Comments to reviewer 1 and modifications to the manuscript.

We thank the reviewer for the constructive comments. We are pleased that that he/she liked the sequencing part of our study. The reviewer raises several concerns regarding our IPL analysis to which we would like to respond below.

1. I urge the authors to re-analyze their data using more appropriate reversed phase HPLC-MS methods. Previous studies have shown that diol column HPLC-MS and other normal phase methods lead to severe underestimation of the abundances of glycosidic GDGTs (Wörmer et al., 2013; Zhu et al., 2013). This would explain the unusually high abundances of HPH-GDGT reported in the present study and the lack of other ubiquitous compounds such as hydroxylated and unsaturated GDGT (Liu et al., 2012; Zhu et al., 2014b) and BDGT/PDGT (Becker et al., 2016; Meador et al., 2015; Zhu et al., 2014a). Alternatively, re-analysis of the samples along with authentic standards would help to correct for different ionization efficiencies based on headgroup types. If the authors did perform such a correction (and if they did not do this) they need to acknowledge this in the methods and discussion sections

   We are familiar with the reversed phase HPLC-MS method published by Wörmer et al., (2013). This is certainly an elegant addition to the lipidomics toolbox, however, we do not agree with the complete dismissal of the LC-MS method originally introduced by Sturt et al., (2004) using a diol column. We have introduced several improvements on the original method (see response to comment below for details), which have considerably improved the performance of the method. The reviewer hints at a systematic bias in the used diol column HPLC-MS method that would lead to an extreme overrepresentation of HPH-GDGTs vs. glycosidic GDGTs. Indeed, it is well known that identical core lipids with different head groups can have very different response factors. Additional drivers of response factors are structural features such as length of the core lipid, degree of unsaturation, number of rings or functionalities like hydroxylation. Van Mooy and Fredricks (2010) already published an estimation of these response differences using the diol column-MS method, based on diacylglycerols with various head groups. Interestingly, they report very similar response factors for diacyls with either a phosphoglycerol or a hexose headgroup, while a diacyl with a dihexose head group has a response factor of one third of the PG and half of the monohexose. While we acknowledge that HPH-GDGTs are likely to have a higher response factor than MHGDGTs or DH-GDGTs, these differences are unlikely to explain the 40 fold difference in apparent relative abundance we observed in e.g. the surface sample at 3000 m depth. The other IPLs mentioned, i.e. unsaturated GDGT and BDGT/PDGT core lipids with varying head groups, have been reported to be absent in several settings (Becker et al., 2016; Zhu et al., 2014b) or were linked to archaeal producers not detected in our reported archaeal community, and it is, therefore, not a given fact that we should detect them in the Arabian Sea sediments analyzed in our study. Unsaturated GDGT IPLs are not commonly detected; Zhu et al. (2014a) reported that unsaturated GDGT IPLs were not detected in sediments in the upwelling region of NW Africa and only in trace amounts in the eastern Mediterranean Sea. The absence of unsaturated GDGT IPLs in our samples is therefore not unsuspected and certainly does not prove that our analysis method is inappropriate as stated by the referee. We screened our extracts for a variety of IPL BDGTs and PDGTs with an extensive list of possible polar head group combinations as indicated in our supplementary Table S1.

However, none of these IPL compounds were detected. Becker et al, (2016) reported that BDGTs and PDGTs were only detected in Methanomassiliicoccales and not in the other studied archaeal lipidomes. Meador et al. (2015) detected MH-GDGTs in an estuarine sediment sample and linked this to a high relative abundance of MCG. However, the notable 13C depletion of the BDGTs reported by Meador et al. (2015) links this compound to relatives of members of Methanomassiliicoccales, members of Methanomassiliicoccales were not detected in our sequencing data. Becker et al. (2015) did not detect BDGTs in samples with relative low concentrations of total organic carbon (TOC), our samples have relative low TOC which can explain the absence of BDGTs. Regarding hydroxylated GDGT-IPLs, we routinely detect these in biomass as well as in environmental samples (data not yet published). We are, therefore, confident that, would they have been present, we would have detected them. The reviewer suggests to reanalyze the samples against standards to aid in quantification. We have purposefully refrained from quantifying the detected GDGT-IPLs because authentic quantitative standards truly representative of the GDGT-IPLS are not available and quantifying these compounds based on standards that have very different core lipids and much less complex head groups, in our opinion, does not lead to a valid quantification. We do use standards to continuously monitor the performance of the analytical system.

We have added text to the materials and methods section clarifying that we for correct peak area using the PAF internal standard but do not correct for any differences in response factors (line 114 – 116)

Finally, re-analysis of the samples at this time would probably lead to biased results due to the storage time of 2-3 years. In our hands, samples that are re-analyzed after such a lengthy storage period do show very different ratios for several IPL-GDGTs.

2. Judging from the supplementary spreadsheet, the authors may have considered a wider range of archaeal lipids, not exclusively traditional intact polar tetraethers, than reported in the manuscript. If so, where are these data? However, it is questionable if this comprehensive analysis is possible with the author's choice of chromatography. A more comprehensive analysis of the archaeal lipidome (e.g., by analyzing intact polar archaeols including core lipid structural modifications) in these samples would enhance source assignments and comparison to previous studies (Meador et al., 2015; Yoshinaga et al., 2015), and provide new biomarkers or source assignment for groups such as MCG and Woesearchaeota. This analysis would then satisfy the author's conclusion that the known diversity of archaeal IPLs was significantly expanded.

For our analysis, we made use of high resolution MS with a mass range of 375- 2000 m/z. Our analytical routine involves a data dependent analysis of the 10 most abundant masses in each MS1 spectrum with a dynamic exclusion window of 6 sec (we have added this information to the method section), maximizing our capacity to identify lipids. In addition, our method works with an inclusion list, where we specify masses of compounds that, if detected, will be targeted for a product spectrum. This inclusion list contains the exact mass of most GDGT and archaeol-IPLs known at the time of analysis of this sample set. We have used all this available data to search for the compounds in the inclusion list and for additional unknown IPLs with a GDGT or archaeol core. We have successfully detected additional GDGTs as well as archaeol-based IPLs in other environmental sample sets and archaeal biomass (data not yet published). We are of the opinion that there is also sometimes information in

the absence of a certain set of compounds and therefore we reported our inclusion list to indicate the breadth of our search. In contrast to what the referee suggests, we do not conclude that we are expanding the known diversity of archaeal IPLs. We wrote that we expanded the screening for IPLs in our lab based on the use of UHPLC chromatography and high resolution mass spectrometry (quadrupole-orbitrap hybrid). We adjusted the manuscript to make this point more clear (line 444-447).

