Answers to os-2019-113_RC2

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 John Kirkpatrick,

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.

Methods: Section 2.1: The coring equipment and handling should be described. Was this a piston core? What diameter? How was it handled? Were the cores halved onboard? Did the authors use an archive half or a working half that had previously been used for other sampling activities? Most importantly, how was the core stored – at what temperature? If at -80 C, this should be clearly stated. If not at -80 C, it should be acknowledged that DNA preservation ex situ may be imperfect.

The core was taken in 2007 with a Kastenlot corer (gravity) that has a diameter of 30 x 30 cm. As this is sufficient material, about 1m long subcores are taken and stored at 4°C. The core has to be opened and subsampled on board. However, the subcores had also been used for other sampling activities. We will refer to the supplement of Gersonde et al. (2012) in the manuscript where the procedure is explained in detail.

We will change the sentence to:

L77-80: "The kastenlot core MSM05/5-712-2 (N 78.915662, E 6.767167, water depth 1487 m) was collected from the western continental slope of Svalbard during the cruise of Maria S. Merian (Budéus, 2007) in the eastern Fram Strait in 2007 (Fig. 1). On board, subsections of 1 m length are taken in square plastic boxes as explained in the supplement of Gersonde et al. (2012) and stored at 4°C. Therefore, DNA preservation might be imperfect."

Do the authors have any tracers or controls for post-coring seawater influence on the sediment core? I am willing to believe that the authors, working in a dedicated lab, were meticulous about their sampling and extraction. Did they use tracer DNA as described by Epp et al.?

At the time of coring (2007), it was not planned to analyze ancient DNA. Hence, no tracers were added. We understand such concerns, but the chance of post-coring contamination deep inside a Kastenlot core seems quite low. A supporting argument is that if we had such contamination, we would not see distinct changes among our samples that are consistent with past climatic changes.

Section 2.2: This is the most significant concern I have about this dataset. The primer set used was designed by Dulias et al. for freshwater lakes. Its specificity for diatoms appears to be very high, as the authors highlight, which is good to address. However, it contains multiple mismatches on both the forward and reverse primers for the taxa noted to contribute to the IP25 proxy (Haslea,

Pleurosigma; please see attached "supplement"). (I did not look at other marine taxa, but that should be investigated.)

Please also note the supplement to this comment:

https://www.ocean-sci-discuss.net/os-2019-113/os-2019-113-RC2-supplement.pdf

The primers were designed to function as a general diatom-specific marker by Stoof-Leichsenring et al. (2012), but since then they were mostly applied on lacustrine sediments. We added this information in the introduction:

L67-69: "We used sedaDNA metabarcoding by applying the diatom-specific *rbcL_76* marker (Stoof-Leichsenring et al., 2012) which has already proved successful in low-productivity lakes of northern Siberia (Dulias et al., 2017; Stoof-Leichsenring et al., 2014, 2015), but so far has not been tested on marine sediments."

However, these primers were not tagged for parallel high-throughput sequencing. Tagging and adjustment of PCR conditions was performed by Dulias et al. (2017). We therefore added this information in the methods section:

L106-107: "The PCR reaction mixes and conditions were prepared following the adjusted protocol for tagged *Diat_rbcL_705F* and *Diat_rbcL_808R* primers as described in Dulias et al. (2017) with the exception that 3 µl DNA (DNA concentration 3 ng/µl) was used as a template."

We agree with your concern. Mismatches in primer binding sites are not ideal, but using DNA metabarcoding always comes with a trade-off. The primer binding sites were chosen to maximize (1) amplification success by having the amplicon short, (2) retrieving as many diatom (pennates and centrics) species as possible with the best resolution possible and (3) it should co-amplify as few non-diatom groups as possible.

