Authors’ Response to Anonymous Referee #1

General comments

Referee’s Comment: Bhattacharya and colleagues obtained a very impressive set of metagenomic and metatranscriptomic data in combination with physicochemical and isotope geochemical measurements. The authors investigate differences in the sediment communities of permanent and seasonal oxygen minimum zones and present shifts in community structure and gene transcripts with depth as well as between the sites. The sampling, the generation, processing and analysis of the data, as well as the sequence archiving was performed carefully. Yet the presentation of the results needs to be improved. In my opinion, the manuscript needs to be shortened (substantially), better structured, and jargon needs to be avoided. The interesting and relevant messages of the study are buried in details, percentages and values. The study is difficult to access in part because the manuscript lacks visualizations (e.g. of the gene transcript abundance). It is recommended to structure all manuscript sections based on compelling objectives, find a red thread that guides the reader through these amazing data, and present only those results that are important to the major findings and conclusions and remove everything else. Please find detailed suggestions below.

Authors’ Response: We thank the reviewer for appreciating the phenomenology and scientific approaches underlying the study. We also agree with the concerns of the reviewer, so have now dealt with each one of them in the revised manuscript.

Authors’ Changes in Manuscript: In the Revised Manuscript we have

- re-structured the text by simplifying the language,
- improved the presentation of the results involving percentages and values (by incorporating new figures for better visualization of metatranscriptomic and slurry incubation results),
- moved all the supportive details to the SI, and focused the manuscript towards the main findings of the study.

Specific comments

Referee’s Comment: A picture speaks a thousand words! That is also true for figures. Instead of listing all the taxa, concentrations, abundances etc, please try to visualize. It is so much easier to grasp the critical information that way. For example, the transcript data: The authors have such an incredible dataset. I am wondering why the impressive gene transcript abundance data is not shown in a dedicated figure. The trends become much clearer when all these abundances are shown along the depth profile, or across samples. For inspiration on representation of transcriptomic data check, e.g., these papers:
https://sfamjournals.onlinelibrary.wiley.com/doi/10.1111/1462-2920.14806
https://sfamjournals.onlinelibrary.wiley.com/doi/full/10.1111/1462-2920.13895
https://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-019-0664-z

Authors’ Response: We agree, and have now edited the manuscript as suggested.

Authors’ Changes in Manuscript: In the Revised Manuscript we have now

- incorporated new figures for the visualization of the results pertaining to the metatranscriptomic analyses and slurry incubation data (abundances, concentrations, etc.),
- All the lists of taxa that were included to describe the database compositions have now been shifted to the SI,

Referee’s Comment: L557ff: Can those methane concentrations be shown in a figure?

Authors’ Response: We agree, now fixed.

Authors' Changes in Manuscript: New figures have been included to graphically represent all the concentration values that were involved in the slurry incubation data.

Referee's Comment: L583ff: For me, and I think most other readers, it is very difficult to grasp all these percentages and make sense of them, compare them, interpret them. Can this be visualized as well?

Authors' Response: We agree, now fixed.

Authors' Changes in Manuscript: New figures have been included to graphically represent all the percentage values that were involved in the metatranscriptome analyses.

Referee's Comment: Please try to avoid all terms that are very specific for a small area of science (jargon). L479-L489: This section is full of jargon: discontinuous reduction, Gaussian distribution, exponential-decay zone, probability density function, etc. Please avoid using these terms, because most of the readers will not know what they mean. Text sections become very difficult if not impossible to understand when they are full of jargon, and then the entire manuscript becomes hard to digest. In addition, when a reader gets lost in these details the overall message gets lost too. The rule of thumb here is: If for example the Gaussian distribution is really important for the study then it needs to be elaborated and explained with words, concepts and examples everybody understands. If it is not critical for the study, remove it, or move it to the SI for the experts that may be interested in the details.

Authors' Response: We agree, so have conceptually elaborated and explained the mathematical terminologies by adding simpler descriptions.
In this context it is noteworthy that the probability density functions (equation fitting) determined and presented for the individual data series critically authenticate the phenomenon "synchronous population fluctuation of multiple metabolic types along the perennial OMZ cores", a main finding of this paper.
If we don't check the population fluctuation trends for potential equations and interpret them mathematically, people would not accept what we are calling increases, decreases or a bell-shaped (Gaussian) distributions are indeed so. From a statistical point of view, people would argue that if still finer resolution data are generated then what is the probability that the trends we are seeing now would still remain so. To answer this it is indispensable to check out probability density functions for the trends observed; if the equations were not obtainable even after carrying out the statistical tests then the envisioned phenomenon of synchronous population fluctuation for the different anaerobic metabolic types along the perennial OMZ sediment cores would have been completely on shaky ground.

