Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2020-11-AC1, 2020
© Author(s) 2020. This work is distributed under
the Creative Commons Attribution 4.0 License.

Biogeosciences Discussions
Open Access
EGU

Interactive comment on "Microbial communities associated with sediments and polymetallic nodules of the Peru Basin" *by* Massimiliano Molari et al.

Massimiliano Molari et al.

mamolari@mpi-bremen.de

Received and published: 4 April 2020

Authors' Response ('AR') to interactive comment on "Microbial communities associated with sediments and polymetallic nodules of the Peru Basin" by Massimiliano Molari et al.

Review by Beth N. Orcutt and Tim D'Angelo ('RC1')

RC1> This study documents the composition and relative diversity of bacterial and archaeal microbial communities inhabiting polymetallic nodules and surrounding sediment of the Peru Basin collected in 2015. The motivations for this study are

Printer-friendly version

Discussion paper

to determine if polymetallic nodules have unique microbial communities, as such seabed mineral deposits may be targeted for deep-sea mining. While there have been similar prior studies of microbial community composition of polymetallic nodules, those studies focused on areas in the northern and central Pacific Ocean where organic carbon deposition rates are lower. Thus, the new study from closer to an equatorial region with higher organic carbon export rates allows an analysis of how broader oceanographic properties impact microbial community diversity. The first major claim of the current study is that microbial diversity is higher in the surrounding sediment than in the polymetallic nodules. This finding is different from a recent survey of available data from polymetallic nodules and sediments of the comparable Clairon Clipperton Zone, which indicated that nodules and sediments had comparable levels of diversity: https://ran-s3.s3.amazonaws.com/isa.org.jm/s3fs-public/files/documents/deep_ccz_biodiversity_synthesis_workshop_report_-_final.pdf. We encourage the authors to consider the implications of these differences between studies, and if data processing steps could be part of this difference.

AR> We thank the Reviewers for their thorough and very helpful revision, and for pointing us to the results of the recent meta-analysis of microbial diversity data available for CCZ, which was not available at the moment of submission.

In the revised MS we will include the outcome of the workshop in the discussion by adding the following statement (added/replaced text in italics): 'Microbial communities associated to nodules are significantly less diverse than those in the sediments, and the decrease in diversity was observed both in rare and abundant bacterial types (Figure 1 and S1). This seems to be a common feature of polymetallic nodules (Wu et al., 2013; Tully and Heidelberg, 2013; Zhang et al., 2014; Shulse et al., 2016; Lindh et al. 2017). However, a recent meta-analysis of 16S rRNA gene diversity reports no significant differences between microbial biodiversity between nodules and sediments within the studied habitats (Church et al., 2019). Church and colleagues also pointed out that the findings are so far not conclusive due to the limited number of studies and

Printer-friendly version

Discussion paper

CC BY

differences in methods (e.g. PCR primers, sequencing approaches) which may also be a reason for the differences between the meta-analysis and the results of this study.'

RC1> Related to this part of this study, we caution that the workflow described in the methods may lead to inflated diversity metrics. The workflow described in L143-144 may allow lower quality sequence reads to pass the QC step, as most published workflows don't allow for sliding window PHRED scores of less than 28-30. For example, Dorado Outcrop basalt samples have around 1500 OTUs after filtering out low abundance/prevalence OTUs (described in Lee et al., 2016). We would expect a similar diversity on nodule samples exposed to bottom seawater but the samples described in this study have 5 - 14K OTUs per sample. Low quality reads can result in artificially large number of OTUs when using clustering-based methods. This has been documented by the developers of MOTHUR as a problem with low quality reads associated with old problematic Illumina chemistry kits. Even if there are true biological differences between Dorado Outcrop basalts and the samples in the current study that translate to different alpha diversity patterns, the presence of 525,169 singletons (as seen in Table 2) is a sign that there are likely issues with the QC steps of this workflow. We recommend that the authors revisit the sequence processing steps and consider using higher quality thresholds, and also consider using an algorithm that produces unique sequence variants (i.e. ASVs) instead of OTU clustering. Moreover, we wonder if there is a more streamlined way to present the information included in Figure 1, or if some of this information could be moved to supplemental materials? It seems like a bit of overkill to have 10 plots essentially showing the same information.

