Dear Associate editor, Dr. Axel Gerdes

We appreciate all comments made by the reviewers and by you, they indeed helped to clarify and improve this manuscript. We wish to submit our revised version of the manuscript with all changes made by us, highlighted in the manuscript body. We specify in this letter point-by-point replies to comments of both reviewers.

Review #1

We thank the reviewer for his comments and suggestions for improving this manuscript. Please see below our response and corrections made for each comment.

Comment: "I think that the manuscript, as it stands, suffers from a somewhat selective referencing of the literature. ...While the present manuscript certainly provides a more detailed analysis of relevant analytical issues I think it is important not to lose sight of the fact that previous work has been conducted in this area."

Response: As suggested by the reviewer, we included in the introduction and further discussed references on previous U–Pb dating of dolomite, for both bulk and in-situ analyses:

Lines 33-37: "Previous U–Pb dating of dolomites on whole-rock samples of U-rich dolostones, conducted in the highest level of cleanroom standards, yielded scattered results along the isochron (Winter and Johnson, 1995; Hoff et al., 1995; Ovchinnikova et al., 2007; Polyak et al., 2016). These studies suggested that in-situ dating of dolomites should be feasible, and indeed several studies recently reported on successful in-situ age determination of dolomites using the LA-ICP-MS methodology (Burisch et al., 2018; Salih et al., 2019; Hu et al., 2020; Mueller et al., 2020)."

Comment: "It is of course disappointing, but nevertheless important, that a number of the samples used in this study provide ages which are seemingly inconsistent with known stratigraphic relationships. I think that this part of the manuscript in particular would benefit from some further thought/exploration of potential mechanisms."

Response: In the revision, we improved the discussion on possible mechanisms for inconsistent age results. However, we think that our data is preliminary and in order to make such far-fetched geological conclusions, a more detailed geochemistry study is required.

Lines 349-355: "The use of CL imaging can help to establish how homogeneous the samples are in terms of precipitation conditions. Micritic material are very hard to study by simple microscopy and slight differences in luminescence may suggest superimposed precipitation events. In such cases, early events that left very small remnant material, but with high U content, and a later dominant event with low U-content, can easily produce mixed age that is shifted towards and old ages. In such cases, it might be useful to implement the methodology described in Drost et al. (2018), in which 2-D elemental and isotopic ratio maps are used for targeting subdomains in carbonate samples with complex geological histories, such as diagenetic overprinting."

Comment: "The discussion of crater morphology is perhaps least convincing. ...Unfortunately, this is not a consistent observation in this study and many of the determined ages are in fact younger than anticipated, not older (with the exception of the two syngenetic Cretaceous samples). ...I can see how crater roughness might equate to ablation inefficiency but how does this translate into an age bias rather than simply larger age uncertainties? It would be very useful to have more discussion here..."

Response: This section of the manuscript was significantly improved. Following reviewer's comment, the morphology of the craters was studied in more detail and is presented in Fig. 5 with additional images. In the revision we show representative crater images for each sample and a classification of the samples according to the quality of their craters. Following the observations made by Guillong et al. (this issue), we were able to correlate observations on crystal size and ablation efficiency with crater morphology and resulted ages. We also better demonstrate that micritic samples tend to produce older ages than expected, up to 20% off towards older ages than expected (see lines 307-315 in the revised paper). Lines 219-236: "Most craters display relatively similar patterns, but consistently, sparry dolomites tend to ablate much better than micritic dolomites. Dolomitic rocks with crystal size >10 µm behave similarly to one another, with smooth bottoms and only minor imperfections along the rims. Although only the age of sample MU-2 is consistent with its stratigraphic appearance, many of the determined ages are younger than anticipated. However, according to the morphology of their analytical craters, these younger ages might be true ages and may reflect late-stage dolomitization. On the other hand, samples Tm-MU-2 and Tm-DV-1, which contain elevated Pb and REE concentrations, tend to produce moderate roughness along crater bottom and some imperfections along crater rims. Nevertheless, the poorly constrained ages of these samples seem to be related to their trace-element chemistry, rather than differences in crater morphology. Crater bottom morphology of micritic dolomites tend to be exceedingly rough and crater rims display numerous imperfections. Samples MAM-3 and MAM-7 are composed of <5 µm crystals and display anomalously rough crater morphology, with deeper and rougher bottoms. These samples were expected to produce Cretaceous ages between 100 to 90 Ma but resulted in much older ages of 137 ± 14 and 173 ± 11 Ma. According to their crater morphology, the older ages in these samples seem to be the result of differences in ablation volume and the amount of material introduced to the plasma. Guillong et al. (this issue) suggest that differences in crystal size may result in ablation rate of up to 160% higher relative to WC-1, with age offsets of 4–8 % for dolomite. Variations in crater size between RM and unknowns for identical laser parameters might shift $^{238}U/^{206}Pb$ ratios to lower values and older ages (Guillong et al., this issue). Differences in particle size and volume of mass introduced to the ICP-MS torch may significantly affect the resulted ages and lead to overestimation of the standard correction. Such differences must be considered when defining a matrix-matched standard for dolomite.

