Review rebuttal for:

Mg/Ca and $\delta^{18}O$ in multiple species of planktonic foraminifera from 15 Ma to Recent

By Boscolo-Galazzo et al.

We thank the reviewers and editor for their comments and feedback on the manuscript. We agree with almost all of the points raised and have revised the text accordingly. Below we give a point-by-point response to the reviewers comments.

Reviewer 1, Lennart de Nooijer

The statistical approaches should be explicitly described. This includes the exact way the regression analyses were done and the way the MC simulation was applied. Please provide the necessary details.

To clarify the text we have rephrased the part which describes the Monte Carlo simulations specifically (lines 259-260). We give the details of uncertainty propagation and the sources of uncertainty that are propagated in Sec. 2.4.2.

Line 254: do the 'calibration coefficients' refer to the uncertainty in the estimated constants from the regression analysis?

We now clarify this on line 258.

It would be interesting to know which of the sources of the uncertainty contribute most to the overall CI.

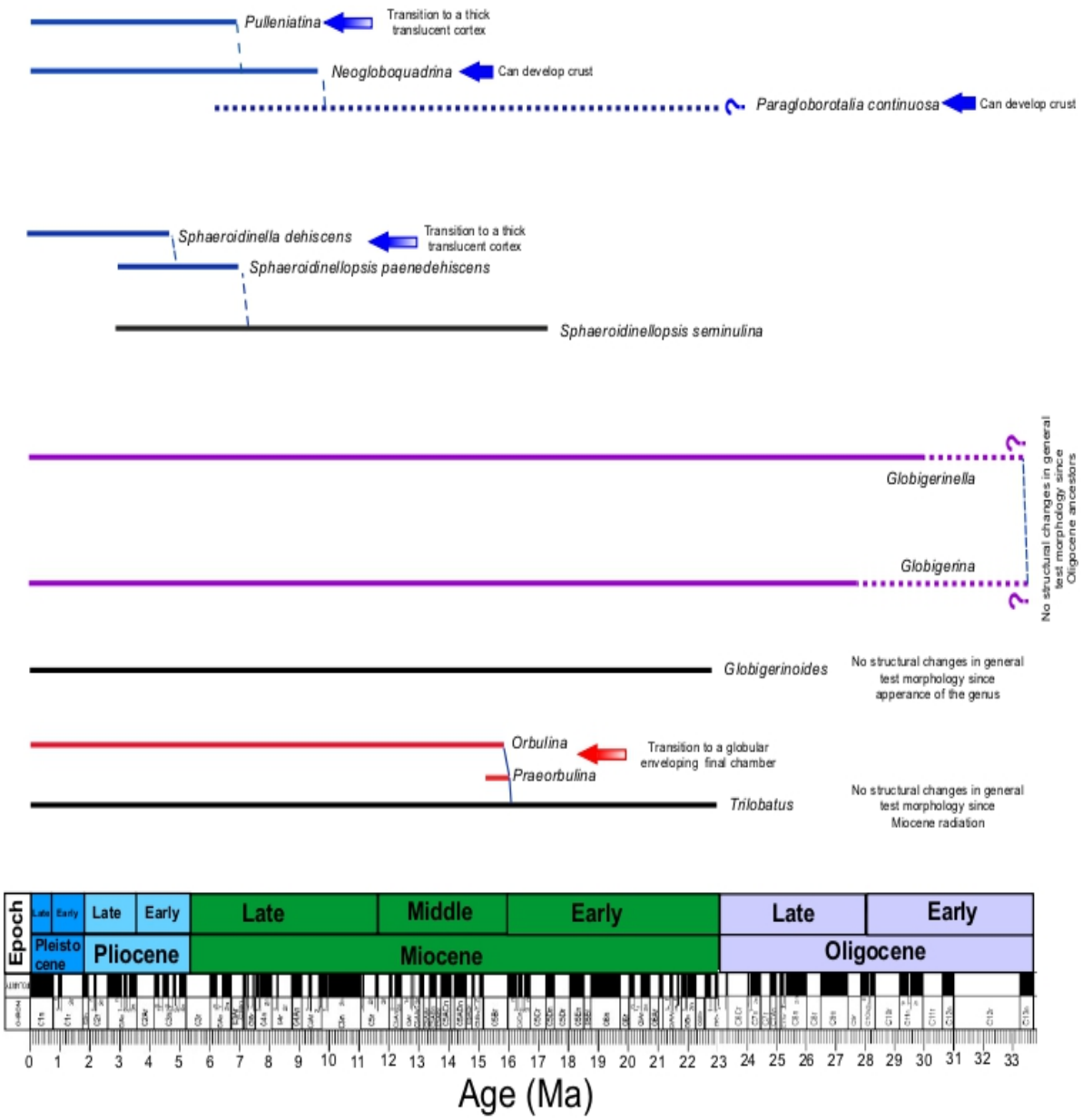
The main sources of uncertainties are derived from: the calibration coefficients, Mg/Casw, pH, and in the case of d18O, the d18Osw. In the case of Mg/Ca, the magnitude of these uncertainties is approximately split between the uncertainty on the calibration coefficients and the uncertainty on the Mg/CaSW reconstruction (~1.5°C each in the mid-Miocene). The uncertainty on salinity is negligible (<0.1°C) and included here only for the sake of completeness. This information is now given in the text on lines 307-310.

