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

Associate Editor Dr. Denise Akob ('AE') Comments

AE> When referring to the 16S rRNA gene be sure to write out "16S rRNA gene" and do not use rDNA or shorten to "16S rRNA"

AR> we checked and changed where necessary. [Line 106]

AE> Make sure all taxa names are spelled correctly

AR> in the revised MS all taxa names have been double-checked and should now be correctly spelled

AE> Consider moving Table 2 to the supplemental information. Optional.

AR> we thank the Editor for the suggestion. In the revised MS Table 2 is now in the supplementary information as Table S1.

AE> Make sure OTUs is correctly abbreviated throughout the manuscript. In some places it is written OUTs.

AR> we checked and corrected this where it occurred (i.e., in the Table captions).

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 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 the CCZ, which were not available at the moment of submission.

In the revised MS we included 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 in microbial biodiversity between nodules and sediments within the studied habitats in the CCZ (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 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*.' [Lines 398-406]

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 was reported in "Methods" of the original MS may have been 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. All sequences used in the MS successfully passed the FastQC quality control, with an 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 is 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).' [Lines 157-166]

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 (as shown in Figure 2). Furthermore, according to our experience in other studies >90% of 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 the 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) provides overall information 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. In our initial response to reviewers, we expressed our willingness to move plots for Archaea to the supplementary information if requested by the editor because Archaea contribute only minor to the total diversity as compared to Bacteria. There was, however, no such request so far.

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 has been replaced by a fold-change plot showing genera enriched in nodules compared to those found in sediments. The original Figure 4 and Table 3 have been moved to the supplementary information (now Figures S2 and Table S3, respectively).

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 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. 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 the limited number of studies and different sequencing methods applied to investigate microbial communities in different 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 microbes with potentially important functions, 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 is highlighted in “Conclusions” of the revised MS by adding the following statement (added/replaced text in *italics*):

‘However remarkable differences in *microbial* community composition (e.g. Mn-cycling bacteria, nitrifiers) between the CCZ and the Peru Basin also show that environmental settings (*e.g.* POC flux) and features of FeMn nodules (e.g. metal content, *nodule-attached fauna*) 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 should address regional differences*, to determine the spatial turnover and *its environmental drivers*, and the *consequences regarding* endemic types.’ [Lines 485-493]

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 at the time of the publication of this study (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 clarified this issue in the revised MS and point the reader to the Vonnahme et al. publication. [Lines 125 and Lines 133-136]

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" (previously 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 improved 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 compared to the surrounding sediments'

Primary Objective [discussion part 4.2]: *elucidates* differences in diversity and in microbial community structure between sediments and nodules

3) Secondary aim [introduction]: 'Secondary aim of this study was to

investigate the nodule features that may play a major role in shaping microbial community composition and microbially-mediated functions.'

Secondary objective [discussion primarily part 4.3]: *investigates potential microbially-mediated functions* and the major drivers in shaping microbial communities associated with nodules

The modified hypotheses / aims in the introduction [Lines 108-113] are now easily associated with the titles of sections 4.1, 4.2, and 4.3 of the discussion. Lines 298-304, 305 382, and 428]

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 is now:

"*The contribution of microbial communities in polymetallic nodules to the diversity of the deep-sea microbiome of the Peru Basin (4130 − 4198 meter depth)*"

RC1> L15 - consider removing "need to"

AR> We removed "need to". [Line 18]

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)

We left this unchanged.

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 corrected this by removing punctuation. [Lines 82-83]

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

AR> Reviewers are correct - we changed this accordingly. [Line 85]

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. [Lines 143-145 of the current version]

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 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 in the general comments section above, and – as mentioned there – we provided additional information in the revised MS. [Lines 157-166]

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

AR> The revised version of the manuscript includes information on where workflow and scripts used for sequence analysis can be found. [Lines 169-170]

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 we corrected 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). *Low*-quality *archaeal* sequences *were removed after merging the reads* in order to enhance the number of retained reads *due to the increase in* archaeal 16S fragment *length*.' [Lines 160-163]

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 with the code used to perform these steps?

AR> Revier 1 is 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 corrected this mistake as follows (added/replaced text in *italics*):

"Beta-diversity in samples from different substrates and 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." [Lines 194-197]

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 has beeen 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*." [Lines 200-202]

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 a Roche 454 Titanium platform for sequencing. Shulse and colleagues extracted DNA with the FastDNA Spin 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 an 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 an 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 limited 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 [Lines 364-368]. This, however, does not change the overall difference in relative percentages of archaea in this study compared to previous work. We further added the statement that we cannot rule out that the slight differences in methodology between Shulse et al. (2016) and our study could be a possible explanation for the observed differences:

We cannot rule out that the observed differences in microbial community structure partly reflect the different sets of primers used in our study and by Shulse et al. (2016). As both primer sets amplified the same hypervariable region of 16S rRNA gene (V4) we assume that biases are small enough to justify the comparison. [Lines 368-371]

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 revised 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).*’ [Lines 388-391]

RC1> L413 - "reductive"

AR> We replaced “reducers” with “reductive” [Line 440]

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

Review by Anonymous Referee ('RC2')

RC2> General comments to authors:
The manuscript by Molari et al. describes the microbial community structure associated with sediments and manganese nodules from 3 and 2 sites, respectively, within the Peru Basin.

The authors find that Gammaproteobacteria and Alphaproteobacteria are the dominant bacterial classes in sediments and manganese nodules while all archaeal communities investigated were dominated by Thaumarchaeota. However, sediment and nodule communities were found to differ significantly at the OTU level, as assessed by calculating Jaccard dissimilarity. The authors note differences in the nodule community composition (specifically, a lower relative abundance of Archaea, and a different nitrifier community) in their study in the Peru Basin as compared with communities in the Clarion-Clipperton Fracture Zone (CCZ), where previous work on microbial community composition of nodules has been done.

The strengths of the manuscript include the following:
i. There is a lack of studies of the prokaryotic diversity in the surface sediments and nodules of the Peru Basin, which has different environmental conditions than the relatively well-studied CCZ. ii. The molecular and bioinformatic methods are well-documented and the microbial community analysis is thorough.

Weaknesses of the manuscript include the following:
i. The lack of metadata associated with the various sites makes interpretation of the differing community structures among sites difficult.

AR> We appreciate the suggestion of the reviewer. However, the primary aim of this study was not to investigate and explain the variability of microbial community between sites, but between habitats (nodules and sediments). Only sedimentary metadata (e.g. pigments and organic carbon content, porewater profiles, and porosity) are available for sites investigated, and not for nodules, which precludes the quantitative characterization of the nodule habitat setting. Thus, sedimentary setting alone does not help to understand differences in microbial community structure and diversity that are observed between sediments and nodules. The metadata available for sediments in the study area as well as the discussion of variability of sedimentary environmental settings and microbial communities will soon be published (scheduled publication at the 29th of April) in "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 point the reader to this publication in the revised version of the MS:

"*Sedimentary metadata (e.g. cell counts, pigments and organic carbon content, porewater profiles, and porosity) and a map of the study area are available in Vonnahme at al. (in press). Focusing entirely on sediments, that publication also includes a discussion of the variability of environmental settings and microbial communities.*" [Lines 133-136]

RC2> Specific comments to the authors:

RC2> Major concerns:

1. Page 3 – 4. Somewhere in this discussion of the CCZ versus the Peru Basin I think it would be helpful to briefly let the reader know the state of hypothetical mining in each of these regions. In the CCZ, the ISA has entered into contracts with various contractors for exploration for polymetallic nodules. Is this the case in the Peru Basin as well?

AR> We clarified this point in the revised MS modifying the introduction as follows (added/replaced text in *italics*):

*"*Nodule accumulations of economic interest have been found in four geographical locations: the Clarion-Clipperton Fracture Zone (CCZ) and the Penrhyn Basin in the central north and south Pacific Ocean, respectively; the Peru Basin in the south-east Pacific; and in the center of the north*ern* Indian Ocean (Miller et al., 2018). *To our knowledge the Peru basin is the only region that does not have exploration activities and plans for mining so far.*" [Lines 60-64]

RC2> 2. Page 5, line 113. "Samples were collected at three sites…" For clarity I think the authors should explicitly state in the text that nodules were only collected at 2 of these 3 sites.

AR> The Reviewer is right and we clarified this issue in the revised MS as follows (added/replaced text in italics):

"Manganese nodules where sampled, using a TV-MUC, or a Remotely Operated Vehicle (*ROV KIEL6000, GEOMAR, Germany): one nodule at Reference West and four nodules at Reference South*." [Lines 126-128]

RC2> 3. Page 5, line 115. "… called "Reference Sites." I suggest directly listing the Reference Sites here in the text instead of making the reader consult Table 1, especially since the authors refer to Reference South later in the text. Could change to "… called "Reference Sites": Reference East, Reference West, and Reference South."

AR> We modified the revised MS according to the Reviewer's suggestion as follows (added/replaced text in italics):

"Samples were collected at three sites outside the seafloor area selected in 1989 for a long-term disturbance and recolonization experiment (DISCOL;

Thiel et al., 2001), for this reason they were called “Reference Sites”*: Reference East, Reference West, and Reference South*.” [Lines 120-122]

RC2> 4. Page 5, line 116. Here a map of the Peru Basin (in addition to the Table already provided), with the study sites and DISCOL experiment sites marked, would be very helpful to the reader.

AR> An appropriate map is available in the Vonnahme et. al (in press) study mentioned above. We would suggest pointing the reader to that publication to avoid duplication. In our initial response to the reviewers we expressed our willingness to to provide a similar map also for inclusion in the revised version of the MS if requested by the editor. There was, however, no such request so far.

RC2> 5. Page 8, lines 226 – 232. “… significant differences were detected in sediment microbial community structure among the different sites… “Site” defined by geographic location and “Substrate” … explained a similar proportion of variation in bacterial community structure…” This was a bit surprising to me and this is where I think some physical/chemical/biological metadata about each site would be really helpful. If any is available, perhaps from other groups on the cruise, it would help add context to some of the observations here.

AR> A detailed environmental characterization of sites investigated and focused discussion of baseline condition (i.e. variability of environmental settings, community structure, and diversity between “Reference sites”) will be soon available in the Vonnahme et al.(in press) paper mentioned above. Primary aims of this study were: i) to compare the microbes of nodules fields with the microbiome of other deep-sea sediments in order to identify specific features of microbial communities of nodule fields; ii) to elucidate differences in diversity and in microbial community structure between sediments and nodules, and their potential implications for microbially-mediated functions. Thus, we believe that in order to achieve these aims it is neither needed nor beneficial to provide and discuss sedimentary metadata. However, as mentioned above, we point out in the revised MS that this information can be found in “Vonnahme et al.” in case the reader seeks a better understanding of the effect of environmental settings on microbial community structure in sediments of the sites investigated. [Lines 133-136]

RC2> 6. Page 8, lines 226 – 229. “…significant differences were detected in sediment microbial community structure … between communities associated with nodules and sediments at Reference South.” I think it is important to state directly in the text that this site, Reference South, was the only site that had enough nodule sampling to allow the authors to do this analysis (at least I assume this is what occurred). Otherwise this sentence could be taken to mean that differences in community structure between nodules and sediments were also investigated at the other 2 sites, and no differences were found.

AR> We thank the Reviewer for highlighting this point.

We clarified this issue in the revised MS as follows (added/replaced text in *italics*):
”Also, significant differences were detected in sediment microbial community structure among the different sites (PERMANOVA; Bacteria: R2 = 0.384; p = 0.003; F2,8 = 1.87; Archaea: R2 = 0.480; p = 0.013; F2,8 = 2.31; Table S1), and between communities associated with nodules and sediment at Reference South (PERMANOVA; Bacteria: R2 = 0.341; p = 0.023; F1,6 = 2.59; Archaea: R2 = 0.601; p = 0.029; F1,6 = 7.53; Table S1), *which was the only site where the number of samples allowed for the test*.“ [Lines 249-254]

RC2> Minor issues to be addressed:

RC2> 7. Page 8, line 238. “Aphaproteobacteria” should be “Alphaproteobacteria”.

AR> This has been corrected. [Line 262]

RC2> 8. Page 9, line 282. “Aphaproteobacteria” should be “Alphaproteobacteria”.

AR> This has been corrected. [Line 307]

~~Microbial communities associated with sediments and polymetallic nodules of the Peru Basin~~ The contribution of microbial communities in polymetallic nodules to the diversity of the deep-sea microbiome of Peru the Basin (4130 − 4198 meter depth)

Massimiliano Molari[1], Felix Janssen[1,2], Tobias R. Vonnahme[1,a], Frank Wenzhöfer[1,2] and Antje Boetius[1,2]

[1] Max Planck Institute for Marine Microbiology, Bremen, Germany
[2] HGF-MPG Joint Research Group on Deep Sea Ecology and Technology, Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, Germany

[a] Present address: UiT the Arctic University of Tromsø, Tromsø, Norway

Correspondence to: Massimiliano Molari (mamolari@mpi-bremen.de)

Abstract. Industrial-scale mining of deep-sea polymetallic nodules will ~~need~~ remove nodules in large areas of the seafloor. The regrowth of the nodules by metal precipitation is estimated to take millions of years. Thus, for future mining impact studies, it is crucial to understand the role of nodules in shaping microbial diversity and function in deep-sea environments. Here we investigated microbial community composition based on 16S rRNA gene sequences retrieved from sediments and nodules of the Peru Basin (4130 − 4198 m water depth). The nodule field of the Peru Basin showed a typical deep-sea microbiome, with dominance of the classes Gammaproteobacteria, Alphaproteobacteria, Deltaproteobacteria, and Acidimicrobiia. Nodules and sediments host distinct bacterial and archaeal communities, with nodules showing lower diversity and a higher proportion of sequences related to potential metal-cycling bacteria (i.e. Magnetospiraceae, Hyphomicrobiaceae), bacterial and archaeal nitrifiers (i.e. *AqS1*, unclassified Nitrosomonadaceae, *Nitrosopumilus*, *Nitrospina*, *Nitrospira*), and bacterial sequences found in ocean crust, nodules, hydrothermal deposits and sessile fauna. Sediment and nodule communities overall shared a low proportion of Operational Taxonomic Units (OTU; 21 % for Bacteria and 19 % for Archaea). Our results show that nodules represent a specific ecological niche (i.e. hard substrate, high metal concentrations, and sessile fauna), with a potentially relevant role in organic carbon degradation. Differences in nodule community composition (e.g. Mn-cycling bacteria, nitrifiers) between the Clarion-Clipperton Fracture Zone (CCZ) and the Peru Basin suggest that changes in environmental setting (e.g. sedimentation rates) play also a significant role in structuring the nodule microbiome.

1 Introduction

Polymetallic nodules (or manganese nodules) occur in abyssal plains (4000−6000 m water depth) and consist primarily of manganese and iron, as well as many other metals and rare earth elements (Crerar and Barnes, 1974; Kuhn et al. 2017). Nodules are potato- or cauliflower-shaped structures with typical diameters of 4−20 cm and are typically found at the sediment surface or occasionally buried in the uppermost 10 cm sediment horizon. The mechanisms of nodule formation are not completely elucidated. The current understanding is that they are formed via mineral precipitations from bottom waters (*Hydrogenetic* growth) or pore waters (*Diagenetic* growth) involving both abiotic and microbiological processes (Crerar and Barnes, 1974; Riemann, 1983; Halbach et al., 1988; Wang et al., 2009). The formation of nodules is a slow process that is estimated to range between thousands and millions of years per millimetre growth (Kerr, 1984; Boltenkov, 2012).

Rising global demand for metals has renewed interests in commercial mining of deep-sea nodule deposits. Mining operations would remove nodules, disturb or erode the top decimeters of sediment, and create near bottom sediment plumes that will resettle and cover the seafloor (Miller et al., 2018). Although the first nodules have been discovered in the 1870's (Murray, 1891), only little is known about the biodiversity, biological processes and ecological functions of the nodules and their surrounding sediments as specific deep-see habitat. Major questions remain, for example as to spatial turnover on local and global scales, the role of the microbial community in and around nodules, and the role of nodules as substrate for endemic species. Hence, there is the need to thoroughly characterize baseline conditions as a requirement for any mining operations as these will require assessments of impacts associated with mining.

Extensive and dense nodule fields are found in different areas of the Pacific and Indian Oceans ~~deep seas~~. Nodule accumulations of economic interest have been found in four geographical locations: the Clarion-Clipperton Fracture Zone (CCZ) and the Penrhyn Basin in the central north and south Pacific Ocean, respectively; the Peru Basin in the south-east Pacific; and in the center of the northern Indian Ocean (Miller et al., 2018). To our knowledge the Peru basin is the only region that does not have exploration activities and plans for mining so far. Previous work on the structure of microbial communities of nodule fields by 16S rRNA gene sequencing focused on the CCZ and the central South Pacific Ocean (Xu et al., 2007; Wu et al., 2013; Tully and Heidelberg, 2013; Blöthe et al., 2015; Shulse et al., 2016; Lindh et al., 2017). All studies showed that polymetallic nodules harbour ~~harbor~~ microorganisms that are distinct from the surrounding sediments and overlying water. They indicate that nodule communities show a pronounced spatial variability, but these results are so far not conclusive. Similar microbial communities were observed in nodules collected at distances of 6000 km and 30 km (Wu et al., 2013; Shulse et al. 2016), while Tully and Heidelberg (2013) found that nodule communities varied among sampling sites (<50 km). Besides, potential Mn-oxidizers and -reducers such as *Alteromonas*, *Pseudoalteromonas*, *Shewanella* and *Colwellia* were proposed as a core of the nodule microbiome involved in the formation of nodules (Wu et al. 2013; Blöthe et al., 2015), but they were not found in all nodules sampled so far (Tully and Heidelberg, 2013; Shulse et al. 2016). The lack of knowledge on the diversity and composition of microbial assemblages of other nodule provinces

makes it difficult to assess whether observed differences within the CCZ may reflect regional differences in environmental conditions (e.g. input of organic matter, bathymetry, topography, sediment type), or in abundance and morphology of nodules, or in the colonization of the nodules by epifauna and protozoans.

