Geochemistry, Geophysics, Geosystems

RESEARCH ARTICLE

10.1029/2018GC007961

Special Section:
Clumped Isotope Geochemistry: From Theory to Applications

Key Points:
- There are no vital effect offsets of benthic foraminifera from abiogenic carbonate materials for clumped isotope temperature proxy
- There are no discernible species-specific offsets for benthic foraminifera
- The clumped isotope proxy is robust through deep time for paleoceanographic reconstructions

Supporting Information:
- Supporting Information S1
- Figure S1
- Figure S2
- Figure S3
- Figure S4
- Figure S5
- Figure S6
- Figure S7
- Tables S1 and S2

Correspondence to:
A. Piasecki,
piasecki@fas.harvard.edu

Citation:
Piasecki, A., Bernasconi, S. M., Grauel, A.-L., Hannisdal, B., Ho, S. L., Leutert, T. J., et al. (2019). Application of clumped isotope thermometry to benthic foraminifera. *Geochemistry, Geophysics, Geosystems*, *20*, 2082–2090. https://doi.org/10.1029/2018GC007961

Received 17 SEP 2018
Accepted 28 MAR 2019
Accepted article online 1 APR 2019
Published online 25 APR 2019

Application of Clumped Isotope Thermometry to Benthic Foraminifera

Alison Piasecki[1,2], Stefano M. Bernasconi[3], Anna-Lena Grauel[4], Bjarte Hannisdal[1], Sze Ling Ho[1,5], Thomas J. Leutert[1], Thomas M. Marchitto[6], Niklas Meinicke[1], Amandine Tisserand[7], and Nele Meckler[1]

[1]Department of Earth Sciences, University of Bergen and Bjerknes Centre for Climate Research, Bergen, Norway, [2]Now at Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, USA, [3]Department of Earth Sciences, Geological Institute, ETH Zürich, Zürich, Switzerland, [4]Institute of Applied Geosciences, Technical University of Gräz, Gräz, Austria, [5]Now at Institute of Oceanography, National Taiwan University (NTU), Taipei, Taiwan, [6]Department of Geological Sciences, University of Colorado Boulder, Boulder, CO, USA, [7]NORCE Norwegian Research Centre AS and Bjerknes Centre for Climate Research, Bergen, Norway

Abstract Obtaining absolute temperatures of the ocean in deep time is complicated by the lack of constraints on seawater chemistry. Seawater salinity, carbonate ion concentration, δ^{18}O, and elemental abundance changes may obscure widely applied paleoproxies. In addition, with foraminifera-based proxies applied over long time scales or through major transitions, taxonomic turnover can impair the robustness of a record. While requiring larger sample sizes than most other proxies, the clumped isotope method is independent of seawater chemistry. Here we test if small benthic foraminifera precipitate their carbonate in equilibrium with respect to the clumped isotope thermometer and if there are any species-specific vital effects. We find that benthic foraminifera fall on the same calibration line as the majority of carbonate minerals including inorganic calcite. In addition, we find no offsets that can be attributed to a species-specific for any of the samples. This finding implies that a necessary amount of sample material can be obtained by aggregating over multiple taxa of benthic foraminifera and allows for the application of this proxy over major climatic transitions that coincide with seawater chemistry changes and foraminifera extinctions.

1. Introduction

Clumped isotopes are a powerful tool that can record more information than just single isotope enrichments. Clumped isotopes involve measuring two different isotopic enrichments (i.e., δ^{18}O and δ^{13}C) adjacent to one another structurally within the same molecule. The method was first used on CO_2 derived from carbonate material (Eiler, 2007) and is now applied to a wide range of other molecules like O_2 (Yeung et al., 2009), methane, and other organic molecules (Clog et al., 2018; Piasecki et al., 2016; Stolper et al., 2014) and to small portions of larger molecules relevant to a wide range of geologic applications. Clumped isotope thermometry, specifically in this case applied to carbonate minerals, operates under the assumption that the material is precipitated in equilibrium, and its isotope distribution, which reflects formation temperature, is not altered by diagenesis until it is measured in the laboratory. Benthic foraminifera, which have many important paleoclimatic applications, have carbonate shells, but these are precipitated through a biologically mediated pathway. In this study, we aim to demonstrate that benthic foraminifera have the same response to temperature of their clumped isotope value as inorganically precipitated carbonates. Due to previous limitations on sample size, this question has not been adequately probed, and it has not been shown if small benthic foraminifera precipitate their carbonate tests in equilibrium with respect to the clumped isotope temperatures relevant to most benthic foraminifera. In addition, we measured 10 different species to test for any species-specific vital effects in benthic foraminifera.

2. Background

One of the most powerful aspects of clumped isotopes, especially in comparison to other paleoceanographic temperature proxies, is that the expected value can be explicitly calculated using first principles (Wang et al., 2004). While these quantum mechanical calculations can yield a direction and order of magnitude of the expected fractionation, they are not sufficiently accurate in solid phases, and therefore, a calibration is needed to accurately transform measured values of CO_2 evolved from carbonate in the laboratory to an absolute temperature scale. More traditional temperature proxies applied to benthic foraminifera cannot explicitly calculate temperature over long time scales due to other dependencies. In addition to temperature, δ^{18}O is dependent on the isotopic composition of seawater (McCrea, 1950); Mg/Ca is affected by seawater Mg/Ca and carbonate ion concentration (Lear et al., 2010), but both proxies are influenced by vital effects (Lear et al., 2002; Marchitto et al., 2014). These proxies have their advantages, mostly that they can generate high-resolution records in a variety of different environments within the recent geological past, but there are complications applying them in geological deep time, when seeking absolute temperatures, due to the requirement of species-specific calibrations and assumptions of seawater composition. In addition, taxonomic turnover can pose problems when applying proxies affected by vital effects over long-term studies or across transitions characterized by assemblage changes. While there are calibrations to compare δ^{18}O between different species and calibrations to estimate temperature using Mg/Ca in different species, by applying these proxies to extinct species or through transitions with assemblage changes, large uncertainties are introduced.

