagarajan N., Passey B., Tripati A., Daeron M. and Came R. (2009) Methods and limitations of 'clumped' CO2 isotope (Delta47) analysis by gas-source isotope ratio mass spectrometry. *J. Mass Spectrom.* **44**, 1318–1329.

Hut, G. (1987) Consultants' group meeting on stable isotope reference samples for geochemical and hydrological investigations.

Hönisch B. and Hemming N. G. (2004) Ground-truthing the boron isotope-paleo-pH proxy in planktonic foraminifera shells: partial dissolution and shell size effects. *Paleoceanography* **19**.

Jentzen A., Nürnberg D., Hathorne E. C. and Schönfeld J. (2018) Mg/Ca and δ18O in living planktic foraminifers from the Caribbean, Gulf of Mexico and Florida Straits. *Biogeosciences* **15**, 7077–7095.

John C. M. and Bowen D. (2016) Community software for challenging isotope analysis: first applications of 'Easotope' to clumped isotopes. *Rapid Commun. Mass Spectrom.* **30**, 2285–2300.

Kele S., Breitenbach S. F. M., Capezzuoli E., Meckler A. N., Ziegler M., Millan I. M., Kluge T., Deák J., Hanselmann K., John C. M., Yan H., Liu Z. and Bernasconi S. M. (2015) Temperature dependence of oxygen- and clumped isotope fractionation in carbonates: a study of travertines and tufas in the 6–95°C temperature range. *Geochim. Cosmochim. Acta* **168**, 172–192.

Kelson J. R., Huntington K. W., Schauer A. J., Saenger C. and Lechler A. R. (2017) Toward a universal carbonate clumped isotope calibration: diverse synthesis and preparatory methods suggest a single temperature relationship. *Geochim. Cosmochim. Acta* **197**, 104–131.

Kim S.-T. and O'Neil J. R. (1997) Equilibrium and nonequilibrium oxygen isotope effects in synthetic carbonates. *Geochim. Cosmochim. Acta* **61**, 3461–3475.

Laskar A. H. and Liang M.-C. (2016) Clumped isotopes in near-surface atmospheric CO 2 over land, coast and ocean in Taiwan and its vicinity. *Biogeosciences* **13**, 5297–5314.

Lea D. W., Mashiotta T. A. and Spero H. J. (1999) Controls on magnesium and strontium uptake in planktonic foraminifera determined by live culturing. *Geochim. Cosmochim. Acta* **63**, 2369–2379.

LeGrande A. N. and Schmidt G. A. (2006) Global gridded data set of the oxygen isotopic composition in seawater. *Geophys. Res. Lett.* **33**.

Leutert T. J., Sexton P. F., Tripati A., Piasecki A., Ho S. L. and Meckler A. N. (2019) Sensitivity of clumped isotope temperatures in fossil benthic and planktic foraminifera to diagenetic alteration. *Geochim. Cosmochim. Acta* **257**, 354–372.

Locarnini R., Mishonov A., Antonov J., Boyer T., Garcia H., Baranova O., Zweng M. and Johnson D. (2010) *World Ocean Atlas 2009, Volume 1: Temperature. S. Levitus, Ed. NOAA Atlas NESDIS 68*. US Government Printing Office, Washington, DC, p. 184.

Marchitto T. M., Curry W. B., Lynch-Stieglitz J., Bryan S. P., Cobb K. M. and Lund D. C. (2014) Improved oxygen isotope temperature calibrations for cosmopolitan benthic foraminifera. *Geochim. Cosmochim. Acta* **130**, 1–11.

Meckler A. N., Ziegler M., Millan M. I., Breitenbach S. F. and Bernasconi S. M. (2014) Long-term performance of the Kiel carbonate device with a new correction scheme for clumped isotope measurements. *Rapid Commun. Mass Spectrom.* **28**, 1705–1715.

Molina-Kescher M., Frank M. and Hathorne E. C. (2014) Nd and Sr isotope compositions of different phases of surface sediments in the South Pacific: Extraction of seawater signatures, boundary exchange, and detrital/dust provenance. *Geochem. Geophys. Geosyst.* **15**, 3502–3520.

Mulitza S., Dürkoop A., Hale W., Wefer G. and Stefan Niebler H. (1997) Planktonic foraminifera as recorders of past surface-water stratification. *Geology* **25**, 335.

Niebler H.-S., Hubberten H.-W. and Gersonde R. (1999) Oxygen isotope values of planktic foraminifera: a tool for the reconstruction of surface water stratification. In *Use of Proxies in Paleoceanography*. Springer, pp. 165–189.

Nürnberg D., Bijma J. and Hemleben C. (1996) Assessing the reliability of magnesium in foraminiferal calcite as a proxy for water mass temperatures. *Geochim. Cosmochim. Acta* **60**, 803–814.

O'Neil J. R., Clayton R. N. and Mayeda T. K. (1969) Oxygen isotope fractionation in divalent metal carbonates. *J. Chem. Phys.* **51**, 5547–5558.

