



**HAL**  
open science

## **InterCarb: A community effort to improve inter-laboratory standardization of the carbonate clumped isotope thermometer using carbonate standards**

Stefano M Bernasconi, Mathieu Daëron, Kristin D Bergmann, Magali Bonifacie, Anna Nele Meckler, Hagit Affek, Noah Anderson, David Bajnai, Eugeni Barkan, Emily Beverly, et al.

### ► To cite this version:

Stefano M Bernasconi, Mathieu Daëron, Kristin D Bergmann, Magali Bonifacie, Anna Nele Meckler, et al.. InterCarb: A community effort to improve inter-laboratory standardization of the carbonate clumped isotope thermometer using carbonate standards. 2021. hal-03094122

**HAL Id: hal-03094122**

**<https://hal.science/hal-03094122>**

Preprint submitted on 4 Jan 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **InterCarb: A community effort to improve inter-laboratory**  
 2 **standardization of the carbonate clumped isotope**  
 3 **thermometer using carbonate standards**

4  
 5 S.M. Bernasconi<sup>1</sup>, M. Daëron<sup>2</sup>, K.D. Bergmann<sup>3</sup>, M. Bonifacie<sup>4</sup>, A.N. Meckler<sup>5</sup>, H.P. Affek<sup>6</sup>, N.  
 6 Anderson<sup>3</sup>, D. Bajnai<sup>7</sup>, E. Barkan<sup>6</sup>, E. Beverly<sup>8†</sup>, D. Blamart<sup>2</sup>, L. Burgener<sup>9</sup>, D. Calmels<sup>4†</sup>, C.  
 7 Chaduteau<sup>4</sup>, M. Clog<sup>10</sup>, B. Davidheiser-Kroll<sup>11</sup>, A. Davies<sup>12††</sup>, F. Dux<sup>13†</sup>, J. Eiler<sup>14</sup>, B. Elliott<sup>15</sup>,  
 8 A. C. Fetrow<sup>11</sup>, J. Fiebig<sup>7</sup>, S. Goldberg<sup>3</sup>, M. Hermoso<sup>4,23</sup>, K.W. Huntington<sup>16</sup>, E. Hyland<sup>9</sup>, M.  
 9 Ingalls<sup>14†</sup>, M. Jaggi<sup>1</sup>, C.M. John<sup>12</sup>, A. B. Jost<sup>3</sup>, S. Katz<sup>8</sup>, J. Kelson<sup>8</sup>, T. Kluge<sup>12†</sup>, I.J. Kocken<sup>17</sup>,  
 10 A. Laskar<sup>18</sup>, T.J. Leutert<sup>5†</sup>, D. Liang<sup>18</sup>, J. Lucarelli<sup>15</sup>, T. J. Mackey<sup>3†</sup>, X. Mangenot<sup>4,14</sup>, N.  
 11 Meinicke<sup>5</sup>, S.E. Modestou<sup>5</sup>, I.A. Müller<sup>17</sup>, S. Murray<sup>19</sup>, A. Neary<sup>8</sup>, N. Packard<sup>8</sup>, B.H. Passey<sup>8</sup>, E.  
 12 Pelletier<sup>8</sup>, S. Petersen<sup>8</sup>, A. Piasecki<sup>5††</sup>, A. Schauer<sup>16</sup>, K.E. Snell<sup>11</sup>, P.K. Swart<sup>20</sup>, A. Tripathi<sup>15</sup>, D.  
 13 Upadhyay<sup>15</sup>, T. Vennemann<sup>21</sup>, I. Winkelstern<sup>8††</sup>, D. Yarian<sup>8</sup>, N. Yoshida<sup>22</sup>, N. Zhang<sup>22</sup>, M.  
 14 Ziegler<sup>17</sup>

15 <sup>1</sup>Geological Institute, ETH Zürich, Sonneggstrasse 5, 8092 Zürich, Switzerland

16 <sup>2</sup>Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ,  
 17 Université Paris-Saclay, Orme des Merisiers, F-91191 Gif-sur- Yvette, France

18 <sup>3</sup>Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of  
 19 Technology, Cambridge, MA 02139, USA

20 <sup>3†</sup>now at Earth and Planetary Sciences, University of New Mexico, Albuquerque, New Mexico,  
 21 USA

22 <sup>4</sup>Université de Paris, Institut de Physique du Globe de Paris, CNRS, F-75005 Paris, France

23 <sup>4†</sup>now at Geosciences Paris Sud (GEOPS), Université Paris-Saclay, CNRS, 91405 Orsay, France

24 <sup>5</sup>Bjerknes Centre for Climate Research and Department of Earth Science, University of Bergen,  
 25 Bergen, Norway

26 <sup>5†</sup>now at Max Planck Institute for Chemistry, Mainz, Germany; thomas.leutert@mpic.de

27 <sup>5††</sup>now at Department of Earth Sciences, Dartmouth College, Hanover, NH, USA

28 <sup>6</sup>Institute of Earth Sciences, Hebrew University of Jerusalem, Jerusalem, Israel

29 <sup>7</sup>Institute of Geosciences, Goethe University Frankfurt, Frankfurt am Main, Germany

30 <sup>8</sup>Department of Earth and Environmental Sciences, University of Michigan, Ann Arbor,  
 31 Michigan, USA

32 <sup>8†</sup>now at Department of Earth and Atmospheric Sciences, University of Houston, Houston,  
 33 Texas, USA

34 <sup>8††</sup>now at Geology Department, Grand Valley State University, Allendale, Michigan, USA

35 <sup>9</sup>Marine, Earth and Atmospheric Sciences, North Carolina State University, Raleigh, USA

36 <sup>10</sup>Scottish Universities Environmental Research Centre (SUERC), Glasgow, Scotland

- 37 <sup>11</sup>University of Colorado Boulder, Boulder, Colorado, USA  
 38 <sup>12</sup>Imperial College, London, UK  
 39 <sup>12†</sup>now at Karlsruher Institut für Technologie KIT  
 40 <sup>12††</sup>now at Stockholm University  
 41 <sup>13</sup>School of Geography, University of Melbourne, Australia  
 42 <sup>13†</sup>now at School of Earth and Life Sciences, University of Wollongong, Australia  
 43 <sup>14</sup>Geological and Planetary Sciences, California Institute of Technology, Pasadena, California,  
 44 USA  
 45 <sup>14†</sup>now at Department of Geosciences, The Pennsylvania State University, University Park, PA,  
 46 USA  
 47 <sup>15</sup>Department of Earth, Planetary, and Space Sciences, University of California Los Angeles, Los  
 48 Angeles, CA, USA  
 49 <sup>16</sup>University of Washington, Seattle, Washington, USA  
 50 <sup>17</sup>Department of Earth Sciences, University of Utrecht, Utrecht, Netherlands  
 51 <sup>18</sup>Institute of Earth Sciences, Academia Sinica, Taipei, Taiwan  
 52 <sup>19</sup>Macquarie University, Australia  
 53 <sup>20</sup>RSMAS/MGS University of Miami, Miami, FL, USA  
 54 <sup>21</sup>Institute of Earth Surface Dynamics, University of Lausanne, Lausanne, Switzerland  
 55 <sup>22</sup>Earth-Life Science Institute, Tokyo Institute of Technology, Tokyo, Japan  
 56 <sup>23</sup>Univ. Littoral Côte d’Opale, Univ. Lille, CNRS, Laboratoire d’Océanologie et de Géosciences  
 57 (UMR 8187 LOG), 62930 Wimereux, France.  
 58

59 Corresponding authors: Stefano M. Bernasconi (stefano.bernasconi@erdw.ethz.ch) Mathieu  
 60 Daëron (daeron@lsce.ipsl.fr)

61

62 **Key Points:**

- 63           • The exclusive use of carbonate reference materials is a robust method for the  
 64           standardization of clumped isotope measurements  
 65           • Measurements using different acid temperatures, designs of preparation lines, and  
 66           mass spectrometers are statistically indistinguishable  
 67           • We propose new consensus values for a set of 7 carbonate reference materials and  
 68           updated guidelines to report clumped isotope measurements  
 69

70 **Abstract**

71 Increased use and improved methodology of carbonate clumped isotope thermometry has greatly  
72 enhanced our ability to interrogate a suite of Earth-system processes. However, inter-laboratory  
73 discrepancies in quantifying carbonate clumped isotope ( $\Delta_{47}$ ) measurements persist, and their  
74 specific sources remain unclear. To address inter-laboratory differences, we first provide  
75 consensus values from the clumped isotope community for four carbonate standards relative to  
76 heated and equilibrated gases with 1,819 individual analyses from 10 laboratories. Then we  
77 analyzed the four carbonate standards along with three additional standards, spanning a broad  
78 range of  $\delta^{47}$  and  $\Delta_{47}$  compositions, for a total of 5,329 analyses on 25 individual mass  
79 spectrometers from 22 different laboratories. Treating three of the materials as known standards  
80 and the other four as unknowns, we find that the use of carbonate reference materials is a robust  
81 method for standardization that yields inter-laboratory discrepancies entirely consistent with in-  
82 laboratory analytical uncertainties. Carbonate reference materials, along with measurement and  
83 data processing practices described herein, provide the carbonate clumped isotope community  
84 with a robust approach to achieve inter-laboratory agreement as we continue to use and improve  
85 this powerful geochemical tool. We propose that carbonate clumped isotope data normalized to  
86 the carbonate reference materials described in this publication should be reported as  $\Delta_{47}$  (I-  
87 CDES) for Intercarb-Carbon Dioxide Equilibrium Scale.

88

## 89 **1 Introduction**

90 Carbonate clumped isotope ( $\Delta_{47}$ ) thermometry is the most developed branch of the  
91 rapidly evolving field of clumped isotope geochemistry. Given the broad range of applications in  
92 Earth Sciences (e.g. Affek & Eiler, 2006; Eagle et al., 2010; Ferry et al., 2011; Ghosh, Adkins, et  
93 al., 2006; Ghosh, Garzzone, et al., 2006; Grauel et al., 2013; Guo & Eiler, 2007; Huntington et  
94 al., 2011; Passey & Henkes, 2012, Dale et al, 2014, Mangenot et al., 2018, Veillard et al, 2019)  
95 and the improvement of analytical methods including automation (Bernasconi et al., 2013, 2018;  
96 Defliese & Lohmann, 2015; Dennis et al., 2011; Fiebig et al., 2019; Ghosh, Adkins, et al., 2006;  
97 He et al., 2012; Hu et al., 2014; Huntington et al., 2009; Meckler et al., 2014; Müller, Fernandez,  
98 et al., 2017; Passey et al., 2010; Petersen et al., 2019; Petersen et al., 2016; Schmid &  
99 Bernasconi, 2010, Adlan et al, 2020), the last 5-10 years have seen an increasing number of  
100 laboratories implementing this technique. The great potential of this thermometric method can  
101 only be fully exploited if the precision and accuracy of the data is sufficient to resolve  
102 differences of a few degrees in formation temperatures. In addition, like for any other  
103 geochemical method, widely available reference materials that match the sample matrices are  
104 necessary so that data can be robustly compared across laboratories (Meier-Augenstein &  
105 Schimmelmann, 2019). Currently the situation in the field of carbonate clumped isotope  
106 geochemistry is far from satisfactory. Published values for the ETH reference materials, the only  
107 carbonates that have been recently measured in many different laboratories worldwide, show a  
108 range of values for the same sample of up to 0.053 ‰ (see Bernasconi et al. 2018, Thaler et al.  
109 2020 for recent comparisons). This clearly calls for better standardization procedures to improve  
110 laboratory comparability.

111  
112 The data normalization scheme currently used in clumped isotope geochemistry of carbonates in  
113 most laboratories is based on the comparison of the composition of the CO<sub>2</sub> liberated from  
114 carbonates by reaction with phosphoric acid with that of a set of CO<sub>2</sub> gases with different bulk  
115 and clumped isotope compositions (Dennis et al., 2011). These gases are prepared either by  
116 heating CO<sub>2</sub> at 1000 °C (heated gases; HG) or by CO<sub>2</sub> equilibration with water at low  
117 temperatures (equilibrated gases; EG, at *e.g.* 25 °C, 50 °C). By comparing the measured  
118 compositions with the theoretical predictions of the equilibrium thermodynamic abundance of  
119 multiply substituted isotopologues in heated and equilibrated gases (Wang et al. 2004 and

120 updates in Petersen et al. 2019), the measurements are standardized to the scale that was named  
121 the “absolute reference frame” (ARF) by Dennis et al., (2011). In more recent publications the  
122 ARF is often referred to as the “Carbon Dioxide Equilibration Scale” (CDES), a terminology  
123 introduced by Passey and Henkes (2012). This approach was designed to allow different  
124 laboratories to link their measurements to an internationally recognized scale firmly anchored to  
125 theory using relatively easy and established laboratory protocols to produce CO<sub>2</sub> standard gases  
126 of known composition. Early comparisons of  $\Delta_{47}$  CDES values for carbonates run in different  
127 laboratories and normalized with HG/EG normalization were promising (Dennis et al., 2011).  
128 While Bonifacie et al. (2017) reported similar  $\Delta_{47}$  CDES values for nine dolomite samples covering  
129 a range of almost 0.4‰ measured at both at Caltech and IPGP laboratories with HG/EG  
130 normalization, Spooner et al. (2016) found that carbonate standardization improved agreement  
131 between data they obtained on samples ran both at Caltech and WHOI laboratories, compared to  
132 when they were using HG/EG normalization. Such recurrent cases of poor inter-laboratory  
133 reproducibility (see also Bernasconi et al., 2018 and Thaler et al., 2020) suggest that there is still  
134 unexplained differences in the results among laboratories (see Petersen et al. (2019) for a recent  
135 review).

136  
137 Apart from preservation problems such as diagenesis and solid-state reordering, two known  
138 issues still limiting the reliability of this method to yield accurate temperature reconstructions  
139 are: [1] the lack of internationally recognized carbonate reference materials for a precise inter-  
140 laboratory calibration, and [2] that published  $\Delta_{47}$ -temperature calibrations produced in different  
141 laboratories have differed in both temperature dependence (slope), and absolute values  
142 (intercept). Possible reasons for the differences in slope and intercepts of the  $\Delta_{47}$  temperature  
143 dependence have been widely discussed in the literature (e.g. Kluge et al, 2015; Bonifacie et al.,  
144 2017; Daëron et al., 2016; Fernandez et al., 2017; Katz et al., 2017; Kelson et al., 2017; Petersen  
145 et al., 2019; Schauer et al., 2016). Discrepancies have been attributed to analytical artefacts such  
146 as CO<sub>2</sub>-acid re-equilibration at different acid digestion temperatures (see Wacker et al. 2013;  
147 Swart et al., 2019, for a recent discussion). Other factors proposed to influence the calculated  
148 slope of the calibrations are the limitations of the datasets used in the individual studies, in  
149 particular in terms of the number of samples and replicates and of the temperature range covered  
150 by the available samples (Bonifacie et al., 2017; Fernandez et al., 2017). However, the

151 discrepancies in the intercept of the calibrations, for example between Kelson et al. (2017) and  
152 Peral et al. (2018), and a generally poor laboratory comparability remain problems that should be  
153 mitigated by using a more robust standardization method.

154  
155 Petersen et al. (2019), in a recent effort to resolve differences in calibrations, compiled raw data  
156 of a number of published temperature calibrations and recalculated them all in a consistent way  
157 using the revised “IUPAC” correction parameters to correct for the  $^{17}\text{O}$  abundance (Daëron et al.,  
158 2016; Schauer et al., 2016). The goal was to test whether data processing differences and/or the  
159 use of consistent but incorrect  $^{17}\text{O}$ - correction parameters in the calculations were the root causes  
160 of inconsistencies. The result of this study was that differences among calibrations were reduced  
161 but not eliminated by the recalculation, implying that other factors must be responsible for the  
162 remaining discrepancies. These differences have pushed many laboratories to use laboratory-  
163 specific calibrations performed with the same analytical approach, as they at least partially take  
164 into consideration possible procedural differences (Petersen et al., 2019). However, if a  
165 laboratory changes analytical procedures or has not generated a robust in-house calibration, this  
166 approach is problematic. Achieving an inter-laboratory reproducibility at the level of accuracy  
167 necessary for meaningful interpretations of the observed variations is a requirement for  $\Delta_{47}$   
168 thermometry to reach its potential as a mature analytical method with broad acceptance and  
169 quantitative usefulness.

170  
171 While the definition of the CDES was a major milestone (Dennis et al., 2011), a known problem  
172 with this approach is that while the  $\text{CO}_2$  standard gases equilibrated at known temperature (HG  
173 or EG) can be confidently used for correction of mass spectrometric fractionations/nonlinearities  
174 and for effects of the purification procedures, they cannot account for the effects of the  
175 phosphoric acid reaction on the composition of the produced  $\text{CO}_2$ . Among the factors responsible  
176 for discrepant calibrations and laboratory comparability, two important ones cannot be tested  
177 with a gas-based standardization: (1) the absolute value and temperature dependence of the  
178 phosphoric acid fractionation factor (see Petersen et al., 2019 for a recent compilation) and (2)  
179 possible  $\text{CO}_2$  equilibration effects during acid digestion of the sample. Swart et al. (2019)  
180 presented evidence that equilibration of  $\text{CO}_2$  with water or hot metal surfaces during phosphoric  
181 acid reaction and transfer of the  $\text{CO}_2$  to the mass spectrometer could be a factor leading to the

182 alteration of the apparent temperature dependence of clumped isotopes in carbonates and on the  
183 absolute value of calculated  $\Delta_{47}$ . As many laboratories use custom built extraction lines with  
184 different designs and volumes of tubing and of acid vessels, these factors are impossible to  
185 precisely quantify for each laboratory and may further contribute to inter-laboratory  
186 discrepancies.