AR> We thank the Reviewers for the opportunity to clarify the bioinformatics workflow. We recognize that how it is reported in "Methods" of the MS may be misleading. As a standard procedure, we applied a score of 10 for bacteria and 13 for Archaea in quality trimming, but then the quality of sequences was assessed with the software FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). If the sequences did not pass the quality check, then they were filtered again with an appropriate quality score.

CC BY

All sequences used in the MS successfully passed the FastQC quality control, with average quality score per sample >34 for Bacteria, and >22 for Archaea. Thus, we believe that the high numbers of OTUs per sample was not caused by the introduction of low-quality sequences in the analysis.

In the revised MS the sequences workflow will be clarified as follows (added/replaced text in italics): 'Subsequently the TRIMMOMATIC software (Bolger et al., 2014) was used to remove low-quality sequences starting with the following settings: SLIDINGWINDOW:4:10 MINLEN:300 (for Bacteria); SLIDINGWINDOW:6:13 MINLEN:450 (for Archaea). In case of bacteria data this step was performed before the merging of reverse and forward reads with PEAR (Zhang et al., 2014). Merging of the archaeal reads was done before removing low-quality sequences in order to enhance the number of retained reads due to long archaeal 16S fragments. All sequences were quality controlled with FastQC (Andrews, 2010). Where necessary, more sequences were removed with TRIMMOMATIC with larger sliding window scores until the FastQC quality control was passed (average quality score per sample >34 for Bacteria and >22 for Archaea).'

We agree that the ASV approach has a higher taxonomic resolution than OTU clustering. However, for the purpose of this paper the resolution returned by SWARM (i.e. "species" level) appears appropriate, as it allowed to distinguish microbial communities associated with nodules and sediments (i.e. Figure 2). Furthermore, according to our experience in other studies >90% OTUs generated by SWARM overlap with variants (ASVs) identified with Dada2.

Regarding Figure 1, we do not fully agree with the Reviewers' view that the panels repeatedly show the "same information". Diversity indices and unique OTUs are reported for Bacteria and Archaea in upper and lower panels, respectively – hence the upper and lower rows of panels refer to independent data sets. As the reviewers are certainly aware, the diversity indices presented in the first four plots or each row differ in their ecological meaning: i) total number of OTUs (H0) provide overall information

CC BY

about alpha-diversity, ii) exponential Shannon (H1) considers species richness and equitability, iii) inverse Simpson (H2) accounts for dominant taxa, and iv) chao1 accounts for rare taxa. Here this was calculated with the same number of sequences for each sample, thus it is not affected by sequencing depth. The last plot shows the contribution of unique OTUs to the total number of OTUs and, hence again has a different focus. While the pattern shown in the different panels may be visually similar, we are still convinced that each panel contains important information and should be presented to the reader.

Therefore, we would like keep the figure in the revised version of the MS. Upon specific request by the editor we would, however move plots for Archaea to the supplementary information. They contribute only minor to the total diversity as compared to Bacteria, but we would stick to the full set of panels for Bacteria.

RC1> A second major effort of this work is to identify taxa that are differentially abundant between nodules and sediments. While the text in Lines 244-261 describes these differences, and Table 4 includes the result of Aldex2 analysis, we don't find that Figure 4 visually conveys these differences in an easily digestible way and suggest using differential log abundance plots to more clearly show which taxa vary between the sample types.

AR> We thank the reviewers for sharing their thoughts about improving the data representation in Fig. 4. In the revised MS Figure 4 will be replaced by a fold- change plot showing genera enriched in nodules compared to those found in sediments.