The drawback of clumped isotope thermometry is that the typical sample size per replicate in conventional extraction lines is around 3–7 mg of carbonate, with several replicates needed. This is difficult to achieve in a paleoceanographic study using single benthic foraminiferal species, except in the most special circumstances, without aggregating samples across long time periods which diminishes the temporal range of the proxy. There have been new advancements with the technique however (Hu et al., 2014; Meckler et al., 2014; Müller et al., 2017; Schmid & Bernasconi, 2010) that allow to complete multiple replicates (20 or more) of smaller size (100–140 μg) of cleaned carbonate sample, greatly reducing the total mass of carbonate needed to obtain a single robust temperature estimate (Meckler et al., 2014). These improvements now allow us to probe the potential species-specific offsets of benthic foraminifera. Clumped isotope thermometry can now be applied to a wide variety of paleoceanographic and paleoclimatic questions (Breitenbach et al., 2018; Grauel et al., 2013; Rodríguez-Sanz et al., 2017).

There are two recent calibration studies of carbonate minerals that are directly relevant to this study. Kele et al. (2015) used a Kiel preparation system (described in detail in following sections) set up for small sample sizes in order to conduct the calibration. The inorganically precipitated travertines and tufa used in their calibration cover a large temperature range of 0–95 °C (Kele et al., 2015). However, their study does not address potential complications of biologically precipitated carbonate minerals from foraminifera with a sufficient number of samples. Recently, due to adjusted parameters, the Kele et al. (2015) study was recalculated, using the same standard values as our data treatment (Bernasconi et al., 2018). Bonifacie et al. (2017) incorporated most other existing calibration studies (including the one of Kele et al., 2015) with dolomite samples spanning a temperature range up to 350 °C and proposed a single common calibration line across sample types. Due to the very large temperature range and due to the fact that it contains other relevant calibration studies, we use the Bonifacie et al. (2017) study as the benchmark to which we compare our data. Their study includes samples that are prepared with different analytical methods from a number of different laboratories and a variety of different sample types.

Three calibration studies focus more on clumped isotope thermometry of foraminifera. The first study looking at benthic foraminifera includes a mixed sample, together with planktic foraminifera and coccoliths, and was greatly limited by sample size due to it using a common acid bath dissolution for the carbonate measurements (Tripati et al., 2010). The second study (Grauel et al., 2013) included a few benthic foraminifera samples, with two species of benthic foraminifera from two sites in the Mediterranean Sea and the Atlantic. We repicked two Mediterranean sites from this study and included them in our analysis (Table 1). Finally, there is a recent study that has two benthic points, while the rest are planktic species (Peral et al., 2018). The benthic species fall on the same line as the planktics and are within error of our results as well. The studies suggested that benthic and planktic foraminifera are indistinguishable from the bulk carbonate calibrations, but there were not enough measurements of different species over wide temperature ranges to assess whether benthic foraminifera have any species-dependent offsets. Comparison of this published benthic data with our benthic data is shown in supporting information Figure S3.

Table 1
Clumped Isotope Averages for Species by Site (Including $\delta^{18}O$, $\delta^{18}O$ Standard Deviation; $\delta^{13}C$, $\delta^{13}C$ Standard Deviation; Δ_{47}, Δ_{47} Standard Error, and Number of Replicates

Site	Region	Latitude	Longitude	Depth (m)	Age (year BP)	Temp (°C)	Mg/Ca, mmol/mol	Reference
13MC-G	Bahamas	24.37°N	83.24°W	348	3040	9.7	2.39	Marchitto et al. (2007)
19MC-G	Bahamas	24.42°N	83.21°W	173	1320	12.7	2.41	Marchitto et al. (2007)
50MC-G	Bahamas	24.41°N	83.22°W	198	1080	12.1	2.43	Marchitto et al. (2007)
53MC-G	Bahamas	24.38°N	83.23° W	302	1800	10	2.28	Marchitto et al. (2007)
89MC-G	Bahamas	24.56°N	79.24°W	353	2280	17.8	3.44	Marchitto et al. (2007)
94MC-G	Bahamas	24.57°N	79.23°W	259	215	18.5	3.41	Marchitto et al. (2007)
GS06-144-19	Norwegian Sea	63.83°N	5.27°E	830	—	−0.74	—	O.A.
GS07-150-17-2	Brazil	4.47°S	37.21°W	1000	796	4.17	1.91	Tisserand et al. (2013)
GS07-150-22-1	Brazil	4.33°S	37.16°W	598	2897	6.06	2.47	Tisserand et al. (2013)
MP43	Mediterranean	39.72°N	16.97°E	246	—	13.9	—	Grauel et al. (2013)
MP46	Mediterranean	39.54°N	17.25°E	582	—	13.9	—	Grauel et al. (2013)
SO213-54-4	Eastern Pacific	43.72°S	120.67°W	3840	—	1.42	—	O.A.
SO213-71-2	Western Pacific	45.58°S	157.90°W	689	—	3.77	—	World Ocean Atlas 2009, volume 1: Temperature (2010)

Note. O.A. is World Ocean Atlas 2009, volume 1: Temperature (2010).

