

Dear Jack Middleburg,

Please find attached our revised submission to BG. This included the mark-up draft of the paper along with the final copy for our review. We have endeavoured to follow the advice of our reviewers. Many of the comments were highly insightful and we appreciate the feedback. There were several remarks however, that we did not understand and of these we did not try to infer what the reviewer meant and left the text in its original form. The location of these edits is indicated in the specific itemized responses provided below.

Additionally, we also took the liberty to add data to the tables that we thought were an oversight on our part during the last submission. These data include the summed totals of IPL and core lipids of GDGTs and those that were used to tally the TEX86 ratios (see mark-up of table 1). We further added a table to the Supplemental that includes the brGDGT data. Although this lipid class does not feature in our study, we do state why we do not consider the proxies based on these lipids. Lastly, we found a computation error in our model. Equation 6 (line 576) uses the standard deviation of lipid degradation rates between GDGTs in its correction of thermogenic and heterotrophic loss impacts on TEX86. The S^2 value had incorrectly been set to 0.11 (the measure of mean variance). It has since been changed to 0.20 (see Table S4). This error, coupled with the recalculation of the model to include 2G-GDGTs (as requested by reviewer 2) has resulted in Figure 6 having slightly different r^2 values and a shift in the grouped reconstructed TEX86 SSTs. No other errors in the computations were identified. Reviewer 2 highlighted a disagreement for a time step delay in our overprint calculation. We maintain that this theoretically and computationally is more realistic (please see below for further discussion).

A final point, reviewer 2 requested a detailed discussion of archaeal community structures in the vent. This, however, was already done for Bentley et al. 2022. The overlap would not be beneficial to this paper. We have instead mentioned that further discussion on the point is available in the sister paper. We have indicated this in the below response.

Sincerely,

Todd

Report #1

Submitted on 22 Mar 2022

Anonymous referee #1

Anonymous during peer-review: **Yes** **No**

Anonymous in acknowledgements of published article: **Yes** **No**

Recommendation to the editor

1) Scientific significance

Does the manuscript represent a substantial contribution to scientific progress within the scope of this journal (substantial new concepts, ideas, methods, or data)?

Excellent **Good** **Fair** **Poor**

2) Scientific quality

Are the scientific approach and applied methods valid? Are the results discussed in an appropriate and balanced way (consideration of related work, including appropriate references)?

Excellent Good Fair Poor

3) Presentation quality

Are the scientific results and conclusions presented in a clear, concise, and well structured way (number and quality of figures/tables, appropriate use of English language)?

Excellent **Good** Fair Poor

For final publication, the manuscript should be
accepted as is
accepted subject to technical corrections
accepted subject to minor revisions
reconsidered after major revisions
rejected

Were a revised manuscript to be sent for another round of reviews:

I would be willing to review the revised manuscript.

I would not be willing to review the revised manuscript.

Suggestions for revision or reasons for rejection (will be published if the paper is accepted for final publication)

See comments in pdf

Referee Report: bg-2021-245-referee-report.pdf

This study focuses on the preservation and in situ production of isoprenoid GDGTs near hydrothermal vents. It shows that GDGTs are quickly recycled but that sedimentary production of GDGTs causes an overprint, impacting proxies such as the TEX86. The authors propose a way to correct for this bias in modern systems.

I found this study presenting some interesting data which are suitable for publication. I have a few comments, however, which I like the authors to address and I detail these below. My main comment is that the concept presented is not new, i.e. the paper of Schouten et al. 2003 already showed that pelagic GDGTs are rapidly degraded in the Guaymas basin and that there is in situ production of GDGTs by thermophilic archaea causing a change in sedimentary GDGT distribution. This should be more discussed at several places in the manuscript.

We realize the oversight and have made several modifications to the text. For example, line 148 of the introduction now reads, “Building on the results of Schouten et al. (2003), it was observed that these lipids can become heavily turned over in the hotter portions of the vent site where they rarely survive long enough to become cracked into hydrocarbon biomarkers such as biphytanes and derivatives of biphytanes.”