RC1> Another major focus of this work is the comparison of the microbial community structures between the Peru Basin nodules and those of the CCZ. I think that the paper could be improved by providing some kind of summary graphic or schematic that visually explains the differences, and their causes, as described in the text. For example, a cartoon illustrating that the lower OC flux in the CCZ leads to nodules that look like X with communities that look like Y and perform Z functions, versus how

CC BY

those conditions are different at the Peru Basin. Such a summary graphic could really help simplify the presentation of the major recommendations from this work in a way that is easy to grasp, which will be especially helpful for policy makers thinking about deep-sea mining.

AR> We appreciate the suggestion of the reviewers and we generally agree with them. However, such a scheme would need to be supported by a deeper analysis that would require a more comprehensive dataset. Such a generalized analysis unfortunately cannot be carried out at the moment due to limited number of studies and different sequencing methods applied to investigate microbial communities between nodules fields. The aims of this study were to explore the role of nodules in deep-sea microbial diversity, their potential role in ecosystem functions, and a comparison of our results with data available from other nodule field regions (i.e. CCZ). Our results highlight the importance of nodules in hosting specific and potentially functional important microbes, which differ from those reported for CCZ. While the number of samples and differences in methods do not allow a generalization, we felt the need to point out important ecological questions and hypotheses that are relevant in the deep-sea mining context, but that are not yet solved and should be addressed in future studies.

This consideration will be highlighted in “Conclusions” of the revised MS by adding the following statement (added/replaced text in italics): ‘However remarkable differences in community composition (e.g. Mn-cycling bacteria, nitrifiers) between the CCZ and the Peru Basin in microbial community composition also show that environmental settings (i.e. POC flux) and features of FeMn nodules (e.g. metal content) may play a significant role in structuring the nodule microbiome. Due to limitations in the available datasets and methodological differences in the studies existing to date, findings are not yet conclusive and cannot be generalized. However, they indicate that microbial community structure and function would be impacted by nodule removal. Future studies need to look at these impacts in more detail and need to address regional differences, to determine the spatial turnover and its environmental drivers, and the consequences

Printer-friendly version

Discussion paper

CC BY

regarding endemic types.'

RC1> A question: in the methods, there is mention of collecting samples for cell abundance determination, but such data are not presented in this paper. Is it possible to include such data? This would help to evaluate if the "hot spot" idea discussed in the paper correlates to cell biomass - i.e. is lower diversity correlated to higher biomass?

AR> Unfortunately, we do not have cell counts for manganese nodules. Originally, we planned to include cell counts (AODC and CARD-FISH) for sediments. However, these data are reported already in another study on the impact of mining on sediment microbial communities and their biogeochemical functions which was under revision at the moment of MS submission but will be available soon (Vonnahme T.R, Molari M., Janssen F., Wenzhöfer F., Haeckel M., Titschack J., Boetius A. Effects of a deep-sea mining experiment on seafloor microbial communities and functions after 26 years. Science Advances, in press). We will clarify this issue in the revised MS and point the reader to the Vonnahme et al. publication.

RC1> A suggestion: there is some mismatch between the 3 hypotheses posed in the introduction and the three objectives posed in the discussion section. The discussion text follows the outline of the objectives, but there is not explicit "testing" of the hypotheses proposed at the beginning of the paper, and also the discussion does exactly follow the objectives as proposed. For example, discussion section 4.3 discusses metabolisms inferred from the amplicon data, not what environmental factors structure the community, as would be assumed by how objective three is worded. We recommend bringing better alignment between the hypotheses/objectives and what the data actually address.

AR> We thank the reviewers for pointing out these inconsistencies.

The first and second hypothesis (Hp1 and Hp2) have been tested using statistical tests (as described in the Methods section). Hypotheses and "primary aims/objectives" are discussed in section 4.1 and 4.2, respectively. The "secondary aim" (previ-

Printer-friendly version

Discussion paper

CC BY

ously Hp3) was addressed by deducing potential metabolic functions and habitat features/preferences from the taxa that were significantly enriched in the nodules (based on statistical testing) and what we know from descriptions of closely related organisms. From these results and comparison with CCZ microbial data we suggested potential environmental factors that can have a major role in shaping the microbial community on nodules. Based on the limited available data, however, these suggestions cannot be rigorously tested. The "secondary objective" was mainly discussed in section 4.3, but also partially addressed in the previous two sections.