3. Methods

3.1. Sample Selection

All samples in this study are from core top sediments, and the majority are from previous calibration studies for Mg/Ca thermometry. Since the previous calibration studies usually only focused on a single species of foraminifera that is no longer abundant in the samples, it is difficult to directly compare results. The Mg/Ca data which do exist are included in Table 1. We use the published instrumental temperatures reported by those studies as the presumed growth temperatures of the foraminifera. The temperatures are typically a combination of World Ocean Atlas temperatures or cruise conductivity temperature depth measurements when the samples were collected. Some of the samples have been radiocarbon dated previously (Table 1). The largest sample suite within this study is from the Bahamas, and the samples span 9.7–18.5 °C (Marchitto et al., 2007). These samples contained multiple abundant benthic foraminifera species in both the >355 and the 250- to 355-μm size fraction (Species listed in Table 2). Another two samples from the Mediterranean Sea come from the clumped isotope calibration study of Grauel et al. (2013). The two sites from this study that were used here experience bottom water temperatures of 13.9 °C and contained two benthic species (*Cibicidoides mundulus* and *Uvigerina mediterranea*) in sufficient abundance. A third sample set is from off the coast of Northern Brazil, where the bottom water temperature range is 4–6 °C (Tisserand et al., 2013) and three benthic species could be selected (*Pyrgo spp.* from two sites and *Cibicidoides pachyderma*). In addition, two samples with no prior information on Mg/Ca were obtained from off the Norwegian Coast and from east of New Zealand (Table 1). Temperatures for these sites are from Ocean Atlas mean annual bottom water temperatures and are −0.7 and 3.7 °C, respectively (World Ocean Atlas 2009, volume 1: Temperature, 2010). The locations of all sites are shown in Figure 1. All species and the size fractions measured are shown in Table 2. Overall growth temperatures of the tests are reasonably well constrained. Unlike for planktic species, we do not have to consider additional uncertainties related to seasonal effects or depth habitat changes, as bottom water temperatures normally vary little over space and time.

3.2. Mass Spectromemtry

At the University of Bergen, we run clumped isotopes on small carbonate samples using a Thermo Fisher Scientific Kiel IV carbonate preparation device coupled to a MAT-253 Plus isotope ratio mass spectrometer. The setup closely follows previous work (Hu et al., 2014; Meckler et al., 2014; Müller et al., 2017; Schmid & Bernasconi, 2010), and our method is summarized here. Small samples, currently 100–140 μg (dependent on instrument sensitivity), are loaded into an automatic sample carousel where a mix of 42 samples and standards are run in a 24-hr period. The samples are digested individually with phosphoric acid at 70 °C, and the generated CO_2 is collected immediately in a cold trap. Water is trapped and then the sample is passively

Table 2
Averaged Clumped Isotope Results by Site Then Species (Including $\delta^{18}O$, $\delta^{18}O$ Standard Error; $\delta^{13}C$, $\delta^{13}C$ Standard Error; Δ_{47}, Δ_{47} Standard Error, and Number of Replicates)

Site	Type	Size (μm)	$\delta^{18}O$	$\delta^{18}O_{ster}$	$\delta^{13}C$	$\delta^{13}C_{ster}$	Δ_{47}	Δ_{47ster}	Num	T
13MC-G	*Cibicidoides pachyderma*	>355	1.526	0.135	0.606	0.024	0.741	0.008	14	9.7
13MC-G	*Hoeglundina elegans*	250–355	2.420	0.000	1.820	0.000	0.723	0.000	1	9.7
13MC-G	*Hoeglundina elegans*	>355	2.695	0.039	1.918	0.031	0.749	0.014	14	9.7
13MC-G	*Lenticulina convergens*	>355	1.729	0.044	−0.048	0.033	0.721	0.011	9	9.7
13MC-G	*Lenticulina* iota	>355	1.521	0.056	0.104	0.098	0.719	0.009	10	9.7
13MC-G	*Pyrgo serrata*	>355	2.533	0.027	0.564	0.047	0.729	0.014	10	9.7
13MC-G	*Uvigerina peregrina*	>355	1.835	0.011	0.087	0.014	0.735	0.009	15	9.7
13MC-G	*Uvigerina peregrina*	250–355	1.481	0.164	0.239	0.118	0.736	0.007	33	9.7
19MC-G	*Amphisterigina radiata*	>355	−0.283	0.050	0.279	0.062	0.692	0.006	15	12.7
19MC-G	*Hoeglundina elegans*	>355	1.700	0.000	2.000	0.000	0.777	0.000	1	12.7
19MC-G	*Lenticulina convergens*	>355	0.298	0.058	−0.568	0.132	0.728	0.010	13	12.7
19MC-G	*Lenticulina iota*	>355	1.014	0.020	0.145	0.042	0.728	0.005	13	12.7
19MC-G	*Pyrgo spp*	250–355	1.328	0.179	0.770	0.097	0.704	0.021	5	12.7
50MC-G	*Cibicidoides pachyderma*	>355	1.242	0.012	0.586	0.033	0.734	0.006	33	12.1
50MC-G	*Lenticulina convergens*	>355	1.092	0.027	−0.581	0.113	0.752	0.009	11	12.1
50MC-G	*Pyrgo spp*	250–355	1.363	0.054	0.800	0.035	0.689	0.010	3	12.1
53MC-G	*Cibicidoides pachyderma*	>355	1.432	0.025	0.556	0.038	0.725	0.009	14	10
53MC-G	*Lenticulina iota*	>355	1.418	0.031	−0.154	0.061	0.732	0.004	43	10
53MC-G	*Lenticulina convergens*	>355	1.594	0.043	0.041	0.084	0.740	0.010	11	10
53MC-G	*Pyrgo serrata*	>355	2.439	0.021	0.430	0.081	0.765	0.009	20	10
89MC-G	*Cibicidoides pachyderma*	250–355	0.292	0.020	1.296	0.054	0.698	0.024	5	17.8
89MC-G	*Hoeglundina elegans*	250–355	−0.077	0.278	1.088	0.129	0.695	0.007	26	17.8
89MC-G	*Lenticulina spp*	250–355	0.258	0.049	0.672	0.111	0.713	0.010	5	17.8
89MC-G	*Pyrgo spp*	250–355	0.850	0.220	1.530	0.060	0.712	0.013	2	17.8
94MC-G	*Amphistegina lessoni*	250–355	−1.814	0.180	0.388	0.084	0.679	0.005	27	18.5
94MC-G	*Hoeglundina elegans*	250–355	0.532	0.105	1.354	0.072	0.695	0.007	21	18.5
94MC-G	*Lenticulina iota*	250–355	0.020	0.126	−0.147	0.175	0.702	0.027	3	18.5
94MC-G	*Oridorsalus umbonatus*	250–355	0.123	0.021	1.437	0.112	0.714	0.015	7	18.5
94MC-G	*Uvigerina peregrina*	250–355	0.570	0.000	0.430	0.000	0.680	0.000	1	18.5
GS06-144-19	*Melonis barleannum*	250–355	3.926	0.029	−2.345	0.053	0.778	0.008	20	−0.743
GS07-150-17-2	*Pyrgo spp*	250–355	3.058	0.079	0.992	0.028	0.758	0.006	21	4.17
GS07-150-22-1	*Cibicidoides pachyderma*	250–355	2.190	0.019	1.030	0.015	0.779	0.013	13	6.06
GS07-150-22-1	*Planulina ariminensis*	250–355	1.963	0.113	1.560	0.023	0.727	0.008	3	6.06
GS07-150-22-1	*Pyrgo spp*	250–355	2.347	0.108	0.814	0.058	0.747	0.005	42	6.06
MP43-BC	*Uvigerina mediterranea*	250–355	2.123	0.020	0.462	0.091	0.714	0.008	12	13.9
MP46-MC	*Cibicidoides mundulus*	250–355	1.631	0.016	1.540	0.037	0.742	0.010	11	13.9
MP46-MC	*Hoeglundina elegans*	250–355	2.380	0.000	2.040	0.000	0.716	0.000	1	13.9
MP46-MC	*Cibicidoides pachyderma*	250–355	1.675	0.015	1.173	0.065	0.750	0.010	6	13.9
MP46-MC	*Uvigerina mediterranea*	250–355	2.102	0.018	0.347	0.069	0.714	0.004	24	13.9
MP46-MC	*Melonis spp*	250–355	1.752	0.219	0.615	0.248	0.715	0.013	6	13.9
SO213-54-4	*Cibicidoides pachyderma*	250–355	2.930	0.024	0.380	0.039	0.747	0.012	6	1.42
SO213-71-2	*Cibicides lobatus*	250–355	1.937	0.120	1.313	0.015	0.733	0.006	3	3.77
SO213-71-2	*Uvigerina peregrina*	250–355	4.511	0.048	0.427	0.020	0.748	0.006	29	3.77

