Ocean Sci. Discuss.,
https://doi.org/10.5194/os-2019-113-RC2, 2020

Ocean Science Discussions

Interactive comment on "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 et al.

John B. Kirkpatrick (Referee)

kirkpatj@evergreen.edu

Received and published: 15 April 2020

General comments: Zimmerman et al. present a novel dataset that relates previously published information on climate records from the waters near Svalbard to newly obtained diatom chloroplast DNA sequences found in the sediment dating up to 30,000 years ago. The attempt not just to detect and sequence old DNA but relate it to known proxies and paleoenvironmental conditions is an exciting step toward realizing the promise these techniques have been hoped to provide. That said, I have some questions about the samples and methods, and the choice of primer sequences limits

Printer-friendly version

Discussion paper

the utility of data to make inferences about the paleo community structure and connect to the IP25 proxy.

Specific comments:

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.

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.?

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.) 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 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.

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?

Line 110 – what version (chemistry)?

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.

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

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.

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.)

Section 3: Line 157 – What do you mean by "outperformed"? Are you referring to the specificity for diatoms? Please clarify. Line 164 – "striking proportion" – I don't think it's that striking, in fact, 4.8% seems pretty low! Line 177 – can you give a numeric range for the copy number variation in published work?

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?).

Section 5: Lines 228-299: This inverse relationship, if significant, is intriguing! Please test statistically for significance. 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. 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. 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) Line 274: "peak proportions", "Lower proportions" – can you please quantify. (Also e.g. line 301.) Line 289: Richness and diversity are not the same; please be specific.

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).

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.

Please also note the supplement to this comment:
https://www.ocean-sci-discuss.net/os-2019-113/os-2019-113-RC2-supplement.pdf

---

Interactive comment on Ocean Sci. Discuss., https://doi.org/10.5194/os-2019-113, 2019.

Printer-friendly version

Discussion paper