In the revised MS we will improve the alignment of hypotheses/aims with objectives by slightly refocusing the first and second hypothesis (Hp1 and Hp2) and by turning the third hypothesis into a secondary aim.

1) Hp1 [introduction]: 'nodules shape deep-sea microbial diversity and functions'

Primary Objective [discussion part 4.1]: compares the microbes of nodule fields with microbiota of deep-sea sediments in other ecosystems in order to identify specific features of microbial diversity of nodule fields.

2) Hp2 [introduction]: 'nodules host a specific microbial community and functions compared to the surrounding sediments'

Primary Objective [discussion part 4.2]: elucidates differences in diversity and in microbial community structure between sediments and nodules, and the potential implications for microbially-mediated functions

3) Secondary aim [introduction]: 'Another aim of this study was to investigate the nodule features that may play a major role in shaping microbial community composition.'

Secondary objective [discussion primarily part 4.3]: investigates the major drivers in shaping microbial communities associated with nodules

The modified hypotheses / aims are easily associated with the titles of sections 4.1, 4.2, and 4.3 of the discussion but we will make sure to refer back to the hypotheses /

CC BY

aims also in the text.

RC1> minor suggestions:

RC1> L1: Title could be more descriptive of what the study discovered

AR> According to reviewers' suggestion, the title of revised MS will be: "The contribution of microbial communities in polymetallic nodules to the diversity of the deep-sea microbiome of Peru Basin (4130 – 4198 meter depth)"

RC1> L15 - consider removing "need to"

AR> We will remove "need to".

RC1> L22 - Acidomicrobia, only one "i". To update throughout the manuscript.

AR> We are referring here to the class Acidimicrobiia within the phylum Actinobacteria in accordance with the NCBI taxonomy (https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=84992)

RC1> L78-79 - need consistency in the presentation of thousands of kilometers. In one instance, there is no punctuation; in the second instance, there is punctuation.

AR> We will correct this by removing punctuation.

RC1> L80 - missing a decimal point in 0.2-0.6%?

AR> Reviewers are correct - we will change this accordingly.

RC1> L123 - were any negative DNA extraction controls included in this study, since low biomass might have been expected? If yes, please describe.

AR> Yes, we had negative controls. This is mentioned in lines 129-130 and table 2 of the submitted version of the MS.

RC1> L141 - Is there a reference that shows why these trimmomatic SLIDINGWINDOW parameters were used? They seem relaxed and would allow for sub-par quality

Printer-friendly version

Discussion paper

CC BY

reads to pass the QC step. Most workflows don't allow for sliding window PHRED scores of less than 28-30.

AR> We give detailed information above in the general comments section, and we will provide additional information the revised MS.

RC1> L141 - recommendation to deposit your data processing pipeline to github or similar repository.

AR> We are currently exploring this with experts in our group. Once the code is made available we will add the information to the revised manuscript.

RC1> L144 - There is a comparative "while" statement describing the differences between how bacterial and archaeal sequences were merged, but the way it is worded, it appears to describe the same order of operations.

AR> We agree and will correct in the revised version of the manuscript as follows (added/replaced text in italics): 'In case of bacteria data this step was performed before the merging of reverse and forward reads with PEAR (Zhang et al., 2014). Merging of the archaeal reads was done before removing low-quality sequences in order to enhance the number of retained reads due to long archaeal 16S fragments.'