Further nuancing should also be that this overprint is likely unique for hydrothermal systems due to the exceptionally high degradation rates and unique lipid profile of thermophilic archaea and that the correction suggestion is likely rarely possible or needed.

That may well be the case, but we suggest that this latter point is still conjecture. Bentley et al (2022) presents a compelling case that archaeal lipids have a global preservation anomaly compared with other lipid classes. Ecological evolutionary changes may account for some, but not all, of the lack of archaeal lipids in paleo sediments.

Detailed comments:

L. 60 (abstract) and 663 (conclusion). I think you can use the correction in modern systems (provided all info is there) but not in paleoclimate studies due to a lack of constraints of several parameters. Furthermore, calibration studies use surface sediments (upper 1 cm), which also in your case is hardly affected by sedimentary overprint. So, which studies would actually really benefit from this correction? Can you provide an example from the literature and eg perform this correction?

The correction has been made to limit the expected use of this model to modern systems has been made.

L. 64-138. The introduction could be shortened in my view. The TEX86 was initially proposed as an SST proxy but ever since the 2007 publication of Huguet et al. in Paleoceanography this notion has been nuanced and plenty of publications as cited by the authors have shown that the signal is likely generated from (upper) subsurface waters (though might still be used to infer surface conditions). So lines 64 to 138 can be shortened to less than a page by just summarizing that TEX86 was initially proposed as an SST proxy but plenty of studies have shown that it is a subsurface signal. Then you can quickly move to the topic of your paper, i.e. pelagic versus sedimentary source for GDGTs.

L. 116: Sinninghe Damsté It is unclear what is being requested here.

L. 139-145. I challenge the notion that these two studies showed that in situ sedimentary production changed TEX86 values. In fact other studies (eg Lengger studies, Schouten et al., 2010) have shown that the IPLs in the study of eg Lips and Hinrichs are not derived from living Archaea but fossil IPLs because of the excellent preservation of pelagic glycosidic GDGTs (eg Xie et al., 2013, PNAS). This aspect should perhaps be more clear from the introduction, i.e. the simple presence of glycosidic GDGTs in sediments is in itself not proof of in situ sedimentary production.

Point taken and the information is now corrected.

L.162-181. The justification of using the Guaymas basin to determine in situ production is not only the active subsurface population but also that the GDGT fingerprint of these thermophilic sedimentary archaea is different from the pelagic background, as first established by Schouten et al. (2003) for this basin. Would be useful to clarify this here. If the distributions would have been similar, the signal would have likely been difficult to pick up.

The point has been added to the introduction.

I. 299. Outliers. We do not know what this is referring to.

I. 302. What is 'relatively high' ? Would be good to get some quantitative information on this.

The rate is provided.

L. 315. The reported concentration is in too many significant numbers. With an error of +/- 4 ug/g sediment you should report concentrations rounded to whole numbers. Change throughout the manuscript.

While this may be true for many instrumental set-ups. We feel it is better to stay consistent with Bentley et al. (2022), which underwent a long and extensive peer review process.

L. 318-324. Would be useful to mention the estimates of degradation of GDGTs of Schouten et al. (2003) here for the same site: they actually found a difference between crenarchaeol versus other GDGTs. Indeed, what I am struck by is that the drastic change in GDGT distribution of GDGTs in Guaymas basin, in particular the dominance of GDGT4 with high temperatures, of Schouten et al. 2003 is not observed here. GDGT4 may co-elute with crenarchaeol under your UHPLC conditions so only mass spectra can distinguish the two. I presume this was done and also that if there was a coelution that the amount of GDGT-4 was corrected for the contribution of the $[M+2+H]^+$ ion of crenarchaeol. Can you speculate why there is this difference between the two studies?

All detected lipids are quantified based on the two prominent adducts $[M+2+H]^+$ and $[M+2+NH4]^+$. Schouten et al. 2003 did not examine lipids from Cathedral Hill. As Guaymas Basin is host to many vent sites with different microbial communities and heat flow conditions it would be rather surprising that the two studies would find identical lipidomes. Please see Bentley et al. 2022 for further information regarding archaeal lipid diversities.

