A new calcium carbonate nano-particulate pressed powder pellet NFHS-2-NP for LA-ICP-OES, LA-(MC)-ICP-MS and µXRF

Wim Boer (1)*, Simon Nordstad (4)*, Michael Weber (3), Regina Mertz-Kraus (3), Bärbel Hönisch (5), Jelle Bijma (6), Markus Raitzsch (6), Dorothee Wilhelms-Dick (6), Gavin L. Foster (7), Heather Goring-Harford (7), Dirk Nürnberg (8), Folkmar Hauff (8), Henning Kuhnert (9), Federico. Lugli (10), Howie Spero (11), Martin Rosner (12), Piet van Gaever (1), Lennart J. de Nooijer (1), Gert-Jan Reichart (1,2).

(1) Department of Ocean Systems, NIOZ Royal Netherlands Institute for Sea Research, Texel, 1790 AB 59, Netherlands

(2) Department Earth Science, Utrecht University, Utrecht, 3584 CS 8, Netherlands.

(3) Institute for Geosciences, Johannes Gutenberg-Universität Mainz, 55128 Mainz, Germany

(4) myStandards GmbH, Kiel, 24118, Germany

(5) Lamont-Doherty Earth Observatory and Department of Earth and Environmental Sciences of Columbia University, Palisades, NY 10964, United States

(6) Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung (AWI), Bremerhaven, Germany

(7) School of Ocean and Earth Science, University of Southampton, National Oceanography Centre Southampton, Southampton, United Kingdom

(8) GEOMAR Helmholtz Centre for Ocean Research Kiel, 24148 Kiel, Germany

(9) MARUM, University of Bremen, 28334 Bremen, Germany

(10) Department of Chemical and Geological Sciences, University of Modena and Reggio Emilia, Via G. Campi, 103 - 41125 Modena, Italy

(11) Stable Isotope Laboratory, UC Davis, United States.

(12) IsoAnalysis UG, Berlin, Germany

Corresponding authors. E-mails: wim.boer@nioz.nl; nordstad@my-standards.com

Abstract

A new matrix-matched reference material has been developed – NFHS-2-NP (NIOZ Foraminifera House Standard-2-Nano-Pellet) – with element mass fractions, and isotope ratios resembling that of natural foraminiferal calcium carbonate. A 180–355 µm size fraction of planktic foraminifera was milled to nano-particles and pressed to pellets. We report reference and information values for mass fractions of 46 elements measured by six laboratories as well as for $^{87}Sr/^{86}Sr$ (three laboratories), $\delta^{13}C$, $\delta^{18}O$ (five laboratories), and $^{206,207,208}Pb/^{204}Pb$ isotope ratios (one laboratory) determined by ICP-MS, ICP-OES, MC-ICP-MS, IRMS, WD-XRF and TIMS. Inter- and intra-pellet elemental homogeneity was tested using multiple LA-ICP-MS analyses in two laboratories applying spot sizes of 60 and 70 µm. The LA-ICP-MS results for most of the elements relevant as proxies for paleoclimate research show RSD values < 3%, demonstrating a satisfactory homogeneous composition. Homogeneity of $^{87}Sr/^{86}Sr$ ratios of the pellet of two laboratories was verified by repeated LA-MC-ICP-MS. Information values are reported for Pb isotope ratios and $\delta^{13}C$, $\delta^{18}O$ values. The homogeneity for these isotope systems remains to be tested by LA-MC-ICP-MS and SIMS. Overall, our results confirm the suitability of NFHS-2-NP for calibration or monitoring the quality of in-situ geochemical techniques.

Keywords: laser ablation-multi collector inductively coupled plasma-mass spectrometry, reference materials, carbonate, homogeneity, in-situ techniques.

Introduction

Robust application of natural carbonates as paleoclimate and environmental archives critically depends on accurately measured geochemical proxy signals. Elemental mass fractions, element to calcium ratios, isotope ratios, or δ-values of these carbonates are typically used in these contexts (Nürnberg et al. 1996; Zachos et al. 2001; Hathorne et al. 2013). Some approaches rely on proxy data collected at micrometer-scale spatial resolution with high accuracy, for example when using bivalves (Bougeois et al. 2014), otoliths (Phillis et al. 2011; Willmes et al.2016), single foraminiferal shells and/or chambers (Titelboim et al. 2018; Pracht et al. 2019), corals (Tarique et al. 2021, Hathorne et al. 2011), or speleothems (Burns et al. 1998, Weber et al. 2021) to reconstruct inter- or intra-annual variability of past climate. Several in-situ state-of-the-art techniques offer excellent analytical precision and provide the desired high spatial resolution for multiple elements and/or isotope ratio analyses. Fossil biogenic carbonates have also often undergone chemical alteration during diagenesis and may carry overgrowths at their surface -whether on the outside or on the inside- that can bias the element/Ca ratio of the pristine signal (Hoogakker et al. 2009). Cleaning protocols have been developed for bulk analysis to remove post-depositional contaminants (e.g., Barker et al. 2003) but testing the effect of these cleaning methods may also benefit from accurate and precise determination of element/Ca in skeletons and shells at a high spatial resolution. Several previous studies have highlighted the need for appropriate reference materials (RMs) for these applications to ensure accuracy (e.g., Greaves et al. 2005 and 2008, Hathorne et al. 2013).

Solution-based techniques generally deliver accurate and precise data via, for example, isotope dilution and standard addition methods using RMs with known chemical composition to convert measured intensities into accurate elemental mass fractions and isotope ratios. If the RM is of a different material

than the samples, matrix effects arising from differential ionization behavior and interferences may bias the results (Arslan and Paulson, 2002; Marchitto, 2006). This can be overcome by dilution of both the digested samples and RMs with, for example, nitric acid and hence increasing the similarity in the composition of their matrices (e.g. de Villiers et al. 2002), or by removal of the matrix altogether with chromatography such as the commercially available seaFAST system (e.g., Hathorne et al. 2012; Arslan and Paulson, 2002)). For in-situ techniques, such as laser ablation (multi-collector) inductively coupled plasma mass spectrometry (LA-(MC)-ICP-MS), laser induced breakdown spectroscopy (LIBS) or micro X-ray fluorescence (μXRF), dilution/separation is not possible and matrix effects can reduce the accuracy of the analysis. Because commonly used RMs (e.g., NIST SRM 610 and 612) are SiO_2 glasses with only about 8 % m/m Ca compared to approximately 40 % m/m Ca in carbonates, they can introduce matrix effects occurring in the laser ablation system and ICP-MS (e.g., Sylvester, 2008, Jochum et al. 2019) when used to calibrate carbonate analyses. Although accurate results of lithophile refractory elements can be obtained using these glasses, for instance for chalcophile/siderophile elements such as Cd, Zn, Pb, P, Ni, Cu, matrix-matched calibration is recommended, and for this carbonate RMs are preferable (Jochum et al. 2012). Unfortunately, previous work showed that the currently available carbonate RMs can be more heterogeneous at the microscale than glass RMs, as demonstrated for the synthetic carbonate USGS MACS-3 (Jochum et al. 2019) when using small spot sizes of 55 μm or test portions. Moreover, the large demand of such RMs may cause a shortage in terms of accessibility; for example, USGS MACS-3 is no longer commercially available, with consequences for future inter-laboratory calibrations and comparisons. The powder of coral *Porites* JCp-1 and giant clam *Tridacna gigas* JCt-1 produced by the Japanese Geological Survey (Okai et al. 2001; Inoue et al. 2004) could be pressed into pellets, but international shipping of this material is restricted by the Convention on International Trade in Endangered Species of Wild Flora and Fauna (CITES) and these RMs too are not commercially available. Therefore, suitable RMs for in-situ carbonate analysis are sparse.

In the absence of widely available solid RMs for carbonate analysis, many laboratories press their own pellets of carbonate powder. These pellets, like other existing solid RMs, are also heterogeneous at the sub-100 µm scale. However, nano-milling is a relatively recently developed way of overcoming such homogeneity issues. Garbe-Schönberg and Müller (2014) developed a nano-milling or nano-particulate (NP) pressed powder method, which produces pellets with excellent cohesion without adding any binder. The wet milling technique results in smaller grain sizes (typically median grain size <1.5 µm for any material) than dry milling techniques with a minimum grain size of 5-10 µm. Wet milling improved homogeneity substantially (Garbe-Schönberg and Müller, 2014). For carbonates, this technique allows even smaller median grain size of 275nm (personal communication). Jochum et al. (2019) analysed MACS-3 NP, JCp-1 NP and JCt-1 NP as well as the precursor material, i.e. USGS MACS-3 and pellets pressed from the original powder of JCp-1 (Okai, 2001), JCt-1 (Inoue, 2004). They found that the NP pellets are more homogeneous by a factor of approximately 2–3 than the original precursor materials. The homogeneity of these NPs is similar to that of the NIST SRM 61x glasses and therefore more suitable for microanalytical purposes (Jochum et al. 2019).

In addition to the limited availability of matrix-matched and homogeneous RMs for mass fraction measurements, the in-situ measurement of isotope ratios, for example, $^{87}Sr/^{86}Sr$, $^{207,208}Pb/^{206}Pb$, $\delta^{13}C$ or $\delta^{18}O$ in carbonates by LA-MC-ICP-MS, Isotope Ratio Mass Spectrometry (IRMS) and secondary-ion mass spectrometry (SIMS) is also hampered by the availability of suitable microanalytical RMs. For a decade, only one reported value of the $^{87}Sr/^{86}Sr$ ratio (Ohno and Hirata, 2007) was available for the original powder JCt-1, but recently this has been complemented by measurements using multiple techniques in different laboratories (Weber et al. 2018) to establish a robust Sr isotope ratio for this RM. Particularly, there is a limited availability of RMs matching the sample mass fraction of interest. Recently, Weber et al. (2020) presented a new synthetic carbonate nanopowder RM called NanoSr, with a Sr mass fraction of ~500 µg g^{-1} for in-situ Sr isotope analyses. Furthermore, a new RM (GBW04481) for calcite oxygen and

carbon isotopic microanalysis has been developed with $\delta^{18}O_{VPDB}$ = −23.12 ± 0.15 ‰ (1s) and $\delta^{13}C_{VPDB}$ = −5.23 ± 0.06 ‰ (1s) for the Oka calcite (Tang et al. 2019). Availability of such well-characterized, high-quality carbonate RM is a prerequisite for high-precision SIMS O and C isotopic microanalysis (Tang et al. 2019).

In elemental analysis using solution measurements, calibration curves are often used with multiple RMs with a range of elemental mass fractions. In-situ techniques often use a single RM for calibration. A disadvantage of this strategy is the lack of control on linearity across a range of elemental mass fractions and the assumption that the calibration line goes through the origin, which is not the case for an interfered element. Hence, accuracy of LA-ICP-MS could be increased with the same calibration strategies as solution measurements, when more suitable carbonate RMs would be available. Despite the findings of e.g., Jochum et al. (2012) showing that using the glasses of the NIST SRM 61x series can yield accurate results when calibrating measurements of carbonates by LA-ICP-MS, it would be preferable to use in addition homogeneous carbonate RMs such as JCp-1 NP, JCt-1 NP, MACS-3 NP and for instance the here presented new RM NHFS-2-NP.

To improve the limited availability of matrix-matched RMs for elemental mass fraction and isotope ratios for in-situ techniques, we present here the new carbonate RM NFHS-2-NP (NIOZ Foraminifera House Standard-2-Nano-Pellet) made from a nano-milled (Garbe-Schönberg and Müller,2014) mixed assemblage of planktic foraminifera. In this study, we present an inter-laboratory compilation of data from solution ICP-MS, ICP-OES, LA-ICP-(MC)-MS, XRF, bulk-analyses of 46 element mass fractions (including elements relevant for environmental research such as B, Ba, Mg, Mn, Na, Sr) as well as isotope ratios ($^{87}Sr/^{86}Sr$, $^{206,207,208}Pb/^{204}Pb$, $\delta^{13}C$, $\delta^{18}O$), to establish the composition and verify the homogeneity of the NFHS-2-NP pellets. Seven laboratories contributed data for elemental mass fractions, five laboratories for $\delta^{13}C$ and $\delta^{18}O$, four laboratories for $^{87}Sr/^{86}Sr$ analysis, and one laboratory for Pb isotopes.

Material and methods

NFHS-2-NP (NIOZ Foraminifera House Standard-2-Nano-Pellet)

NFHS-2-NP (Fig. 1) was prepared from a mixed assemblage of planktic foraminifera shells, sieved from a calcareous ooze collected in section#1 (5.3-6.5 m) of a piston core (PE174-15PC), retrieved by Royal Netherlands Institute for Sea Research (NIOZ) from Walvis Ridge (2878 m water depth; 29°85S, 2°3405E). The sediments were wet sieved over 63, 180 and 355 µm mesh sizes. The 180-355 µm size fraction was subsequently cleaned according to the protocol of Barker et al. (2003), scaled up to clean ~300 g of material. The cleaned material was pre-milled with a Retsch Planetary Ball Mill PM 100 using cleaned zirconium oxide beakers. The powder NFHS1 is used as an in-house standard for $\delta^{13}C$ and $\delta^{18}O$, and solution ICP-MS for many years. To improve the homogeneity for pellets for LA-ICP-MS measurements, the powder was subsequently wet-milled using agate milling gear by myStandards GmbH, Kiel, Germany to obtain nano particles. Using a programmable hydraulic press, these particles were then pressed to a pellet (Fig. 1) with a diameter of 5 mm (4 tons) or 13 mm (10 tons). A binder was not added.

Fig. 1: Photo of a nano-milled pellet (blue arrow) made from a mixed assemblage of planktic foraminifera (in front of pellet) from the 180-355 µm size fraction of a piston core from Walvis Ridge in the South Atlantic Ocean. Photo made by Ellen R. Pruiksma.

Analytical techniques

In this study, twelve laboratories were involved in the characterization process of the new carbonate RM NFHS-2-NP (Table 1).

Table 1: Institutions involved in the characterization process of NFHS-2-NP.

