Geochemistry, Geophysics, Geosystems

AGU ADVANCING EARTH AND SPACE SCIENCE

Key Points:
- Comparison of two different techniques: Laser Ablation Inductively Coupled Plasma Mass Spectrometry and electron probe microanalyzer (EPMA) for obtaining Mg/Ca ratios in individual foraminifera shells
- The goal is to assess the use of EPMA as an alternative technique for Mg/Ca analyses of foraminiferal calcite shells
- Results indicate that it is possible to use EPMA to collect Mg/Ca data, if the ratios are lower than ~13 mmol mol^{-1}

Supporting Information:
- Table S1
- Table S2

Correspondence to:
B. Balestra,
balestra@american.edu

In Situ Mg/Ca Measurements on Foraminifera: Comparison Between Laser Ablation Inductively Coupled Plasma Mass Spectrometry and Wavelength-Dispersive X-Ray Spectroscopy by Electron Probe Microanalyzer

B. Balestra[1,2,3] , **T. Rose**[4], **J. Fehrenbacher**[5] , **K. D. Knobelspiesse**[6], **B. T. Huber**[2], **T. Gooding**[4], and **A. Paytan**[3]

[1]Environmental Science Department, American University, Washington, DC, USA, [2]Department of Paleobiology, MRC-121, National Museum of Natural History, Washington, DC, USA, [3]Institute of Marine Science, University of California Santa Cruz, Santa Cruz, CA, USA, [4]Department of Mineral Sciences, National Museum of Natural History, Washington, DC, USA, [5]Oregon State University, College of Earth, Ocean, and Atmospheric Sciences, Corvallis, OR, USA, [6]Ocean Ecology Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA

Abstract We present a comparison of two different techniques: Laser Ablation Inductively Coupled Plasma Mass Spectrometry (LA-ICP-MS) and wavelength-dispersive X-Ray Spectroscopy by electron probe microanalyzer (EPMA) for obtaining Mg/Ca ratios in individual foraminifera shells. The goal is to assess the use of EPMA as an alternative technique for Mg/Ca analyses of single foraminiferal calcite shells. Foraminifera obtained from sediments (benthic, *Uvigerina* spp.) and from plankton tows (planktonic, *Orbulina universa*) were analyzed. All specimens were prepared in epoxy mounts and exposed in cross-section such that multiple high-resolution analyses could be completed on the shells using both techniques. We examined our data using statistical methods designed for the assessment and comparison of measurement techniques. In the case of *Uvigerina*, the mean difference for ratios obtained using EPMA and LA-ICP-MS is very small (−0.046 mmol mol^{-1}) and scale independent. The *Limits of Agreement* (*LoA*, the standard deviation of the bias plus the mean bias) is [−0.315, 0.223] mmol mol^{-1}. For samples with ratios lower than 13 mmol mol^{-1}, we found a mean EPMA–LA-ICP-MS bias of −2.44 mmol mol^{-1} and a corresponding *LoA* of [−3.85, −1.04] mmol mol^{-1}. For ratios higher than 13 mmol mol^{-1}, there appears to be a scale dependent bias, meaning that the EPMA measured ratios become progressively larger than those of LA-ICP-MS as the Mg/Ca ratio increases, so the mean bias and *LoA* metrics are not meaningful. Results indicate that it is possible to use EPMA to collect Mg/Ca data, if the ratios are lower than ~13 mmol mol^{-1}.

Plain Language Summary Analysis of the chemical composition of individual foraminifers using different techniques is necessary to compare data obtained using different methods, to increase the user base of single foraminifer analyses, and to enhance understanding of past paleoceanographic and paleoclimate events. We present a comparison of two different techniques: Laser Ablation Inductively Coupled Plasma Mass Spectrometry (LA-ICP-MS) and wavelength-dispersive X-Ray Spectroscopy by electron probe microanalyzer (EPMA) for obtaining Mg/Ca ratios in individual foraminifera shells. The goal is to assess the use of EPMA as an alternative technique for Mg/Ca analyses of foraminiferal calcite shells. Data obtained using these methods for two different species of foraminifera, the benthic *Uvigerina* and the planktonic *Orbulina* were evaluated. To assess our data, we used a statistically robust analysis technique designed for paired assessment of different methodologies, the so-called Bland-Altman analysis (Altman & Bland, 1983, https://doi.org/10.2307/2987937; Bland & Altman, 1986, https://doi.org/10.1016/S0140-6736(86)90837-8; Bland & Altman, 1999, https://doi.org/10.1177/096228029900800204). The Bland-Altman statistical analysis of the data indicate that both instruments can be utilized to retrieve Mg/Ca ratios if 13 mmol mol^{-1} is set as the upper limit for these methodologies.

1. Introduction

Mg/Ca ratios of foraminiferal calcite are one of the most commonly used geochemical proxies for past temperature reconstructions (Lear et al., 2000; Mekik, 2018 and reference therein; Nürnberg et al., 1996;

Rosenthal et al., 1997; Shackleton et al., 1984; Wejnert et al., 2013). This ratio is related to water temperature and calcite saturation in the environment in which the organisms calcify (Hathorne et al., 2003; Reichart et al., 2003; Wu & Hillaire-Marcel, 1995). Although temperature is the primary control on the incorporation of Mg into shell calcite, pH and salinity also play minor roles (Dissard et al., 2010; Gray & Evans, 2019). Classical measurement techniques, such as solution-based Inductively Coupled Plasma Mass Spectrometry (ICP-MS), used to gather Mg/Ca ratios, require destruction/consumption of multiple foraminifera shells. Microanalytical methods such as Laser Ablation Inductively Coupled Plasma Mass Spectrometry (LA-ICP-MS) are useful when a limited number of specimens are available or when an assessment of chemical heterogeneity within or among individual shells is required (Fehrenbacher et al., 2015; Glock et al., 2012, and references therein). One clear advantage of microanalytical in situ techniques is that single foraminifer tests can be re-analyzed multiple times by the same or different in situ techniques (Balestra et al., 2020). LA-ICP-MS instruments however are expensive and not widely available in many universities and research institutes hence, a diverse range of analytical approaches that can be integrated to better characterize the environmental signals recorded in single foraminiferal shells or to refine understanding of biomineralization or post deposition alteration processes, are being assessed (Balestra et al., 2020). It is important to compare different measurement techniques that preserve the samples, and that provide accurate and precise Mg/Ca data. Wavelength-dispersive X-Ray Spectroscopy by electron probe microanalyzer (EPMA) is a well-established method for the determination of major, minor and trace elements and EMPA facilities are available in many geology departments and research facilities (Dai et al., 1995; Donovan et al., 2011; Fehrenbacher et al., 2020). However, few studies have demonstrated the accuracy of this method for obtaining element/Ca ratio in calcite shells (Chamberlayne et al., 2020; Fehrenbacher & Martin, 2014; Galante-Oliveira et al., 2014). Here, we assess the utility of EPMA as an alternative procedure for the analysis of Mg/Ca ratios in carbonate shells (i.e., foraminifera) and compare the data to results obtained by LA-ICP-MS. The EPMA methodology used to map element ratios can also improve our understanding of the typical (i.e., banded) distribution of Mg within the shells. We use single-shells of benthic and planktonic foraminifera commonly used for paleoclimate studies (i.e., *Uvigerina* spp. and *Orbulina universa*) and compare the results using a statistical technique designed for the assessment of different measurement methods (mean bias and *Limits of Agreement* [*LoA*]; Altman & Bland, 1983; Bland & Altman, 1986, 1999).