Only 3 species are known to produce IP_{25}: *Pleurosigma stuxbergii* var. *rhomboides*, *Haslea spicula* and *Haslea kjellmanii* (Limoges et al., 2018). Of those, only *Pleurosigma stuxbergii* has a reference sequence containing our marker and we detected only 1 mismatch in the forward and one mismatch in the reverse primer-binding site. *Haslea nusanatra* and *Haslea howeana* which you provided in your supplement have a tropical distribution. On the other hand, *Haslea avium*, which was also detected in our study, shows only 1 mismatch in the forward and one mismatch in the reverse primer-binding site. We checked the number of mismatches in the forward and the reverse primers with the 15 *Haslea* spp. reference sequences in our database and found between 1-5 mismatched in forward and 0-2 mismatches in reverse primers. The number of mismatches *Haslea spicula* and *Haslea kjellmanii* have is certainly in that range. However, most sequences that we retrieve show mismatches: e.g. *Thalassiosira* species mostly have 2-3 mismatches with the forward and 0-2 with the reverse primers. And up to 5 mismatches with the forward and 1 mismatch with the reverse primer can be found for *Chaetoceros socialis* (assigned to GenBank Accession FJ002154). Hence, detection might also be

As a result this dataset should not be considered to be a well-rounded assessment of diatoms in this setting. Any of the increases / decreases in richness, or the absence of certain taxa, or overall assessment of diatom diversity are not valid as we are effectively blinded to many of the taxa known to be present in this location (and presumably throughout the sediment record).

This concern is based on the assumption that the primers were specifically designed for freshwater lakes, which they were not. We agree that absences in our data does not necessarily mean absence in the past.

We include the following sentence:

L217: "The recovery of taxa by *sed*aDNA metabarcoding is prone to false presences or absences. As our study lacks a morphological diatom record, false absences cannot be assessed and true presences cannot be confirmed, which means that absence in our record not necessarily translates to a true physical absence in the past."

This could perhaps be partially addressed by compiling an alignment of diatom chloroplast rbcL sequences documented to be or have been present at this location and noting how many taxa you would expect to find do or don't have mismatches with the primers. Maybe the authors are aware of this (?). At the end of the day, the detection of the sequences found is real data, and valuable. However, a broad assessment of shifts in diversity and community responses to climate is very problematic.

We agree with your concern. Unfortunately we only know a fraction of the species that occur(red) in the area. The region is overall understudied by molecular surveys and dissolution leads to incomplete microfossil records. Hence, we feel that we cannot provide such an alignment, as it would be incomplete as well. About the mismatches, please have a look 2 questions above. This is a known issue and wobble bases are included in the primers for mitigation of this bias. If it is desired, we can prepare a figure to visualize the primer mismatches.

On a different note, the cycle numbers and annealing temperatures, in my opinion, should be included even if they are included in the given citation. 50 cycles is a very large number. Did the authors ever visualize bands in their extraction or PCR negative controls?

We agree, that 50 cycles are a very large number, but large cycle numbers are not uncommon in ancient DNA studies (up to 65 cycles: (Willerslev et al., 2014) or 45 cycles: (Voldstad et al., 2020)). We point these issues out in lines 190-194. For future studies, reducing cycle numbers to 45 is planned, but we cannot change this for the current project. However, reducing cycle numbers also comes with a trade-off, as this has shown to reduce replicability (Krehenwinkel et al., 2017; Nichols et al., 2018).

L201: We included the reference Nichols et al., 2018.

Yes, after each PCR, we first evaluated bands on an agarose gel. We will include the following sentences:

L108-110: "PCRs were performed with the following settings: 5 minutes at 94°C (initial denaturation), then 50 cycles at 94°C (denaturation), 49°C (annealing) and 68°C (elongation) and a final elongation step at 72°C for 5 minutes. Subsequently the PCR success was checked with gel-electrophoresis."

Line 110 – what version (chemistry)?

We change the sentence to:

L114: "The sequencing library was prepared with the Mid Output kit v 2 according to the Fasteris Metafast protocol for low complexity […]"

Section 2.3: I'm OK with the use of ASVs and I appreciate the authors taking time to explain why they didn't use OTUs. Regarding their reference database, how many references did this method produce? Did these references include diatoms documented in other studies from the waters

around Svalbard? The reference database is a frequent scapegoat in the discussion; release 138 is a couple years ago. The taxonomy seems pretty robust to me but it could be updated if need be.

This method produced 2,320 references, and yes, it contained many taxa which are documented in waters around Svalbard, such as *Nitzschia frigida, Chaetoceros socialis, Thalassiosira antarctica* or *Thalassiosira nordenskioeldii* (to name a few that were also detected in our record).

The ENA/EMBL release 138 was released in November 2018, and thus up-to-date when we started the analyses. The reviewer probably confused this with the GenBank release 138 from 2003.

