Answers to os-2019-113_RC1

Changes in the composition of marine and sea-ice diatoms derived from sedimentary ancient DNA of the eastern Fram Strait over the past 30,000 years

By Heike H. Zimmermann, Kathleen R. Stoof-Leichsenring, Stefan Kruse, Juliane Müller, Rüdiger Stein, Ralf Tiedemann, and Ullrike Herzschuh

Dear Jessica Ray,

thank you very much for your comments and questions which helped to improve the quality of this manuscript. Please, find below your comments and questions in bold letters, while our answers are placed below and changes that will be made in the text are underlined. New references that will be included are placed at the end of this document. Line numbers are referring to changes made in the revised manuscript.

l.127 - Less than 50% of 3 replicates?

We agree that this is confusing and changed the sentence from:

"[…](3) were present at least 3 times among the different replicates, (4) showed taxonomic resolution below the phylum level "Bacillariophyta" and (5) were tagged as "internal" by obiclean in less than 50 % of the different replicates to reduce PCR and sequencing artefacts."

L131-133:

"[…](3) were present at least 3 times among <u>all sequenced PCR products</u>, (4) showed taxonomic resolution below the phylum level "Bacillariophyta" and (5) were tagged as "internal" by obiclean in less than 50 % of <u>all sequenced PCR products</u> to reduce PCR and sequencing artefacts."

l.147 - What was the constraining factor used for CONISS analysis? Please include.

The constraining factor was depth. With regard to your concerns below we decided to omit the CONISS analysis.

l.201 - What are the rare ASVs? A quick perusal of the Table S1 indicates many zeros. Would it be possible or useful to include a supplementary figure showing rank-abundance curves, either for the entire dataset or for the individual samples?

The majority of ASVs are rare (less than 1% per sample). We will provide an additional supplemental file including the rank abundance curves for each depth (see preliminary Fig. 1 below). We decided to also include several stratigraphic diagrams containing the ASVs without grouping.

L209-2011: The replicates of each sample show some variations (Fig. 3) in the presence and abundance of ASVs, especially for ASVs amounting to less than 1 % per sample (Suppl. Table 2, Suppl. File 1).