Authors' Changes in Manuscript: We have now conceptually elaborated and explained the mathematical terminologies by adding simpler descriptions.

Referee's Comment: L490-517: This section too is very hard to digest. Is all this detail really necessary? To be honest, I got lost in the equations right away, and was unable to say what the message of the paragraph is even after reading it repeatedly (I couldn't see the proverbial forest from all the trees). Please consider a broad audience.

Authors' Response: As stated above, the section between lines 490 and 517 narrates a central finding of the paper: sulfate-reducers, methanogens, ANME and acetogens predominate in the top-layers of perennial OMZ sediments and decline via synchronous fluctuations along the cores. These trends are reverse in the seasonal shallow coastal OMZ sediment horizon, so are crucial for microbiome discrimination in the two sedimentation / oxygenation regimes. Whilst the probability density functions (equation fitting) determined and presented for the individual data series critically authenticate the phenomena, we agree that the statistical concepts introduced needed simpler explanations for general readership.

Authors' Changes in Manuscript: We have now conceptually elaborated and explained the mathematical terminologies by adding easier descriptions.

L520-545: Same in this paragraph

Authors' Response: The anaerobic sulfur-oxidizing bacteria not only have crucial roles in the biogeochemistry of these sediment horizons but also hold critical implications for microbiome differentiation in the two sedimentation / oxygenation regimes. Accordingly, it was necessary to explore whether potential statistical correlations existed between their population fluctuations and the population fluctuations of the other metabolic types; also, their population fluctuation trends needed to be authenticated by searching for probability density functions, i.e. finding potential equations that define the trends.

Authors' Changes in Manuscript: We have now conceptually elaborated and explained the statistical concepts and mathematical terminologies by adding easier descriptions.

Referee's Comment: In addition to an amazing dataset the authors have some very interesting findings. E.g. (L635:"coexistence and covariance of sulfate-reducers with methanogens, ANME and acetogens"). However, those too get lost in all the details. It is highly recommended to shorten and declutter the manuscript. Find out what are the most interesting findings, what is the red thread that keeps the reader's (and writers) attention. Then start by introducing compelling objectives, and then follow these objectives in the results and discussion. Everything that is not needed for the core story should be removed or moved to the SI. There is a lot of potential to shorten the manuscript, I would say it should be at least a third shorter, if not half.

Authors' Response: We agree, and have now considerably streamlined the discussion and shortened the manuscript.

Authors' Changes in Manuscript: As stated in Authors' Response

Referee's Comment: The authors often include results in the discussion section. These sections should either be separated, i.e. the discussion only refers to the results, but contains very little values and numbers anymore. Or they should be combined in a Results and Discussion section. In any case, it would greatly help if the discussion was structured based on objectives that were introduced earlier. Presently the discussion is very broad, and I could not see where the readers attention/focus is drawn towards, i.e. what are the major take home messages.

Authors' Response: We agree, and have now considerably streamlined the discussion section.

In this context, however, it is noteworthy that there are no major, or actual, results in the discussion section.

A few current data that were only relevant in the background (perspective) of the discussions (phenomena envisioned) have been provided here in the form of Supplementary Tables.

We also agree that the whole body of geochemical data available for the three sediment cores have been mentioned in the discussion section. But since these are all published data (adequately cited as Fernandes et al., 2018, 2020) they cannot be put under Results; this said, their recapitulation is also central to the current revelation of distinct microbiome structures and functions, amidst comparable pore fluid chemistries, in the sediment horizons underlying the seasonal (shallow coastal) and perennial (deep sea) oxygen minimum zones, so they cannot also be one away with.

Authors' Changes in Manuscript: We have now considerably streamlined the discussion section as per the objectives spelt out in the Introduction, and cut back particularly on the detailed values for the geochemical data.

Referee's Comment: L615-630: This reads like a summary of the results section. In general, this is welcomed, as now the trends become clear, but it is still a results part, a discussion goes beyond that.

Authors' Response: We thank the reviewer for appreciating the importance of having this recapitulative paragraph to spell out the final trends clearly.

Authors' Changes in Manuscript: We have further streamlined this paragraph in sync with the objectives of the work and eventual phenomena discovered.