Lab ID	Technique	Analyte	Institution	Acronym	Comment
1A	SF-ICP-MS	Trace elements	Alfred-Wegener-Institut, Helmholtz-	AWI	
1B	KIEL/IRMS	$\delta^{13}C$, $\delta^{18}O$	Zentrum für Polar- und Meeresforschung,		
			Lamont-Doherty Earth Observatory and		
2	Q-ICP-MS	Trace elements	Columbia University, US	LDEO	
3A	SF-ICP-MS	Trace elements	Royal Netherlands Institute for Sea	NIOZ	
3B	LA-Q-ICP-MS	Trace elements	Research, NL		
3C	KIEL/IRMS	$\delta^{13}C$, $\delta^{18}O$			
			School of Ocean and Earth Science, University of Southampton, National		
4	ICP-MS	Trace elements	Oceanography Centre Southampton	Southhampton	
					Accredited in
5	ICP-OES/Q-ICP-MS	Trace elements	Anonymous		accordance with ISO
		Major- and trace			Accredited in
6A	ICP-OES	elements	Anonymous		accordance with ISO
6B	QQQ-ICP-MS	Major- and trace			
6C	fusion XRF	Major- and trace			
7A	LA-Q-ICP-MS	Trace element	Institute for Geosciences, Johannes	JGU	
7B	MC-ICP-MS	Sr isotopes	Gutenberg-Universität, Mainz, DE		
7C	LA-MC-ICP-MS	Sr isotopes			
					Accredited in
8	TIMS	Sr isotopes	IsoAnalysis UG, Berlin, DE	IsoAnalysis	accordance with ISO
9A	TIMS	Sr, Pb isotopes	GEOMAR Helmholtz Centre for Ocean	GEOMAR	
9B	KIEL/IRMS	$\delta^{13}C$, $\delta^{18}O$	Research Kiel, DE		
			Geochemistry Lab, University of Modena		
10	LA-MC-ICP-MS	Sr isotopes	and Reggio Emilia, IT	Modena	
11	KIEL/IRMS	$\delta^{13}C$, $\delta^{18}O$	MARUM, University of Bremen, DE	Marum	
12	KIEL/IRMS	$\delta^{13}C$, $\delta^{18}O$	Stable Isotope Laboratory, UC Davis, US	Davis	

Seven laboratories were involved in trace, minor and major elemental mass fraction analysis, five in stable isotopes analysis, four in strontium isotope analysis, and one laboratory in Pb isotope analysis. Data resulting from TIMS, MC-ICP-MS, sector field (SF)-ICP-MS, Q-ICP-MS, triple quadrupole (QQQ)-ICP-MS, ICP-OES, LA-MC-ICP-MS, LA-Q-ICP-MS, XRF, KIEL-IRMS are presented and discussed. In the following, analytical procedures and most important features of the instruments used in this study are described.

Elemental mass fractions

Mass fractions of the powder of NFHS-2-NP were determined for 46 elements (Al, B, Ba, Cd, Ce, Co, Cr, Cs, Cu, Dy, Er, Eu, Fe, Ga, Gd, Hf, Ho, K, La, Li, Lu, Mg, Mn, Na, Nd, Ni, P, Pb, Pr, Rb, S, Sc, Si, Sm, Sn, Sr, Tb, Th, Ti, Tm, V, U, Y, Yb, Zn, Zr) by AWI, LDEO, NIOZ, University of Southampton and two anonymous laboratories (Lab 5 and 6) using SF-ICP-MS, Q-ICP-MS, and ICP-OES after HNO_3/HF, HNO_3/HF/HCl,

HNO_3/HF/$HClO_4$ or HNO_3 digestion. At Lab 6, the powder was analysed by WD-XRF in addition to ICP-OES and QQQ-ICP-MS analysis.

Participants were asked to not clean the samples again before analysis, but to analyse them as provided.

At AWI (Lab 1A), a subsample (~5 mg) of the NFHS-2-NP powder was dissolved in double-distilled 1 mol l^{-1} HNO_3, from which a small aliquot was diluted with double-distilled 0.3 mol l^{-1} HNO_3 to gain a target [Ca] of ~10 µg g^{-1}. The sample measurements were bracketed by measurements of a 5-point calibration block made from an in-house artificial foraminiferal RM. A solution of JCp-1 was run along with the sample as a quality control to ensure the accuracy of the measurements. The liquids were introduced into a Nu Instruments Attom high-resolution ICP-MS via a Savillex PFA nebulizer (self-aspirating at ~100 µL min^{-1}) and an ESI quartz cyclonic spray chamber. The RMs and samples were analysed two times in quintuplicate for the isotopes ^{7}Li, ^{11}B, ^{25}Mg, ^{27}Al, ^{43}Ca, ^{51}V, ^{87}Sr, ^{111}Cd, ^{137}Ba, ^{146}Nd, and ^{238}U in low-resolution mode. From the 10 measurements in total of the NFHS-2-NP, the average concentrations and repeatability are reported.

At LDEO (Lab 2), cleaned carbonate samples were dissolved in 2 % HNO_3 and analysed for [Ca] and element/Ca ratios on a Thermo Scientific iCAP-Q Q-ICP-MS. First, a small aliquot (10-20 %) of this solution was used to determine the [Ca]. Upon [Ca] determination, each sample solution was diluted to 50 µg g^{-1} [Ca]. Signal intensities of trace elements were then measured using a Thermo Scientific iCAP Q ICP-MS following methods adapted from Yu et al. (2005). Signals of low-mass isotopes ^{11}B were measured in standard mode; ^{23}Na, ^{27}Al, ^{43}Ca, ^{55}Mn, ^{56}Fe and ^{111}Cd were measured in kinetic energy discrimination mode using a He collision cell to minimize polyatomic interferences. To minimize matrix effects and to account for instrumental drift, signal intensities were corrected with an internal multi-element quality

control reference solution (containing 50 µg g^{-1} Ca) measured every ten samples. Concentrations and element ratios (relative to Ca) were calculated from signal intensities measured on a multi-element stock solution prepared gravimetrically from trace element-grade stock solutions. These reference solutions contained 50 µg g^{-1} Ca and were prepared with a range of trace element concentrations reflective of typical foraminiferal samples (e.g. Mg/Ca between 0.5 - 5.5 mmol mol^{-1}), creating calibrations of element intensity to concentration.

At NIOZ (Lab 3A), 20 mg oven dried (110 °C for 2 hours) powder of NFHS-2-NP (n=5) was weighted on a 6 decimal balance and was digested in 4 ml Teflon rocket beakers after a total digestion procedure using concentrated ultrapure HNO_3/HF/HCl (Durand et al. 2016), adjusted to a weight of 20 mg of the RM. After removal of the HNO_3/HF/HCl matrix in a class 5 flow bench in a cleanroom class 7, the digested RM was dissolved in ultrapure 0.1 mol l^{-1} HNO_3 prepared with low boron ultrapure water (>18.2 MW cm^{-1}). In addition to the total digestion and for comparison, the NFHS-2-NP (n=2) was dissolved in ultrapure 0.1 mol l^{-1} HNO_3. To determine the [Ca] and element/Ca ratios in the digested RM, we used a SF-ICP-MS Thermo Fischer Scientific Element-2. For the [Ca] analysis, intensities of masses ^{44}Ca, ^{45}Sc and $^{87}Sr^{++}$, ^{87}Sr, and ^{88}Sr were measured in medium resolution. The intensity of ^{44}Ca was corrected for an $^{88}Sr^{++}$ interference (Sr^{++}/Sr was 1.4 %). ^{45}Sc was used as an internal standard to correct for drift during the measurement. Accordingly, to these [Ca], solution of NFHS-2-NP was diluted to obtain 100 µg g^{-1} Ca. We measured samples in replicate by monitoring the isotopes ^{7}Li, ^{11}B, ^{23}Na, ^{25}Mg, ^{43}Ca, ^{66}Zn, ^{85}Rb, ^{87}Sr, ^{88}Sr, ^{112}Cd, ^{121}Sb, ^{138}Ba, ^{208}Pb, ^{238}U in low resolution, the isotope ^{27}Al, ^{31}P, ^{32}S, ^{43}Ca, ^{51}V, ^{52}C, ^{55}Mn, ^{56}Fe, ^{59}Co, ^{60}Ni, ^{63}Cu in medium resolution and the isotopes ^{39}K, ^{43}Ca, ^{75}As in high resolution. To determine element/Ca (mmol/mol) ratios, we used a ratio calibration method (De Villiers et al. 2002) with all RMs, QCMs JCp1 and JCt1, drift standards and six calibration standards having a similar matrix of 100 ± 5 µg g^{-1} Ca in 0.1 mol l^{-1} HNO_3.

At the University of Southampton (Lab 4), a 75 mg subsample of the RM was dissolved in an acid washed PFA vial with 5 ml 0.5mol l^{-1} HNO_3 that was prepared using thermally distilled HNO_3 and low boron Milli-Q water. Following Henehan et al. (2015), analysis of element/Ca ratios was performed with a Thermo Scientific Element II SF-ICP-MS with a Teflon barrel spray chamber introduction system, into which ammonia gas was added to aid the washout of boron (Al-Ammar et al. 2000). Standards and samples were signal-matched at 1 mmol l^{-1} Ca, and a standard bracketing technique was employed to correct mass bias drift within the run, using an in-house multi-element standard which was measured between every 3 samples/standards. All isotopes (^{7}Li, ^{11}B, ^{23}Na, ^{24}Mg, ^{25}Mg, ^{27}Al, ^{43}Ca, ^{48}Ca, ^{55}Mn, ^{86}Sr, ^{87}Sr, ^{111}Cd, ^{138}Ba, ^{146}Nd, ^{238}U, ^{56}Fe) were measured in low resolution mode except for ^{56}Fe which was measured in medium resolution. Isobaric interferences from doubly charged ^{48}Ca on ^{24}Mg and ^{86}Sr on ^{43}Ca were corrected using measurements of doubly charged ^{43}Ca and ^{87}Sr, respectively. In-house consistency standards with varying concentrations of elements relative to Ca were used to check the validity of the analysis. Five measurements of the RM were made during the analytical session and the reproducibility (RSD) was better than 2% for all element/Ca ratios except for those involving ^{11}B, ^{56}Fe, ^{111}Cd and ^{238}U, which were better than 4%.

At Lab 5, 250 mg of sample (n=5) was dissolved in perchloric ($HClO_4$)-, nitric (HNO_3), hydrochloric (HCl)- and hydrofluoric acid (HF). The remaining residue after drying was dissolved with dilute HCl and analysed by ICP-MS and ICP-OES. Results are corrected for spectral and interelement interferences. Further information was omitted.

At Lab 6, the powder was analysed by WD-XRF as a fused bead after lithiumborate fusion as well as by QQQ-ICP-MS and ICP-OES after aqua regia (HNO_3+HCl, 1 to 3 ratio) and 4-acid (see ALS) digestion. Sample weights for XRF, aqua regia and 4-acid were 1 g, 1 g and 250 mg respectively. Further information was omitted.

The participants report their data either as element to calcium ratios (µmol mol^{-1} Ca or mmol mol^{-1} Ca) obtained by an intensity ratio calibration method (e.g. de Villiers et al. 2002) or as elemental mass fractions (% m/m or µg g^{-1}). To allow better intercomparison of the data, we used the mean [Ca] of 38.5±0.6 % m/m (CaO 53.9±0.9 % m/m) determined after a three-acid total digestion by SF-ICP-MS (NIOZ), after a four-acid total digestion by ICP-OES (Lab 6B) and XRF (Lab 6C). This value was subsequently used to convert elemental ratios into elemental mass fractions, which we report throughout the rest of the manuscript.

Elemental mass fraction homogeneity

Testing for elemental homogeneity was done according to ASTM-Guide E826-14 by dividing the surface of the test pellets into distinct zones (Fig. 2). Analyses were performed at NIOZ (Lab 3B) and JGU (Lab 7A). Operation parameters are provided in Table 2.

Table 2: Operating parameters LA-Q-ICP-MS at NIOZ and JGU.

	NIOZ	JGU
Laser ablation system:	ESI NWR193UC	ESI NWR193
Sample cell	Two-Vol2	Two-Vol2
Wavelength	193 nm	193 nm
Pulse duration	4 ns	4 ns
Laser fluence	3 J cm^{-2}	3 J cm^{-2}
Laser spot size	60 µm	70 µm
Laser repetition rate	10 Hz	10 Hz
Carrier gas flow rate (He)	0.6 l min-1	0.7 l min-1
Ablation time	40 s	40 s
ICP-MS:	Thermo Fisher Scientific iCAP-Q	Agilent 7500ce
Measuring mode	Standard (no collision gas)	Standard (no collision gas)
Cooling gas flow rate	17.5 l min-1	15 l min-1
RF power	1550 W	1200 W
Auxilliary gas flow rate	0.75 l min^{-1}	1 l min^{-1}
Sample gas flow rate (Ar)	0.45 l min^{-1}	0.8 l min^{-1}
Monitored isotopes	^{11}B, ^{23}Na, ^{25}Mg, ^{27}Al, ^{29}Si, ^{43}Ca, ^{55}Mn, ^{57}Fe, ^{66}Zn, ^{88}Sr, ^{89}Y, ^{137}Ba, ^{208}Pb, ^{238}U	^{7}Li, ^{9}Be, ^{11}B, ^{23}Na, ^{25}Mg, ^{27}Al, ^{29}Si, ^{31}P, ^{39}K, ^{43}Ca, ^{45}Sc, ^{47}Ti, ^{49}Ti, ^{51}V, ^{53}Cr, ^{55}Mn, ^{56F}e, ^{57}Fe, ^{59}Co, ^{60}Ni, ^{63}Cu, ^{65}Cu, ^{66}Zn, ^{67}Zn, ^{71}Ga, ^{75}As, ^{85}Rb, ^{86}Sr, ^{88}Sr, ^{89}Y, ^{90}Zr, ^{93}Nb, ^{98}Mo, ^{111}Cd, ^{118}Sn, ^{133}Cs, ^{135}Ba, ^{137}Ba, ^{139}La, ^{140}Ce, ^{141}Pr, ^{143}Nd, ^{147}Sm, ^{153}Eu, ^{157}Gd, ^{159}Tb, ^{163}Dy, ^{165}Ho, ^{167}Er, ^{169}Tm, ^{173}Yb, ^{175}Lu, ^{177}Hf, ^{184}W, ^{208}Pb, ^{232}Th, ^{238}U
Integration time	0.01 s all, but ^{11}B (0.1 s), ^{25}Mg (0.02 s) and ^{57}Fe (0.02 s)	0.01 s
Time per pass	280 ms	660 ms
ThO/Th	0.6%	0.4%
U/Th	1.06	1.1
Lab ID.	3B	7A

The larger the spot size, the more material and potential inhomogeneities are averaged (Jochum et al. 2012). To minimize such adverse smoothing, a spot size of 60 and 70 µm was chosen.