2. Materials and Methods

2.1. Fossil Samples, *Uvigerina* spp.

The benthic foraminifer *Uvigerina* spp. was obtained from sediments from Site ODP 1015 in the California Margin, representing the last 1,000 years (Balestra et al., 2018). The samples were soaked overnight in deionized water (pH = ~8.0, buffered), washed over 63 and 250 µm sieves, and oven dried at 45°C (after Kozdon et al., 2013). The >250 µm fraction was then examined for benthic foraminifera (i.e., *Uvigerina* spp.). Sample cleaning consisted of multiple ultrasonication steps in MilliQ water and methanol, with additional rinses in MilliQ water between and after sonication. The cleaned tests were embedded along with two grains of UWC-3 calcite standard (Kozdon et al., 2009) in a single 25-mm diameter, round epoxy mount. This mount was made with Buehler EpoxiCure 2 hardener and resin. After embedding, the epoxy mount was ground to the level of best sample exposure (Kozdon et al., 2011, 2013) and polished to a topographic relief of less than 1 µm (Kita et al., 2009).

We performed LA-ICP-MS analysis with a Teledyne Photon Machines Analyte Excite laser (193 nm) coupled to a Thermo X Series II ICP-MS (University of California, Santa Cruz). We used a 12 µm spot size, which is smaller than the typical 50 µm spot used for nonepoxy mounted LA-ICP-MS analyses. This size was necessary because the measurements were taken on the thin section of the foraminifera (i.e., wall thickness between 15 and 20 µm). Since the genus *Uvigerina* adds a new layer of calcite to all existing chambers every time that a new chamber is formed (Grunlund & Hansen, 1976), we acquired high-resolution SEM images of all samples before LA-ICP-MS analysis to facilitate targeting appropriate calcite layers. These high-resolution images helped avoid contamination due to the presence of pores or channels. The 12 µm spots were carefully targeted to be in test areas with no pores and away from the edge of the test walls to avoid contamination bias in the measurements. Preablation of the spots (5 Hz, five shots, 0.4 J/cm^3) was performed

to clean the surface area. Approximately two to five laser spots (10 Hz, 200 shots, 1.83 J/cm^2), 12 μm in diameter, were ablated on each of the 14 tests (for a total of 45 single measurements; Table S1). NIST 610 was analyzed using the same parameters and used for elemental calibration. Long-term reproducibility of the Mg/Ca ratio in NIST 610 (GeoRem, 6/2011), was 8.66 ± 0.14 mmol mol^{-1} (1σ, n = 150 over 5 analytical days). We used the Thermo PlasmaLab software, which permits screening of raw data and filters each laser profile to remove suspect contaminate interferences (Balestra et al., 2020). The mean Mg/Ca for each profile was calculated by normalization to the known trace element concentrations in the drift-corrected NIST SRM standards (Jochum et al., 2011).

After LA-ICP-MS analysis, the sample mounts were carbon-coated for analyses using a JEOL 8530F EPMA in the Department of Mineral Sciences, Smithsonian Institution. These analyses were carried out using an accelerating voltage of 15 kV with 10 nanoamp beam current and a defocused 10 μm beam. Only cations were analyzed with carbon and oxygen calculated stoichiometrically. Time-dependent intensity was monitored during the set of analyses using Probe for EPMA software. Smithsonian reference materials calcite (NMNH 136321), dolomite (NMNH R10057), and strontianite (NMNH R10065) were used as primary standards. UWC-3 calcite standard (Kozdon et al., 2009) was used as secondary standard. All the measurements taken on the reference materials were typically within the 2% of the total amount for the major elements. Between two and five 10 μm spots were analyzed by the EPMA on each of the 14 tests correspondent to the LA-ICP-MS measurements (i.e., in paired) (for a total of 45 single measurements; Table S1). EPMA measurements were placed as close as possible to the laser spots of the LA-ICP-MS, and in the same chambers. To compare the data from the two instruments, EPMA data collected as oxide weight per cent were converted to elemental weight per cent and then to molar ratios.

2.2. Plankton Tow Specimens, *Orbulina universa*

The second set of samples contained five tests of the planktonic foraminifer *Orbulina universa*. Plankton tow specimens were collected near Green Island, Taiwan in May 2019 using a 1 m plankton net (Sea-Gear, 150 μm mesh, 0–50 m depth). *Orbulina universa* was chosen because the geochemistry of the shell is consistent around the entire test, and fragments of the same specimen can be analyzed using multiple techniques (Spero et al., 2015; Vetter et al., 2014). This species is among the most comprehensively studied in terms of test calcification, physiology, and foraminiferal-algal symbiosis, (Eggins et al., 2004 and ref. therein). Specimens were picked immediately upon return to the laboratory. One gametogenic (cytoplasm barren) specimen was rinsed in DI water, dried, and then placed in a micropaleontology slide for later analyses. Five live *Orbulina universa* were placed in different 120 ml borosilicate glass jars where they remained until undergoing gametogenesis. We followed previously established culture protocols for this species (e.g., Vetter et al., 2013). Briefly, specimens were maintained at ambient temperature under 12:12 h light/dark conditions (lights set to 12,000 lumens [approximately 276 par]). Specimens were fed a day-old *Artemia nauplii* every other day until gametogenesis. After gametogenesis, specimens were rinsed in DI water and placed in micropaleontology slides for later analyses. To prepare *Orbulina universa* for trace element analysis, specimens were cracked open using a scalpel to expose inner surfaces. Specimens were then oxidatively cleaned to remove remnant organic matter by submerging shells in a 50:50 mixture of hot (80°C) 30% hydrogen peroxide and 0.1N NaOH for 10 min. The oxidative solution was then aspirated, and the shell fragments were rinsed 3x in ultrapure water, which was aspirated between each rinse. Fragments were then air dried.