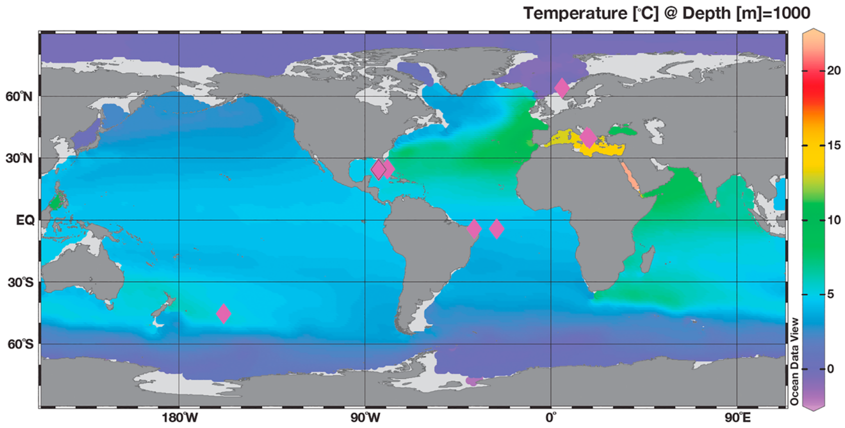

Figure 1. Map of sample locations with more than 10 replicates. There are multiple points in the Bahamas, the Mediterranean, and off New Zealand. Locations are given in Table 1. Figure made with Ocean Data View.

passed over silver wool and a PoraPak QTM (ethylvinylbenzene and divinylbenzene copolymer bead) trap in a sulfinert coated tube held at −20 °C. The PoraPak QTM column is heated to 120 °C before a run for 1 hr to clean it in between runs. After passing over the PoraPak QTM, the sample is cryofocused in a microvolume before it is passively expanded through a crimped stainless steel capillary into the mass spectrometry. The sample is measured with the long-integration dual-inlet method (Hu et al., 2014) for a total 400 s with signals typically decreasing from 18 to 10 V over the course of the measurement. After the sample is measured, the standard gas at a matching initial pressure is measured from a different depleting microvolume.

The pressure baseline is corrected based on daily peak scans at five different intensities (Meckler et al., 2014) on the day of measurement, and the samples are standardized with four carbonate standard powders of differing δ^{13}C, δ^{18}O, and Δ_{47} following Meckler et al. (2014). We made some modifications to the procedure described in Meckler et al. (2014). First, three of the standards (ETH1, ETH3, and ETH4) are used to put the samples in the absolute reference frame (Dennis et al., 2011), and the fourth (ETH2) is used as a tracer to monitor performance. The conversion into the absolute reference frame is done by a single step with three standards (listed above) without initial offset correction (we use the recalculated vaules from Bernasconi et al., 2018). In addition, we use 80 individual standard measurements to convert samples into the absolute reference frame (discussed further in the supporting information with references, e.g., Fernandez et al., 2017). The standard deviation per run of five measurements of each standard type is typically < 0.03 %0 for Δ_{47}. Furthermore, we isolate blocks of varying length (weeks to months) called correction intervals, defined by major maintenance events like source cleaning or reference gas change. For these longer-term correction intervals, the reproducibility of the standards is typically < 0.035%0 (1 SD) for Δ_{47} (supporting information Tables S1 and S2). All data are processed using the Easotope software (John & Bowen, 2016) and utilizing the Brand correction parameters (Daëron et al., 2016).