The evolutionary perspective on differences in Mg/Ca-temperature relationships is very interesting (and slightly annoying for myself: I was just working on something similar…). However, the analysis of offsets in Mg/Ca and del18O versus evolutionary kinships is not making things clearer. Figures 4 and 5 actually suggest that offsets occur here and there, with no single 'branch' containing all the offset results.

We thank the reviewer for providing the opportunity for us to clarify this point. We analysed multiple species belonging to 18 genera from 8 localities. While it may appear that offsets are here and there on a very broad scale (e.g. it is not the case that lineage-specific Mg/Ca-T relationships only appear in the non-spinose species), systematic offsets from the multispecies calibration is in many cases consistent within a lineage. Specifically, we observe that offsets tend to appear with the origination of a new lineage and to continue to the modern representative(s) of that lineage. While with the current dataset our lineage dependent offset interpretation cannot be conclusively tested, the within-lineage consistency of the offsets that we observe is what you would expect if the appearance of an offset is linked to evolution. In order to make this clearer, we have combined figure 4 and 5, as suggested by the reviewer, and added a summary figure to help visualize the consistent occurrence of the observed offsets.

It may help to show which lineages fall towards the left-upper side of the average line and which towards the right-lower side (Figure 6). This would mean that three groups can be compared. But with the current figures, it remains very difficult to see if the species with a larger offset from the average (Figure 6) cluster together. I think this needs to be shown more rigorously if the conclusion of this exercise (i.e. that evolution is somehow responsible for the variability in Mg/Ca versus del18O) can hold.

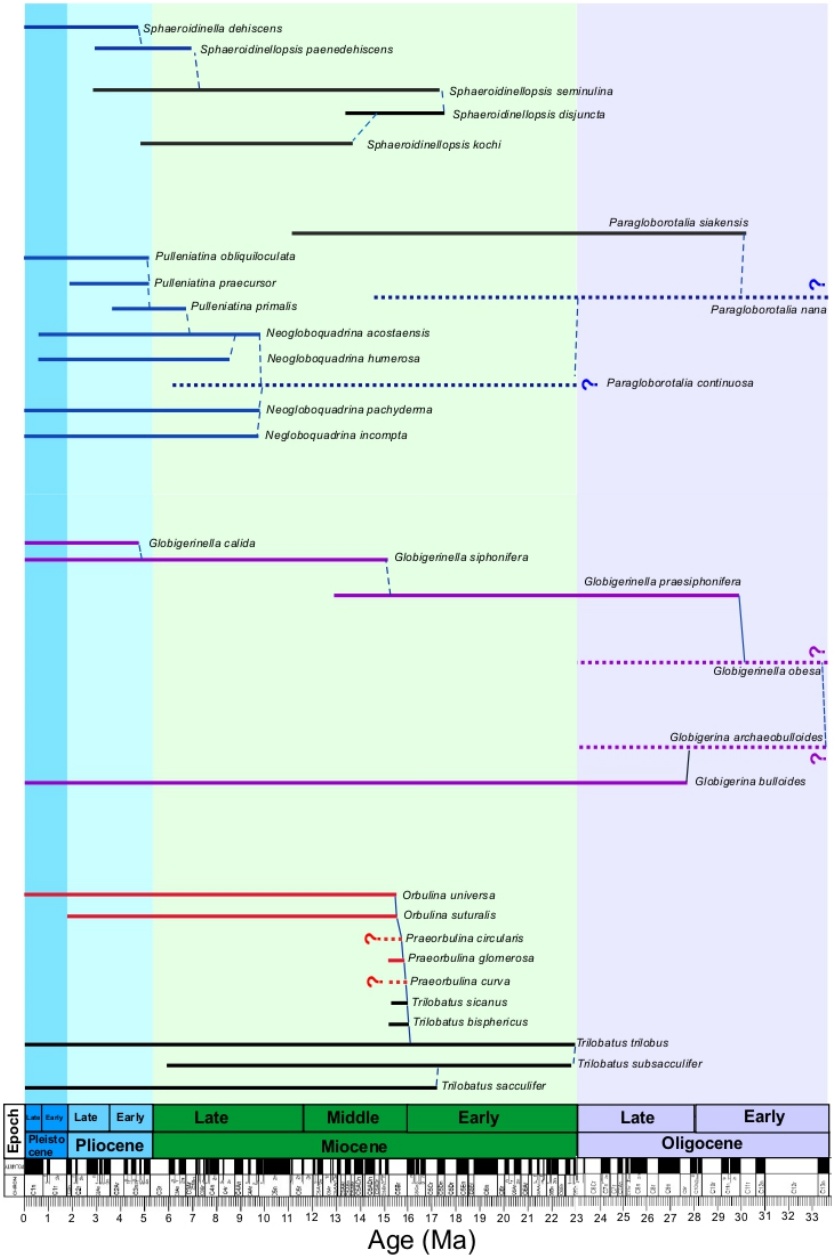
With the current dataset we cannot conclusively claim that evolution is responsible for the occurrence of offsets and we did not mean to (and in our view do not) phrase it in these terms in the manuscript. However, the consistency of the pattern suggests that this is most likely the case, pending further studies. We have now made this clearer in the text and added a summary figure to better display the offset distribution among lineages (see below).