Pearson P. N. (2012) Oxygen isotopes in foraminifera: overview and historical review. *Paleontol. Soc. Pap.* **18**, 1–38.

Peral M., Daëron M., Blamart D., Bassinot F., Dewilde F., Smialkowski N., Isguder G., Bonnin J., Jorissen F., Kissel C., Michel E., Vázquez Riveiros N. and Waelbroeck C. (2018) Updated calibration of the clumped isotope thermometer in planktonic and benthic foraminifera. *Geochim. Cosmochim. Acta*.

Petersen S. V., Defliese W. F., Saenger C., Daëron M., Huntington K. W., John C. M., Kelson J. R., Bernasconi S. M., Coleman A. S., Kluge T., Olack G. A., Schauer A. J., Bajnai D., Bonifacie M., Breitenbach S. F. M., Fiebig J., Fernandez A. B., Henkes G. A., Hodell D., Katz A., Kele S., Lohmann K. C., Passey B. H., Peral M. Y., Petrizzo D. A., Rosenheim B. E., Tripati A., Venturelli R., Young E. D. and Winkelstern I. Z. (2019) Effects of improved 17O correction on inter-laboratory agreement in clumped isotope calibrations, estimates of mineral-specific offsets, and temperature dependence of acid digestion fractionation. *Geochem. Geophys. Geosyst.*.

Piasecki A., Bernasconi S. M., Grauel A.-L., Hannisdal B., Ho S. L., Leutert T. J., Marchitto T. M., Meinicke N., Tisserand A. and Meckler N. (2019) Application of clumped isotope thermometry to benthic foraminifera. *Geochem. Geophys. Geosyst.*.

Polik C. A., Elling F. J. and Pearson A. (2018) Impacts of paleoecology on the TEX86 sea surface temperature proxy in the pliocene-pleistocene mediterranean sea. *Paleoceanogr. Paleoclimatol.* **33**, 1472–1489.

Prahl F. G., Muehlhausen L. A. and Zahnle D. L. (1988) Further evaluation of long-chain alkenones as indicators of paleoceanographic conditions. *Geochim. Cosmochim. Acta* **52**, 2303–2310.

Raddatz J., Nürnberg D., Tiedemann R. and Rippert N. (2017) Southeastern marginal West Pacific Warm Pool sea-surface and thermocline dynamics during the Pleistocene (2.5–0.5 Ma). *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **471**, 144–156.

Regenberg M., Nürnberg D., Steph S., Groeneveld J., Garbe-Schönberg D., Tiedemann R. and Dullo W.-C. (2006) Assessing the effect of dissolution on planktonic foraminiferal Mg/Ca ratios: Evidence from Caribbean core tops. *Geochem. Geophys. Geosyst.* **7**.

Regenberg M., Steph S., Nürnberg D., Tiedemann R. and Garbe-Schönberg D. (2009) Calibrating Mg/Ca ratios of multiple planktonic foraminiferal species with δ18O-calcification temperatures: paleothermometry for the upper water column. *Earth Planet. Sci. Lett.* **278**, 324–336.

Rink S., Kühl M., Bijma J. and Spero H. (1998) Microsensor studies of photosynthesis and respiration in the symbiotic foraminifer Orbulina universa. *Mar. Biol.* **131**, 583–595.

Rippert N., Nürnberg D., Raddatz J., Maier E., Hathorne E., Bijma J. and Tiedemann R. (2016) Constraining foraminiferal calcification depths in the western Pacific warm pool. *Mar. Micropaleontol.* **128**, 14–27.

Saenger C., Affek H. P., Felis T., Thiagarajan N., Lough J. M. and Holcomb M. (2012) Carbonate clumped isotope variability in shallow water corals: temperature dependence and growth-related vital effects. *Geochim. Cosmochim. Acta* **99**, 224–242.

Schauble E. A., Ghosh P. and Eiler J. M. (2006) Preferential formation of 13C–18O bonds in carbonate minerals, estimated using first-principles lattice dynamics. *Geochim. Cosmochim. Acta* **70**, 2510–2529.

Schauer A. J., Kelson J., Saenger C. and Huntington K. W. (2016) Choice of (17) O correction affects clumped isotope (Delta47) values of CO2 measured with mass spectrometry. *Rapid Commun. Mass Spectrom.* **30**, 2607–2616.

Schiebel R. and Hemleben C. (2017) *Planktic Foraminifers in the Modern Ocean*. Springer.

Schlitzer, R. (2018) Ocean data view. http://odv.awi.de.

Schmid T. W. and Bernasconi S. M. (2010) An automated method for 'clumped-isotope' measurements on small carbonate samples. *Rapid Commun. Mass Spectrom.* **24**, 1955–1963.

Schmuker B. and Schiebel R. (2002) Planktic foraminifers and hydrography of the eastern and northern Caribbean Sea. *Mar. Micropaleontol.* **46**, 387–403.

Schouten S., Hopmans E. C., Schefuß E. and Sinninghe Damsté J. S. (2002) Distributional variations in marine crenarchaeotal membrane lipids: a new tool for reconstructing ancient sea water temperatures? *Earth Planet. Sci. Lett.* **204**, 265–274.