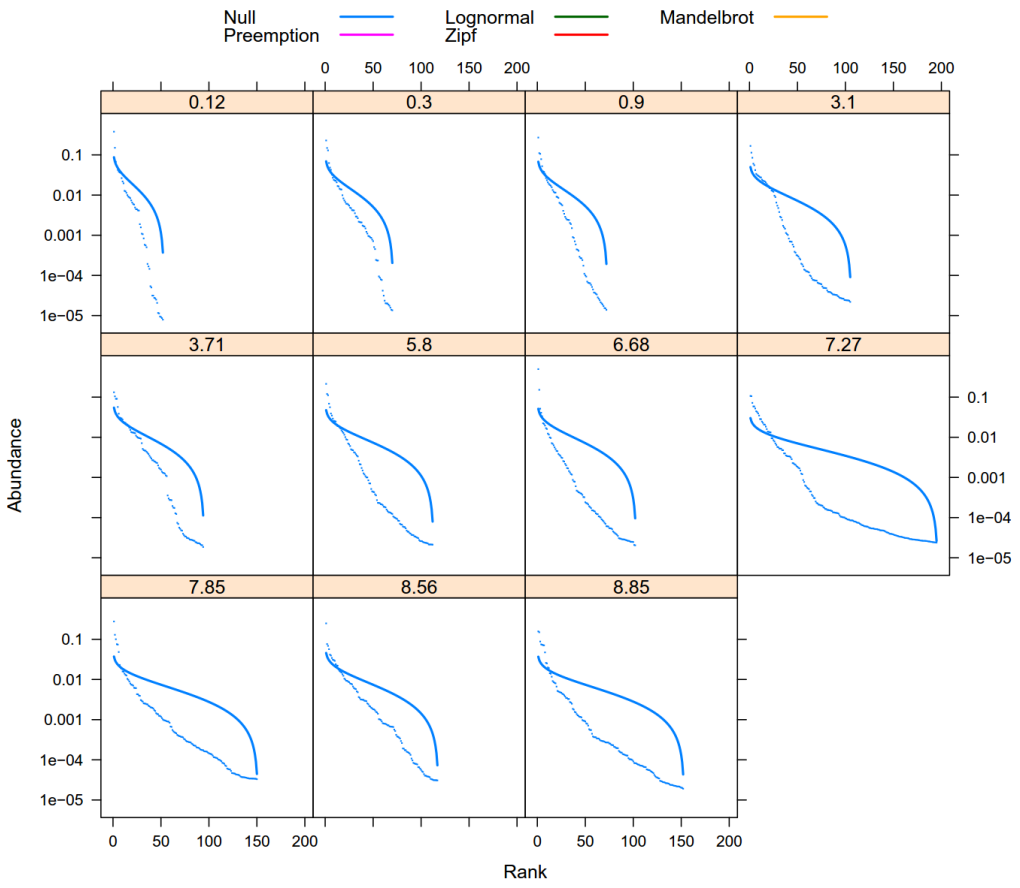


Figure 1: Rank-abundance plot for each sample (depth in m) based on amplicon sequence variants (ASVs).

l.205 - Interesting theory, but why would inhibition only appear in some samples? I think this might be easy to test by a dilution PCR of a few representative samples. Also, if there are differences in the relative amount of diatom DNA in each sedaDNA sample, would it be possible to do a quick qPCR check of rbcL target abundance using the same primer set?

Unfortunately, we could not go back to the lab as work there was reduced to a minimum the past months. We discussed this and decided that it is most likely due to different amounts of diatom template molecules and deleted the part referring to inhibition.

L228-229: It is possible that higher dissimilarities between some replicates are the result of the low amount of template DNA ~~or due to PCR inhibition by compounds that were not sufficiently removed during DNA extraction~~.

l.209 (Section 5) - My first impression from reading the results is that the authors have some difficulty in interpreting diatom community assemblage differences according to the results of the CONISS analysis, i.e. division into five aggregate "zones". In particular because trends in the relative abundance of specific sympagic diatom taxa between different CONISS "zones" are mentioned but not statistically tested. This leads me to question whether CONISS analysis is useful given the present dataset. Were any other analyses attempted to identify discriminant ASVs/taxa? And how did the authors conclude that there are two diatom assemblage reorganizations (l.340) when the CONISS analysis identifies four?

We agree, that this method might not be ideal and removed the CONISS analysis. We have deleted all references to this method and thus slightly adjusted the text.

l.223 - Could you please elaborate on what is meant by "richness of taxonomic names". According to l.184 different ASVs can be assigned to the same name, so how might this affect the apparent richness? And why use "taxonomic names" for richness calculations when taxonomic rank assignment is not uniform across all ASVs?

With "richness of taxonomic names" we mean grouped ASVs assigned to the same taxonomic name. We compared rarefied taxonomic richness based on (1) ASVs and (2) also for grouped ASVs assigned to the same taxonomic name. While changes of ASV-derived richness vary stronger between the samples in comparison to the changes of ASVs grouped by their taxonomic name, the trends are similar: Richness is highest in the last glacial samples, lower in the deglacial samples and lowest in the Holocene samples.

L244-245: We changed the sentence to:

“Generally, the richness of both ASVs and unique taxonomic names (ASVs grouped based on identically assigned taxonomic names) is higher in samples dated to the last glacial in comparison to those dated to the Holocene (Fig 4).”

Furthermore. we changed the captions of Fig. 4 and 5 to:

“Figure 4: Barplots showing the rarefied (a) number of amplicon sequence variants (ASVs) per sample and (b) ~~number of unique taxonomic names~~ grouped ASVs assigned to the same taxonomic name for each sample with depth (m) of the sediment core MSM05/5-712-2.”

“Figure 2: Taxonomic composition with relative proportions (%) of the 360 detected sequence variants grouped into 75 unique taxonomic names based on identically assigned taxonomic names of sediment core MSM05/5-712-2. Taxonomic names are sorted according to the weighted average with depth.”

