We thank Reviewer 4 for their thoughtful comments on our manuscript, and give our responses below. Reviewer comments are in italics.

The reproducibility of trace elemental ratios presented in this study for six individual sample sets between ASE-treated and non-ASE treated sample sets is remarkable, but the authors somewhat oversell the reproducibility for Li/Ca and Mg/Ca. For both elemental ratios, for two out of six (i.e., one third) of the samples the reproducibility is in fact not good. The authors argue for smaller sample sizes (256-257, 265-267) but this is not apparent from the plots. As long as no other geochemical criterion can be identified to distinguish a reliable from an unreliable elemental ratio for ASE-treated samples such elemental ratios from ASE-treated samples should not be used in such a manner, and the authors should make this clear in the manuscript. The other ratios are either remarkably (!) good (B/Ca, Na/Ca, Sr/Ca) or satisfactory (U/Ca).

We appreciate the reviewer sharing their thoughts on this. A large range in Mg/Ca variability between foraminifera grown under the same conditions is well documented, for one example see the ~1 mmol/mol spread in values of samples of 20-30 *G. ruber* individuals at ~28 C from plankton tows in the Mozambique Channel (Weldeab et al., 2014). Spread such as this are likely driven by large interspecimen variability in Mg/Ca. Such variability was demonstrated by Rongstad et al. (2017), who performed single-foraminifera analyses of samples consisting of between 66-70 individuals across three species of foram (*G. ruber, N. dutertrei, P. obliquiloculata*), 9 samples in total. The spread in Mg/Ca values for individual foraminifera that they found for each sample ranged from 1.92 to 4.31 mmol/mol, with the biweight standard deviation (which reduces the effect of outliers) ranging from 0.37 to 0.83. Given this, it makes sense that we see an offset in Mg/Ca between our samples of *O. universa*, numbering 13 and 16 individuals.

We also note that for Li/Ca, there is no consistency in terms of the direction of the offset. Of the two samples which have an offset greater than the uncertainty on the measurement, one is observed in an ASE samples, and one in a pre-ASE sample.

Feedback from a previous reviewer is that the order of samples within the tables and plots should follow a consistent sequence, and for the ease of the reader, we agree with this
approach. However, we suggest that we indicate on the plot which of the sample pairs includes a sample which consists of <70 individuals, and hope that this will be acceptable to the reviewer. We are happy to expand on these points about interspecimen heterogeneity and the impact when analysing small samples in the revised manuscript.

No reductive cleaning was carried out. While I understand why the authors did not carry out this step (avoidance of signal bias for Mg/Ca as well as sample loss) this renders the Mn/Ca signal at the very least ambivalent if not pointless. The authors discuss the features of the Mn/Ca results but should not place too much weight on recovered ratios. It is an interesting observation that ASE treatment did not affect the sample Mn/Ca for five out of six sample sets, but such a ratio should probably not be interpreted in a palaeoceanographic fashion anyways since the largest proportion of the Mn in the signal is likely derived from Fe-Mn oxyhydroxides attached to the foraminifera. I also have reservations towards the argument that Fe-Mn oxide phases are not detectable in SEM images. These may simply be too fine-scaled (284-285) to be detected with SEM imaging. If a sediment contains Fe-Mn oxyhydroxides as well as authigenic Mg-containing mineral phases I would not expect a correlation (285) since these are two separate properties of the sediment with independent origin.

We thank the reviewer for raising this important point, and agree that this section would benefit from rewording. We lay out our main points here, and trust that incorporating these into this section will address the reviewer’s concerns.

We agree that as a palaeo-signal, the Mn/Ca is not of palaeo-proxy interest given the lack of a reductive clean. We also tried to avoid placing too much weight on these ratios given the presence of the unusual pink crystalline growths shown in Figure 1 which are potentially linked to the elevated Mn/Ca in samples from core 926B. We appreciate that we failed to make the (potential) link between the pink crystals/authigenic coating and elevated Mn/Ca, which we will amend in our revised manuscript.

Our main argument for a carbonate phase hosting Mn rather than Fe-Mn oxyhydroxides in the samples from core 926B is the remarkable agreement between the Mn measurements (coupled with the authigenic coating observed under SEM, and visible crystalline growths), and the fact that there is this agreement within the species pairs but not between them. We believe that this is unlikely to occur in these measurements without a constant ratio of Mn:Ca, such as would be found in a carbonate, and would be unlikely in the case of Fe-Mn oxyhydroxides. It is plausible that different geometry or starting geochemistry (organics or trace elements) in the foram species might explain the consistent offset between *G. menardii* and *T. trilobus* from 926B.