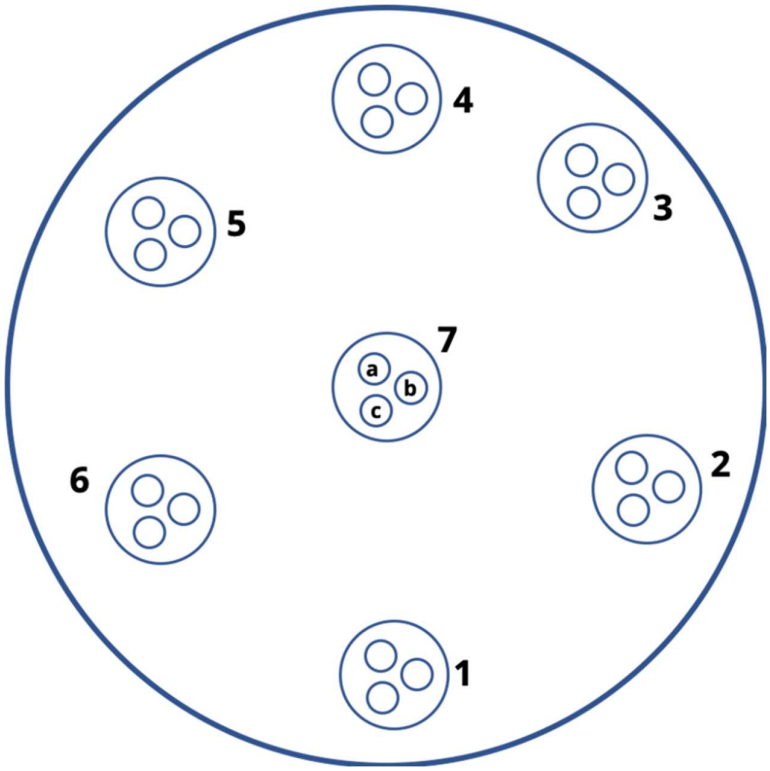


Fig. 2: Laser ablation spots (7 times 3) used for testing homogeneity. Letters a,b,c are only shown in section 7 as an example and omitted in the other sections for image clarity.

At NIOZ (Lab 3B), four pellets of NFHS-2-NP were analysed with 7x3 spots programmed on every pellet in a pattern shown in Fig. 2. The distance between the 3 spots that were close to each cluster was around 200 µm. Spots were measured in a random order to avoid systematic offsets. Ablation was performed with helium as carrier gas and mixed with argon before entering the ICP-MS. A self-made squid following the design of Eggings et al. (1998) was used to produce a more stable signal. Two interface vacuum pumps were used to increase the sensitivity. Backgrounds were measured for 40 s and ablation time was 40 s, followed by 5 s of wash out. Data reduction was performed using an adapted version of the data reduction software SILLS (Signal Integration for Laboratory Laser Systems; Guillong et al. 2008) in Matlab. Data from the first 5 seconds of every laser ablation profile were excluded to avoid potential surface contamination. To calculate element mass fractions, relative sensitivity factors using ^{43}Ca as internal

standard were calculated using both NIST SRM 610 and NIST SRM 612 as calibrants using the reference values reported in the GeoReM database (Jochum et al. 2005, 2011). The intensity of ^{43}Ca was corrected for an $^{86}Sr^{++}$ interference (Sr^{++}/Sr was 0.39 %). QCMs JCp1 NP (n=21), JCt1 NP (n=21), ECRM752 NP (n=21), RS3 NP (n=21), and MACS-3 (n=24) were measured in the same sequence to monitor precision and accuracy of the applied protocol. JCp-1 NP made by myStandards GmbH was used to quantify drift and was measured in triplicate after every seventh sample.

At JGU (Lab 7B), one pellet was analysed with 7x3 patterns (repeated three times) as shown in Fig. 2. Ablation was carried out under a He atmosphere and the sample gas was mixed with Ar before entering the plasma. Backgrounds were measured for 20 s prior to each ablation. Ablation time was 40 s, followed by 20 s of wash out. Signals were recorded in time-resolved mode and processed using an in-house Excel spreadsheet (Jochum et al. 2007). Details of the calculations are given in Mischel et al. (2017). NIST SRM 610 and 612 were used as calibration material, applying the reference values reported in the GeoReM database (http://georem.mpch-mainz.gwdg.de/, Application Version 27; Jochum et al. 2005; Jochum et al. 2011) to calculate the element concentrations of the sample measurements. During each run, basaltic USGS BCR-2G, synthetic carbonate USGS MACS-3 and biogenic carbonate JCp-1 NP and JCt-1 NP were analysed repeatedly as QCMs to monitor precision and accuracy of the measurements as well as calibration strategy. All RMs were analysed at the beginning and at the end of a sequence and after 21 spots on the samples. For all materials ^{43}Ca was used as internal standard applying for NFHS-2-NP the preferred values reported in the GeoReM database for NIST SRM 610 and 612, the QCMs USGS BCR-2G and MACS-3 and for JCp-1 and JCt-1 the values provided by Okai (2002) and Aizawa (2008), respectively. Resulting element mass fractions for the QCMs agreed typically within uncertainties with published values.

$^{87}Sr/^{86}Sr$ isotope ratios

Sr isotope ratios of the powder were measured by three different laboratories (IsoAnalysis UG, GEOMAR, JGU) using two isotopic techniques: TIMS and MC-ICP-MS.

At JGU (Lab 7B), the solution based $^{87}Sr/^{86}Sr$ ratios were obtained by coupling the Thermo Neptune Plus MC-ICP-MS with an ESI Apex Omega HF desolvator system, following the methods described in Weber et al. 2020. Prior to analysis, about 8 mg of NFHS-2-NP was dissolved in 5 ml of 3 mol l^{-1} HNO_3 in a closed PFA beaker overnight on a hotplate at 120 °C. Ten aliquots of 300 µL from the digested material were processed together with a modern seashell following the methods described in Lugli et al. (2020) and Weber et al. (2018). All solutions were prepared at 10 ng g^{-1} Sr in 0.8 mol l^{-1} HNO_3 and analysed in a standard sample bracketing sequence using NIST SRM 987 for bracketing and normalization to $^{87}Sr/^{86}Sr$ = 0.710248. Within the same analytical session, NIST SRM 987 yielded an average $^{87}Sr/^{86}Sr$ of 0.71026 ± 0.00003 (2s, n = 57). The modern seashell yielded an $^{87}Sr/^{86}Sr$ of 0.70917 ± 0.00001 (2SE, n = 1). Data evaluation at JGU was performed offline, using an in-house R script (R Core Team, 2020).

At IsoAnalysis UG (Lab 8) about 60 mg of sample were dissolved in 6 mol l^{-1} HNO_3, using PFA vials, at 120 °C. Sr was separated ion chromatographically (Eichrom® Sr-Spec resin) and analysed in three separate aliquots sevenfold using TIMS. The raw data were converted using a $^{86}Sr/^{88}Sr$ ratio of 0.1194 and a best practice value of 0.71025 for NIST SRM 987. Rb was measured on mass 85 and the data were corrected for a potential proportion of ^{87}Rb.

At GEOMAR (Lab 9A) about 100 mg of NFHS-2-NP was dissolved in 3 ml of 6 mol l^{-1} HCl at 100 °C overnight in closed 7 ml Savillex® Teflon beakers. Samples beakers were then opened, and samples dried

at 110 °C. Sr separation used a single pass through BioRad Micro Bio-Spin™ columns filled with ca 150 µL of Eichrom® Sr-Spec resin (50-100 µm) equilibrated in 6 mol l^{-1} HNO_3. Sr was cleaned by fivefold addition of 500 µL increments of 6 mol l^{-1} HNO_3 followed by a final washout with 2 ml 0.05 mol l^{-1} HNO_3. USGS BCR-2 (~100 mg) QCM, leached in 2 mol l^{-1} HCl at 70 °C prior to dissolution, was routinely processed along with NFHS-2-NP but followed the silicate digestion procedure outlined in Hoernle et al. (2008). Sr isotope ratios were determined on a Thermo Scientific® TRITON Plus TIMS operating multi-dynamic mode for Sr. Within run mass bias correction uses $^{86}Sr/^{88}Sr = 0.1194$. Possible mass interference of ^{87}Rb on ^{87}Sr were monitored by the ^{85}Rb signal, but none was detected in any of the measurements. Reference material NIST SRM 987 (Sr) was repeatedly measured along with the samples and gave $^{87}Sr/^{86}Sr = 0.710262 \pm 0.000007$ (2s, n=8). The laboratory routinely reports sample data relative to $^{87}Sr/^{86}Sr = 0.710250$ for NIST SRM 987 with a long-term uncertainty of ±0.000009 (2s, n=834 since 2014). Total chemistry blanks were < 100 pg Sr and therefore considered negligible relative to the amounts of sample used. Leached reference material USGS BCR-2 gave $^{87}Sr/^{86}Sr = 0.705001 \pm 0.000005$ (2SE) and lies within those of Fourny et al. (2016) for unleached USGS BCR-2.

Spatial homogeneity was tested by LA-MC-ICP-MS in two laboratories (JGU and Modena). JGU (Lab 7C) used a line scan with circular spot of 70 µm and line length of 500 µm (n=21). Modena (Lab 10) employed two different sampling strategies: line scans with circular spot of 100 µm and a line length of 600 µm (n=12), and spot scans with a size of 80 µm and a dwell time of 50 s (n=3). Operation parameters are provided in Table 3.

Table 3: Operating parameters LA-MC-ICP-MS at Modena and JGU.

	JGU	Modena
Laser ablation system:	ESI NWR193	New Wave UP 213
Sample cell	Two-Vol^{2}	Single volume
Wavelength	193 nm	213 nm
Pulse duration	4 ns	4 ns
Laser fluence	5 J cm^{-2}	10 J cm^{-2}
Spot size - line scans	70 µm	100 µm
Line length	500 µm	600 µm
Spot size - depth profiling	N/A	80 µm
Laser repetition rate	10 Hz	10 Hz
Carrier gas flow rate (He)	0.8 L min^{-1}	0.6 L min^{-1}
Ablation time	100 s	Spots 50 s, Lines 120 s
ICP-MS:	Thermo Neptune Plus	Thermo Neptune
Resolution	Low	Low
Cooling gas flow rate	15 L min^{-1}	15 L min^{-1}
RF power	1200 W	1200 W
Auxiliary gas flow rate	0.7 L min^{-1}	0.8 L min^{-1}
Sample gas flow rate (Ar)	0.9 L min^{-1}	1 L min^{-1}
Lab ID.	7C	10

At Modena (Lab 10), analyses were performed using a NewWave UP 213 nm laser ablation system, coupled to a Neptune MC-ICPMS, following Lugli et al. (2020) and Weber et al. (2018). Potential interferences of Kr on the isotopes of interest were corrected by subtracting the on-peak baseline acquired during the laser warm-up (~ 30-60 s). To correct for the presence of isobaric Rb on m/z 87, the ^{85}Rb was measured and a ^{87}Rb/^{85}Rb ratio of 0.3856656 was used. Mass bias normalization of Rb and Sr was performed using the exponential law and a stable ^{88}Sr/^{86}Sr ratio of 8.375209 (Berglund and Wieser 2011). Rare-earth element and Ca dimers/argides were monitored, but not corrected because their interference with the analyses was negligible. A modern seashell (*Acanthocardia* sp., Adriatic Sea; expected seawater ^{87}Sr/^{86}Sr ratio: ~0.70917 from e.g. Palmer and Edmond, 1992) and the NanoSr RM (Weber et al. 2020, reference ^{87}Sr/^{86}Sr ratio: **0.70756 ± 0.00003**) were also analysed during the session,

yielding a $^{87}Sr/^{86}Sr$ ratio of 0.70918 ± 0.00002 (2s, n = 5) and of 0.70753 ± 0.00001 (2s, n = 2), respectively.

At JGU (Lab 7C), a Neptune Plus (Thermo Scientific) MC-ICP-MS was coupled to an ArF Excimer 193 nm laser ablation system (ESI NWR193) with a TwoVol^2 ablation cell. To introduce N_2 and thereby enhance sensitivity, a CETAC Aridus 3 desolvating system was coupled to the sample line (Weber et al. 2020). The analytical methods are similar as described for Modena and references therein. Sampling positions for NFHS-2-NP were chosen according to Fig. 2 and analysed in a random order, resulting in a total number of 21 measurements. To account for potential offsets in $^{87}Sr/^{86}Sr$, JCt-1 was used as RM ($^{87}Sr/^{86}Sr$ = 0.70917 ± 0.00003, 2s, n = 15). Within the same analytical session JCp-1 was analysed as QCM ($^{87}Sr/^{86}Sr$ = 0.70916 ± 0.00002, 2s, n = 6), as well as two bioapatite QCMs to check for matrix dependent offsets in $^{87}Sr/^{86}Sr$. A modern shark tooth yielded an $^{87}Sr/^{86}Sr$ of 0.70919 ± 0.00002 (2s, n = 9, Weber et al. 2020) and the African elephant molar AGLOX yielded an $^{87}Sr/^{86}Sr$ of 0.70992 ± 0.00012 (2s, n = 6, Weber et al. 2021). All resulting $^{87}Sr/^{86}Sr$ ratios agree, within uncertainties, with their respective literature value.