RC1> L174 - Transforming count matrices using the center-log ratio requires a strategy for replacing zeros with a pseudo count because the presence of zeros produces NA values. There is no zero-replacement strategy described in this workflow. The Bray-Curtis distance cannot be computed on data matrices that contain negative numbers. A Center- log-Ratio transformed count matrix contains negative numbers. CLR transformed data is usually ordinated using the Aitchison distance metric or the Euclidean distance. I am unclear on how these analyses were performed in the way that they are described. Was the data log10(x+1) transformed? That transformation is compatible with the bray- curtis distance. The resulting ordinations looks correct, but I think the description in the methods section is inaccurate. Could the authors provide a document

Printer-friendly version

Discussion paper

CC BY

with the code used to perform these steps?

AR> Totally right, indeed the Euclidean distance matrix was used and not Bray-Curtis (as also specified in caption of Figure 2). In the revised MS we will correct this mistake as follows (added/replaced text in italics): Beta-diversity in samples from different substrates and from the substrate in samples from different sites was quantified by calculating an Euclidean distance matrix based on centred log-ratio (CLR) transformed OTU abundances (function clr in R package compositions) and Jaccard dissimilarity based on a presence/absence OTU table.

RC1> L237 - these percentages are for all nodules in aggregate as an average, but does not show the variation between samples. I recommend including standard deviation plus/minus for each percentage.

AR> The percentage reported is not the average between/within groups, but it is the result of hierarchical clustering (function hclust in R package vegan) using the complete linkage method (data reported in Figure 4) and, hence, standard deviation cannot be reported. This information will be added to the revised MS as follows (added/replaced text in italics): The Jaccard dissimilarity coefficient was used to perform hierarchical clustering (function hclust in R package vegan, using the complete linkage method), and the dissimilarity values for cluster nodes were used to calculate the number of shared OTUs between/within groups.

RC1> L338 - could the differences in relative percentages of archaea between this study and prior studies be due to difference in DNA extraction, primers used, or sequencing approach?

AR> Previous data for Archaea are only reported by Tully and Heidelberg (2013) and Shulse at al. (2016) in the CCZ. In the first study a modified phenol-chloroform extraction method was applied for DNA extraction, universal primers (U515/U1048; targeting the V4 region of the 16S rRNA gene) for PCR amplification and Roche 454 Titanium platform for sequencing. Shulse and colleagues extracted DNA with the FastDNA Spin

CC BY

Kit for Soil (MP Biomedicals, USA), PCR amplification was carried out with universal primers (515f/805r; targeting the V4 region of the 16S rRNA gene) and sequencing with illumina MiSeq platform. These pipelines indeed differ from those applied in our study and reported in the Methods section: DNA extraction with FastDNA Spin Kit for Soil (MP Biomedicals, USA), PCR with bacteria (341F/785R; targeting the V3-V4 region of the 16S rRNA gene) and archaea primers (349F/915R; targeting the V3-V5 region of the 16S rRNA gene), and sequencing with illumina MiSeq platform. We agree with the reviewers that different methods applied make the comparison difficult, especially with data from Tully and Heidelberg (2013) where differences in methodology appear most pronounced. In the revised MS we will limit the comparison to data from Shulse at al. (2016) because differences in methods are limited to the choice of primers, which however amplified the same hypervariable region of 16S rRNA gene (V4) reducing biases in the comparison. This, however, does not change the overall difference in relative percentages of archaea in this study compared to previous work. We will further add the statement that we cannot rule out that the slight differences in methodology between the Shulse et al. and our study could be a possible explanation for the observed differences.

RC1> L359 - suggestion to add a clause to the end– of the sentence regarding nodules and sediments have distinct communities, stating that this observation is consistent with what has been found in earlier studies, and cite a few examples.

AR> We will revise the MS accordingly (added/replaced text in italics): ‘Analysis of community composition at OTU level shows that nodules and sediments host distinct bacterial and archaeal communities (Figure 2), as previously reported also for CCZ (Wu et al., 2013; Tully and Heidelberg, 2013; Shulse et al., 2016; Lindh et al. 2017).’

RC1> L413 - "reductive"

AR> We will replace “reducers” with “reductive”

CC BY