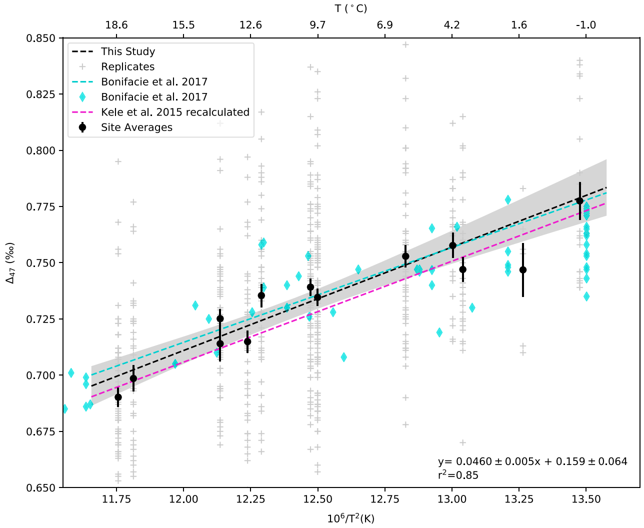

Figure 2. Data averaged per site shown in black circles with error bars indicating one standard error. Individual replicates are shown as gray crosses. Compared to the most comprehensive existing calibration of Bonifacie et al. (2017) and Kele et al. (2015) recalculated by Bernasconi et al. (2018). Data match within error.

4. Results

A total of 43 individual species measurements was made at 13 different sites, of which 27 could be replicated more than nine times, which was determined to be the threshold for a species-specific data point using counting statistics (i.e., the standard error of the replicates approaches the shot noise limit given the cumulative amount of time spent measuring mass 47). However, all replicates were used to determine the regression shown in Figure 2, which is based on averages over the results from all species from a given site. Note that there are two separate sites at 13.9 °C, both in the Mediterranean Sea.

Table 3
Statistics of our Regression and Other Recent Relevant Calibrations From the Literature

Regression	Slope	Intercept	r^2	Slope Error	Intercept Error
This study	0.046	0.159	0.85	0.005	0.064
Bonifacie et al. (2017)	0.0422	0.208	0.99	0.0019	0.0207
Kele et al. (2015) recalculated	0.0449	0.167	0.96	0.001	0.01

The site averages fit well with the overall regression. Plotted in Figure 2 are the site averages with one standard error bars as is typical with clumped isotope measurements. We also calculated error bars with 99% bootstrap confidence interval, and in that case all averages including error bars fall within the confidence interval of the line (supporting information Figure S5). Our slope (0.0460 ± 0.005, Table 3) is steeper than either the Bonifacie et al. (2017) compilation study or the Bernasconi et al. (2018) recalculation of travertine samples from Kele et al. (2015). We believe that this is due to the narrow temperature range over which our samples occur, as discussed further below.

The Bonifacie et al. (2017) curve is also shown in Figure 2. This calibration line was selected as a reference since it is currently the clumped isotope calibration that has the widest temperature range of all calibrations (Bonifacie et al., 2017). Although most of the data were obtained at a digestion 90 °C and a common acid bath, the compilation contains data points that were run at 70 °C with a Kiel preparation system (Kele et al., 2015) like our measurements, and the study converted all data to the 90 °C reference frame using the acid digestion fractionation factors determined by Defliese et al. (2015). As we are converting our data to a hypothetical acid fractionation at 25 °C (by adding 0.062 to convert from 70 °C), currently common practice in the clumped isotope community, we also converted the data from the (Bonifacie et al., 2017) calibration to the 25 °C reference frame by adding 0.082%0 (to convert from her 90 °C reference frame back to 25 °C Defliese et al., 2015). Our regression between Δ_{47} and temperature is within error of the (Bonifacie et al., 2017) calibration and the recalculated Kele et al. (2015) from Bernasconi et al. (2018). Since both calibrations cover a large range and our range is only 20 °C, their slope is better constrained and therefore is a more robust calibration. Importantly, our data suggest that benthic foraminifera fall within error of the universal calibration and do not have a measurable biological offset.

Species-specific data, using all samples with more than nine replicates, are plotted in Figure 3 to these data points in terms of species, size fraction, age, or location of sampling (Table 2). We investigate the possibility of such effects further using taxon-specific regression, which is elaborated below.

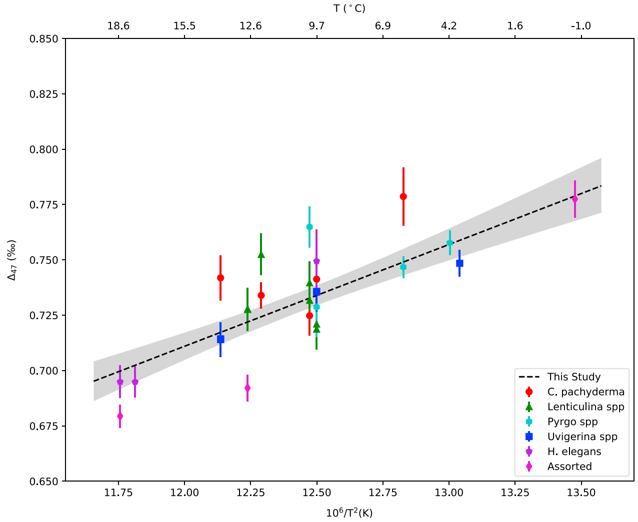

Figure 3. Figure shows genus-specific data averaged by site. Each point is a single species, but they are grouped in color by genus. Assorted refers to data from species that were only measured in one site; see Table 2. Regression is the site-averaged regression also shown in Figure 2. Inorganic calibrations fall within the gray confidence interval which are also shown in Figure 2