We changed the sentence to:

L124-126: To generate the reference database for the taxonomic assignment of the sequences we downloaded the EMBL release 138 (released November 2018) and used *ecopcr* (Ficetola et al., 2010) according to the descriptions of Dulias et al. (2017) and Stoof-Leichsenring et al. (2012) containing 2320 reference sequences.

Going back to the negative controls, on line 122 how many exactly is "the majority" that were singletons?

We provide an additional Supplementary table which contain the negative controls from extractions and PCRs. Furthermore, we will change the sentence to:

L127-129: "Of the 204 different sequence variants detected in extraction and PCR negative controls 83% of their occurrences were singletons and […]"

On line 125, why 10 read counts? How many sequences, and how many ASVs, were removed out of the total using these criteria? As the supplemental table only has the kept sequences, one can't tell.

This is only a first threshold we use for denoising so we do not lose too much information in the beginning. We included the following information in the text:

L134-135: "Filtering with R reduced the number of read counts from 7,536,449 to 6,199,984."

Section 2.4 (and throughout): Please make sure to always use the term "PCR replicate" or "PCR replication", as these are not sample replicates. PCR replication is useful but the terminology should be clear as later on it may get confusing for readers who aren't methods geeks and think that these are sample or extraction replicates.

Thank you for pointing this out. We changed this accordingly throughout the manuscript and figure captions.

Section 2.5: Resampling 100 times might be overkill but I guess it can't hurt. The new minimum number of sequence counts doesn't quite make sense to me through – can you clarify "according to descriptions in the preceding paragraph"? (It's not quite 12 *25,601.)

We change the following sentence "For taxonomic composition and richness calculations, we first summed up the sequence counts for each ASV of the replicates belonging to the same sample…"

To:

L152-153: “For taxonomic composition and richness calculations, we combined the PCR replicates of the corresponding sample. This resulted in a new minimum number of sequence counts (300,415 counts) that was used for resampling the dataset 100 times according to descriptions in the preceding paragraph.”

Section 3: Line 157 – What do you mean by "outperformed"? Are you referring to the specificity for diatoms? Please clarify.

Thank you for pointing that out. We changed the sentence to:

L174-175: “Possibly owing to a much shorter marker size, the *rbcL_76* marker also surpassed the usually amplified 18S rDNA markers with regard to specificity for diatoms (Coolen et al., 2007; De Schepper et al., 2019; Kirkpatrick et al., 2016)”

Line 164 – "striking proportion" – I don’t think it’s that striking, in fact, 4.8% seems pretty low!

We deleted the word striking.

Line 177 – can you give a numeric range for the copy number variation in published work?

Unfortunately we cannot give a numeric range, as most papers we found mention this as vaguely as we did. However, we included another reference, where the authors showed that rbcL copy number was significantly different in the 8 tested benthic freshwater diatoms and that this variation is positively correlated with cell biovolume (Vasselon et al., 2018).

We changed the sentence to:

L195: “Furthermore, the enrichment of centric diatoms in the *sed*aDNA record could be the result of copy number variation of the *rbcL* gene between different species and cell biovolume (Bedoshvili et al., 2009; Round et al., 1990; Vasselon et al., 2018).“

Section 4: Here I think again it is important to be clear that the replicates being discussed are PCR replicates of the same template DNA. Also, what happened to the two sample replicates from the deepest sample? How similar or dissimilar were they and does the sample variation exceed the variation found in PCR replication? While two duplicates aren’t statistically as useful as a triplicate, they need to be discussed in the context of reproducibility. Doing so would, I think, strengthen the authors’ arguments. This should be addressed in revision (unless I missed it somewhere?).

We already handled PCR replicates, as mentioned above.

In the NMDS plot (see below the improved version of the figure), you can see, that the 3 PCR replicates of the two samples at 8.85 m depth cluster so narrowly, that it is really hard to identify the labels at all. Hence, they were very similar. However, as this is an ordination biplot, we cannot move around the labels. Therefore, we included the colours of the polygons in the caption and hope that this makes it clearer.