187  
188 We propose that these issues can be circumvented if carbonates, which undergo the same acid  
189 digestion as the samples, are used for normalization instead of or in addition to gases, consistent  
190 with the principle of identical treatment of sample and standards (Carter & Fry, 2013; Werner &  
191 Brand, 2001a). In addition, normalizing results to accepted carbonate reference material values,  
192 as is commonly done with conventional carbon and oxygen isotope analysis in carbonates,  
193 removes the requirement to precisely quantify acid fractionation factors at different temperatures  
194 (Bernasconi et al., 2018).

195 A carbonate standardization approach was introduced by Schmid & Bernasconi (2010) and  
196 improved by Meckler et al. (2014), with the following benefits: (1) the use of carbonates can  
197 more easily be fully automated, eliminating time-consuming and possibly error-prone manual  
198 preparation of CO<sub>2</sub> standard gases (equilibrated at known temperature) by individual users on  
199 separate extraction lines; (2) in some automated systems designed for small sample sizes, the  
200 gases had to be measured through a different capillary than the carbonates with potential biases  
201 that would go unrecognized; and (3) in those systems the equilibrated and heated gases are  
202 measured at constant beam intensity, whereas the samples are measured with decreasing beams.  
203 These features argue in favor of carbonate standardization *a priori*, but it remains critical to  
204 assess *a posteriori* whether the results of this approach are as robust and accurate as expected,  
205 and whether they significantly improve the inter-laboratory reproducibility of  $\Delta_{47}$  measurements.  
206 Discussions at the Sixth International Clumped Isotope Workshop (Paris, 2017) led to the present  
207 inter-laboratory comparison exercise (InterCarb) to evaluate the benefits and drawbacks of a  
208 carbonate-based standardization approach as an alternative to the use of gas standards.

209  
210 The primary goal of this study was to test whether the exclusive use of carbonate reference  
211 materials can minimize inter-laboratory discrepancies and provide an alternative to the  
212 measurement of heated and equilibrated gases for the entire community. This is particularly



213 important because of the increasing number of laboratories using commercial small-sample  
 214 automated devices which cannot easily be standardized using the HG-EG approach. The  
 215 InterCarb exercise also provides an opportunity to define the best community-derived absolute  
 216  $\Delta_{47}$  values for the ETH standards of Meckler et al. (2014). Although these standards are already  
 217 used in many laboratories, their current nominal  $\Delta_{47}$  values are based on measurements from the  
 218 ETH laboratory only. The InterCarb exercise can similarly establish community accepted values  
 219 for other common carbonate reference materials, some of which have been in use for several  
 220 years, in order to provide the community with a self-consistent set of carbonate reference  
 221 materials with a broad range of bulk and clumped isotope compositions.

222

### 223 **1.1. Nomenclature and data processing**

224

225 Clumped isotope compositions are reported as an excess abundance of the CO<sub>2</sub> isotopologue of  
 226 cardinal mass 47 (dominantly the isotopologues <sup>13</sup>C<sup>18</sup>O<sup>16</sup>O) compared to a stochastic distribution  
 227 according to the formula:

228

$$R_{47} = R^{47} / R^{47*} - 1$$

229

230 where  $R^{47}$  is the abundance of the minor isotopologue with mass 47 relative to the most abundant  
 231 isotopologue with mass 44. The stochastic ratio  $R^{47*}$  is calculated using the measured abundance  
 232 of <sup>13</sup>C and <sup>18</sup>O in the sample (Affek & Eiler, 2006). According to the IUPAC guidelines the  
 233 formula does not include the factor 1000 (Coplen, 2011). Also we omit here the classically  
 234 included terms involving  $R_{45*}$  and  $R_{46*}$ , which are assumed to be zero by definition when  
 235 computing  $\delta^{13}\text{C}$  and  $\delta^{18}\text{O}$ , and in practice never exceed  $\pm 0.00002$  ‰ in our calculations (Daeron  
 236 et al. 2016). The measured abundance of isotopologues with  $m/z$  47 in the sample with respect to  
 237 the working gas (WG) in the mass spectrometer is reported in the traditional delta notation as:

238

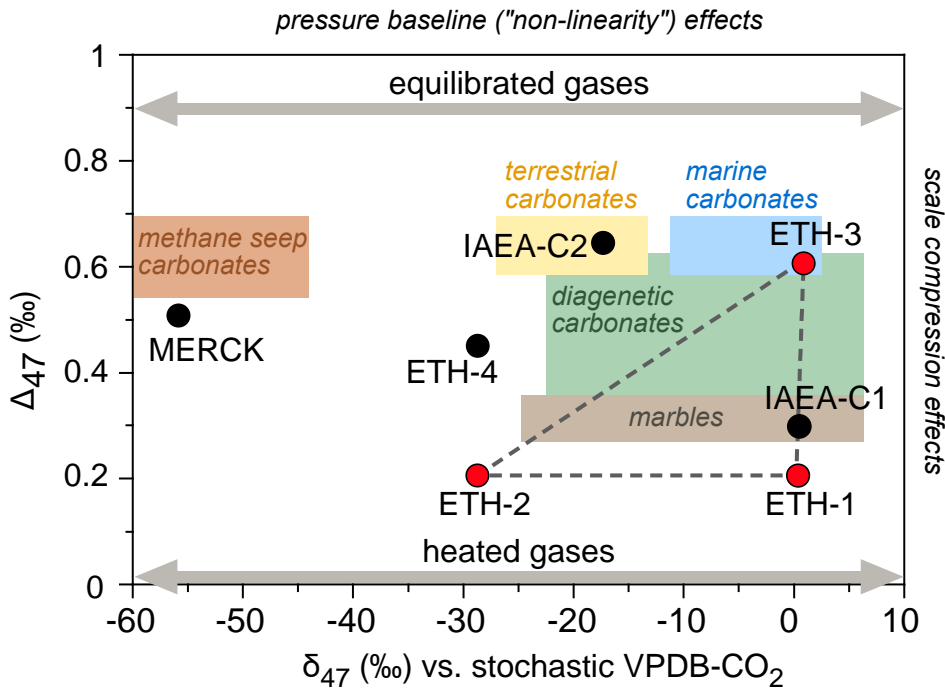
$$\delta^{47} = R^{47} / R_{WG}^{47} - 1$$

239

240 The  $\delta^{47}$  scale is a measure of the difference between the sample of interest and the WG of the  
 241 specific instrument, therefore, it cannot be compared across laboratories. The same notation is  
 242 used for masses 45, 46, 48 and 49.

243  
 244 The CO<sub>2</sub> gas-based standardization scheme for clumped isotope thermometry in carbonates relies  
 245 on a set of CO<sub>2</sub> standard gases with different bulk compositions ( $\delta^{47}$ ), preferably chosen by the  
 246 user to encompass the  $\delta^{47}$  values of unknown samples, that have been (1) heated at 1000°C to  
 247 reach a near-stochastic distribution of all isotopologues, and (2) equilibrated with water at low  
 248 temperatures to reach equilibrium enrichments in the mass-47 isotopologues (Dennis et al.,  
 249 2011). The heated gases, having a stochastic distribution of the heavy isotopes among all  
 250 isotopologues, define the zero point of the CDES scale ( $\Delta_{47} = 0.0266$  ‰) and the equilibrated  
 251 gases define a high point (e.g., at 25 °C  $\Delta_{47} = 0.9196$  ‰). The theoretical values linking  
 252 measurements to theory were calculated by Wang et al. (2004), and revised by Petersen et al.  
 253 (2019). A wide range in  $\delta^{47}$  values of gases used for normalization is generally chosen to allow  
 254 for accurate correction for an apparent dependence of  $\Delta_{47}$  on  $\delta^{47}$  which is caused by inaccurate  
 255 pressure-dependent background corrections on the  $m/z$  47 collector observed on many  
 256 instruments (Bernasconi et al., 2013; He et al., 2012). The large range in  $\Delta_{47}$  (i.e., 25 °C, 1000  
 257 °C), on the other hand, is necessary to correct for scale compression caused by processes of  
 258 scrambling and molecule recombination in the source of the mass spectrometer or elsewhere in  
 259 the sample preparation, transfer lines and/or the capillaries (Dennis et al., 2011, Swart et al.  
 260 2019). With properly chosen CO<sub>2</sub> standard gases with widely varying  $\delta^{47}$  values it is possible to  
 261 cover the entire range of natural carbonate compositions, avoiding extrapolations in the  $\delta^{47}$  vs.  
 262  $\Delta_{47}$  compositional space (Fig. 1). Note that with measurement errors (typically not better than  
 263 0.010 ‰) being relatively large compared to the natural compositional range (less than 0.5 ‰;  
 264 Fig. 1), the large (0.9 ‰) difference in  $\Delta_{47}$  of the CO<sub>2</sub> standard gases minimizes errors introduced  
 265 by uncertainties resulting from the measurement of HG and EG.

266



267

268

269 **Figure 1.**  $\delta^{47}$  vs.  $\Delta_{47}$  values of carbonate standards ( $\Delta_{47}$  on the I-CDES scale proposed here) and heated  
 270 and equilibrated gases in comparison to the compositional ranges of typical natural carbonates. The  
 271 observed range in measured clumped isotope compositions in natural carbonates can be completely  
 272 bracketed by heated and equilibrated  $\text{CO}_2$  standard gases from which  $\delta^{47}$  values have been chosen by the  
 273 user.  $\delta^{47}$  values for the anchor samples used in InterCarb (red) and the unknowns (black) are reported for  
 274 a theoretical working gas with stochastic isotope distribution, derived from VPDB. Actual  $\delta^{47}$  values will  
 275 vary by laboratory depending on the composition of the working gas. Note the smaller achievable range  
 276 in both  $\delta^{47}$  and  $\Delta_{47}$  values when using carbonate standards compared to heated and equilibrated gases, and  
 277 the large extrapolation necessary for the determination of the composition for MERCK. Heated and  
 278 equilibrated  $\text{CO}_2$  standard gases have a larger  $\Delta_{47}$  range, allowing for more robust stretching calculations  
 279 with identical numbers of standard:sample analyses.

280

281 Meckler et al. (2014) attempted to achieve a similar framework as the  $\text{CO}_2$  gas-based  
 282 standardization but with carbonate standards. They described four carbonates that were  
 283 developed at ETH Zürich to serve as replacements for HG-EGs and demonstrated that good  
 284 long- and short-term reproducibility can be achieved using only carbonates for data correction.  
 285 Bernasconi et. al (2018) discussed in detail these standards and postulated, based on a limited  
 286 inter-laboratory dataset, that carbonate standardization should generally improve inter-laboratory

287 data comparability. This claim seems arguably strengthened by the results of Meinicke et al.  
288 (2020), Peral et al. (2018), Piasecki et al. (2019), Kele et al. (2015) as recalculated by Bernasconi  
289 et al. 2018), and Jautzy et al., (2020). The first three studies produced independent foraminifera-  
290 based the fourth a travertine and the fifth a synthetic carbonate-based  $\Delta_{47}$ -temperature calibration  
291 anchored to the same set of carbonate standards, yielding statistically indistinguishable slopes  
292 and intercepts despite the use of independent sample sets and in the case of Peral et al. (2018), a  
293 different analytical system. In addition, a reanalysis of samples from five previous calibrations  
294 by Anderson et al. (2020) using carbonate standardization revealed no significant differences in  
295 temperature dependence of  $\Delta_{47}$  between the different sample sets.

296  
297 A possible limitation of carbonate standardization is that available carbonates have a smaller  
298 range in  $\delta^{47}$  and, perhaps more importantly, a smaller range in  $\Delta_{47}$  values than what is achievable  
299 with heated and equilibrated gases. In some specific cases, standardization procedures may  
300 require extrapolation to compositions that are not within the  $\delta^{47}$ - $\Delta_{47}$  space created by carbonate  
301 standards (Fig. 1). In addition, the range of  $\Delta_{47}$  values for carbonates is only on the order of 0.5  
302 ‰ between 0 and 1000 °C. The smaller range in  $\Delta_{47}$  compared to HG-EG requires higher  
303 precision and thus high standard replication and a ~50:50 standard to sample replicate ratio to  
304 keep normalization errors small (see Daëron (this volume) and Kocken et al., 2019 for details).

305

## 306 **1.2. InterCarb goals and design:**

307

308 InterCarb was designed with the aim to carefully evaluate the potential of carbonates to serve as  
309 a standardization scheme that improves inter-laboratory agreement for ‘unknown’ carbonates  
310 both inside and outside of the  $\delta^{47}$ - $\Delta_{47}$  space defined by the anchor samples (Fig. 2). The main  
311 questions posed are:

312

- 313 1. Is it possible to produce consistent carbonate clumped isotope measurements across  
314 laboratories using carbonate reference materials exclusively? In other words, does the  
315 observed inter-laboratory scatter in  $\Delta_{47}$  values match that expected from intra-laboratory  
316 analytical precision?

- 317 2. How well does the carbonate standardization approach perform when extrapolating  
 318 beyond the  $\delta^{47}$ - $\Delta_{47}$  compositional space sampled by a set of carbonate reference  
 319 materials?  
 320 3. Do carbonate reference materials fully correct effects arising from different reaction  
 321 temperatures, sample preparation protocols, and analytical equipment?  
 322 4. Can we define a self-consistent set of widely available reference materials with  
 323 community-agreed compositions accurately anchored to the CDES scale?  
 324 5. Does the use of carbonate reference materials for standardization improve the inter-  
 325 laboratory reproducibility compared to HG-EG standardization?

326

327 **1.3 Approach:**

328

329 Seven carbonate standards with a large range of  $\delta^{47}$  and  $\Delta_{47}$  values (Fig. 1) were distributed  
 330 among participating laboratories and analyzed, treating three carbonates as “anchors” (whose  
 331  $\Delta_{47}$  values are assigned *a priori*) and the remaining four as “unknowns” (whose  $\Delta_{47}$  values are  
 332 unknown, to be determined by comparison with the anchors). Due to their relatively widespread  
 333 use in different laboratories, the three reference materials ETH-1, ETH-2 and ETH-3 (Meckler et  
 334 al. 2014; Bernasconi et al. 2018) were chosen as anchors. They are still available today in  
 335 relatively large quantities (>600 g), have been in use at ETH since 2013 and in many other  
 336 laboratories for several years. Importantly, they have been thoroughly tested for homogeneity  
 337 based on thousands of measurements in 80 to 150 $\mu$ g aliquot sizes in different laboratories and no  
 338 changes in composition have been noticed at ETH in the 7 years they have been in use.

339

340 The “unknown” InterCarb reference materials were chosen to cover a wide natural range in  $\delta^{47}$   
 341 and  $\Delta_{47}$  values. These samples had to be available in large quantities, inexpensive, and if possible  
 342 distributed by an organization with a long-term perspective in order to ensure future data quality  
 343 and availability for the increasing number of laboratories globally.

344

## 345 **2 Materials and Methods**

### 346 **2.1. Sample description**

347

348 The anchor samples ETH-1 (Carrara marble heated at 600 °C), ETH-2 (synthetic carbonate  
349 heated at 600 °C) and ETH-3 (Upper cretaceous chalk) are described in detail in Bernasconi et  
350 al. (2018).

351

352 IAEA-C1 (marble from Carrara, Italy) is distributed by the International Atomic Energy Agency  
353 (IAEA) as a mechanically crushed and milled product with grains ranging from 1.6 to 5 mm. All  
354 50 g provided were ground and thoroughly homogenized in a ball mill at ETH Zürich to a grain  
355 size of less than 100 µm and transferred in 0.5 g aliquots to plastic vials for distribution.

356

357 IAEA-C2 is a freshwater travertine from Bavaria distributed by IAEA as a powder which was  
358 treated identically to IAEA-C1.

359

360 ETH-4 is a commercially available synthetic calcite (Riedel-De Haën; calcium carbonate Puriss.  
361 p.a.; Lot No. 30800) with intermediate formation temperature and the same bulk isotope  
362 composition as ETH-2 (see Bernasconi 2018 for details).

363

364 MERCK (Catalog No. 1.02059.0050; lot no. B1164559 515) is an ultra-pure, commercially  
365 available synthetic calcite and was chosen for its very low  $\delta^{13}\text{C}$  and  $\delta^{18}\text{O}$  values, of  
366 approximately -41.7 ‰ and -15.5 ‰ (VPDB), respectively. This sample represents an extreme  
367 case of extrapolation from the  $\delta^{47}\text{-}\Delta_{47}$  space defined by the anchor materials (Fig. 2).

368

### 369 **2.2 Instrumentation:**

370

371 The reported data were produced with a variety of preparation systems including custom built  
372 (13 laboratories) and commercial systems (11 laboratories; Protium MS IBEX, ThermoFisher  
373 Scientific Kiel IV device and Nu Instruments Nucarb). Reaction temperatures were generally 90

374 °C for “large-sample” custom preparation systems and 70 °C for the Kiel and the NuCarb. Four  
375 mass spectrometer types were used: Thermo Fisher Scientific MAT253 and 253Plus, Nu  
376 Instruments Perspective, and Elementar Isoprime 100. All participants contributed results they  
377 considered to be of a “publication-grade” quality, based on their existing quality-control  
378 procedures.

### 379 **2.3 Clumped isotope compositions of the ETH anchor materials.**

380  
381 The clumped isotope compositions of the four ETH reference materials relative to the CO<sub>2</sub>  
382 reference frame CDES were first reassessed based on new data provided by 9 laboratories that  
383 also provided HG and EG data measured during the same sessions as the ETH reference  
384 materials. Data processing was performed with the same Python script that was used to process  
385 the carbonate data in order to avoid any differences in data processing (see section 2.4).