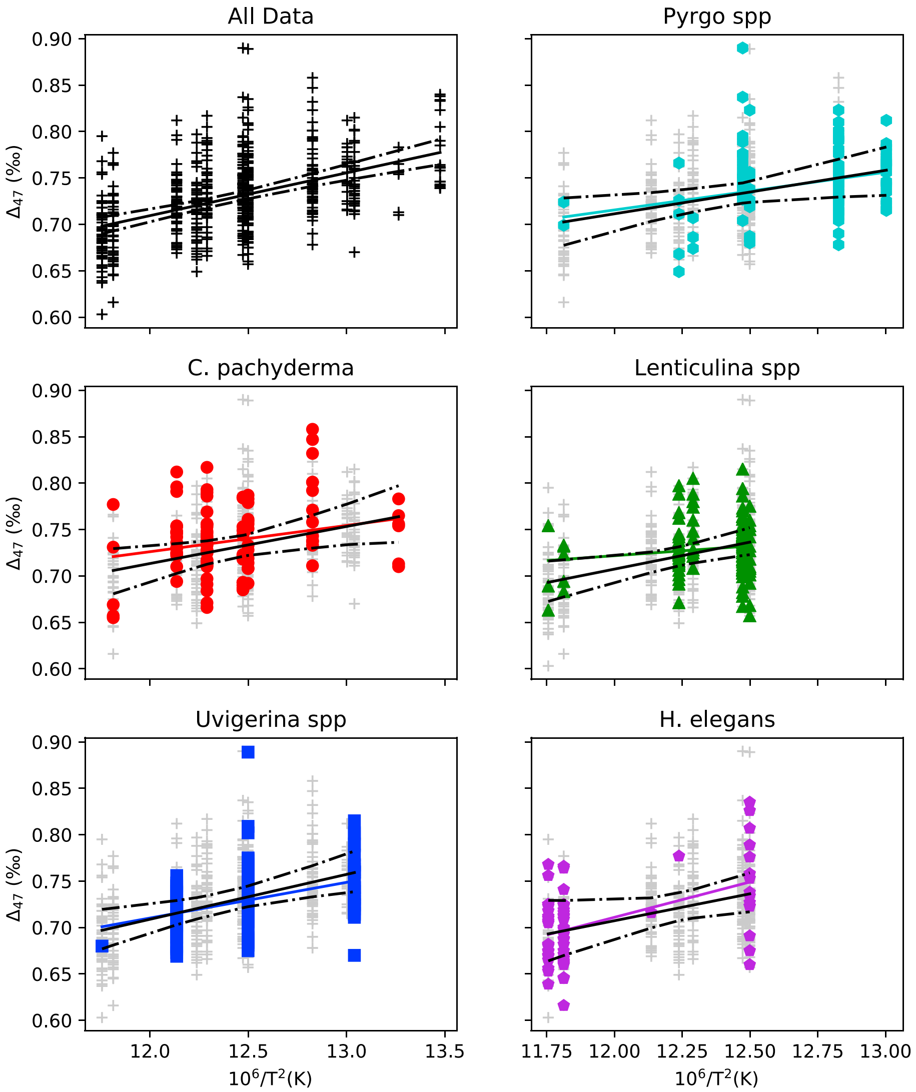

Figure 4. Plot showing the regression of all species (as in Figure 2; black solid line) and its uncertainty (99% confidence interval; black dashed line) versus a given species over the temperature range of the single-species occurrence. Gray crosses represent all species within the temperature range used to calculate the black regression and confidence interval within the panel. Regressions of individual species (colored lines) are each the median of 1,000 iterations.

5. Discussion

5.1. Species-Specific Considerations

It is necessary to determine if any there are any taxon-specific differences or if any taxon-specific regressions show obvious deviations from the average of all taxa. In order to do this, we performed a Monte Carlo resampling procedure to identify taxon-specific offsets. First, we separated the samples by genus (Figure 4). There are only a few species represented; for example, all *Hoeglundia* are *Hoeglundia elegans*, and all *Cibicidoides* are *Cibicidoides pachyderma*. *Lenticulina* is evenly split between *Lenticulina iota* and *Lenticulina convergens*. And *Pyrgo* is a mixture of *Pyrgo serrata* and others. *Uvigerina spp* is *Uvigerina peregrina* for most samples, and *U. mediterranea* for the two Mediterranean samples. In order to test for genus-specific offsets, we examined specifically that the temperature interval that each genus was measured over. We calculated a regression and bootstrap confidence interval (99%) using all data from all species measured over that temperature range. All regressions were performed using iteratively reweighted least squares with a bisquare weighting function (the robustfit function in Matlab). We estimated 0.99 confidence envelopes on regression and average values using percentiles of 10,000 bootstrap samples (randomly resampling with replacement from the original data maintaining the same samples size). From those resampled replicates we selected the median regression for each genus and compared it to the data for all species occurring over that temperature interval. For example, *Pyrgo* occur over the temperature interval of 4–18 °C, and we compared its median regression to the regression of all replicates of any species occurring within that temperature interval. As illustrated in Figure 4, all of the individual genera fall within the Δ_{47} confidence interval for the regression represented by all samples over that temperature range. Therefore, even though there is variability among a given species, no species is statistically significantly different from the average of all species over the temperature occurrence range. We suggest that the variation seen in our samples is within the range of individual replicates for other calibration studies (Figure 2). The absolute value of our regression and intercept does vary from other studies, but this is due to the narrow temperature range of our study, and therefore, we still find our data to be consistent with the other commonly used current calibration. Our method of averaging numerous small samples appears at first glance to have increased error, but when robust statistical methods are employed, the individual species's regressions all fall within error of the inorganic and all species's mean.