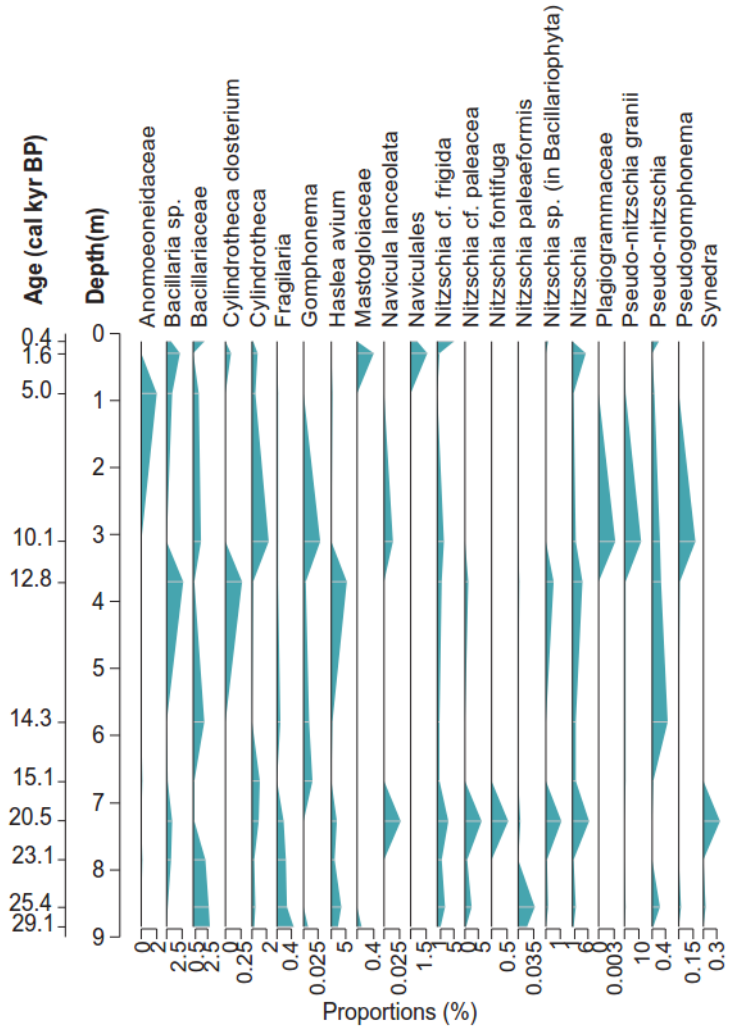
l.224 - I am not quite comfortable with the use of "turnover" in the context it is used, since the samples are discrete and therefore a discontinuous representation of time. In my opinion, "turnover" suggests a biological/ecological linkage from one sample to the next, while in this study the samples compared are isolated snapshots in time. Can the authors comment on the choice to use this term?

We agree, that the term is misleading. We chose the term as we see a distinct change in the taxonomic composition, e.g. after the LGM. But since we did not calculate beta-diversity (exactly because we only have a few samples) we changed “turnover” to “shift” in all occurrences.

Fig.5 - I wonder if the reader might find this figure somewhat difficult to interpret given that relative taxon abundance at multiple taxonomic ranks are presented for each sample. Could the authors please the reasoning for presenting the data in this way?

Stratigraphic diagrams are a common representation in paleoecology, for example for pollen or microfossil analysis. As taxonomic resolution differs in different families or genera (for example due

to lack of morphological differences or here due to sequence similarity of closely related species in the databse), showing the detected organisms on the lowermost taxonomic level as possible was our goal. Out intentions was to used wa.order=”topleft” option in strat.plot, which sorts the taxa according to the weighted average with depth to better visualize the change over time. However, we see your point and will omit the weighted average and make the plot more informative, e.g. by separating centrics from pennates. Again, in the manuscript we want to show the grouped version to make the plot as complete as possible. However, we made a new supplementary file with several figures showing strat.plots of all 360 ASVs.



