Editor Decision: Reconsider after major revisions (08 Jan 2020) by Erin McClymont
Comments to the Author:
Constructive comments have been received by two reviewers and a short online comment. I thank the authors for responding to each of these and considering how to incorporate the suggested changes and clarifications. On the whole the reviewer concerns have been addressed, but there are several where clarification is still required and/or edits to the manuscript should be considered.

1) Reviewer 1's first comment indicated that there could be greater value in introducing more of the context and rationale for the study in the Introduction. The authors have replied that they prefer to keep a short introduction and leave the detail in the Discussion section. I agree with Reviewer 1 that some further details on the palaeo-context and application of these morphological changes would be valuable, and so ask the authors to consider increasing the detail (either in the Introduction or Discussion, as they see fit): for example, can the authors explain how these previous papers have used morphological changes for palaeoclimate reconstruction, rather than just citing the papers? Or where have debates emerged in the literature about the interpretations of morphological changes? I agree with the Reviewers that at present the manuscript needs some more depth, so that readers can see the link to the important issues in palaeoclimate interpretation. For example, in response to Review 1 comment 6 the authors note that "frequently made assumptions" in the literature are being tested (and see reply to Reviewer 2 lines 47-48), but I did not find these assumptions to be clearly explained nor cited in the manuscript. Reviewer 2 (line 41-43) also suggests citing a source to justify the assumption being tested: if this can't be done (as the author reply seems to suggest: the response does not clarify this), then I agree that a few lines of additional explanation would help the reader to understand this assumption. The authors do a good job where they address Reviewer 2 concerns about the malformations under higher CO2 (response to lines 235-242); this approach could be applied elsewhere. In summary, by addressing these concerns by the reviewers and explaining more about the literature context, I think a stronger foundation for the study will in turn lead to increased clarity about how these results can be applied.

We thank the Editor for her recommendations that allow us to reconsider both the introduction and discussion chapters: we confess that we discuss a lot about all issues that the Editor point out and especially about point 1 and 3. Finally, we decided to add some lines in the introduction (L 37-44) to clarify the application of morphological changes in the geological record making a few examples. We also added a sentence in the discussion part (L230-232) to explain the process of associating morphological studies with paleo-environmental conditions.
However, we still think that a short introduction and discussion that focused on the main question of our work could make this article more interesting, readable and appealing. The primary aim of our work was to understand if physiological experiments with contemporary species are a valid tool to reconstruct responses of ancient coccolithophores to environmental change in the geological record and this is also the reason for which we didn't dig into species-specific morphological variations.

To our knowledge, this work for the first time evidences the weakness of associating nannofossil morphologies and environmental parameters supporting this approach with observations and data from living coccolithophore species. This practice assumes that coccolithophore algae preserve the same mechanism through time towards one (or more) environmental stress/parameter. As far as we know, with this work, for the first time, someone looked deeply into coccolith morphologies under environmental parameters that were stated to induce coccolith variations in the fossil record. Since now, the idea was that both living and fossil species react and reacted in the same way to extreme oceanic conditions given the fact that they all belong to the same group (Haptophyte).
We would like to emphasize that one of the authors of this manuscript, applied the same approach in two papers that are cited in the text (Faucher et al, 2017a; Faucher et al., 2017b) and therefore, since the beginning of these experiments, we firstly questioned ourselves the validity of the approach that we also previously used.

Having said that, we tried to give more information on the paleo-context we took inspiration from to perform our experiments, in both the introduction and discussion. We wrote a further sub-chapter that was intended to be placed before the (already written) discussion chapter, where we talked about the possible influence of the five tested parameters for coccolith production in the past ocean. However, we still think that this detracts from the main message of our manuscript and we finally decided to delete it. We think that, eventually, this part could be included in the Supplementary file as a "State of the art" paragraph to show the background of our work and underline that we consider most of the literature on the topic of morphometry and environmental condition in the past ocean. We ask the Editor her opinion for this decision hoping that she agrees on our point of view. We leave the written paragraph down here and we also added it in the supplementary file (S1).

Geological background - state of the art
Several authors tried to link nannofossil morphological variations with detected environmental conditions: the process assumes to analyze nannofossil species through a sedimentary succession and to evaluate the presence of shape or size anomalies in the considered interval of time. The eventual detected morphological variations are then linked to

independent paleo-environmental data (e.g. sea surface temperature reconstruction (SST), CO_2) to find the environmental driver for the identified morphological variation