3. Further, because the authors used a column different from the one described in their method reference, the suitability of the chromatography for IPL analysis cannot be assessed. The authors should provide annotated TIC or EIC traces for representative samples in supplementary figures. Importantly, the type of chromatography used here does not provide chromatographic separation of IPLs by core lipid type (e.g., GDGT-0, GDGT-1), but the way the authors present their data implies the contrary. If the different core lipid types of each IPL co-eluted in their analyses, the relative abundances reported would be questionable. At the very least the authors would need to do an isotope peak correction and then report in detail how this was done (correction factors for each IPL-core lipid combination). It would be much better to use a method than can actually chromatographically separate IPL-GDGTs by core lipid structure (Zhu et al., 2013). Additionally, the method used for this study likely cannot distinguish between the core lipid crenarchaeol and it's regioisomer. However, crenarchaeol more abundant than the regioisomer in MH GDGT and HPH GDGT but lower than its regioisomer in DH GDGT in thaumarchaeal cultures (Elling et al., 2017). This limitation needs to be addressed either by using a different chromatographic method or at least by discussing this issue in the manuscript

   Indeed, the column used for the work described here is a different column from the original method. It is a UHPLC silica based diol column used with the same solvent system as originally described by Sturt et al., (2004) and the overall chromatographic patterns have thus not changed but the resolution of the chromatography has significantly improved. We added information on the improved chromatography to the supplemental information. Due to the improved separation power of the column, we can separate GDGT-IPLs with different cores but the same head group. The suggested isotopic corrections are, therefore, not necessary. The reviewer is correct that we cannot separate crenarchaeol from crenarchaeol regio-isomer with identical head groups. We are aware that Zhu et al. (2013) reported a separation between IPL crenarchaeol and crenarchaeol regio-isomer with the same head group. With our method this is not feasible. We have added a statement to the material and method section that IPL crenarchaeol is the sum of crenarchaeol and crenarchaeol regio-isomer (line 116-120). We stress that for the purpose of our paper this separation is not required. Lengger et al. (2012) reported core lipid content in these samples and did not find dominance of crenarchaeol regio-isomer over crenarchaeol. We, therefore, do not think that the culture study the reviewer refers to is relevant in this context. We have added EIC traces for MH-GDGTs, DH GDGTs and HPH-GDGTs to the supplementary data to clarify our peak separation.

4. Given these fundamental limitations, major parts of the methods, results, and discussion sections need to be revised. However, HPLC-MS (re-)analysis with a different method could be achieved within a day or two, given that the current study only includes 8 samples. I hope that the authors will agree that this approach has the potential to greatly expand the significance of their work with only moderate additional effort.

As discussed above, we do not agree with the dismissal of the diol column based HPLC-MS method. As explained above, we feel that reanalysis of the samples at this time with the reversed phase method would not lead to valid results due to the age of the extracts. We respectfully disagree with the reviewer that our results, and therefore our interpretation and discussion, are invalid.

5. Other comments: Line 27: DeLong et al. report only on Antarctic samples. You may want to choose a reference that discusses a more diverse set of environments (e.g. DeLong and Pace, 2013; Schleper et al., 2005) or a collection of references e.g., (DeLong, 1992; Fuhrman et al., 1992; Teske and Sørensen, 2008).

The addition of literature would benefit the introduction, we have added references to the manuscript as suggested by the reviewer. We assumed that the reviewer referred to the paper of DeLong and Pace (2001). (line 27-28)

6. Line 31: Lloyd et al. only provide data on two archaeal groups. You may rather cite a paper that actually discusses archaeal diversity, such as Teske and Sørensen (2008) or Teske (2013) some of the more recent literature, e.g., Hug et al. (2016) or Spang et al. (2017).

We modified the manuscript and added some of the proposed literature references. However, we feel that it is also important to cite original references where pioneering efforts have been described. (line 32-33)

7. Line 40-42: What are the metabolisms of the other archaeal groups?

We mention the metabolism of detected groups in our discussion, however for most of the detected groups metabolism is unknown as neither cultures nor metagenomes are available to predict their metabolic capabilities. For the rest i.e. MBG-B and E, potential metabolism is derived from (meta)genome information as shown already in lines 334-336.

8. Line 59-61: What about the Lincoln et al. (Lincoln et al., 2014) paper?

Lincoln et al. (2014) indeed reported the marine group II as potential producers of crenarchaeol. However, there is still an ongoing discussion on this hypothesis (Lincoln et al., 2014b; Schouten et al., 2014). We added this reference and discussion into our introduction (lines 62-64).

9. Line 86: Why was this standard added? Was it used for any correction?

This standard (PAF; 1-O-hexadecyl-2-acetoyl-sn-glycero-3-phosphocholine) was added to continuously monitor the performance of the MS response and to correct for matrix effects. We also corrected peak areas accordingly. This has been clarified in the materials and methods section. (line 114-116)

10. Line 90: Specify the modifications.

The modifications are already described in detail in the text after the referral to the original Sturt et al. (2004) method, lines 93-120. We have added more details as mentioned above in the response to previous comments

11. Line 159-161: Two DH isomers were also reported earlier by Elling et al. (2014; 2017).

Indeed, Elling et al. (2014) reported an "early eluting " DH-GDGT and Elling et al. (2017) reports early eluting isomers of DH- and DH-OH-GDGTs. We changed our manuscript to acknowledge these reports. However, although Elling et al. (2017) speculates on the nature of the isomerisation, no spectral details are given. We show, for the first time, spectral data on these isomers and based on these spectra, interpret the nature of the isomerization. We, therefore, feel it is worthwhile to include this information in the supplementary information.

12. Line 166-171: Did you detect crenarchaeol regioisomer? If not, why not? Coelution with crenarchaeol? This should be pointed out here or elsewhere.

    As stated above, the crenarchaeol regio-isomer IPLs co-elude with crenarchaeol with the same head group. We have added a statement regarding this to our material and method section. (line 116-120)

13. Line 172-173: The way the % values are used in this section is very confusing (. . .37.6% of 36.6%) etc. It could help to provide a further table in addition to Table 2 that shows the samples arranged by depth or headgroup type instead of sorting by core GDGT type.

    We clarified this in the manuscript to make our data more clear. We also refer to the original Table 3 in our manuscript where the data are sorted according to the different headgroup types per sampling depth.

14. Line 196: What does "p=1.00" represent? The correlation between our IPL derived CL-GDGTs and the ones reported by Lengger et al., (2012).

    The noted correlation is adjusted in the manuscript. We also added some sentences to the material and method section in order to clarify our goal with the analysis. (material and method section: line 121-125, corrected correlation: line 213-215)

15. Line 254-260: How were the transcripts analyzed? I did not find this information in the methods section. How long were the samples stored before analysis? The transcripts were analysed as described in Pitcher et al., (2011).

    We refer to this in our material and method section. As mentioned, the sediments were sliced on board and directly stored at -80C. They were kept at -80C until lipid and DNA/RNA extractions. After the extractions the DNA/RNA and the cDNA were stored at -80C until analyses.

16. Line 270-274: This sentence is a bit long and complex. Revise?

    The referee is correct, we have adjusted this in the manuscript. (line 290-294)

17. Line 281: "Differences"?

    The referee is correct, it should be differences instead of difference. We altered this. (line 302)

18. Line 297: Rather "overlying"?

    The referee is correct, it should be overlying. We altered this. (line 318)

19. Line 319-321: Could you point out whether there is any evidence for the occurrence of Woesearchaeota in the water column and whether you would expect this group and the other archaeal groups to be present in oxic vs. anoxic environments (or both).