Pb isotope ratios

At GEOMAR (Lab 9A), the Pb chemistry protocol used the same dissolution as for the Sr isotope ratio measurements. Pb-separation carried out a two pass 150 µl microcolumn procedure filled with BioRad AG®1x8 Anion Exchange Resin (Hoernle et al. 2008). Pb isotope ratios were determined on a Thermo Scientific® TRITON Plus TIMS operating in static multi-collection mode and mass bias correction used Pb double spike (DS) of Hoernle et al. (2011).

The long term Pb-DS corrected NIST SRM 981 values are $^{206}Pb/^{204}Pb$ = 16.9408 ± 0.0018, $^{207}Pb/^{204}Pb$ = 15.4975 ± 0.0018, $^{208}Pb/^{204}Pb$ = 36.7207 ± 0.0047, $^{207}Pb/^{206}Pb$ = 0.914804 ± 0.000049 and $^{208}Pb/^{206}Pb$ =

2.167583 ± 0.000098 (2s, n = 169) and lie well within the recommended NIST SRM 981 poly-spike values of Taylor et al (2014). Total chemistry blanks were <30 pg Pb and this considered negligible relative to the amounts of sample used. QCM of leached BCR-2 (USGS) gave $^{206}Pb/^{204}Pb$ = 18.8033 ± 0.0005, $^{207}Pb/^{204}Pb$ = 15.6242 ± 0.0005, $^{208}Pb/^{204}Pb$ = 38.8271 ± 0.00016 (2SE repeatability for all) and are within those of Fourny et al. (2016) and Todd et al. (2015) for leached BCR-2.

$\delta^{13}C$ and $\delta^{18}O$ values

$\delta^{13}C$ and $\delta^{18}O$ values were determined by five laboratories, AWI (Lab 1B), GEOMAR (Lab 9B), MARUM (Lab 11), NIOZ (Lab 3C), UC Davis (Lab 12), using automated carbonate preparation device of where the principle applies to all instruments, regardless of manufacturer (Fisons Optima IRMS, KIEL IV systems coupled to a MAT252, MAT253(plus)).

In the automated carbonate preparation device, samples and RMs react with phosphoric acid at temperatures of 70-90 °C (Table S1) to form $CaHPO_4$ (precipitate), H_2O and CO_2 (Swart et al. 1991). After separating the CO_2 from water vapor using a cryogenic trap, the CO_2 is measured using a sector field gas IRMS. Three laboratories measured NFHS-2-NP against NBS-19, while two laboratories against their respective house standards, themselves calibrated against NBS-19 (Table S1).

To examine the potential different behavior of the very fine grained NFHS-2-NP, NIOZ (Lab 3C) performed an experiment leaving the RM in prolonged contact with humid air (in contrast to the storing in drying cabinet or desiccator). The increased specific surface area of the powder makes it potentially more

susceptible to adsorption of moisture from air. Consequently, we cannot exclude a priori a slow reaction/recrystallization in the presence of absorbed water and changing $\delta^{18}O$ values by oxygen exchange. NFHS-2-NP was first dried at 105 °C for 3 hours and split into six aliquots and filled into six glass bottles. One bottle was closed directly and kept in a desiccator. The other five bottles were kept open in a laminar flow hood class 5 (preventing contamination from external sources), and in contact with air with a humidity of 45-57 %. Each week a bottle was closed with aluminum caps with rubber seals, sealed with parafilm, and stored in a desiccator upon analysis. Each subsequent bottle had one week additional and potential reaction time. The sixth and final sample had accordingly been in contact with humid air for five weeks. All samples were measured five or six times in a random order in two sequences.

Data evaluation

To yield reference values for the mass fractions, data obtained by bulk analyses of 46 elements by laboratories 1-6 were evaluated following ISO 17034:2017 (see references) and ISO Guide 35:2017 (see references). A prerequisite for this approach is that the values in the dataset are approximately normally distributed. This was tested here by applying a Kolmogorov-Smirnov test and a Grubb's test to identify outliers. Identified outliers were not removed unless there was a clear, technical/analytical reason that explained the abberant value(s). Accordingly, the assigned values were calculated by applying eq. 1, ISO Guide 35:2017 Annex A.2.4. equation A.1. We call these assigned values the *mean of means* throughout this paper.

$$Mean\ of\ means = \frac{\sum y_i}{p} \qquad (1)$$

Where y is the arithmetic mean and *p* the number of data sets.

Secondly, the uncertainty component for this value was calculated using eq. (2), ISO Guide 35:2017 Annex A.2.5.3, equation A.4. We call this uncertainty component, *reproducibility* throughout the paper.

$$Reproducibility = \frac{Standard\ deviation\ (y)}{\sqrt{p}} \quad (2)$$

Then, the *expanded reproducibility* from characterization was calculated using eq. 3.

$$Expanded\ reproducibility = Reproducibility \times k \quad (3)$$

The expansion factor (k) is determined by the effective degrees of freedom within our dataset assuming a normal distribution. The degrees of freedom are obtained by subtracting the number of observations for each dataset. Thus, if two laboratories provide 5 data points, the number of observations is 10 and degrees of freedom is 8. If the effective degrees of freedom are >10, this factor is k = 2, ISO Guide 35:2017 10.4. If the effective degrees of freedom are < 10 the appropriate factor from a Student's-t distribution should be applied. In this case, the reproducibility is represented at the 95 % confidence level using a factor of k=2 because all assigned values have >10 degrees of freedom. No fully ISO compliant homogeneity or stability test was performed on the pellets. Therefore, the real reproducibility will be higher. However, homogeneity and stability usually contribute the least to the full reproducibility.

We chose to use the Horwitz test (Horwitz and Albert, 1995) to evaluate the reproducibility of all mass fractions determined for the 46 elements by laboratories 1-6. This test calculates an expected relative standard deviation (RSD) (eq.4) based on the analyte's mass fraction C expressed as a decimal fraction (i.e. 1 g $100g^{-1}$ = 0.01), which is subsequently compared to the measured RSD (eq.5).

$$RSD\%_{expected} = 2 \times C^{(-0.1505)} \quad (4)$$

$$RSD\%_{measured} = \frac{Expanded\ reproducibility}{Mean\ of\ means} \times 100 \quad (5)$$

$$HORRAT = \frac{RSD\%_{measured}}{RSD\%_{expected}} \quad (6)$$

The resulting ratio, called the *Horwitz-Ratio (HORRAT*, eq.6), indicates whether precision is acceptable (HORRAT ≤ 1) or not (HORRAT close to 2). The calculated expanded reproducibility is based on all participants that provided data for a specific analyte, excluding the LA-ICP-MS data.

In summary, we followed recommendations outlined in ISO Guides (see references). Practically, after applying the Kolmogorov-Smirnov test to check for a normal distribution and a Grubbs test for identifying outliers, we calculated the mean of means, the reproducibility and the expanded reproducibility and finally performed a Horwitz test to compare the measured RSD with the expected RSD. The criteria for classification of the mean of means into reference and information values are (1) data from more than one laboratory, and (2) useable reproducibility using a Horwitz-Test.

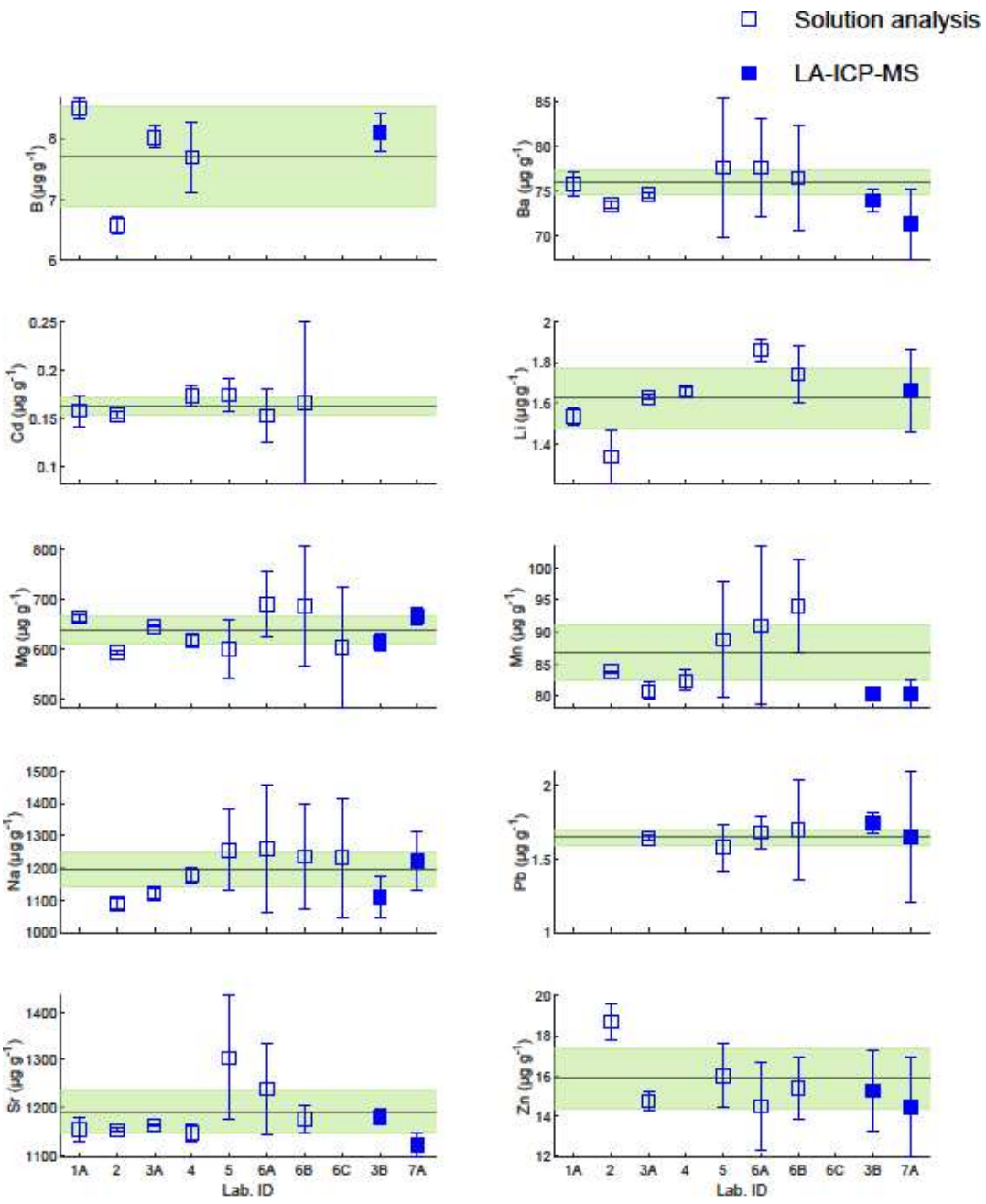
Pellet homogeneity was evaluated according to the ASTM Guide E826-14. Homogeneity can be evaluated on element to calcium ratios or on elemental mass fractions. This will result in equal results. Details of data evaluation of pellet homogeneity are given in supplement 1.

Results and discussion

Elemental mass fraction – reference and information values

The mass fractions of 46 elements are listed in Table S2 and those of 10 selected elements relevant for paleoclimate studies and proxy application (i.e. B, Ba, Cd, Li, Mg, Mn, Na, Pb, Sr, and Zn) are shown in figure 3. Element mass fraction values are in overall agreement a) when comparing results between laboratories and b) across methods (Fig. 3). For the selected elements, the mean of means and expanded reproducibility are shown as a horizontal line and green band, respectively (Fig. 3). In general, the mean of means and reproducibility were derived from solution techniques. In this case, however, the Li-borate fusion and XRF analysis is counted among them, as the sample matrix was decomposed. In contrast, the results of LA-ICP-MS (ID. 3B and 7A, Fig. 3) were not included in the calculation as all methods should be traceable to the SI-unit and derived following protocols in compliance with ISO 17025 (see references). It should be noted that laser ablation data are in close agreement with the mean of means.

Fig. 3: Element mass fraction of NFHS-2-NP reported by the various participating laboratories (1-7). Laboratory 1-5 used solution ICP-MS, 6A and 6B applied two different total digestion methods and applied ICPMS, ICP-OES, 6C applied XRF, and 3B and 7A determined the composition of the NFHS-2-NP using LA-ICP-MS. Error bars represent the repeatability *(95% CL) of the individual laboratories. The horizontal line represents the mean of means. The green band represents the expanded reproducibility (eq. 3).*



A HORRAT test is one way to divide elements into reference or information values. Of all 46 determined element mass fractions, 36 have a HORRAT <1 (Fig. 4) and are therefore sufficiently well-constrained to serve as reference values. Considering the age of the publication (Horwitz and Albert, 1995) and the increase in analytical instrument quality and hence data produced since then, we deliberately set the maximum precision limit at HORRAT=1 (instead of 2). In addition, we here also used the expanded reproducibility (eq. 3) to calculate the RSD. Without the application of this coverage factor (k), the

reproducibility would be lower by a factor of 2. For the elements Al, Cr, Fe, Hf, K, P, S, Ti, U, V, and Zr HORRAT values > 1 indicate insufficient reproducibility for the mass fractions to be used as reference values (Fig. 4) and are therefore stated as information values (supplement 2).