I. 369. Schouten et al. (2010) was not the first one to use this equation to model degradation rates of lipids. Please refer to the original literature.

The citation is now prefaced as (e.g.).

L. 463-467. Note that hydrous pyrolysis experiments typically last 3 days only. Pretty sure the Guaymas sediments were exposed to longer time periods and hence thermal cracking may have occurred at lower temperatures.

The point would seem to be pretty obvious. Nonetheless we have tried to include the comment in the text.

L. 479-489. Would be good to mention here that the same was observed in the Guaymas Basin by Schouten et al. (2003) who observed an increase in the RI of core lipid GDGTs with in situ temperature. Interestingly, they got a higher RI than observed here due mainly to a much lower relative abundances of crenarchaeol and a very high GDGT4. Can you speculate why your RI remains lower except for the deepest point in core 4?

Bentley et al. (2022) talks much more about the RI with respect to crenarchaeol.

L.527-529. I apologize for repeating myself here, but was this not already shown in Schouten et al., 2003, i.e. a replacement of pelagic GDGTs with GDGTs of thermophilic archaea?

We would contend that it was inferred based on lipid distributions, which otherwise could be due to community composition.

Other studies which have documented overprints of pelagic GDGTs are those of sedimentary sulfate-methane transition zones where GDGTs of AOM archaea (GDGTs 1-3) overprint GDGTs and thereby impact TEX86 values. This is one of the reasons why the MI index was developed by Zhang et al (2011) to check for

this overprint. Would be good to mention this example here.

The references and statement to original discovery have been added to the concluding sentences of this section.

L. 626-627 and 645-647. It has been demonstrated here that in hydrothermal systems there is an overprint of GDGTs in sediments. This is a very important nuance as in these systems degradation rates are substantially higher due to the higher temperature (both biotic and abiotic degradation) as well as production of archaeal GDGTs with a distinctly different profile.

The conclusion has been re-written to hopefully convey this point.

Report #2

Submitted on 06 Apr 2022

Anonymous referee #2

Anonymous during peer-review: **Yes** No

Anonymous in acknowledgements of published article: **Yes** No

Recommendation to the editor

1) Scientific significance

Does the manuscript represent a substantial contribution to scientific progress within the scope of this journal (substantial new concepts, ideas, methods, or data)?

Excellent Good **Fair** Poor

2) Scientific quality

Are the scientific approach and applied methods valid? Are the results discussed in an appropriate and balanced way (consideration of related work, including appropriate references)?

Excellent Good **Fair** Poor

3) Presentation quality

Are the scientific results and conclusions presented in a clear, concise, and well structured way (number and quality of figures/tables, appropriate use of English language)?

Excellent Good **Fair** Poor

For final publication, the manuscript should be

accepted as is

accepted subject to technical corrections

accepted subject to minor revisions

reconsidered after major revisions

rejected

Were a revised manuscript to be sent for another round of reviews:

I would be willing to review the revised manuscript.

I would not be willing to review the revised manuscript.

Suggestions for revision or reasons for rejection (will be published if the paper is accepted for final publication)

"The influence of near surface sediment hydrothermalism on the TEX86 tetraether lipid-based proxy and a new correction for ocean bottom lipid overprinting" by Bentley et al.

In this manuscript, Bentley et al. use data recently published elsewhere to assess the potential overprint of sedimentary tetraether production on the TEX86 proxy near a hydrothermal vent system. The authors detect substantial loss of tetraethers in some of the investigated sediment cores, but find no selective degradation of specific tetraethers. Nonetheless, TEX86 does not report the expected SST and instead seems to be influenced by benthic production. Finally, the authors suggest a model that corrects for the influence of benthic production and diagenesis/catagenesis.

In my opinion, this manuscript needs major revisions before it can be published. The text suffers from a lack of specificity when defining terms or laying out arguments, which often makes very difficult to assess any implications of arguments. Furthermore, although benthic production is identified as the major control on TEX86 overprints, there is a surprising lack of discussing ecological controls on TEX86 (i.e., community composition and growth conditions/substrate availability/habitat; these factors are indeed mentioned in the introduction). This aspect appears to be quite important given that the suggested correction – while successful at removing a certain degree of overprint – fails at reconstructing the actual measured SSTs, an aspect not given any further attention.

