The work by Frieling and colleagues is strong framework and a much-needed study that will open a new opportunity for applications of organic microfossil 13C analysis. Like single species foram analyses (the benchmark for modern carbon and oxygen isotope studies) single or several organic microfossil 13C limits the breadth of sources to sedimentary organic matter and limits the degrees of freedom in a highly advantageous way. This study is the gateway to the deeper geologic record that will allow broad application of the dinocyst proxy to ancient carbon cycle studies. The questions below are meant to enhance the discussion, but the work, as it is, stands on its own as it is presented.

Author response:

We thank the reviewer for the positive and constructive review of our work. The review highlights a number of aspects of the methodology which often do not appear in the published literature but in this case will be helpful to provide a baseline for further work.

From a methodological perspective I appreciate the details provided here. Controlling for size and process length is great approach but do you see relationships between $d^{13}C_{cyst}$ and cyst size?

Author response:

The reason we aim to exclude size-dependent $\delta^{13}C$ differences is that for e.g. foraminifera, coccoliths and living dinoflagellates a size-dependent ^{13}C fractionation has been observed previously (Burkhardt et al., 1999; Hoins et al., 2015). Unfortunately, a size-dependent $\delta^{13}C$ relation in modern ocean dinocysts is beyond what we can reasonably test with our method – because of the analytical uncertainty when measuring such small (30-40 µm diameter) individuals the number of required repeat measurements would become unpractically large. However, we fully agree this is a logical next step, once the methodology is sufficiently developed to achieve the precision needed to distinguish between individual cysts' $\delta^{13}C$ signature of modern species.

To what degree do you feel that the time averaging affected your data? Do you have access to any 14C dates of the surface sediments? From here you could potentially model the expected range of 13C values of DIC accounting for Suess Effect. More details in the manuscript on your rough correction would be helpful.

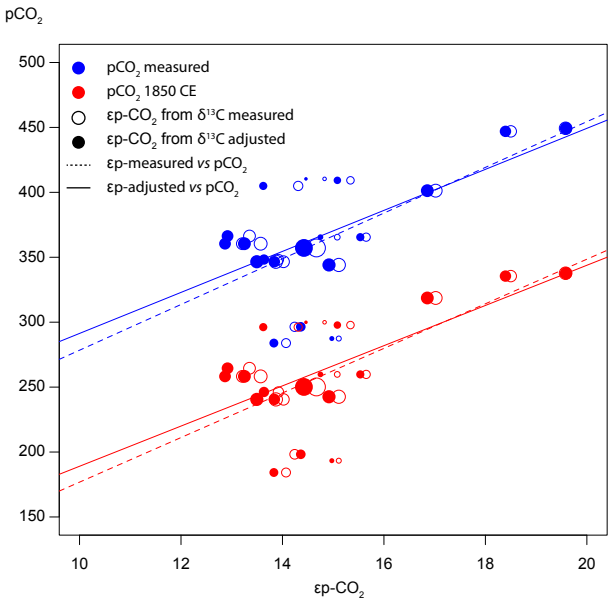
Author response:

This point is in line with one of the points raised by the other reviewer. In addition to the explanation below, we refer to the other reponse for further details on our considerations. We will elaborate on this in the revised manuscript.

We fully agree that better age-control on sediments, but especially of the individual cysts would be helpful and also modelling $\delta^{13}C_{DIC}$ (subtracting the Suess Effect) as the reviewer points out is a step we would like to take if feasible. However, the use of ^{14}C dates (often carbonate, otherwise bulk or macro-scale organic matter) is complicated as they cannot be measured on the cysts themselves – not even when concentrating large amounts of cysts. Therefore, these analyses cannot represent the true dinocyst age, and certainly not an age distribution as the individual cysts in our data represent. We feel that incorporating any sample 'age' correction for now would mainly result in erroneous corrections and hence we

prefer to use a correction as a sensitivity test (what could be the maximum error), which we will elaborate on in the (revised) manuscript.

Specifically, in our revised manuscript we will illustrate the maximum error that may result from the Suess-effect by constructing calibrations with uncorrected data for each of the parameters (measured $\delta^{13}C$ and CO_2) and compare these to the original calibration (see Figure below). However, we stress the difference between these calibrations is very much a worst-case scenario and unlikely an accurate reflection of true uncertainty caused by the Suess-effect.