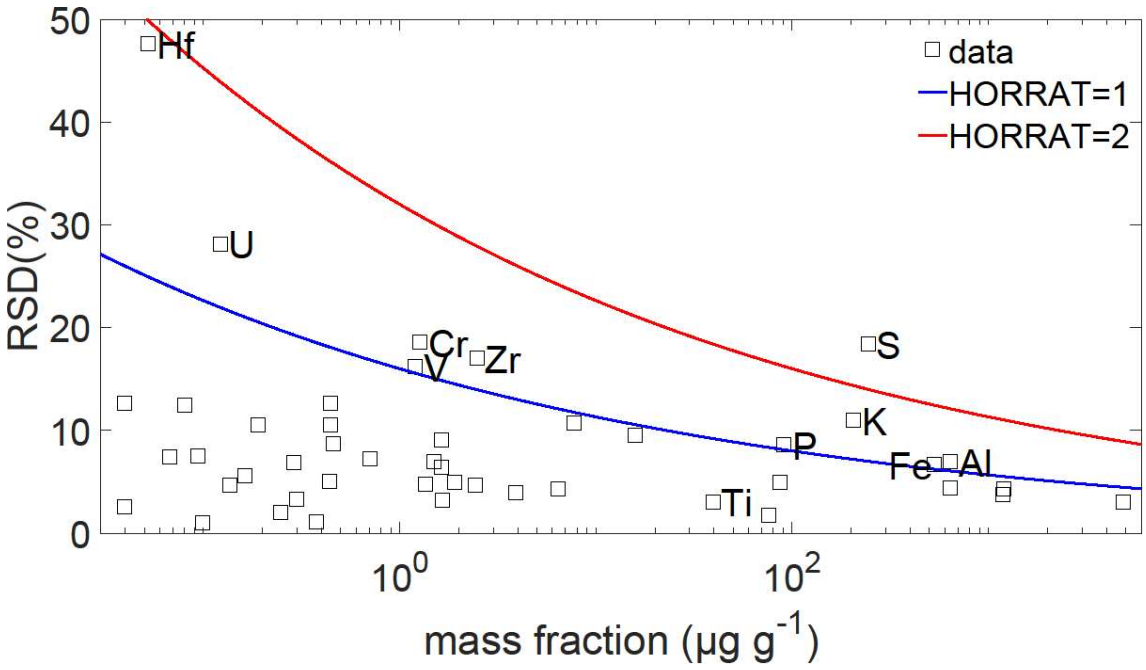


Fig. 4: Measured relative standard deviation *against mass fraction (log scale). Several elements with HORRAT above 1 (blue curve, the upper red curve indicated HORRAT=2) indicate insufficient reproducibility based on the Horwitz classification. The 36 elements with an approved reproducibility (below the blue curve) are not labeled.*

In addition to those elements that show insufficient reproducibility, the Al and Fe mass fractions of 641±45 and 532±36 µg g^{-1}, respectively (supplement 2), are high when compared to the values expected for planktic foraminiferal carbonate (Johnstone et al. 2016). Both elements are often used to indicate potential exogenous contaminations (e.g., by adhered clay particles or oxide overgrowths). Following observations of insufficient reproducibility of several elements and an offset of Al and Fe between laboratories using HNO_3 dissolution and total digestion, NIOZ directly compared 23 elements from NFHS-

2-NP using two dissolution methods. We found that the values were similar for B, Ba, Mg, Mn, Na, P, Pb, Sr, and Zn using both methods (Table S3). However, the total digestion values were higher for Cu, Li, and U, being considerably higher for S, Rb, Fe, K, V, Al, and Cr (10, 14, 17, 21, 26, 30, and 32 % respectively) than using HNO_3 only (Table S3). These elements also showed the poorest reproducibility when comparing results from the different participants using different dissolution methods (total digestion, Li-borate fusion, HNO_3 only) (Fig. 4).

HNO_3 dissolution, in this case, is believed to only extract the HNO_3-soluble fraction from the milled powder. Milling is always an abrasive process and since the foraminifera were milled in both zirconium oxide and agate (SiO2) beakers, a contamination of the resulting powder of NFHS-2-NP with SiO_2 and ZrO_2 and lattice-bound elements in both materials is likely. Since neither of these are soluble in HNO_3, it is very likely that mass fractions will be underestimated if only HNO_3 is used for dissolution. Based on these observations, we hypothesize the presence of a refractory fraction in the powder. Al and Fe could originate from aluminum- and iron silicates from the zirconia milling gear, however currently there is no available information that this milling gear contains S, Rb, K, V, and Cr. The refractory fraction may also originate from clays. The relatively low boron and magnesium mass fractions, however, contradict this hypothesis. In conclusion, there is a refractory fraction or phase of yet unclear origin in the powder, which can only be fully accounted for when using a more severe dissolution method than nitric acid. The generally good correlation between LA-ICP-MS and total digestion for these refractory elements (e.g. Al (Fig. 5)) confirms this.

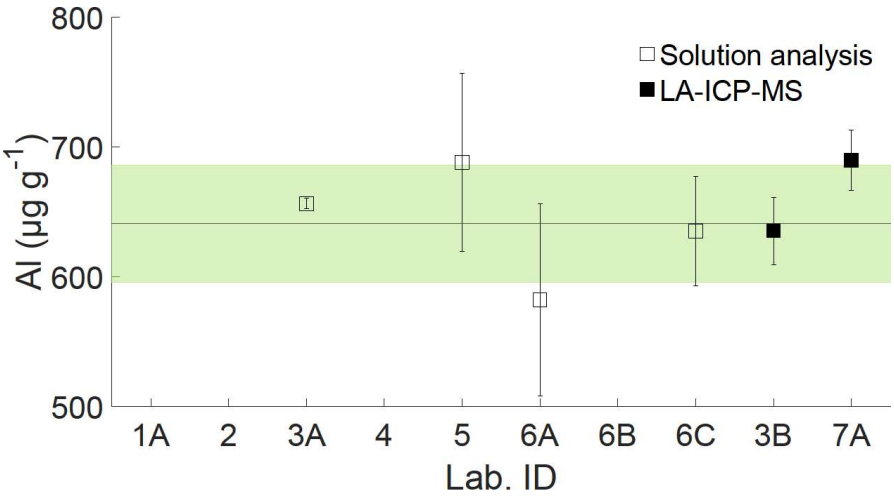


Figure 5: Mass fraction of Al reported by the various participating laboratories (1-7). Lab 3A and 5 used ICP-MS after total digestion, 6A, 6B applied two different total digestion methods and 6C applied XRF, 3B and 7A determined the composition of the NFHS-2-NP using LA-ICP-MS (Lab 1A, 2, and 4 used ICP-MS after digestion with nitric acid only and data are removed). Error bars represent the repeatability (95% CL) of the individual laboratories. The green band represents the expanded reproducibility (eq. 3) of the mean of means derived from total digestion and XRF techniques.

Total digestion using HF in combination with other strong acids is more comparable to laser ablation, which can be considered a bulk analytical technique where all the material ablated is analysed by ICP-MS. As a result of the observation of all laboratories and the NIOZ comparison, the data of Al, Fe, and V from the laboratories using only HNO_3 to dissolve the nano-milled carbonate powder of NFHS-2-NP were not included in the calculation. The data of S, Rb, K, and Cr, with lower reproducibility, were all included as they were obtained from participants using total digestion. Of the elements relevant for environmental research B, Ba, Cd, Li, Mg, Mn, Na, Sr, and Zn will be published as reference values whereas U with a low reproducibility will be published as an information value.

Elemental mass fraction – homogeneity

An ASTM Guide 826 E14 test was applied on all the LA-ICP-MS mass fraction data to check for inter- and intra-pellet homogeneity. All elements passed this test and the pellets showed homogeneity for all 46 monitored elements when analysed using a spot size of 60-70µm.

The level of homogeneity per element can be expressed as the relative standard deviation for the measured concentration within a pellet. All measured elements are shown in Fig. 6 and Table S4. The RSD of elements relevant as proxies for paleoclimate research (i.e., B, Ba, Li, Mg, Mn, Na, Sr, and Zn) are <3 % and similar to the variability of those elements in the NIST SRM 61x glasses (Jochum et al. 2011). Elements P, Ti, V, and Zr are the ones with poor reproducibility but are considered as less important elements for paleoclimate research.

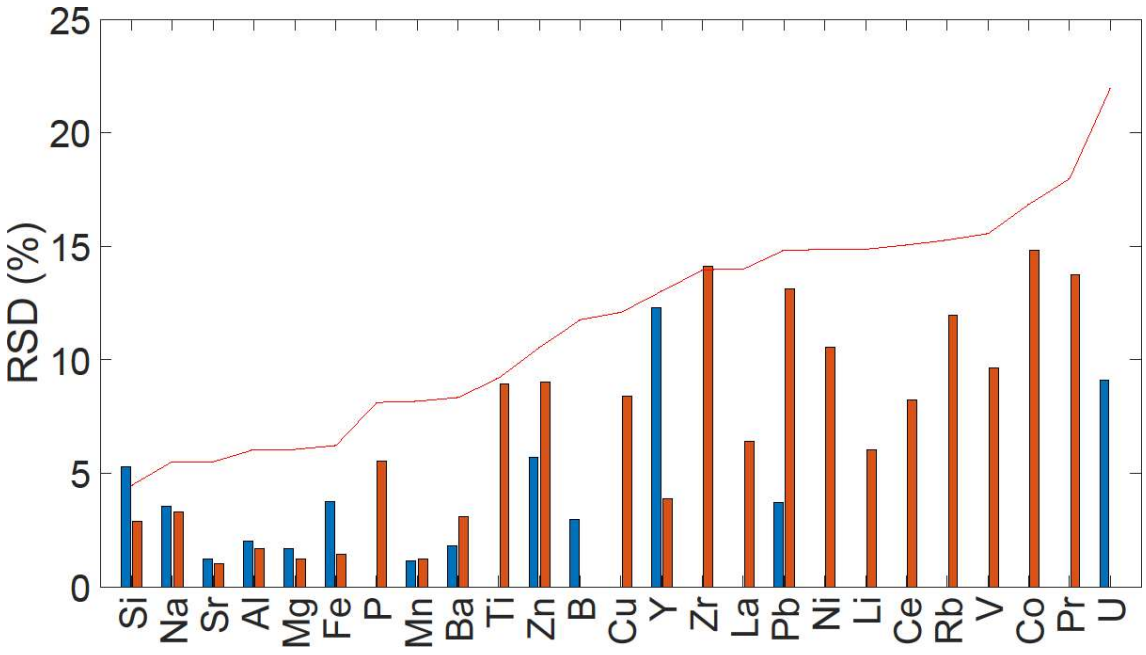


Figure 6: Relative standard deviation of elements measured by laboratories 3B (in blue) and 7A (in red) using LA-ICP-MS using 60 and 70 µm spot sizes, respectively. The red line represents the expected (maximum) relative standard deviation calculated using the Hortwitz equation (eq. 4).

Figure 7 shows a comparison of the repeatability (RSD%) as an indicator for homogeneity of three NPs (JCp-1 NP, JCt-1 NP, MACS-3 NP) and the original USGS MACS-3 (data from Jochum et al. 2019) analysed with a spot size of 55 µm, and the repeatability (RSD%) of NFHS-2-NP of this study determined by two laboratories. The lower RSDs of MACS-3 NP compared to USGS MACS-3 shows the improved homogeneity because of the nano-milling process. Homogeneity of NFHS-2-NP measured with a spot size of 60 and 70 µm is similar to that of JCp-1 NP and JCt-1 NP measured with a spot size of 55 µm. MACS-3 NP seems to be more homogeneous than JCp-1 NP and JCt-1 NP. Looking at a threshold of <3 % RSD, we can conclude that most data corresponding with MACS-3 NP and NFHS-2-NP fall in this range. These data of NFHS-2-NP in the range <3 % are the elements relevant for paleoclimate research. Some RSDs in NFHS-2-NP of in general elements with mass fractions < 10 µg g^{-1} are relatively high in comparison to the other RMs and are thus less well distributed in the pellet. These elements include the elements P, Ti, V, and Zr with poor reproducibility (Fig. 6).

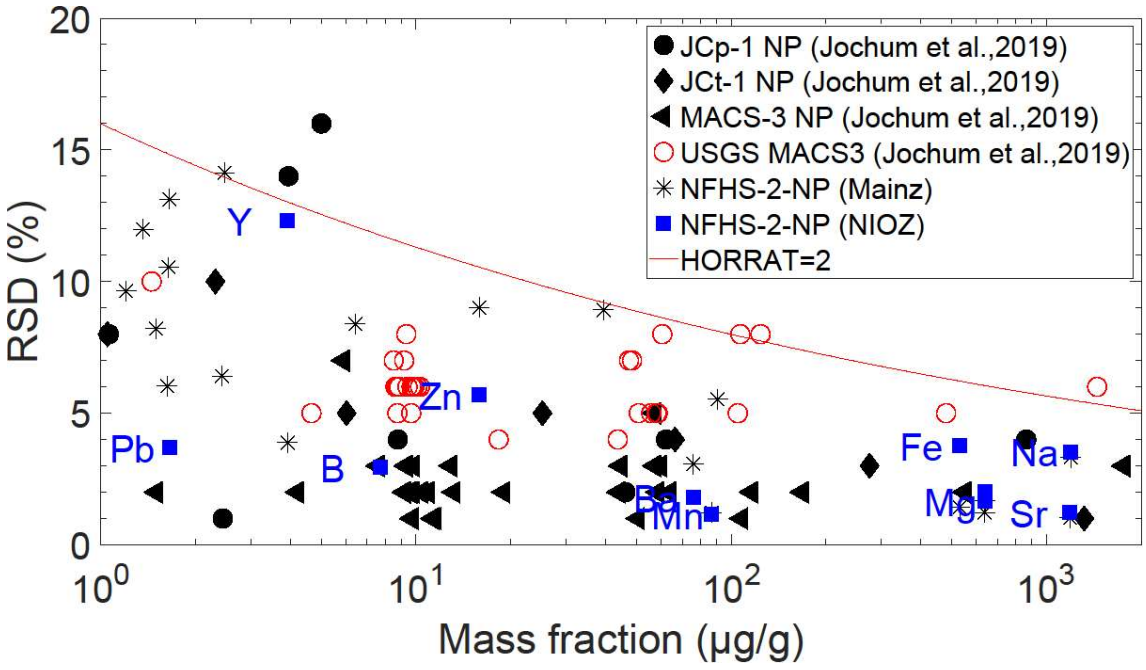


Figure 7: Relative standard deviation of JCp-1 NP, JCt-1 NP, MACS-3 NP, USGS MACS-3, and NFHS-2-NP against elemental mass fraction in µgg^{-1} (log-scale). For the simplicity of the plot only the NIOZ values are labelled and trace elements with extreme low mass fractions <1 µg g^{-1} are not shown. Elements relevant for environmental research measured by NIOZ and Mainz were in close agreement (Fig. 6).