The degree of dissimilarity is visualized in the NMDS plot. We agree that more replicates always help to decrease the level of uncertainty and include this statement in the text:

L225-227: “For the oldest sample at 8.85 m depth we additionally processed a sample replicate. The PCR replicates of both sediment samples at 8.85 m depth were highly similar and clustered tightly

together in the NMDS plot (Fig. 3). Although a higher number of replicates would improve the robustness of our analysis […]"

We changed the following sentence to make it clear, that the PCR replicates are based on the same template:

L 219-220: "The PCR replicates (different PCR products of the same DNA extract) of each sample show some variations (Fig. 3) […]"

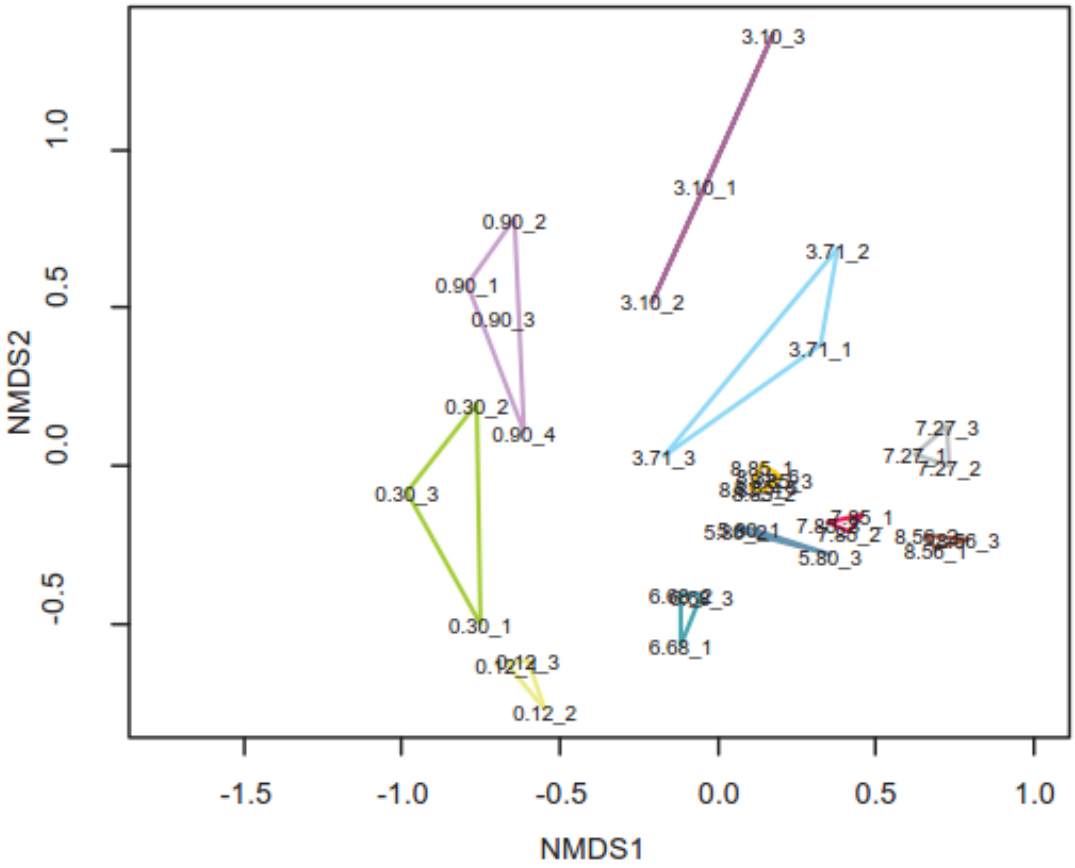


Figure 3: Non-metric multidimensional scaling plot based on the filtered and resampled diatom dataset with the PCR replicates (indicated by an underscore with a number) of a sample (depth in m) linked in a polygon of sample-specific colour: light green = 0.12m, green = 0.3m, light violet = 0.9 m, violet = 3.1 m, light blue = 3.71 m, dark blue = 5.8 m, turquoise = 6.68 m, grey =7.27 m, red = 7. 85 m, brown = 8.56 m, orange = 8.85 m (2 sample replicates with 3 PCR-replicates each).

Section 5: Lines 228-299: This inverse relationship, if significant, is intriguing! Please test statistically for significance.