Summary figure. Simplified phylogeny for the offset lineages discussed in the text with the occurrence of morphological changes associated with their evolution highlighted. Colored lines indicate species offset in Mg/Ca in the study dataset relative to a multispecies calibration approach, black lines indicate non-offset species. Red lines indicate offset species with higher Mg/Ca, blue lines offset species with lower Mg/Ca, and purple lines species displaying both types of offsets. Question marks indicate the lack of Mg/Ca data for a given species in the dataset presented here. Phylogeny after Aze et al. (2011) and Spezzaferri et al. (2018). The phylogenetic chart was generated using Mikrotax (Huber et al., 2016; www.mikrotax.org/pforams). The reference time scale in the figure is the Astronomical time scale of Lourens et al. (2004), until base of Chron C6Cn.2n, and Pälike et al. (2006), from top Chron C6Cn.3n until base C13n.

Also: the black color of some lines in figure 4 and 5 suggest that they are not offset from the average Mg/Ca-del18O, but with the red question mark beside them, we actually don't know, right? This gives the false impression that more species than is maybe the case, fall around the average Mg/Ca-del18O relationship. Can figures 4 and 5 not be combined? If done, it will become clear that the spinose/ non-spinose divide doesn't match the offset/ non-offset divide. This actually makes me most suspicious of the conclusion that there is a link between phylogeny and calcite chemistry (at least on this level).

We combined figures 4- and 5 and edited them following the reviewer suggestion (see below). We agree it's more complex than spinose/non-spinose as offsets appear in both groups in parallel with the evolution of new lineages. Also, the observed offsets in the different spinose and non-spinose lineages do not necessarily have the same mechanism. We suspect that large structural changes in the text morphology/wall texture (e.g. appearance of *Praeorbulina-Orbulina* or *Pulleniatina*) may be linked to the appearance of an offset through changes in e.g. biomineralization pathways. Vice versa, for *Globigerina* and *Globigerinella*, where test morphology/texture is a lot more conservative through geological time, the occurrence of the offset may be linked to changes in the ecology. We tried and showed this in the new summary figure. However, to conclusively test these interpretations we would need to track the occurrence/absence of offsets down to the Oligocene ancestors of modern/Neogene offset species.



Combined figure 4&5. Phylogenetic relationships of offset spinose and non-spinose species. Shown here are the species discussed in the text, their most closely related species and the ancestors. Colored lines indicate species offset in Mg/Ca in the study dataset relative to a multispecies calibration approach, black lines indicate non-offset species. Red lines indicate offset species with higher Mg/Ca, blue lines offset species with lower Mg/Ca, and purple lines species displaying both types of offsets. Question marks indicate the lack of Mg/Ca data for a given species in the dataset presented here. Phylogeny after Aze et al. (2011) and Spezzaferri et al. (2018). The phylogenetic chart was generated using Mikrotax (Huber et al., 2016; www.mikrotax.org/pforams). The reference time scale in the figure is the Astronomical time scale of Lourens et al. (2004), until base of Chron C6Cn.2n, and Pälike et al. (2006), from top Chron C6Cn.3n until base C13n.