Indeed, the past oceans were characterized by episodes of anomalous or extreme sea-water conditions that could have possibly influenced the phytoplanktonic communities. A good example is the oscillations between "calcite seas" and "aragonite seas" (Sandberg, 1983) that possibly influenced the productivity of calcareous nannoplankton in different times (Erba, 2006). The amount of massive amount of chalk deposited during the Late Cretaceous is a good illustration of a high productivity time for calcareous nannoplankton probably permitted by a shift in seawater chemistry towards a very high level of Ca. In parallel, rising Mg/Ca ratio during the Cenozoic and up to present days is correlated to a reduction in coccolithophore diversity and coccolith thickness (Bown et al., 2004). Also, locally, light could have played a major role for coccolithophore calcification in the past ocean: it was documented that during episodes of intensified continental weathering, more clastic particles were transported into the sea and in the coastal area might have diminished the depth of the photic zone (Lechler et al., 2015). The reduction in light availability was associated with habitat changes of the photoautotrophic primary producers that produced smaller coccoliths to be able to dwell in shallower depth and compensate for the reduction in sunlight (Lübke and Mutterlose, 2016). The calcareous phytoplanktonic communities in past oceans were also disturbed by intervals with excess CO_2 concentrations related to intense volcanic activity. Modifications in size and morphology of calcareous nannofossil during times were CO_2 reached up to 1000-2000 ppm, were interpreted as a transient response to survive progressively increasing surface-water acidification (Erba et al., 2010, Lübke et al., 2015; Faucher et al., 2017). Besides, environmental constraints for calcareous nannoplankton growth, involve the ocean trophic level: in the fossil record, some authors linked the decrease in nutrient availability, with reductions of abundances and sizes of some calcareous nannofossil species (Linnert and Mutterlose, 2012). Other authors, on the contrary, detected similar size reductions in several ocean areas characterized by both oligotrophic and mesotrophic seawater conditions (Bornemann et al., 2006; Faucher et al., 2017). Finally, the ocean in its history was subjected to variation in temperature: a strong hydrothermal activity on one hand and an intense continental weathering, on the other hand, were the main triggers of respectively CO_2 released and CO_2 sequestration, that in turns, often produced a concomitant increase or decrease in SST. Episodes characterized by relatively low SST were sometimes related to small coccolith sizes (Bornemann and Mutterlose, 2006). However, opposite behaviors were also observed (size decreases under extremely warm conditions, Erba et al., 2010; Lübke et al., 2015) in the same species in different geological intervals.

References:

Bornemann, A. and Mutterlose J.: Size analyses of the coccolith species Biscutum constans and Watznaueria barnesiae from the Late Albian "Niveau Breistroffer"(SE France): taxonomic and palaeoecological implications, Geobios, 39(5), 599-615, https://doi.org/10.1016/j.geobios.2005.05.005, 2006.

Bown, P. R., Lees, J. A., Young, J. R.: Calcareous nannoplankton evolution and diversity through time: In Coccolithophores, pp. 481-508. Springer Berlin Heidelberg, 2004.

Erba, E.: The first 150 million years history of calcareous nannoplankton: biosphere–geosphere interactions, Palaeogeography, Paleoclimatology, Paleoecology, 232, 2, 237-250, https://doi.org/10.1016/j.palaeo.2005.09.013, 2006.

Erba, E., Bottini, C., Weissert, H. J., Keller, C. E.: Calcareous nannoplankton response to surface-water acidification around Oceanic Anoxic Event 1a, Science, 329(5990), 428-432, https://doi.org/10.1126/science.1188886, 2010.

Faucher, G., Erba, E., Bottini, C., Gambacorta, G.: Calcareous nannoplankton response to the latest Cenomanian Oceanic Anoxic Event 2 perturbation, Rivista Italiana di Paleontologia e Stratigrafia (Research In Paleontology and Stratigraphy), 123(1), 2017a.

Lechler, M., von Strandmann, P. A. P., Jenkyns, H. C., Prosser, G., & Parente, M.: Lithium-isotope evidence for enhanced silicate weathering during OAE 1a (Early Aptian Selli event). Earth and Planetary Science Letters, 432, 210-222. 2015 https://doi.org/10.1016/j.epsl.2015.09.052.

Linnert, C., Mutterlose, J.: Biometry of Cenomanian–Turonian placoliths: a proxy for changes of fertility and surface-water temperature? Lethaia, 46(1), 82-97, 2012 https://doi.org/10.1111/j.1502-3931.2012.00323.x.

Lübke, N. and Mutterlose, J.: The impact of OAE 1a on marine biota deciphered by size variations of coccoliths, Cretaceous Research, 61, 169-179, https://doi.org/10.1016/j.cretres.2016.01.006, 2016.

Lübke, N., Mutterlose, J., Bottini, C.: Size variations of coccoliths in Cretaceous oceans, a result of preservation, genetics and ecology? Marine Micropaleontology, 117, 25-39, https://doi.org/10.1016/j.marmicro.2015.03.002, 2015.

Sandberg, P. A.: An oscillating trend in Phanerozoic non-skeletal carbonate mineralogy. Nature, 305(5929), 19-22, 1983.

2) Reviewer 1 queried the different control conditions exhibited in Figure 2. It would be useful for the authors to include in their main text or figure caption, the information about the different timing of the experiments to account for the differences. I'm not clear from the reply whether the same control sample was used for all experiments, but stored in between, or if different control samples were generated each time an experiment was started. Could the authors please clarify?