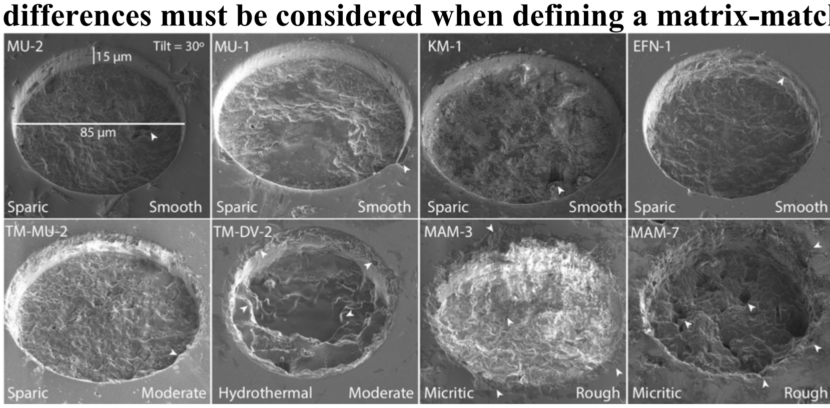

Figure 5: Ablation craters of studied samples, arranged by crater geometry and bottom roughness, from smooth (MU-2, sparic dolomite), via moderate (TM-MU-2, sapric dolomite), to rough (MAM-7, micritic dolomite). Uneven mass removal from crater rims are marked by white arrows.”

Comment: "It is also argued that mineralogical/textural controls may results in mixed ages but, once again, the evidence provided does not seem to back up these assertions. The inclusion of remnant (pre-dolomitisation) calcite grains (section 3.3) in the analysis would surely bias the ages towards the existing stratigraphic constraints, not make them younger than expected?"

Response: Indeed, in some samples the resulted ages are younger than the stratigraphic ages and in two samples they are older. We rephrased this section and clarified that the younger ages are the result of mixing between the stratigraphic age (residual calcite) and epigenetic dolomite overgrowths. We argue that it is difficult under the current analytical resolution to distinguish between the stratigraphic age and a later dolomitization event.

Lines 290-293: “These clusters are probably remnants of primary calcite that was later replaced by dolomite (Fig. 8C, D). The WDS mapping could be therefore used for detecting zoning and remnant calcite impurities in the dolomite sample, which in case of late dolomitization event(s), might shift the determined age towards the stratigraphic age of the sample. It is therefore highly recommended to use WDS elemental mapping for samples with sparry grains.”

Comment: "I can’t help but wonder in all of this if many of these younger ages are in fact analytically just fine – and simply reflect the time of closure during late-stage dolomitization ie. it is the existing interpretation of the timing of dolomitization (not stratigraphic age) that is incorrect?"

Response: Since the age of ~55 Ma repeats over different samples, we could just interpret it as the age of a true dolomitization event. However, we also mentioned that there are no supportive geological evidences for a dolomitization event at that time. A follow-up study is needed to proclaim either this is the actual time of dolomitization or a mixture between the stratigraphic and a much younger epigenetic dolomitization event.