The inverse relationship between *Chaetoceros* and *Thalassiosira* was tested for significance on family level. The Pearson's correlation efficient is -0.61126971, with a p-value of 0.045701848. Hence, there is a moderate significant negative correlation. We included the information in brackets into the sentence:

L251: "A general trend that can also be observed at the family level is an inverse relationship (r=-0.61, p= 0.046) of the dominant families Thalassiosiraceae and Chaetoceraceae (Fig 6)"

Furthermore, we will include a sentence into the methods section (see answer after the comment regarding section 6).

Lines 242-244... etc: This sort of speculation about how ice conditions may have affected diatom community diversity, etc., is undermined by the primer issue.

Is this concern based on the assumption that the primers were specifically designed for freshwater lakes (which they were not)? Based on the sedaDNA composition (both sympagic and pelagic diatoms were detected) and by taking into consideration the results of other proxies during this time, it is conceivable that conditions were heterogeneous throughout the seasons. We changed the sentence slightly:

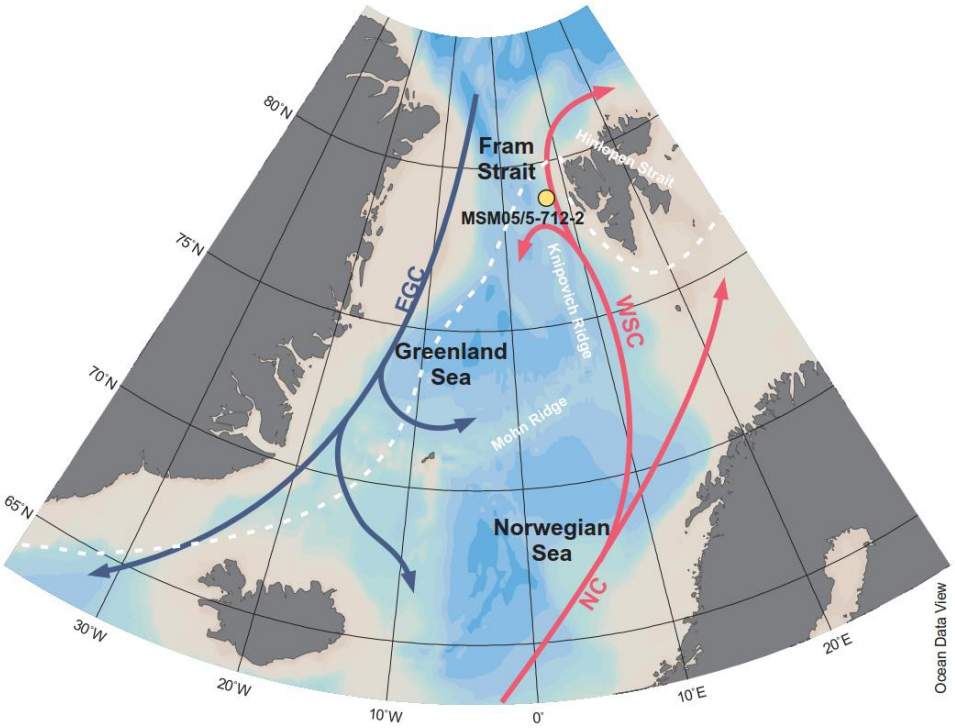
L264: "It is conceivable that a heterogeneous and dynamic environment produced by winter sea-ice cover with ice-free conditions during summer probably allowed for diverse diatom communities to develop in the different habitats and over the seasons, which is reflected in the highest numbers overall of ASVs in samples dated to this time span."

Lines 253 and elsewhere: "Low proportions of Nitzchia cf. frigida..." and similar statements should be reworded, as we do not know the proportion of diatoms. We know the proportions of the sequences found in the dataset after extensive amplification (50 cycles!). Please be careful to keep this distinction clear.

We included the proportions in brackets after each statement.

Line 265: How far away is that (Hinlopen Strait)? In this section in general, there are numerous and interesting comparisons to published records from related locations. It would be appreciated if the proximity of these (in km) were clear for the less-familiar reader. (E.g. also line 276, 278, 291, 304, 323)

We added an annotation of the places mentioned in our text into the map.



Line 274: "peak proportions", "Lower proportions" – can you please quantify. (Also e.g. line 301.)

Yes, we included the proportions in brackets after each statement.

Line 289: Richness and diversity are not the same; please be specific.

We will consistently refer to richness instead of diversity as recommended.