386  
387 Although, strictly speaking, <sup>13</sup>C-<sup>18</sup>O clumping in carbonate represents a mass-63 anomaly, the  
388 clumped isotope composition of carbonate minerals is reported as  $\Delta_{47}$  i.e. as the mass-47 excess  
389 in the CO<sub>2</sub> produced by acid digestion of these minerals, including the respective temperature-  
390 dependent isotopic fractionation. As initially all reactions were carried out at 25°C (Ghosh et al.  
391 2006), the  $\Delta_{47}$  values have traditionally been reported for a 25°C acid temperature. With the  
392 advent of automated extraction lines, reaction temperatures have been raised to 70 or 90°C. To  
393 account for the temperature dependence of the acid fractionation factor (Guo et al., 2009,  
394 Petersen et al. 2019) and to project these results back to the original 25°C acid reaction, various  
395 acid temperature correction values have been reported over time, based on experimental  
396 observations and/or theoretical predictions. Given that here seven out of ten laboratories reacted  
397 carbonates at 90 °C, two at 70 °C, and only one at 25 °C, our redetermination of the  $\Delta_{47}$  values  
398 of ETH-1/2/3/4 relative to the CDES projected to 25°C would substantially rely on the accuracy  
399 of these acid temperature corrections (which typically range between 60–90 ppm). Here we  
400 report instead the  $\Delta_{47}$  values of CO<sub>2</sub> produced by reacting ETH-1/2/3/4 at 90 °C. With this  
401 choice the numerical effect of poorly known acid corrections is minimized, because the data  
402 from 70 °C and 25 °C reactions have relatively little influence on the final, error-weighted  
403 average  $\Delta_{47}$  values (cf statistical weights in Fig. 2). We thus propose to break with tradition and

404 define the nominal  $\Delta_{47}$  values of the anchor standards as those of CO<sub>2</sub> produced at 90 °C,  
405 providing the most robust relationship to the CDES.

#### 406 **2.4 Data processing, correction and error assessment.**

407  
408 It should be stressed that the InterCarb experiment, by design, is not intended to grade the  
409 analytical “performance” of individual laboratories. Each participating laboratory (or mass  
410 spectrometer, in the case of laboratories with several instruments) was thus randomly assigned an  
411 anonymous identifying number. Within each laboratory, analyses were grouped in different  
412 analytical sessions defined by the participants themselves. An analytical session is generally  
413 defined by a time in which the behavior of the analytical system (preparation system, source  
414 tuning, backgrounds, isotope scrambling in the source) is considered to be similar. The database  
415 record of each analysis consists of a laboratory identifier; a session identifier; an analysis  
416 identifier; the name of the analyzed sample; the mass spectrometer model; the acid reaction  
417 temperature; the mass of reagent carbonate; and background-corrected  $\delta^{45}$ ,  $\delta^{46}$  and  $\delta^{47}$  values.

418  
419 The only instrumental corrections to the raw data applied independently by each participating  
420 laboratory were background corrections (“Pressure Baseline Correction” or PBL) to the ion  
421 currents/voltages (Bernasconi et al., 2013; He et al., 2012). The PBL is strongly dependent on  
422 instrument design (it is not observed in some instruments) and configuration, and varies  
423 temporally depending on many factors. This correction, therefore, can only be carried out by  
424 each participating laboratory according to its own established procedures and monitoring.

425  
426 To avoid artefacts arising from different calculation/standardization procedures, rounding errors,  
427 and <sup>17</sup>O correction parameters, raw data from all laboratories were processed by a single Python  
428 script (<http://doi.org/10.5281/zenodo.4314449>) based on data reduction, standardization and  
429 error propagation methods described in detail in the companion paper (Daëron, this volume).  
430 Here we briefly summarize these calculations.

431  
432



433 Session-averaged, background-corrected  $\delta^{45}$  and  $\delta^{46}$  values for each of the three anchor samples  
 434 were first used to calculate the bulk isotope composition of the working gas used in each session,  
 435 based on (a) previously reported  $\delta^{13}\text{C}_{\text{VPDB}}$  and  $\delta^{18}\text{O}_{\text{VPDB}}$  values of ETH-1, ETH-2, and ETH-3  
 436 (Bernasconi et al., 2018), (b) the IUPAC  $^{17}\text{O}$  correction parameters of Brand et al. (2010), and  
 437 (c) a temperature-dependent oxygen-18 acid fractionation factor between  $\text{CO}_2$  and calcite of Kim  
 438 et al. (2015). This recalculation of working gas bulk compositions avoids (small) discrepancies  
 439 potentially introduced by inaccuracies in the nominal compositions of the working gases.

440

441 Raw  $\Delta_{47}$  values were computed according to:

442

$$R_{47}^{\square\square\square} = R^{47} / R^{47*} - 1$$

443

444 Where  $R^{47}$  is the measured ratio and  $R^{47*}$  the calculated stochastic ratio of mass 47 over mass 44  
 445 of  $\text{CO}_2$ , assuming perfectly linear IRMS measurements and a stochastic working gas. Values are  
 446 then normalized to “absolute”  $\Delta_{47}$  values (noted  $R_{47}^{\square\square\square}$  in the equation below, and simply  $\Delta_{47}$   
 447 thereafter) using session-specific relationships of the form:

448

$$R_{47}^{\text{raw}} = a R_{47}^{\square\square\square} + b \delta^{47} + c$$

449

450 For each session, the best-fit standardization parameters (a, b, c) are computed from an  
 451 unweighted least-squares regression, treating  $R_{47}^{\text{raw}}$  as the response variable, only considering the  
 452 three anchor samples ETH-1, ETH-2, and ETH-3. Note the advantage of this form over that in  
 453 Dennis et al. 2011 is the ability to have three standards with distinct  $\Delta_{47}$  values whilst being able  
 454 to solve for b (compositional nonlinearity) (Daëron et al. 2016). Absolute  $\Delta_{47}$  values are then  
 455 computed for all replicates within that session. Standardization parameters for all sessions are  
 456 listed in Table S1 (supplementary information).

457

458 In figure labels and in tables, measurement uncertainties are reported as 1SE of the mean,  
 459 considering fully-propagated errors taking into account reference frame corrections. In Figures 2  
 460 and 4, different types of error bars reflect either these full analytical errors, or those only  
 461 considering uncertainties in the analyses of a given sample. In both cases, the analytical error  
 462 assigned to each individual raw  $\Delta_{47}$  analysis is equal to the pooled “external” repeatability of raw

463  $\Delta_{47}$  measurements for all samples (anchors and unknowns) within each session. This treatment of  
464 error is a new approach that more fully accounts for error in both the sample measurement and  
465 reference frame.

### 466 **3 Results and Discussion**

#### 467 **3.1 Redetermination of nominal $\Delta_{47}$ values for the ETH standards relative to heated and** 468 **equilibrated CO<sub>2</sub> gases**

469 The weighted averages of the 4 standards (projected to 90 °C for the reactions at 25 and 70°C  
470 using the acid temperature correction suggested by Petersen et al. 2019), comprising 619  
471 analyses of the carbonate standards and 828 heated and equilibrated gases from 9 different  
472 laboratories, are reported in Table 1 and Fig 4. The large number of analyses and the appropriate  
473 consideration of the errors on the anchors (CO<sub>2</sub> gas analyses) distinguishes this effort from  
474 previous work, and allows a robust redetermination of the accepted values of the ETH reference  
475 materials with 1SE uncertainties of 2 ppm or less.

476

477

Laboratory	all	A	B	C	D	E	F	G	H	I	J	
<b>N of sessions</b>	<b>34</b>	4	4	11	7	1	2	1	1	1	2	
<b>N of H/E CO2</b>	<b>946</b>	44	193	257	85	47	21	38	192	13	56	
<b>ETH-1</b>	<b>N of analyses</b>	<b>232</b>	34	14	5	54	4	4	19	58	8	32
	$\Delta_{47}$ (90 °C acid)	<b>0.2052</b>	0.2016	0.1926	0.2108	0.1940	0.1601	0.2013	0.2143	0.1932	0.2183	0.2152
	$\pm$ 1SE	<b>0.0016</b>	0.0046	0.0058	0.0069	0.0042	0.0245	0.0107	0.0032	0.0045	0.0109	0.0036
	Statistical weight		0.118	0.074	0.053	0.146	0.004	0.022	0.241	0.124	0.021	0.197
<b>ETH-2</b>	<b>N of analyses</b>	<b>215</b>	23	13	11	51	4	4	18	51	8	32
	$\Delta_{47}$ (90 °C acid)	<b>0.2085</b>	0.2077	0.1840	0.2225	0.1978	0.1374	0.1650	0.2141	0.1968	0.2172	0.2170
	$\pm$ 1SE	<b>0.0015</b>	0.0047	0.0070	0.0046	0.0050	0.0233	0.0101	0.0029	0.0043	0.0154	0.0033
	Statistical weight		0.105	0.047	0.108	0.092	0.004	0.023	0.272	0.125	0.010	0.213
<b>ETH-3</b>	<b>N of analyses</b>	<b>264</b>	55	15	20	54	4	5	15	59	8	29
	$\Delta_{47}$ (90 °C acid)	<b>0.6132</b>	0.6156	0.5975	0.6169	0.6102	0.5950	0.6143	0.6159	0.6094	0.6428	0.6124
	$\pm$ 1SE	<b>0.0014</b>	0.0037	0.0056	0.0033	0.0038	0.0237	0.0099	0.0033	0.0042	0.0103	0.0035
	Statistical weight		0.140	0.062	0.175	0.134	0.003	0.020	0.179	0.110	0.018	0.158
<b>ETH-4</b>	<b>N of analyses</b>	<b>162</b>	10	12	5	55	4	4	12	47	7	6
	$\Delta_{47}$ (90 °C acid)	<b>0.4505</b>	0.4438	0.4230	0.4624	0.4506	0.4230	0.4454	0.4560	0.4414	0.4831	0.4646
	$\pm$ 1SE	<b>0.0018</b>	0.0058	0.0071	0.0068	0.0049	0.0226	0.0095	0.0032	0.0042	0.0161	0.0057
	Statistical weight		0.093	0.064	0.068	0.133	0.006	0.035	0.314	0.177	0.012	0.097

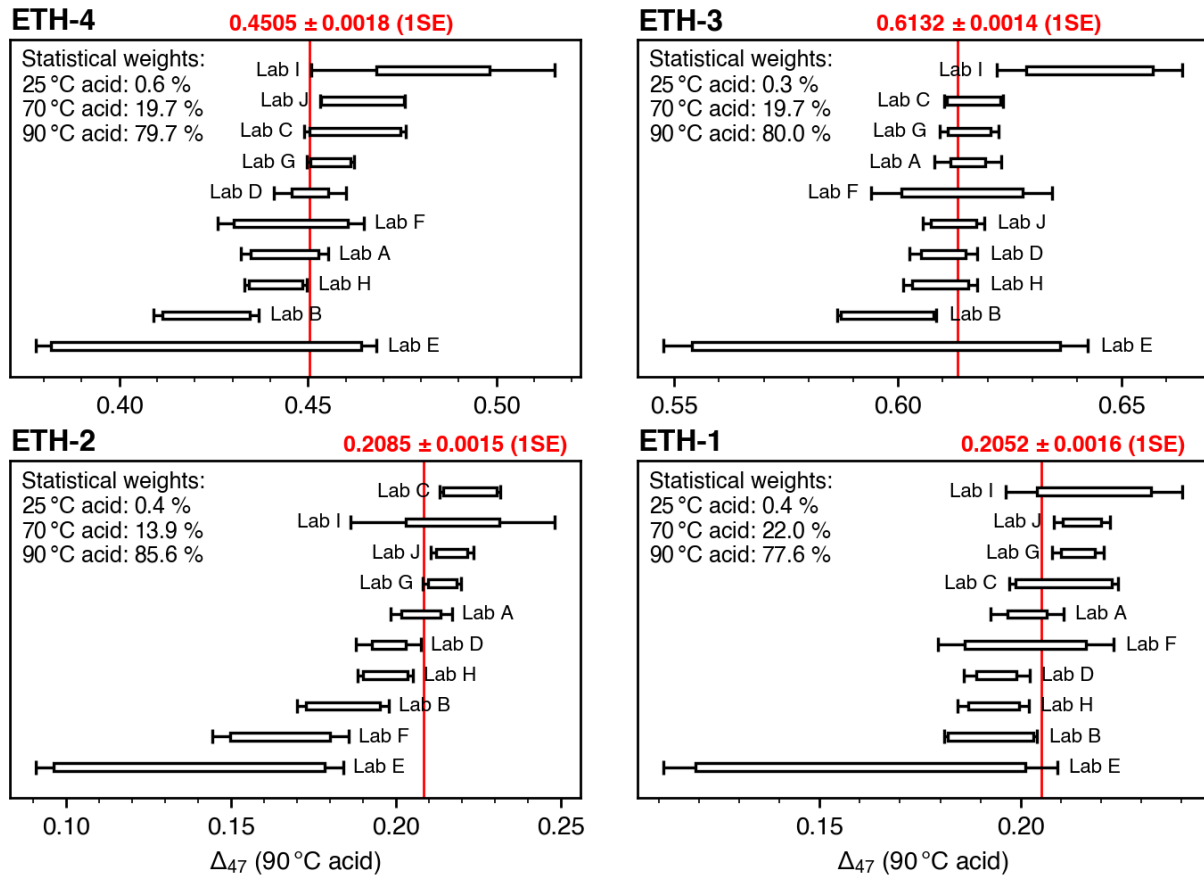
478

479 **Table 1.** Newly determined nominal  $\Delta_{47}$  values of the ETH standards projected to 90 °C acid reaction  
 480 using acid correction factors of -0.088 ‰ and -0.022 ‰ for 25 °C and 70 °C reactions, respectively  
 481 (Petersen et al. 2019). Reported standard errors represent analytical uncertainties associated both with  
 482 reference frame errors (HG/EG) and carbonate sample reproducibility (Daëron, this volume).

483

484 When compared with the published values in Bernasconi et al. (2018) the average  $\Delta_{47}$  values  
 485 ETH-1 and ETH-2, projected back to 25 °C (+0.088 ‰), are respectively 0.035 and 0.040 ‰  
 486 greater than the original values, whereas ETH-3 increases by 0.010 and ETH-4 by 0.031 ‰.  
 487 Such positive offset of  $\Delta_{47}$  values reported in Bernasconi et al. (2018) versus other laboratories  
 488 has also been reported in Fiebig et al. (2019) and Thaler et al. (2020).

489



490

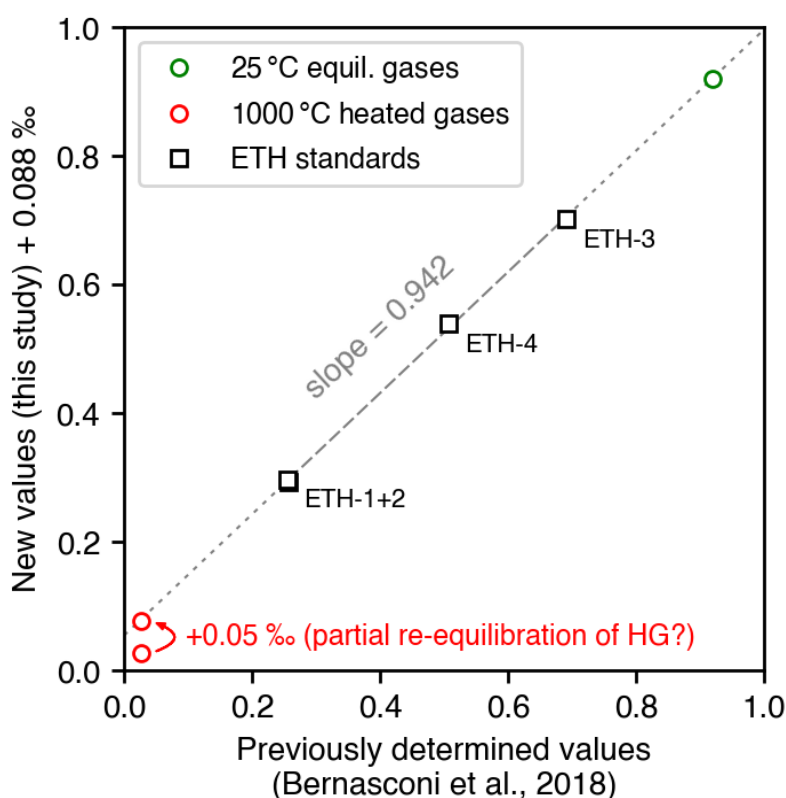
491 **Figure 2.** New determination of  $\Delta_{47}$  values for the four ETH standards relative to the CDES using updated  
 492  $\text{CO}_2$  equilibrium values. These measurements, using acid reaction temperatures of 90 °C, 70 °C or 25 °C,  
 493 are projected to 90 °C using acid corrections of -0.088 ‰ and -0.022 ‰ for 25 °C and 70 °C reactions,  
 494 respectively (Petersen et al. 2019). Error bars correspond to 95 % confidence limits taking into account  
 495 fully propagated errors (i.e. taking into account errors in both unknown and anchor analyses). Boxes  
 496 correspond to 95 % confidence limits not accounting for normalization errors (i.e. only taking into  
 497 account errors in unknown analyses). Red numbers are the error-weighted average values (with statistical  
 498 weights summarized in upper-left corners). Note different horizontal scales for the different samples.