Most elements and especially elements relevant as proxies for paleoclimate research are distributed homogenously when considering the data obtained by LA-ICP-MS of this study, implying that the pellet is suitable for in-situ techniques targeting these elements using spot sizes of approximately 60 µm and above. Also, when refractory elements (e.g., Al and Fe) are used to indicate detrital contamination in foraminifera shells, this pellet may serve as a QCM.

$^{87}Sr/^{86}Sr$ – reference values and homogeneity

A compilation of all Sr isotope data is provided in Table S5. Figure 8 shows the results of $^{87}Sr/^{86}Sr$ isotope ratio measurements of NFHS-2-NP. $^{87}Sr/^{86}Sr$ isotope ratios of the powder were analysed by three different laboratories using TIMS and solution MC-ICP-MS (ID 1-3, Fig. 8) and yield a mean of means and expanded reproducibility of $^{87}Sr/^{86}Sr$ of 0.709163 ± 0.000008, which is in good agreement with $^{87}Sr/^{86}Sr$ of 0.70917 Pleistocene seawater values (McArthur et al. 2012). The foraminifera were collected from a section of a piston core with estimated age of 450 ka-540 ka based on six ^{14}C measurements of a box core taken at the same location, assuming that the sedimentation rate was constant (Lončarić et al. 2007). In addition to the mean $^{87}Sr/^{86}Sr$ isotope ratios, we report individually measured $^{87}Sr/^{86}Sr$ isotope ratios for all participants (Table S6). Replicate analysis of NFHS-2-NP of GEOMAR (Lab 9A) by means of a second

digests are within 2s of NIST SRM 987 mentioned and indicate homogeneity of NFHS-2-NP at a sample size of 100mg (Fig. 8, Table S6).

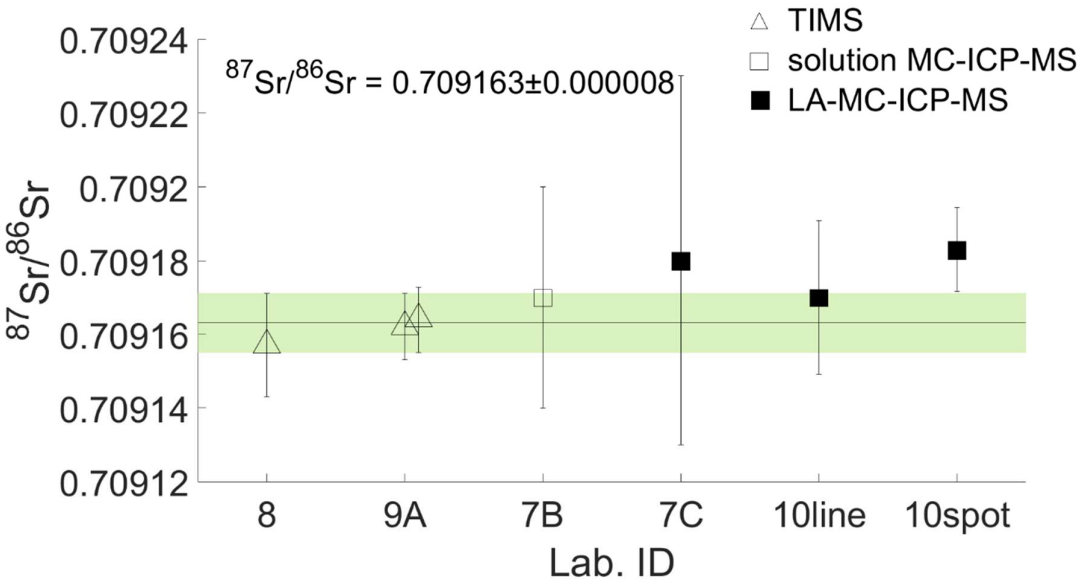


Figure 8: $^{87}Sr/^{86}Sr$ ratio against laboratory ID. The horizontal line represents the mean of means derived from TIMS and solution MC-ICP-MS. The green band represents the expanded reproducibility (eq. 3) of the mean of means. Error bars represent the 2s repeatability *of the individual laboratories.*

In addition, JGU and Modena used LA-MC-ICP-MS to evaluate the homogeneity of NFHS-2-NP for $^{87}Sr/^{86}Sr$. The LA-MC-ICP-MS analyses (Fig. 8) yield average $^{87}Sr/^{86}Sr$ ratios and repeatability of 0.70918 ± 0.00005 (2s, n = 21, Lab 7C, line scan with a line length of 500 µm and spot size of 70 µm), 0.70917 ± 0.00002 (2s, n = 12, Lab 10line, line scan with a length of 600 µm and spot size of 100µm), and 0.70918 ± 0.00001 (2s, n = 3, Lab 10spot, spot analysis with a spot size of 80 µm) respectively. These mean LA-MC-ICPMS values agree well with TIMS and solution MC-ICP-MS analyses. The repeatability (2s) of 0.00005, 0.00002, and 0.00001 is low and shows homogeneity for the Sr isotope ratios at the spot sizes used.

Pb isotope ratios – information values

The results of the Pb isotope ratios analysed at GEOMAR can be found in Table 4. Replicate analysis of NFHS-2-NP by means of a second digests are within 2s of NIST SRM 981 and prove high precision and indicate homogeneity of NFHS-2-NP at a sample size of 100mg. These data are useful to perform matrix-matched external mass bias correction during LA-(MC)-ICP-MS for Pb isotope ratios using these values from bulk analysis of NFHS-2-NP.

Table 4: Results of Pb isotope ratio measurements of NFHS-2 NP by Thermo Scientific® TRITON Plus TIMS at GEOMAR (Lab 9A) along with reference material USGS BCR-2.

Ratio	NFHS-2-NP digest A unleached		NFHS-2-NP digest B unleached		BCR-2 2M HCl @ 70°C for 1hr	
	Mean	2SE	Mean	2SE	Mean	2SE
$^{206}Pb/^{204}Pb$	17.8572	0.0007	17.8584	0.0009	18.8033	0.0005
$^{207}Pb/^{204}Pb$	15.5907	0.0006	15.5911	0.0007	15.6242	0.0005
$^{208}Pb/^{204}Pb$	37.7701	0.0016	37.7706	0.0019	38.8271	0.0016
$^{207}Pb/^{206}Pb$	0.873075	0.000009	0.873040	0.000011	0.830931	0.000008
$^{208}Pb/^{206}Pb$	2.115121	0.000021	2.115008	0.000021	2.064913	0.000043

$\delta^{13}C$ and $\delta^{18}O$ – information values

The $\delta^{13}C$ and $\delta^{18}O$ values were determined by five laboratories and are shown in Table S1. The number of individual analyses per laboratory ranged from 5-90. Figure 9 shows for each laboratory the results of the mean and repeatability (1s). The mean of means of all laboratories with an expanded reproducibility (95% CL) of $\delta^{13}C_{VPDB}$ and $\delta^{18}O_{VPDB}$ is 0.79±0.04 ‰ and -0.04±0.05 ‰, respectively. The significantly lower $\delta^{18}O$ values (2nd Table of S1) of NFHS-2-NP stored at the institutes for longer than one year (discussed below) were excluded from the mean of means.

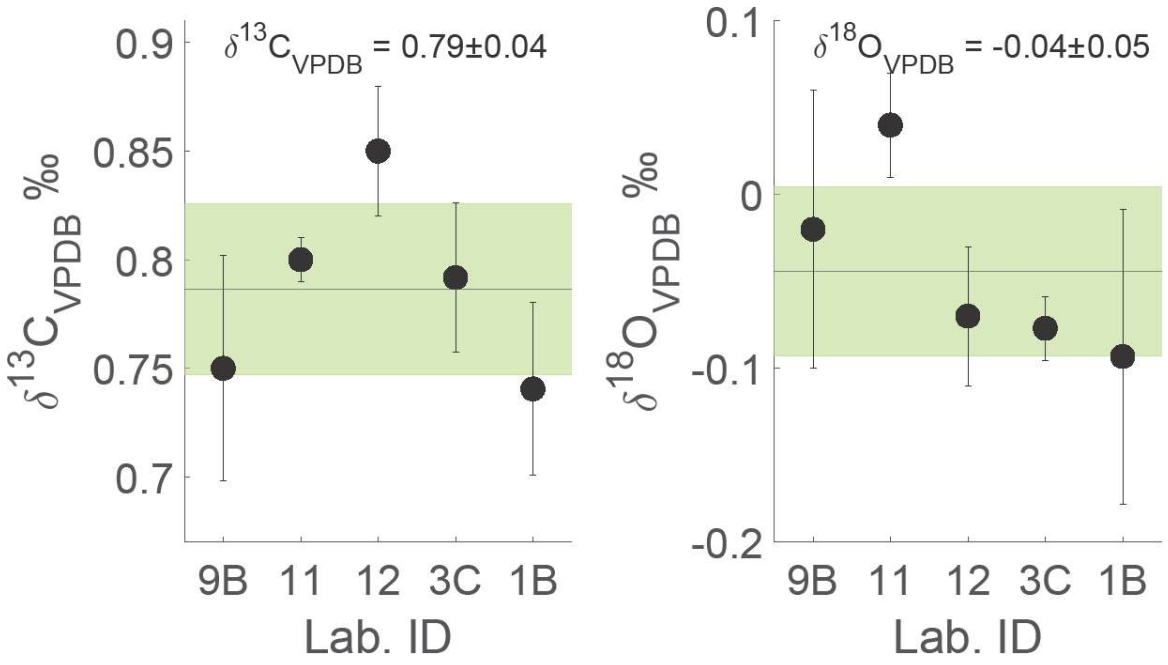


Figure 9: $\delta^{13}C$ and $\delta^{18}O$ values obtained for NFHS-2-NP against laboratory ID. The horizontal line represents the mean of means *derived from five participants. The green band represents the expanded reproducibility (eq. 3). The individual error bars reflect the repeatability (1s) of each laboratory.*

After longer than one year of storage of NFHS-2-NP in a desiccator or sealed bag, analyses were repeated by NIOZ, AWI and MARUM using the original sample aliquots and storage bottles. This showed that the original $\delta^{18}O$ value of -0.04±0.05 ‰ drifted to -0.25, -0.39, and -0.44 ‰ measured by NIOZ, MARUM and AWI, respectively (2nd Table S1, data are not included in Fig. 9). The QCMs measured by these three laboratories showed high accuracies (Table S1) for all measurements excluding that this offset was caused by instrumental issues. Moreover, measurement by MARUM of a second batch, stored at myStandards GmbH for two years, reproduced the original data of MARUM from the first batch (Table S1, batch 2, not included in the mean of means), confirming the long-term $\delta^{18}O$ drift of the first batch stored at MARUM. A cross calibration between MARUM and NIOZ by exchanging vials confirmed (a) that the $\delta^{18}O$ value of MARUM's batch drifted to lower values (-0.39±0.05 ‰ (n=8) measured by MARUM and -0.50±0.07 (n=6)

measured by NIOZ), (b) that the $\delta^{18}O$ value of NIOZ's batch also drifted to lower values (-0.24±0.13 ‰ (n=2) measured by MARUM and -0.25±0.07 (n=6) measured by NIOZ).

A reason for the observed offset could be contact with humid air during the weighing step in the laboratory in combination with the high specific surface area of NFHS-2-NP. NIOZ tested the effect of exposing NFHS-2-NP to moisture for a duration of 0 (t=0) to 36 days (t=36, ~ 5 weeks). Results show that such exposure does not alter the $\delta^{13}C$ or $\delta^{18}O$ values at NIOZ (Fig. 10). Over the duration of the experiment the values of the dry powder (t= 0) of both stable isotopes remained constant, within the repeatability of the measurement. This does not exclude resetting of isotopes when stored over longer time periods in the presence of moisture and hence we still recommend storing the material in a desiccator, as is common practice for RMs. However, we concluded from this experiment that contact with humid air ,during a time span of five weeks, was not the reason for the observed offset.

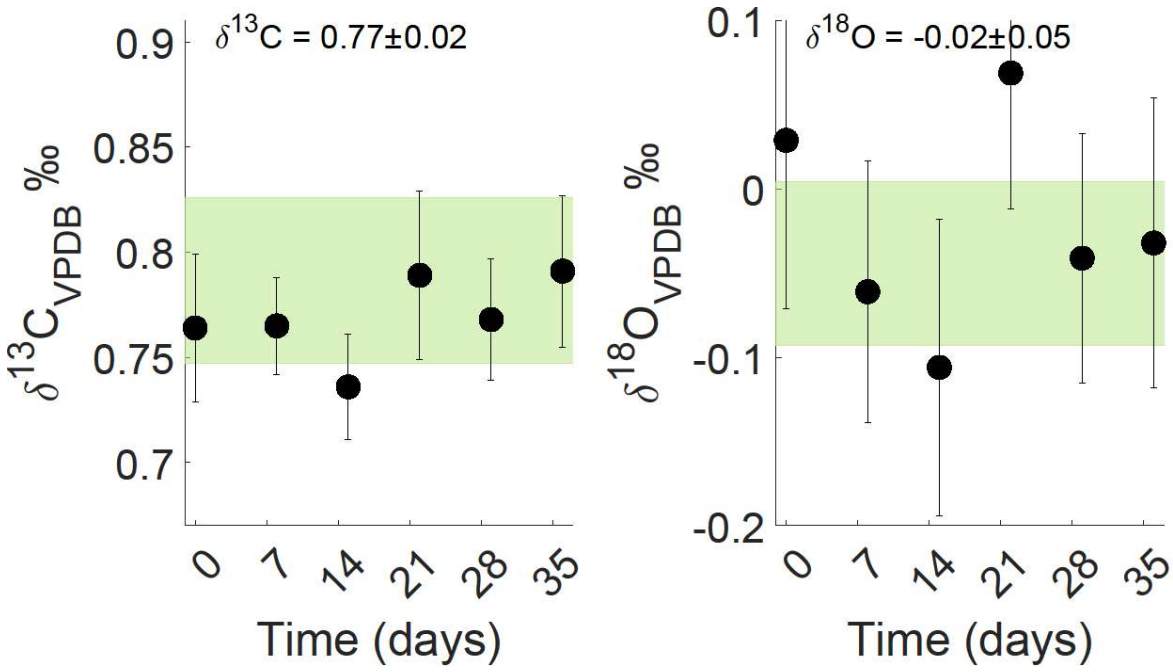


Figure 10: $\delta^{13}C$ and $\delta^{18}O$ against time in days to test for potential impact of moisture absorption. NFHS-2-NP was analysed 5 or 6 times per batch by NIOZ. The green band represents the expanded reproducibility (eq. 3) derived from the five participants.