One fragment from each specimen was transferred to double sided carbon tape with the inner surface facing up to prepare them for LA spot analyses. The remaining fragments were embedded in epoxy resin for EMP analyses as described in the previous section. To compare the data from the two instruments, EPMA data collected as oxide weight percent was converted to elemental weight per cent and then to molar ratios. *Orbulina universa* laser analyses were conducted at the Keck Collaboratory in the College of Earth, Ocean, and Atmospheric Sciences at Oregon State University. Shells were analyzed using thin section spot analysis. For the analyses, acquisitions were performed using a 20 μm round spot size, a 6 Hz repetition rate, and a laser fluence of 1.02–1.26 J/cm^2. Laser He gas flow was set to 0.8 lpm set to achieve high count rates while maintaining ThO$^+$/Th$^+$ oxides below 0.02%. Two to five repeat analyses were obtained for each specimen. The thin section spot analyses were placed near the EPMA analyses, when possible. Isotopes (^{24}Mg, ^{25}Mg, ^{27}Al,

^{43}Ca, ^{44}Ca, ^{55}Mn, ^{88}Sr, and ^{138}Ba) were measured using a rapid peak hopping procedure. Trace element to calcium ratios (TE/Ca) were calculated offline in LA-Tools (Branson et al., 2016), a Python based laser data reduction program. LA-Tools follows established data reduction protocols (Longerich et al., 1996) including screening for outliers, drift correction, and subtracting average background counts from each data point. The mean TE/Ca for each profile is calculated by normalization to the known trace element concentrations in the drift-corrected NIST SRM (610 and 612) glass standards (Jochum et al., 2011). ^{43}Ca was used as an internal standard. Individual spot TE/Ca averages were obtained by integrating TE/Ca ratios throughout the depth profile, after excluding high TE/Ca ratios often encountered at the beginning of the ablation (e.g., Figure S2B in Fehrenbacher et al., 2018). Between five and fifteen 10 µm spots were analyzed for the EPMA on each of the 5 tests correspondent to the LA-ICP-MS measurements (i.e., in paired for a total of 38 single measurements; Table S2).

2.3. Statistical Analysis

The paired LA-ICP-MS and EPMA measurements that were performed on the fossil and tow samples were assessed to determine the degree to which those methods agree, and if they can be used interchangeably. To assess our data, we used a statistically robust analysis technique designed for paired assessment of different methodologies (Altman & Bland, 1983; Bland & Altman, 1986, 1999). At the core of these techniques is the so-called Bland-Altman (also known as a Tukey mean-difference) plot. Measurements are treated as paired LA-ICP-MS/EPMA observations, where the mean value of each pair is represented on the abscissa, and the bias on the ordinate axis.

This is a simple, graphical, illustration of two parameters of relevance to an analysis of measurement techniques: the mean bias, and the *LoA* (a measure of scatter in the bias). Unlike commonly used regression correlation parameters, the mean bias and the *LoA* are expressed in the units of the measurement. Furthermore, they are also less subject to the measurement numerical distribution and the existence of outliers, a common problem with regression techniques (Seegers et al., 2018). We followed these steps to perform the analysis:

1. Calculate the paired mean ($m_i = (r_{i,\text{ EPMA}} + r_{i,\text{ LA-ICP-MS}})/2$) for all measurements (indicated by the index i). In our case, there were 45 overall paired measurements for fossil *Uvigerina* and 38 for planktonic tow *Orbulina universa*
2. Calculate the paired bias ($b_i = r_{i,\text{ EPMA}} - r_{i,\text{ LA-ICP-MS}}$)
3. For visualization purposes, create a *m* versus *b* plot, as we have in Figures 1 and 2
4. Assess the independence of *b* from *m*. This is a requirement of this technique for the mean bias and *LoA* to be meaningful. We used a two tailed test that the absolute value of the Pearson correlation coefficient was below a critical value. For a 90% confidence, this critical value was 0.126 for both cases
5. If the above test is passed, calculate the mean bias ($\bar{b} = \frac{1}{n}\sum_{i=0}^{n}b_i$, where n is the number of measurements)
6. Calculate the *LoA*, which is the standard deviation of *b* plus or minus the bias. $LoA = \bar{b} \pm \sqrt{\frac{1}{n}\sum_{i=0}^{n}\left(b_i - \bar{b}\right)}$.

   If the numerical distribution of the mean bias is Gaussian distributed (which is more likely for differences than original data values), then we should expect that 68% of all paired measurements should fall in this range