499

500

501 The observation that these changes in nominal values decrease as  $\Delta_{47}$  increases suggests a simple  
 502 hypothesis to explain this discrepancy: in the original study of Meckler et al. (2014), the  
 503 carbonate samples and the heated/equilibrated  $\text{CO}_2$  gases experienced different analytical  
 504 procedures. The HGs were measured as large samples at constant beam intensity through a

505 different capillary than the carbonates, which were measured using the microvolume and a  
 506 decreasing beam). The potential effects of partial re-equilibration for the heated gases in the gas  
 507 preparation line or in the capillaries of the mass spectrometer could be significant whereas it  
 508 would be minuscule for the gases equilibrated at 25°, leading to an overestimation of  $\Delta_{47}$  scale  
 509 compression and thus of the stretching applied to the  $\Delta_{47}$  scale towards theoretical values. The  
 510 observed changes in apparent ETH-1 and ETH-2  $\Delta_{47}$  values may therefore simply reflect partial  
 511 re-equilibration of heated gases at the time of measurements at ETH (and reported in Meckler et  
 512 al., 2014), increasing their values in the original study by about 0.05 ‰ (Fig. 3).  
 513



514  
 515 **Figure 3.** New nominal  $\Delta_{47}$  values for the ETH standards compared to previously reported ones. The  
 516 dashed gray line is a linear regression through the new versus old values of ETH-1/2/3/4, whose  
 517 extrapolation coincides with 25 °C equilibrated  $\text{CO}_2$  but not with heated gases. Apparent changes in the  
 518 ETH-1/2/3/4 values thus scale linearly with the  $\Delta_{47}$  difference between carbonate samples and 25 °C  
 519 equilibrated  $\text{CO}_2$ , suggesting that  $\Delta_{47}$  values of heated gases in the original study may have been biased by  
 520  $\sim +0.05$  ‰ through partial re-equilibration at room temperature between the quenching of heated  $\text{CO}_2$  and  
 521 its ionization in the IRMS source.  
 522

523 It has been suggested previously that ETH-1 and ETH-2 should be indistinguishable in  $\Delta_{47}$  and  
524 close to stochastic distribution (Müller, Violay, et al., 2017). This is because  $\Delta_{47}$  values of ETH-  
525 1 and ETH-2, originally heated to 600 °C, were found to be higher by only around 0.006 ‰ from  
526 the same carbonates heated at 1000 °C to achieve stochastic distribution of the isotopes.  
527 However, additional test measurements in multiple laboratories of samples heated at >1000 °C  
528 are necessary to confirm this observation.

529  
530 One laboratory (Lab F) did however observe a large difference in the value for ETH-1 and ETH-  
531 2, although their values of ETH-3 and ETH-4 are similar to other laboratories. The reason for  
532 these inconsistencies is probably due to the fact that ETH-1 was only measured four times with a  
533 limited number of HG/EG, and ETH-2 and ETH-4 were not measured in the same session. In  
534 addition, the labs with the smallest number of replicate measurements have uncertainties that are  
535 systematically larger (Table 1). These results highlight the importance of correction procedures  
536 in clumped isotope analysis. Sufficient replication of both standards and samples is critical and,  
537 if insufficient, offsets can arise when comparing results from different sessions. Due to these  
538 difficulties it is good practice to spread replicates of the same sample in different sessions over  
539 longer periods of time to obtain accurate results and follow a ~50:50 standard to sample replicate  
540 ratio.

541  
542 Based on the results above, the difference between the average of ETH1/2 and ETH-3 is reduced  
543 by 0.028 ‰, thus leading to a compression of the scale by about 5.6 % compared to the values  
544 reported by Bernasconi et al. (2018). As a consequence, the slopes of published temperature  
545 calibrations produced with carbonate standardization (Kele et al, 2015; Bernasconi et al., 2018;  
546 Jautzy et al., 2021; Meinicke et al., 2020; Peral et al., 2018; Piasecki et al., 2019) will become  
547 slightly shallower, with more positive y-intercepts. If  $\Delta_{47}$  results from previous publications are  
548 also recalculated with the new standard values (see section 3.4), however, changes in calculated  
549 formation temperatures will be negligible. For this reason, when comparing data from  
550 publications using old accepted values of the ETH standards for standardization (either those  
551 published by Meckler et al. (2014) or those recalculated with the IUPAC parameters by  
552 Bernasconi et al. (2018)) to newer data, it is recommended to directly compare the reconstructed  
553 temperatures rather than recalculating  $\Delta_{47}$ . Full recalculation of old measurements requires the

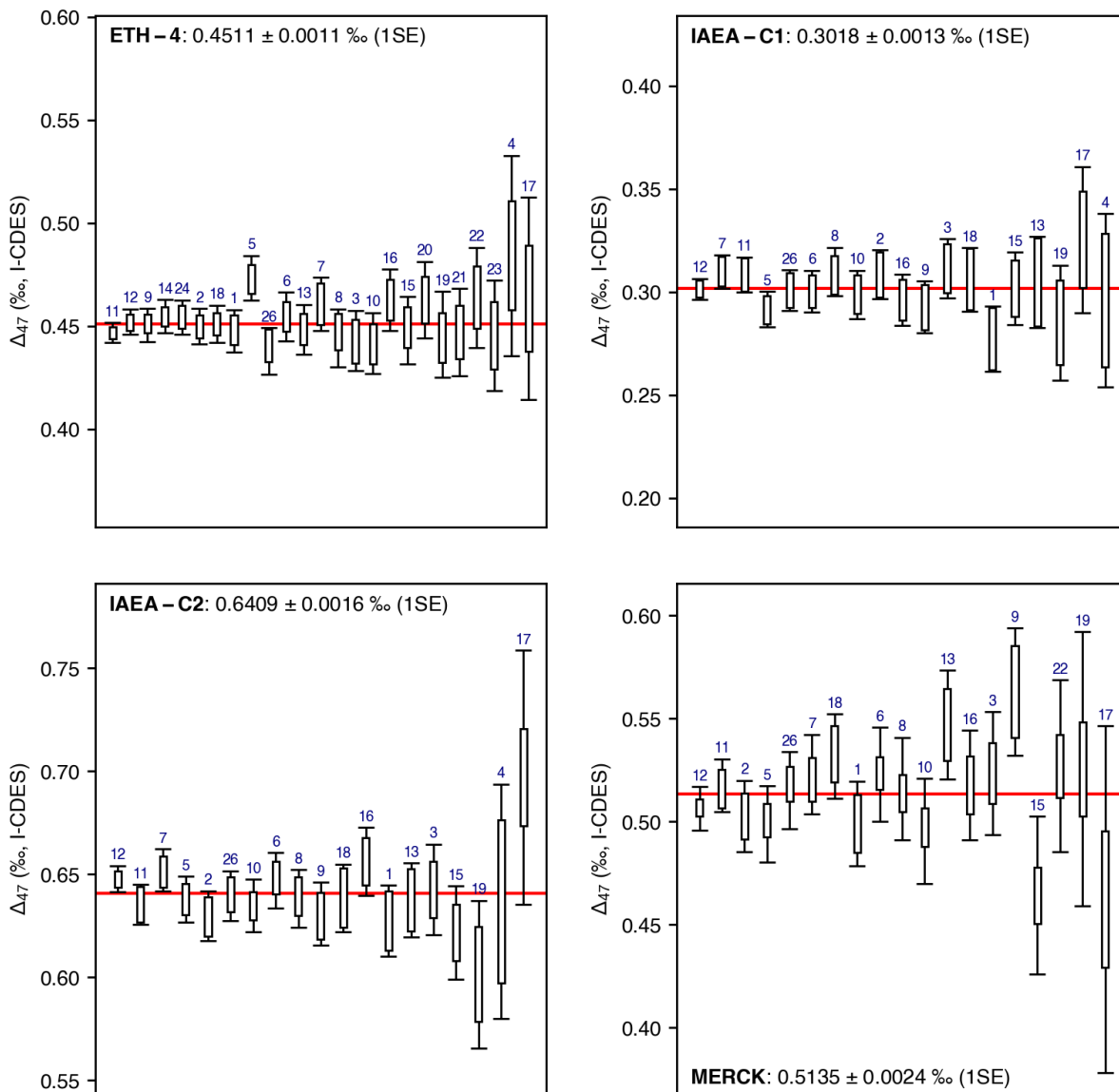
554 availability of the entire dataset including standards and the same correction procedures (e.g.,  
555 averaging method) used in the original publications, and is described in Appendix 1.

556

### 557 **3.2 InterCarb results**

558 Results for the unknown carbonate samples were obtained from 26 mass spectrometers in 22  
559 laboratories.  $\Delta_{47}$  values of the 4 unknown samples were normalized to the new community-  
560 derived values of ETH-1, ETH-2, and ETH-3 (Table 1), averaged per individual analytical  
561 session and mass spectrometer (Tables 2 and S1). Sample mean values obtained in each mass  
562 spectrometer are shown in Figure 4. The details of each analytical session, including the number  
563 of samples and standards measured, the composition of the working standard, the scaling  
564 parameters and the reproducibilities of the individual sessions are given in the supplementary  
565 information (Table S1). Some laboratories reported data for only a subset of the unknown  
566 samples, and both replication level and analytical reproducibility vary from laboratory to  
567 laboratory (Table S1).

568 To clearly distinguish  $\Delta_{47}$  values normalized to the carbon dioxide reference scale (CDES) using  
569 carbonates rather than heated and equilibrated gases, we propose the new acronym (I-CDES),  
570 short for InterCarb-CDES, to reflect the use of the proposed InterCarb reference materials for  
571 data standardization (see section 3.5 for more details).



572

573 **Figure 4.** Final InterCarb results by laboratory. Error bars correspond to fully propagated 95 %  
 574 confidence limits, taking into account errors in both unknown and anchor analyses. Boxes correspond to  
 575 95 % confidence limits not accounting for normalization errors (i.e., only taking into account errors in  
 576 unknown analyses). Results are sorted by increasing analytical errors, and laboratories are identified by  
 577 number. Overall error weighted average  $\Delta_{47}$  values are displayed as solid red lines and reported in each  
 578 panel. All plots have the same vertical scale.

579

580 Laboratories with stronger analytical constraints (i.e. better in-lab repeatability of  $\Delta_{47}$   
 581 measurements and/or greater number of analyses) generally converge towards the overall mean



582 value for each unknown sample. This observation suggests that the observed inter-laboratory  
 583 variability is largely due to random errors that can be alleviated by replication, even for  
 584 laboratories with relatively large analytical errors on individual measurements. It is also notable  
 585 that fully propagated analytical errors (i.e., taking into account uncertainties in the  
 586 standardization procedure) can be substantially larger than the errors based only on the  
 587 uncertainty associated with unknown sample analyses, which is what is generally reported in the  
 588 literature. The increase in error is also related to in-lab repeatability and the number of standards  
 589 measured. In addition, error increases for unknown samples whose compositions lie outside the  
 590 “anchor triangle” defined by ETH-1/2/3. This is illustrated by the increased scatter and errors  
 591 associated with MERCK, the carbonate material farthest from the “anchor triangle”, consistent  
 592 with the models of Daëron (this volume, see also Kocken et al., 2019).

593  
 594 As seen in Table S1, there are stark differences in the total number of replicate analyses and the  
 595 typical  $\Delta_{47}$  reproducibility achieved in different laboratories. As a result, final uncertainties in the  
 596 average  $\Delta_{47}$  values of unknown samples vary considerably (Fig. 4). It is thus not very useful to  
 597 quantify inter-laboratory reproducibility in terms of a single, overall scatter of  $\Delta_{47}$  values: inter-  
 598 laboratory variability should be small among laboratories with small analytical uncertainties, and  
 599 larger among laboratories with few replicate analyses and/or poor analytical repeatability. We  
 600 may still, however, assess whether inter-lab discrepancies are significantly larger than expected  
 601 from in-lab analytical uncertainties, i.e., whether we can detect the effects of hypothetical  
 602 unrecognized sources of scatter beyond known analytical errors.

603  
 604  
 605  
 606

607

Lab	ETH-4		IAEA-C1		IAEA-C2		MERCK	
	$\Delta_{47(I-CDES)} (\% \pm 1SE)$	N	$\Delta_{47(I-CDES)} (\% \pm 1SE)$	N	$\Delta_{47(I-CDES)} (\% \pm 1SE)$	N	$\Delta_{47(I-CDES)} (\% \pm 1SE)$	N
1	0.4477 ± 0.0052	80	0.2773 ± 0.0080	19	0.6275 ± 0.0088	21	0.4991 ± 0.0105	22
2	0.4499 ± 0.0044	22	0.3086 ± 0.0060	6	0.6299 ± 0.0061	8	0.5025 ± 0.0089	6
3	0.4430 ± 0.0074	23	0.3114 ± 0.0073	17	0.6427 ± 0.0112	13	0.5235 ± 0.0152	11
4	0.4841 ± 0.0248	9	0.2959 ± 0.0215	6	0.6368 ± 0.0291	4		
5	0.4734 ± 0.0055	22	0.2916 ± 0.0044	18	0.6378 ± 0.0057	17	0.4987 ± 0.0094	14
6	0.4545 ± 0.0060	12	0.3004 ± 0.0051	12	0.6471 ± 0.0069	12	0.5229 ± 0.0116	12
7	0.4607 ± 0.0066	8	0.3099 ± 0.0042	16	0.6520 ± 0.0052	15	0.5231 ± 0.0098	8
8	0.4442 ± 0.0072	73	0.3099 ± 0.0060	67	0.6383 ± 0.0071	62	0.5159 ± 0.0127	70
9	0.4505 ± 0.0041	65	0.2926 ± 0.0064	8	0.6309 ± 0.0078	8	0.5630 ± 0.0158	2
10	0.4416 ± 0.0075	26	0.2987 ± 0.0060	33	0.6348 ± 0.0065	62	0.4954 ± 0.0130	31
11	0.4468 ± 0.0025	222	0.3085 ± 0.0043	31	0.6354 ± 0.0050	27	0.5175 ± 0.0066	25
12	0.4521 ± 0.0032	21	0.3015 ± 0.0026	20	0.6479 ± 0.0032	21	0.5064 ± 0.0054	19
13	0.4484 ± 0.0062	47	0.3048 ± 0.0113	6	0.6376 ± 0.0091	12	0.5470 ± 0.0135	9
14	0.4548 ± 0.0041	46						
15	0.4480 ± 0.0083	8	0.3016 ± 0.0090	4	0.6217 ± 0.0116	4	0.4642 ± 0.0195	4
16	0.4627 ± 0.0076	4	0.2962 ± 0.0063	4	0.6563 ± 0.0084	3	0.5176 ± 0.0136	2
17	0.4634 ± 0.0250	5	0.3254 ± 0.0181	6	0.6971 ± 0.0314	6	0.4623 ± 0.0429	3
18	0.4510 ± 0.0046	196	0.3060 ± 0.0079	26	0.6386 ± 0.0084	28	0.5317 ± 0.0104	31
19	0.4460 ± 0.0106	14	0.2851 ± 0.0142	5	0.6015 ± 0.0183	4	0.5256 ± 0.0339	4
20	0.4627 ± 0.0095	6						
21	0.4470 ± 0.0108	3						
22	0.4639 ± 0.0124	7					0.5269 ± 0.0213	7
23	0.4453 ± 0.0137	6						
24	0.4544 ± 0.0042	12						
26	0.4378 ± 0.0058	8	0.3008 ± 0.0051	6	0.6396 ± 0.0062	6	0.5152 ± 0.0095	6
<b>w. avg</b>	<b>0.4511 ± 0.0011</b>	<b>945</b>	<b>0.3018 ± 0.0013</b>	<b>310</b>	<b>0.6409 ± 0.0016</b>	<b>333</b>	<b>0.5135 ± 0.0024</b>	<b>286</b>

608

609 **Table 2.** Average  $\Delta_{47}$  values ( $\pm 1SE$ , fully propagated uncertainties) obtained by each participating  
 610 laboratory. A full table of average  $\Delta_{47}$  values for all analytical sessions is included in supplementary  
 611 materials (Table S3).