Overall, our data fall within error of the Bonifacie et al. (2017) calibration and within error of the Bernasconi et al. (2018) recalculation of Kele et al. (2015). These other data sets are larger. Specifically, the Bonifacie et al. (2017) compilation is very large encompassing most preparation methods of clumped isotope measurements, multiple different carbonate phases, and both biogenic and abiogenic samples measured at multiple labs. The calibration, due to all of these factors, is more robust than our study, which only has a temperature range of 20 °C and is limited by small samples of benthic foraminifera. Due to this fact, our data confirm that there are no species-specific offsets for clumped isotope temperatures and that benthic foraminifera do not as class fractionate while precipitating carbonate minerals. Since our data are skewed to one end of the available temperature range, we do not provide a good estimate for the calibration over all ranges and suggest that other authors rely on a more universal calibration especially when studying samples of greater than 20 °C. In addition, although the regression of Bonifacie et al. (2017) is more robust than ours, as shown in Figure 2, the overall spread that they see especially at lower temperatures is the same size or larger than the spread that we see for our samples (Figure 2, pale turquoise diamonds). We also compare our data to Bernasconi et al. (2018), who recalculated a previous data set of Kele et al. (2015) using the new, widely used Brand parameters. Our results fall in between these two main calibrations and are within error of both. Most current calibrations are within error of one another, including ours, which points to a universal calibration that benthic foraminifera are not biologically offset from.

5.2. Implications for Paleoceanographic Studies

Overall, this study indicates that clumped isotopes on benthic foraminifera yield reliable estimates of bottom water temperature. While there is variation between tests, it is independent of seawater chemistry, δ^{18}O, and species. It is necessary to have enough replicates such that counting statistics are reached in the measurement compared with fewer larger samples. We suggest that if it is difficult to do so with a single species, measurements done on multiple different species can be averaged at a given site and time. When using a small sample approach (e.g., Kiel device method) with multiple aliquots reacted separately, it would however be beneficial to run aliquots of individual species, as different species have different δ^{18}O values due to species-specific fractionation factors for δ^{18}O. Since Δ_{47} temperature is calculated using the δ^{18}O of the test, there is error introduced if species with very different δ^{18}O are mixed. A possible sampling scheme for paleoceanographic studies using the small sample analytical approach would be to complete few replicate analyses of downcore samples on the same foraminifera species, yielding high-resolution δ^{18}O and δ^{13}C records. And subsequently average over Δ_{47} results from 10 or more samples spaced evenly over time for a lower resolution clumped isotope temperature record (Grauel et al., 2013; Rodríguez-Sanz et al., 2017). The results obtained this way are robust against short-term temperature fluctuations and against aliasing of cyclic variations in climate, if sampled at the right frequency.

6. Conclusions

In this study we measured a range of benthic foraminifera species that occur in modern sediments from many different ocean environments. We found that our samples are consistent with other carbonate clumped isotope calibration curves presented in this work. In addition, there are no discernable species-specific offsets in the benthic foraminifera studied here. Despite variability within the data set, all species fall within the expected relationship for all samples over a given temperature range. This allows for using an aggregation of multiple single-species analyses to obtain a single temperature in cases where benthic foraminifera abundances are low.

Acknowledgments
This work was funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (agreement 638467) and by the Bergen Research Foundation. All data including every replicate and standards used to put samples in absolute reference frame are included in the supporting information.

References

Bernasconi, S. M., Müller, I. A., Bergmann, K. D., Breitenbach, S. F. M., Fernandez, A., Hodell, D. A., et al. (2018). Reducing uncertainties in carbonate clumped isotope analysis through consistent carbonate-based standardization. *Geochemistry, Geophysics, Geosystems*, *19*, 2895–2914. https://doi.org/10.1029/2017GC007385

Bonifacie, M., Calmels, D., Eiler, J, Horita, J., Chaduteau, C., Vasconcelos, C., et al. (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. *Geochimica et Cosmochimica Acta*, *200*, 255–279.

Breitenbach, SebastianF. M., Mleneck-Vautravers, M. J., Grauel, A.-L., Lo, L., Bernasconi, S. M., Müller, I. A., et al. (2018). Coupled Mg/Ca and clumped isotope analyses of foraminifera provide consistent water temperatures. *Geochimica et Cosmochimica Acta*, *236*, 283–296.

Clog, M., Lawson, M., Peterson, B., Ferreira, A. A., Neto, E. V. S., & Eiler, J. (2018). A reconnaissance study of 13 C-13 C clumping in ethane from natural gas. *Geochimica et Cosmochimica Acta*, *223*, 229–244.

Daëron, M., Blamart, D., Peral, M., & Affek, H. P. (2016). Absolute isotopic abundance ratios and the accuracy of Δ 47 measurements. *Chemical Geology*, *442*, 83–96.

Defliese, W. F., Hren, M. T., & Lohmann, K. C. (2015). Compositional and temperature effects of phosphoric acid fractionation on Δ 47 analysis and implications for discrepant calibrations. *Chemical Geology*, *396*, 51–60.

Dennis, K. J., Affek, H. P., Passey, B. H., Schrag, D. P., & Eiler, J. (2011). Defining an absolute reference frame for isotope studies of CO_2. *Geochimica et Cosmochimica Acta*, *75*, 7117–7131.

Eiler, J. (2007). "Clumped-isotope" geochemistry-The study of naturally-occurring, multiply-substituted isotopologues. *Earth and Planetary Science Letters*, *626*, 309–327.