We should note that the mean bias and *LoA* indicate the paired differences between measurements and cannot indicate the source of bias or scatter (if they exist). Comparative analysis of measurement techniques, either by regression analysis or Bland-Altman, demonstrate the agreement between measurement methods, but cannot indicate individual measurement method uncertainty. To be clear on terminology, "error" is the difference between the true and measured value in an individual measurement, while "uncertainty" is a statistical description of that error (Povey & Grainger, 2015). What these techniques do demonstrate is how well different methods agree. Thus, if a user has confidence in one measurement technique, and has expectations of measurement uncertainty for that technique, comparative analysis reveals how similar another technique will be. The advantage of Bland-Altman, for these purposes, is that it expresses agreement in terms of different types of uncertainty (and in the units of the measurement). Mean bias ($\bar{b}$) indicates the
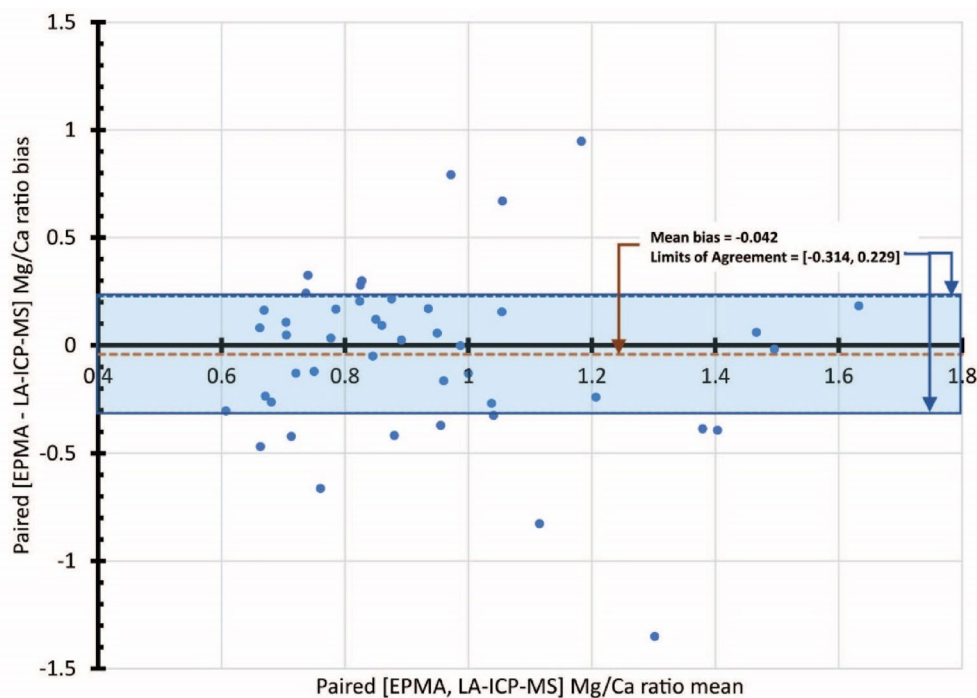
Figure 1. Bland-Altman plot for EPMA–LA-ICP-MS Mg/Ca ratios for *Uvigerina* spp. demonstrating the paired bias between these determinations of Mg/Ca plotted against the mean. Because the bias indicates no relationship with that mean, we can calculate its mean value (−0.046) (red dashed line) and the *LoA* (*Limits of Agreement*), [−0.315,0.225] (area within the blue dashed lines). Under the assumption that the bias is normally distributed, we can expect 68% of paired EMPA versus LA-ICP-MS comparisons to fall within this range. EPMA, wavelenght-dispersive X-Ray Spectroscopy by electron probe microanalyzer; LA-ICP-MS, Laser Ablation Inductively Coupled Plasma Mass Spectrometry.

amount of systematic differences between measurement techniques, which could be related to errors such as calibration. The *LoA* are the amount of scatter in the relationship, whose source could be random errors such as electronic shot noise. Scale dependence, by which different amounts of bias or scatter are present with varying measurement value, interferes with the meaning of the bias or *LoA*. For this reason, we verify that there is no relationship between the paired mean and paired bias before performing the Bland-Altman analysis (step 4, above). We do not, and cannot, know the measurement error, but this technique relates measurement uncertainty of EPMA to that of LA-ICP-MS. It assesses the use of EPMA as an alternative to LA-ICP-MS by telling us the amount of similarity of the measurements.

Additionally, linear regression and correlation coefficients considering the average of the measurements taken for each test were generated using PAST software (Hammer et al., 2001) (Figures 3 and 4). This was done to place all results in context of the more commonly used (but inappropriate for these purposes) regression analysis techniques.

3. Results

3.1. *Uvigerina* spp.

Uvigerina spp. is a benthic genus commonly used for paleoclimate studies. Several works have been conducted on this genus because of the limited influence of $\Delta[CO_3^{2-}]$ on shell Mg/Ca of it (Elderfield et al., 2010). When analyzing these samples, particular care has been taken to place the EPMA beam as close as possible to the LA-ICP-MS analysis location to avoid capturing differences within a test. Values for Mg/Ca of individual specimens for the EPMA spot analyses ranged from 0.4 to 1.7 mmol mol^{-1}. For the LA-ICP-MS the
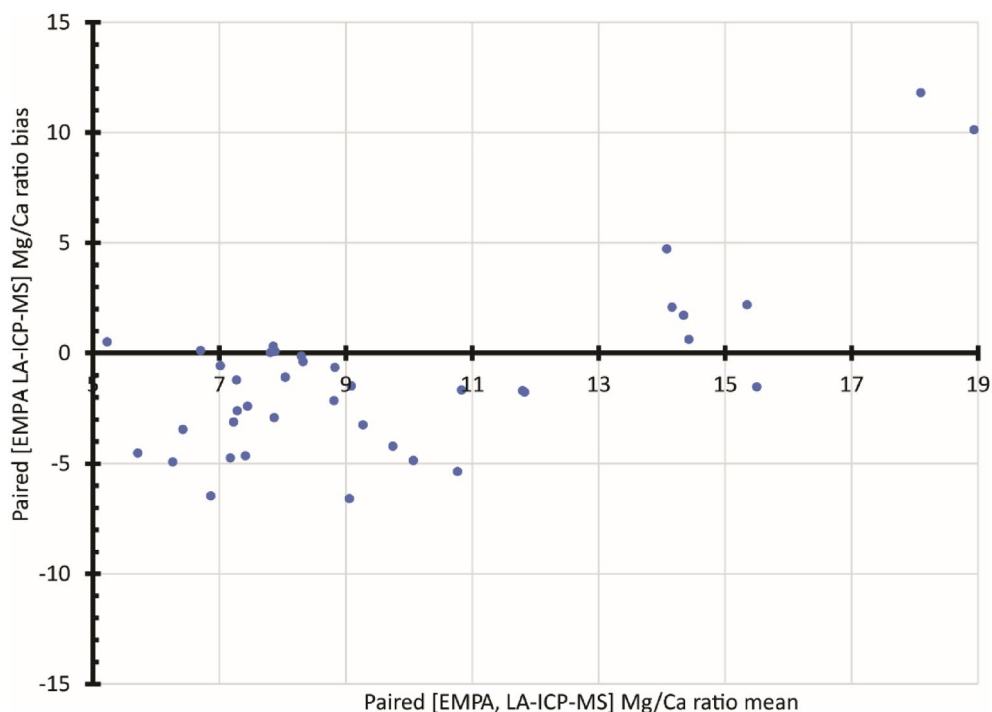
Figure 2. Bland-Altman plot for EPMA–LA-ICP-MS Mg/Ca ratios for *Orbulina universa* demonstrating the paired bias between these determinations of Mg/Ca plotted against the mean. Because there is a relationship between paired bias and mean, a mean bias value and the *LoA* (*Limits of Agreement)* were not calculated. Qualitatively, it appears the bias changes for ratio values that exceed 13 mmol mol^{-1}, see text for more details. EPMA, wavelenght-dispersive X-Ray Spectroscopy by electron probe microanalyzer; LA-ICP-MS, Laser Ablation Inductively Coupled Plasma Mass Spectrometry.

values range from 0.5 to 1.9 mmol mol^{-1} (Table S1). For the paired EPMA–LA-ICP-MS Bland-Altman analysis, we found a mean bias of -0.046 mmol mol^{-1} and *LoA* of $[-0.315, 0.223]$ mmol mol^{-1}.