612

613

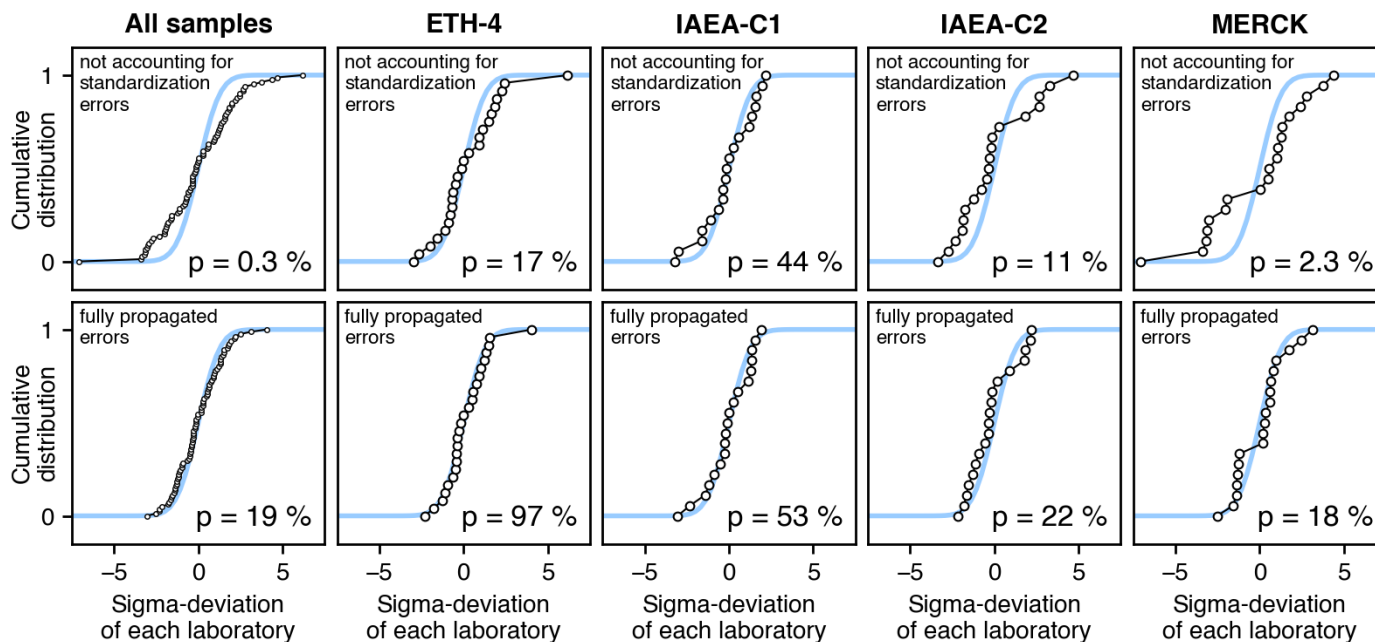
614

615 In order to do so, we compute the “number-of-sigma” deviation obtained by each laboratory for  
616 each unknown sample, relative to that sample’s overall weighted average value. For example, the  
617 sigma-deviation for sample ETH-4 and Lab01 is equal to  $(0.4477 - 0.4511) / 0.0052 = -0.66$  and  
618 that for MERCK and Lab13 is equal to  $(0.5470 - 0.5135) / 0.0135 = +2.48$ . If the analytical  
619 errors reported in Table 2 are reasonably accurate, we expect the population of sigma-deviations  
620 among all labs to be distributed as the canonical Gaussian distribution ( $\mu = 0$ ;  $\sigma = 1$ ), and we can  
621 test this prediction using established statistical methods such as a Kolmogorov–Smirnov test of  
622 normality (Massey, 1951). We carried out this test for two cases: only considering the error of  
623 sample replication (Fig. 5, upper row) and secondly including the normalization error, (i.e. the  
624 fully propagated error (Fig. 5, lower row). If we neglect uncertainties arising from  
625 standardization (the “allogenic” errors of Daëron, this volume), the sigma-deviations are no  
626 longer normally distributed ( $p = 0.003$ , upper-left panel). As shown in the lower-left panel of  
627 Figure 6, the distribution of sigma-deviations for all labs and all samples is statistically  
628 indistinguishable from the expected normal distribution ( $p = 0.19$ ) when considering fully  
629 propagated analytical errors. Figure 6 also illustrates that neglecting standardization errors does  
630 not strongly affect the normality of sigma-deviations for ETH-4 and IAEA-C1, both of which  
631 have  $\delta^{47}$  and  $\Delta_{47}$  values within the range covered by the three anchor samples. By contrast,  
632 sigma-deviations for unknowns with “exotic” isotopic compositions (IAEA-C2 and especially  
633 MERCK) are only normally distributed if standardization uncertainties are correctly accounted  
634 for.

635

636 Based on these tests, we conclude that the inter-laboratory scatter observed in the InterCarb data  
637 set is neither smaller nor larger than expected from the analytical uncertainties computed within  
638 each laboratory, as long as standardization errors are taken into account. This important finding  
639 implies that, at least for the time being, we can rule out any systematic inter-laboratory  
640 discrepancies in carbonate-standardized  $\Delta_{47}$  measurements, which constitutes an important  
641 milestone in the progress of clumped isotope measurement techniques.

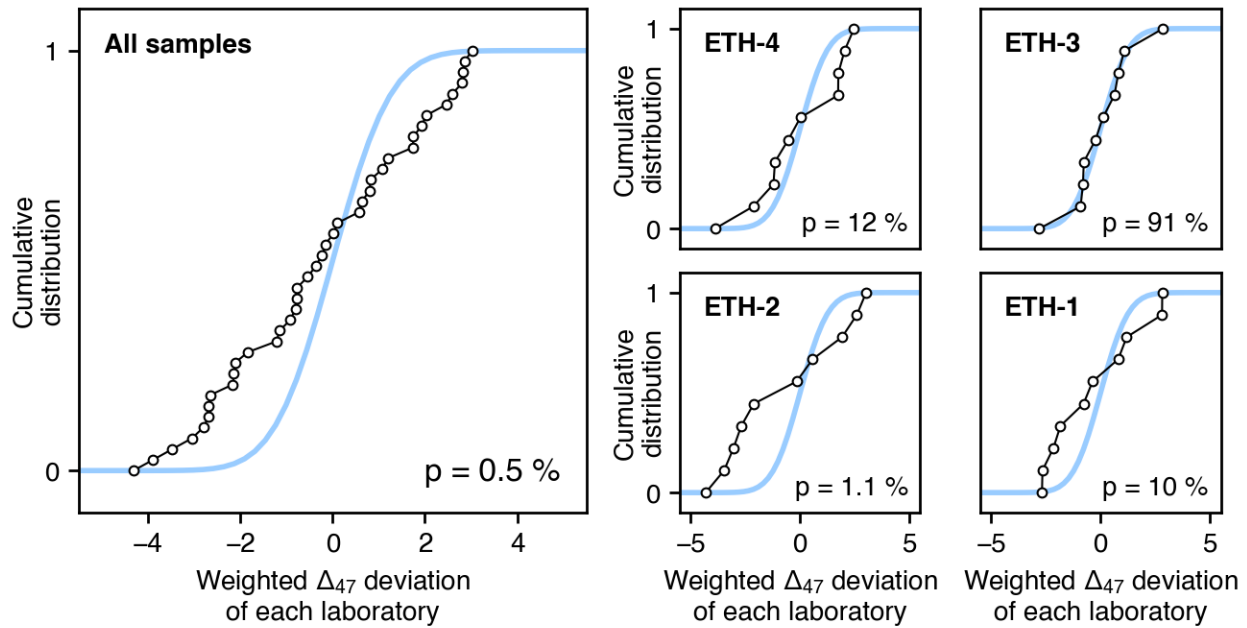
642



644  
 645 **Figure 5.** Kolmogorov–Smirnov tests of normality for the sigma-deviations obtained in each laboratory  
 646 participating in the Intercarb effort (circular markers), either neglecting standardization uncertainty (upper  
 647 row) or considering fully propagated analytical errors (lower row). Lower-right corner Kolmogorov–  
 648 Smirnov p-values correspond to the null hypothesis that the sigma-deviations are normally distributed  
 649 with a mean of zero and a standard deviation of 1. Blue lines correspond to the canonical Gaussian  
 650 distribution ( $\mu = 0$ ;  $\sigma = 1$ ).

651  
 652 On demonstrating that we can fully account for inter-laboratory error using carbonate  
 653 standardization, we revisit the results obtained for ETH1-4 using HG and EG (Fig. 2). Applying  
 654 the same Kolmogorov–Smirnov test of normality yields unambiguous evidence that the inter-  
 655 laboratory scatter observed here, using HG/EG standardization, is significantly greater than  
 656 predicted from known analytical errors alone ( $p = 0.005$ , Fig. 6), contrary to the opposite finding  
 657 for the InterCarb results, using carbonate-based standardization ( $p = 0.19$ , Fig. 5).

658



659

660

661 **Figure 6.** Kolmogorov–Smirnov tests of normality for the sigma-deviations, considering fully propagated  
 662 analytical errors (accounting for uncertainties associated with conversion to the CDES reference frame),  
 663 obtained in each laboratory participating in the ETH-1/2/3/4 determination using HG and EG (circular  
 664 markers). Lower-right corner Kolmogorov–Smirnov p-values correspond to the null hypothesis that the  
 665 sigma-deviations are normally distributed with a mean of zero and a standard deviation of 1. Blue lines  
 666 correspond to the canonical Gaussian distribution ( $\mu = 0$ ;  $\sigma = 1$ ).

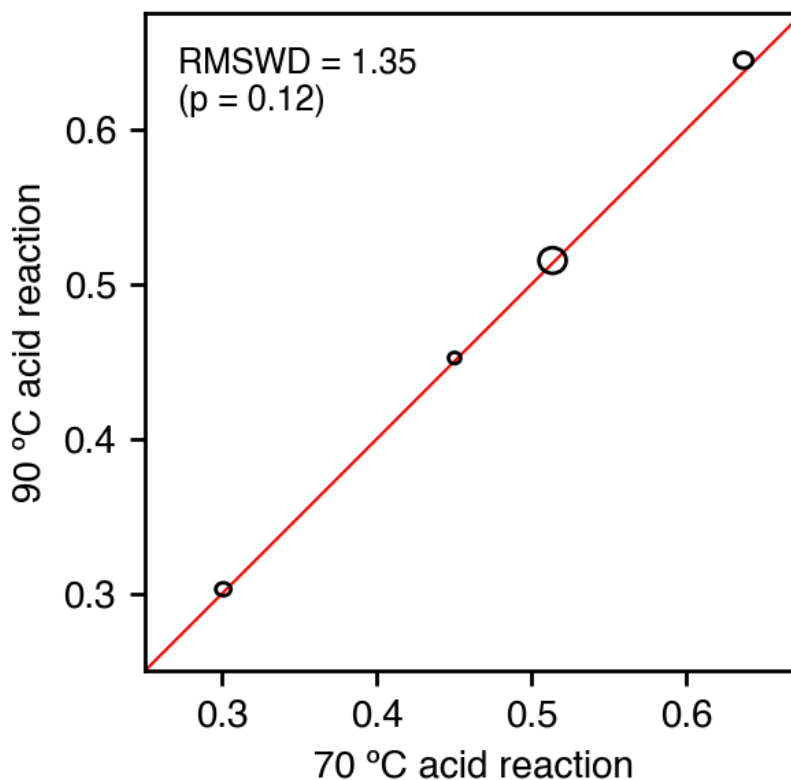
667

### 668 3.3 Effects of acid reaction temperature and IRMS models

669

670 Out of 25 sample preparation systems, 10 convert samples to  $\text{CO}_2$  by acid reactions at  $70^\circ\text{C}$  and  
 671 14 at  $90^\circ\text{C}$ . To test for the possible effect of acid temperature, a commonly discussed cause for  
 672 different slopes in the published temperature calibration curves (Came et al., 2014; Fernandez et  
 673 al., 2014; Swart et al., 2019), we plot the  $\Delta_{47}$  values of unknowns obtained by laboratories  
 674 reacting at  $70^\circ\text{C}$  vs. those obtained at  $90^\circ\text{C}$  (Fig. 8, Table 3). Because acid fractionation effects  
 675 equally affect anchors and unknowns, carbonate-standardized results can be compared directly  
 676 without acid temperature correction.  $\Delta_{47}$  values averaged by acid temperature are statistically  
 677 indistinguishable for all of the unknowns. This implies that relative  $\Delta_{47}$  differences between  $\text{CO}_2$

678 evolved from different samples are independent of acid reaction temperature within the range of  
 679 experimental conditions covered here, and for a very wide range of  $\Delta_{47(I-CDES)}$  values spanning  
 680 0.302 ‰ (marbles) to 0.641 ‰ (carbonates formed at ambient temperatures).  
 681



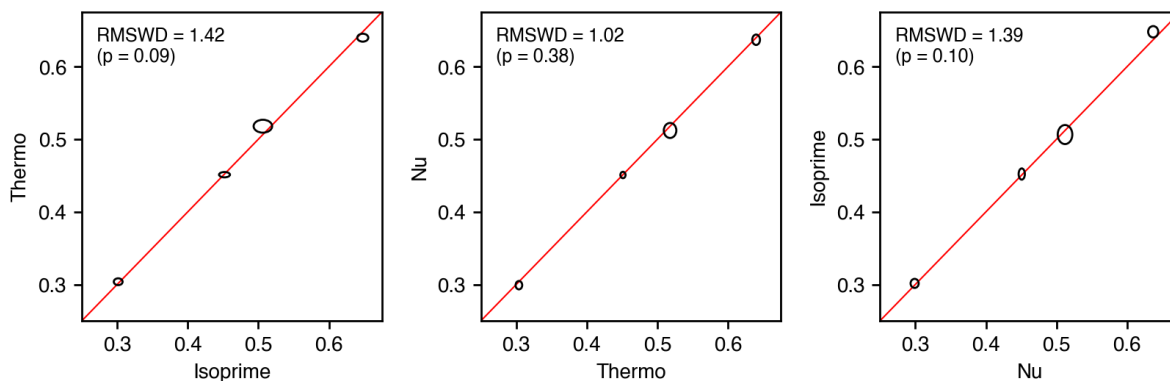
682  
 683  
 684 **Figure 7.** Error-weighted average  $\Delta_{47(I-CDES)}$  values of unknowns obtained from acid reactions at  
 685 90 °C vs 70 °C. Solid black ellipses correspond to 95 % confidence limits. (see also Table 3).  
 686

	70 °C acid	90 °C acid	Difference ( $\pm 1SE$ )
<b>ETH-4</b>	0.4501 $\pm$ 0.0016	0.4521 $\pm$ 0.0015	0.0020 $\pm$ 0.0022
<b>IAEA-C1</b>	0.3006 $\pm$ 0.0020	0.3026 $\pm$ 0.0017	0.0020 $\pm$ 0.0026
<b>IAEA-C2</b>	0.6369 $\pm$ 0.0024	0.6445 $\pm$ 0.0021	0.0076 $\pm$ 0.0032
<b>MERCK</b>	0.5134 $\pm$ 0.0036	0.5151 $\pm$ 0.0034	0.0017 $\pm$ 0.0049
<b>average (all samples)</b>			0.0033 $\pm$ 0.0017

687  
 688 **Table 3.** Error-weighted average  $\Delta_{47(I-CDES)}$  values ( $\pm 1SE$ ) for each unknown as a function of acid reaction  
 689 temperature (see also Fig. 7).

690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709

The error-weighted results separated by mass spectrometer type and design of associated preparation lines, another postulated source of inter-lab disagreement (Swart et al., 2019), are shown in Figure 8 and Table 4. Out of the 22 participating laboratories, 8 use the Nu Perspective, 16 use versions of the Thermo MAT253, and one uses an Isoprime 100. Most results are statistically indistinguishable across instruments. Only IAEA-C2 yielded a significantly ( $>2\sigma$ ) higher mean  $\Delta_{47}$  value ( $\Delta\Delta_{47}$  of +0.0110 and +0.0081 ‰ vs Nu perspective and MAT253, respectively) when analyzed with an Isoprime 100 versus other instrument models (but note that all of the Isoprime 100 data comes from a single laboratory). Inter-instrument differences averaged over all four samples (bottom row of Table 4) remain, however, indistinguishable from zero. Thus, any potential biases introduced by the use of different mass spectrometer models and/or the design of the preparation line which could cause partial equilibration of the produced  $\text{CO}_2$  with the acid and/or heated metal surfaces (Swart et al.2019) are undetectable when using carbonate standardization. Sample sizes used for individual measurements ranged from 90-120  $\mu\text{g}$  for the Kiel IV to  $\sim 500 \mu\text{g}$  for the NuCarb individual acid vial preparation systems, and to 3-12 mg for samples reacted in common acid bath custom-built extraction lines. The fact that small sample measurements are carried out at 70 °C and large ones at 90 °C, also suggests that there is no significant effect of sample sizes and variations in sample to acid ratios in these results.



710  
711  
712  
713

**Figure 8.** Error-weighted average  $\Delta_{47(I-CDES)}$  values of unknowns obtained using different mass spectrometer types. Solid black ellipses correspond to 95 % confidence limits.

	<b>MAT 253 vs Isoprime 100</b>	<b>Nu Perspective vs MAT 253</b>	<b>Isoprime 100 vs Nu Perspective</b>
<b>ETH-4</b>	-0.0009 ± 0.0035	-0.0004 ± 0.0024	0.0013 ± 0.0036
<b>IAEA-C1</b>	0.0023 ± 0.0032	-0.0048 ± 0.0030	0.0025 ± 0.0035
<b>IAEA-C2</b>	-0.0081 ± 0.0039	-0.0029 ± 0.0037	0.0110 ± 0.0043
<b>MERCK</b>	0.0115 ± 0.0065	-0.0059 ± 0.0056	-0.0056 ± 0.0068
<b>average (all samples)</b>	0.0012 ± 0.0022	-0.0035 ± 0.0019	0.0023 ± 0.0024

714

715 **Table 4.** Error-weighted average  $\Delta_{47(i-CDES)}$  values ( $\pm 1SE$ ) for each unknown as a function of mass  
716 spectrometer types.

717 **3.4. Guidelines for improving data quality in the future.**

718

719 The results of this inter-laboratory exercise support the use of carbonate standardization for  
720 clumped isotope measurements. When considering all laboratories, the standard deviation of  
721 laboratory averages for the 4 unknowns range from 0.011 ‰ to 0.018 ‰ for the samples  
722 requiring no to moderate extrapolation, to 0.025 ‰ for MERCK, an extreme case of  
723 extrapolation. The spread across laboratories is still relatively large, and of the same magnitude  
724 as those obtained by HG-EG normalization if we consider either (1) what has been reported on 4  
725 carbonate standards for 4 laboratories in Dennis et al. (2011) or (2) the scatter in the values  
726 reported by the 10 laboratories that provided data for the re-determination of the accepted values  
727 of ETH-1 to ETH-3 in this study. However, based on this study, we can clearly state that the  
728 large scatter is dominated by random errors, and there are strategies to improve the repeatability  
729 within each laboratory.