Section 6: Beyond the methods concerns, there are still some interesting points here. These could be strengthened by a statistical analysis testing the correlation between the new data to the published IP25 proxy (lines 319 - 323).

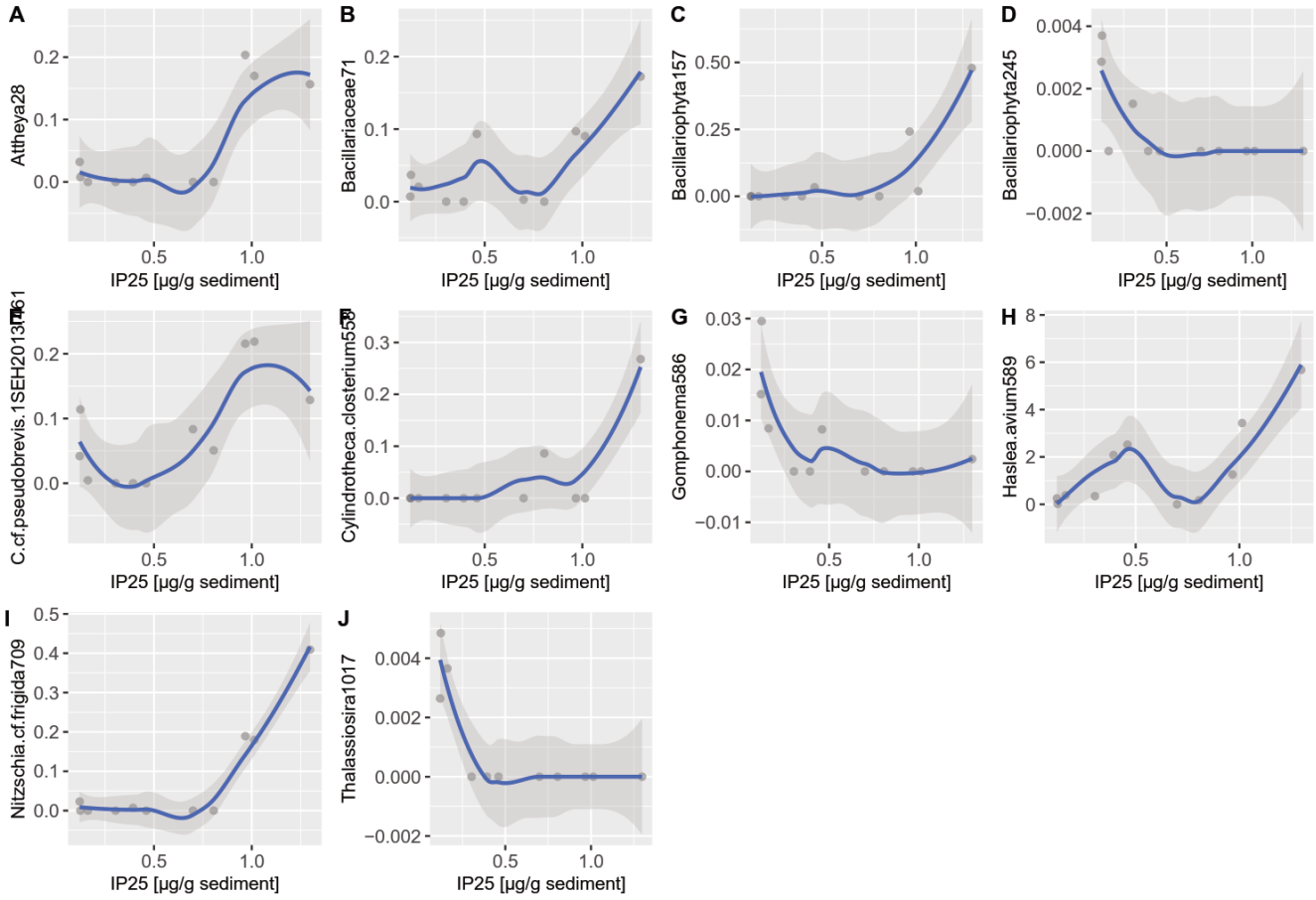
We have tested for significance, first on family level and second on ASV level and will add this information in the text, the plots showing significantly correlated families or ASVs (see Figure below) will be added in the Supplement. As we had to interpolate IP_{25} values for the depths that were analysed in this study, we include a paragraph describing the process in the methods section. We plotted the significantly correlated ASVs (see table below and figure below) and found that they do not show a linear relationship with IP_{25}, which is why we will not include the table in the manuscript. However, some of the patterns are quite interesting (see figures below the table) and we have re-structured section 6 and included new results and parts of the discussion about primer mismatches in Haslea and Pleurosigma.