There is evidence for Woesearchaeota in marine pelagic oxygenated environments, (Liu et al., 2017). The Arabian Sea contains an oxic environment and an oxygen minimum zone, so far it is uncertain if the Woesearchaeota could reside in the anoxic pelagic zone. Woesearchaeota were recently also detected, in relative high abundances, in oxygenated surface sediments (Lipsewers et al., 2017)

20. Line 353-357: I disagree with the statement that the diversity of detected (detectable) IPLs was greatly enhanced by this study. There are two IPL types described in addition to those reported by (Lengger et al., 2012). However, the diversity of IPLGDGTs reported by previous studies is much higher, e.g. Yoshinaga et al. (2015), with respect to both the headgroup types as well as the structural modifications in the core lipid such as hydroxylation, unsaturation, methylation, monoalkylation, trialkylation, or substitution of glycerol with butanetriol or pentanetriol. Many of these would likely be detectable in the presented samples using different methods.

In contrast to what the referee suggests, we do not conclude that we are expanding the known diversity of archaeal IPLs (see remark on point 2). However, we adjusted the sentences in order to be more clear about our findings. (line 445-447)

21. Line 359-360: GDGT-0 and crenarchaeol are also dominant core lipids of MHGDGT in Thaumarchaeota (Elling et al., 2014; Elling et al., 2015; Elling et al., 2017; Schouten et al., 2008). MH-GDGT also appears to be the dominant membrane lipid in many Thaumarchaeota. Also, how does your interpretation fit to the results of Lincoln et al. (2014) who suggested production of crenarchaeol by other archaeal groups?

We also see a high relative abundance of IPL GDGT-0 and IPL crenarchaeol in the surface sediment samples coinciding with a high relative abundance of thaumarchaeota. Lincoln et al. (2014) suggested that marine group II (MGII) may be significant contributors of crenarchaeol in oceanic surface waters based on the combination of core lipid (CL)-GDGT and metagenomic analyses. However, the lack of culture representatives of this group and the difference in the resilience times between CL-GDGTs (attributed to dead biomass) and DNA weakens the arguments of MGII as important GDGTs producers (Lincoln et al., 2014b; Schouten et al., 2014). Members of the MGII are, so far, only found in the pelagic and not in the benthic environment therefore we did not discuss them in our manuscript. The archaeal groups that we detected are, currently, lacking cultured representatives and therefore their membrane lipid composition is still uncertain.

Second comments to reviewer 1 and modifications to the manuscript.

We thank the reviewer again for the additional comments. Here are our replies.

1. I agree with many of the modifications proposed by the authors. Most importantly, since they use an unpublished method, they will need to demonstrate in detail the separation of

the different IPL-GDGT cyclized GDGT-headgroup combinations, and exclude, e.g. the co-elution of unsaturated GDGTs with regular GDGTs.

As stated in our previous reply to comments regarding the use of a modifications to the original normal phase IPL method, we do not agree with the reviewer that this represents an "unpublished method". The basis of the method is unchanged and thus we do not feel the functioning of the method deserves a lengthy discussion in a paper that focusses on different issues. As indicated before, we illustrated the improvements made in the method by adding illustrative figures in the supplemental information. We cannot fully exclude the co-elution of unsaturated GDGTs-IPLs with regular GDGT-IPLs. Per suggestion of the reviewer we re-analyzed the anoxic sediment samples (surface at 885 mbsl, and subsurface of all the stations) using a reversed phase LC/MS method to estimate the contribution of unsaturated core GDGTs to the GDGT-CL pool. As expected after a couple of years of storage the IPLs were partially decayed and could therefore not be reinterpreted. We did observe a trace of potentially unsaturated CL-GDGTs (e.g. GDGT-0:1) but these represented <0.25% of the total CL pool.

We have added figures (Supplemental Figs. 3a, b and c) to our supplementary information to illustrate our used method.

2. I appreciate that determining response factors for GDGTs is very challenging due to the lack of commercial standards but I think that this would have been perfectly feasible for the small number of compounds the authors have analyzed. I urge the authors to clearly state that they have not used response factors and how this affects interpretation of the results. It is not acceptable to cite the response factors from Van Mooy & Fredricks (2010) since they were measured on a different chromatographic method and different instruments and on compounds that are not comparable to GDGT IPLs. GDGT-IPLs are twice the size of the standards measured in that paper and thus the relative ionization efficiencies will be dramatically different (this will also change the relative effect of one versus two glycosidic moieties). Further, ionization of GDGT-IPLs will heavily depend on elution time, since the method chosen by the authors uses a strong polarity gradient. This will strongly affect the efficiency of ionization of GDGTs relative to their retention time and will likely lead to a systematic underestimation of more apolar GDGT-IPLs (e.g. MH) that elute earlier when the mobile phase consists mostly of hexane (which is not beneficial for ionization).

We do not agree with the referee that isolation of standards to the purity required for use as standards (i.e. >95%) to determine response factors is "perfectly feasible". In fact, it is a large project in itself that requires a large amount of source material (hard to culture archaea), many months of experimental work and large amounts of solvents to do the purification as well as proofing of the material by NMR and mass spectrometry. Given the fact that isolated IPLs are also unstable, and the fact that absolute quantitation does not contribute to the discussion (as we focus on relative distributions in comparisons with the genetic fingerprint of the sediments), we do not feel the effort was or is justified. We have already indicated in our response to the reviewer's earlier comments that we clarified in our materials and methods that we do not use response factors to correct peak areas. We therefore also do not plan to cite the Van Mooy & Fredricks (2010), or apply their published response factors, but merely used this work to illustrate what the impact of the difference in response factors might be.

We have assessed the difference between the HPLC-MS method using a diol-column used for the study here and the one that uses a reversed phase LC column (Wörmer et al., 2013) using a fresh sample of North Atlantic SPM. Looking at the ratio of MH:DH:HPH for cren and GDGT-0, the NP IPL method may underestimate the MH IPLs by a factor of 10. However, even when taking this into account, this does not change the conclusion of this study. We have focused our discussions on the HPH variety of the IPLs as this is the best life marker among the IPLs. MH-GDGTs are certainly produced directly by archaea but can also have a sizable fossil contribution and are decay products of DH- and HPH-GDGTs, especially in stored extracts.

For the anoxic surface sediment at 885 mbsl we observed a discrepancy between the relative abundance of IPL derived CL-GDGT-0 and the IPL-GDGT-0 relative abundance reported in our manuscript. Even if our method underestimates the relative abundance of MHs, this discrepancy cannot be fully explained. We, therefore, remained this part in our discussion. However, we also highlight the fact that we underestimate the relative abundance of MHs in our study.