In this study we investigated the diversity and composition of bacterial and archaeal communities associated with manganese nodule fields of the Peru Basin. The Peru Basin is located about 3000 km off the coast of Peru and covers about half of the size of the CCZ, which is 5000–9000 km away. The present-day organic carbon flux in this area is approximately two times higher than in the CCZ, resulting in higher content of organic carbon in the surface sediments (>1 % *vs* 0.2–0.6 % in the CCZ), and a shallower oxic-suboxic front (10 cm *vs* tens of meters sediments depth in the CCZ; Müller et al., 1988; Heackel et al., 2001; Volz et al., 2018). As a consequence of differences in environmental conditions (e.g. organic carbon flux, carbonate compensation depth, sediment type, topography and near-bottom currents), the Peru basin and the CCZ host manganese nodules with different geological features (Kuhn et al. 2017). This includes: i) nodules from the Peru Basin are often larger, with a typical cauliflower shape, compared to those in CCZ that have a discoidal shape and a size of 2-8 cm (Kuhn et al. 2017); ii) average nodule abundance in the Peru Basin is lower (10 kg m^{-2}) than in CCZ (15 kg m^{-2}; Kuhn et al. 2017); iii) Mn nodules from the Peru Basin are thought to be mainly formed by suboxic diagenesis, whereas CCZ nodules apparently exhibit a mixture of diagenetic and hydrogenetic origin (von Stackelberg 1997; Chester and Jickells 2012); iv) while Peru Basin and CCZ nodules consist of the same type of mineral (disordered phyllomanganates), they have a different metal content (Wegorzewski and Kuhn 2014; Wegorzewski et al. 2015).

An increasing number of studies and policy discussions address the scientific basis of ecological monitoring in deep-sea mining, highlighting the need to identify appropriate indicators and standards for environmental impact assessments and ecological management. A key aspect is avoiding harmful effects to the marine environment, which will have to include loss of species and ecosystem functions. The primary aims of this study were to assess the structure and similarity of benthic microbial communities of nodules and sediments of the Peru Basin nodule province, and to compare them with those of other global deep-sea sediments and nodules in the CCZ. The focus was on similarity comparisons in order to investigate endemism and potential functional taxa that could be lost due to the removal of manganese nodules by mining activities. To achieve this, the hypervariable 16S rRNA gene regions V3-V4 for Bacteria and, V3-V5 for Archaea were amplified from DNA extracted from nodules and surrounding sediments and sequenced using the Illumina paired-end MiSeq platform. The hypotheses tested were i) nodules shape deep-sea microbial diversity ~~and functions~~, ii) nodules host a specific microbial community compared to the surrounding sediments. Secondary aim of this study was to investigate the nodule features that may play a major role in shaping microbial community composition and microbially-mediated functions. ~~and iii) environmental setting and nodule features impact microbial community composition.~~

2 Methods

2.1 Sample collection

Sediment samples and polymetallic nodules were collected as a part of the MiningImpact project of the Joint Programming Initiative JPI Healthy and Productive Seas and Oceans (JPI Ocean) on board of R/V Sonne (expedition SO242/2; 28th of August - 1st of October 2015) in the Peru Basin around 7° S and 88.5° W. Samples were collected at three sites outside the seafloor area selected in 1989 for a long-term disturbance and recolonization experiment (DISCOL; Thiel et al., 2001), for this reason they were called "Reference Sites": Reference East, Reference West, and Reference South. Sediment samples were collected using TV-guided MUltiple Corer (TV-MUC) at three stations per site (Table 1). The cores were sliced on board in a temperature-controlled room (set at *in situ* temperature), and aliquots of sediment were stored at −20 °C for DNA extraction ~~and prepared for cell counts (see sections below)~~. Manganese nodules where sampled, using a TV-MUC, or a Remotely Operated Vehicle (~~ROV Kiel6000~~ ROV KIEL6000, GEOMAR, Germany): one nodule at Reference West and four nodules at Reference South. The nodules were ~~where~~ partly located at the surface or buried down to 3 cm below the seafloor (bsf) with diameters of a few cm. Nodules were gently rinsed with 0.22-µm filtered cold bottom seawater to remove adhering sediment, stored in sterile plastic bags at −20 °C and crushed before DNA extraction in the home lab. From the nodules collected with the ROV, only the surface layer was scraped off using a sterile spoon, and subsequently crushed and frozen (−20 °C). Sedimentary metadata (e.g. cell counts, pigments and organic carbon content, porewater profiles, and porosity) and a map of the study area are available in Vonnahme at al. (in press). Focusing entirely on sediments, that publication also includes a discussion of the variability of environmental settings and microbial communities.

2.2 DNA extraction and sequencing

~~The nodules collected with the MUC where crushed and stored at −20 °C. From the nodules collected with the ROV, only the surface layer was scraped off using a sterile spoon, and subsequently crushed and frozen (−20 °C).~~ The DNA was extracted from 1 g of wet sediment (0-1 cm layer) and from 1 g of wet nodule's fragments using the FastDNA™ SPIN Kit for Soil (Q-BIOgene, Heidelberg, Germany) following the ~~manual~~ protocol provided by the manufacturer. An isopropanol precipitation was performed on the extracted DNA, and DNA samples were stored at −20 °C. As control for DNA contamination (negative control), DNA extraction was carried out on purified water after being in contact with sterile scalpel and plastic bag.

Amplicon sequencing was done at the CeBiTec laboratory (Centrum für Biotechnologie, Universität Bielefeld) on an Illumina MiSeq machine. For the 16S rRNA gene amplicon library preparation we used the bacterial primers 341F (5′-CCTACGGGNGGCWGCAG-3′) and 785R (5′-GACTACHVGGGTATC TAATCC-3′), and the archaeal primers Arch349F (5′-GYGCASCAGKCGMGAAW-3′) and Arch915R (5′-GTGCTCCCCCGCCAATTCCT-3′) (Wang and Qian, 2009; Klindworth et al., 2013), which amplify the 16S rRNA gene hypervariable region V3-V4 in Bacteria (400−425 bp fragment length) and the V3-V5 region in Archaea (510 bp fragment length). The amplicon library was sequenced with the MiSeq v3 chemistry, in a 2x300 bp paired run with

>50,000 reads per sample, following the standard instructions of the 16S rRNA gene Metagenomic Sequencing Library Preparation protocol (Illumina, Inc., San Diego, CA, USA).

The quality cleaning of the sequences was performed with several software tools. CUTADAPT (Martin, 2011) was used for primer clipping. 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). Low-quality archaeal sequences were removed after merging the reads in order to enhance the number of retained reads due to the increase in archaeal 16S rRNA gene fragment length. 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). Clustering of sequences into OTUs (operational taxonomic units) was done using the SWARM algorithm (Mahé et al., 2014). The taxonomic classification was based on the SILVA rRNA reference database (release 132), at a minimum alignment similarity of 0.9, and a last common ancestor consensus of 0.7 (Pruesse et al., 2012). Workflow and scripts applied in this study can be found in Hassenrück et al. (2016). Raw sequences with removed primer sequences were deposited at the European Nucleotide Archive (ENA) under accession number PRJEB30517 and PRJEB32680; the sequences were archived using the service of the German Federation for Biological Data (GFBio; Diepenbroek et al., 2014).

The total number of sequences obtained in this study is reported in Table S1 ~~table 2~~. Absolute singletons (SSOabs), i.e. OTUs consisting of sequences occurring only once in the full dataset (Gobet et al., 2013) were removed (Table S1 ~~Table 2~~). Similarly, contaminant sequences (as observed in the negative control) and unspecific sequences (i.e., bacterial sequences in the archaeal amplicon dataset, and archaeal, chloroplast, and mitochondrial sequences in the bacterial dataset) were removed from amplicon data sets before the analysis (Table S1 ~~Table 2~~). The dominant OTU sequences and OTU sequences highly abundant in the nodules were subjected to BLAST search (BLASTn; Gene Bank nucleotide database 12/06/2019) in order to identify in which others habitats the closest related (i.e. >99 %) sequences have been previously reported.

2.3 Data analysis

The first three Hill Numbers, or the effective number of species, were used to describe alpha-diversity: species richness (H_0), the exponential of Shannon entropy (H_1), and the inverse Simpson index (H_2; Chao et al., 2014). Coverage-based and sample-size-based rarefaction (based on actual number of sequences) and extrapolation (based on double number of sequences) curves were calculated for the Hill's numbers using the R package iNEXT (Hsieh et al., 2018). Calculation of the estimated richness (Chao1) and the identification of unique OTUs (present exclusively in one sample) were based on repeated (n = 100) random subsampling of the amplicon data sets. Significant differences in alpha-diversity indices between substrates (i.e. manganese nodules and sediments) were determined by

analysis of variance (ANOVA), or by non-parametric Kruskal-Wallis test (KW) when ANOVA's assumptions were not satisfied.

Beta-diversity in samples from different substrates and 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. The latter was calculated with 100 sequence re-samplings per sample on the smallest dataset (40613 sequences for Bacteria and 1835 sequences for Archaea). Euclidean distance ~~Bray-Curtis dissimilarity~~ was used to produce non-metric multidimensional scaling (NMDS) plots. 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. The permutational multivariate analysis of variance (PERMANOVA; Anderson, 2001) was used to test difference in community structure and composition.

Differentially abundant OTUs and genera were detected using the R package ALDEx2 (Fernandes et al. 2014) at a significance threshold of 0.01 and 0.05 for Benjamini–Hochberg (BH) adjusted parametric and non-parametric (KW) P-values, respectively. We only ~~discussed~~ discuss the taxa that were at least two times more abundant in nodules than in sediments (i.e. (Log2(Nodule/sediment) ≥1) and with a sequences contribution of total number of sequences ≥1 % (for genera) or ≥0.1 % (for OTUs).

All statistical analyses were conducted in R using the core distribution with the additional packages vegan (Oksanen et al. 2015), compositions (Van den Boogaart et al., 2014), iNEXT (Hsieh et al., 2018), and ALDEx2 (Fernandes et al. 2014).

3 Results

3.1 Microbial alpha-diversity

Bacterial and archaeal communities in 5 nodules and 9 sediment samples (Table 1) were investigated using specific sets of primers for Bacteria and Archaea on the same extracted pool of DNA per station. The number of bacterial sequences retrieved from DNA extracted from sediments and nodules was on average 5±5 and 25±14 times higher, respectively, than those obtained for archaea (t-test: $p<0.001$, df=11, t=4.5).

Table S1 shows the statistics of sequence abundance and proportion of singletons and cosmopolitan types. Sequence abundances of bacteria were comparable between sediments and nodules. Cosmopolitan OUT, i.e. those present in 80 % of the sediments and nodule samples, were only 9 % of all taxa (77 % of all sequences), whereas rare OTU occurring only in <20 % of all samples represented 50 % the taxa (4 % of all sequences). Sediments *vs* nodules contained only 4 and 2 %, respectively, of endemic taxa, defined as those were abundant in either substrate but rare in the other. Thus the contribution of unique OTUs to the total number of OTUs was lower in manganese nodules than in sediments samples (Table 2, Figure 1a). Bacterial and archaeal diversity was investigated calculating the total number of OTUs (Hill number q=0; H_0) and the estimated richness (Chao1), and the unique

OTUs (present exclusively in one station). For this analysis, the latter were calculated with sequence re-sampling, to overcome differences in sequencing depth. Abundance-based coverage estimators, exponential Shannon (Hill number q=1; H_1) and inverse Simpson (Hill number q=2; H_2), were also calculated. The rarefaction curve indicates that the richness (H_0) of the less abundant and rare OTUs was somewhat underestimated both in nodules and in sediments (Figure S1 a-b). However, the bacterial and archaeal diversity was well described for the abundant OTUs (H_1 and H_2; Figure S1 a-b); with more than 90 % of the estimated diversity covered (Figure S1 c-d). Both in sediments and nodules the alpha-diversity indices were higher for Bacteria than for Archaea (t-test: $p<0.0001$, df=12, t=8.0−16.0), while the contribution of unique OTUs to the total number of OTUs was comparable (Table 2). Bacterial communities in manganese nodules have lower Hill numbers and Chao1 indices compared to those associated to sediments (Table 2, Figure 1a). Archaeal communities showed the same patterns for diversity indices and unique OTUs, with exception for the H_2 index that did not show significant difference between nodules and sediments (Table 2, Figure 1b).

3.2 Patterns in microbial community composition

The changes in microbial community structure at OTU level (beta-diversity) between substrates and samples were quantified by calculating ~~Bray-Curtis dissimilarities~~ Euclidean distances from CLR transformed OTU abundance. Shared OTUs were estimated by calculating Jaccard dissimilarity from OTU presence/absence based on repeated random subsampling of the amplicon data sets. Microbial communities associated with manganese nodules differed significantly from those found in the sediments (Figure 2, Table S2). Also, significant differences were detected in sediment microbial community structure among the different sites (PERMANOVA; Bacteria: $R^2 = 0.384$; $p = 0.003$; $F_{2,8} = 1.87$; Archaea: $R^2 = 0.480$; $p = 0.013$; $F_{2,8} = 2.31$; Table S2), and between communities associated with nodules and sediment at Reference South site (PERMANOVA; Bacteria: $R^2 = 0.341$; $p = 0.023$; $F_{1,6} = 2.59$; Archaea: $R^2 = 0.601$; $p = 0.029$; $F_{1,6} = 7.53$; Table S2), which was the only site where the number of samples allowed for the test. “Site” defined by geographic location, and “Substrate”, i.e. origin from sediment or nodule, explained a similar proportion of variation in bacterial community structure (27 % and 23 %, respectively). “Substrate” had a more important role in shaping archaeal communities than “Site” (explained variance 35 % and 19 %, respectively; Table S2). The number of shared OTUs between nodules and sediments (Bacteria: 21 %; Archaea: 19 %) was lower than those shared within nodules (Bacteria: 30 %; Archaea: 30 %) and within sediments (Bacteria: 31 %; Archaea: 32 %) (Figure S2).

Bacterial communities in manganese nodules and sediments were dominated by the classes Gammaproteobacteria (26 %), Alphaproteobacteria (19 %), Deltaproteobacteria (9 %), Bacteroidia (5 %), Acidimicrobiia (4 %), Dehalococcoidia (4 %), Planctomycetacia (4 %), Nitrospinia (3 %), and Phycisphaerae (3 %), which accounted for more than 75 % of the total sequences (Figure 3). All archaeal communities were dominated by Thaumarchaeota (*Nitrosopumilales*), which represented more than 95% of all sequences. The remaining small proportion of sequences was taxonomically assigned to *Woesearchaeia* (Figure S2b). Nodule and sediment samples showed similar compositions of most abundant bacterial genera (contribution to total number of sequence ≥1 %; Figure S2a). 69 bacterial

genera (9 % of all genera) were differentially abundant in the nodules and in the sediment, accounting for 36 % and 21 % of total sequences retrieved from nodules and sediments, respectively (ALDEx2: ANOVA adjusted $p<0.01$ and KW adjusted $p<0.05$; Figure 4 and Table S3). Of those only one unclassified genus within the family of Sphyngomonadaceae and the genus *Filomicrobium* were exclusively found in nodules and not in the sediment samples, and their contribution to the total number of sequences was less than 0.06 %. Genera that were more abundant in the nodules than in the sediments included: unclassified Alphaproteobacteria (7 %), *Nitrospina* (4 %), unclassified SAR324 clade (Marine group B; 3 %), unclassified Hyphomicrobiaceae (3 %), Pirellulaceae Pir4 lineage (2 %), unclassified Methyloligellaceae (1 %), unclassified Pirellulaceae (1 %), *Acidobacteria*, unclassified Subgroup 9 (1 %) and Subgroup 17 (1 %), Nitrosococcaceae *AqS1* (1 %), Calditrichaceae *JdFR-76* (1 %), and *Cohaesibacter* (1 %) (Figure 4 and Table S3). In the sediment we identified 21 genera that were more abundant than in the nodules, but all together they represented only 3 % of total sequences recovered from sediments. 128 OTUs were highly abundant in nodules (ALDEx2: ANOVA adjusted $p<0.01$ and KW adjusted $p<0.05$), which accounts for 24 % of total sequences retrieved from nodules (Table 3a). The closest related sequences (≥99 % similarity) were retrieved from ocean crusts (30 %), from nodule fields (26 %), from hydrothermal/seep sediments and deposits (21 %), from worldwide deep-sea sediments (16 %), and associated to invertebrates (7 %; Table 3b and Figure 5).