Referee's Comment:
L700ff: This contains a lot of results again.
L715-742: Now this section even contains isotope geochemical results that have not yet been mentioned before. There is no mention of a context, so this entire section needs to be moved to the results section. Also, there is way too much information, too many numbers, percentages, etc.
L757ff: Results

Authors' Response: As we have already mentioned above, the whole body of geochemical data available for the three sediment cores is given in the discussion section. Since these are all published data (adequately cited as Fernandes et al., 2018, 2020) they cannot be put under Results; yet, their recapitulation is also central to the current revelation of distinct microbiome structures and functions, amidst comparable pore fluid chemistries, in the sediment horizons underlying the seasonal (shallow coastal) and perennial (deep sea) oxygen minimum zones, so they cannot also be removed altogether.

Authors' Changes in Manuscript: We have now removed the detailed values for the geochemical data.

Technical Corrections

Referee's Comment: L38: the influence of bottom water oxygen on benthic microbial communities is not entirely unknown (e.g. Jessen et al. Sci Adv 2017), so it is recommended to write "are less understood" or "are still unclear" etc.

Authors' Response: We agree, now fixed.

Authors' Changes in Manuscript: As mentioned in the authors' response.

Referee's Comment: L81: This section should include mention of the study by Jessen et al 2017 Sci Adv, as it is presenting findings on microbial communities and organic matter in sediments along an oxygen gradient in the Black Sea

Authors' Response: We agree, Jessen et al. 2017 now mentioned.

Authors' Changes in Manuscript: As mentioned in the authors' response.

Referee's Comment: L105: what is meant by hectic? Do you mean substantial?

Authors' Response: "Hectic chemoorganoheterotrophic activities" now reworded as "high levels of chemoorganoheterotrophic activities".

Authors' Changes in Manuscript: As mentioned in the authors' response.

Referee's Comment: L119: ANME only refers to the ANaerobic MEthane oxidizing archaea.

Authors' Response: We agree, now fixed.

Authors' Changes in Manuscript: "bacteria symbionts" deleted.

Referee's Comment: L167: Thank you for depositing the sequence data prior to submission, and also thanks for including all raw data and not just MAGs! That is best practice!

Authors' Response: We thank the reviewer for this positive reception.

Authors' Changes in Manuscript: Not applicable.

Referee's Comment: L227ff: These lists are a good place to start shortening the MS. It could be included in a supplementary table. The idea of summing up the taxa of a target guild may be clear to the reader without the need to present an exhaustive list.

Authors' Response: We agree, so have now moved these lists to the SI.

Authors' Changes in Manuscript: As mentioned in the authors' response.

Referee's Comment: L251: What about ANME-2d (Methanoperedens). It becomes increasingly apparent that this clade may be important at marine systems as well. E.g. in this very recent publication https://www.nature.com/articles/s41396-021-00918-w

Authors' Response: We have now upgraded all the metagenomic and metatranscriptomic data pertaining to ANME with analyses based on the searching of our sequence datasets against a comprehensive database of ANME genomes (MAGs) curated non-redundantly from NCBI and GTDB. That curated database contains all the ANME-2d MAGs sequenced thus far.

Authors' Changes in Manuscript: As mentioned in the authors' response.

Referee's Comment: L270: The equations for the reported Simpson and Shannon indices seem to be the standard equation used for these respective indices. Is there a reason why the authors chose to present the math in detail? Here too it may suffice to cite the respective papers and maybe elaborate on these metrics in the SI. For the reader it would be more helpful to explain the meaning of these metrics, their end members (min/max values) and how to interpret them. In my opinion it is more necessary to explain the terms, .e.g. evenness/equitability than show the equation. As an example, for me it helps to interpret a given index, if I am aware of the max/min values. E.g. that the minimum value Inverse
Simpson can be equals 1 (a pure culture) and the maximum value equals the number of observed species (which is the case when all species occur at exactly the same relative proportion).

Authors' Response: We agree, so have now shifted the mathematical details to the SI and provided conceptual explanations and examples in the main text.

Authors' Changes in Manuscript: As mentioned in the authors' response.

Referee's Comment: L312: From which studies did the two ANME genomes come from? Please reference. There should be quite a number of available ANME MAGs by now. Maybe these can be used as a database for your next studies.

Authors' Response: As mentioned above, we have now upgraded all the metagenomic and metatranscriptomic data pertaining to ANME with analyses based on the searching of our sequence datasets against a comprehensive database of ANME genomes (MAGs) curated non-redundantly from NCBI and GTDB. Accordingly, the web address (https://gtdb.ecogenomic.org/) and reference for GTDB (Parks et al., 2018) has been cited in the revised manuscript.

Authors' Changes in Manuscript: As mentioned in the authors' response.

Referee's Comment: L328: Was hydrogen added as an electron donor for hydrogenotrophic methanogens?