3. Lastly, the authors should to assess why the fractional abundances of IPLs are so different between the current study and their earlier study on the same samples (Lengger et al., 2012)

Lengger et al. (2012) studied only IPLs with a crenarchaeol core using an SRM-MS method, therefore in fact quantifying the fragments and not the parent molecules. Lengger et al. (2014) isolated various IPLs and quantified core GDGTs after hydrolysis. These methods are very different from the more direct detection method used in the present study where signals are assessed in the MS1 signal. It should also be noted that Lengger et al. (2014) studied at a different sampling resolution (surface sediment used was 0-2 compared to 0-0.5 cm in our case). It is, therefore, not unexpected that results may vary. However, the distribution of IPLs described in our manuscript, when excluding the MH IPLS (for the reasons discussed above), actually resembles the data shown in the 2014 paper very closely. We added some sentences to our discussion comparing the data published by Lengger et al. (2012; 2014) with our data. We also highlighted that differences were likely resulted from ionization efficiencies. (line 378-380, 409-410)

Comments to Dr. Lipp (reviewer 2) and modifications to the manuscript.

We thank Dr. Lipp for his comments. We have responded to these below. Many comments are a reiteration of those made by reviewer 1. In those cases, we referred back to our earlier replies made to reviewer 1.

1. The method has not been described before (with this choice of column) and it is not completely clear how the new results compare to previous installments. Most importantly, it is not clear which compounds are captured in the analytical window as according to the authors some important GDGT-0-based lipid which can be found during APCI analysis of hydrolysate cannot be found during IPL analysis. Also, the comparison of relative proportions of lipids between samples in the current way is misleading. The ionization behavior of lipids varies strongly according to their structure and accordingly using the sum of peak areas for "total lipid abundance" and relative abundance calculated from this value is problematic.

Since the authors have not used standards (and do not plan to) for correction of response factors, its perhaps best to compare absolute peak areas and report them in a table. The methodological shortcomings need to be addressed in a revised version (see also comments below).

As stated in the replies to reviewer 1, we do not agree with the reviewer that we used an undescribed method. All compounds captured with the previous type of diol column (standard analytical HPLC size) are also captured using the UHPLC version of the diol column. We also would like to add that no analytical method is fully comprehensive. There will always be compounds that will fall out of the analytical method due to polarity, molecular weight or other properties. The notion that there may be a source of GDGT-0 that is perhaps outside of the current analytical window is, therefore, not a sign that the method does not work, but simply an acknowledgement of the fact that, despite the efforts of the IPL community, we might not know everything yet. We, therefore, feel it is important to retain a discussion on this observation in the manuscript but will remove the statements regarding this "missing" GDGT-0 from the conclusion. With regards to the comparison of IPLs in relative abundances, we feel it is a fair representation of the observed peak areas and their differing contributions to the overall lipid profile to create a "summed area" and then calculate the relative contribution (expressed as a percentage) of a peak to this summed area. We don't see the difference with what the referee proposes (absolute peak areas) except that the numbers obtained are more easy to handle. As this normalization is done per depth (necessary to compare the lipid profile to the DNA profile), we agree that this way the differences in abundance between depths is unclear.

We added an additional column to table 2, indicating the total IPL peak area per sample.

2. The data only compare poorly to previous results reported in Lengger et al. 2014 where absolute quantities and relative proportions of MH, DH, and HPH-GDGTs were reported. In this study, the hydrolysis method with subsequent APCI quantification was used and problematic response factors are avoided. However, the numbers just do not add up (see detailed comments). I am very surprised that this study is not included in the discussion at all. This also needs to be addressed thoroughly in a revised version.

As we state in our previous reply to reviewer 1, we acknowledge that MH has a reduced response in the NP IPL method compared to the RP IPL method. However when compared with the study of Lengger et al. (2014), we do see a good correspondence between the datasets when it comes to the DH and HPH IPLs. Again, it should be noted that Lengger et al. (2014) used a different sampling resolution (surface sediment used was 0-2 compared to our 0-0.5 cm). It is, therefore, not unexpected that results may vary. We have added a large part in our discussion to compare the differences IPL distributions between our study and the previously published studies of Lengger et al., (2012; 2014). In this section we also raised the issue of varying ionization efficiencies among IPLs. (line 378-380, 409-410)

3. Comment to author's reply to response factors raised by reviewer 1: The study by Van Mooy and Fredricks (2010, GCA) explicitly states that ". . .these RFs are not applicable for use in any future quantitative analyses of IP-DAGs using a ThermoFinnigan LCQ Deca XP ion-trap mass spectrometer or any other mass spectrometer. . .". I urge that no conclusions should be drawn from these data. It is also speculative that a 40-fold different is unlikely, as the two ion sources are different (heated vs unheated ESI) and the technology of the two mass

spectrometers is also different (ion trap vs orbitrap). We probably all can agree that quantification with standards analyzed on the same machine as the environmental samples would be more appropriate here. I am more worried that the quantitative data here does not match previous data from the same samples (Lengger et al., 2014, see comments below). This hints to some methodological bias and supports suspicions from reviewer 1. These issues should be discussed.

As we state in our reply to the second comment of reviewer 1 we merely used this reference to discuss the possible differences in response factors between IPLs. We do not intend to use it in the revised manuscript. The ion source used in this experiment is basically the same ion source that we have used for all our previous IPL work with the difference that the ESI probe now has a heating option. The heater in this experiment was set to a minimal T of 50 °C. Without active heating the source already reaches a temperature of ca. 40 C due to passive heating from the MS. However, this heating is dependent on ambient temperature and therefore fluctuates. The applied temperature of 50 °C is merely applied to stabilize this temperature and any effect it could have on ionization to make the method more reproducible but does not affect response or ionization behavior.

Detailed comments:

4. Line 70: typo "repertoire ".

   Modified in the manuscript (line 73)

5. Line 86: was the PAF standard used somehow for quantification?

   The response of the PAF standard was used to normalize for matrix effects and shifts in performance of the MS. The reported peak areas are after this correction. We have added a couple of sentences on the use of PAF as internal standard and the correction. (line 114-117)

6. Lines 93-98: this is a novel LCMS method that has not been published before. As reviewer 1 noted more information would be desirable in this case (the best way would be to report the new method in another peer-reviewed manuscript). It seems that the individual GDGTs with rings can be separated (judging from the supplemental figure), here some more mass chromatograms as supplemental figures would indeed be good for illustration. What other compounds can be separated, are the BDGTs/other methylated GDGTs, unsaturated GDGTs, hydroxylated GDGTs in the analytical window? Especially OMZ sediments should have abundant unsaturated GDGTs (Zhu et al., 2014, RCM). How do quantitative results compare to other published methods? Is the method suitable to comprehensively capture the archaeal lipidome?

   This comment is a repeat of earlier comments by reviewer 1 and we therefore refer to our rebuttal to reviewer 1.

7. Line 101: the unit of resolution is not ppm. I also suggest to use "resolving power" instead of "resolution" as it is a better term.

   We changed “resolution” into “resolving power” in our manuscript. (line 104-106)

8. Line 166: according to Tab 2 it is 44.7% - I am sure this is due to rounding. However it would be great to have consistent numbers in the table and the text. Other values are also different in text and table, please check.

We have checked the manuscript and changed the numbers accordingly.