4 Discussion

Industrial-scale mining of deep-sea polymetallic nodules may remove nodules and the active surface seafloor layer at a spatial scale ranging from ca. ~~50,000–75,000~~ 50000−75000 km^2 per claim to ca. 1 million km^2 including all current exploration licences (Miller et al., 2018). The regrowth of nodules will take millions of years, thus it is unknown if the associated biota could recover at all (Simon-Lledo et al., 2019). The response of microbial communities to the loss of nodules and seafloor integrity is largely unknown. It may play an important role in the ecological state of the seafloor habitat due to the many functions bacteria and archaea hold in the food-web, element recycling, and biotic interactions, beyond representing the largest biomass in deep-sea sediments (Joergensen and Boetius 2007). It is thus crucial to understand the role of nodules in shaping microbial diversity and in hosting microbes with important ecological functions. So far, only few studies were carried out to investigate the microbiota of nodule fields, and most of them were focused on identifying microbes involved in metal cycling. Here, we investigated similarity of microbial community structures in sediments and nodules retrieved from the Peru Basin. The objectives of this study were: i) compares the microbes of nodules fields with microbiota of deep-sea sediments, in order to identify specific features of microbial diversity of nodule fields; ii) elucidates differences in diversity and in microbial community structure between sediments and nodules; ~~and their relapses on potential microbially-mediated functions;~~ iii) ~~understand~~ investigates potential microbially-mediated functions and the major drivers in shaping microbial communities associated to the nodules.

4.1 Microbial diversity of nodule fields is distinct from other deep-sea areas

306 Benthic bacterial assemblages in sediments and nodules of the Peru basin showed the typical
307 dominance of the classes Gammaproteobacteria, Alphaproteobacteria, Deltaproteobacteria, and
308 Acidimicrobiia, as reported for ~~worldwide~~ deep-sea sediments worldwide (Bienhold et al., 2016;
309 Figure 3) and in the Pacific Nodule Province (Wang et al., 2010; Wu et al., 2013; Shulse et al., 2016;
310 Lindh et al., 2017). ~~But~~ However at higher taxonomic resolution we detected substantial differences to
311 the microbial community composition of other deep-sea regions. Sediments of the Peru Basin bacteria
312 classes were depleted in sequence abundances of Flavobacteria, Gemmetimonadetes and Bacilli,
313 whereas sequence abundances of the Chloroflexi (i.e. Dehalococcoidia), Planctomycetes (i.e.
314 Pirellulaceceae, Phycisphaeraceae) and the genus *Nitrospina* were higher compared to other deep-sea
315 regions (Bienhold et al., 2016, Varliero et al., 2019). Dehalococcoidia and Planctomycetes were
316 previously reported as important component of benthic microbial assemblages in the Pacific Ocean
317 (Wang et al., 2010; Wu et al., 2013; Blöthe et al., 2015; Walsh et al., 2016; Lindh et al., 2017). Their
318 contribution to the total community was found to increase in organic matter depleted subsurface
319 sediments (Durbin and Teske, 2011; Walsh et al., 2016).
320
321 Dominant OTUs (>1 %) belonged to unclassified Actinomarinales, Gammaproteobacteria, Subgroup
322 21 (phylum Acidobacteria), and to genus *Woesia* (family Woeseiaceae). Members of Actinomarinales
323 and Woeseiaceae are cosmopolitan types in deep-sea sediments (Bienhold et al., 2016). For
324 Actinomarinales there are no cultivates, and the function of this group remains unknown. In the case of
325 Woeseiaceae, one representative is in culture (*Woeseia ocaeni*). *W. ocaeni* is an obligate
326 chemoorganoheterotroph (Du et al., 2016), suggesting a role in organic carbon remineralization for
327 members of that family, as recently confirmed by analysis of deep-sea assembled genomes (Hoffmann
328 et al., 2020 ~~in revision~~). Closest related sequences of Subgroup 21 have been reported in deep-sea
329 sediments (Schauer et al. 2010) and across Pacific nodule fields (Wu et al., 2013), but also in
330 association with deep-sea benthic giant foraminifera (Xenophyophores) and in surrounding sediments
331 (Hori et al., 2013). The subgroup 21-like OTU was also one of the 10 most abundant OTUs retrieved
332 from nodules (0.9 %). Xenophyophores have agglutinated tests and can grow to decimetre size,
333 suggesting that members of Subgroup 21 may be colonists of biological and/or hard substrates.
334
335 Within the class Alphaproteobacteria the most abundant OTUs (>0.5 %) belonged to unclassified
336 genera of the families Magnetospiraceae (order Rhodospirillales), Hyphomicrobiaceae (order
337 Rhizobiales), and Kiloniellaceae (order Rhodovibrionales). Magnetospiraceae and Hyphomicrobiaceae
338 are the most abundant families in nodules with >2 % of OTUs. Closely related sequences have been
339 reported previously across Pacific Nodule Provinces (Xu et al., 2007; Shulse et al., 2016). The family
340 of Magnetospiraceae includes microaerophilic heterotrophs, able of magnetotaxis and iron reduction
341 (i.e. genus Magnetospirillum; Matsunaga et al. 1991; Schuler and Frankel, 1999), and thus the
342 members of this family could play a role in Fe(III) mobilization, affecting its bioavailability.
343 Hyphomicrobiaceae-like sequences found in this study are related to genera *Hyphomicrobium* and
344 *Pedomicrobium* (sequence identity 97 %), which have been reported to be involved in manganese
345 cycling (Tyler, 1970; Larsen et al., 1999; Stein et al., 2001). A potential contribution of these groups in

346 metal cycling in manganese nodules is also suggested by the presence of closest related sequences in
347 ocean crust (Santelli et al., 2008; Lee et al., 2015), which typically hosts epilithic and endolithic
348 microbial communities of chemolithotrophic metals-oxidizers (Staudigel et al., 2008). Similarly,
349 Kiloniellaceae related OTUs might be involved in metal-cycling as closely related sequences were
350 found in marine basalts (Mason et al., 2007; Santelli et al., 2008) and inside other manganese nodules
351 (Blöthe et al., 2015). Most of the marine cultivates in the family Kiloniellaceae belong to genus
352 *Kiloniella*, that have been isolated from marine macroalga (Wiese et al., 2009), the guts of Pacific
353 white shrimp (Wang et al., 2015), marine sponge (Yang et al., 2015), spider crab and clam (Gerpe et
354 al., 2017), and from the surface water of a polynia in the Western Antarctic Sea (Si et al., 2017).
355 Besides, Kiloniellaceae-like sequences were found in sponges (Cleary et al., 2013), sea star ~~start~~ larvae
356 (Galac et al., 2016) and in seamount's iron mats (Scott et al., 2017). The presence of rich sessile and
357 mobile metazoan communities associated to nodules offers various potential hosts for members of
358 Kiloniellaceae. *Kiloniella* is a chemoheterotrophic aerobe, and the draft genome of an isolate from the
359 gut of a Pacific white shrimp shows potential for denitrification and iron acquisition and metabolism
360 (Wang et al., 2015). Thus, either as free-living or host-associated life, the potential contribution of
361 Kiloniellaceae in metal cycling requires further investigation.
362
363 Archaea were also present in sediments of the Peru Basin, with Nitrosopumilaceae (phylum
364 Thaumarchaeota) dominating the archaeal communities (Figure S2b). ~~In contrast to what was reported~~
365 ~~for CCZ (Tully and Heidelberg, 2013; Shulse et al., 2016),~~ Archaeal sequences comprised a lower
366 portion of total sequences retrieved from sediments (6 – 45 %) and nodules (<1 – 7 %) of Peru basin
367 ~~(ca. 10 %)~~, and they were lower in nodules compared to the sediments. Our data differed from what
368 was reported by Shulse et al. (2016) for CCZ, especially for nodules (ca. 20%). We cannot rule out that
369 the observed differences in microbial community structure partly reflect the different sets of primers
370 used in our study and by Shulse et al. (2016). As both primer sets amplified the same hypervariable
371 region of 16S rRNA gene (V4) we assume that biases are small enough to justify the comparison. The
372 majority of member of Nitrosopumilaceae are believed to be capable of oxidation of ammonia to
373 nitrite, the first step of nitrification (Offre et al., 2013). Archaeal ammonia oxidizers have a higher
374 affinity for ammonia than bacterial ammonia oxidizers, and they are favoured in environments with
375 low ammonia concentrations (Martens-Habbena et al., 2009). The Peru Basin has higher particulate
376 organic-carbon fluxes as compared to central Pacific Ocean (Haeckel et al., 2001; Mewes et al., 2014),
377 which results in higher remineralisation rates and higher ammonia fluxes. These limit the thickness of
378 oxygenated sediments to 10 cm in the Peru Basin while they can reach up to 2-3 m depth in the CCZ
379 (Haeckel et al., 2001; Mewes et al., 2014; Volz et al., 2018). Hence differences observed between CCZ
380 and Peru nodule fields in the contribution of archaeal sequences to microbial assemblages are likely
381 due to ammonia availability, which is controlled by organic matter fluxes.

382 **4.2 Microbial community structure differs between sediments and nodules**

383 Beta-diversity of microbial community structure in the Peru Basin sediments showed remarkable OTU
384 turnover already on a local scale (<60 km; Figure S2), which is at the higher end for turnover rates

from previous microbial beta-diversity estimates for bathyal and abyssal seafloor assemblages (Jacob et al., 2013; Ruff et al., 2015; Bienhold et al., 2016; Walsh et al., 2016; Varliero et al., 2019). Here we focused specifically on the contribution of nodules to diversity, which could be a critical parameter in the ecological assessment of nodule removal. 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). Albeit the microbial communities in the sediment showed significant differences between sites, the low number of shared OTUs between sediments and nodules <20 % supports the presence of specific bacterial and archaeal communities associated with polymetallic nodule habitats ~~(Wu et al., 2013; Tully and Heidelberg, 2013; Shulse et al., 2016; Lindh et al. 2017)~~. However, the proportion of truly endemic, unique nodule OTUs was also low (~~Table 3,~~ Figure 1a, Table S1), nonetheless it is relevant to highlight that nodule removal would lead to a loss of specific types of microbes in a mined deep-sea region (Blöthe et al., 2015).

Microbial communities associated ~~to~~ with nodules are ~~generally~~ 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 Figure 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 in microbial biodiversity between nodules and sediments within the studied habitats in the CCZ (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 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. Tully and Heidelberg (2013) suggested that lower microbial diversity in the nodules might be due to less availability of potential energy sources (e.g. organic matter) compared to sediments. Despite that the sedimentation rate exceeds the growth rate of nodules, the nodules are typically exposed to bottom water and not covered by sediments (Peukert et al., 2018). Although, it is unknown whether physical mechanisms (e.g. current regime or seasonal events) or biological processes (e.g. grazing, active cleaning) are responsible for lack of sediments accumulation on nodules, the decrease of microbial diversity with the decrease of organic matter availability is in accordance with positive energy-diversity relationship reported for deep-sea sediments (Bienhold et al., 2012). However, the presence of foraminiferal assemblages (Gooday et al., 2015) and specific sessile metazoan communities (Vanreusel et al., 2016) on the surface of nodules may represent a potential source of transformed organic matter (e.g. dissolved organic matter) and catabolic products, which may represent a much more valuable energy source for microbes than refractory particulate organic matter sinking from the water column. Furthermore, higher microbial diversity in the sediments than in the nodules could be the result of the accumulation of allochthonous microbes, as suggested by the higher proportion of rare and unique OTUs in the sediments. Lastly, the nodules offer hard substrate and presence of metals, which can select for specific Bacteria and Archaea. Similarly, hydrothermal deposits have typically lower bacterial diversity than deep-sea sediments despite chemical energy sources being highly available (Ruff et al., 2015; Wang et al., 2018). We propose that the decreased diversity of abundant OTUs in

nodules, observed especially for Bacteria, suggests selection for colonists adapted to specific ecological niches associated with nodules (e.g. high metals concentration, hard substrate, presence of sessile fauna).

4.3 Potential functions of microbial communities associated to nodules

The presence of a large proportion of bacterial community with low abundance in the sediments, but enriched ~~by~~ in the nodules ~~nodule environment~~ both at the level of genera (35 %) and OTUs (24 %) (Figure 4, Table S3 and Table 3a) indicates niche specialization. The most abundant OTUs (13 % of the bacterial community) in nodules include unclassified Hyphomicrobiaceae, Magnetospiraceae, Alphaproteobacteria, Arenicellaceae and SAR324, *Nitrospina*, *AqS1*, Methyloligellaceae, Subgroup 9, Subgroup 17, Kiloniellaceae, *Cohaesibacter* and JdFR-76, which closest related sequences have been retrieved from Pacific nodules (e.g. Wu et al. 2013; Blöthe et al., 2015), basaltic rocks (e.g. Mason et al 2007; Santelli et al 2008; Mason et al., 2009; Lee et al., 2015), sulfide and carbonate hydrothermal deposits (e.g. Sylvan et al., 2012; Kato et al., 2015), and giant foraminifera (Hori et al., 2013; Table 3b and Figure 5). There are currently no cultivated representatives and metabolic information for these members of the Bacteria, and it is not known whether they have metal tolerance mechanisms or they are actively involved in metal cycling. The high abundance of potential metal ~~reducers~~ reductive (i.e. Magnetospiraceae) and oxidizers (i.e. Hyphomicrobiaceae), and presence of encrusting protozoans (Gooday et al., 2015), microbial eukaryotes (Shulse et al., 2016) and metazoans (Vanreusel et al., 2016) create specific ecological niches, which may be at least partially responsible for the observed selection of microbial taxa in nodules. Overall, these findings suggest that bacterial groups adapted to lithic or biological substrates preferentially colonize nodules, likely favoured by manganese and iron availability, formation of biofilms and presence of sessile fauna communities.

The reduction and dissolution of Mn oxides by dissolved organic matter (e.g. humic compounds) occurs typically in photic or reducing aquatic environments (Sunda et al., 1983; Stone and Morgan, 1984; Stone, 1987; Sunda and Huntsman, 1994). However reductive ~~reductively~~ dissolution of Mn oxides by dissolved organic substrates has been observed also in dark oxygenated seawater (Sunda et al., 1983; Sunda and Huntsman, 1994), suggesting that it could be a relevant abiotic process in manganese nodules. Indeed, this reaction yields manganese(II) and low-molecular-weight organic compounds (Sunda and Kieber, 1994), which potentially may favour Mn-oxidizing Bacteria and microbial exploitation of refractory dissolved organic matter. Intense extracellular enzymatic activities have been reported for seafloor-exposed basalts (Meyers et al., 2014), raising the question of whether the ~~close~~ closely related microbes associated ~~to~~ with nodules might have comparable degradation rates. Furthermore, nodules host diversified communities of suspension feeders such as serpulid tubeworms, sponges, corals and crinoids (Vanreusel et al., 2016), which filter microbes and POC from the bottom water and release DOM and catabolic metabolites (e.g. ammonia). Thus, nodules may act as hot spots of organic carbon degradation. Albeit metabolic activity has never been quantified on nodules and sequence abundances are lower, the increased abundance of nitrifiers in nodules compared to the sediments reported for Pacific Nodule Province (Tully and Heidelberg, 2013; Shulse et al., 2016) and in this study could indicate a high metabolic activity. Nitrifiers catalyse the oxidation of ammonia, a

catabolic product of heterotrophic metabolism, to nitrite and eventually to nitrate. In the CCZ the nitrifier community was composed of archaeal ammonia-oxidizing *Nitrosopumilus*, which represented a large portion of the microbial assemblages (up to 20 %), and a minor contribution of bacterial nitrite-oxidizing *Nitrospira* (Tully and Heidelberg, 2013; Shulse et al., 2016). Peru Basin sediments and nodules showed more diversified nitrifier communities, which are enriched by ammonia oxidizing *AqS1* (1 %) and unclassified Nitrosomonadaceae (1 %) and by nitrite-oxidizing *Nitrospina* (4 %) and *Nitrospira* (1 %; Figure 4, Table S3 and Table 3a). *Nitrospina* are not commonly reported for deep-sea sediments, but they are the dominant nitrite oxidizers in the oceans (Luecker et al., 2013). They have recently been reported as symbiont of deep-sea glass sponges (Tian et al., 2016), which also commonly colonize FeMn nodules (Vanreusel et al., 2016). The Nitrospina-related OTUs detected in the nodules showed only low similarity with pelagic *Nitrospina gracilis* and *Nitrospina*-like sequences found in deep-sea glass sponge (sequence identity of 93 %), but were closely related with sequences recovered from marine basalts (Mason et al., 2007; Santelli et al. 2008; Mason et al. 2009), suggesting nodules as a native habitat.

5 Conclusions

The sediments of nodule fields in the Peru Basin host a specific microbial community ~~composition~~ of bacterial taxa reported for organic carbon poor environments (i.e. Chloroflexi, Planctomycetes) and potentially involved in metal-cycling (i.e. Magnetospiraceae, Hyphomicrobiaceae). Nodule communities were distinct from sediments and showed a higher proportion of sequences from potential Mn-cycling bacteria including bacterial taxa found in ocean crust, nodules and hydrothermal deposits. Our results are in general agreement with previous studies in the CCZ, confirming that nodules provide a specific ecological niche. However remarkable differences in microbial community composition (e.g. Mn-cycling bacteria, nitrifiers) between the CCZ and the Peru Basin also show that environmental settings (e.g. POC flux) and features of FeMn nodules (e.g. metal content, nodule attached fauna) 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 should address regional differences, to determine the spatial turnover and its environmental drivers, and the consequences regarding endemic types.