3.2. *Orbulina universa*

The range in Mg/Ca ratios within a shell was much larger for these *Orbulina universa* measurements. Values for Mg/Ca of individual specimens for the EPMA spot analyses ranged from 2.5 to 24.5 mmol mol^{-1}. For the LA-ICP-MS the values range from 4.9 to 16.2 mmol mol^{-1} (Table S2). The entire data set did not meet the criteria for the Bland-Altman analysis because of a scale dependent bias, where EPMA becomes progressively larger than LA-ICP-MS when the Mg/Ca ratio exceeds $\sim$13 mmol mol^{-1}. The subset of data below that value do not show a scale dependent bias, and for these data the mean bias is -2.44 mmol mol^{-1} and *LoA* = $[-3.85, -1.04]$ mmol mol^{-1}.

4. Discussion

The Bland-Altman statistical approach that we used is designed to assess the similarity of two different instrumental techniques, which we apply for measurements of Mg/Ca ratios of calcite shells. We tested this for two different foraminifera, one a benthic obtained from a core-top sediment sample, the other a planktonic specimen gathered from a tow. In the medical literature, Bland and Altman (Altman & Bland, 1983; Bland & Altman, 1986, 1999, and many others) investigated the most statistically appropriate means to assess if different measurement techniques are interchangeable. They noted that no single parameter can be used for this task, and instead devised a method based on the bias between paired measurements, where the bias (difference) between each pair is plotted against the mean (Figures 1, 2, and 5). If the bias has no dependence upon the mean (i.e., differences are independent of measurement value), then two parameters can be
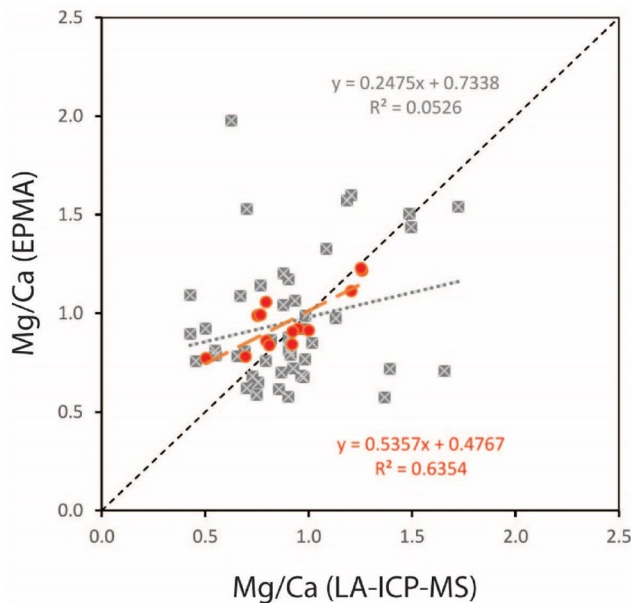
Figure 3. Bivariate least square regression of the in situ Mg/Ca paired analyses using LA-ICP-MS (12 μm spot) versus EPMA (10 μm spot) for fossil *Uvigerina* spp. The gray colors represent the total number of measurements while the red represent the averaged values. EPMA, wavelenght-dispersive X-Ray Spectroscopy by electron probe microanalyzer; LA-ICP-MS, Laser Ablation Inductively Coupled Plasma Mass Spectrometry.

Figure 4. Bivariate least square regression of the in situ Mg/Ca paired analyses using LA-ICPMS (20 μm spot) versus EPMA (10 μm spot) for cultured *Orbulina universa*. The gray colors represent the total number of measurements while the red represent the averaged values. EPMA, wavelenght-dispersive X-Ray Spectroscopy by electron probe microanalyzer; LA-ICP-MS, Laser Ablation Inductively Coupled Plasma Mass Spectrometry.
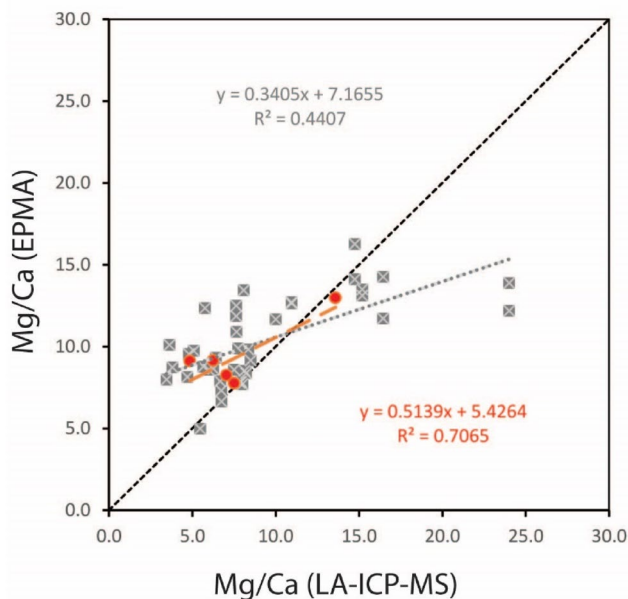
derived. The first is the mean bias, which describes, on average, the difference between the measurements—zero is obviously the best, and nonzero values can perhaps be used to inform calibration adjustments. The second parameter is the *LoA*, which simply describes the scatter about the mean bias, and is calculated as the standard deviation of the biases. This parameter represents the random variation in the bias. If feasible (i.e., passing a two tailed hypothesis test that there is no relationship between mean and bias), this method is preferable to the more common method of calculating linear regression and correlation coefficients between the measurements. The latter is an assessment of the association between paired measurements, but not their agreement. Furthermore, metrics such as the Pearson correlation coefficient assume that the data are numerically Gaussian distributed without outliers, which is not always the case (Seegers et al., 2018). Regression can also mask relevant differences between measurements (Knobelspiesse et al., 2019). Finally, the mean bias and *LoA* from the Bland and Altman method are expressed in the units of the measurements (in this case the Mg/Ca ratio) and so can be compared to overall expectations of measurement uncertainty and scientific need. Below we discuss each species separately and provide insight into the method comparison in general with suggestions for future work.

4.1. *Uvigerina* spp.