Authors' Response: The headspaces of the culture bottles contained hydrogen (10% v/v) as a mixture of N_2:CO_2:H_2 = 80:10:10 (v/v/v) was used to deoxygenate the medium, as well as to preset the Whitley H35 Hypoxystation (inside which inoculum addition, culture incubation and culture transfer were carried out) at 0% partial pressure of O_2.

Authors' Changes in Manuscript: The above details are already there in the reference that was cited for the methanogen-specific medium used (Whitman et al., 2006); still we have now mentioned these points in the revised text over and above the existing reference.

Referee's Comment: L356ff: A lot of the results in this section are redundant to the Sampling section in the Materials and Methods. Please shorten, or remove, at either section.

Authors' Response: We agree, now fixed.

Authors' Changes in Manuscript: Redundant information now removed from the Results section.

Referee's Comment: L364: What does BP stand for, please write out each abbreviation at first use.

Authors' Response: We agree, now fixed.

Authors' Changes in Manuscript: Full form of BP (before present) has now been provided.

Referee's Comment: L379: Which database and version was used for the taxonomic classification? I might have missed this in the methods, but if it is not mentioned, please include.

Authors' Response: For taxonomic classification of metagenomic reads the datasets were searched separately against the NCBI non-redundant (*nr*) protein sequence database (last access: 14 April 2020) as well as the four distinct databases of single-copy conserved marker proteins, which were specially curated from CheckM version 0.7.060 (Parks et al., 2015, last access: 22 December 2020).

Authors' Changes in Manuscript: This was already mentioned in Methods.

Referee's Comment: L387: Please refer to "relative sequence abundance" to avoid confusion with other measures of relative abundance.

Authors' Response: We agree, now fixed.

Authors' Changes in Manuscript: Text edited as suggested.

Referee's Comment: L392f: Generally, it is recommended to phrase it the other way around: "…xx % of metagenomic reads affiliated with Alphaproteobacteria…"

Authors' Response: We agree, now fixed.

Authors' Changes in Manuscript: Text edited as suggested.

Referee's Comment: L393ff: Personally, I think it is not necessary to list all these taxa. It makes the manuscript very long and voluminous and disrupts the reading flow. Unless a specific taxon needs to be highlighted, it will suffice to point the reader to a figure showing the clades and their relative sequence abundances.

Authors' Response: We agree and have now reduced the text as suggested.

Authors' Changes in Manuscript: Mention of all the relatively-less abundant taxa have been removed from the revised text and the overall difference in the microbiome profiles of perennial and seasonal OMZ sediments has been pointed out by referring to the heat map of Figure 2.

Referee's Comment: L480: I think microbial guilds would be a more widespread and better term than "metabolic-types"

Authors' Response: We agree that microbial guilds is a widespread terminology but since it does not specify that the groups here were unified based on particular metabolisms we prefer to retain the term "metabolic-types".

Authors' Changes in Manuscript: Not applicable.

Referee's Comment: L579: The number of reads is important, but not here. 1. Large numbers break the reading flow and 2.) including the read numbers does not add info that is necessary for the argument at this point. This section is about mcrA transcripts and so it should be focused on this.

Authors' Response: We agree, and have now moved the information on read numbers to Methods section.

Authors' Changes in Manuscript: As mentioned in the authors' response.

Referee's Comment: L580ff: redundant to materials section. Please remove.

Authors' Response: We agree, now fixed.

Authors' Changes in Manuscript: This redundant portion has been removed from revised manuscript.

Referee's Comment: L652: Could the absence of methane not be explained also by aerobic methane oxidation?
Did you check if you have aerobic methane oxidizers in the dataset (eg Methylococcales).

Authors' Response: We agree and thank the reviewer very much for vindicating our previous findings regarding aerobic methane oxidation in the same sulfidic cores SSK42/5 and SSK42/6 (please see Bhattacharya et al.: Aerobic microbial communities in the sediments of a marine oxygen minimum zone. FEMS Microbiol. Letts., 367, fnaa157, 2020).

Authors' Changes in Manuscript: The revised text now factors in the role of aerobic methane oxidizers by citing Bhattacharya et al., 2020.

Referee's Comment: L659: Is there a reason why a database like GTDB (https://www.nature.com/articles/nbt.4229?testing) was not used? It should contain plenty of ANME genomes/MAGs.

Authors' Response: We have now upgraded all the metagenomic and metatranscriptomic data pertaining to ANME with analyses based on the searching of our sequence datasets against a comprehensive database of ANME genomes (MAGs) curated non-redundantly from NCBI and GTDB.

Authors' Changes in Manuscript: As mentioned in the authors' response.