9. Lines 156-196: the "relative abundance" is given as proportion of peak area. This should at least be reported and the shortcomings of not using authentic standards need to be discussed thoroughly (see also major comments and line 381-397). Perhaps add absolute peak area values to the table?

As discussed above in our reply to Dr. Lipp's general comment, we have modified this in our manuscript. We added a statement in the revised manuscript that no absolute quantifications were made and that the data cannot be interpreted as such.

10. Lines 194-196: Consistency with Lengger et al. 2012 study: please show the data of the comparison (table?). The Greek letter rho should be used for Pearson correlation coefficients - I assume that is what has been calculated here. Please add information on what is compared and how it is calculated. The letter "p" is usually used for statistical significance, p=1 would be really bad.

We do not agree with the reviewer that adding the Lengger et al. data would really add additional information to the manuscript. We analyzed the anoxic surface sediment sample at 885 mbsl with the same method as Lengger et al. (2012), where they studied IPL-derived GDGTs in similar sediment samples. We described this in our material and method section, lines 121-125. We separated the IPL fraction from the core lipids with the use of a silica column and flushing with MeOH. This IPL fraction was hydrolyzed for 3 h. This was done to determine if the IPL-derived CL-GDGT distribution was altered due to degradation during storage. We compared the distribution of IPL derived GDGTs published by Lengger et al., 2012 to our data and found a significant correlation (ρ = < 0.001) between the two analyses. We changed the notation of the statistical significance into: "(r= 0.99, ρ = < 0.001)". We have added a couple of sentences to the material and method section to clarify the aim of this analysis. (line 121-125)

11. Line 360: add Elling et al. 2014, 2015, 2017 and Schouten et al. 2008 references for more complete IPL inventory of Thaumarchaeota.

We have added some of the suggested references to the manuscript. It is undoable to list for every statement a complete list of references. We selected the ones that are most original or appropriate. (line 383)

12. Line 362-363 and 367-369: the cited studies have not looked at stability of HPH-GDGTs. DH-GDGT stability also has not been experimentally assessed by the cited Lengger et al. 2012 and 2014 studies. Please discuss the stability of phosho vs glycolipids (and possibly ester vs ether lipids as no study has compared purely phospho vs glyco ETHER lipids, cf. Logemann et al. 2010) in a more balanced way and refrain from speculation without evidence; (additional) useful references in this context are Lipp and Hinrichs 2009, Logemann et al. 2010, Schouten et al. 2010, Xie et al 2013.

We acknowledge that Harvey et al. (1986) and Schouten et al. (2010) did not study the degradation of HPH-GDGT but sedimentary phospholipids in general. We clarified this in the manuscript. Lengger et al. (2012) and (2014) did observe that the abundance of HPH-

GDGTs decreased with increasing sediment depth in contrast to the MH- and the DH-GDGTs that remained equal or even increased in abundance with increasing sediment depth. This was interpreted to reveal the less stable nature of HPH-GDGTs. We altered the manuscript concerning the references Harvey et al, (1986) and Schouten et al. (2010), the studies describing the labile nature of sedimentary phospholipids. We assumed that the reviewer was referring to Logemann et al. (2011), we have added this reference and Xie et al. (2013) into our discussion. We also raised the topic of phospho vs glyco ether lipids in our discussion. (line 393-396)

13. Line 371: What about the possibility of other archaeal sources for crenarchaeol, e.g. Lincoln et al. 2014?

Lincoln et al. (2014) describes the possibility of the (archaeal) Marine Group II as a potential source for crenarchaeol. We did not detect members of the Marine Group II in our sequencing data, therefore, we do not find it relevant to discuss this here.

14. Line 381-397: regarding the unknown IPL type for GDGT-0, what other compounds might have been missed? It seems like potentially a major proportion! There is a lot of speculation regarding the source of an undetected GDGT-0, but how sure can you be that there is no methodological problem with the method, especially as it seems to be used for the first time? Lengger et al. 2014 have done semi-preparative IPL separation into head group classes and found abundant GDGT-0 connected to MH, DH and HPH headgroups (Fig. 5 and Table A3 in the supplemental material, here station P900 0-2 cm). Comparison of these values and Tab. 2 for the 885 mbsl surface sample shows major differences and multiple values do not match: e.g. DH-GDGT-0 and DH-GDGTcren are 63.7 and (144+86.5=230.5; incl. regioisomer) ng/g sed, respectively, a ratio of ~1:3. According to Tab 2 in the current manuscript the corresponding values are "ND" and 43.1 %. Why is the DH-GDGT-0 not detected here, assuming a ratio of 1:3 it should have roughly 10% contribution? Another example is found in Tab 3, it seems that the 885 mbsl surface sample is dominated by 98.1% of DH lipids. However, Lengger et al. 2014 report much higher MH abundances than DH (roughly 800 ng/g vs 400 ng/g). What is the authors' explanation for this major difference? Please discuss all data in comparison to Lengger et al. 2014. Is the diol column method not sensitive enough to capture what had been seen with the prep-hydrolysis-APCI method before? Are some compounds not detected with similar efficiency as suspected by reviewer 1? Can the discrepancies be due to that what the authors report as "relative abundance" is in fact the relative proportion of total peak area for compounds which are known to behave dramatically different during ionization and cannot be simply summed together? A meaningful comparison is only possible with standards and correction of response factors. Again, this comparison is important and needs to be discussed, especially before invoking unidentified and undetected IPL headgroup types for GDGT-0 which are somehow not in the analytical window of a new analytical method which has not been previously published.

We addressed the comments raised on the analytical issues in previous replies. We would like to point out that Lengger et al. (2012) and (2014) performed an indirect quantification of the GDGT-IPLs. Since this previous work used different approaches than we applied in the current study, it is not unexpected to find some difference. MH-IPL is also a degradation product of other IPLs and a portion of the differences may simply be due to differences in sample handling, which we keep to an absolute minimum. Additionally, and as stated earlier, we acknowledge the underestimation of MH-GDGTs but as stated above, we would like to retain a paragraph of discussion on this. We had a detailed look into our

analyses with regards to the discrepancy between the DH-GDGT-0 abundance between our study and Lengger et al. (2014). However, we truly did not detect any DH-GDGT-0 within our sample, the surface sediment at 885 mbsl, whereas our method is clearly able to do so. This discrepancy could be explained by resolution differences in these two studies. Lengger et al., 2014 analyzed the top 0-2 cm, while we analyzed the surface with a resolution of 0-0.5 cm. It is, therefore, not unexpected that results may vary. (line 378-380, 401-410)

15. Line 384: is the reference Lengger et al. 2014 correct? Or do the authors want to cite the 2012 study? As explained above, the 2014 manuscript is probably as important as the 2012 study.