The figure above will be added to our supplementary information. The figure shows the difference between measured and adjusted pCO_2 and $\delta^{13}C$ ($\varepsilon p\text{-}CO_2$). Open symbols indicate measured $\delta^{13}C$, closed symbols represent data after eliminating small signals (<0.2 Vs) and outliers. Blue dots represent measured CO_2 values and red dots indicate the CO_2 around 1850 CE. For each dataset a simple linear regression, weighted to the number of measurements, is given. Dashed lines utilize measured $\delta^{13}C$, solid lines are from final $\delta^{13}C$ data. The red solid line is used in Figure 5A.

What do you think is the background blank source? Is it from atmospheric aerosols that adhere to all surfaces regardless of precautions or is it from within the nickel plate? (Does the nickel plate show scoring from the laser?). Regardless, the approach to signal size to noise, considerations of the blank and other corrections seem reasonable. These consideration are important not only for your study and approach but for the future potential of this kind of analysis for sample return from Mars and elsewhere.

Author response:

We will expand our methods section with a few statements covering these aspects:

The origin of the blank source is currently unknown and proved difficult to constrain. When setting up the system, we used a liquid N_2 cooled trap to pre-concentrate CO_2, before releasing it to the IRMS. With that system inherently the 'blank' $\delta^{13}C$ was also much larger, as it also concentrates the blanc signal, which is why we returned to the current true continuous flow system. However, even with that concentrated blank it remained difficult to constrain it's C-isotopic value and therefore also the opportunity to confidently identify the source (see also e.g., Van Roij et al., 2015).

Still, we considered potential sources to be (1) part atmospheric CO_2 (air or particles that come into the system when opening the sample cell) and (2) part residual material from earlier measurements (wall sorption) and (3) possibly a minor amount of additional C from the water the cysts are isolated from. We assume that the first source are the atmospheric particles the reviewer hints at. The nickel plate is not scored or etched during ablation and testing with clean plates (before any sample is added) shows the nickel plate itself does not release C. Similarly, repeated test with water droplets being added showed this a minor source for a blank only (< 10% of total blank). Until better constraints become available, we prefer to refrain from further speculation on the potential blank source, other than stated above.

Regardless of the source, the combined contribution of these factors proved to be small and stable, so that we were able to correct for it (as can be seen in Fig. 2).

Line 280: From this discussion I think I favor your argument that intercyst variability reflects individual differences. One can envision that individual cells or cysts have significantly different 13C values owing to the randomness of cellular growth, changes in microenvironments of growth that also affect DIC and CO2 13C. Add in the time averaging from core top sample collection it is not a surprise that you see large variance. In fact, I would be worried if you did not. Your suggestion of controlling for size, as much as one is able, is a good idea.

Author response:

We thank the reviewer for their view – indeed this is also our preferred scenario for explaining the intra-sample variability. We will include the reviewer's points on potential for ^{13}C-impact of cell-microenvironments and growth-induced randomness to the $\delta^{13}C$ of the cyst in our revised manuscript.

Line 280: For standards have you considered dissolving a standard material like caffeine in water and allowing it to dry onto a surface and analyzing that (you could spray it or something). At the very least here you could assume that the starting composition is isotopically uniform. I supposed 13C differences could arise from the drying process, but it may be better than PEF.

Author response:

We have been in search for a sufficiently homogenous standard with similar ablation and material characteristics as the polyethylene plastic used here for some time. The reasons for searching for similar material are mostly operational. For example if the entire sample plate is covered in standard material (as indeed could be done with spray or by submerging), it becomes impossible to insert the sample without contaminating the system, or ablate sample

unknowns within the same run (thus cannot be used for corrections). It seems likely such a technique would also increase the blank. The main issue with solid standard material (e.g. a film or foil) is homogeneity on the scale required (~80 µm spot size). During cooling or drying (crystallisation) of materials (including e.g. glycerine, various corn-based products, monosodium glutamate) we observed that the resulting solid did not remain homogeneous. Work on a new standard continues and, once successful, will be implemented in our methods and report on it in future publications.

Line 300: Have you investigated the compositional differences between cyst and motile cells? I am familiar with the references you report on this issue but what specifically are the differences? What proportion of the carbon from the cell transferred to the cyst? Is this known?

Author response:

These are important outstanding questions and subject of currently running as well as planned work regarding cell compartment derivation of cyst molecules using LC-IRMS, cyst production – excystment experiments to assess cell to cyst fractionation. In short, we here need to compare the core-top cysts to cultured motiles as that is the only currently available data.