All other comments regarding text clarification were addressed accordingly.

Review #2

We thank the reviewer for his comments and suggestions for improving this manuscript. Please find our responses for each comment.

Comment: " I don’t really understand the interest of REE, which don’t bring much to the discussion. The spectra are identical to each other, and close to those expected for marine carbonates. I suggest to remove that part."

Response: We pointed out that the main take-home message from the REE data is that in spite of similarity in patterns, relatively high REE content might indicate anomalously high common-Pb

values and lower chances for successful age. REE analysis in carbonates can be easily done and can help to screen out such samples.
Lines 172-175: "This is also supported by the REE signature of sample Tm-MU-2, showing enrichment in LREE, depletion in HREE and positive Gd anomaly (Figure 4B). This pattern is similar to other dolomite samples in this study, although one order of magnitude higher, suggesting that dolomitic rocks in association with hydrothermal activity contain high common Pb concentrations, and thus, low-chance for successful dating."

Comment: "Also, it would be nice to have higher magnifications for thin section photomicrographs (not "photomicrographic") (Figure 1), as any detail can be hardly seen here."

Response: We provided higher magnification of thin section photomicrographs to Figure 1 and correct the text accordingly.

Comment: "...I am wondering why there are so much difference in the calcite/dolomite ratios calculated between XRD and EBSD?..."

Response: This is a very good question that we are not really sure why. We speculate that the crystallographic parameters used in EBSD interpretation programs are different from those in XRD.

Comment: "Did you perform simple analyses such as optical cathodoluminescence
(CL)? This can be useful to detect several cement generations or recrystallization events on a single carbonate sample and can be done easily in all Earth science labs."

Response: We agree and we have provided CL images of the samples in the revised version of our paper. We added the CL and PPL images on an additional figure (Fig. 2) and included explanation in section 2.1. Lines 99-100:" Cathodoluminescence images of representative carbonate material, used to infer slight changes in fluid composition (e.g. Mn^{2+}, Fe^{2+} content), and/or precipitation conditions, are presented in Fig. 2."
Lines 121-123:" The zoned hydrothermal dolomite grains of sample Tm-DV-1 are slightly zoned under CL (Fig. 2) with very similar luminescence suggesting minimal changes in fluid composition and/or precipitation conditions."
Lines 138-139:" The non-homogenized luminescence of the sparry sample KM-1 (Fig. 2) may indicate a possible mixture of phases that precipitated under different conditions."
Lines 145-147:" The bright luminescence of the cement material of sample EFN-1 suggest a single phase of precipitation that is distinctively different from precipitation conditions of the fragment material."

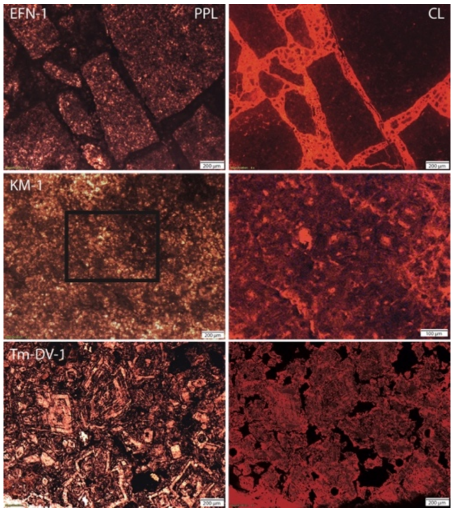

Figure 2. PPL images (left panels) and Cathodoluminescence (CL) images (right panels) of representative studied samples. Note the differences in CL colors of breccia fragments and cement in Sample EFN-1, the non-homogenize CL response in sample KM-1 and the slight zoned dolomite crystals in sample Tm-DV-1.

Comment: "...Have you tried to make spot analyses on polished slabs instead of thin sections?"

Response: We did not try spot analyses on polished slabs, only on 100 µm thick sections; we expect thick sections and polished slab to respond similarly to incoming laser radiation.

Comment: "...It can be expected to find ages younger than the stratigraphic age (as you also mention in the text). This does not mean that age is wrong, the question being to know which event is dated in a complex geological history..."