Minor comments:
Line 30: remove 'of'

Ok, done.

Line 32: The authors argue in the discussion that deviations from the average Mg/Ca-del18O correlation is mainly due to inter-species variability in Mg/Ca (rather than variability in del18O). This may be the case, but interspecies differences in oxygen isotopes may still contribute to the variability shown in figures 6 and 7. Therefore, it may be more accurate here to say 'systematic offsets in the Mg/Ca-del18O relationship' or something similar.

We fully agree and have edited the text accordingly. Specifically, we have edited the sentence on lines 433-435 for clarity and now also compare the range of inter-species offsets in d18O and Mg/Ca.

Line 175: I guess you combined the 'Mg/Ca-temperature calibrations'
Edited accordingly (line 176).

Lines 205-208: it may be easier to simply state that you corrected the Mg/Cacc for past seawater Mg/Ca and pH (see 2.4.2). I guess it doesn't matter for this procedure what Mg/Ca-temperature calibration you use (since you use it twice in 'opposite' directions).

The reviewer is correct and we now clarify that this is the case (line 210).

Line 239: I don't understand this: it seems that the effect of pH, salinity and temperature were taken into account twice. In lines 218-234 it was also described that the long-term changes in salinity and pH were corrected for. Conversion to temperature was already mentioned in line 238.

We have rephrased for clarity (lines 241-242). The corrections were of course only applied once as the reviewer correctly surmises.

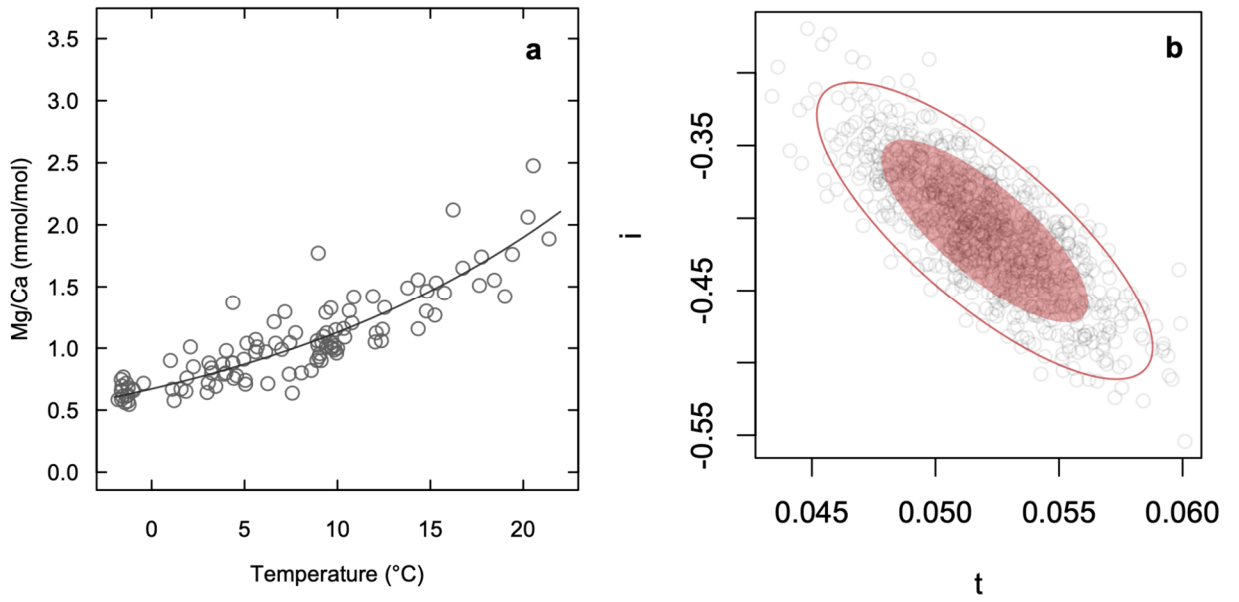
Line 240: what exponential coefficient? For the G. ruber Mg/Ca-pH sensitivity?