Bland-Altman statistical analysis was successfully performed on these data, as the data met the criteria in step 4, above, namely that there is no relationship between paired mean and bias (scale dependent bias). We found that on average there is agreement, with a mean bias of only -0.046 mmol mol^{-1}, indicating low systematic error. Random errors, expressed as scatter in the comparisons and described by the *LoA* are $LoA = [-0.315, 0.223]$ mmol mol^{-1} (Figure 1). This means that for a given pair of EPMA and LA-ICP-MS measurements, we would expect agreement within this range approximately 68% of the time (representing to one standard deviation within a Gaussian normal distribution). Both metrics are expressed in the measurement units (and scale) of the instruments, and as such are more directly applicable than a unitless correlation coefficient. To summarize, EPMA and LA-ICP-MS agree well on average, but they do have random variability. This suggests that repeated measurements on the same foraminifer spot by one or both techniques (if possible) are likely to increase the agreement between procedures. If the source of differences is primarily random errors, and repeated measurements can be made independently, then we would expect the *LoA* to decrease in a manner that follows the Law of large numbers ($\sqrt{n^{-1}}$), where n indicates the number of independent measurements. Lastly, the number of repeated measurements is also link to the thickness of the thin section of the foraminifera analyzed.

We have also averaged all the measurements for each single foraminifer with both methods (Table S1). The classic linear regression calculated considering the single data point measurements show a poor correlation with R^2 of 0.05 whereas instead the linear regression considering the average of the measurement for each single test shows a R^2 of 0.7 (Figure 3). This difference, showing better correlation in the averaged data,
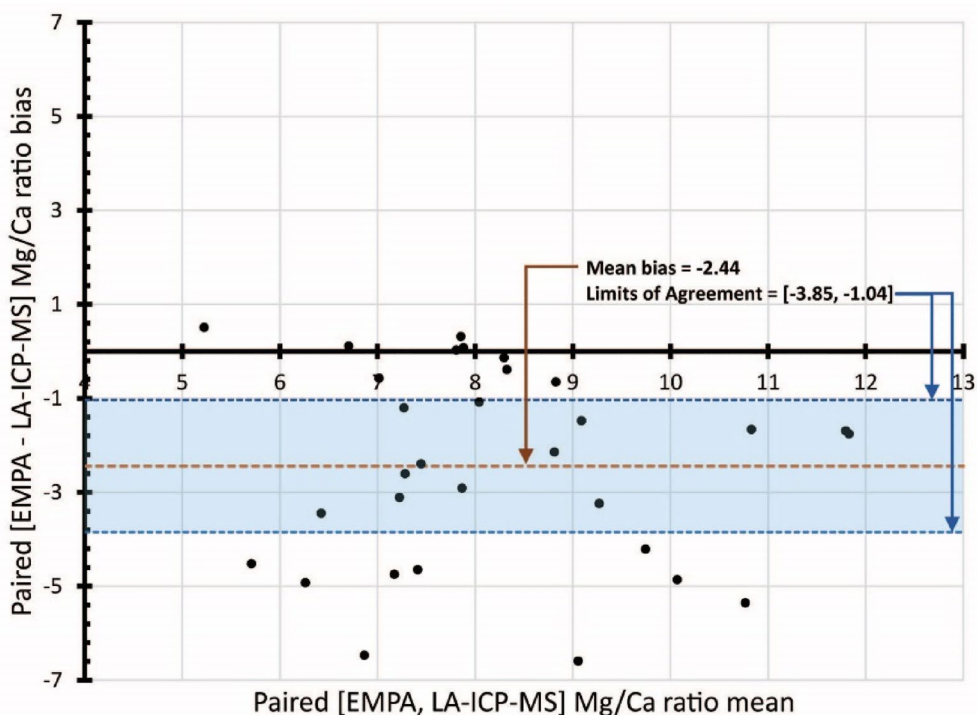
Figure 5. Bland-Altman plot for EPMA–LA-ICP-MS Mg/Ca ratios for *Orbulina universa* demonstrating the paired bias between these determinations of Mg/Ca plotted against the mean. Mg/Ca values that exceed 13 mmol mol^{-1} have been excluded. The mean bias is -2.44 (red dashed line) and $LoA = [-3.85, -1.04]$ (blue area within the dashed lines). EPMA, wavelenght-dispersive X-Ray Spectroscopy by electron probe microanalyzer; LA-ICP-MS, Laser Ablation Inductively Coupled Plasma Mass Spectrometry; LoA, Limits of Agreement.

expresses internal Mg/Ca ratio variability in each foraminifer, likely due to the different growth patterns each specimen as is common in *Uvigerina* (Grunlund & Hansen, 1976).

4.2. *Orbulina universa*

We found that comparison between EPMA and LA-ICP-MS for the measurements on the *Orbulina universa* data set do not abide by the scale independence criteria required for Bland-Altman analysis. Figure 2 shows that for Mg/Ca ratio values less than 13, EPMA ratios are generally smaller than ratios obtained by LA-ICP-MS, and for values above 13 the bias progressively increases. If we exclude Mg/Ca ratios above 13, the Orbulina universa do meet the Bland-Altman independence criteria, and the mean bias is -2.44 mmol mol^{-1} and $LoA = [-3.85, -1.04]$ mmol mol^{-1} (Figure 5). This means that we can expect the measurement techniques to agree with these characteristics for Mg/Ca values less than roughly 13. Interestingly, the LoA in this case is smaller (compared to the total Mg/Ca range) than it was for *Uvigerina* spp., indicating less random scatter. Since we do not expect scale dependent bias from both instruments, we considered that the discrepancies for ratios above 13 could be due to differences in instrument sensitivities, beam saturation or more complex factors that affect the Mg/Ca ratio generated by the two techniques. Furthermore, the values above 13 belong to just one test (test #5 in Table S2). We thus cannot exclude that this high ratio may derive from the conditions of this individual foraminifera. Intrashell Mg/Ca variability is quite high in this species and ratios that exceed 13 mmol mol^{-1} are not uncommon (e.g., Spero et al., 2015). With no definitive source for the high Mg/Ca, our determination of the 13 mmol mol^{-1} threshold is a qualitatively chosen value based upon visual analysis in Figure 2. The ultimate result of this analysis indicates an agreement between EPMA, and LA-ICP-MS as defined by the mean bias and LoA for values less than 13 mmol mol^{-1}. Further analysis is required to confirm if the threshold of 13 mmol mol^{-1} is valid. As for the *Uvigerina* spp., we applied a linear regression calculated for this subset of samples considering the single data point measurements and

the averaged data within the same tests (Figure 4). Both correlations in this set of data are good, $R^2 = 0.4$ for the single paired, and $R^2 = 0.7$ for the averaged measurements.