NFHS-2NP was made by nano milling NFHS-1. NFHS1 is coarser grained and used at NIOZ as an in-house standard for stable isotopes. Information values of the mean of means of this RM analysed by Cardiff University, University of Utrecht, UU and NIOZ for $\delta^{13}C_{VPDB}$ and $\delta^{18}O_{VPDB}$ are 0.84±0.02 ‰ and 1.11±0.06 ‰. The $\delta^{18}O$ values in NFHS1 are stable since 1994. Interestingly, the $\delta^{18}O$ of NFHS-2-NP, thus after nano-milling, is -0.04±0.05 ‰ and significantly lower than $\delta^{18}O$ of NFHS1 of 1.11±0.06 ‰ before milling. Mass fractions of NFHS1 (not shown) were measured at NIOZ and showed an almost identical composition as NFHS-2-NP, however NFHS-2-NP is finer grained and contains more SiO_2 than NFHS1. This SiO_2 cannot cause a change in the $\delta^{18}O$ values because no CO_2 is produced by dissolution with phosphoric acid in the automated carbonate preparation device. Therefore, we have no explanation for the observed difference between the $\delta^{18}O$ values of NFHS1 and NFHS-2-NP. Research is needed to investigate how the nano milling procedure changed the $\delta^{18}O$ value. Research is needed to check whether and how time changed the $\delta^{18}O$ value at AWI, MARUM and NIOZ while the $\delta^{18}O$ value of NFHS-2-NP stored at myStandards GmbH remained stable.

The pellets of NFHS-2-NP with information values for $\delta^{13}C_{VPDB}$ and $\delta^{18}O_{VPDB}$ of 0.79±0.04‰ and -0.04±0.05‰, respectively, are intended to be used for SIMS analysis. However, more testing regarding pellet stability under high-vacuum conditions and ion beam is required before we are confident to distribute it for this purpose. Thus, stable isotopes by SIMS measurements remains to be tested. For the SIMS application, we hope that our $\delta^{13}C$ and $\delta^{18}O$ information values combined with in-situ mass fraction and Sr isotope ratios will be useful.

NFHS-2-NP as a matrix matched calibration RM

At NIOZ, NFHS-2-NP was firstly measured as an unknown material with LA-ICP-MS, together with four different NP QCMs and USGS MACS-3, using NIST SRM 610 as a calibration RM. After reference values of the new RM were established in the context of this study, NFHS-2-NP was used as a calibration RM to obtain data for the QCMs JCp-1 NP, JCt-1 NP, and ECRM752 NP which are close in composition to NFHS-2-NP to test its applicability as calibration material when analysing carbonate samples. It should be noted that the mass fraction of Sr in JCp-1 NP is higher than in NFHS-2-NP and NIST SRM 610, thus Sr data can only be obtained by extrapolation which can lead to inaccuracies. Sr mass fraction of JCp-1 NP is more similar to USGS MACS-3 or MACS-3 NP. For Sr mass fractions in aragonite marine carbonates, USGS MACS-3 or MACS-3 NP could be a better calibration RM. Table 5 shows the mass fraction of those elements measured at NIOZ and where reference values of NFHS-2-NP are known. It shows that data of all elements using NFHS-2-NP as a calibration RM agrees, within uncertainty, with data using NIST SRM 610 as a calibration RM. NFHS-2-NP could therefore be an alternative for or in combination with e.g., NIST SRM 610 and/or NIST SRM 612 to calibrate LA-ICP-MS measurements on carbonate samples. For some matrix sensitive elements (e.g. Na, Pb, Zn) in natural calcite samples, it could be a better alternative than silicate glass RMs.

Table 5: data of JCp-1 and JCt-1 using NFHS-2-NP as a calibration RM.

QCM:	JCp-1 NP				JCt-1 NP				ECRM752 NP			
CAL RM:	NFHS-2-NP		NIST SRM 610		NFHS-2-NP		NIST SRM 610		NFHS-2-NP		NIST SRM 610	
Element	Mean	1s	Mean	1s	Mean	1s	Mean	1s	Mean	1s	Mean	1s
B	42	4	45	5	18	3	19	3	2.1	0.1	2.3	0.1
Ba	12	2	12	1	10	3	10	3	62	5	60	5
Mg	858	93	849	92	271	13	268	13	890	11	882	11
Mn	1.09	0.06	1.01	0.06	0.48	0.07	0.45	0.07	83	2	77	1
Na	4146	469	4184	473	3872	186	3908	188	49	5	49	5
Pb	0.25	0.02	0.25	0.02	0.14	0.02	0.14	0.02	1.5	0.1	1.5	0.1
Sr	7477	76	7418	76	1454	20	1443	20	160	1	159	1
Zn	0.9	0.1	0.9	0.1	0.79	0.08	0.76	0.07	4.9	0.6	4.8	0.6

Compared to non-NP RMs such as USGS MACS-3, NFHS-2-NP could be a better alternative for the analysis of natural calcite samples. Firstly, because of the discussed improved homogeneity. Secondly, because the improved cohesion of NPs is beneficial for laser ablation analysis as already pointed out by Garbe-Schönberg and Müller (2014) who highlighted the smooth surfaces of the tablet, and that crater walls and bottom are similar to ablation pits in glass. Therefore, for mass fractions analyses, the ablation behavior should be quite similar to a solid sample including carbonates. The calcite RM NFHS-2-NP could therefore have similar ablation rates and crater aspect ratio as natural crystalline calcite material (e.g. foraminifera). If there is still a difference in ablation rate between a natural sample and RM, the internal standard Ca could correct for a bias.

However, for isotopic analysis, Guillong et al. (2020) suggested that carbonate U–Pb dating by LA-ICP-MS is only accurate when the laser ablation crater aspect ratio is similar for RM and an unknown sample. The authors concluded that the ablation rates in natural samples of dolomite and aragonite is 1.6 and 2 times larger than in calcite causing different crater aspect ratios. In this study we did not analyse U-Pb, however $^{87}Sr/^{86}Sr$ was studied in detail. Table 6 highlights the QCMs described at the LA-MC-ICP-MS method section of Modena and JGU.

Table 6: Measured and reference $^{87}Sr/^{86}Sr$ ratios of six QCMs analysed at JGU and Modena

		Measured $^{87}Sr/^{86}Sr$			Reference $^{87}Sr/^{86}Sr$		
Lab ID	RM	Mean	1s	n	Mean	1s	Comment
JGU (Lab 7)	JCt-1	0.70917	0.00003	15	0.70916		Weber et al., 2018, GGR
	JCp-1	0.70916	0.00002	6	0.709164		Weber et al., 2018, GGR
	AGLOX	0.70992	0.00012	6	0.70999	0.00006	In-house RM, african elephant molar, Weber et al., 2021, Chemical Geology
	Shark	0.70919	0.00002	9	0.70918		Modern seawater value (shark teeth)
Modena (Lab 10)	modern seashell	0.70918	0.00002	5	0.70917		Modern seawater value
	NanoSr	0.70753	0.00001	2	0.70756	0.00003	Weber et al., 2020, GGR

All resulting $^{87}Sr/^{86}Sr$ ratios agree, within uncertainties, with their respective literature value derived from solution MC-ICP-MS/TIMS or modern seawater values. The high accuracies of the QCM measurements as shown in table 6 together with our NFHS-2-NP results (Fig. 8 and Table S5) where LA-MC-ICP-MS values

agrees with TIMS and solution MC-ICP-MS analyses, leads us to the conclusion that for this kind of carbonate samples matrix matching using crystalline RMs is not crucial for in situ Sr isotope measurements.

While the use of Nano-Pellets in LA-ICP-MS has been increasing in the recent years it remains to be shown, that the matrix-matching is true and how ablation behavior affects other isotope systems and determined ratios (e.g. $\delta^{11}B$). The isotopes ^{10}B and ^{11}B have a relatively larger mass difference than ^{87}Sr and ^{86}Sr. $\delta^{11}B$ could therefore, or for other unknown reasons, be more sensitive to fractionation. NFHS-2-NP could therefore become an important matrix matched calibration RM for $\delta^{11}B$ analysis of natural calcite samples using in-situ techniques.

Conclusions

We present a new carbonate reference material NFHS-2-NP, made from shells of fossil planktic foraminifera that were ground using a nano milling technique and pressed into pellets without any binder. Twelve independent laboratories analysed this new RM (powder and pellets) applying various techniques. The obtained elemental and isotopic data confirm the homogeneity of NFHS-2-NP and therefore its usability for calibration or monitoring the quality of in-situ measurements by techniques such as LA-(MC)-ICP-MS, and µXRF.

We provide reference and information values for 46 elemental mass fractions in a Product information sheet (digital pdf supplement). For $^{87}Sr/^{86}Sr$ we determined a reference value with expanded reproducibility of 0.709163 ± 0.000008. Information values for $\delta^{13}C_{VPDB}$ and $\delta^{18}O_{VPDB}$ are 0.79±0.04 ‰

and -0.04±0.05 ‰ (95% CL), respectively. Regarding Pb isotope ratios we established information values, i.e., $^{206}Pb/^{204}Pb$ = 17.858, $^{207}Pb/^{204}Pb$ = 15.591, $^{208}Pb/^{204}Pb$ = 37.770, $^{207}Pb/^{206}Pb$ = 0.87306, and $^{208}Pb/^{206}Pb$ = 2.11506.

To use NFHS-2-NP as a RM for microanalytical techniques homogeneity at micrometer-scale is essential. LA-ICP-MS results confirm that the relative standard deviation of several elements relevant for climate research (e.g., B, Na, Mg, Mn, Zn, Sr, Ba, and Pb) is <3 % at spot sizes of 60 and 70 µm, indicating a high degree of homogeneity similar to NIST SRM 61x glasses and JCp-1 NP and JCt-1 NP measured at a spot size of 55µm (Jochum et al. 2019). The LA-ICP-MS values agree within reproducibility, with the solution-based reference values for the monitored elements.

In addition to the element mass fractions, $^{87}Sr/^{86}Sr$ ratios determined by LA-MC-ICP-MS agree with solution-based MC-ICP-MS and TIMS analyses and display low 2s values of 0.00005 (n=21, 70 µm spot size) and 0.00002 (n=12, 100 µm spot size), showing that micro homogeneity of the nano-pellet is indeed ensured for this RM.

Acknowledgments

We thank the cruise leader Prof. G.J. Brummer (NIOZ) for providing piston core PE174-15PC from Walvis Ridge. Dr. Inge van Dijk (NIOZ) is thanked for help during the sieving and cleaning of the foraminifera and Patrick Laan (NIOZ) for the help with the SF-ICP-MS analysis at the NIOZ. Pellets are available on request from myStandards GmbH (info@my-standards.com).

References

ASTM E826-14, Standard Practice for Testing Homogeneity of a Metal Lot or Batch in Solid Form by Spark Atomic Emission Spectrometry, ASTM International, West Conshohocken, PA, 2014, www.astm.org.

Barker, S., et al. (2003). "A study of cleaning procedures used for foraminiferal Mg/Ca paleothermometry." Geochemistry Geophysics Geosystems 4(9): 2003GC000559

Bougeois, L., De Rafelis, M., Reichart, G.-J., De Nooijer, L.J., Nicollin, F., Dupont-NIvet, G. (2014). "A high resolution study of trace elements and stable isotopes in oyster shells to estimate Central Asian Middle Eocene seasonality." Chemical Geology 363: 200-212.

Berglund M. and Wieser M.E. (2011). "Isotopic compositions of the elements 2009 (IUPAC Technical Report)". Pure and Applied Chemistry, 83: 397–410.

Burns, S.J. et al. (1998). "Speleotherm-based paleoclimate record from Northern Oman." Geology 26(6): 499-502.

De Villiers S., Greaves M., Elderfield H. (2002). "An intensity ratio calibration method for the accurate determination of Mg/Ca and Sr/Ca of marine carbonates by ICP-AES." Geochemistry Geophysics, Geosystems, 3: 2001GC000169

Durand, A., et al. (2016). "Improved methodology for the microwave digestion of carbonate-rich environmental samples." International Journal of Environmental Analytical Chemistry,96(2): 119-136.

Eggins, S.M., Kinsley, L.K., Shelley JMG, J.M.G. (1998). Deposition and element fractionation processes occurring during atmospheric pressure laser sampling for analysis by ICPMS. Appl. Surf. Sci. 127–129, 278–286.

Fourny, A., Weis, D., and Scoates, J. S. (2016) Comprehensive Pb-Sr-Nd-Hf isotopic, trace element, and mineralogical characterization of mafic to ultramafic rock reference materials: Geochemistry, Geophysics, Geosystems, 17(3): 739-773.

Hoernle K., Hauff F., Kokfelt T.F., Haase K., Garbe-Schönberg C.-D. and Werner R. (2011) On- and off-axis chemical heterogeneities along the South Atlantic Mid-Ocean Ridge (5-11°S): Shallow or deep recycling of ocean crust and/or intraplate volcanism? Earth Planetary Sci. Lett. 306: 86-97.