730

731 The scatter in the InterCarb data (Fig. 4) is especially influenced by laboratories that show the  
732 largest errors in the individual sample reproducibility and have a significantly magnified  
733 normalization error induced by a small number of replicates of anchors. This observation  
734 underscores the necessity of increased replication, ideally in separate sessions, and of the  
735 measurement of a sufficient number of standards to produce data of the quality that is required  
736 for meaningful interpretations. Special care needs to be taken when reconstructing small  
737 temperature changes, such as for the reconstruction of climate change and/or the study of high



738 temperature processes where the small temperature sensitivity of the clumped isotope  
739 thermometer requires highest analytical precision.

740

741 We can, for example, define a target laboratory internal repeatability of 0.010 ‰ (1SE) as a  
742 desirable goal with the currently available analytical systems. This arbitrary target number is  
743 within the shot-noise limits of modern IRMS instruments, and is necessary for applications in  
744 paleoclimate reconstructions, one of the main applications of clumped isotopes, as it corresponds  
745 to an uncertainty of approximately  $\pm 3$  °C. If for each sample we select only the results from the  
746 laboratories that have provided data with a combined error of less than 0.010 ‰ (1SE; Table 2),  
747 inter-laboratory standard deviation (1SD) becomes  $\leq 10$  ppm for ETH-4 (N=22), IAEA-  
748 C1(N=15) and IAEA-C2 (N=13) and  $\leq 0.015$  ‰ for MERCK (N=11; with SE < 0.0135). It can  
749 be noted that this does not significantly change the average value of the reference materials.

750

751 Based on this example, it appears that with modern instrumentation from all manufacturers and  
752 with both custom built and commercially available systems used by the laboratories involved in  
753 this study, it is possible to reach this data quality and inter-laboratory consistency. The main  
754 factor to be taken into consideration is sufficient replication of both sample and standards (see  
755 Daëron, this volume, Bonifacie et al., 2017; Fernandez et al., 2017; Kocken et al., 2019).

756

757 With standardization using carbonates, the principle of identical treatment of samples and  
758 standards (Werner & Brand, 2001b) is fulfilled, in contrast to the use of HG/EG standardization.  
759 Two important outcomes of this study are that acid reaction temperature and instrument and  
760 preparation line design are not a cause for differences among laboratories when standardization  
761 is based on carbonates. The lack of resolvable differences observed in our dataset, indicates that  
762 if preparation line differences affect  $\Delta_{47}$  measurements, standardization with carbonates corrects  
763 any such effects whereas HG-EG standardization may fail to do so (Swart et al. 2019). Thus,  
764 when using carbonate standardization, these factors can be ignored provided the carbonate  
765 standards cover a large range in  $\Delta_{47}$ .

766 As the acid digestion conditions (e.g., reaction times, temperatures) vary with the carbonate  
767 mineralogy, possible effects on  $\Delta_{47}$  might be mineral-specific. For this reason, it would be highly  
768 desirable to produce reference materials for other minerals such as dolomite, aragonite,

769 magnesite and siderite. For dolomite, three samples were proposed by Müller et al. (2019) as  
770 possible reference materials and are available upon request.

771  
772 Standardization errors could be reduced to some extent by extending the range of bulk  
773 composition of the anchor samples (e.g., as illustrated by Fig. 1 of *Daëron*, this volume),  
774 especially when samples are measured that require significant extrapolation. A sample with an  
775 extreme bulk composition like MERCK would be a useful addition as an anchor, regardless of its  
776  $\Delta_{47}$  value. While with a three-anchor system, two heated standards for normalization are not  
777 strictly necessary, a “heated MERCK” anchor in combination with ETH-1 would furthermore  
778 allow verifying the PBL correction with greater confidence and with less replication than with  
779 ETH-2 (keeping in mind that small quadratic components to PBL correction might introduce a  
780 significant bias over a  $\delta^{47}$  range of 60 ‰, e.g., Fig. 7 from He et al., [2012]).

781  
**782 3.5 Reporting data normalized to carbonates: definition of the “InterCarb” Carbon**  
**783 Dioxide Equilibrium Scale (I-CDES)**

784  
785 There is a need for a community consensus on how to report clumped isotope measurements,  
786 both to promote data comparability and to reduce confusion stemming from the different scales  
787 used in the literature. Currently  $\Delta_{47}$  data are reported for different temperatures of phosphoric  
788 acid digestion, mostly projected to 25 °C but also to 70 °C or 90 °C reactions, and generally  
789 labeled respectively as  $\Delta_{47\text{CDES}25}$ ,  $\Delta_{47\text{CDES}70}$  and  $\Delta_{47\text{CDES}90}$ , a terminology introduced in Bonifacie  
790 et al. (2017). In the literature, phosphoric acid correction factors used by different research  
791 groups to convert results from 90 to 25 °C reaction temperatures have varied between 0.069 ‰  
792 (Wacker et al., 2014) to 0.092 ‰ (e.g., Bonifacie et al., 2017), thus representing a significant  
793 source of uncertainty and confusion.

794  
795 The direct standardization to accepted values of solid phases, on the other hand, removes the  
796 need for a phosphoric acid correction, yielding results which are independent of the temperature  
797 at which the samples were reacted. In InterCarb this has only been tested for calcites, the

798 mineralogy of all standards used here. Further studies are necessary in particular for dolomite  
799 and siderite, as these minerals require longer reaction times and there are contrasting findings in  
800 the literature on whether they require different phosphoric acid fractionations (Bonifacie et al.,  
801 2017; Müller et al. 2019, van Dijk et al., 2019; Petersen et al. 2019), possibly affecting the  
802 absolute values of samples with such mineralogies. Nevertheless, the use of the InterCarb  
803 reference carbonates, allows for a consistent correction of instrumental effects and normalization  
804 to the I-CDES, independent of the mineralogy. Possible effects related to the longer reaction time  
805 remain to be tested, and require intercomparison samples of different mineralogy. In any case for  
806 all calcites carbonate standardization eliminates different phosphoric acid correction factors as a  
807 source of uncertainty and provides a consistent framework to report  $\Delta_{47}$  without uncertainties  
808 related to the reaction temperature.

809  
810 For these reasons, we recommend that in the future, carbonate clumped isotope values should be  
811 reported relative to a carbonate reference frame uniquely defined by the absolute  $\Delta_{47}$  values  
812 reported in Table 1 for ETH-1, ETH-2, and ETH-3. Unknown samples may be anchored to this  
813 reference frame either (a) directly by comparison to ETH-1/2/3, (b) indirectly by comparison  
814 with several of the four other carbonate standards used here (Figure 4), or (c) by comparison  
815 with a set of in-house laboratory standards whose composition is well-constrained relative to the  
816 materials reported here. To clearly distinguish this data normalization scheme from previous  
817 ones the denomination I-CDES (InterCarb - Carbon Dioxide Equilibration Scale) should be used,  
818 and, where applicable, the notation  $\Delta_{47(I-CDES)}$ . This approach is analogous to the change from  
819 the PDB to the VPDB scale, which was accomplished by assigning a consensus offset of +1.95  
820 ‰ between the original PDB reference material and the NBS19 carbonate. This was  
821 subsequently, albeit temporarily, improved by defining a second anchor point with the L-SVEC  
822 Lithium carbonate standard (Coplen et al., 2006). We note that because the carbonate  $\Delta_{47}$  values  
823 in Table 1 are firmly anchored to the CDES scale via HG/EG measurements in multiple  
824 laboratories, the two scales are in principle equivalent. However, I-CDES has three major  
825 advantages (1) it follows the principle of equal treatment of sample and standards (2) it removes  
826 uncertainties related to fractionation effects due to different acid reaction temperatures and  
827 designs of the preparation lines and (3) it is based on traceable, stable materials (calcium  
828 carbonates) that are widely available to interested laboratories. . What is more, the results

829 summarized in Fig. 5 imply that I-CDES standardization yields consistent  $\Delta_{47}$  values  
830 independent of laboratory and/or analytical protocols, so that inter-laboratory scatter is  
831 accurately predicted by the fully-propagated analytical uncertainties computed within each  
832 laboratory. By contrast, this does not always seem to be the case for gas-based standardization to  
833 the CDES reference frame (Fig. 6), suggesting the existence of poorly understood sources of  
834 inter-lab discrepancies in that approach. The broad availability of carbonate reference materials  
835 with widely varying bulk and clumped isotope compositions is an important step to help  
836 establish a worldwide equivalence among laboratories, and to help new laboratories establish and  
837 verify their analytical procedures. .

838  
839 The decision that  $\Delta_{47(I-CDES)}$  values are reported for an acid reaction temperature of 90 °C implies  
840 that the I-CDES value of any given sample will be almost 0.1 ‰ lower than  $\Delta_{47CDES25}$  values  
841 which is currently the most common, but not the sole, convention used to report clumped  
842 isotopes. At first sight it may seem a disadvantage to lose the ability to intuitively compare new  
843 results to those obtained through different standardization approaches in the published literature.  
844 However, it will make it immediately obvious that I-CDES-normalized values cannot directly be  
845 compared to data standardized to the ETH-1/2/3/4 values reported by Bernasconi et al. (2018).  
846 To directly compare previous carbonate-normalized  $\Delta_{47}$  values and the already published  
847 calibration equations produced with ETH standard normalisation ((Kele et al, 2015; Bernasconi  
848 et al., 2018; Jautzy et al., 2020; Meinicke et al., 2020; Peral et al., 2018; Piasecki et al., 2019) to  
849 future data reported on the I-CDES scale, previous data have to be recalculated as described in  
850 detail in the Appendix. We emphasize again, however, that the previously reported temperatures  
851 calculated from the original calibrations using the original ETH standard values are directly  
852 comparable to temperatures calculated for samples normalized and calibrations recalculated to  
853 the I-CDES. This is because only the nominal values of the standards have changed, and thus the  
854 data normalization is internally consistent and traceable to the same solid standards.

855  
856 In principle, I-CDES data are directly comparable to data produced by phosphoric acid reaction  
857 at 90°C using the HG/EG approach. With InterCarb, the nominal values of the ETH standards  
858 are robustly linked to the CDES, as they are now based on the average results of 10 laboratories,  
859 and are not only based on the values determined at ETH in 2013. Some laboratories may still

860 want to continue measuring HG and EG to keep established laboratory procedures and/or to  
861 cover ranges in bulk compositions that require large extrapolations. However, the results of  
862 InterCarb, and the discussions in the literature (e.g. Petersen et al. 2019) show that with the  
863 HG/EG approach there are still poorly understood interlaboratory discrepancies (as suggested by  
864 Fig. 6) which are absent in the carbonate-based normalization (see Fig 5). For this reason, it is of  
865 paramount importance that several of the InterCarb reference carbonates are incorporated in the  
866 laboratory procedures to ensure interlaboratory data compatibility. The use of matrix-matched  
867 reference materials is necessary so that delta values can be unambiguously compared on a like-  
868 for-like basis (see Meier-Augenstein and Schimmelmann, 2019 for a recent discussion).

869 For laboratories using large sample common-acid bath methodologies and preferring gas-based  
870 data correction, we recommend that in addition to HG/EG a minimum of two of the InterCarb  
871 reference materials (or in-house standards with values calibrated to ETH standards) should be  
872 measured within the same analytical sessions as the samples and used in the data correction  
873 scheme. We recommend choosing two standards with a large difference in  $\Delta_{47}$ , for example  
874 ETH-1 and ETH-3 or ETH-2 and IAEA-C2, depending on the bulk composition of the unknown  
875 samples (see Fig. 1). In the case of samples very low  $\delta^{47}$  we additionally recommend the use of  
876 MERCK. The comparison of I-CDES data to older HG/EG normalized data, is in principle  
877 possible, with the caveat that in absence of measurements of widely available carbonates, a direct  
878 comparison remains uncertain. A community effort to robustly anchor the composition of legacy  
879 standards measured in the first laboratories, may alleviate this problem in the near future.

880 For laboratories with commercially available single acid vial, small sample preparation systems  
881 we discourage the use of HG/EG, as the gases would most probably be measured under different  
882 conditions than the samples, an approach which is prone to error. In addition, InterCarb shows  
883 that carbonate-based standardization can provide robust and accurate data without the use of  
884 gases.

885

### 886 **3.6. Outlook for future improvements**

887

888 InterCarb has shown that with carbonate standardization it is possible to reach an inter-laboratory  
889 reproducibility that is as good as what can be reached within a single laboratory. The use of

890 carbonate standards allows for a better monitoring of the performance of the entire preparation  
891 system, including acid reaction and sample purification. Regular and systematic measurement of  
892 carbonate reference material distributed within a run or measurement interval is a prerequisite to  
893 correct for short- and long-term variations in the analytical systems. In this study, analyses were  
894 grouped in measurement intervals, and all data were processed assuming no short-term variation  
895 in the instrumentation. However, especially with "small sample approaches" relying on short  
896 (~30-45 min) measurements of many replicates, one can observe short-term variations  
897 (Bernasconi et al. 2018, Fig. 4). Thus a moving window correction with variable window size  
898 may be desirable in these cases.

899  
900 Replicate sample measurements should be spread in time as much as possible. The number and  
901 distribution of standards in a run are also important parameters that can improve reproducibility  
902 and reduce errors. This has been discussed in detail by Kocken et al. (2019) who concluded that  
903 carbonate standards with bulk and clumped-isotope compositions similar to those of unknowns  
904 should be analyzed with greater frequency than the other anchors, while preserving a minimal  
905 level of replication for each anchor.

906  
907 For InterCarb, all distributed aliquots of IAEA-C1, IAEA-C2 and MERCK originated from  
908 single bottles. The results suggest that after milling, these carbonates were homogeneous within  
909 these bottles, but we strongly recommend verifying that additional bottles purchased from IAEA  
910 and MERCK are identical to the ones tested here. Finally, as there is currently no consensus in  
911 the literature on whether different carbonate mineralogies exhibit the same phosphoric acid  
912 fractionation effects (e.g. Müller et al. 2017; van Dijk et al. 2019; Bonifacie et al. 2017, Petersen  
913 et al. 2019) and on possible differences in the temperature-dependence of the phosphoric acid  
914 fractionation for dolomite (Defliese et al. 2015; Murray et al. 2016), there should be a  
915 community effort to identify and characterize reference materials for other carbonates such as  
916 aragonite and siderite, similar to the dolomite reference materials proposed by Müller et al.  
917 (2019).

#### 918 **4 Conclusions**

919

- 920
- 921
- 922
- 923
- 924
- 925
- 926
- 927
- 928
- 929
- 930
- 931
- 932
- 933
- 934
- 935
- 936
- 937
- 938
- 939
- 940
- 941
- 942
- 943
- 944
- 945
- 946
- 947
- 948
- 949
- 950
- This study demonstrates that carbonate-based standardization of clumped isotope measurements solves many open questions so far limiting the application of carbonate clumped isotope thermometry as a mature and reliable tool in Earth sciences.
  - Carbonate-based standardization is robust in spite of the smaller range of isotopic compositions compared to the heated/equilibrated gas approach, even when extrapolation is necessary..
  - Inter-laboratory discrepancies observed in this study among 22 laboratories are not greater than those predicted from intra-laboratory analytical uncertainties.
  - We propose a set of 7 widely available calcite reference materials for normalization of carbonate clumped isotope measurements with new community-accepted values: the 4 ETH standards, two samples distributed by the IAEA (C1 and C2), and a synthetic carbonate produced by MERCK.
  - Standardization with ETH-1, ETH-2 and ETH-3 provides a robust framework for converting carbonate clumped isotope data to the CDES using reference materials matching the sample matrix and, thus, following the principle of equal treatment of samples and standards.
  - Data standardized with this method should be reported as I-CDES to explicitly signify that the data were normalized using these reference materials or a combination of HG/EG and at least two of the reference carbonates.
  - Carbonate standardization removes the need to apply an acid digestion fractionation factor, eliminating uncertainties due to poorly known acid fractionation factors and different preparation systems and thus reduces differences between laboratories.
  - The use of a reference material with an extreme bulk isotope composition such as MERCK is recommended, as the large range of compositions reduces standardization errors for samples with “exotic” bulk compositions (as previously suggested by some authors but precisely quantified in Daëron, this volume).
  - Data expressed in the I-CDES are directly comparable to samples reacted at a temperature of 90°C normalized to the CDES with HG/EG. However, we emphasize that carbonate standardisation is preferred to pure HG/EG

951 normalization because it is based on traceable carbonate samples that can be  
 952 measured in every laboratory. Reporting the measured compositions of carbonate  
 953 reference materials together with the samples is the only way to ensure inter-  
 954 laboratory consistency.

- 955 • Robust standardization of clumped isotope measurements requires the analysis of  
 956 a sufficient number of replicates of both samples and standard materials (either  
 957 gases or carbonates) alongside unknowns to minimize error and obtain accurate  
 958 measurements within a single laboratory.
- 959 • Replicate analyses should be distributed in time and ideally span different  
 960 correction intervals
- 961 • When reporting data, error should include both replication error and reference  
 962 frame error in proper statistical data treatment and reporting of uncertainties

963 **Appendix:**

964 **A. Converting older, carbonate-anchored  $\Delta_{47}$  values to the I-CDES**

965  
 966 This section describes the steps necessary to convert existing  $\Delta_{47}$  measurements to the I-CDES,  
 967 provided that they were either standardized using carbonate anchors or analyzed simultaneously  
 968 with several carbonate anchors. This mathematically exact approach is a simpler alternative to  
 969 fully reprocessing the original raw data (with the caveat that this conversion will not provide, by  
 970 itself, any error estimates).