Comments

I. 34: The sentence should be rephrased to read "...of tetraether lipids produced by archaea and bacteria in soils and sediments..." or something similar. Thaumarchaeal GDGT proxies are not readily applied to soil samples (I guess the idea here is a comparison with bacterial brGDGTs?).

The correction to the wording in the sentence has been made.

I. 38: please specify "environmental and/or geochemical variations". **Done.**

I. 39: do temperature changes have to be driven by climate change? Other mechanisms could affect seasonality or temperature depth gradients (expansion/compression of the thermocline). **Correct. We have changed the wording of this sentence to include oceanographic effects.**

I. 45: the phrase "...across TEX86 GDGT lipid classes" is confusing here since IPLs have not been introduced yet. **The word classes has been deleted.**

I. 48: delete "when it is"? **Done.**

I. 50: TEXH86 has not yet been introduced – **The sentences above do discuss the ratio.**

I. 59: please specify "geochemical and physical conditions".

I. 63: how do you assess that TEX86 is "the most widely used" archaeal paleotemperature proxy? **The word most has been deleted.**

I. 64: the reference to Tab. 1 should probably be deleted here? **Done.**

I. 68: given that the entire introductory paragraph places an emphasis on marine systems (SSTs, ocean bottom sediments, marine planktonic archaea...), I would suggest to delete the reference to lake sediments. **Done.**

I. 72: Blessing et al.? **The typo has been corrected.**

I. 72 and 97: "TEX86-based lipids"? The index is based on lipids, not vice versa. **The change has been made.**

I. 77: is there a reference that would support this statement? **The sentence is linked to references and more follow in the proceeding discussion.**

- I. 82: change to “Mg/Ca or clumped ...”? **Done.**
- I. 84: introduce alkenone paleothermometry for readers not familiar with UK’37. **Done.**
- I. 92: Ho & Laepple (2016) do not discuss benthic production of GDGTs and there is substantial criticism of the 0-900m depth calibration suggested by these authors (see <https://doi.org/10.1038/ngeo2997>)
- I. 99-101: highly controversial. **That is the reason for this sentence being prefaced as “may also”.**
- I. 107: the use of “deeper waters” may be confusing for some readers who understand the qualifier to indicate depths below the thermocline/below the ammonium maximum. **The sentences has been changed to read: Here TEX_{86} dissimilarities appear to be driven by increases in the relative abundances of the GDGT-2 and isomers of crenarchaeol (see Lui et al., 2018; Sininghe Damsté et al., 2018) coupled with decreasing abundances of GDGT-1 and GDGT-3 below the thermocline and the ammonium maxima of the water column thereby producing a systematic reconstructed SST bias for deep-water surface sediments.**
- I. 116: this sentence should be more specific, Kim et al. do not calibrate against 0-900m depth. **The citation is removed.**
- I. 120: delete “by” **Done.**
- I. 122: abbreviation IPL not yet introduced in main text. **Corrected.**
- I. 130-134: sentence could be split into two more digestible sentences. **Done.**
- I. 139: “elevated sedimentation rates” in comparison to? **Elevated is changed to high.**
- I. 153-154: This sentence suggests that additional measurements were conducted for this study rather than specifically stating that the data from Bentley et al. (2022) are used. **The sentence has been qualified to improve clarity.**
- I. 213: The statement that TEXH86 is “for sediments outside the polar region” is somewhat misleading given the 15°C threshold. **Changed to low latitudes.**
- I. 235: why would MilliQ water be used for the final 3 “washes”, if the mixture is not liquid-liquid extracted again? **This section has been slightly modified for improved clarity.**
- I. 236: “gentle stream”? **The description seems OK to us.**
- I. 255-256: which mode was used? Targeted/untargeted? And what is the resolution of the Q-ToF? **The sentence has been modified to now read: The mass spectrometer was set to a 100–3000 m/z scan range in positive mode in an untargeted method with 10 ppb resolution to simultaneously resolve both archaeal IPLs and CLs.**
- I. 274: please express the reproducibility in %. Given that many GDGTs or IPLs have concentrations below 1 $\mu\text{g/g}$ sediment (Tab. S1 and S2), the absolute deviation is not informative (there are few IPLs that occur at >4 $\mu\text{g/g}$). **We do not understand what is being asked here.**
- I. 282: heterotrophic loss is not addressed in this section. **Now change to turnover.**
- I. 285-287: brGDGTs not mentioned above and numbering (typically Roman numerals) may be confusing here. **This has been corrected and a table with the data has been added to the supplemental.**
- I. 287-288: in Fig. 2, the concentrations in core 3 do not decrease systematically. **The working has change to nearly all.**
- I. 290: please add information how the age of the sediments was determined, fallout radionuclides? **Age estimates were based on expected basin sedimentation rates (Bentley et al. 2022).**
- I. 299: specify “extreme vent fluid conditions” **This has been changed to Due to the high temperature conditions of the vent fluids at Cathedral Hill.**
- I. 303: maybe rather “contain 0-4 cyclopentyl moieties...” **We are not seeing the purpose of this change.**
- I. 306: please add regression/correlation coefficient (or other statistical means) to demonstrate the “tight control” of porewater temperature. **Done.**
- I. 310-315: why are 1G-GDGTs with abundances of ~8 $\mu\text{g/g}$ DW in core 4 used for the study, but 2G-GDGTs with abundances of ~7 $\mu\text{g/g}$ DW are not further considered? Maybe expressing the abundances in % would be more convincing here. **This has been corrected to include the additional lipid class.**