Fernandez, A., Müller, I. A., Rodríguez-Sanz, L., van Dijk, J., Looser, N., & Bernasconi, S. M. (2017). A Reassessment of the precision of carbonate clumped isotope measurements: Implications for calibrations and paleoclimate reconstructions. *Geochemistry, Geophysics, Geosystems*, *18*, 4375–4386. https://doi.org/10.1002/2017GC007106

Grauel, A.-L., Leider, A., Goudeau, M.-L. S., Müller, I. A., Bernasconi, S. M., Hinrichs, K. U., et al. (2013). What do SST proxies really tell us? A high-resolution multiproxy ($U^{K'}_{37}$, TEX^{H}_{86} and foraminifera δ^{18} O) study in the Gulf of Taranto, central Mediterranean Sea. *Quaternary Science Reviews*, *73*, 115–131.

Hu, B., Radke, J., Schlüter, H.-J., Heine, F. T., Zhou, L., & 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 Communications in Mass Spectrometry*, *28*, 1413–1425.

John, C. M., & Bowen, D. (2016). Community software for challenging isotope analysis: First applications of "Easotope" to clumped isotopes. *Rapid Communications in Mass Spectrometry*, *30*, 2285–2300.

Kele, S., Breitenbach, S. F. M., Capezzuoli, E., Meckler, A. N., Ziegler, M., Millan, I. M., et al. (2015). Temperature dependence of oxygen- and clumped isotope fractionation in carbonates: A study of travertines and tufas in the 6-95 °C temperature range. *Geochimica et Cosmochimica Acta*, *168*, 172–192.

Lear, C. H., Mawbey, E. M., & Rosenthal, Y. (2010). Cenozoic benthic foraminiferal Mg/Ca and Li/Ca records: Toward unlocking temperatures and saturation states. *Paleoceanography*, *25*, PA4215. https://doi.org/10.1029/2009PA001880

Lear, C. H., Rosenthal, Y., & Slowey, N. (2002). Benthic foraminiferal Mg/Ca-paleothermometry: A revised core-top calibration. *Geochimica et Cosmochimica Acta*, *66*, 3375–3387.

Marchitto, T. M., Bryan, S. P., Curry, W. B., & McCorkle, D. C. (2007). Mg/Ca temperature calibration for the benthic foraminifer Cibicidoides pachyderma. *Paleoceanography*, *22*, PA1203. https://doi.org/10.1029/2006PA001287

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

McCrea, J. M. (1950). On the isotopic chemistry of carbonates and a paleotemperature scale. *The Journal of Chemical Physics*, *18*, 849–857.

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

Müller, I. A., Fernandez, A., Radke, J., van Dijk, J., Bowen, D., Schwieters, J., & Bernasconi, S. M. (2017). Carbonate clumped isotope analyses with the long-integration dual-inlet (LIDI) workflow: Scratching at the lower sample weight boundaries. *Rapid Communications in Mass Spectrometry*, *31*, 1057–1066.

Müller, I. A., Violay, M. E. S., Storck, J.-C., Fernandez, A., van Dijk, J., Madonna, C., & Bernasconi, S. M. (2017). Clumped isotope fractionation during phosphoric acid digestion of carbonates at 70 °C. *Chemical Geology*, *449*, 1–14.

Peral, M., Daëron, M., Blamart, D., Bassinot, F., Dewilde, F., Smialkowski, N., et al. (2018). Updated calibration of the clumped isotope thermometer in planktonic and benthic foraminifera. *Geochimica et Cosmochimica Acta*, *239*, 1–16.

Piasecki, A., Sessions, A. L., Lawson, M., Ferreira, AA, Neto, E. V. S., & Eiler, J. (2016). Analysis of the site-specific carbon isotope composition of propane by gas source isotope ratio mass spectrometer. *Geochimica et Cosmochimica Acta*, *188*, 58–72.

Rodríguez-Sanz, L., Bernasconi, S. M., Marino, G., Heslop, D., Müller, I. A., Fernandez, A., et al. (2017). Penultimate deglacial warming across the Mediterranean Sea revealed by clumped isotopes in foraminifera. *Nature Scientific Reports*, *7*, 16572.

Schmid, T. W., & Bernasconi, S. M. (2010). An automated method for "clumped-isotope" measurements on small carbonate samples. *Rapid Communications in Mass Spectrometry*, *24*, 1955–1963.

Stolper, D. A., Sessions, A. L., Ferreira, A. A., Santos Neto, E. V., Schimmelmann, A., Shusta, S. S., et al. (2014). Combined 13C-D and D-D clumping in methane: Methods and preliminary results. *Geochimica et Cosmochimica Acta*, *126*, 169–191.

Tisserand, A. A., Dokken, T. M., Waelbroeck, C., Gherardi, J.-M., Scao, V., Fontanier, C., & Jorissen, F. (2013). Refining benthic foraminiferal Mg/Ca-temperature calibrations using core-tops from the western tropical Atlantic: Implication for paleotemperature estimation. *Geochemistry, Geophysics, Geosystems*, *14*, 929–946. https://doi.org/10.1002/ggge.20043

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

Wang, Z., Schauble, E. A., & Eiler, J. (2004). Equilibrium thermodynamics of multiply substituted isotopologues of molecular gases. *Geochimica et Cosmochimica Acta*, *68*, 4779–4797.

World Ocean Atlas 2009, volume 1: Temperature (2010).

Yeung, L. Y., Affek, H. P., Hoag, K. J., Guo, W., & Wiegel, A. A. (2009). Large and unexpected enrichment in stratospheric $^{16}O\,^{13}C\,^{18}O$ and its meridional variation. *Proceedings of the national academy of sciences*, *106*, 11,496–11,501.