Shackleton, N. (1974) Attainment of isotopic equilibrium between ocean water and the benthonic foraminifera genus Uvigerina: isotopic changes in the ocean during the last glacial.

Shackleton N. J., Wiseman J. D. H. and Buckley H. A. (1973) Non-equilibrium isotopic fractionation between seawater and planktonic foraminiferal tests. *Nature* **242**, 177.

Spencer C. and Kim S.-T. (2015) Carbonate clumped isotope paleothermometry: a review of recent advances in CO2 gas evolution, purification, measurement and standardization techniques. *Geosci. J.* **19**, 357–374.

Spero H. J., Bijma J., Lea D. W. and Bemis B. E. (1997) Effect of seawater carbonate concentration on foraminiferal carbon and oxygen isotopes. *Nature* **390**, 497.

Steph S., Regenberg M., Tiedemann R., Mulitza S. and Nürnberg D. (2009) Stable isotopes of planktonic foraminifera from tropical Atlantic/Caribbean core-tops: Implications for reconstructing upper ocean stratification. *Mar. Micropaleontol.* **71**, 1–19.

Takagi H., Kimoto K., Fujiki T., Saito H., Schmidt C., Kucera M. and Moriya K. (2019) Characterizing photosymbiosis in modern planktonic foraminifera. *Biogeosciences* **16**, 3377–3396.

Tang J., Dietzel M., Fernandez A., Tripati A. K. and Rosenheim B. E. (2014) Evaluation of kinetic effects on clumped isotope fractionation (Δ47) during inorganic calcite precipitation. *Geochim. Cosmochim. Acta* **134**, 120–136.

Tripati A. K., Eagle R. A., Thiagarajan N., Gagnon A. C., Bauch H., Halloran P. R. and Eiler J. M. (2010) 13C–18O isotope signatures and 'clumped isotope' thermometry in foraminifera and coccoliths. *Geochim. Cosmochim. Acta* **74**, 5697–5717.

Tripati A. K., Hill P. S., Eagle R. A., Mosenfelder J. L., Tang J., Schauble E. A., Eiler J. M., Zeebe R. E., Uchikawa J., Coplen T. B., Ries J. B. and Henry D. (2015) Beyond temperature: clumped isotope signatures in dissolved inorganic carbon species and the influence of solution chemistry on carbonate mineral composition. *Geochim. Cosmochim. Acta* **166**, 344–371.

Tripati A. K., Sahany S., Pittman D., Eagle R. A., Neelin J. D., Mitchell J. L. and Beaufort L. (2014) Modern and glacial tropical snowlines controlled by sea surface temperature and atmospheric mixing. *Nat. Geosci.* **7**, 205.

Turich C., Freeman K. H., Bruns M. A., Conte M., Jones A. D. and Wakeham S. G. (2007) Lipids of marine archaea: patterns and provenance in the water-column and sediments. *Geochim. Cosmochim. Acta* **71**, 3272–3291.

Urey H. C. (1947) The thermodynamic properties of isotopic substances. *J. Chem. Soc. (Resumed)*, 562–581.

Wacker U., Fiebig J., Tödter J., Schöne B. R., Bahr A., Friedrich O., Tütken T., Gischler E. and Joachimski M. M. (2014) Empirical calibration of the clumped isotope paleothermometer using calcites of various origins. *Geochim. Cosmochim. Acta* **141**, 127–144.

Wang L. (2000) Isotopic signals in two morphotypes of Globigerinoides ruber (white) from the South China Sea: implications for monsoon climate change during the last glacial cycle. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **161**, 381–394.

Wang Z., Schauble E. A. and Eiler J. M. (2004) Equilibrium thermodynamics of multiply substituted isotopologues of molecular gases. *Geochim. Cosmochim. Acta* **68**, 4779–4797.

Watkins J. M. and Hunt J. D. (2015) A process-based model for non-equilibrium clumped isotope effects in carbonates. *Earth Planet. Sci. Lett.* **432**, 152–165.

York D., Evensen N. M., Martınez M. L. and De Basabe Delgado J. (2004) Unified equations for the slope, intercept, and standard errors of the best straight line. *Am. J. Phys.* **72**, 367–375.

Yu J., Thornalley D. J., Rae J. W. and McCave N. I. (2013) Calibration and application of B/Ca, Cd/Ca, and δ11B in Neogloboquadrina pachyderma (sinistral) to constrain CO2 uptake in the subpolar North Atlantic during the last deglaciation. *Paleoceanography* **28**, 237–252.

Zaarur S., Affek H. P. and Brandon M. T. (2013) A revised calibration of the clumped isotope thermometer. *Earth Planet. Sci. Lett.* **382**, 47–57.

Zachos J., Pagani M., Sloan L., Thomas E. and Billups K. (2001) Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* **292**, 686–693.

Zeebe R. E. (1999) An explanation of the effect of seawater carbonate concentration on foraminiferal oxygen isotopes. *Geochim. Cosmochim. Acta* **63**, 2001–2007.

Associate editor: Cedric Michael John