I. 317-320: BIT, CBT, MBT not introduced yet. Overall, these two sentences have very limited meaning, please be more specific. Which paleoclimate records (these proxies are used for different purposes), what is “environmentally scaled loading” (which parameters), what are “other ocean floor sediment systems” (this could mean anything from lateral influx to benthic production, I guess)?

I. 323: when referring to “parallel lines”, you actually mean to refer to similar slopes? **This has now been made more clear.**

I. 332: there seems to be something wrong with the numbering of the equations throughout the text. **This has been fixed by eliminating the equation numbers from Table 1**

I. 339-341: it would be helpful to quickly summarize the results here/provide some statistical analysis, so the reader can follow the argument made.

I. 340: “similar slopes” **Correction made.**

I. 342: RI and MI not introduced yet. **Correction made.**

I. 346-347: this trend is primarily driven by the TEX86 values obtained for core 1 whereas core 2 displays TEX86 values mostly in agreement with the values determined for the habitable zone. It would be nice to see this aspect and possible explanations for this pattern addressed in the text. Do these data points correspond to the highest porewater temperatures? **This comment does not reflect our data. Core 2 has 3 samples that are in the habitable zone. The rest of the core samples fall outside that condition. The trend is predominantly formed by cores 1 and 2.**

I. 353-354: please be more specific, which isotope systems? **Carbon has been added for clarification.**

I. 356: cGDGTs not introduced yet. **Now corrected.**

I. 358: MI has been used in other than cold seep systems as well. Somewhere in this paragraph, it should be mentioned that contributions from Euryarchaeota are implied in such high-MI settings. **Noted.**

I. 366: please be more specific, sulfate reduction rates? **Not clear what the reviewer is looking for in this revision. AOM to our knowledge is a coupled system of methanogenesis and microbial sulfate reduction. So the metabolism is drawing on carbon and sulfur.**

I. 367-369: I would not define 0.2-0.38 as “very low” values. How is any thermal control tested, please be more specific (include simple statistical analysis such as regression/correlation coefficients). **The word very has been deleted. Regression statistics are presented on the figures.**

I. 369: what does “this” refer to, the MI values or the lack of thermal controls? **This has been replaced with low values.**