We included:

L161-166: "For correlation analysis we interpolated IP_{25} values using the methods described in Reschke et al. (Reschke et al., 2019). Therefore, the IP_{25} data were transformed using function *zoo* from the "zoo" package (Zeileis and Grothendieck, 2005) and used in the function *CorIrregTimser* using the package "corit" (https://github.com/EarthSystemDiagnostics/corit). The correlation between Chaetocerotaceae and Thalassiosiraceae as well as between IP_{25} and all ASVs was tested for significance using R function *rcorr* (method Pearson) from the package "Hmisc" (Hollander and Wolfe, 1975; Press et al., 1988)(Quelle)."

Table 1: Amplicon sequence variants (ASVs) with significant (p-value) positive correlation.

ASV	Correlation coefficient	p-value
Attheya28	0.7542327	0.00732029
Bacillariaceae71	0.6970434	0.01713238
Bacillariophyta157	0.7159116	0.01322211
Bacillariophyta245	-0.6217257	0.0411348
Chaetoceros.cf.pseudobrevis.1.SEH.2013461	0.6619875	0.02649079
Cylindrotheca.closterium558	0.6387489	0.03439247
Gomphonema586	-0.6077406	0.04731908
Haslea.avium589	0.6899115	0.01880586
Nitzschia.cf.frigida709	0.8093591	0.0025435
Thalassiosira1017	-0.673798	0.02301182



Technical corrections: Figure 3: Labels are not readable. Figure 5: This appears upside-down. More importantly, the labels and proportion bars are very, very small; but much of the graph is empty whitespace. Please try making the bars wider. Also, consider organizing the taxa by e.g. rank order (based on proportion) and breaking into 2 or 3 separate figures. Figure 56: Please define "higher level" in the caption.

Fig. 3: As we don't want to manipulate the biplot by spreading apart the labels, we attributed distinct colours and will refer to PCR replicates in the caption. We placed the figure to another answer further above (Figure 3).

Fig. 5: Yes, we split the graph into 2 graphs and included further recommendations from Jessica Ray (Reviewer1). Originally we used a rank order (based on proportion): wa.order="topleft" option in strat.plot, which sorts the taxa according to the weighted average with depth.

We defined higher level in the caption as: "Higher level contains sequences assigned between family and phylum level."

References:

Coolen, M. J. L., Volkman, J. K., Abbas, B., Muyzer, G., Schouten, S. and Damsté, J. S. S.: Identification of organic matter sources in sulfidic late Holocene Antarctic fjord sediments from fossil rDNA sequence analysis, Paleoceanography, 22(2), doi:10.1029/2006PA001309, 2007.

De Schepper, S., Ray, J. L., Skaar, K. S., Sadatzki, H., Ijaz, U. Z., Stein, R. and Larsen, A.: The potential of sedimentary ancient DNA for reconstructing past sea ice evolution, ISME J, doi:10.1038/s41396-019-0457-1, 2019.

Dulias, K., Stoof-Leichsenring, K. R., Pestryakova, L. A. and Herzschuh, U.: Sedimentary DNA versus morphology in the analysis of diatom-environment relationships, J Paleolimnol, 57(1), 51–66, doi:10.1007/s10933-016-9926-y, 2017.

Ficetola, G. F., Coissac, E., Zundel, S., Riaz, T., Shehzad, W., Bessière, J., Taberlet, P. and Pompanon, F.: An In silico approach for the evaluation of DNA barcodes, BMC Genomics, 11(1), 434, doi:10.1186/1471-2164-11-434, 2010.

Gersonde, R.: The expedition of the research vessel "Sonne" to the subpolar North Pacific and the Bering Sea in 2009 (SO202-INOPEX), Berichte zur Polar- und Meeresforschung = Reports on polar and marine research, 643, doi:https://doi.org/10.2312, 2012.

Hollander, M. and Wolfe, D. A.: Nonparametric statistical methods., Biometrische Zeitschrift, 17(8), 526–526, doi:10.1002/bimj.19750170808, 1975.