971  
 972 By way of example, let us consider measurements originally standardized using ETH-1/2/3 with  
 973 the nominal  $\Delta_{47}$  values reported by Bernasconi et al. (2018). The “old”  $\Delta_{47}$  values of these  
 974 measurements are noted  $^{\text{old}}\Delta_{47}$ , and we wish to compute the “new”  $\Delta_{47}$  values, noted  $^{\text{new}}\Delta_{47}$ , that  
 975 would be obtained if the same data were standardized to the I-CDES.

976  
 977 Both old and new  $\Delta_{47}$  values are derived from the same set of raw measurements using affine  
 978 relationships of the form:

979

$$\Delta_{47}^{\text{old}} = x_1 + y_1\delta^{47} + z_1\Delta_{47}^{\text{raw}}$$



980 (A.1)

981

982 (A.2)  $\Delta_{47}^{new} = x_2 + y_2\delta^{47} + z_2\Delta_{47}^{raw}$

983

984 We can rearrange the above equations to express  $\Delta_{47}^{new}$  as

985 a function of  $\delta^{47}$  and  $\Delta_{47}^{old}$  :

986 (A.3)  $\Delta_{47}^{new} = a + b\delta^{47} + c\Delta_{47}^{old}$

987

988

989 Computing the numerical values of (a,b,c) is thus all that is required to compute  $\Delta_{47}^{new}$  for any

990 sample whose ( $\delta^{47}$ ,  $\Delta_{47}^{old}$ ) values are known. This computation only requires knowing ( $\delta^{47}$ ,  $\Delta_{47}^{old}$

991 ,  $\Delta_{47}^{new}$ ) for three different anchors, for instance ETH-1/2/3. In matrix form, eq. (A.3) then

becomes:

993

994 (A.4) Thus: 
$$\begin{pmatrix} 1 & \delta_{(ETH1)}^{47} & \Delta_{47(ETH1)}^{old} \\ 1 & \delta_{(ETH2)}^{47} & \Delta_{47(ETH2)}^{old} \\ 1 & \delta_{(ETH3)}^{47} & \Delta_{47(ETH3)}^{old} \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \Delta_{47(ETH1)}^{new} \\ \Delta_{47(ETH2)}^{new} \\ \Delta_{47(ETH3)}^{new} \end{pmatrix}$$

996

997 (A.5) 
$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 & \delta_{(ETH1)}^{47} & \Delta_{47(ETH1)}^{old} \\ 1 & \delta_{(ETH2)}^{47} & \Delta_{47(ETH2)}^{old} \\ 1 & \delta_{(ETH3)}^{47} & \Delta_{47(ETH3)}^{old} \end{pmatrix}^{-1} \begin{pmatrix} \Delta_{47(ETH1)}^{new} \\ \Delta_{47(ETH2)}^{new} \\ \Delta_{47(ETH3)}^{new} \end{pmatrix}$$

998

1000

1001 In this example, the  $\Delta_{47}^{old}$  and  $\Delta_{47}^{new}$  values are the old and new nominal values of ETH-1/2/3. If

1002 the  $\delta^{47}$  values are defined as usual by reference to a working gas, (a,b,c) will vary when different

1003 working gases are used. However, the above equations remain valid if  $\delta^{47}$  is defined instead by

1004 reference to a fixed, hypothetical CO<sub>2</sub> composition, e.g., stochastic VPDB-CO<sub>2</sub> ( $\delta^{13}C_{VPDB} = 0$  ‰;

1005  $\delta^{18}O_{VSMOW} \approx 41.5$  ‰;  $\Delta_{47} = 0$  ‰,  $R^{47} = 4.834 \times 10^{-5}$ ). In that case, numerical values of the

1006 parameters (a,b,c) can be determined once and then used to compute  $\Delta_{47}^{new}$  for any unknown

1007 sample based only on its  $\Delta_{47}^{old}$  and  $\delta^{47}$  values (the latter being defined relative to VPDB-CO<sub>2</sub> and

1008 computed by applying an acid <sup>18</sup>O/<sup>16</sup>O fractionation factor of 1.01025 to said sample).

1009 Importantly, this particular transformation applies to all data sets originally standardized in the  
1010 reference frame defined by the “old” nominal values of ETH-1/2/3.

1011

1012 For instance, the relationship linking the ETH-1/2/3 reference frame of Bernasconi et al. (2018)  
1013 to the I-CDES is defined by:

1014

1015

1016 (A.6)

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 & 0.010 & 0.258 \\ 1 & -28.375 & 0.258 \\ 1 & 0.538 & 0.691 \end{pmatrix}^{-1} \begin{pmatrix} 0.2052 \\ 0.2085 \\ 0.6132 \end{pmatrix} = \begin{pmatrix} -0.037997 \\ -0.000182 \\ 0.942590 \end{pmatrix}$$

1017

1018

1019 Thus:

1020

$$(A.7) \quad \Delta_{47}^{\text{new}} = -0.037997 - 0.000182 \delta^{47} + 0.942590 \Delta_{47}^{\text{old}}$$

1022

1023

1024

1025

1026 In this case, it is clear that the conversion from the old reference frame to the new one is not very  
 1027 sensitive to  $\delta^{47}$  values: for unknown samples with  $\delta^{47}$  values within  $\pm 6$  ‰ of ETH-1 (or  
 1028 stochastic VPDB-CO<sub>2</sub>), the effect of the second term in Eq. (A.7) is less than 1 ppm, and less  
 1029 than 3 ppm for unknowns within  $\pm 18$  ‰ of ETH-1. In such cases, the conversion may be  
 1030 simplified as a simpler affine transformation, akin to a more traditional two-point normalization  
 1031 (e.g., VSMOW-VSLAP standardization):

1032

$$(A.8) \quad \Delta_{47}^{\text{new}} = 0.942590 \Delta_{47}^{\text{old}} - 0.037997$$

1034

1035 For instance, to convert the  $\Delta_{47}^{\text{old}}$  value of ETH-4 reported by Bernasconi et al. (2018) to the I-  
 1036 CDES, we only need to know that  $\delta^{47}_{(\text{ETH4})} = -28.8$  ‰ and  $\Delta_{47}^{\text{old}} = 0.507 \pm 0.004$  ‰. The  $\Delta_{47}^{\text{new}}$   
 1037 value predicted by eq. (A.7) is then  $0.445 \pm 0.004$  ‰, to be compared with the independently  
 1038 constrained values reported here in Table 1 ( $0.450 \pm 0.002$  ‰) and Table 2 ( $0.451 \pm 0.001$  ‰).

1039

1040 The above computation could also be performed using any arbitrary set of three carbonate  
 1041 materials whose  $(\delta^{47}, \Delta_{47}^{\text{old}}, \Delta_{47}^{\text{new}})$  values are known, provided that they span a wide enough  
 1042 range in  $\delta^{47}$  and  $\Delta_{47}$ . This is true even if the carbonates in question were not originally used to  
 1043 standardize the raw data, as would be the case for CO<sub>2</sub>-standardized measurements. In the case  
 1044 where only two suitable carbonate standards X and Y are available, an acceptable approach

1045 would be to neglect effects (equivalent to setting the value of  $b$  to zero in eq. (A.3)) by solving  
 1046 the following equation:

1047

1048 (A.9) 
$$\begin{pmatrix} a \\ c \end{pmatrix} = \begin{pmatrix} 1 & \Delta_{47(X)}^{\text{old}} \\ 1 & \Delta_{47(Y)}^{\text{old}} \end{pmatrix}^{-1} \begin{pmatrix} \Delta_{47(X)}^{\text{new}} \\ \Delta_{47(Y)}^{\text{new}} \end{pmatrix}$$

1049

1050

1051

## 1052 **Acknowledgments, Samples, and Data**

1053 SMB acknowledges instrumentation funding from ETH Zürich and support from Swiss National  
 1054 Science Foundation grants 200021\_143485, 200020\_160046. MD acknowledges the clumped-  
 1055 isotope facility at LSCE is part of PANOPLY (Plateforme Analytique Géosciences Paris-Saclay)  
 1056 and was supported by the following institutions: Région Ile-de-France; Direction des Sciences de  
 1057 la Matière du Commissariat à l’Energie Atomique; Institut National des Sciences de l’Univers,  
 1058 Centre National de la Recherche Scientifique; Université de Versailles/Saint-Quentin-en-  
 1059 Yvelines. KDB and the MIT carbonate clumped isotope facility acknowledge support from the  
 1060 Packard Foundation, the Agouron Foundation and NASA Exobiology Grant 80NSSC19K0464.  
 1061 MB acknowledges the program Emergences Ville de Paris for early funding that allowed  
 1062 building her clumped isotope laboratory at Institut de Physique du Globe de Paris. MB also  
 1063 thanks IPGP and Université de Paris for support for the organization of the VIth Clumped  
 1064 Isotope Workshop in Paris in 2017. Funding was provided to CJ and the Imperial College  
 1065 Clumped Isotope Lab by Qatar Petroleum, Shell, and the Qatar Science and Technology Panel.  
 1066 Funding was provided to KH at University of Washington from NSF EAR grants 1933130,  
 1067 1713275, and 1156134. TV acknowledges the support of the Swiss National Science Foundation  
 1068 grant 206021-164032. ANM acknowledges funding from the Trond Mohn Foundation and the  
 1069 European Research Council (ERC) under the European Union's Horizon 2020 research and  
 1070 innovation programme (grant agreement no. 638467). Funding was provided to the Tripathi Lab at  
 1071 UCLA for these analyses from DOE BES grant DE-FG02-13ER16402. Funding was provided to  
 1072 the Hyland Lab at NCSU for these analyses from NSF EAR-FRES grant 1925973. Funding was  
 1073 provided to the Yoshida Lab at Tokyo Tech. for these analyses from JSPS grant  
 1074 JP17H06105. Funding was provided to the CLIMB Lab at the University of Bergen for these

1075 analyses by the European Research Council (ERC) under the European Union's Horizon 2020  
1076 research and innovation programme (grant agreement No 638467) and by the Trond Mohn  
1077 Foundation. Funding was provided to the UU Clumped Lab at Utrecht University by the  
1078 Netherlands Research Organization (NWO) through VIDI grant 016.161.365. Funding was  
1079 provided to the CUBES-SIL at CU Boulder by startup funds from the University of Colorado  
1080 Boulder and NSF EAR grant 1524785. The UM SCIPP Lab was supported by startup funds from  
1081 the University of Michigan. JR Kelson was supported by NSF PRF grant 1854873.

## 1082 **Data**

1083 All data analyzed as part of this study, along with all code used to process them, are available at  
1084 <https://github.com/mdaeron/InterCarb> and <http://doi.org/10.5281/zenodo.4314449>.

1085

## 1086 **Author contributions**

1087 The community designed the study at the Paris 2017 International Clumped Isotope Workshop.  
1088 S.M.B. prepared materials, gathered results and wrote the first draft of the manuscript. M.D.  
1089 created all code and conducted all analysis of the results. S.M.B., M.D., K.D.B., N.M., and M.B.  
1090 organized the effort and contributed to early drafts. All authors contributed to the analyses and  
1091 curation of data shared in this study. All authors contributed to the final version of the  
1092 manuscript.

1093

## 1094 **References**

1095

1096 Adlan, Q., Davies, A.J., John, C.M., 2020. Effects of oxygen plasma ashing treatment on  
1097 carbonate clumped isotopes, *Rapid Communications in Mass Spectrometry*, DOI:

1098 <https://doi.org/10.1002/rcm.8802>

1099 Affek, H. P., & Eiler, J. M. (2006). Abundance of mass 47 CO<sub>2</sub> in urban air, car exhaust,  
1100 and human breath. *Geochimica et Cosmochimica Acta*, 70(1), 1–12.

1101 <https://doi.org/10.1016/j.gca.2005.08.021>

1102 Anderson N.T., Kelson J.R., Kele S., Daëron M., Bonifacie M., Mackey T.J., John C.M.,

1103 Huntington K.W., Bernasconi S.M. and Bergman K.D. (2020) Reanalyzing carbonate

- 1104 clumped isotope calibration materials using robust carbonate standardization. AGU  
 1105 Fall Meeting 2020 Abstracts
- 1106 Bernasconi, S. M., Hu, B., Wacker, U., Fiebig, J., Breitenbach, S. F. M., & Rutz, T. (2013).  
 1107 Background effects on Faraday collectors in gas-source mass spectrometry and  
 1108 implications for clumped isotope measurements. *Rapid Communications in Mass  
 1109 Spectrometry*, 27(5), 603–612. <https://doi.org/10.1002/rcm.6490>
- 1110 Bernasconi, S. M., Müller, I. A., Bergmann, K. D., Breitenbach, S. F. M. M., Fernandez, A.,  
 1111 Hodell, D. A., et al. (2018). Reducing uncertainties in carbonate clumped isotope  
 1112 analysis through consistent carbonate-based standardization. *Geochemistry,  
 1113 Geophysics, Geosystems*, 19(9), 1–20. <https://doi.org/10.1029/2017GC007385>
- 1114 Bonifacie, M., Calmels, D., Eiler, J. M., Horita, J., Chaduteau, C., Vasconcelos, C., et al.  
 1115 (2017). Calibration of the dolomite clumped isotope thermometer from 25 to 350 °C,  
 1116 and implications for a universal calibration for all (Ca, Mg, Fe)CO<sub>3</sub> carbonates.  
 1117 *Geochimica et Cosmochimica Acta*, 200, 255–279.  
 1118 <https://doi.org/10.1016/j.gca.2016.11.028>
- 1119 Brand, W. a., Assonov, S. S., & Coplen, T. B. (2010). Correction for the 17O interference in  
 1120  $\delta(13C)$  measurements when analyzing CO<sub>2</sub> with stable isotope mass spectrometry  
 1121 (IUPAC Technical Report). *Pure and Applied Chemistry*, 82(8), 1719–1733.  
 1122 <https://doi.org/10.1351/PAC-REP-09-01-05>
- 1123 Came, R. E., Brand, U., & Affek, H. P. (2014). Clumped isotope signatures in modern  
 1124 brachiopod carbonate. *Chemical Geology*, 377, 20–30.  
 1125 <https://doi.org/10.1016/j.chemgeo.2014.04.004>
- 1126 Carter, J. F., & Fry, B. (2013). Ensuring the reliability of stable isotope ratio data - Beyond  
 1127 the principle of identical treatment. *Analytical and Bioanalytical Chemistry*, 405(9),  
 1128 2799–2814. <https://doi.org/10.1007/s00216-012-6551-0>
- 1129 Daëron, M. (submitted) Full propagation of analytical uncertainties in  $\Delta_{47}$  measurements.  
 1130 Submitted to *Geochemistry, Geophysics, Geosystems (Companion Paper)*
- 1131 Daëron, M., Blamart, D., Peral, M., & Affek, H. P. (2016). Absolute isotopic abundance  
 1132 ratios and the accuracy of  $\Delta_{47}$  measurements. *Chemical Geology*, 442, 83–96.  
 1133 <https://doi.org/10.1016/j.chemgeo.2016.08.014>

- 1134 Dale, A., John, C. M., Mozley, P. S., Smalley, P. C., & Muggeridge, A. H. (2014). Time-  
 1135 capsule concretions: Unlocking burial diagenetic processes in the Mancos Shale using  
 1136 carbonate clumped isotopes. *Earth and Planetary Science Letters*, 394.  
 1137 <https://doi.org/10.1016/j.epsl.2014.03.004>
- 1138 Defliese, W. F., & Lohmann, K. C. (2015). Non-linear mixing effects on mass-47  
 1139 CO<sub>2</sub> clumped isotope thermometry: Patterns and implications. *Rapid*  
 1140 *Communications in Mass Spectrometry*, 29(9), 901–909.  
 1141 <https://doi.org/10.1002/rcm.7175>
- 1142 Defliese, W. F., Hren, M. T., & Lohmann, K. C. (2015). Compositional and temperature  
 1143 effects of phosphoric acid fractionation on D47 analysis and implications for  
 1144 discrepant calibrations. *Chemical Geology*, 396, 51–60.  
 1145 <https://doi.org/10.1016/j.chemgeo.2014.12.018>
- 1146 Dennis, K. J., Affek, H. P., Passey, B. H., Schrag, D. P., & Eiler, J. M. (2011). Defining an  
 1147 absolute reference frame for “clumped” isotope studies of CO<sub>2</sub>. *Geochimica et*  
 1148 *Cosmochimica Acta*, 75(22), 7117–7131. <https://doi.org/10.1016/j.gca.2011.09.025>
- 1149 Eagle, R. a, Schauble, E. a, Tripathi, A. K., Tütken, T., Hulbert, R. C., & Eiler, J. M. (2010).  
 1150 Body temperatures of modern and extinct vertebrates from (13)C-(18)O bond  
 1151 abundances in bioapatite. *Proceedings of the National Academy of Sciences of the*  
 1152 *United States of America*, 107(23), 10377–10382.  
 1153 <https://doi.org/10.1073/pnas.0911115107>
- 1154 Fernandez, A., Tang, J., & Rosenheim, B. E. (2014). Siderite ‘clumped’ isotope  
 1155 thermometry: A new paleoclimate proxy for humid continental environments.  
 1156 *Geochimica et Cosmochimica Acta*, 126, 411–421.  
 1157 <https://doi.org/10.1016/j.gca.2013.11.006>
- 1158 Fernandez, A., Müller, I. A., Rodríguez-Sanz, L., van Dijk, J., Looser, N., & Bernasconi, S.  
 1159 M. (2017). A Reassessment of the Precision of Carbonate Clumped Isotope  
 1160 Measurements: Implications for Calibrations and Paleoclimate Reconstructions.  
 1161 *Geochemistry, Geophysics, Geosystems*, 18(12).  
 1162 <https://doi.org/10.1002/2017GC007106>