Furthermore, our results suggest that the removal of nodules, and potentially also the blanketing of nodules with plume sediments ~~suspended~~ resuspended during the mining operations may affect the cycling of metal and other elements. Future work is needed to characterize metabolic activities on and in nodules, and to understand factors and processes controlling nodule colonization. Specifically, restoration experiments should take place to test whether artificial substrates favour the recovery of microbial and fauna communities, and their related ecological functions.

Data availability

Raw sequences with removed primer sequences were deposited at the European Nucleotide Archive (ENA) under accession number PRJEB30517 and PRJEB32680.

Author contributions

A.B, **F.J**, **F.W.** and **M.M** conceived the study. **A.B**, **F.J**, **F.W.** and **T.R.W.** performed sampling activities. **M.M.** compiled and analysed the data. **M.M.** wrote the paper with the contribution from all Authors.

Competing interest

The authors declare that they have no conflict of interest.

Special issue statement

Acknowledgements

The Authors want to thank captain and crew of the SO242/1 and /2 expedition, and M. Alisch, J. Bäger, J. Barz, A. Nordhausen, F. Schramm, R. Stiens and W. Stiens, (HGF-MPG Joint Research Group on Deep Sea Ecology and Technology) for technical support. The Authors are also grateful to Dr. H. Tegetmeyer for support with Illumina sequencing (CeBiTec laboratory; HGF-MPG Joint Research Group on Deep Sea Ecology and Technology). Lastly, the Authors would like to thank the Dr. Denise Akob, Dr. Beth N. Orcutt, Dr. Tim D'Angelo and an anonymous reviewer, which comments and suggestions great helped in improving the quality of the manuscript.

This work was funded by the German Ministry of Research and Education (BMBF grant no. 03F0707A-G) as part of the MiningImpact project of the Joint Programming Initiative of Healthy and Productive Seas and Oceans (JPIOceans). We acknowledge further financial support from the Helmholtz Association (Alfred Wegener Institute Helmholtz Center for Polar and Marine Research, Bremerhaven) and the Max Planck Society (MPG), as well as from the ERC Advanced Investigator Grant ABYSS (294757) to AB for technology and sequencing. The research has also received funding from the European Union Seventh Framework Program (FP7/2007- 2013) under the MIDAS project, grant agreement number 603418.

References

Anderson, M. J.: A new method for non-parametric multivariate analysis of variance, Austral Ecol., 26(1), 32–46, 2001.

Andrews, S.: FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (2010).

Bienhold, C., Boetius, A. and Ramette, A.: The energy – diversity relationship of complex bacterial communities in Arctic deep-sea sediments, ISME J., 6, 724–732, doi:10.1038/ismej.2011.140, 2012.

Bienhold, C., Zinger, L., Boetius, A. and Ramette, A.: Diversity and Biogeography of Bathyal and Abyssal Seafloor Bacteria, PLoS One, 11(1), 1–20, doi:10.1371/journal.pone.0148016, 2016.

536 Blöthe, M., Wegorzewski, A., Müller, C., Simon, F., Kuhn, T. and Schippers, A.: Manganese-Cycling
537 Microbial Communities Inside Deep-Sea Manganese Nodules, Environ. Sci. Technol., 49(13),
538 7692–7700, doi:10.1021/es504930v, 2015.
539 Bolger, A. M., Lohse, M. and Usadel, B.: Trimmomatic: a flexible trimmer for Illumina sequence data,
540 Bioinformatics, 30(15), 2114–2120, doi:10.1093/bioinformatics/btu170, 2014.
541 Boltenkov, B. S.: Mechanisms of formation of deep-sea ferromanganese nodules: Mathematical
542 modeling and experimental results, Geochemistry Int., 50(2), 125–132,
543 doi:10.1134/S0016702911120044, 2012.
544 Chao, A., J. Gotelli, N., Hsieh, T. C., L. Sander, E., H. Ma, K., Colwell, R. and M. Ellison, A.:
545 Rarefaction and extrapolation with Hill numbers: A framework for sampling and estimation in
546 species diversity studies, Ecol. Monogr., 84, 45–67, doi:10.1890/13-0133.1, 2014.
547 Chester, R. and Jickells, T.: Marine Geochemistry, 3rd ed Wiley-Blackwell, Oxford., 2012.
548 Church, M. J., Wear, E. K., Orcutt, B. N., Young, C. R. and Smith, J. M.: Taxonomic diversity of
549 Bacteria and Archaea in the Clarion-Clipperton Zone of the North Pacific Ocean. *Annex V* in Deep
550 CCZ Biodiversity Synthesis Workshop, Friday Harbor, Washington, USA, 1-4 October 2019.
551 Cleary, D. F. R., Becking, L. E., Voogd, N. J. De, Pires, A. C. C., Ana, R. M. P., Egas, C. and Gomes,
552 N. C. M.: Habitat- and host-related variation in sponge bacterial symbiont communities in
553 Indonesian waters, FEMS Microbiol. Ecol., 85, 465–482, doi:10.1111/1574-6941.12135, 2013.
554 Crerar, D. A. and Barnes, H. L.: Deposition of deep-sea manganese nodules, Geochim. Cosmochim.
555 Acta, 38(2), 279–300, doi:10.1016/0016-7037(74)90111-2, 1974.
556 Diepenbroek, M., Glöckner, F. O., Grobe, P., Güntsch, A., Huber, R., König-Ries, B., Kostadinov, I.,
557 Nieschulze, J., Seeger, B., Tolksdorf, R. and Triebel, D.: Towards an Integrated Biodiversity and
558 Ecological Research Data Management and Archiving Platform : The German Federation for the
559 Curation of Biological Data (GFBio), Inform. 2014 – Big Data Komplexität meistern. GI-Edition
560 Lect. Notes Informatics - Proc., 1711–1724, 2014.
561 Du, Z., Wang, Z., Zhao, J.-X. and Chen, G.: Woeseia oceani gen. nov., sp. nov., a novel
562 chemoheterotrophic member of the order Chromatiales, and proposal of Woeseiaceae fam. nov, Int.
563 J. Syst. Evol. Microbiol., 66, doi:10.1099/ijsem.0.000683, 2015.
564 Durbin, A. M. and Teske, A.: Microbial diversity and stratification of South Pacific abyssal marine
565 sediments, Environ. Microbiol., 13(12), 3219–3234, doi:10.1111/j.1462-2920.2011.02544.x, 2011.
566 Fernandes, A. D., Reid, J. N. S., Macklaim, J. M., McMurrough, T. A., Edgell, D. R. and Gloor, G. B.:
567 Unifying the analysis of high-throughput sequencing datasets: characterizing RNA-seq, 16S rRNA
568 gene sequencing and selective growth experiments by compositional data analysis, Microbiome,
569 2(1), 15, doi:10.1186/2049-2618-2-15, 2014.
570 Galac, M. R., Bosch, I., and Janies, D. A.: Bacterial communities of oceanic sea star (Asteroidea:
571 Echinodermata) larvae. Mar. Biol. 163, 162, DOI 10.1007/s00227-016-2938-3, 2016.
572 Gerpe, D., Buján, N., Diéguez, A. L., Lasa, A. and Romalde, J. L.: Kiloniella majae sp . nov ., isolated
573 from spider crab (Maja brachydactyla) and pullet carpet shell clam (Venerupis pullastra) ☙, Syst.
574 Appl. Microbiol., 40, 274–279, doi:10.1016/j.syapm.2017.05.002, 2017.
575 Gobet, A., Boetius, A. and Ramette, A.: Ecological coherence of diversity patterns derived from
576 classical fingerprinting and Next Generation Sequencing techniques, Environ. Microbiol., 16(9),
577 2672–2681, doi:10.1111/1462-2920.12308, 2014.
578 Gooday, A. J., Goineau, A. and Voltski, I.: Abyssal foraminifera attached to polymetallic nodules from
579 the eastern Clarion Clipperton Fracture Zone: a preliminary description and comparison with North
580 Atlantic dropstone assemblages, Mar. Biodivers., 45(3), 391–412, doi:10.1007/s12526-014-0301-9,
581 2015.
582 Haeckel, M., Konig, I., Reich, V., Weber, M. E. and Suess, E.: Pore water profiles and numerical
583 modelling of biogeochemical processes in Peru Basin deep-sea sediments. Deep-Sea Res Pt II 48,
584 3713-3736, doi.org/10.1016/S0967-0645(01)00064-9, 2001.
585 Halbach, P., Friedrich, G. and von Stackelberg, U.: The manganese nodule belt of the Pacific Ocean.
586 Enke, Stuttgart, p 254, 1988.
587 Hassenrück, C., Quast, C., Rapp, J., and Buttigieg, P.: Amplicon. GitHub repository,
588 https://github.com/chassenr/NGS/tree/master/AMPLICON, 2016.
589 Hoffmann, K., Bienhold, C., Buttigieg, P. L., Knittel, K., Laso-Pérez, R., Rapp, J. Z., Boetius, A. and
590 Offre, P.: Diversity and metabolism of Woeseiales bacteria, global members of deep-sea sediment
591 communities. The ISME Journal 14, 1042-1056, doi.org/10.1038/s41396-020-0588-4, 2020.
592 Hori, S., Tsuchiya, M., Nishi, S., Arai, W. and Takami, H.: Active Bacterial Flora Surrounding
593 Foraminifera (Xenophyophorea) Living on the Deep-Sea Floor, Biosci. Biotechnol. Biochem.,
594 77(2), 381–384, doi:10.1271/bbb.120663, 2013.

595 Hsieh, T.C., Ma, K. H. and Chao, A. iNEXT: iNterpolation and EXTrapolation for species diversity. R
596 package version 2.0.17 URL, 2018: http://chao.stat.nthu.edu.tw/blog/software-download/.
597 J. Müller, P., Hartmann, M. and Suess, E.: The chemical environment of pelagic sediments, in Halbach
598 P, Friedrich G, von Stackelberg U (eds) The manganese nodule belt of the Pacific ocean:
599 geological, environment, nodule formation, and mining aspects. Enke, Stuttgart, pp. 70–90., 1988.
600 Jacob, M., Soltwedel, T., Boetius, A. and Ramette, A.: Biogeography of Deep-Sea Benthic Bacteria at
601 Regional Scale (LTER HAUSGARTEN, Fram Strait, Arctic), PLoS One, 8(9), e72779 [online]
602 Available from: https://doi.org/10.1371/journal.pone.0072779, 2013.
603 Jacobson Meyers, M. E., Sylvan, J. B. and Edwards, K. J.: Extracellular enzyme activity and microbial
604 diversity measured on seafloor exposed basalts from Loihi seamount indicate the importance of
605 basalts to global biogeochemical cycling, Appl. Environ. Microbiol., 80(16), 4854–4864,
606 doi:10.1128/AEM.01038-14, 2014.
607 Kato, S., Ikehata, K., Shibuya, T., Urabe, T., Ohkuma, M. and Yamagishi, A.: Potential for
608 biogeochemical cycling of sulfur, iron and carbon within massive sulfide deposits below the
609 seafloor, Environ. Microbiol., 17(5), 1817–1835, doi:10.1111/1462-2920.12648, 2015.
610 Kerr, R. A.: Manganese Nodules Grow by Rain from Above, Science (80-.)., 223(4636), 576 LP –
611 577, doi:10.1126/science.223.4636.576, 1984.
612 Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M. and Glöckner, F. O.:
613 Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation
614 sequencing-based diversity studies, Nucleic Acids Res., 41(1), e1–e1, doi:10.1093/nar/gks808,
615 2012.
616 Kuhn, T., Wegorzewski, A., Rühlemann, C. and Vink, A.: Composition, Formation, and Occurrence of
617 Polymetallic Nodules, in Deep-Sea Mining: Resource Potential, Technical and Environmental
618 Considerations, edited by R. Sharma, pp. 23–63, Springer International Publishing, Cham., 2017.
619 Larsen, E. I., Sly, L. I. and Mcewan, A. G.: Manganese (II) adsorption and oxidation by whole cells
620 and a membrane fraction of Pedomicrobium sp . ACM 3067, Arch. Microbiol., 171(4), 257–264,
621 1999.
622 Lee, M. D., Walworth, N. G., Sylvan, J. B., Edwards, K. J. and Orcutt, B. N.: Microbial Communities
623 on Seafloor Basalts at Dorado Outcrop Reflect Level of Alteration and Highlight Global Lithic
624 Clades, Front. Microbiol., 6, 1470, doi:10.3389/fmicb.2015.01470, 2015.
625 Lindh, M. V., Maillot, B. M., Shulse, C. N., Gooday, A. J., Amon, D. J., Smith, C. R. and Church, M.
626 J.: From the surface to the deep-sea: Bacterial distributions across polymetallic nodule fields in the
627 clarion-clipperton zone of the Pacific Ocean, Front. Microbiol., 8(SEP), 1–12,
628 doi:10.3389/fmicb.2017.01696, 2017.
629 Luecker, S., Nowka, B., Rattei, T., Spieck, E. and Daims, H.: The Genome of Nitrospina gracilis
630 Illuminates the Metabolism and Evolution of the Major Marine Nitrite Oxidizer , Front. Microbiol.
631 , 4, 27 [online] Available from: https://www.frontiersin.org/article/10.3389/fmicb.2013.00027,
632 2013.
633 Mahé, F., Rognes, T., Quince, C., de Vargas, C. and Dunthorn, M.: Swarm: robust and fast clustering
634 method for amplicon-based studies, PeerJ, 2, e593, doi:10.7717/peerj.593, 2014.
635 Martens-Habbena, W., Berube, P., Urakawa, H., De la Torre, J. and Stahl, D.: Ammonia oxidation
636 kinetics determine niche separation of nitrifying Archaea and Bacteria, Nature, 461, 976–979,
637 doi:10.1038/nature08465, 2009.
638 Martin, M.: Cutadapt removes adapter sequences from high-throughput sequencing reads,
639 EMBnet.journal; Vol 17, No 1 Next Gener. Seq. Data Anal. - 10.14806/ej.17.1.200 [online]
640 Available from: https://journal.embnet.org/index.php/embnetjournal/article/view/200, 2011.
641 Mason, O. U., Di Meo-Savoie, C. A., Van Nostrand, J. D., Zhou, J., Fisk, M. R. and Giovannoni, S. J.:
642 Prokaryotic diversity, distribution, and insights into their role in biogeochemical cycling in marine
643 basalts, ISME J., 3(2), 231–242, doi:10.1038/ismej.2008.92, 2009.
644 Mason, O. U., Stingl, U., Wilhelm, L. J., Moeseneder, M. M., Di Meo-Savoie, C. A., Fisk, M. R. and
645 Giovannoni, S. J.: The phylogeny of endolithic microbes associated with marine basalts, Environ.
646 Microbiol., 9(10), 2539–2550, doi:10.1111/j.1462-2920.2007.01372.x, 2007.
647 Matsunaga, T.: Applications of bacterial magnets Mognet, Trends Biotechnol., 9(March), 91–95, 1991.
648 Mewes, K., Mogollón, J. M., Picard, A., Rühlemann, C., Kuhn, T., Nöthen, K. and Kasten, S.: Impact
649 of depositional and biogeochemical processes on small scale variations in nodule abundance in the
650 Clarion‐Clipperton Fracture Zone, Deep Sea Res. Part I Oceanogr. Res. Pap., 91, 125–141,
651 doi:https://doi.org/10.1016/j.dsr.2014.06.001, 2014.
652 Miller, K. A., Thompson, K. F., Johnston, P. and Santillo, D.: An Overview of Seabed Mining
653 Including the Current State of Development, Environmental Impacts, and Knowledge Gaps, Front.
654 Mar. Sci., 4(January 2018), doi:10.3389/fmars.2017.00418, 2018.

Murray, J. and Renard, A. F.: Report on deep sea deposits based on the specimens collected during the voyage of H.M.S. Challenger in the years 1872 - 1876, Rep. Sci. results Voyag. H.M.S. Chall. Dur. years 1873 - 1876, 1–688, doi:https://doi.org/10.1594/PANGAEA.849018, 1891.

Offre, P., Spang, A. and Schleper, C.: Archaea in biogeochemical cycles. Annu. Rev. Microbiol. 67, 437–457. doi: 10.1146/annurev-micro-092412-155614, 2013.

Oksanen, J., Blanchet, F. G., Kindt, R. et al.: vegan: Community Ecology Package. R package version 2.3-0, 2015.

Peukert, A., Schoening, T., Alevizos, E., Köser, K., Kwasnitschka, T. and Greinert, J.: Understanding Mn-nodule distribution and evaluation of related deep-sea mining impacts using AUV-based hydroacoustic and optical data, Biogeosciences, 15(8), 2525–2549, doi:10.5194/bg-15-2525-2018, 2018.

Pruesse, E., Peplies, J. and Glöckner, F. O.: SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes, Bioinformatics, 28(14), 1823–1829, doi:10.1093/bioinformatics/bts252, 2012.