Kirkpatrick, J. B., Walsh, E. A. and D'Hondt, S.: Fossil DNA persistence and decay in marine sediment over hundred-thousand-year to million-year time scales, Geology, 44(8), 615–618, doi:10.1130/G37933.1, 2016.

Krehenwinkel, H., Wolf, M., Lim, J. Y., Rominger, A. J., Simison, W. B. and Gillespie, R. G.: Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding, Scientific Reports, 7(1), 17668, doi:10.1038/s41598-017-17333-x, 2017.

Limoges, A., Massé, G., Weckström, K., Poulin, M., Ellegaard, M., Heikkilä, M., Geilfus, N.-X., Sejr, M. K., Rysgaard, S. and Ribeiro, S.: Spring succession and vertical export of diatoms and IP_{25} in a seasonally ice-covered high Arctic fjord, Front. Earth Sci., 6, doi:10.3389/feart.2018.00226, 2018.

Nichols, R. V., Vollmers, C., Newsom, L. A., Wang, Y., Heintzman, P. D., Leighton, M., Green, R. E. and Shapiro, B.: Minimizing polymerase biases in metabarcoding, Molecular Ecology Resources, 18(5), 927–939, doi:10.1111/1755-0998.12895, 2018.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P., Eds.: Numerical recipes in C: the art of scientific computing, Cambridge University Press, Cambridge [Cambridgeshire] ; New York., 1988.

Reschke, M., Kunz, T. and Laepple, T.: Comparing methods for analysing time scale dependent correlations in irregularly sampled time series data, Computers & Geosciences, 123, 65–72, doi:10.1016/j.cageo.2018.11.009, 2019.

Stoof-Leichsenring, K. R., Epp, L. S., Trauth, M. H. and Tiedemann, R.: Hidden diversity in diatoms of Kenyan Lake Naivasha: a genetic approach detects temporal variation, Molecular Ecology, 21(8), 1918–1930, doi:10.1111/j.1365-294X.2011.05412.x, 2012.

Vasselon, V., Bouchez, A., Rimet, F., Jacquet, S., Trobajo, R., Corniquel, M., Tapolczai, K. and Domaizon, I.: Avoiding quantification bias in metabarcoding: Application of a cell biovolume correction factor in diatom molecular biomonitoring, Methods in Ecology and Evolution, 9(4), 1060–1069, doi:10.1111/2041-210X.12960, 2018.

Voldstad, L. H., Alsos, I. G., Farnsworth, W. R., Heintzman, P. D., Håkansson, L., Kjellman, S. E., Rouillard, A., Schomacker, A. and Eidesen, P. B.: A complete Holocene lake sediment ancient DNA record reveals long-standing high Arctic plant diversity hotspot in northern Svalbard, Quaternary Science Reviews, 234, 106207, doi:10.1016/j.quascirev.2020.106207, 2020.

Willerslev, E., Davison, J., Moora, M., Zobel, M., Coissac, E., Edwards, M. E., Lorenzen, E. D., Vestergard, M., Gussarova, G., Haile, J., Craine, J., Gielly, L., Boessenkool, S., Epp, L. S., Pearman, P. B., Cheddadi, R., Murray, D., Brathen, K. A., Yoccoz, N., Binney, H., Cruaud, C., Wincker, P., Goslar, T., Alsos, I. G., Bellemain, E., Brysting, A. K., Elven, R., Sonstebo, J. H., Murton, J., Sher, A.,

Rasmussen, M., Ronn, R., Mourier, T., Cooper, A., Austin, J., Moller, P., Froese, D., Zazula, G., Pompanon, F., Rioux, D., Niderkorn, V., Tikhonov, A., Savvinov, G., Roberts, R. G., MacPhee, R. D. E., Gilbert, M. T. P., Kjaer, K. H., Orlando, L., Brochmann, C. and Taberlet, P.: Fifty thousand years of Arctic vegetation and megafaunal diet, Nature, 506(7486), 47–51, doi:10.1038/nature12921, 2014.

Zeileis, A. and Grothendieck, G.: zoo: S3 Infrastructure for Regular and Irregular Time Series, Journal of Statistical Software, 14(1), 1–27, doi:10.18637/jss.v014.i06, 2005.