Here we refer to the pre-exponential coefficient of a regression in the form Mg/Ca = BeAT, which is now clarified on lines 244-245.

Line 250: what does this re-fit mean precisely? I understood that the Mg/Cacc were already corrected for salinity and pH…

MgCaRB requires a covariance matrix of parameter uncertainty in order to implement the uncertainty propagation (Gray and Evans, 2019). In order to generate the covariance matrix, we performed a bootstrap of the Tierney et al. (2019) *Neogloboquadrina pachyderma* dataset,

following the method outlined in Gray and Evans (2019), prescribing the calibration sensitivities and uncertainties from Tierney et al 2019. The calibration dataset and uncertainty covariance are shown in figure X. As a sensitivity experiment we performed the same analysis, however prescribing the 'generic' pH sensitivity of Gray and Evans (2019).



We propose to add the figure above as supplementary figure: (a) *Neogloboquadrina pachyderma* calibration dataset of Tierney et al 2019 (showing the temperature sensitivity determined in that study) and **(b)** covariance between temperature sensitivity and the 'intercept' term implemented in MgCaRB.

Figure 2: so, since the different corrections are plotted 'on top' of each other, the order in which they are stacked matters for the suggestion that only part of the data could be accounted for with these corrections. The order could also be such that the upper left part of the data cloud falls within the calibration lines. It is not important for the paper, but perhaps the authors could try to change this? The applied corrections are actually vectors pushing the original calibration in different directions…

We agree with the reviewer that this could be a more intuitive way of displaying the relative importance of the various nonthermal controls on the proxies, and we did not do so only because in several cases it is not possible - some effects cannot be described as vectors in Mg/Ca-d18O space. For example, pH impacts both Mg/Ca and d18O and has the effect of rotating the slope of the relationship between the two proxies when they are plotted against each other. That is, the slope of the line becomes shallower and the vector applicable to the bottom-left of the diagram would not be the same as that applicable to the top-right.

Figure 2: I don't understand the Mg/Casw =2.5. That is (almost) half that of today, but in the methods there lacks a justification for this. Instead, the methods describe how the reconstructed change in Mg/Ca is used to correct the Mg/Cacc. It makes me wonder in figure 2, then what this line actually signifies: the caption suggests that the change in Mg/Casw from 15 Ma till now was used to correct the Mg/Ca of the foraminifera, but this would not (necessarily) result in a constant offset (as the orange line here suggests).

The logic for choosing Mg/CaSW = ~50% present day is that this is the lowest possible value that could apply to a sample presented here, that is, it is roughly equivalent to the lower uncertainty bound on early-middle Miocene Mg/CaSW. We chose the largest possible change to make the effect easier to see and do not mean to apply that this is the value that is applicable to any specific sample - the idea of this figure in general is to visualise the sensitivity of the proxy values to the various controls. This is now clarified in the figure caption (lines 288-292).

Figure 2B: what does the 'm = -2.08' mean?
We agree that the meaning of this line was not clear, and have now clarified in the figure caption (lines 296-301).

Line 284-285: 'out of' should be 'from'.
Now corrected.

Line 311: I tried to look it up, but couldn't find how many datapoints were generated in total. The 62% sounds a higher percentage than it looks to me, but there may simply be many datapoints obscured by the density of the center of the cloud.
Yes, they all cluster together, all data are available in the supplementary table.

Line 352: should be 'addition'
Now corrected.

Line 434 and on: what do the ratios between parentheses mean?
Now clarified.

Line 442: reference to Fig 5 is incorrect here. Also, references to figure 5 in the text of the manuscript appear after references to figure 6.

We double checked this and the references to the figures appear correct to us. Figure 5 is mentioned for the first time in line 436, Figure 6 in line 439, so this also looks correct. In any case, now the figure numbering in the text has changed as a result of the inclusion of a new figure and all call outs to figures have been updated accordingly.

Figure 4: what are 'offset' spinose species? If it refers to species that have a del18O-MgCa relationship that deviates too much from the average relationship (figure 6), then what exactly makes it offset? What is the threshold value for such a distinction and what is this based on?

We define a species offset when it plots outside the combined error envelope of the two proxies, hence when the difference between the reconstructed temperature from the two proxies is >~5°C (although this is evaluated on a case-by-case basis using the specific uncertainties for a given data point). This is described in lines 266-268 in the Methods and lines 498-501 in the Discussion. We also now explain this in the figure captions of Figure 6 and 7.

Line 539: should be 'as in Figure 6'.
Now corrected.