Riemann, F.: Biological aspects of deep-sea manganese nodule formation, Oceanol. acta, 6(3), 303–311, 1983.

Ruff, S. E., Biddle, J. F., Teske, A. P., Knittel, K., Boetius, A. and Ramette, A.: Global dispersion and local diversification of the methane seep microbiome, Proc. Natl. Acad. Sci., 112(13), 4015 LP – 4020, doi:10.1073/pnas.1421865112, 2015.

Santelli, C. M., Orcutt, B. N., Banning, E., Bach, W., Moyer, C. L., Sogin, M. L., Staudigel, H. and Edwards, K. J.: Abundance and diversity of microbial life in ocean crust, Nature, 453(May), 5–9, doi:10.1038/nature06899, 2008.

Schauer, R., Bienhold, C., Ramette, A. and Harder, J.: Bacterial diversity and biogeography in deep-sea surface sediments of the South Atlantic Ocean, ISME J., 4, 159–170, doi:10.1038/ismej.2009.106, 2010.

Schüler, D. and Frankel, R.: Bacterial magnetosomes: Microbiology, biomineralization and biotechnological applications, Appl. Microbiol. Biotechnol., 52, 464–473, doi:10.1007/s002530051547, 1999.

Scott, J. J., Glazer, B. T. and Emerson, D.: Bringing microbial diversity into focus: high-resolution analysis of iron mats from the Lō'ihi Seamount, Environ. Microbiol., 19(1), 301–316, doi:10.1111/1462-2920.13607, 2017.

Shulse, C. N., Maillot, B., Smith, C. R. and Church, M. J.: Polymetallic nodules, sediments, and deep waters in the equatorial North Pacific exhibit highly diverse and distinct bacterial, archaeal, and microeukaryotic communities, Microbiologyopen, 6(2), 1‑16, doi:10.1002/mbo3.428, 2017.

Si, O., Yang, H., Hwang, C. Y., Kim, S., Choi, S., Kim, J., Jung, M., Kim, S., Roh, S. W. and Rhee, S.: Kiloniella antarctica sp . nov ., isolated from a polynya of Amundsen Sea in Western Antarctic Sea, Int. J. Syst. Evol. Microbiol., 67, 2397–2402, doi:10.1099/ijsem.0.001968, 2017.

Simon-Lledo, E., Bett, B. J., Huvenne, V. A. I., Köser, K., Schoening, T., Greinert, J., and Jones, D. O. B.: Biological effects 26 years after simulated deep-sea mining. Sci. Rep. 9, 8040, doi:10.1038/s41598-019-44492-w, 2019.

Staudigel, H., Furnes, H., McLoughlin, N., Banerjee, N. R., Connell, L. B. and Templeton, A.: 3.5 billion years of glass bioalteration: Volcanic rocks as a basis for microbial life?, Earth-Science Rev., 89(3), 156–176, doi:https://doi.org/10.1016/j.earscirev.2008.04.005, 2008.

Stein, L. Y., Duc, M. T. La, Grundl, T. J. and Nealson, K. H.: Bacterial and archaeal populations associated with freshwater ferromanganous micronodules and sediments, Environ. Microbiol., 3(1), 10–18, 2001.

Sylvan, J. B., Toner, B. M. and Edwards, K. J.: Life and Death of Deep-Sea Vents: Bacterial Diversity and Ecosystem Succession on Inactive Hydrothermal Sulfides, edited by M. A. Moran, MBio, 3(1), e00279-11, doi:10.1128/mBio.00279-11, 2012.

Thiel, H., Schriever, G., Ahnert, A., Bluhm, H., Borowski, C. and Vopel, K.: The large-scale environmental impact experiment DISCOL - Reflection and foresight, Deep. Res. Part II Top. Stud. Oceanogr., 48(17–18), 3869–3882, doi:10.1016/S0967-0645(01)00071-6, 2001.

Tian, R.-M., Sun, J., Cai, L., Zhang, W.-P., Zhou, G., Qui, J.-W. and Qian, P.-Y.: The deep‐sea glass sponge Lophophysema eversa harbours potential symbionts responsible for the nutrient conversions of carbon, nitrogen and sulfur, Environ. Microbiol., 18(8), 2481–2494, doi:10.1111/1462-2920.12911, 2016.

Tully, B. J. and Heidelberg, J. F.: Microbial communities associated with ferromanganese nodules and the surrounding sediments, Front. Microbiol., 4(JUN), 1–10, doi:10.3389/fmicb.2013.00161, 2013.

Tyler, P. A.: Hyphomicrobia and the oxidation of manganesse in aquatic ecosystems Mg /, Antonie von Leeuwenhoek, 36, 567–578, 1970.

Van den Boogaart, K. G., Tolosana, R. and Bren, M.: compositions: Compositional Data Analysis. R package version 1.40-1, 2014.

Vanreusel, A., Hilario, A., Ribeiro, P. A., Menot, L. and Arbizu, P. M.: Threatened by mining, polymetallic nodules are required to preserve abyssal epifauna, Sci. Rep., 6, 26808 [online] Available from: https://doi.org/10.1038/srep26808, 2016.

Varliero, G., Bienhold, C., Schmid, F., Boetius, A. and Molari, M.: Microbial Diversity and Connectivity in Deep-Sea Sediments of the South Atlantic Polar Front, Front. Microbiol., 10, 1–18, doi:10.3389/fmicb.2019.00665, 2019.

Volz, J. B., Mogollón, J. M., Geibert, W., Arbizu, P. M., Koschinsky, A. and Kasten, S.: Natural spatial variability of depositional conditions, biogeochemical processes and element fluxes in sediments of the eastern Clarion-Clipperton Zone, Pacific Ocean, Deep. Res. Part I Oceanogr. Res. Pap., 140(December 2017), 159–172, doi:10.1016/j.dsr.2018.08.006, 2018.

Von Stackelberg, U.: Growth history of manganese nodules and crusts of the Peru Basin, Geol. Soc. London, Spec. Publ., 119(1), 153 LP – 176, doi:10.1144/GSL.SP.1997.119.01.11, 1997.

Vonnahme, T. R., Molari, M., Janssen, F., Wenzhöfer, F., Haeckel, M., Titschack, J. and Boetius, A.: Effects of a deep-sea mining experiment on seafloor microbial communities and functions after 26 years. Science Advances, in press.

Walsh, E. A., Kirkpatrick, J. B., Rutherford, S. D., Smith, D. C., Sogin, M. and Hondt, S. D.: Bacterial diversity and community composition from seasurface to subseafloor, ISME J., 10, 979–989, doi:10.1038/ismej.2015.175, 2016.

Wang, C., Liao, L., Xu, H., Xu, X., Wu, M. and Zhu, L.: Bacterial Diversity in the Sediment from Polymetallic Nodule Fields of the Clarion-Clipperton Fracture Zone, J. Microbiol., 48(5), 573–585, doi:10.1007/s12275-010-0151-5, 2010.

Wang, L., Li, X., Lai, Q. and Shao, Z.: Kiloniella litopenaei sp . nov ., isolated from the gut microflora of Pacific white shrimp, Litopenaeus vannamei, Antonie Van Leeuwenhoek, 108, 1293–1299, doi:10.1007/s10482-015-0581-5, 2015.

Wang, L., Yu, M., Liu, Y., Liu, J., Wu, Y., Li, L., Liu, J., Wang, M. and Zhang, X.-H.: Comparative analyses of the bacterial community of hydrothermal deposits and seafloor sediments across Okinawa Trough, J. Mar. Syst., 180, 162–172, doi:https://doi.org/10.1016/j.jmarsys.2016.11.012, 2018.

Wang, X. H., Gan, L. and Müller, W. E. G.: Contribution of biomineralization during growth of polymetallic nodules and ferromanganese crusts from the Pacific Ocean, Front. Mater. Sci. China, 3(2), 109–123, doi:10.1007/s11706-009-0033-0, 2009.

Wang, Y. and Qian, P.-Y.: Conservative Fragments in Bacterial 16S rRNA Genes and Primer Design for 16S Ribosomal DNA Amplicons in Metagenomic Studies, PLoS One, 4(10), e7401 [online] Available from: https://doi.org/10.1371/journal.pone.0007401, 2009.

Wegorzewski, A. V. and Kuhn, T.: The influence of suboxic diagenesis on the formation of manganese nodules in the Clarion Clipperton nodule belt of the Pacific Ocean, Mar. Geol., 357, 123–138, doi:10.1016/j.margeo.2014.07.004, 2014.

Wegorzewski, A. V., Kuhn, T., Dohrmann, R., Wirth, R. and Grangeon, S.: Mineralogical characterization of individual growth structures of Mn-nodules with different Ni+Cu content from the central Pacific Ocean, Am. Mineral., 100(11–12), 2497–2508, doi:10.2138/am-2015-5122, 2015.

Wiese, J., Thiel, V., Ga, A., Schmaljohann, R. and Imhoff, J. F.: alphaproteobacterium from the marine macroalga Laminaria saccharina, Int. J. Syst. Evol. Microbiol., 59, 350–356, doi:10.1099/ijs.0.001651-0, 2009.

Wu, Y. H., Liao, L., Wang, C. S., Ma, W. L., Meng, F. X., Wu, M. and Xu, X. W.: A comparison of microbial communities in deep-sea polymetallic nodules and the surrounding sediments in the Pacific Ocean, Deep. Res. Part I Oceanogr. Res. Pap., 79, 40–49, doi:10.1016/j.dsr.2013.05.004, 2013.

Xu, M., Wang, F., Meng, J. and Xiao, X.: Construction and preliminary analysis of a metagenomic library from a deep-sea sediment of east Pacific Nodule Province, FEMS Microbiol. Ecol., 62(3), 233–241, doi:10.1111/j.1574-6941.2007.00377.x, 2007.

Yang, S., Seo, H., Lee, J., Kim, S. and Kwon, K. K.: Kiloniella spongiae sp . nov ., isolated from a marine sponge and emended description of the genus Kiloniella Wiese et al . 2009 and Kiloniella laminariae, Int. J. Syst. Evol. Microbiol., 65, 230–234, doi:10.1099/ijs.0.069773-0, 2015.

Zhang, G., He, J., Liu, F. and Zhang, L.: Iron-Manganese Nodules Harbor Lower Bacterial Diversity and Greater Proportions of Proteobacteria Compared to Bulk Soils in Four Locations Spanning from North to South China, Geomicrobiol. J., 31(7), 562–577, doi:10.1080/01490451.2013.854428, 2014.

Zhang, J., Kobert, K., Flouri, T. and Stamatakis, A.: PEAR: a fast and accurate Illumina Paired-End reAd mergeR, Bioinformatics, 30(5), 614–620, doi:10.1093/bioinformatics/btt593, 2013.

Figure captions

Figure 1. Comparison of diversity indices and unique OTUs between manganese nodules and sediments for (a) bacterial and (b) archaeal communities. H_0: number of OTUs (q=0); H_1: exponential Shannon (q=1); H_2: inverse Simpson (q=2); Unique: OTUs present exclusively in each station (percentage relative to total OTUs of whole dataset). Chao1 and Unique OTUs were calculated with 100 sequence re-samplings per sample to the smallest dataset (40613 sequences for Bacteria and 1835 sequences for Archaea). Red line shows the median. F: statistic *F-ratio*, with subscript numbers reporting the degrees of freedom between groups and within groups, respectively; p: probability level; KW-test: Kruskal-Wallis test; χ^2: Chi square test value, with subscript numbers reporting the degrees of freedom between groups and sample size, respectively.

Figure 2**.** Non-metric multidimensional scaling (NMDS) plot based on Euclidean distance similarity matrix of bacterial (a) and archaeal (b) community structure at OTU level. Sequence abundances of OTUs were centre log-ratio transformed. Permutational multivariate analysis of variance (PERMANOVA) showed significant differences between nodule and sediment associated microbial communities (for details see Table S1). Each sample (dot) is connected to the weighted averaged mean of the within group distances. Ellipses represent one SD of the weighted averaged mean.

Figure 3. Bacterial community structure at dominant class level (cut-off ≥1 %). MN: manganese nodules; MUC: sediments.

~~Figure 4. Bacterial (a) and Archaeal (b) dominant genera (cut-off ≥1 %) for surface nodules and sediments. Cluster on top of barplot showed dissimilarity in OTUs composition, as defined by Jaccard dissimilarity index based on presence/absence OTU table and calculated with 100 sequence re-samplings per sample on the smallest dataset (40613 sequences for Bacteria and 1835 sequences for Archaea). un.: unclassified. * due to extremely low number of sequences (n=182), this sample was not included in analysis requiring sequence re-samplings. MN: manganese nodules; MUC: sediments.~~

Figure 4. Genera highly abundant in nodules (ALDEx2: glm adjusted p<0.01; KW adjusted p<0.05). Base 2 logarithm of the ratios between geometric mean centred sequences number of nodule (Nod) and sediment (Sed), and average of the sequences contribution of total number of sequences (%) retrieved in nodules and in sediments are shown. For details see Table S3.

Figure 5. Habitats coverage for the closest related sequences (≥99 % similarity) to OTUs highly abundant in the nodules. For details see Table 4a-b.

Table captions

Table 1. Stations list and description of investigated sites/substrates.

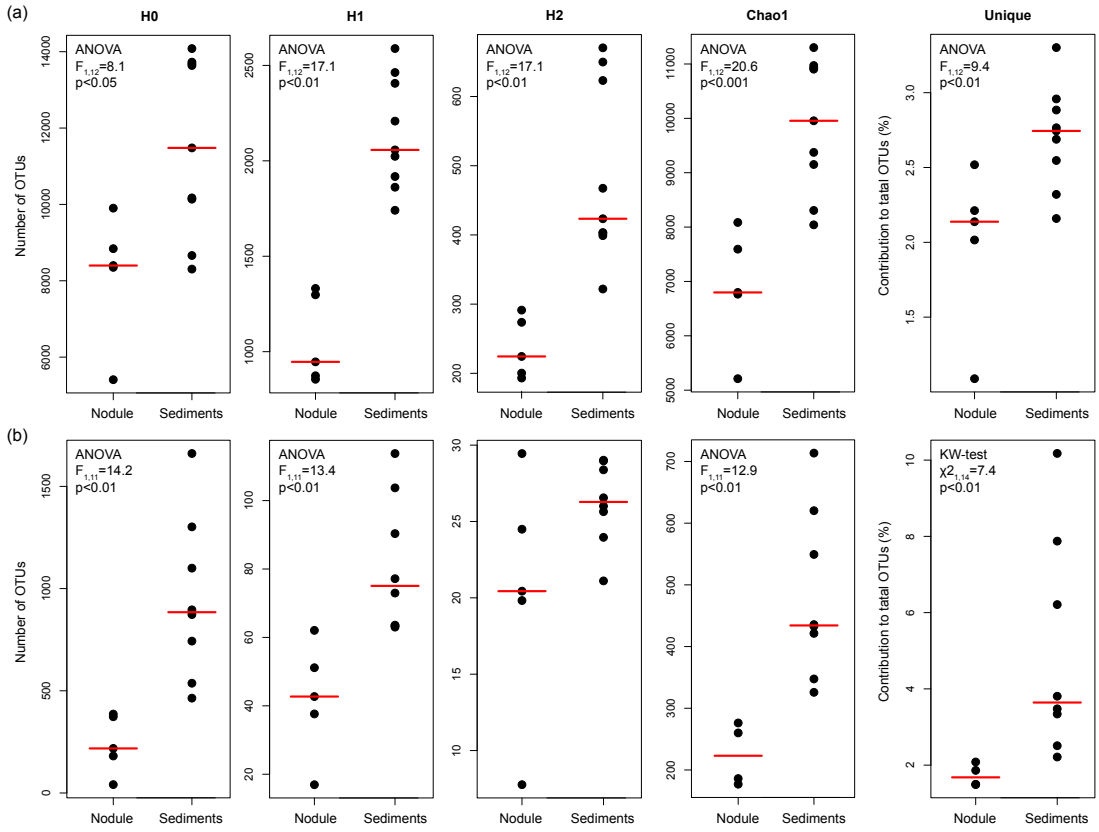
~~Table 2. Statistics of sequence and OTUs abundance, and proportion of absolute singletons, cosmopolitans and endemics for sediments (n=9) and nodule (n=5 for Bacteria, n=4 for Archaea) samples collected in Peru Basin. Absolute singletons: OTUs consisting of sequences occurring only once in the entire dataset; Cosmopolitan: OUTs present in 80 % of sediments and 80 % of nodule samples; Endemics: OTUs exclusively present only in 80 % sediments (and <20 % of nodule samples) or in 80 % nodule samples (and <20 % of sediments samples).~~

Table 2. Bacterial and archaeal diversity indices and unique OTUs for all nodules and sediment samples. Indices and unique OTUs were calculated without singletons.