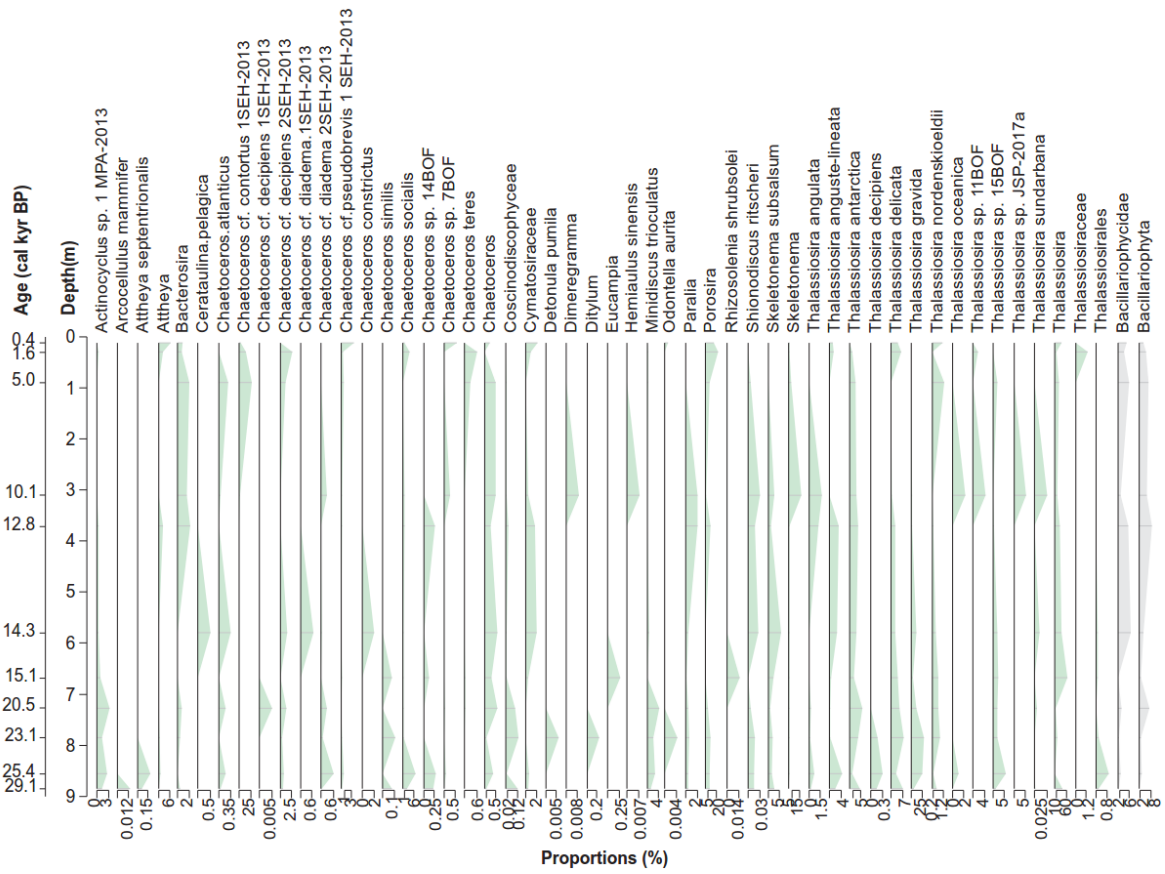


Table 1 - The Paleoenvironmental conditions descriptions seem somewhat arbitrary. For example, how is "sea-ice retreat" (3.1 m depth) different from "Reduced sea-ice cover allowing spring sea-ice algal and summer phytoplankton productivity" (7.85 m depth).

We agree. We cited the descriptions given in the original research. We decided to omit the table.

l.239 – foraminifer

Changed. Thank you.

Fig.6 - Family-level taxonomy is presented but l.170-174 states that this collation may mask functional differences.

We provided stratigraphic diagrams for all 360 ASVs in the supplement. This is figure was supposed to give an alternative representation of the data on family level in context with IP_{25}. We also sorted the taxa according to centrics and pennates, so that the figure will be easier to understand.