Garbe-Schönberg, D. and S. Müller (2014). "Nano-particulate pressed powder tablets for LA-ICP-MS." J. Anal. At. Spectrom. 29(6): 990-1000.

Greaves, M., et al. (2005). "Accuracy, standardization, and interlaboratory calibration standards for foraminiferal Mg/Ca thermometry." Geochemistry, Geophysics, Geosystems, 6(2): 2004GC000790

Greaves, M., et al. (2008). "Interlaboratory comparison study of calibration standards for foraminiferal Mg/Ca thermometry." Geochemistry Geophysics Geosystems 9(8): 1-27.

Guillong, M., Meier, D. L., Allan, M. M., Heinrich, C. A., and Yardley, B. W. (2008): SILLS: A MATLAB-based program for the reduction of laser ablation ICP-MS data of homogeneous materials and inclusions. In P.J. Sylvester (ed), *Laser ablation ICP-MS in the Earth Sciences: Current practices and outstanding issues.* Mineralogical Association of Canada Short Course Series, 40, 328333. Publisher Mineralogical Association of Canada.

Guillong, Marcel, et al. (2020). "Evaluating the reliability of U-Pb laser ablation inductively coupled plasma mass spectrometry (LA-ICP-MS) carbonate geochronology: matrix issues and a potential calcite validation reference material." Geochronology, 2(1): 155.

Hathorne, E.C., Thomas Felis, Rachael H. James, Alex Thomas (2011), “Laser ablation ICP-MS screening of corals for diagenetically affected areas applied to Tahiti corals from the last deglaciation”, Geochimica et Cosmochimica Acta,75(6): 1490-1506.

Hathorne, E. C., et al. (2013). "Interlaboratory study for coral Sr/Ca and other element/Ca ratio measurements." Geochemistry Geophysics Geosystems, 14(9): 3730-3750.

Henehan, M. J., Foster, G. L., Rae, J. W. B., Prentice, K. C., Erez, J., Bostock, H. C., Marshall, B. J., and Wilson, P. A. (2015), "Evaluating the utility of B/Ca ratios in planktic foraminifera as a proxy for the carbonate system: A case study of Globigerinoides ruber", Geochem. Geophys. Geosyst., 16, 1052– 1069,

Hoernle, K., Abt, D.L., Fischer, K.M., Nichols, H., Hauff, F., Abers, G.A., van den Bogaard,P., Heydolph, K., Alvarado, G., Protti, M. and Strauch, W. (2008) Arc-parallel flow in the mantle wedge beneath Costa Rica and Nicaragua. Nature 451, 1094-1097.

Hoogakker, B.A.A., Klinkhammer, G.P., Elderfield, H., Rohling, E.J., Hayward, C. (2009). "Mg/Ca paleothermometry in high salinity environments." Earth and Planetary Science Letters, 284: 583-589.

Horwitz, W., and R. Albert (1995). "Precision in analytical measurements: Expected values and consequences in geochemical analyses." Fresenius Journal of Analytical Chemistry, 351: 507-513.

Inoue, M., et al. (2004). "Concentrations of Trace Elements in Carbonate Reference Materials Coral JCp-1 and Giant Clam JCt-1 by Inductively Coupled Plasma-Mass Spectrometry." Geostandards and Geoanalytical Research, 28(3): 411-416.

ISO Guide 35:2017. Reference materials-Guidance for characterization and assessment of homogeneity and stability. International Organisation for Standardisation (ISO), Geneva, 2017.

ISO 17034:2017. General requirements for the competence of reference material producers. International Organisation for Standardisation (ISO), Geneva, 2017.

ISO 17025:2018. General requirements for the competence of testing and calibration laboratories. International Organisation for Standardisation (ISO), Geneva, 2018.

Jochum, K.P., et al. (2005). GeoReM: a new geochemical database for reference materials and isotopic standards. Geostand. Geoanal. Res. 29, 333–338. doi:10.1111/j.1751–908X.2005.tb00904.x

Jochum, K.P., et al. (2007). Validation of LA-ICP-MS trace element analysis of geological glasses using a new solid-state 193 nm Nd:YAG laser and matrix-matched calibration. J. Anal. At. Spectrom. 22, 112. doi:10.1039/b609547j

Jochum, K. P., et al. (2011). "Determination of Reference Values for NIST SRM 610-617 Glasses Following ISO Guidelines." Geostandards and Geoanalytical Research, 35(4): 397-429.

Jochum, K. P., et al. (2012). "Accurate trace element analysis of speleothems and biogenic calcium carbonates by LA-ICP-MS." Chemical Geology, 318-319: 31-44.

Jochum, K. P., et al. (2019). "Nano-Powdered Calcium Carbonate Reference Materials: Significant Progress for Microanalysis?" Geostandards and Geoanalytical Research, 43(4): 595-609.

Johnstone, H. J. H., et al. (2016). "Automated cleaning of foraminifera shells before Mg/Ca analysis using a pipette robot." Geochemistry Geophysics Geosystems, 420: 3502-3511.

McArthur, J., Howarth, R. J., and Shields, G. A., 2012. Strontium isotope stratigraphy. In The Geological Time Scale. Elsevier, Chap. 7: 127–144.

Marchitto, T. M. (2006), “Precise multielemental ratios in small foraminiferal samples determined by sector field ICP-MS”, Geochem. Geophys. Geosyst., 7: 1-10.

Mischel, S. A., Mertz-Kraus, R., Jochum, K. P., and Scholz, D. (2017). ”TERMITE: An R script for fast reduction of laser ablation inductively coupled plasma mass spectrometry data and its application to trace element measurements”. Rapid Commun. Mass Spectrom., 31: 1079– 1087.

Lončarić, N., Jolanda van Iperen, Dick Kroon, Geert-Jan A. Brummer (2007), “Seasonal export and sediment preservation of diatomaceous, foraminiferal and organic matter mass fluxes in a trophic gradient across the SE Atlantic”, Progress in Oceanography, 73(1): 27-59.

Lugli, F., Weber, M., Giovanardi, T., Arrighi, S., Bortolini, E., Figus, C., ... & Cipriani, A. (2020). Fast offline data reduction of laser ablation MC-ICP-MS Sr isotope measurements via an interactive Excel-based spreadsheet ‘SrDR’. Journal of Analytical Atomic Spectrometry, 35(5): 852-862.

Nürnberg, D., Bijma, J., Hemleben, C. (1996). “Assessing the reliability of magnesium in foraminiferal calcite as a proxy for water mass temperatures. Geochmica et Cosmochimica Acta 60(5): 803-814.

Okai, T., et al. (2002). "Preparation of a New Geological Survey of Japan, Geochemical Reference Material: Coral JCp-1." Geostandards Newsletter.

Palmer, M. R., & Edmond, J. M. (1992). “Controls over the strontium isotope composition of river water”. Geochimica et Cosmochimica Acta, 56(5): 2099-2111.

Pracht, H., Metcalfe, B., Peeters, F.J.C. (2019). “Oxygen isotope composition of the final chamber of planktic foraminifera provides evidence of vertical migration and depth-integrated growth.” Biogeosciences, 16: 643-661.

R Core Team (2020). R: A Language and Environment for Statistical Computing.
Vienna: R Foundation for Statistical Computing. Available online at: https://
www.R-project.org/

Swart, P.K., Burns, S.J. and Leder, J.J. (1991). Fractionation of the stable isotopes of oxygen and carbon in carbon dioxide during the reaction of calcite with phosphoric acid as a function of temperature and technique. Chem. Geol. 86: 89-96.

Sylvester P.J. (2008). Matrix Effects in Laser Ablation-ICP-MS. In: Sylvester, P.J. (ed), *Laser Ablation ICP-MS in the Earth Sciences: Current Practices and Outstanding Issues*. Mineralogical Association of Canada, Short Course Series, 40: 67-78.Publisher Mineralogical Association of Canada.

Tang, G. Q., et al. (2019). "A new Chinese national reference material (GBW04481) for calcite oxygen and carbon isotopic microanalysis." Surface and Interface Analysis 52(5): 190-196.

Tarique, M., Rahaman, W., Fousiya, A.A., Lathika, N., Thamban, M., Achyuthan, H., Misra, S. (2021). "Surface pH record (1990-2013) of the Arabian Sea from boron isotopes of Lakshadweep corals – Trend, variability and control." Journal of Geophysical Research: Biogeosciences 126(7).

Taylor R, Ishizuka O, Michalik A, Milton J, Croudace I (2014). "Evaluating the precision of Pb isotope measurement by mass spectrometry". J. Anal. At. Spectrom. 30.

Titelboim, D., Sadekov, A., Hyams-Kaphzan, O., Almogi-Labin, A., Herut, B., Kucera, M., Abramovich, S. (2018). "Foraminiferal single chamber analyses of heavy metals as a tool for monitoring permanent and short term anthropogenic footprints." Marine Pollution Bulletin 126: 65-71.

Todd, E., Stracke, A., and Scherer, E. E. (2015) Effects of simple acid leaching of crushed and powdered geological materials on high-precision Pb isotope analyses: Geochemistry, Geophysics, Geosystems, 16(7): 2276-2302.

Weber, M., et al. (2018). "Calcium Carbonate and Phosphate Reference Materials for Monitoring Bulk and Microanalytical Determination of Sr Isotopes." Geostandards and Geoanalytical Research 42(1): 77-89.

Weber, M., et al. (2020). "NanoSr – A New Carbonate Microanalytical Reference Material for

In Situ Strontium Isotope Analysis." Geostandards and Geoanalytical Research 44(1): 69-83.

Weber, M., et al. (2020). "Strontium uptake and intra-population 87Sr/86Sr variability of bones and teeth—controlled feeding experiments with rodents (Rattus norvegicus, Cavia porcellus)." *Frontiers in Ecology and Evolution* 8: 569940.

Weber, K. et al. (2021). “Diagenetic stability of non-traditional stable isotope systems (Ca, Sr, Mg, Zn) in teeth – An in-vitro alteration experiment of biogenic apatite in isotopically enriched tracer solution”. Chemical Geology, 572: 120196.

Weber, M.l, et al. (2021) "Opposite Trends in Holocene Speleothem Proxy Records From Two Neighboring Caves in Germany: A Multi-Proxy Evaluation." Frontiers in Earth Science 9: 154.

Yu, J., Day, J., Greaves, M. & Elderfield, H. (2005). “Determination of multiple element/calcium ratios in foraminiferal calcite by quadrupole ICP-MS”. Geochem. Geophys. Geosyst, 6(8): 2005GC000964.

Zachos, J., Pagani, M., Sloan, L., Thomas, E., Billups, K. (2001). “Trends, rythms, and abberations in global climate 65 Ma to present.” Science 292: 686-693.

Supporting information

The following supporting information can be found in the online version of this article:

Table S1. $\delta^{13}C_{VPDB}$ *and* $\delta^{18}O_{VPDB}$ *of NFHS-2-NP obtained from five laboratories (Lab 1B, 3C, 9B, 11, and 12).*

Table S2. *Mass fractions of NFHS-2-NP obtained from seven laboratories (Lab 1-7).*

Table S3: *Comparison total digestion and* HNO_3 *dissolution by NIOZ (Lab 3A).*

Table S4. *Mass fractions and RSD values obtained from analyses NFHS-2-NP measured by NIOZ (Lab 3B) and JGU (Lab 7A) compared with USGS MACS-3, MACS-3 NP, JCp-1 NP, and JCt-1 NP (Jochum et al. 2019).*

Table S5. $^{87}Sr/^{86}Sr$ *isotopes ratios obtained from four laboratories (Lab 7B, 7C, 8, 9A, 10).*

Table S6. *Individual* $^{87}Sr/^{86}Sr$ *isotopes ratios obtained from four laboratories (Lab 7B, 7C, 8, 9A, 10).*

Table S7. *Example of the B/Ca ratio [mmol/mol] results from a homogeneity test on a single Nano-Pellet.It shows the mean, standard deviation (1s), and RSD% both within and between the seven analytical Zones (Fig. 2). The within-zone evaluation is in the rows below Run 1-3, the between-zone evaluation is in the columns to the right of Zone 1-7.*

Table S8*: Example of a homogeneity test with calculations according to ASTM Guide E826-14. LA-ICP-MS B/Ca ratio [mmol/mol] data for a single Nano-Pellet are shown. The highlighted box in green shows a passed homogeneity test. The test passes if the absolute difference of t' is < the w value.*

Supplement 1: Data evaluation of pellet homogeneity

Homogeneity testing of solid samples using LA-ICP-MS

The homogeneity testing of solid candidate reference materials is less straightforward than that of powdered materials. ASTM-Guide E826-14 (https://www.astm.org/Standards/E826.htm, citation format at bottom of page) provides guidance on this issue. The guide states, that the sample surface should be divided into zones in which repeated measurements should be performed. The problem is, that the measurements are supposed to be performed on exactly the same spot. This is not possible for typical LA-ICP-MS analyses. An adaptation of the guide's approach is shown in Fig. 2 (material and method section). The circle in Fig. 2 represents the pellet's sample surface and shows its subdivision into seven analytical zones in which three spot-measurements in the closest feasible proximity to one another should be performed. In total this yields 21 analyses. Each run should be bracketed by a suitable reference material for drift correction, ensuring that any observed variation can be linked to sample composition and not instrumental error. Additionally, the analyses should be carried out in random order by generating 3 randomised sequences from 1 to 7. Ideally, this procedure should be carried out on several pellets. Traditionally, and following ISO guidelines the number of units required for a homogeneity test can be calculated using eq. 7.

$$N° \; of \; units_{Homogeneity \; Study} = \sqrt[3]{Number \; of \; produced \; units} \qquad (7)$$

An example of B/Ca ratio [mmol/mol] results from a homogeneity test on a single Nano-Pellet is shown in Table S7. It shows the averages and RSD both within and between the zones. Another example of

calculations according to ASTM Guide E826-14 are shown in Table S8. The highlighted box in green shows a passed homogeneity test. The test passes if the absolute difference of t' is < the w value.