~~Table 3. Genera differentially abundant in nodules and sediments (ALDEx2: glm adjusted p<0.01; KW adjusted p<0.05). In bold the most abundant genera (≥1 %) at least two times more abundant in nodule than in sediment; in italic the genera exclusively present (i.e. unique) in nodules. Base 2 logarithm of the ratios between geometric mean centred sequences number of nodule (Nod) and sediment (Sed), and average of the sequences contribution of total number of sequences (%) retrieved in nodules and in sediments are shown.~~

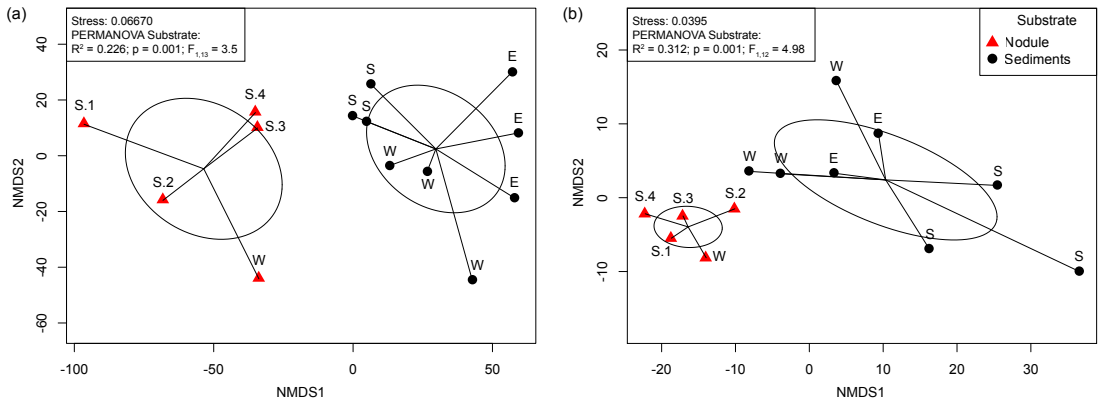
Table 3. (a) OTUs highly abundant in nodules (ALDEx2: glm adjusted p<0.01; KW adjusted p<0.05). Only OTUs ≥0.1 % are reported. Base 2 logarithm of the ratios between geometric mean centred sequences number of nodule (Nod) and sediment (Sed), and average of the sequences contribution of total number of sequences (%) retrieved in nodules and in sediments are shown. (b) Closest related sequences as indemnified with BLASTn (NCBI nucleotide database 12/06/2019).

829 **Figure 1.**



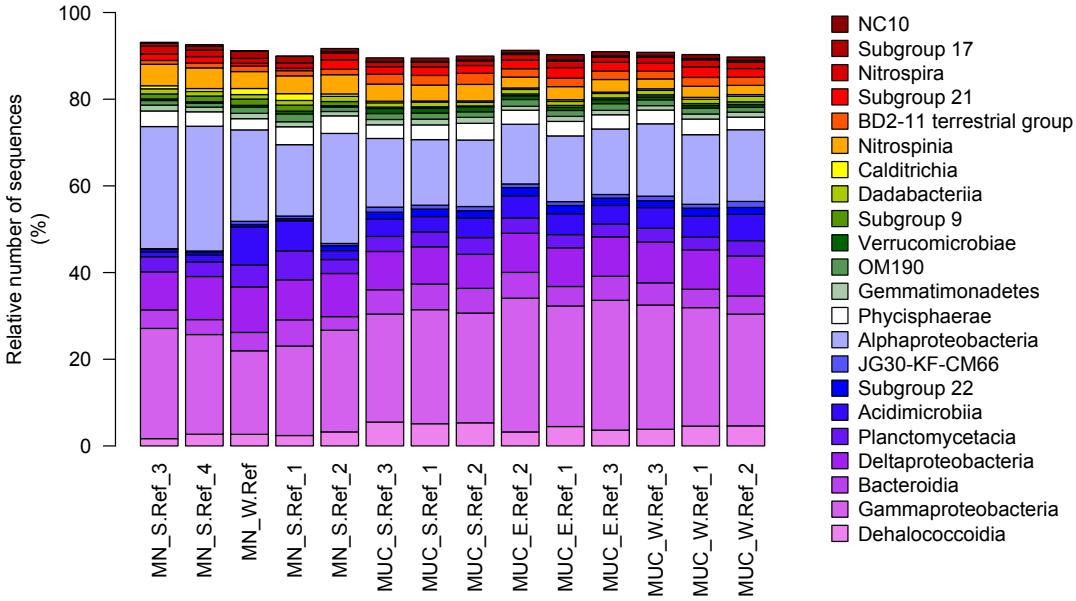
830
831

832 **Figure 2.**



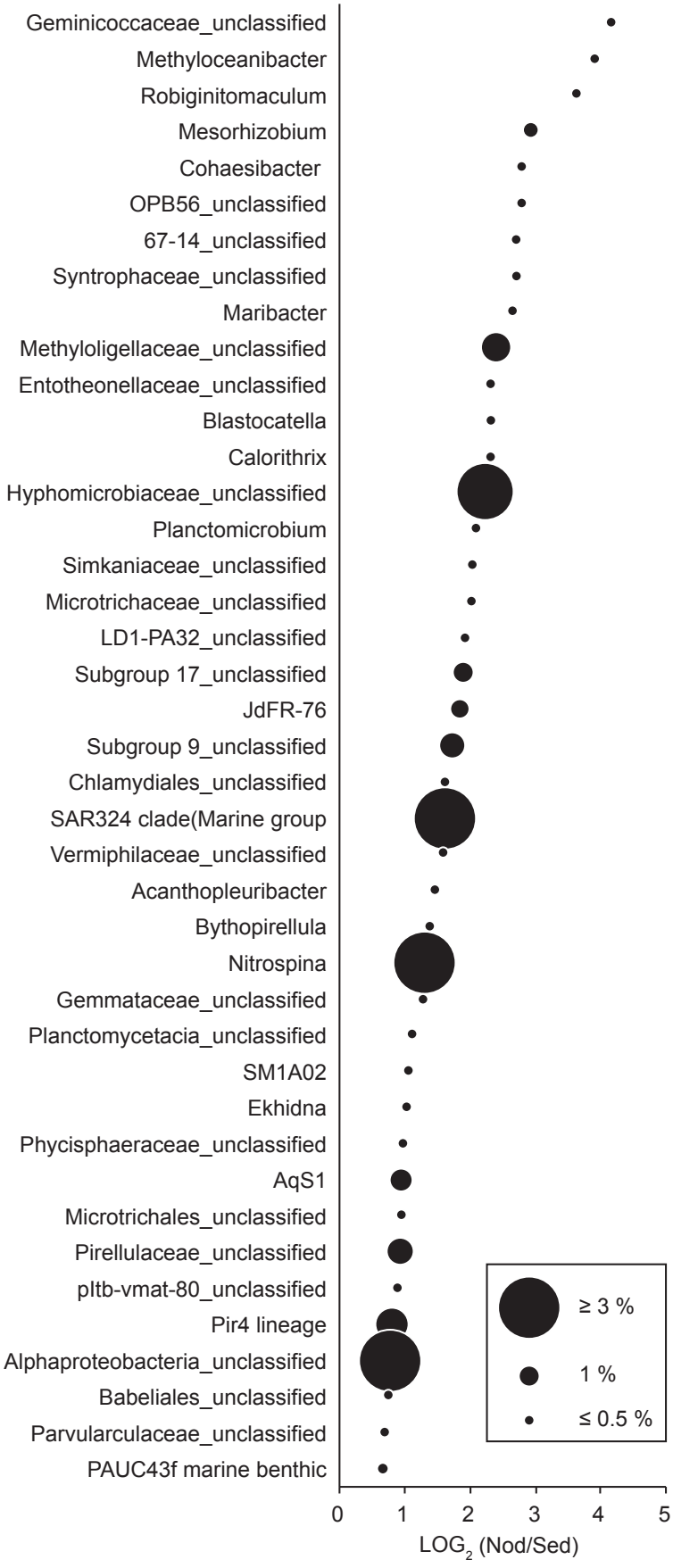
833
834

835 **Figure 3.**



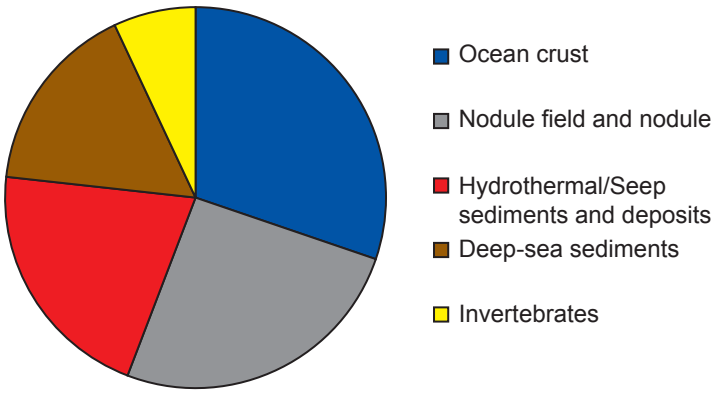
836

837 **Figure 4.**



838

839 **Figure 5.**



840
841

Table 1.

Station	Sample ID	Sampling Time	Latitude (N)	Longitude (E)	Depth (m)	Device	Site	Sediment layer (cm bsf)	Substrate
SO242/2_147	MUC_E.Ref_1	02.09.15	−7.1007	−88.414	4198.2	MUC	Reference East	0-1	sediments
SO242/2_148	MUC_E.Ref_2	02.09.15	−7.1006	−88.414	4195.8	MUC	Reference East	0-1	sediments
SO242/2_151	MUC_E.Ref_3	03.09.15	−7.1006	−88.414	4197.8	MUC	Reference East	0-1	sediments
SO242/2_194	MN_W.Ref	15.09.15	−7.0761	−88.526	4129.5	MUC	Reference West	surface	nodule
SO242/2_194	MUC_W.Ref_1	15.09.15	−7.0761	−88.526	4129.5	MUC	Reference West	0-1	sediments
SO242/2_194	MUC_W.Ref_2	15.09.15	−7.0761	−88.526	4129.5	MUC	Reference West	0-1	sediments
SO242/2_194	MUC_W.Ref_3	15.09.15	−7.0761	−88.526	4129.5	MUC	Reference West	0-1	sediments
SO242/2_198	MN_S.Ref_1	16.09.15	−7.1262	−88.450	4145.6	ROV	Reference South	surface	nodule
SO242/2_198	MN_S.Ref_2	16.09.15	−7.1262	−88.450	4145.6	ROV	Reference South	surface	nodule
SO242-2_208	MN_S.Ref_3	19.09.15	−7.1256	−88.450	4150.7	MUC	Reference South	surface	nodule
SO242-2_208	MN_S.Ref_4	19.09.15	−7.1256	−88.450	4150.7	MUC	Reference South	surface	nodule
SO242/2_208	MUC_S.Ref_1	15.09.15	−7.0761	−88.526	4129.5	MUC	Reference South	0-1	sediments
SO242/2_208	MUC_S.Ref_2	15.09.15	−7.0761	−88.526	4129.5	MUC	Reference South	0-1	sediments
SO242/2_208	MUC_S.Ref_3	15.09.15	−7.0761	−88.526	4129.5	MUC	Reference South	0-1	sediments

MUC: TV-guided MUltiple Corer; ROV: Remote Operated Vehicle (Kiel 6000); bsf: below seafloor.

Table 2.

Bacteria	Sequences n. [a]	Sequences n. [b]	H_0	H_1	H_2	Chao1 [c]	sd	Unique (%) [c]	sd
MUC_E.Ref_2	226078	161443	13638	2024	402	10930	201.6	2.7	0.1
MUC_E.Ref_3	218324	166847	13680	2057	423	10972	200.1	2.9	0.1
MUC_E.Ref_1	222924	164985	14082	2208	467	11302	166	3.0	0.1
MN_W.Ref	209563	159724	9902	1296	290	8085	143.6	2.2	0.1
MUC_W.Ref_1	137990	104301	11480	1918	403	9955	164	2.8	0.1
MUC_W.Ref_2	112259	81103	10171	1862	399	9151	148.3	2.3	0.1
MUC_W.Ref_3	236896	178985	13727	1741	322	10905	198.9	3.3	0.1
MN_S.Ref_1	313418	236498	8841	853	192	6798	138.3	2.5	0.1
MN_S.Ref_2	220364	172668	8399	872	199	6766	132.8	2.1	0.1
MN1_S.Ref_3	114074	43932	5409	945	223	5211	73.28	1.1	0.1
MN2_S.Ref_4	64218	76729	8351	1329	272	7594	124.2	2.0	0.1
MUC_S.Ref_1	77424	65890	10137	2588	670	9374	110.5	2.7	0.1
MUC_S.Ref_2	58575	45832	8662	2406	623	8306	93.85	2.2	0.1
MUC_S.Ref_3	59503	40613	8306	2463	650	8041	84.73	2.5	0.1

Archaea	Sequences n. [d]	Sequences n. [b]	H_0	H_1	H_2	Chao1 [c]	sd	Unique (%) [c]	sd
MUC_E.Ref_2	40952	34494	896	63	21	433	52	3.5	0.5
MUC_E.Ref_3	25090	20215	743	73	26	421	50	3.3	0.6
MUC_E.Ref_1	na	na	na	na	na	na	na	na	na
MN_W.Ref	11737	12623	373	51	24	260	33	2.1	0.4
MUC_W.Ref_1	18097	14878	537	63	24	348	36	2.5	0.5
MUC_W.Ref_2	37656	31192	873	77	28	436	53	3.8	0.6
MUC_W.Ref_3	13031	10444	464	64	26	326	31	2.2	0.4
MN_S.Ref_1	7423	5384	218	38	20	186	26	1.5	0.3
MN_S.Ref_2	15314	9472	386	62	29	276	29	1.5	0.4
MN1_S.Ref_3	6099	1835	181	43	20	177	15	1.9	0.3
MN2_S.Ref_4 [e]	722	182	41	17	8	na	na	na	na
MUC_S.Ref_1	34166	29221	1100	90	27	549	66	6.2	0.6
MUC_S.Ref_2	41633	36378	1302	104	29	620	71	7.9	0.8
MUC_S.Ref_3	75344	64433	1661	114	29	714	75	10.2	0.9

H_0: number of OTUs; H_1: exponential Shannon; H_2: inverse Simpson; Unique: OTUs present exclusively in one station (percentage relative to total OTUs of whole dataset); na: not available.

[a] after the merging of forward and reverse reads;

[b] after removal of un-specific and contaminants sequences (see Methods for details);

[c] calculated with 100 sequence re-samplings per sample to the smallest dataset (40613 sequences for Bacteria and 1835 sequences for Archaea), average data and standard deviation (sd) are given;

[d] after quality trimming of merged forward and reverse reads;

[e] due to extremely low number of sequences, this sample was not included in analyses requiring sequences re-sampling.

Table 3.

(a)