I. 378: please add reference when citing the upper thermal limit of life. **Done.**

I. 396: McClymont et al. (2012) used POM from sediment traps for this purpose, not sediments. **Correction made.**

I. 408: if they were determined by vertical fluxes alone. **Unclear what the requested change is.**

I. 411: the “period of ~37.5 to 75 yrs” refers to which depth interval? **This was provided.**

I. 412: the “shallow-surface sample” is typically described as core-top. **The change is made.**

I. 415: please specify “larger ranges”. **The range is now added to the text.**

I. 417: 10–21 cmbsf, i.e., the non-habitable zone mentioned above. **Unclear what is being asked here.**

I. 419-420: this could also be triggered by other processes than temperature that influence the archaeal community composition. **The sentence is now changed to read “The fundamental driver for the proxy’s is likely influenced by the archaeal community composition that is responding to their exposure to *in situ* vent fluid temperatures (Figure 4).”**

I. 442-443: alternative ecological controls should be listed and discussed here (I assume data such as O₂ or NH₄ concentrations are not available to be compared to the TEX86 values?). **Correct.**

I. 451: abbreviation IPLTEX86 should be introduced. **Done.**

I. 454-459: the 1G-GDGT-derived TEX86 values can again also be influenced by other ecological factors than temperature alone. **We consider the strong thermal gradients of the porewater to be the primary**

controls on the ratio, how that intern is reflected in the subsurface archaeal community taxonomy is outside the scope of this project.

I. 458: the “living lipid pool” – it has been well established that IPLs are much more persistent than initially thought and that sedimentary IPL pools can contain substantial fossil components. At vent temperatures of +100°C in sulfidic conditions, headgroups will not be stable outside the host cell for any length of time. Therefore it is correct to call this the living lipid pool.

I. 461-480: I am missing a discussion about the other potential ecological controls on TEX86 values and, as is, the discussion only distinguishes ‘living’ and ‘dead’ archaea, but does not address the influence on TEX86 through changes in the contributions of Thaumarchaeota / Euryarchaeota (ANME?)/Bathyarchaeota etc. as well as potential ecotype contributions (even if the observed TEX86 trends are entirely determined by temperature, these archaea may not have the same response to adapt their membrane to temperature). The sentence has been reworded to read: In this regard, 1G-iGDGT IPLTEX86 ratio appears to also largely measure in situ porewater temperatures but may also be further impacted by the archaeal community ecology of the vent system.

I. 479: if TEX86 “may entirely record ocean bottom sediment porewater temperatures”, why are the calculated SSTs not maxed out in all samples (e.g., at 0.9=40°C)? We did not say that. If it does, however, than TEX86 could also be used (calibrated) as a porewater thermometer, which would be immensely useful.

I. 481: it would be very helpful if the symbols in Fig. 5C would allow the reader to differentiate the origin of the samples in the clusters. One would assume that cluster C is only made up of samples from the non-habitable zone, but this does not seem to be the case (includes 4 samples from core 3, which is ‘habitable’ in its entirety. Are there any trends, are these the lower samples or is there no pattern whatsoever? We consider the attempt to draw out extra information within this plot to be a reach resulting in over-interpretation of our data.

I. 490: above, an argument is made that diagenesis is not selective. The data largely support non-selective turnover of lipids.

I. 498: Is d0-n an additional depth scaling parameter or a descriptor? Some initial info should be added here. This is a descriptor.

I. 505-509: what is the definition of the “deeper water column” here? Changed to upper water column. This information is essential to understand the implication here. Does “deeper water column” refer to the bathypelagic or indeed to the thermocline (as is done in the introduction). In case of the latter, this is a major pitfall of the suggested correction since the majority of IPLs/GDGs is synthesized in the thermocline/in the ammonium maximum. In case of a bathypelagic origin, the contribution of GDGTs may indeed be negligible, but then I wonder why this point is not made earlier in this section and the term bWCTEX86 is not simply omitted from equation 7. From where in the water column most of the lipids are sourced is indeed not tractable in this study. It is instead lumped as being from the water column or from the sediments. We separate the two into different parts of the water column in equation 6 and 7 on theoretical grounds for future studies.

I. 514: in line 314, the number is ~7 µg/g DW. We do not follow this comment.

I. 514-515: I understand the mass balance argument, but the fact that 2G-GDGs only come with 1 or 2 cyclopentyl moieties could actually bias TEX86 much more profoundly (thus, is not a good argument). This has been eliminated by incorporating the 2G-GDGs.