5. Summary

Analysis of the chemical composition of individual foraminifers using different techniques is necessary for ensuring consistency in data and enabling comparison of data obtained using different methods, to better understand paleoceanographic and paleoclimate events. In this study we compare Mg/Ca ratio in calcite shells of foraminifera using two different instruments, EPMA and LA-ICP-MS. Both instruments utilize microanalytical techniques to determine TE/Ca ratios and they are particularly useful when a limited number of specimens are available or when information regarding intra species and comparison between individual species is needed (Balestra et al., 2020; Glock et al., 2012). Data obtained using these methods for two different species of foraminifera, the benthic *Uvigerina* and the planktonic *Orbulina* were evaluated. The Bland-Altman statistical analysis of the data indicate that both instruments can be utilized to retrieve Mg/Ca ratios if 13 mmol mol^{-1} is set as the upper limit for these methodologies. The majority of those measurements came from a single foraminifer, so it is unclear if this is evidence for systematic measurement obstacles above 13 mmol mol^{-1} in either or both instruments, or a uniquely divergent sample. Data sets show less agreement when evaluated using mean bias and *LoA*. Further investigation using more tests is needed to unravel the reason for reduced agreement between methods for samples with higher than 13 mmol mol^{-1} ratios.

Data Availability Statement

The data used in this study (Tables S1 and S2) have being archived in the Pangaea data repository https://doi.org/10.1594/PANGAEA.926484 and https://doi.org/10.1594/PANGAEA.926486.

References

Altman, D. G., & Bland, J. M. (1983). Measurement in medicine: The analysis of method comparison studies. *The Statistician*, *32*, 307–317. https://doi.org/10.2307/2987937

Balestra, B., Krupinski, N. B. Q., Erohina, T., Fessenden-Rahn, J., Rahn, T., & Paytan, A. (2018). Bottom-water oxygenation and environmental change in Santa Monica Basin, Southern California during the last 23 kyr. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *490*, 17–37.

Balestra, B., Orland, I. J., Fessenden-Rahn, J., Gorski, G., Franks, R., Rahn, T., & Paytan, A. (2020). Paired analyses of oxygen isotope and elemental ratios within individual shells of benthic foraminifera genus *Uvigerina*. *Chemical Geology*, *533*, 119377.

Bland, J. M., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, *327*(8476), 307–310. https://doi.org/10.1016/S0140-6736(86)90837-8

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, *8*(2), 135–160. https://doi.org/10.1177/096228029900800204

Branson, O., Bonnin, E. A., Perea, D. E., Spero, H. J., Zhu, Z., Winters, M., et al. (2016). Nanometer-scale chemistry of a calcite biomineralization template: implications for skeletal composition and nucleation. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, 12934–12939.

Chamberlayne, B. K., Tyler, J. J., & Gillanders, B. M. (2020). Environmental controls on the geochemistry of a short-lived bivalve in southeastern Australian estuaries. *Estuaries and Coasts*, *43*(1), 86–101.

Dai, Z., Ren, C., Zhang, J., Ma, G., & Yang, F. (1995). Comparison of quantitative PIXE and EPMA microanalysis of mineral samples. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, *104*(1–4), 489–493.

Dissard, D., Nehrke, G., Reichart, G. J., & Bijma, J. (2010). Impact of seawater pCO2 on calcification and Mg/Ca and Sr/Ca ratios in benthic foraminifera calcite: Results from culturing experiments with Ammonia tepida. *Biogeosciences*, *7*, 81–93.

Donovan, J. J., Lowers, H. A., & Rusk, B. G. (2011). Improved electron probe microanalysis of trace elements in quartz. *American Mineralogist*, *96*(2–3), 274–282.

Eggins, S. M., Sadekov, A., & De Deckker, P. (2004). Modulation and daily banding of Mg/Ca in Orbulina universa tests by symbiont photosynthesis and respiration: A complication for seawater thermometry? *Earth and Planetary Science Letters*, *225*(3–4), 411–419.

Elderfield, H., Greaves, M., Barker, S., Hall, I. R., Tripati, A. K., Ferretti, P., et al. (2010). Mg/Ca and stable oxygen isotope ratios of Uvigerina spp. from the Southern Ocean. *Quaternary Science Reviews*, *29*(1–2), 160–169. https://doi.org/10.1016/j.quascirev.2009.07.013

Fehrenbacher, J. S., Marchitto, T., & Spero, H. J. (2020). Comparison of laser ablation and solution based ICP-MS results for individual foraminifer Mg/Ca and Sr/Ca analyses. *Geochemistry, Geophysics, Geosystems*, *21*(12), e2020GC009254. https://doi.org/10.1029/2020GC009254

Fehrenbacher, J. S., & Martin, P. (2014). Exploring the dissolution effect on the intrashell Mg/Ca variability of the planktic foraminifer *Globigerinoides ruber*. *Paleoceanography*, *29*, 854–868. https://doi.org/10.1002/2013PA002571

Fehrenbacher, J. S., Russell, A. D., Davis, C. V., Spero, H. J., Chu, E., & Hönisch, B. (2018). Ba/Ca ratios in the non-spinose planktic foraminifer Neogloboquadrina dutertrei: Evidence for an organic aggregate microhabitat. *Geochimica et Cosmochimica Acta*, *236*, 361–372.

Fehrenbacher, J. S., Spero, H. J., Russell, A. D., Vetter, L., & Eggins, S. (2015). Optimizing LA-ICP-MS analytical procedures for elemental depth profiling of foraminifera shells. *Chemical Geology*, *407–408*, 2–9. https://doi.org/10.1016/j.chemgeo.2015.04.007

Galante-Oliveira, S., Marçal, R., Guimarães, F., Soares, J., Lopes, J. C., Machado, J., & Barroso, C. (2014). Crystallinity and microchemistry of Nassarius reticulatus (Caenogastropoda) statoliths: towards their structure stability and homogeneity. *Journal of Structural Biology*, *186*(2), 292–301.