    Lengger et al. (2014) is a valid study on the same suite of samples but uses an entirely different method. Also the sampling resolution in the 2014 paper is different compared to the current study. Nevertheless, we see the point of the referee and we compared the Lengger et al (2014) data more closely with our results in the revised discussion. (line 405-410)

16. Line 403: A known source are the Sulfolobales. Add to discussion, see also line 414- 417 comment.

    This is mentioned further on in the manuscript as indicated by the reviewer (lines 414-416): "GDGT-0 with a cyclopentanetetraol head group has been previously detected in cultures of the hyperthermophilic crenarchaeal Sulfolobales (Langworthy et al., 1974; Sturt et al., 2004)." We did not find any evidence in our sequencing data for the presence of these hyperthermophilic archaea. We have added to the discussion that we found no evidence of Sulfolobales in our samples. (line 436)

17. Line 409: is there evidence for selective preservation of water column GDGTs at this station? This would require a dramatic degradation of GDGT-0 which is abundant in SPM within the OMZ (Line 404). It seems more likely that the lipids are produced within the sediments. Rearrange.

    We compared the anoxic surface sediments with the overlying oxygenated surface sediments. We observe a shift in IPL-GDGT distributions, with lower GDGT-0 and higher GDGT-1, 2, 3 and 4 relative abundances within the subsurface compared to the surface sediments. This could be due to in-situ production with a preference towards IPL-GDGT-1, 2, 3 and 4 or selective preservation. This could be a factor because the majority of IPL-GDGT-0 within the surface sediments consisted with the headgroup HPH whereas the other IPL-GDGTs where dominantly detected with DHs. This is discussed (line 425-429).

18. Line 410-412: the DH isomers have been reported previously in Elling et al. 2014 and 2017. Elling et al. 2017 studied the lipidome of several thaumarchaeal cultures and found isomers not only for DH-GDGTs but also for hydroxylated DH-GDGTs. Also, they found these structures not only for the acyclic structure (GDGT-0) but also for core lipids with more rings. Please add references to prior studies and discuss why only GDGT-0 has been found in this study.

    We have altered the manuscript as also indicated in the comment to reviewer 1 (first comments, point 11)

19. Line 413-414: which moieties? And what cannot be compared? Please clarify.

We are refereeing to the DH moieties, as mentioned in the previous sentences. We could not compare this because of the missing archaeal community compositions in these depths. We have altered this in the manuscript.

20. Line 414-417: the GDGT-0 with a cyclopentanetetraol (formerly called GDNT) has not only been found in pure cultures of Sulfolobales but was also detected in sediments (Sturt et al., 2014; Lipp and Hinrichs, 2009). The sources in the present study are likely the same as in these two studies. I assume the authors mean "microbial sources" and not "sources" (Line 416). Please rephrase. Have Sulfolobales been found in the molecular biological data?

    We have altered this in our manuscript, see earlier comment by this reviewer (point 16) (line 436)

21. Lines 418-437: I do not think there should be references in the conclusions section. If you want to have references here, add all relevant ones (see below and reviewer 1 comments).

    We agree with the reviewer; we removed the references from the conclusion section.

22. Line 420: what does "specially" mean here? I would rephrase to ". . .we have unraveled the high diversity of benthic archaea harbored in anoxic sediments of the Arabian Sea, as well as. . .". Remove "specially" and add "Arabian Sea".

    With specially we referred to the high diversity of Archaeal groups in the anoxic sediments compared to the oxygenated sediments. We rephrased the sentence according to the reviewers suggestion. (line 445-447)

23. Line 421: "increasing the repertoire of archaeal intact polar lipids detected" sounds as if many new archaeal IPLs were found. However, all the described lipids (and many many more that have been found in similar environmental samples) have already been described in the literature (see also reviewer 1 comments). Please rephrase or remove statement.

    We agree with the reviewer on the ambiguity of the sentence. We altered the sentence to clarify the broadening of our analytical window and not the detection of novel IPLs. The focus of our manuscript is not the identification of new archaeal IPLs but to provide more information on their potential sources by a comparison with genetic data.

24. Line 424: add Elling et al. 2014 reference.

    We have added the reference to our manuscript, but removed the references from the conclusion section as suggested by the reviewer (point 21).

25. Line 426: change "important" to "abundant".

    We removed this part of the conclusion, because this section contained speculation of an unknown IPL-GDGT-0 source. We removed this as discussed in earlier comments.

26. Line 429: rephrase to ". . .which could either be attributed to a fossil signal. . . or being IPLs synthesized. . ." (add "either").

    We changed the manuscript according to the reviewers suggestion. (line 455)

27. Line 431: why are these GDGT-0 derivatives "unusual"? They have been described in the literature before (e.g. Elling et al. 2014, 2017 for DH-GDGT isomers, HCP (then labelled as GDNT) in Sturt et al. 2004 and Lipp and Hinrichs, 2009).

We altered the manuscript accordingly. (line 456-457)

28. Line 437: this implies the authors have only assumed a very low diversity before they analyzed the samples. Why is that? It is known that there is a large variety of IPLs in the environment (see all the details that reviewer 1 has posted in his/her comments), what was the rationale for a low diversity in these samples? I suggest removing this statement as it is not relevant what the authors have assumed.

    We removed the statement and altered the sentence. (line 463)

29. Table 2/3: use consistent names "subsurface/deep".

    We altered this in our manuscript.

30. Fig 1: add "10-12 cm" to "subsurface", also in tables and other figures.

    We added this to our figures and figure legends.

References

Becker, K. W., Lipp, J. S., Versteegh, G. J. M., Wörmer, L. and Hinrichs, K.-U.: Rapid and simultaneous analysis of three molecular sea surface temperature proxies and application to sediments from the Sea of Marmara, Org. Geochem., 85, 42–53, doi:10.1016/j.orggeochem.2015.04.008, 2015.

Becker, K. W., Elling, F. J., Yoshinaga, M. Y., Söllinger, A., Urich, T. and Hinrichs, K. U.: Unusual butane- and pentanetriol-based tetraether lipids in Methanomassiliicoccus Luminyensis, a representative of the seventh order of methanogens, Appl. Environ. Microbiol., 82(15), 4505–4516, doi:10.1128/AEM.00772-16, 2016.

Delong, E. F. and Pace, N. R.: Environmental Diversity of Bacteria and Archaea, Soc. Syst. Biol., 50(4), 470–478, doi:10.1080/10635150118513, 2001.

Elling, F. J., Könneke, M., Nicol, G. W., Stieglmeier, M., Bayer, B., Spieck, E., de la Torre, J. R., Becker, K. W., Thomm, M., Prosser, J. I., Herndl, G. J., Schleper, C. and Hinrichs, K. U.: Chemotaxonomic characterisation of the thaumarchaeal lipidome, Environ. Microbiol., 19(7), 2681–2700, doi:10.1111/1462-2920.13759, 2017.

Lengger, S. K., Hopmans, E. C., Reichart, G.-J., Nierop, K. G. J., Sinninghe Damsté, J. S. and Schouten, S.: Intact polar and core glycerol dibiphytanyl glycerol tetraether lipids in the Arabian Sea oxygen minimum zone. Part II: Selective preservation and degradation in sediments and consequences for the TEX86, Geochim. Cosmochim. Acta, 98, 244–258, doi:10.1016/j.gca.2012.05.003, 2012.