We do however suspect Fe-Mn oxyhydroxides may be the source of high Mn in the samples from 1406B, based on i) the elevated Mn/Ca, ii) the offset between these samples and iii) the lack of any visible contamination either under light microscope or SEM; as the reviewer states, these may not be detectable using these tools.

The reviewer raises an important point regarding our interpretation of the Mg/Ca:Mn/Ca crossplot, and we thank them for bringing this to our attention. As we have assumed a single secondary phase, we had not considered that there might be issues with exploring this using a crossplot. However, we do feel that there is utility in the cross-plot; this shows clearly that while Mn might be significantly elevated, Mg/Ca remains reproducible within the species pairs and within the range we would expect for a primary signal, though differs between species (as would be expected by the occupation of different ecological niches; an investigation into these was considered but we decided is outside the scope of this paper). We are happy to amend this discussion to take account of the reviewer’s comment and include our interpretation of these data in the discussion.
The reproducibility of the boron isotopic results is outstanding and a key result of this study. On the other hand, the discussion around it is really short. Is there really nothing to discuss? The authors should at the very minimum also include a d11B vs B/Ca plot. They should also mention that the triple treatment of the ASE samples indeed had no (coupled) isotopic or B/Ca effect. This is not obvious from a theoretical point of view given that boron is volatile and the likelihood of a fraction of boron escaping the carbonate matrix during this pre-treatment is not zero and could hence be mentioned.

We did not feel the need to elaborate further on the boron isotope results, as we observed no effect of the treatment on them and we have submitted this paper as a Technical Note. It is indeed an interesting point that the lack of an ASE impact suggests this treatment really does not release any boron from the carbonate matrix to be volatilised, though this is consistent with the apparent good preservation state of the foraminiferal shells following treatment and boron’s incorporation into the CaCO3 lattice. We are happy to include an additional plot and the discussion points outlined, but are concerned about overly expanding the length of the manuscript and so leave this to the Editor’s discretion.

Minor comments:

Lines 24-27: Why not add two good example references behind the usage of every proxy/parameter?

We thank the reviewer for this suggestion, and will include these.

Line 33: “the 10 Myr residence time of boron in seawater presents a challenge for determining absolute ocean pH values on multi-millennial timescales” – add brief explanation as to why this is the case. It is not immediately evident to every reader.

We can add a brief statement of elaboration e.g. as changes in foram d11B will be a function of both changes in pH and in the d11B of seawater. We also reference papers that describe this effect in detail for those who are interested to pursue the details further. We thank the reviewer for drawing our attention to this sentence, as we have noticed an error: this should read multi-million year timescales.

Line 34-35: “while phytoplankton-based proxies may struggle to capture low-CO2 conditions” - add brief explanation as to why this is the case. It is not immediately evident to every reader.

We are happy to add a brief further explanation for this point.

Line 58: CH2Cl2/MeOH sounds like a mixture of a molecular formula and an ingredient name. The molecular formula would be C2H6Cl2O, the name one of these: DCM methanol, methanol DCM, CH2Cl2 methanol, methanol CH2Cl2, dichloromethane MeOH. So how would this reagent officially be addressed?

We thank the reviewer for pointing this out, and propose to amend this to DCM/MeOH.

Also line 58: I find it remarkable that triple sample treatment at 100°C (how long actually?) does not affect the boron isotopic composition of the carbonates. This needs mentioning in the discussion (see major point above).

Each static cycle took 10 minutes, so a total of 30 minutes was spent at 100°C. We note that no such impact has been noted during oxidative cleaning which typically occurs at rather elevated temperatures itself (here, 80 °C for 15 minutes). As mentioned above, the lack of boron release at these temperatures is consistent with its lattice bound position.
Table 1: Please also add the depth downcore of each sample used. Furthermore, it would be useful to know whether these sedimentary depths were positioned below a possible sulphate-methane transition zone (if present at these sites). I am mentioning this since sediments within the methane stability field may contain authigenic carbonates (e.g., Meister et al., 2007, Sedimentology) which could have an effect on the nature and robustness of an extracted stable isotope (B, C or O) or trace element signal. The authors for example mention the possible presence of an authigenic phase in lines 165-167. Could the presence of authigenic carbonates for example have consequences for some of the non-reproducible Mg/Ca or Li/Ca in the sample set?