Glock, N., Eisenhauer, A., Liebetrau, V., Wiedenbeck, M., Hensen, C., & Nehrke, G. (2012). EMP and SIMS studies on Mn/Ca and Fe/Ca systematics in benthic foraminifera from the Peruvian OMZ: A contribution to the identification of potential redox proxies and the impact of cleaning protocols. *Biogeosciences*, *9*, 341–359. http://dx.doi.org/10.5194/bg-9-341-2012

Gray, W. R., & Evans, D. (2019). Nonthermal Influences on Mg/Ca in Planktonic Foraminifera: A Review of Culture Studies and Application to the Last Glacial Maximum. *Paleoceanography and Paleoclimatology*, *34*, 306–315. https://doi.org/10.1029/2018PA003517

Grunlund, H., & Hansen, H. J. (1976). Scanning electron microscopy of some recent and fossil nodosariid Foraminifera. *Bullettin of the Geological Society of Denmark*, *25*, 49–62.

Hammer, Ø., Harper, D. A., & Ryan, P. D. (2001). PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica*, *4*(1), 9.

Hathorne, E. C., Alard, O., James, R. H., & Rogers, N. W. (2003). Determination of intratest variability of trace elements in foraminifera by laser ablation inductively coupled plasmamass spectrometry. *Geochemistry, Geophysics, Geosystems*, *4*(12), 8408. http://dx.doi.org/10.1029/2003GC000539

Jochum, K. P., Weis, U., Stoll, B., Kuzmin, D., Yang, Q., Raczek, I., et al. (2011). Determination of reference values for NIST SRM 610–617 glasses following ISO guidelines. *Geostandards and Geoanalytical Research*, *35*(4), 397–429. https://doi.org/10.1111/j.1751-908X.2011.00120.x

Kita, N. T., Ushikubo, T., Fu, B., & Valley, J. W. (2009). High precision SIMS oxygen isotope analysis and the effect of sample topography. *Chemical Geology*, *264*(1–4), 43–57. https://doi.org/10.1016/j.chemgeo.2009.02.012

Knobelspiesse, K., Tan, Q., Bruegge, C., Cairns, B., Chowdhary, J., van Diedenhoven, B., et al. (2019). Intercomparison of airborne multi-angle polarimeter observations from the Polarimeter Definition Experiment. *Applied Optics*, *58*(3), 650–669.

Kozdon, R., Kelly, D. C., Kita, N. T., Fournelle, J. H., & Valley, J. W. (2011). Planktonic foraminiferal oxygen isotope analysis by ion microprobe technique suggests warm tropical sea surface temperatures during the Early Paleogene. *Paleoceanography*, *26*, PA3206. https://doi.org/10.1029/2010PA002056

Kozdon, R., Kelly, D. C., Kitajima, K., Strickland, A., Fournelle, J. H., & Valley, J. W. (2013). In situ delta O-18 and Mg/Ca analyses of diagenetic and planktic foraminiferal calcite preserved in a deep-sea record of the Paleocene-Eocene thermal maximum. *Paleoceanography*, *28*, 517–528. http://dx.doi.org/10.1002/palo.20048

Kozdon, R., Ushikubo, T., Kita, N. T., Spicuzza, M., & Valley, J. W. (2009). Intratest oxygen isotope variability in the planktonic foraminifer N. pachyderma: Real vs. apparent vital effects by ion microprobe. *Chemical Geology*, *258*, 327–337. https://doi.org/10.1016/j.chemgeo.2008.10.032

Lear, C. H., Elderfield, H., & Wilson, P. A. (2000). Cenozoic deep-sea temperatures and global ice volumes from Mg/Ca in benthic foraminiferal calcite. *Science*, *287*, 269–272.

Longerich, H. P., Günther, D., & Jackson, S. E. (1996). Elemental fractionation in laser ablation inductively coupled plasma-mass spectrometry. *Fresenius' Journal of Analytical Chemistry*, *355*, 538–542.

Mekik, F. (2018). Do proxies agree? δ18O, δ13C and Mg/Ca from tests of Neogloboquadrina dutertrei in the eastern equatorial Pacific. *Geochimica et Cosmochimica Acta*, *236*, 260–282. https://doi.org/10.1016/j.gca.2018.03.005

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

Povey, A. C., & Grainger, R. G. (2015). Known and unknown unknowns: Uncertainty estimation in satellite remote sensing. *Atmospheric Measurement Techniques*, *8*(11), 4699–4718. https://doi.org/10.5194/amt-8-4699-2015

Reichart, G. J., Jorissen, F., Anschutz, P., & Mason, P. R. D. (2003). Single foraminiferal test chemistry records the marine environment. *Geology*, *31*, 355–358.

Rosenthal, Y., Boyle, E. A., & Slowey, N. (1997). Temperature control on the incorporation of magnesium, strontium, fluorine, and cadmium into benthic foraminiferal shells from Little Bahama Bank: Prospects for thermocline paleoceanography. *Geochimica et Cosmochimica Acta*, *61*, 3633–3643.

Seegers, B. N., Stumpf, R. P., Schaeffer, B. A., Loftin, K. A., & Werdell, P. J. (2018). Performance metrics for the assessment of satellite data products: An ocean color case study. *Optics Express*, *26*(6), 7404–7422.

Shackleton, N. J., Backman, J., Zimmerman, H., Kent, D. V., Hall, M. A., Roberts, D. G., et al. (1984). Oxygen-isotope calibration of the onset of ice-rafting and history of glaciation in the North Atlantic region. *Nature*, *3*, 620–623.

Spero, H. J., Eggins, S. M., Russell, A. D., Vetter, L., Kilburn, M. R., & Hönisch, B. (2015). Timing and mechanism for intratest Mg/Ca variability in a living planktic foraminifer. *Earth and Planetary Science Letters*, *409*, 32–42.

Vetter, L., Kozdon, R., Valley, J. W., Mora, C. I., & Spero, H. J. (2014). SIMS measurements of intrashell δ13C in the cultured planktic foraminifer Orbulina universa. *Geochimica et Cosmochimica Acta*, *139*, 527–539.

Vetter, L., Spero, H. J., Russell, A. D., & Fehrenbacher, J. S. (2013). LA-ICP-MS depth profiling perspective on cleaning protocols for elemental analyses in planktic foraminifers. *Geochemistry, Geophysics, Geosystems*, *14*, 2916–2931. https://doi.org/10.1002/ggge.20163

Wejnert, K. E., Thunell, R. C., & Astor, Y. (2013). Comparison of species-specific oxygen isotope paleotemperature equations: Sensitivity analysis using planktonic foraminifera from the Cariaco Basin, Venezuela. *Marine Micropaleontology*, *101*, 76–88.

Wu, G., & Hillaire-Marcel, C. (1995). Application of LP-ICP-MS to benthic foraminifers. *Geochimica et Cosmochimica Acta*, *59*(2), 409–414. https://doi.org/10.1016/0016-7037(94)00370-2