Lengger, S. K., Hopmans, E. C., Sinninghe Damsté, J. S. and Schouten, S.: Impact of sedimentary degradation and deep water column production on GDGT abundance and distribution in surface sediments in the Arabian Sea: Implications for the TEX86 paleothermometer, Geochim. Cosmochim. Acta, 142, 386–399, doi:10.1016/j.gca.2014.07.013, 2014.

Lipsewers, Y. A., Hopmans, E. C., Sinninghe Damsté, J. S. and Villanueva, L.: Potential recycling of thaumarchaeotal lipids by DPANN Archaea in seasonally hypoxic surface marine sediments, Org. Geochem., doi:10.1016/j.orggeochem.2017.12.007, 2017.

Logemann, J., Graue, J., Köster, J., Engelen, B., Rullkötter, J. and Cypionka, H.: A laboratory experiment of intact polar lipid degradation in sandy sediments, Biogeosciences, 8(9), 2547–2560, doi:10.5194/bg-8-2547-2011, 2011.

Meador, T. B., Bowles, M., Lazar, C. S., Zhu, C., Teske, A. and Hinrichs, K.-U.: The archaeal lipidome in estuarine sediment dominated by members of the Miscellaneous Crenarchaeotal Group, Environ.

Microbiol., 17(7), 2441–2458, doi:10.1111/1462-2920.12716, 2015.

Van Mooy, B. A. S. and Fredricks, H. F.: Bacterial and eukaryotic intact polar lipids in the eastern subtropical South Pacific: Water-column distribution, planktonic sources, and fatty acid composition, Geochim. Cosmochim. Acta, 74(22), 6499–6516, doi:10.1016/j.gca.2010.08.026, 2010.

Sturt, H. F., Summons, R. E., Smith, K., Elvert, M. and Hinrichs, K.-U.: Intact polar membrane lipids in prokaryotes and sediments deciphered by high-performance liquid chromatography/electrospray ionization multistage mass spectrometry—new biomarkers for biogeochemistry and microbial ecology, Rapid Commun. Mass Spectrom., 18(6), 617–628, doi:10.1002/rcm.1378, 2004.

Wörmer, L., Lipp, J. S., Schröder, J. M. and Hinrichs, K. U.: Application of two new LC-ESI-MS methods for improved detection of intact polar lipids (IPLs) in environmental samples, Org. Geochem., 59, 10–21, doi:10.1016/j.orggeochem.2013.03.004, 2013.

Xie, S., Lipp, J. S., Wegener, G., Ferdelman, T. G. and Hinrichs, K.-U.: Turnover of microbial lipids in the deep biosphere and growth of benthic archaeal populations, Proc. Natl. Acad. Sci., 110(15), 6010–6014, doi:10.1073/pnas.1218569110, 2013.

Zhu, C., Lipp, J. S., Wörmer, L., Becker, K. W., Schröder, J. and Hinrichs, K. U.: Comprehensive glycerol ether lipid fingerprints through a novel reversed phase liquid chromatography-mass spectrometry protocol, Org. Geochem., 65, 53–62, doi:10.1016/j.orggeochem.2013.09.012, 2013.

Zhu, C., Yoshinaga, M. Y., Peters, C. A., Liu, X. L., Elvert, M. and Hinrichs, K. U.: Identification and significance of unsaturated archaeal tetraether lipids in marine sediments, Rapid Commun. Mass Spectrom., 28(10), 1144–1152, doi:10.1002/rcm.6887, 2014a.

Zhu, C., Meador, T. B., Dummann, W. and Hinrichs, K. U.: Identification of unusual butanetriol dialkyl glycerol tetraether and pentanetriol dialkyl glycerol tetraether lipids in marine sediments, Rapid Commun. Mass Spectrom., 28(November 2013), 332–338, doi:10.1002/rcm.6792, 2014b.

Benthic Archaea as potential sources of tetraether membrane lipids in sediments across an oxygen minimum zone

Marc A. Besseling[1*], Ellen C. Hopmans[1], Jaap S. Sinninghe Damsté[1,2], and Laura Villanueva[1]

[1]NIOZ Royal Netherlands Institute for Sea Research, Department of Marine Microbiology and Biogeochemistry, and Utrecht University. P.O. Box 59, NL-1790 AB Den Burg, The Netherlands.

[2]Utrecht University, Faculty of Geosciences, Department of Earth sciences, P.O. Box 80.021, 3508 TA Utrecht, The Netherlands

Correspondence to: Marc A. Besseling (marc.besseling@nioz.nl)

Abstract. Benthic Archaea comprise a significant part of the total prokaryotic biomass in marine sediments. Recent genomic surveys suggest they are largely involved in anaerobic processing of organic matter but the distribution and abundance of these archaeal groups is still largely unknown. Archaeal membrane lipids composed of isoprenoid diethers or tetraethers (glycerol dibiphytanyl glycerol tetraether, GDGT) are often used as archaeal biomarkers. Here, we compare the archaeal diversity and intact polar lipid (IPL) composition in both surface (0–0.5 cm) and subsurface (10–12 cm) sediments recovered within, just below, and well below the oxygen minimum zone (OMZ) of the Arabian Sea. Archaeal 16S rRNA gene amplicon sequencing revealed a predominance of Thaumarchaeota (Marine Group I, MG-I) in oxygenated sediments. Quantification of archaeal 16S rRNA and ammonia monoxygenase (*amo*A) of Thaumarchaeota genes and their transcripts indicated the presence of an active *in situ* benthic population, which coincided with a high relative abundance of hexose phosphohexose crenarchaeol, a specific biomarker for living Thaumarchaeota. On the other hand, anoxic surface sediments within the OMZ and all subsurface sediments were dominated by archaea belonging to the Miscellaneous Crenarchaeota Group (MCG), the Thermoplasmatales and archaea of the DPANN superphylum. Members of the MCG were diverse with a dominance of subgroup MCG-12 in anoxic surface sediments. This coincided with a high relative abundance of IPL GDGT-0 with an unknown polar head group. Subsurface anoxic sediments were characterized by higher relative abundance of GDGT-0, 2 and 3 with dihexose IPL-types, as well as GDGT-0 with a cyclopentanetetraol molecule and a hexose, as well as the presence of specific MCG subgroups, suggesting that these groups could be the biological sources of these archaeal lipids.