Phylum	Class	Order	Family	Genus	OTU	LOG2 (Nod/Sed)	Nodule (%)	Sediment (%)
Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	Hyphomicrobiaceae_unclassified	otu29	2	2.5	0.7
Proteobacteria	Alphaproteobacteria	Rhodospirillales	Magnetospiraceae	Magnetospiraceae_unclassified	otu11	2	2.2	0.7
Proteobacteria	Alphaproteobacteria	Alphaproteobacteria_unclassified	Alphaproteobacteria_unclassified	Alphaproteobacteria_unclassified	otu31	8	1.5	0.0
Proteobacteria	Alphaproteobacteria	Alphaproteobacteria_unclassified	Alphaproteobacteria_unclassified	Alphaproteobacteria_unclassified	otu83	8	0.9	0.0
Proteobacteria	Alphaproteobacteria	Alphaproteobacteria_unclassified	Alphaproteobacteria_unclassified	Alphaproteobacteria_unclassified	otu160	6	0.3	0.0
Proteobacteria	Alphaproteobacteria	Alphaproteobacteria_unclassified	Alphaproteobacteria_unclassified	Alphaproteobacteria_unclassified	otu249	7	0.2	0.0
Proteobacteria	Deltaproteobacteria	SAR324 clade(Marine group B)	SAR324 clade(Marine group B)_unclassified	SAR324 clade(Marine group B)_unclassified	otu66	8	0.7	0.0
Proteobacteria	Deltaproteobacteria	SAR324 clade(Marine group B)	SAR324 clade(Marine group B)_unclassified	SAR324 clade(Marine group B)_unclassified	otu78	2	0.6	0.1
Proteobacteria	Deltaproteobacteria	SAR324 clade(Marine group B)	SAR324 clade(Marine group B)_unclassified	SAR324 clade(Marine group B)_unclassified	otu202	3	0.4	0.0
Proteobacteria	Deltaproteobacteria	SAR324 clade(Marine group B)	SAR324 clade(Marine group B)_unclassified	SAR324 clade(Marine group B)_unclassified	otu317	1	0.2	0.1
Proteobacteria	Deltaproteobacteria	SAR324 clade(Marine group B)	SAR324 clade(Marine group B)_unclassified	SAR324 clade(Marine group B)_unclassified	otu947	4	0.2	0.0
Proteobacteria	Deltaproteobacteria	SAR324 clade(Marine group B)	SAR324 clade(Marine group B)_unclassified	SAR324 clade(Marine group B)_unclassified	otu588	2	0.1	0.0
Proteobacteria	Deltaproteobacteria	SAR324 clade(Marine group B)	SAR324 clade(Marine group B)_unclassified	SAR324 clade(Marine group B)_unclassified	otu425	2	0.1	0.0
Nitrospinae	Nitrospinia	Nitrospinales	Nitrospinaceae	Nitrospina	otu68	3	1.4	0.1
Nitrospinae	Nitrospinia	Nitrospinales	Nitrospinaceae	Nitrospina	otu227	6	0.2	0.0
Nitrospinae	Nitrospinia	Nitrospinales	Nitrospinaceae	Nitrospina	otu215	3	0.2	0.0
Nitrospirae	Nitrospira	Nitrospirales	Nitrospiraceae	Nitrospira	otu636	6	0.2	0.0
Nitrospinae	Nitrospinia	Nitrospinales	Nitrospinaceae	Nitrospina	otu434	3	0.1	0.0
Proteobacteria	Gammaproteobacteria	Arenicellales	Arenicellaceae	Arenicellaceae_unclassified	otu36	2	1.4	0.4
Proteobacteria	Gammaproteobacteria	Arenicellales	Arenicellaceae	Arenicellaceae_unclassified	otu162	5	0.4	0.0
Proteobacteria	Gammaproteobacteria	Steroidobacterales	Woeseiaceae	Woeseia	otu97	4	0.5	0.0
Proteobacteria	Gammaproteobacteria	Steroidobacterales	Woeseiaceae	Woeseia	otu266	2	0.2	0.1
Proteobacteria	Gammaproteobacteria	Steroidobacterales	Woeseiaceae	Woeseia	otu521	6	0.2	0.0
Proteobacteria	Gammaproteobacteria	Steroidobacterales	Woeseiaceae	Woeseia	otu346	4	0.2	0.0
Proteobacteria	Gammaproteobacteria	Steroidobacterales	Woeseiaceae	Woeseia	otu991	5	0.1	0.0
Proteobacteria	Alphaproteobacteria	Rhizobiales	Methyloligellaceae	Methyloligellaceae_unclassified	otu113	2	0.5	0.1
Proteobacteria	Alphaproteobacteria	Rhizobiales	Methyloligellaceae	Methyloligellaceae_unclassified	otu184	7	0.3	0.0
Proteobacteria	Alphaproteobacteria	Rhizobiales	Methyloligellaceae	Methyloligellaceae_unclassified	otu234	3	0.2	0.0
Acidobacteria	Subgroup 9	Subgroup 9_unclassified	Subgroup 9_unclassified	Subgroup 9_unclassified	otu255	6	0.5	0.0
Proteobacteria	Gammaproteobacteria	Nitrosococcales	Nitrosococcaceae	AqS1	otu122	2	0.8	0.3
Acidobacteria	Subgroup 17	Subgroup 17_unclassified	Subgroup 17_unclassified	Subgroup 17_unclassified	otu326	5	0.7	0.0
Acidobacteria	Subgroup 17	Subgroup 17_unclassified	Subgroup 17_unclassified	Subgroup 17_unclassified	otu865	1	0.1	0.1
Calditrichaeota	Calditrichia	Calditrichales	Calditrichaceae	JdFR-76	otu171	1	0.6	0.2
Calditrichaeota	Calditrichia	Calditrichales	Calditrichaceae	JdFR-76	otu541	4	0.2	0.0
Proteobacteria	Alphaproteobacteria	Rhodovibrionales	Kiloniellaceae	Kiloniellaceae_unclassified	otu357	3	0.1	0.0
Proteobacteria	Alphaproteobacteria	Rhodovibrionales	Kiloniellaceae	Kiloniellaceae_unclassified	otu435	4	0.1	0.0
Proteobacteria	Alphaproteobacteria	Rhodovibrionales	Kiloniellaceae	Kiloniellaceae_unclassified	otu370	3	0.1	0.0
Proteobacteria	Alphaproteobacteria	Rhodovibrionales	Kiloniellaceae	Kiloniellaceae_unclassified	otu467	6	0.1	0.0
Proteobacteria	Alphaproteobacteria	Rhodovibrionales	Kiloniellaceae	Kiloniellaceae_unclassified	otu450	2	0.1	0.0
Proteobacteria	Alphaproteobacteria	Rhodovibrionales	Kiloniellaceae	Kiloniellaceae_unclassified	otu519	3	0.1	0.0
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Cohaesibacter	otu71	4	0.7	0.1
Actinobacteria	Acidimicrobiia	Actinomarinales	Actinomarinales_unclassified	Actinomarinales_unclassified	otu163	2	0.3	0.1
Actinobacteria	Acidimicrobiia	Actinomarinales	Actinomarinales_unclassified	Actinomarinales_unclassified	otu532	6	0.2	0.0
Acidobacteria	Subgroup 9	Subgroup 9_unclassified	Subgroup 9_unclassified	Subgroup 9_unclassified	otu342	3	0.3	0.0
Acidobacteria	Subgroup 9	Subgroup 9_unclassified	Subgroup 9_unclassified	Subgroup 9_unclassified	otu674	3	0.1	0.0
Gemmatimonadetes	Gemmatimonadetes	Gemmatimonadales	Gemmatimonadaceae	Gemmatimonadaceae_unclassified	otu203	1	0.4	0.2
Proteobacteria	Alphaproteobacteria	Kordiimonadales	Kordiimonadaceae	Kordiimonas	otu86	2	0.4	0.1
Bacteroidetes	Bacteroidia	Cytophagales	Cyclobacteriaceae	Cyclobacteriaceae_unclassified	otu233	3	0.4	0.0
Dadabacteria	Dadabacteriia	Dadabacteriales	Dadabacteriales_unclassified	Dadabacteriales_unclassified	otu347	3	0.2	0.0
Dadabacteria	Dadabacteriia	Dadabacteriales	Dadabacteriales_unclassified	Dadabacteriales_unclassified	otu1016	3	0.1	0.0
Actinobacteria	Thermoleophilia	Solirubrobacterales	67-14	67-14_unclassified	otu324	3	0.3	0.0
Proteobacteria	Deltaproteobacteria	NB1-j	NB1-j_unclassified	NB1-j_unclassified	otu344	1	0.1	0.0
Planctomycetes	Planctomycetacia	Pirellulales	Pirellulaceae	Pirellulaceae_unclassified	otu538	2	0.1	0.0
Actinobacteria	Acidimicrobiia	Microtrichales	Microtrichaceae	Microtrichaceae_unclassified	otu669	2	0.1	0.0
Acidobacteria	Blastocatellia (Subgroup 4)	Blastocatellales	Blastocatellaceae	Blastocatella	otu489	3	0.1	0.0
Entotheonellaeota	Entotheonellia	Entotheonellales	Entotheonellaceae	Entotheonellaceae_unclassified	otu788	4	0.1	0.0
Acidobacteria	Thermoanaerobaculia	Thermoanaerobaculales	Thermoanaerobaculaceae	Subgroup 10	otu711	3	0.1	0.0
Proteobacteria	Gammaproteobacteria	Oceanospirillales	Kangiellaceae	Kangiellaceae_unclassified	otu744	5	0.1	0.0
Proteobacteria	Gammaproteobacteria	Thiohalorhabdales	Thiohalorhabdaceae	Thiohalorhabdaceae_unclassified	otu571	6	0.1	0.0
Bacteroidetes	Bacteroidia	Cytophagales	Cyclobacteriaceae	Ekhidna	otu651	5	0.1	0.0
Bacteroidetes	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Flavobacteriaceae_unclassified	otu579	6	0.1	0.0
Gemmatimonadetes	BD2-11 terrestrial group	BD2-11 terrestrial group_unclassified	BD2-11 terrestrial group_unclassified	BD2-11 terrestrial group_unclassified	otu1439	3	0.1	0.0

(b)

OTU	NCBI ID ≥ 99% similarity	Habitat(s)
otu29	KT748605.1; JX227334.1; EU491654.1	basaltic crust; nodule fields
otu11	JX227511.1; JQ013353.1; FJ938664.1	nodule fields; deep-sea sediments; cobalt-rich crust
otu31	MG580220.1; KF268757.1	Mariana subduction zone sediments; heavy metal contaminated marine sediments
otu83	MG580220.1; JN621543.1	Mariana subduction zone sediments; manganese oxide-rich marine sediments
otu160	MG580740.1; JX227257.1	Mariana subduction zone sediments; nodule fields
otu249	JQ287236.1; KM051824.1	inactive hydrothermal sulfides; basaltic crust
otu66	JX226721.1 [a]	nodule fields
otu78	JN860354.1; HQ721444.1	hydrothermal vents; deep-sea sediments;
otu202	MG580143.1; JX227690.1; JN860358.1	Mariana subduction zone sediments; nodule fields; hydrothermal vents
otu317	JX227432.1; AY627518.1	nodule fields; deep-sea sediments;
otu947	JX226721.1 [a]	nodule fields
otu588	LC081043.1	nodule
otu425	JX227680.1; FJ938661.1	nodule fields; cobalt-rich crust
otu68	JN886931.1; FJ752931.1; KJ590663.1	hydrothermal carbonate sediments; polychaete burrow environment; biofilm
otu227	MG580382.1; AM997732.1	Mariana subduction zone sediments; deep-sea sediments
otu215	KC901562.1; AB015560.1	basaltic glasses; deep-sea sediments
otu636	HM101002.1; EU491612.1; KC682687.1	Marine Sponge Halichondria; ocean crust;
otu434	EU287401.1; JN977323.1	Subsurface sediments; marine sediments
otu36	JX227383.1; KY977840.1; AM997938.1	nodule fields; Mariana subduction zone sediments; deep-sea sediments
otu162	FN553503.1; AM997671.1	hydrothermal vents; deep-sea sediments
otu97	JX227693.1; FJ024322.1; EU491736.1	nodule fields; ocean crust
otu266	AB694157.1; JX227083.1	deep-sea benthic foraminifera; nodule fields
otu521	KY977757.1; KT336088.1; JX227223.1	Mariana subduction zone sediments; nodules; nodule fields
otu346	KY977757.1; JX227223.1	Mariana subduction zone sediments; nodule fields
otu991	JX227363.1; AM997733.1	nodule fields; deep-sea sediments
otu113	JX226757.1; EU491557.1	nodule fields; ocean crust
otu184	EU491404.1	ocean crust
otu234	EU491604.1	ocean crust
otu255	JX227709.1; FJ437705.1; KM110219.1	nodule fields; hydrothermal deposits
otu122	MG580277.1; AM997814.1; AJ966605.1	Mariana subduction zone sediments; deep-sea sediments; nodule fields
otu326	JN886905.1; KT748584.1	hydrothermal carbonate sediments; basalt crust
otu865	JX227375.1; FJ938651.1; AY225640.1	nodule fields; cobalt-rich crust; hydrothermal sediments
otu171	AM997407.1; FJ205352.1; EU491267.1	deep-sea sediments; hydrothermal vents; ocean crust
otu541	AB694393.1	deep-sea benthic foraminifera
otu357	EU236317.1; GU302472.1	marine sponge; hydrocarbon seep
otu435	KY609381.1; KM051717.1; JX226899.1	Fe-rich hydrothermal deposits; basaltic crust; nodule fields
otu370	EU491648.1 [a]	ocean crust
otu467	FN553612.1; AB858542.1; KM051770.1	hydrothermal vents; sulfide deposits; basaltic crust
otu450	AM997745.1; KM051762.1; EU491108.1	deep-sea sediments; basaltic crust; ocean crust
otu519	GU220747.1; MG580729.1	Fe-rich hydrothermal deposits; Mariana subduction zone sediments
otu71	FJ205181.1; JX226787.1	hydrothermal vents; noduel fields
otu163	JX227427.1; JN886907.1; EU491661.1	nodule fields; hydrothermal carbonate sediments; ocean crust
otu532	EU491402.1; JX227188.1; EU374100.1	ocean crust; nodule fields; deep-sea sediments
otu342	JX227410.1; FJ205219.1; KT336055.1	nodule fields; hydrothermal vents; nodules
otu674	JX227662.1; KT336085.1; FJ938601.1	nodule fields; nodules; cobalt-rich crust
otu203	KP305065.1; FJ938598.1	corals; cobalt-rich crust
otu86	AM997620.1; FJ938474.1	deep-sea sediments; cobalt-rich crust
otu233	JX227464.1; AM997441.1	nodule fields; deep-sea sediments
otu347	JX227062.1; EU491655.1	nodule fields; ocean crust
otu1016	KF616695.1; KM396663.1; EU491261.1	carbonate methane seep; brine seep; ocean crust
otu324	JX226791.1; JN886912.1	nodule fields; hydrothermal carbonate sediments
otu344	EU438185.1; KY977824.1	deep-sea sediments and hydrohtermal vents; ocean crust
otu538	KM356353.1; JX226930.1; DQ996924.1	carbonate methane seep; nodule fields; deep-sea sediments
otu669	EU491619.1; MG580068.1; KT748607.1	ocean crust
otu489	EU491660.1; MG580531.1; AM998023.1	ocean crust; deep-sea sediments
otu788	JN886890.1; MG580099.1	hydrothermal carbonate sediments; ocean crust
otu711	JX193423.1; GU302449.1; AY225643.1	mariculture sediments; hydrocarbon seep; ocean crust
otu744	AB831375.1; EU290406.1; KM454306.1	deep-sea methane-seep sediments; marine sponge; marine sediments
otu571	JQ287033.1; AM911385.1; EU236424.1	hydrothermal sulfides; cold-water corals; sponges
otu651	KT972875.1 [a]	outcrops
otu579	EU491573.1; KT336070.1	ocean crust; nodules
otu1439	JN886922.1; KC747092.1; JN884864.1	hydrothermal carbonate sediments; deep-sea sediments; methane seep

861

862 [a] ≥ 98% similarity.

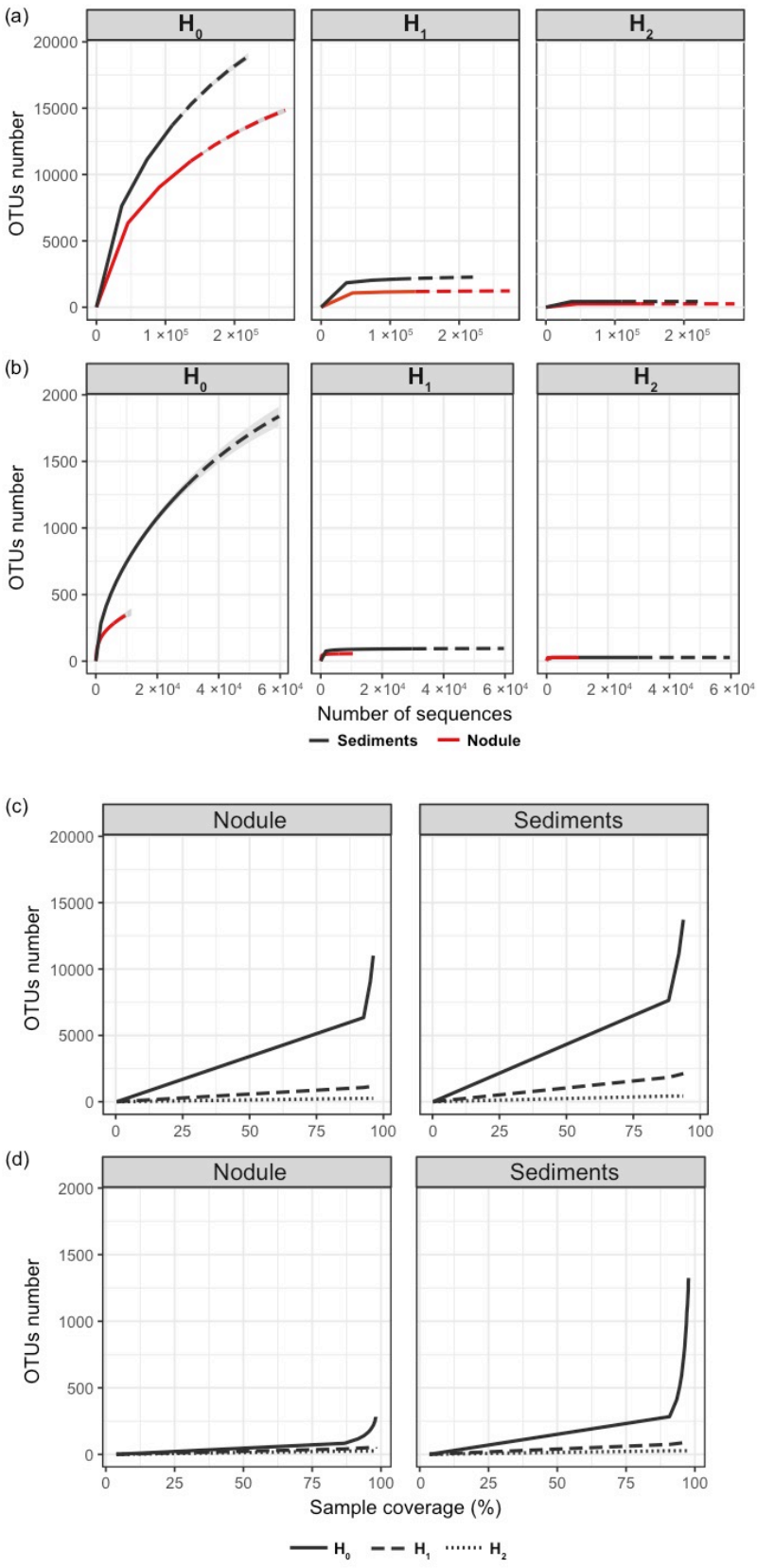
Supplementary Information

Microbial communities associated with sediments and polymetallic nodules of the Peru Basin

Massimiliano Molari, Felix Janssen, Tobias R. Vonnahme, Frank Wenzhöfer and Antje Boetius

Correspondence to: Massimiliano Molari (mamolari@mpi-bremen.de)

Figure S1. Rarefaction curves and diversity coverage in manganese nodules and sediments. a-b) show sample-size-based rarefaction curve for Bacteria and Archaea, respectively; c-d) show coverage-based curves for Bacteria and Archaea, respectively. The solid lines represent the observed accumulation with the number of sequences sampled, and the dashed lines represent the extrapolated accumulation up to double amount of sequences (only in a-b plots). Shaded area showed the 95 % confidence intervals based on 100 bootstrap replications. Knots = 10 for Bacteria, and knots = 40 for Archaea. H0: number of OTUs (q=0); H_1: exponential Shannon (q=1); H_2: inverse Simpson (q=2).