- 1163 Ferry, J. M., Passey, B. H., Vasconcelos, C., & Eiler, J. M. (2011). Formation of dolomite at  
 1164 40-80 °C in the Latemar carbonate buildup, Dolomites, Italy, from clumped isotope  
 1165 thermometry. *Geology*, 39(6), 571–574. <https://doi.org/10.1130/G31845.1>
- 1166 Fiebig, J., Bajnai, D., Löffler, N., Methner, K., Krsnik, E., Mulch, A., & Hofmann, S.  
 1167 (2019). Combined high-precision  $\Delta_{48}$  and  $\Delta_{47}$  analysis of carbonates. *Chemical*  
 1168 *Geology*, 522 (Feb), 186–191. <https://doi.org/10.1016/j.chemgeo.2019.05.019>
- 1169 Ghosh, P., Adkins, J., Affek, H., Balta, B., Guo, W., Schauble, E. A., et al. (2006). 13C-  
 1170 18O bonds in carbonate minerals: A new kind of paleothermometer. *Geochimica et*  
 1171 *Cosmochimica Acta*, 70(6), 1439–1456. <https://doi.org/10.1016/j.gca.2005.11.014>
- 1172 Ghosh, P., Garzzone, C., & Eiler, J. (2006). Rapid Uplift of the Altitplano Revealed Through  
 1173 13C-18O Bonds in Paleosol Carbonates. *Science*, 311(January), 511–515.  
 1174 <https://doi.org/10.1126/science.1119365>
- 1175 Grauel, A. L., Schmid, T. W., Hu, B., Bergami, C., Capotondi, L., Zhou, L., & Bernasconi,  
 1176 S. M. (2013). Calibration and application of the “clumped isotope” thermometer to  
 1177 foraminifera for high-resolution climate reconstructions. *Geochimica et Cosmochimica*  
 1178 *Acta*, 108, 125–140. <https://doi.org/10.1016/j.gca.2012.12.049>
- 1179 Guo, W., & Eiler, J. M. (2007). Temperatures of aqueous alteration and evidence for  
 1180 methane generation on the parent bodies of the CM chondrites. *Geochimica et*  
 1181 *Cosmochimica Acta*, 71(22), 5565–5575. <https://doi.org/10.1016/j.gca.2007.07.029>
- 1182 Guo, W., Mosenfelder, J. L., Goddard, W. A., & Eiler, J. M. (2009). Isotopic fractionations  
 1183 associated with phosphoric acid digestion of carbonate minerals: Insights from first-  
 1184 principles theoretical modeling and clumped isotope measurements. *Geochimica et*  
 1185 *Cosmochimica Acta*, 73(24), 7203–7225. <https://doi.org/10.1016/j.gca.2009.05.071>
- 1186 He, B., Olack, G. A., & Colman, A. S. (2012). Pressure baseline correction and high-  
 1187 precision CO<sub>2</sub> clumped-isotope ( $\Delta_{47}$ ) measurements in bellows and micro-volume  
 1188 modes. *Rapid Communications in Mass Spectrometry*, 26(24), 2837–2853.  
 1189 <https://doi.org/10.1002/rcm.6436>
- 1190 Hu, B., Radke, J., Schlüter, H. J., Heine, F. T., Zhou, L., Bernasconi, S. M., et al. (2014). A  
 1191 modified procedure for gas-source isotope ratio mass spectrometry: The long-  
 1192 integration dual-inlet (LIDI) methodology and implications for clumped isotope



- 1193 measurements. *Rapid Communications in Mass Spectrometry*, 28(13), 1413–1425.  
 1194 <https://doi.org/10.1002/rcm.6909>
- 1195 Huntington, K. W., Eiler, J. M., Affek, H. P., Guo, W., Bonifacie, M., Yeung, L. Y., et al.  
 1196 (2009). Methods and limitations of “clumped” CO<sub>2</sub> isotope ( $\Delta 47$ ) analysis by gas-  
 1197 source isotope ratiomass spectrometry. *Journal of Mass Spectrometry*, 44(9), 1318–  
 1198 1329. <https://doi.org/10.1002/jms.1614>
- 1199 Huntington, K. W., Budd, D. A., Wernicke, B. P., & Eiler, J. M. (2011). Use of Clumped-  
 1200 Isotope Thermometry To Constrain the Crystallization Temperature of Diagenetic  
 1201 Calcite. *Journal of Sedimentary Research*, 81(9), 656–669.  
 1202 <https://doi.org/10.2110/jsr.2011.51>
- 1203 Jautzy, J. J., Savard, M. ., Dhillon, R. ., Bernasconi, S. ., Lavoie, D., & Smirnoff, A. (2021).  
 1204 Clumped isotope temperature calibration for calcite: Bridging theory and  
 1205 experimentation. *Geochemical Perspective Letters*, (2020), 36–41.  
 1206 <https://doi.org/10.7185/geochemlet.2021>
- 1207 Katz A., Bonifacie M., Hermoso M., Cartigny P., Calmels D. (2017) Laboratory-grown  
 1208 coccoliths exhibit no vital effect in clumped isotope ( $\Delta_{47}$ ) composition on a range of  
 1209 geologically relevant temperatures in *Geochimica Cosmochimica Acta*, 208, 335–353.  
 1210 <https://doi.org/10.1016/j.gca.2017.02.025>
- 1211 Kele, S., Breitenbach, S. F. M., Capezzuoli, E., Meckler, A. N., Ziegler, M., Millan, I. M.,  
 1212 Kluge, T., Deák, J., Hanselmann, K., John, C. M., Yan, H., Liu, Z., & Bernasconi, S.  
 1213 M. (2015). Temperature dependence of oxygen- and clumped isotope fractionation in  
 1214 carbonates: A study of travertines and tufas in the 6-95°C temperature range.  
 1215 *Geochimica et Cosmochimica Acta*, 168. <https://doi.org/10.1016/j.gca.2015.06.032>
- 1216 Kelson, J. R., Huntington, K. W., Schauer, A. J., Saenger, C., & Lechler, A. R. (2017).  
 1217 Toward a universal carbonate clumped isotope calibration: Diverse synthesis and  
 1218 preparatory methods suggest a single temperature relationship. *Geochimica et*  
 1219 *Cosmochimica Acta*, 197, 104–131. <https://doi.org/10.1016/j.gca.2016.10.010>
- 1220 Kim, S. T., Coplen, T. B., & Horita, J. (2015). Normalization of stable isotope data for  
 1221 carbonate minerals: Implementation of IUPAC guidelines. *Geochimica et*  
 1222 *Cosmochimica Acta*, 158, 276–289. <https://doi.org/10.1016/j.gca.2015.02.011>

- 1223 Kluge, T., John, C. M., Jourdan, A.-L., Davis, S., & Crawshaw, J. (2015). Laboratory  
 1224 calibration of the calcium carbonate clumped isotope thermometer in the 25-250°C  
 1225 temperature range. *Geochimica et Cosmochimica Acta*, 157, 213-227.  
 1226 <https://doi.org/10.1016/j.gca.2015.02.028>
- 1227 Kocken, I. J., Müller, I. A., & Ziegler, M. (2019). Optimizing the Use of Carbonate  
 1228 Standards to Minimize Uncertainties in Clumped Isotope Data. *Geochemistry,*  
 1229 *Geophysics, Geosystems*, 20(11), 5565–5577. <https://doi.org/10.1029/2019GC008545>
- 1230 Mangenot X., Gerdes A., Gasparrini M., Bonifacie M., Rouchon V. (2018). An emerging  
 1231 thermo-chronometer for carbonate-bearing rocks :  $\Delta_{47}/(U-Pb)$ . *Geology* 46 (12), 1067-  
 1232 1070.
- 1233 Massey, F. J. (1951). The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of the*  
 1234 *American Statistical Association*, 46(253), 68–78.  
 1235 <https://doi.org/10.1080/01621459.1951.10500769>
- 1236 Meckler, A. N., Ziegler, M., Millan, M. I., Breitenbach, S. F. M., Bernasconi, S. M., Millán,  
 1237 M. I., et al. (2014). Long-term performance of the Kiel carbonate device with a new  
 1238 correction scheme for clumped isotope measurements. *Rapid Communications in Mass*  
 1239 *Spectrometry*, 28(15), 1705–1715. <https://doi.org/10.1002/rcm.6949>
- 1240 Meier-Augenstein, W., & Schimmelmann, A. (2019). A guide for proper utilisation of  
 1241 stable isotope reference materials \*. *Isotopes in Environmental and Health Studies*,  
 1242 55(2), 113–128. <https://doi.org/10.1080/10256016.2018.1538137>
- 1243 Meinicke, N., Ho, S. L., Hannisdal, B., Nürnberg, D., Tripathi, A., Schiebel, R., & Meckler,  
 1244 A. N. (2020). A robust calibration of the clumped isotopes to temperature relationship  
 1245 for foraminifers. *Geochimica et Cosmochimica Acta*, 270, 160–183.  
 1246 <https://doi.org/10.1016/j.gca.2019.11.022>
- 1247 Müller, I. A., Fernandez, A., Radke, J., Dijk, J. Van, Bowen, D., Schwieters, J., et al.  
 1248 (2017). Carbonate clumped isotope analyses with the long-integration dual-inlet (LIDI)  
 1249 workflow: scratching at the lower sample weight boundaries. *Rapid Communications*  
 1250 *in Mass Spectrometry*, 31(12), 1057–1066. <https://doi.org/10.1002/rcm.7878>
- 1251 Müller, I. A., Violay, M. E. S., Storck, J.-C. C., Fernandez, A., van Dijk, J., Madonna, C., &  
 1252 Bernasconi, S. M. (2017). Clumped isotope fractionation during phosphoric acid

- 1253 digestion of carbonates at 70 °C. *Chemical Geology*, 449, 1–14.
- 1254 <https://doi.org/10.1016/j.chemgeo.2016.11.030>
- 1255 Müller, I. A., Rodriguez-Blanco, J. D., Storck, J.-C., do Nascimento, G. S., Bontognali, T.
- 1256 R. R., Vasconcelos, C., et al. (2019). Calibration of the oxygen and clumped isotope
- 1257 thermometers for (proto-)dolomite based on synthetic and natural carbonates.
- 1258 *Chemical Geology*, 525(December 2018), 1–17.
- 1259 <https://doi.org/10.1016/j.chemgeo.2019.07.014>
- 1260 Murray, S. T., Arienzo, M. M., & Swart, P. K. (2016). Determining the  $\Delta 47$  acid
- 1261 fractionation in dolomites. *Geochimica et Cosmochimica Acta*, 174, 42–53.
- 1262 <https://doi.org/10.1016/j.gca.2015.10.029>
- 1263 Passey, B. H., & Henkes, G. A. (2012). Carbonate clumped isotope bond reordering and
- 1264 geospeedometry. *Earth and Planetary Science Letters*, 351–352, 223–236.
- 1265 <https://doi.org/10.1016/j.epsl.2012.07.021>
- 1266 Passey, B. H., Levin, N. E., Cerling, T. E., BROWN, F. H., Eiler, J. M., & Turekian, K. K.
- 1267 (2010). High-temperature environments of human evolution in East Africa based on
- 1268 bond ordering in paleosol carbonates. *Proceedings of the National Academy of*
- 1269 *Sciences of the United States of America*, 107(25), 11245–11249.
- 1270 <https://doi.org/10.1073/pnas.1001824107>
- 1271 Peral, M., Daëron, M., Blamart, D., Bassinot, F., Dewilde, F., Smialkowski, N., et al.
- 1272 (2018). Updated calibration of the clumped isotope thermometer in planktonic and
- 1273 benthic foraminifera. *Geochimica et Cosmochimica Acta*, 239, 1–16.
- 1274 <https://doi.org/10.1016/j.gca.2018.07.016>
- 1275 Petersen, S. V., Defliese, W. F. F., Saenger, C., Daëron, M., Huntington, K. W. W., John, C.
- 1276 M. M., et al. (2019). Effects of Improved 17 O Correction on Inter-laboratory
- 1277 Agreement in Clumped Isotope Calibrations, Estimates of Mineral-Specific Offsets,
- 1278 and Temperature Dependence of Acid Digestion Fractionation. *Geochemistry,*
- 1279 *Geophysics, Geosystems*, 20, 3495–3519. <https://doi.org/10.1029/2018GC008127>
- 1280 Petersen, Sierra V., Winkelstern, I. Z., Lohmann, K. C., & Meyer, K. W. (2016). The effects
- 1281 of Porapak<sup>TM</sup> trap temperature on  $\delta 18 O$ ,  $\delta 13 C$ , and  $\Delta 47$  values in preparing samples
- 1282 for clumped isotope analysis. *Rapid Communications in Mass Spectrometry*, 30(1),
- 1283 199–208. <https://doi.org/10.1002/rcm.7438>

- 1284 Piasecki, A., Bernasconi, S. M., Grauel, A., Hannisdal, B., Ho, S. L., Leutert, T. J., et al.  
1285 (2019). Application of Clumped Isotope Thermometry to Benthic Foraminifera.  
1286 *Geochemistry, Geophysics, Geosystems*, 2018GC007961.  
1287 <https://doi.org/10.1029/2018GC007961>
- 1288 Schauer, A. J., Kelson, J., Saenger, C., & Huntington, K. W. (2016). Choice of 17O  
1289 correction affects clumped isotope ( $\delta^{47}$ ) values of CO<sub>2</sub> measured with mass  
1290 spectrometry. *Rapid Communications in Mass Spectrometry*, 30(24), 2607–2616.  
1291 <https://doi.org/10.1002/rcm.7743>
- 1292 Schmid, T. W., & Bernasconi, S. M. (2010). An automated method for “clumped-isotope”  
1293 measurements on small carbonate samples. *Rapid Communications in Mass*  
1294 *Spectrometry*, 24(14), 1955–1963. <https://doi.org/10.1002/rcm.4598>
- 1295 Spooner, P. T., Guo, W., Robinson, L. F., Thiagarajan, N., Hendry, K. R., Rosenheim, B.  
1296 E., & Leng, M. J. (2016). Clumped isotope composition of cold-water corals: A role  
1297 for vital effects? *Geochimica et Cosmochimica Acta*, 179, 123–141.  
1298 <https://doi.org/10.1016/j.gca.2016.01.023>
- 1299 Swart, P. K., Murray, S. T., Staudigel, P. T., & Hodell, D. A. (2019). Oxygen Isotopic  
1300 Exchange between CO<sub>2</sub> and Phosphoric Acid: Implications for the Measurement of  
1301 Clumped Isotopes in Carbonates. *Geochemistry, Geophysics, Geosystems*, (1950),  
1302 2019GC008209. <https://doi.org/10.1029/2019GC008209>
- 1303 Thaler C., Katz A., Bonifacie M., Ménez B., Ader M. Oxygen isotope composition of waters  
1304 recorded in carbonates in strong clumped and oxygen isotopic disequilibrium (2020).  
1305 *Biogeosciences*, 17, 1–14, 2020. DOI 10.5194/bg-17-1-2020
- 1306 van Dijk, J., Fernandez, A., Storck, J. C., White, T. S., Lever, M., Müller, I. A., et al.  
1307 (2019). Experimental calibration of clumped isotopes in siderite between 8.5 and 62°C  
1308 and its application as paleo-thermometer in paleosols. *Geochimica et Cosmochimica*  
1309 *Acta*, 254, 1–20. <https://doi.org/https://doi.org/10.1016/j.gca.2019.03.018>
- 1310 Veillard, C. M. A., John, C. M., Krevor, S., & Najorka, J. (2019). Rock-buffered  
1311 recrystallization of Marion Plateau dolomites at low temperature evidenced by  
1312 clumped isotope thermometry and X-ray diffraction analysis. *Geochimica et*  
1313 *Cosmochimica Acta*, 252, 190–212. <https://doi.org/10.1016/J.GCA.2019.02.012>

- 1314 Wacker, U., Fiebig, J., & Schoene, B. R. (2013). Clumped isotope analysis of carbonates:  
 1315 comparison of two different acid digestion techniques. *Rapid Communications in Mass*  
 1316 *Spectrometry : RCM*, 27(14), 1631–42. <https://doi.org/10.1002/rcm.6609>
- 1317 Wacker, U., Fiebig, J., Tödter, J., Schöne, B. R., Bahr, A., Friedrich, O., et al. (2014).  
 1318 Empirical calibration of the clumped isotope paleothermometer using calcites of  
 1319 various origins. *Geochimica et Cosmochimica Acta*, 141, 127–144.  
 1320 <https://doi.org/10.1016/j.gca.2014.06.004>
- 1321 Wang, Z., Schauble, E. A., & Eiler, J. M. (2004). Equilibrium thermodynamics of multiply  
 1322 substituted isotopologues of molecular gases. *Geochimica et Cosmochimica Acta*,  
 1323 68(23), 4779–4797. <https://doi.org/10.1016/j.gca.2004.05.039>
- 1324 Werner, R. A., & Brand, W. A. (2001a). Referencing strategies and techniques in stable  
 1325 isotope ratio analysis. *Rapid Communications in Mass Spectrometry*, 15(7), 501–519.  
 1326 <https://doi.org/10.1002/rcm.258>
- 1327 Werner, R. A., & Brand, W. A. (2001b). Referencing strategies and techniques in stable  
 1328 isotope ratio analysis. *Rapid Communications in Mass Spectrometry*, 15(7), 501–519.  
 1329 <https://doi.org/10.1002/rcm.258>