Different control samples were generated every time for every experiment from the stock culture. The controls have always the same medium and were treated in the same ways in every experiment. We added in the text the timing of the experiments. See the caption in Fig. 2.

3) Reviewer 1 queried why the wider range of morphological measurement parameters was not outlined in the manuscript. Although the authors indicate that in some way this data might detract from their focus of comparing species (and the main assumption they outline for testing), could this information not be valuable for considering the within-species responses, and providing some interesting discussion on what kinds of morphological changes can be observed? I ask the authors to consider whether this information could be included, even though the measurements are not uniform across all species. It could make the manuscript a more comprehensive discussion of morphological change, and add to the finding that different species have different responses.

We thank the Editor for her suggestion that allows us to review all the data collected and re-checked all collected morphological parameters.
We modified some sentences of the results underlining morphometric information that was already shown in Tab.1.

We had a long and deep discussion among us, about the possibility to dig into the species-specific responses for every experiment but at last, we decided to leave the discussion as it is without adding any further paragraph. In the process of writing this manuscript, in the draft version before the submission process, we organized the discussion with six sub-chapter, one for every experiment and a final part where we summarize the observed responses. However, at the end and after spending time and energy in writing those parts, we realized that this structure, the analysis of the available literature for both living and fossil organisms for every experiment, transformed our manuscript into a review paper that wasn't our purpose. We took the drastic decision to delete that part because we thought at that time and we still think that all these paragraphs would have diverged from the key message of our work and would have made little of the answer to our questions.
To be more explicit, we entirely agree with the Editor that there are some species-specific changes to the tested parameters and that our dataset has lots of potential in digging more on the response of single coccolithophore species towards environmental stress and we, therefore, understand the idea of the Editor to extend the discussion into within-species responses. However, we believe that analyzing the responses of every species towards every environmental parameter and understand the reason why coccolith react in a certain way towards an environmental condition, is not the goal of this paper.
We hope to convince the Editor that a discussion that focuses on the major question we asked ourselves ("Can morphological features of coccolithophores serve as a reliable proxy to reconstruct environmental conditions of the past?) is more desirable.

Minor concerns:
Reviewer 2 noted confusion about the timing of the divergences. I ask the authors to separately state the timing of the divergence of E.huxleyi and G.oceanica compared to the much older divergence of G.oceanica with the other two species. The timing of E.huxleyi / G.oceanica divergence is never stated, but should be somewhere, especially as it is not visible on Figure 1. Mixing the 'hundred thousand years' with millions of years in the suggested edit doesn't provide clarity.

This information was added to the text (line 77).

Reviewer 2 also indicated that on Fig 1 the numbers on the nodes were not clearly explained. The authors have added "Node, representing following divergence, used for calibrating...". I'm not still not clear what this numbers mean: is some text missing here?

The caption was changed. Fig. 1 is modified from Liu et al., (2010). The nodes represent divergences of species. The numbers represent some of the progressive nodes that Liu et al., (2010) wanted to highlight. For every node the time of divergence was calculated. For example, node 57 corresponds to the divergence between Coccolithus pelagicus and Helicosphaera carteri that was calculated as 220 Ma. For the divergence time of all nodes, see Liu et al., 2010.

Reviewer 2 queried (for line 138) why non-linear regression was used. The author reply does not provide clarification: is it because no linear relationship was determined? or because the structure of the changing light intensity experiments does not lend itself to linear regression? or something else?

First of all, for the light experiment we preferred to design an experiment with many treatments in the expense of the number of replicates following Cottingham *et al., (2005). Therefore, a high number of treatment levels was set up with no replication. We used a non-linear regression because we observed non-linearities in the response variable,*

The change to Figure 3 is valuable for showing the detailed structure of the data more clearly. Comments have raised concern about how easy the text on the figures is to read. For Figure 3 I would recommend considering plotting this graphic as 2 x 2 panels rather than 1 x 4, which would allow larger versions of each panel to be displayed. The authors should also consider increasing the font size wherever they are able, across Figures 3 and 4.

Figure 3 was changed following the Editor advice. The font sizes were increased when possible in Fig. 3 and 4.

In a revised version of the manuscript the authors should check that the number of coccoliths observed is clearly stated. This was raised several times and although the authors indicate that "this information was highlighted in the text and in the caption" (reply to reviewer 1) I did not find it in the original manuscript either.

We checked the text. It is now written in: line 142; in the captions in Fig. 3 (lines 481-484), in Fig. 4 (line 505), Tab.1 (line 517), Tab.2 (line 530), Tab. 3 (lines 550), Tab. 4 (line 559) and Tab. 5 (line 589).

All the variation in the text are in:
Red: referees requests.
Underline in grey: Editor's requests.

With kind regards
On behalf of all co-authors
Giulia Faucher