I. 516-517: I find this sentence about ‘increasing accumulation’ confusing considering that there is significant loss of GDGTs in most cores. Increasing has been changed to progressive.

I. 517-524: I am getting lost here, there is too much information missing. What is the definition of “allochthonous”? Changed to water column input of lipids. Why sum the GDGT concentrations for the entire depth interval (as implied by the upper bound n) if the goal is to correct the TEX86 values in a distinct depth interval? Shouldn’t the correction only scale the abundance of GDGTs in that particular

interval against the original input rather than summing the concentration in the total sediment column? No. The subsurface biosphere perpetually exists in all the unconsolidated sediments across time and is living dividing and dying. The deeper the sediment the more subsurface lipids it has collected.

Why would the core-top not already have lost some of the IPLs? The diagenetic time step would be lost. We admit that the offset is not perfect. It is not calibrated to the exact sedimentation rate of this system (which is not determinable by us). So the offset represents the estimated time delay for IPL hydrolysis.

Sedimentation rates of 0.4–0.2 cm/yr imply that a 0-2cm core-top integrates 5-10 years of deposition.

I. 526-532: this section is at odds with the discussion regarding the kinetic model and the last two sentences directly contradict each other "...the selectivity of lipid classes being adsorbed to a protokerogen is undeterminable. More importantly, for this site it is insignificant...". The sentence has been rewritten to remove the ambiguity.

I. 536-538: there is no discussion on IPL turnover rates in section 3.1 and only a sentence linking GDGT turnover rates to hydrothermalism (citing the findings of Bentley et al. 2022). How GDGT and IPL concentrations compare downcore would be a useful addition to Fig. 2 or at least it could be expressed here using statistical means (Kruskal-Wallis). We find this to be excessive given the trend is so obvious.

I. 557: write "partial least squares" Done.

I. 564-565: what is the explanation for this 3°C offset relative to the actual SST? This seems to be an important discussion to be had given that the suggested TEX86-correction effectively accounts for a similar offset (some 3-4°C). This has now been eliminated do the reworking of the model.

I. 566: how is a "high degree of influence" assessed? It will be worthwhile to compare the offset in light of the calibration residuals. This request would entail a great deal more text to explain the statistics.

Residual analysis could be useful, but it would also make for a more complicated write-up for what is not the primary points of the paper.

I. 568-569: please rephrase this sentence to be more meaningful. The fact that the "microbial community influences TEX86 measurements" is the reason it was proposed as a proxy. Be more specific. Correction made to mark benthic community.

I. 575: The caption should state that reconstructed SST are combined SS+WC SSTs at stated in the text. Done.

I. 597-598: what is the meaning of "global changes by 2–4°C mean completed deglaciation" and, more importantly, which reference does this statement refer to?

I. 599: above, it is explicitly stated that the deeper water column is not considered in this study. This is changed to upper water column.

I. 601-604: the overprint in other sedimentary systems may likely be much lower, especially in core-top sediments (less vivid benthic communities in non-seep settings, substantially less diagenesis/catagenesis) The comment is noted, but is a point of reflection for all readers of this paper.

Figures

Overall, the quality of the Excel figures is not great. The data in Figures 3-6 will not be discernible for people with deutanopia, I would recommend either using different colors or different symbols to differentiate cores. The grid lines in Figures 3-6 also obscure the data (I don't think they are needed, but if the authors prefer them, they should recede into the background). References to background colors also do not correspond to the colors displayed (e.g., pink in Fig. 3 or blue in Fig. 4).

Some other remarks

The use of the phrase "lipid classes" throughout the text is incorrect when referring only to differences in the biphytane chains rather than headgroups. The word class has been dropped from the text.

The phrase “loading” has a statistical connotation and may be confusing. **The wording has been changed.**

There is an inconsistent delimiter use throughout the text when reporting numbers (and it often suggests much higher precision than can be expected).

The authors should give the text a final read to check spelling/grammar/punctuation and identify text residuals from previous manuscript versions.