(a)
H0
H1
H2
OTUs number
20000
15000
10000
5000
0
0
1 ×10^5
2 ×10^5
(b)
2000
1500
1000
500
2 ×10^4
4 ×10^4
6 ×10^4
Number of sequences
Sediments
Nodule
(c)
Nodule
Sediments
25
50
75
100
(d)
Sample coverage (%)
H0
H1
H2

878

Figure S2. Statistics of sequence and OTUs abundance, and proportion of absolute singletons, cosmopolitans and endemics for sediments (n=9) and nodule (n=5 for Bacteria, n=4 for Archaea) samples collected in Peru Basin. Absolute singletons: OTUs consisting of sequences occurring only once in the entire dataset; Cosmopolitan: OTUs present in 80 % of sediments and 80 % of nodule samples; Endemics: OTUs exclusively present only in 80 % sediments (and <20 % of nodule samples) or in 80 % nodule samples (and <20 % of sediments samples).

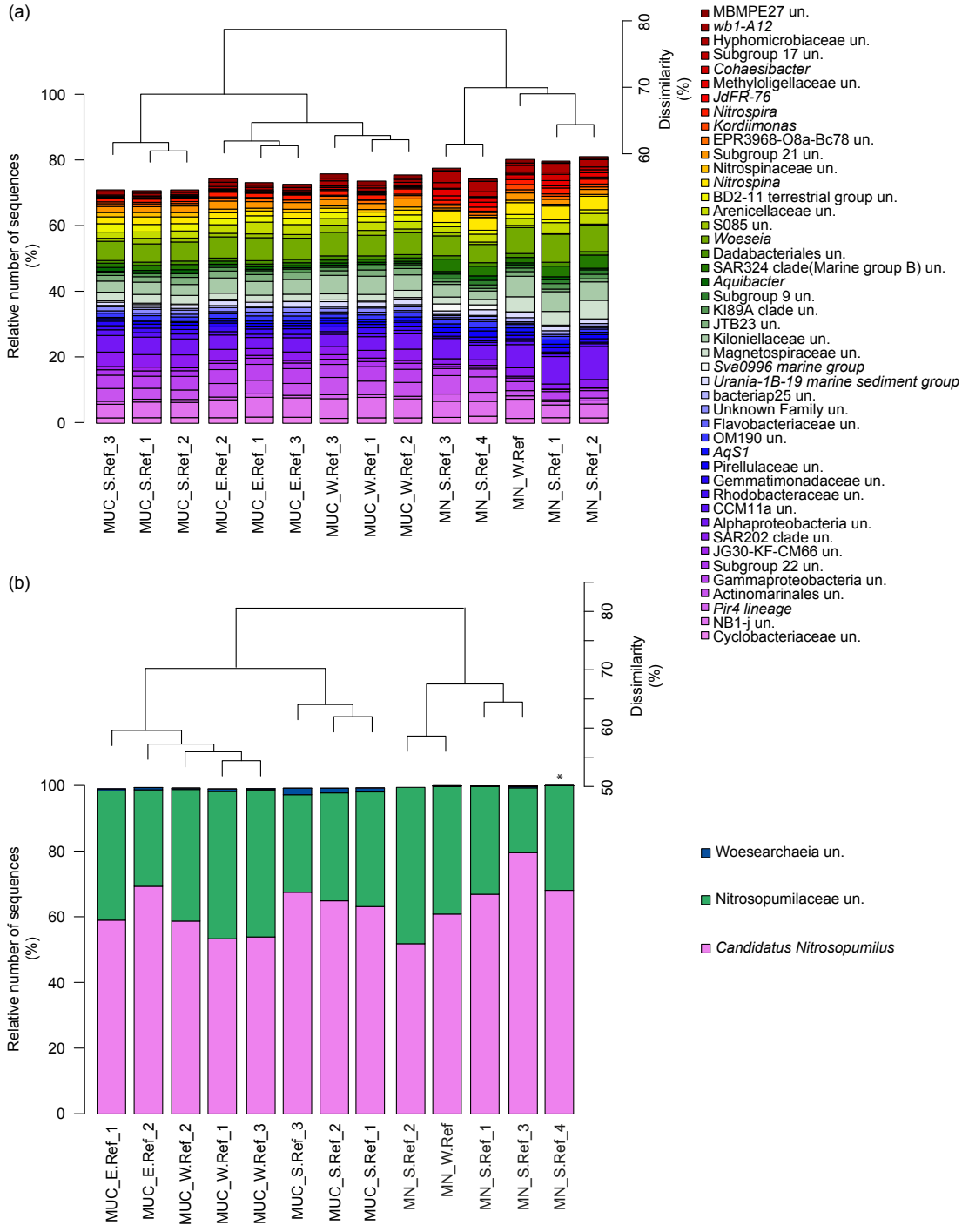


Table S1. Statistics of sequence and OTUs abundance, and proportion of absolute singletons, cosmopolitans and endemics for sediments (n=9) and nodule (n=5 for Bacteria, n=4 for Archaea) samples collected in Peru Basin. Absolute singletons: OTUs consisting of sequences occurring only once in the entire dataset; Cosmopolitan: OTUs present in 80 % of sediments and 80 % of nodule samples; Endemics: OTUs exclusively present only in 80 % sediments (and <20 % of nodule samples) or in 80 % nodule samples (and <20 % of sediments samples).

Bacteria	OTUs		Sequences	
	n.	%	n.	%
Entire dataset	557468		2271610	
Contaminants	20	0.0	15710	0.7
Absolute singletons	525169	94.2 (56 / 39) [b]	525169	23.1 (14 / 9) [b]
Working dataset [a]	32279	5.8	1730731	76.2
Sediments dataset [a]	28666	5.1 (88.8) [c]	1032246	45.4 (59.6) [c]
Nodule dataset [a]	19279	3.5 (59.7) [c]	698485	30.8 (40.3) [c]
Cosmopolitan OTUs [a]	1452	0.5 (8.9) [c]	1167668	58..4 (76.7) [c]
Endemics OTUs sediments [a]	1356	0.2 (4.2) [c]	39895	1.8 (2.3) [c]
Endemics OTUs nodules [a]	599	0.1 (1.9) [c]	52328	2.3 (3.0) [c]

Archaea	OTUs		Sequences	
	n.	%	n.	%
Entire dataset	51856		293098	
Contaminants	0	0.0	0	0.0
Absolute singletons	49482	95.4 (77 / 19) [b]	49482	16.9 (14 / 3) [b]
Working dataset [a]	2372	4.6	243616	83.1
Sediments dataset [a]	2356	4.5 (99.3) [c]	219460	74.9 (90.1) [c]
Nodule dataset [a]	591	1.1 (24.9) [c]	24156	8.2 (9.9) [c]
Cosmopolitan OTUs [a]	112	0.2 (4.7) [c]	194736	66.4 (79.9) [c]
Endemics OTUs sediments [a]	198	0.4 (8.3) [c]	10610	3.6 (4.4) [c]
Endemics OTUs nodules [a]	5	0.01 (0.2) [c]	121	0.04 (0.05) [c]

[a] after removal of contaminants (defined by negative control) and absolute singletons sequences (see Methods for details), percentage calculated on Entire dataset.
[b] contribution of sediments and nodules to absolute singletons, respectively.
[c] percentage calculated on working dataset.

Table S2. Output permutational multivariate analysis of variance on distance matrices (PERMANOVA).

OTUs CLR tranformed and Euclidean distance

	Bacteria							Archaea						
Substrates		Df	SS	MS	F	R^2	P		Df	SS	MS	F	R^2	P
	Substrates	1	21547	21546.7	3.5043	0.22602	0.001	Substrates	1	3014.5	3014.53	4.9801	0.31164	0.001
	Residuals	12	73783	6148.6		0.77398		Residuals	11	6658.5	605.32		0.68836	
	Total	13	95330			1		Total	12	9673			1	
Sites/Sediment		Df	SS	MS	F	R^2	P		Df	SS	MS	F	R^2	P
	Sites	2	18360	9180.2	1.8699	0.38397	0.003	Sites	2	2542.1	1271.05	2.3048	0.47969	0.013
	Residuals	6	29458	4909.6		0.61603		Residuals	5	2757.3	551.47		0.52031	
	Total	8	47818			1		Total	7	5299.4			1	
Reference.South/Substrate		Df	SS	MS	F	R^2	P		Df	SS	MS	F	R^2	P
	Substrates	1	12771	12770.7	2.5848	0.34079	0.023	Substrates	1	3369.1	3369.1	7.5272	0.60087	0.029
	Residuals	5	24703	4940.7		0.65921		Residuals	5	2237.9	447.6		0.39913	
	Total	6	37474			1		Total	6	5607			1	
Sites/Substrates (Strata=Site)		Df	SS	MS	F	R^2	P		Df	SS	MS	F	R^2	P
	Sites	2	25842	12921	2.4356	0.27108	0.005	Sites	2	1837	918.5	1.8506	0.18991	0.021
	Sites:Substrates	2	21743	10871	2.0492	0.22808	0.005	Sites:Substrate	1	3369.1	3369.1	6.7881	0.3483	0.021
	Residuals	9	47745	5305		0.50084		Residuals	9	4466.9	496.3		0.46179	
	Total	13	95330			1		Total	12	9673			1	

OTUs P/A table and Jaccard dissimilarity [A]

	Bacteria							Archaea						
Substrates		Df	SS	MS	F	R^2	P		Df	SS	MS	F	R^2	P
	Substrates	1	0.5986	0.59863	2.7963	0.18899	0.002	Substrates	1	0.45274	0.45274	2.2661	0.18474	0.003
	Residuals	12	2.5689	0.21408		0.81101		Residuals	10	1.9979	0.19979		0.81526	
	Total	13	3.1676			1		Total	11	2.45064			1	
Sites/Sediment		Df	SS	MS	F	R^2	P		Df	SS	MS	F	R^2	P
	Sites	2	0.50624	0.25312	1.3286	0.30693	0.002	Sites	2	0.52048	0.26024	1.4829	0.37231	0.003
	Residuals	6	1.14312	0.19052		0.69307		Residuals	5	0.87749	0.1755		0.62769	
	Total	8	1.64936			1		Total	7	1.39798			1	
Reference.South/Substrate		Df	SS	MS	F	R^2	P		Df	SS	MS	F	R^2	P
	Substrates	1	0.48253	0.48253	2.2875	0.31389	0.035	Substrates	1	0.41752	0.41752	2.0157	0.33507	0.1
	Residuals	5	1.0547	0.21094		0.68611		Residuals	4	0.82856	0.20714		0.66493	
	Total	6	1.53722			1		Total	5	1.24609			1	
Sites/Substrates (Strata=Site)		Df	SS	MS	F	R^2	P		Df	SS	MS	F	R^2	P
	Sites	2	0.5954	0.29772	1.4698	0.18798	0.006	Sites	2	0.46094	0.23047	1.2347	0.18809	0.027
	Sites:Substrates	2	0.7492	0.37458	1.8493	0.23651	0.006	Sites:Substrate	2	0.68307	0.34154	1.8297	0.27873	0.027
	Residuals	9	1.823	0.20255		0.57551		Residuals	7	1.30663	0.18666		0.53318	
	Total	13	3.1676			1		Total	11	2.45064			1	

CLR: centered log-ratio; P/A: presence/absence; Df: degrees of freedom; SS: sum of the squares; F: statistic *F-ratio*; P: probability level.

[A] based on 100 sequence re-samplings per sample to the smallest dataset (40613 sequences for Bacteria and 1835 sequences for Archaea).

Table S3. Genera differentially abundant in nodules and sediments (ALDEx2: glm adjusted p<0.01; KW adjusted p<0.05). In bold the most abundant genera (≥0.5 %) at least two times more abundant in nodule than in sediment; in italic the genera exclusively present (i.e. unique) in nodules. Base 2 logarithm of the ratios between geometric mean centred sequences number of nodule (Nod) and sediment (Sed), and average of the sequences contribution of total number of sequences (%) retrieved in nodules and in sediments are shown.

Enriched in Nodule	LOG2(Nod/Sed)	Nodule (%)	Sediment (%)
Sphingomonadaceae_unclassified	–	*0.04*	*0.00*
Filomicrobium	–	*0.01*	*0.00*
Geminicoccaceae_unclassified	4	0.12	0.01
Methyloceanibacter	4	0.17	0.02
Robiginitomaculum	4	0.09	0.00
Mesorhizobium	3	0.25	0.01
Cohaesibacter	**3**	**0.78**	**0.10**
OPB56_unclassified	3	0.03	0.00
67-14_unclassified	3	0.31	0.06
Syntrophaceae_unclassified	3	0.06	0.01
Maribacter	3	0.06	0.01
Methyloligellaceae_unclassified	**2**	**1.46**	**0.31**
Entotheonellaceae_unclassified	2	0.20	0.04
Blastocatella	2	0.18	0.04
Calorithrix	2	0.03	0.01
Hyphomicrobiaceae_unclassified	**2**	**2.72**	**0.71**
Planctomicrobium	2	0.05	0.01
Simkaniaceae_unclassified	2	0.13	0.03
Microtrichaceae_unclassified	2	0.13	0.03
LD1-PA32_unclassified	2	0.05	0.01
Subgroup 17_unclassified	**2**	**1.03**	**0.27**
JdFR-76	**2**	**0.93**	**0.26**
Subgroup 9_unclassified	**2**	**1.26**	**0.42**
Chlamydiales_unclassified	2	0.16	0.06
SAR324 clade(Marine group B)_unclassified	**2**	**3.12**	**1.10**
Vermiphilaceae_unclassified	2	0.14	0.05
Acanthopleuribacter	1	0.04	0.01
Bythopirellula	1	0.04	0.01
Nitrospina	**1**	**3.79**	**1.72**
Gemmataceae_unclassified	1	0.04	0.02
Planctomycetacia_unclassified	1	0.05	0.03
SM1A02	1	0.29	0.15
Ekhidna	1	0.17	0.09
Phycisphaeraceae_unclassified	1	0.49	0.27
AqS1	**1**	**1.10**	**0.66**
Microtrichales_unclassified	1	0.19	0.08
Pirellulaceae_unclassified	**1**	**1.31**	**0.75**
pltb-vmat-80_unclassified	1	0.05	0.00
Pir4 lineage	**1**	**1.60**	**0.91**
Alphaproteobacteria_unclassified	**1**	**7.15**	**4.44**
Babeliales_unclassified	1	0.10	0.07
Parvularculaceae_unclassified	1	0.06	0.04
PAUC43f marine benthic group_unclassified	1	0.54	0.36
Subgroup 10	0	0.75	0.67
Aquibacter	0	0.83	0.73
Cyclobacteriaceae_unclassified	0	1.76	1.70
Gemmatimonadaceae_unclassified	0	1.17	1.15
Rhodothermaceae_unclassified	0	0.38	0.39
Total		35.37	17.82

Enriched in Sediment	LOG2(Nod/Sed)	Nodule (%)	Sediment (%)
Planctomycetales_unclassified	-0.02	0.44	0.49
Lutibacter	-1	0.00	0.02
Chloroflexi_unclassified	-2	0.03	0.09
AT-s3-28_unclassified	-2	0.03	0.09
Chitinophagales_unclassified	-2	0.05	0.16
Bacteriovoracaceae_unclassified	-2	0.04	0.14
Nannocystaceae_unclassified	-2	0.02	0.07
Cellvibrionaceae_unclassified	-2	0.02	0.08
OM182 clade_unclassified	-2	0.13	0.47
Candidatus Komeilibacteria_unclassified	-2	0.01	0.03
Roseobacter clade NAC11-7 lineage	-2	0.04	0.11
Bacteroidia_unclassified	-2	0.03	0.11
IS-44	-2	0.05	0.20
Oligoflexaceae_unclassified	-2	0.01	0.08
Lentimicrobiaceae_unclassified	-2	0.01	0.04
Marinoscillum	-3	0.02	0.08
Anaerolineaceae_unclassified	-3	0.05	0.36
Colwelliaceae_unclassified	-3	0.01	0.13
Subgroup 7_unclassified	-3	0.00	0.03
Peredibacter	-3	0.01	0.05
Marinimicrobia (SAR406 clade)_unclassified	-4	0.01	0.07
Total		1.00	2.91