Response: Perhaps we did not clarify this properly in the text. For the non-stratigraphic ages, we argue that the date may truly represent the age of a geologic event, but it can also be a mixture of stratigraphic age and an unknown, much younger age related to a dolomitization event. In such a case, one would expect to observe several isochrons superimposed one another as we noted. We rephrase this section in lines 332-338: "
The age of sample MU-2 (93 ±7) is correspond to the expected stratigraphic age for this unit and probably represent early diagenesis. In sample MU-1, on the other hand, a wedge pattern similar to the fragments in sample EFN-1 can be identified. Out of 80 spot analyses, the older 13 dates form a reasonable isochron with age of 91 ± 6 Ma, with MSWD of 1.8. This age falls within the expected stratigraphic age and probably represents early diagenetic age for that sample. The youngest 38 spot analyses yield an age of 53 ± 2 Ma, with MSWD of 2. The older isochron corresponds to the expected stratigraphic age of this sample, while the younger isochron is ~30 Ma younger and may reflect either the time of closure during late-stage dolomitization, or mixed age between stratigraphic age and a much younger dolomitization event (Fig. 10)."

Comment: "I don't know if you have analyzed other dolomite samples than those presented here, but you only present results on micritic or sparitic dolomites which clearly have, according to your EBSD analyses, complex histories involving the presence of mixed calcite / dolomite mineralogy."

Response: Several of our samples (MU-1/2) were previously described in a *Science* paper as primary dolomite (or as syngenetic/penecontemporaneous, i.e. forming by diagenetic replacement of limestone immediately after the deposition of the stratum), and yet, some of the samples did not give stratigraphic age.

Comment: "TW plots: I read on the figures that you always have positive error correlations. I would expect to have also negative ones, notably in a TW diagram?"

Response: This is an interesting point; error correlations in LA data are difficult to assess because of the method by which uncertainties are propagated. We included in lines 74-75: "Error correlations are calculated following Schimdtz and Schoene, 2007."

Other comments on the text were addressed accordingly.

We also provided coordination of each sample to Table 1 and modified Figure 9 (Fig. 8 in the previous version) to better clarify our discussion on down-hole fractionation (section 3.4).

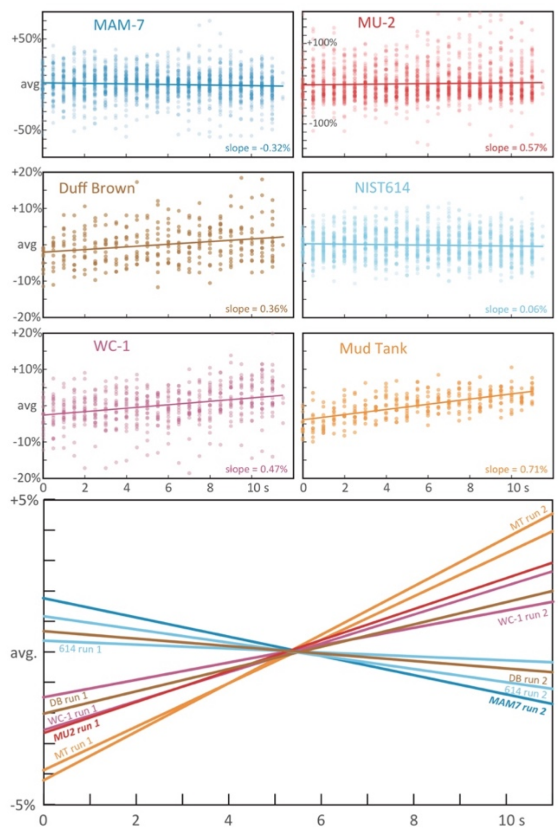

Figure 9. Average down-hole fractionation of RMs and select unknowns. Raw ^{207}Pb-corrected values (corrected for baseline) are normalized to the average value and a linear fit shows different fractionation trends between glass, zircon, calcite and dolomite. Lower panel showing the difference in average down-hole fractionation between unknown samples and reference materials in two different analytical runs.