Thank you for raising this point. Based on this and feedback from a previous reviewer, we propose to amend the sample table to give the Leg, site, hole, core, section, interval, and depth of the cores. The depth of the samples is summarised here:

1168A 25X = 229.2 m
1168A 26X = 238.8 m
1406B = 63.04 m
926B = 185.41 m

The details of the sulfate-methane transition zone (SMTZ) of the cores used here are as follows. The SMTZ it is found at around 225-230 mbsf at ODP Site 1168A (http://www-odp.tamu.edu/publications/189_IR/chap_03/chap_03.htm), so 25X would be at around the SMTZ, and 26X immediately below it. However, the 3 samples from this site had the lowest Mn analysed, and no visibly identified evidence of authigenic carbonates.

For 1406B there is no analysis of methane or sulfate for that site, only for 1406A, but based on stratigraphic correlations between both sites, the SMTZ would be at around 161 m (approx) in 1406A, and in an equivalent depth at 1406B (http://publications.iodp.org/proceedings/342/107/107_f26.htm). Therefore our sampled depth sits above the SMTZ.

For 926B sulfate concentrations decrease by nearly 70% over the sampled sequence at 926B (down to 591.25 mbsf), but sulfate is never fully reduced (http://www-odp.tamu.edu/publications/154_IR/VOLUME/CHAPTERS/ir154_05.pdf)

There appears to be no relationship between the location of the SMTZ and the samples with elevated Mn. It could potentially be a factor in the offsets observed in 1168, given the proximity of these samples to the SMTZ, although we do not observe any indication of an authigenic phase in these samples, and believe the offsets here are much more likely to be due to interspecimen heterogeneity, as laid out above.

We are happy to include a reference to the SMTZ in the text, although we feel that a detailed investigation into the cause of the authigenic phases lies outside the scope of this paper.

Table 1: I also do not understand why for some samples more individuals have been analysed than were apparently weighed in. This does not make sense to me given the explanation in the text.

In the text it is mentioned that further specimens were picked for T. trilobus and G. miotumida; however, we thank the reviewer for drawing our attention to the fact that this point could be made more clearly, and will amend our final manuscript to do so.
Fig. 1: Very interesting figure and feature of these dark crystal overgrowths.

We thank the reviewer for their interest in this point; we agree that this is an intriguing feature and hope that the publication of this paper might stimulate interest in them.

*Lines 107/108: Five seconds is a very short exposure time of foraminifera shells to ultrasonication during cleaning! But the Al/Ca measured on cleaned foraminifera sound encouraging.*

We agree that this is a short period of time, and are also encouraged by the low Al/Ca.

*Line 123: better write “triple quadrupole” than “QQQ”.*

Thank you for pointing this out, we will amend it in our revised manuscript.

*Lines 127-129: Do these reproducibilities represent 1 SD or 2 SD?*

We thank the reviewer for drawing our attention to this, the reproducibilities here are 2SD. We will edit our revised manuscript to include this.

Regarding the uncertainty on Mg, we have noticed an error in the reporting of this. We typically collect data for 24Mg and 25Mg as standard, and report the data for 24Mg, given the greater number of interferences on 25Mg, and that 24Mg tends to be more robust to switches in detector mode, and more stable long term as a result. However, we have noticed that in this case we had erroneously reported the error on 25Mg (0.89 %) instead of 24Mg (1.97 %), and we will amend this in the revised manuscript. The differences between the collected datasets are very small (<0.03 mmol/mol) and this will have no impact on the findings of our paper, but we wanted to ensure that this mistake is noted and corrected at this stage.

*Line 135: Please add molarity of ammonium acetate buffer. What was ammonium acetate buffered with at which concentration?*

Boron isotope column chemistry depends on adjusting the pH of the sample, which is dissolved in 0.5 M HNO₃. Amberlite resin sorbs borate ion, and releases boric acid; therefore, to capture the sample in the column, manipulation of the sample chemistry is necessary. This is achieved here by buffering the sample with a pH 6, 1.1M ammonium hydroxide:1.2 M acetic acid buffer (exact concentrations adjusted to achieve pH ~6). We are happy to amend our methods to include this information.

*Line 143: I am surprised and impressed that the authors still managed to obtain a decent isotopic signal (after blank correction) from such a low-concentration boron solution.*

We thank the reviewer for recognising the quality of the data obtained here. We have a manuscript in preparation that details the methodological developments that have permitted reproducibility at this level.