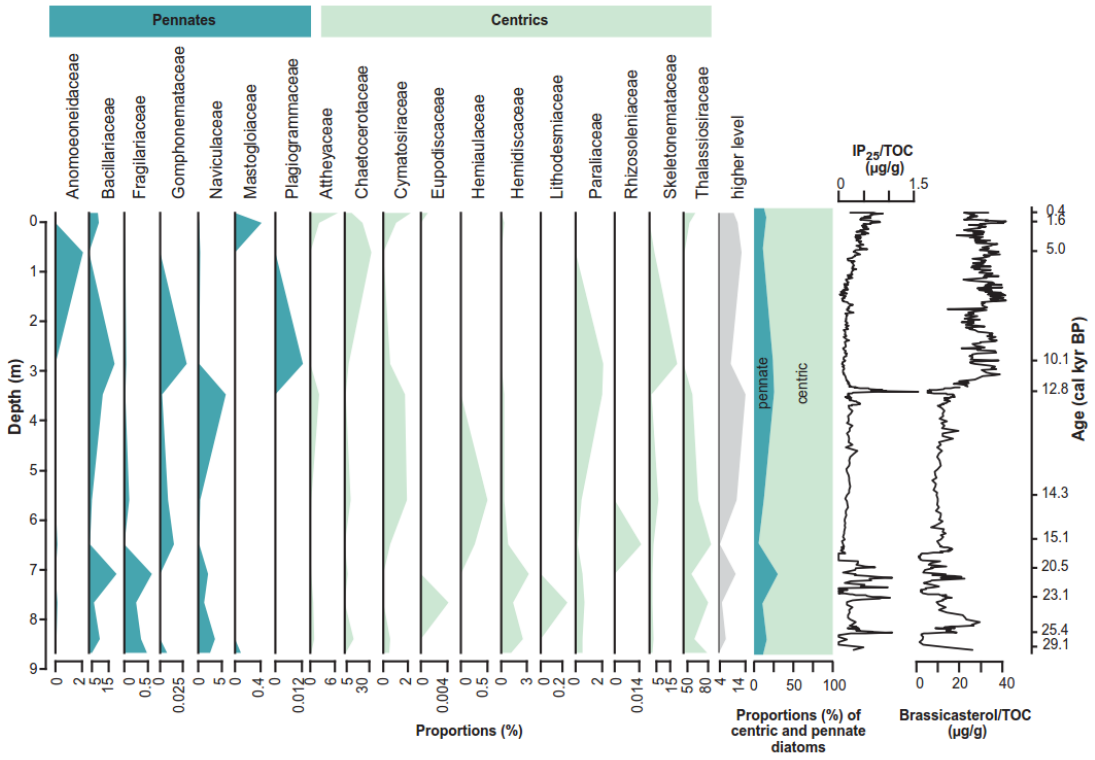


Fig.6 - The double top axes (depth and age) are very helpful, but identify very clear differences in sedimentation rates in the downcore. For example, zone II has higher temporal resolution than zone I. Might not differences in sedimentation rates also affect sedaDNA signal?

Yes, a higher temporal resolution indeed could affect the sedaDNA signal. This could be an additional explanation for the lower richness of zone II in comparison to zone I. However, the distinct shift we see in taxonomic composition is rather an environmental signal. We added the following sentence.

L281-283: “Higher sedimentation rates and thus higher temporal resolution during the Bølling/Allerød phase could have affected the *sed*aDNA signal, yet the distinct shift in taxonomic composition suggests, that this is rather an effect of the changing environmental conditions during this phase.”

l.247 - Again, ambiguous results from CONISS analysis?

Changed

l.336 - "highly detailed taxonomic resolution" depends on what is meant by highly detailed, and what fraction of the data is being referred to. I suggest moderating this statement.

L.336-337: “The rbcL_76 marker is highly diatom specific and provides highly detailed taxonomic resolution, mostly at genus and species level.”

We deleted the word “highly”. We mention in this sentence that we refer to the fraction of data resolved on genus and species level. In total 64.4% of our 360 ASVs of the filtered dataset are resolved on genus or species level (see line 163), and we believe this is quite detailed. But

l.320 - Very interesting that N. cf. frigida sticks out as a possible new sea-ice proxy. However, according to Fig.6, this taxon as highest relative abundance in the most recent sample (0.4 kya BP) when sea-ice cover is low. What about Cylindrotheca closterium? Or Haslea avium? Again, I think it would be very helpful if taxon relative abundances were statistically tested in order to identify ASVs/taxa that contribute significantly the observed diatom diversity in different samples.

Around 0.4 kyr BP, which falls in the temporal phase of the Little Ice Age, IP25 values are increasing. So there is sea ice influence. Results of our statistical analysis can be found in the comments of Reviewer 2. We did not find a linear relationship between any of the ASV in our record and IP_{25}. However, we have re-written the section and included the new results.