INTRODUCTION

Archaea are ubiquitous microorganisms in the marine system ~~(DeLong et al., 1994)~~(DeLong et al., 1994; Delong and Pace, 2001; Schleper et al., 2005). They occur in diverse environments, e.g. hydrothermal vents (Stetter et al., 1990), the marine water column (Karner et al., 2001; Massana et al., 2004), in the underlying sediments (Lloyd et al., 2013; Teske and Sørensen, 2008), and well below the seafloor (Biddle et al., 2006; Lipp et al., 2008), where they are considered key players in diverse biogeochemical processes (Offre et al., 2013, and references cited therein). Specifically marine sediments have been shown to contain a highly diverse archaeal community (Lloyd et al., 2013; Spang et al., 2017; Teske, 2013; Teske and Sørensen, 2008). The ammonia-oxidizing Thaumarchaeota of the marine group I.1a (further referred to as MG-I) is probably the most widely studied archaeal group in marine sediments. However, in comparison with studies of marine pelagic Thaumarchaeota, the diversity and distribution of benthic Thaumarchaeota is still not well established (e.g. Durbin & Teske, 2010; Jorgenson et al., 2012; Learman et al., 2016). Genomic studies have revealed the existence of uncultured archaeal groups other than Thaumarchaeota in marine, predominantly anoxic, sediments such as the Miscellaneous Crenarchaeota Group (MCG; Meng et al., 2014), archaea of the DPANN superphylum (composed of Micrarchaeota, Diapherotrites, Aenigmarchaeota, Nanohaloarchaeota, Parvarchaeota, Nanoarchaeota, Pacearchaeota and Woesearchaeota; Castelle et al., 2015; Rinke et al., 2013) and the Marine Benthic Group (MBG) B (Teske & Sørensen, 2008), and D (Lloyd et al., 2013). In the case of the archaea belonging to the groups of the MCG and MBG-D, metagenomic studies suggest that they are able to degrade extracellular proteins and aromatic compounds (Lloyd et al., 2013; Meng et al., 2014).

Archaeal diversity is currently determined through nucleic acid-based methods but the characterization of other cellular biomarkers such as membrane lipids has proven to be also effective in tracking the presence of archaeal groups in different ecosystems (e.g. Coolen et al., 2004a; Ingalls et al., 2012; Meador et al., 2015; Pitcher et al., 2011b; Sturt et al., 2004). One of the advantages of using lipid-based methods to determine the presence of archaeal groups is that lipids can be preserved in the sedimentary record. Therefore, they can also be used as biomarkers of the presence and metabolic potential of these microorganisms in past environments. On the contrary, other biomolecules like DNA have a more rapid turnover and they cannot be used for this purpose. In recent years, intact polar lipids (IPLs) have increasingly been applied for tracing 'living' bacteria and archaea in the environment (Lipp et al., 2008; Lipp and Hinrichs, 2009; Rossel et al., 2008). IPLs with polar head groups are present in living cells but upon cell lysis the polar head groups are lost, releasing core lipids (CLs) that may be preserved in the fossil record. Since IPLs degrade relatively quickly after cell death (Harvey et al., 1986), it is possible to associate the presence of IPLs in the environment with the occurrence of their living producers (Lipp and Hinrichs, 2009; Schubotz et al., 2009).

Archaeal membrane lipids are typically a variation of two main structures, *sn*-2,3-diphytanylglycerol diether (archaeol) with phytanyl (C_{20}) chains in a bilayer structure, and *sn*-2,3-dibiphytanyl diglycerol tetraether (glycerol dibiphytanyl glycerol tetraether, GDGT), in which the two glycerol moieties are connected by two C_{40} isoprenoid chains, allowing

the formation of a monolayer membrane (Koga and Morii, 2007). GDGTs containing 0–4 cyclopentane moieties (Fig. S1) are usually not exclusive to a specific archaeal group (Schouten et al., 2013) with the exception of the GDGT crenarchaeol, containing 4 cyclopentane and one cyclohexane moiety, which is deemed to be exclusive to the Thaumarchaeota phylum (De La Torre et al., 2008; Sinninghe Damsté et al., 2002, 2012). Recently, (~~,~~Lincoln et al., (2014) proposed the Marine Group II as potential producers of crenarchaeol. However, this is still debated (Lincoln et al., 2014b; Schouten et al., 2014). The newly described archaeal groups detected by genetic methods are yet uncultured, therefore, their membrane lipid composition remains unknown.

In this study, we determined the archaeal diversity in a marine benthic system along a strong gradient in bottom water oxygen concentrations and compared it with the diversity of archaeal lipids. We aimed to characterize changes in the archaeal benthic community under different physicochemical conditions, as well as to provide clues on the potential archaeal lipid biomarkers produced by uncultured benthic archaea. We analyzed sediments (surface 0–0.5 cm, and subsurface 10–12 cm) of the Murray ridge in the Arabian Sea, which is impinged by one of the strongest present-day oxygen minimum zones (OMZ). Previous studies observed changes in the diversity of archaeal lipids in the same environmental setting in sediments under different oxygen and nutrient concentrations (Lengger et al., 2012; 2014). In our study, we expand the ~~repertory~~ repertoire of archaeal lipid diversity previously detected by Lengger et al. (2012; 2014) by analyzing these sediments with High Resolution Accurate Mass/Mass spectrometry (UHPLC-HRAM MS). In addition, we determined the archaeal diversity by means of 16S rRNA gene amplicon sequencing, as well as the abundance and potential activity of specific archaeal groups by quantitative PCR (QPCR) of 16S rRNA and the metabolic gene coding for the ammonia monoxygenase (*amo*A gene) of Thaumarchaeota.

MATERIAL and METHODS

Sampling

Sediments were collected in the Northern Arabian Sea during the PASOM cruise in January 2009 with *R/V Pelagia*. Sediment cores obtained with a multicorer were taken on the Murray ridge at four depths, 885 m below sea level (mbsl) (within the OMZ), at 1306 mbsl (just below the OMZ), at 2470 mbsl and 3003 mbsl (both well below the OMZ) as previously described by Lengger et al. (2012). Upon retrieval the cores were sliced in 0.5 cm resolution for the first 2 cm and at 2 cm resolution beyond 10 cm below the surface, and stored at -80°C until further analysis. For an overview of the surface sediments physicochemical conditions see Table 1.

Lipid extraction and analysis

Total lipids were extracted from surface (upper 0–0.5 cm) and subsurface (10–12 cm) sediments after freeze-drying using a modified Bligh and Dyer method (Bligh and Dyer, 1959) as previously described by Lengger et al. (2014). C_{16}-PAF (1-O-hexadecyl-2-acetoyl-sn-glycero-3-phosphocholine) was added to the extracts as an internal standard and the

extracts were dried under a stream of nitrogen. The extracts with the added standard were then dissolved by adding solvent (hexane:isopropanol:H_2O 718:271:10 [v/v/v/v]) and filtered through a 0.45 µm, 4 mm-diameter True Regenerated Cellulose syringe filter (Grace Davison, Columbia, MD, USA).

IPLs were analyzed according to Sturt et al. (2004) with some modifications. An Ultimate 3000 RS UHPLC, equipped with thermostated auto-injector and column oven, coupled to a Q Exactive Orbitrap MS with Ion Max source with heated electrospray ionization (HESI) probe (Thermo Fisher Scientific, Waltham, MA), was used. Separation was achieved on a YMC-Triart Diol-HILIC column (250 x 2.0 mm, 1.9 µm particles, pore size 12 nm; YMC Co., Ltd, Kyoto, Japan) maintained at 30 °C. The following elution program was used with a flow rate of 0.2 mL min^{-1}: 100% A for 5 min, followed by a linear gradient to 66% A: 34% B in 20 min, maintained for 15 min, followed by a linear gradient to 40% A: 60% B in 15 min, followed by a linear gradient to 30% A: 70% B in 10 min, where A = hexane/2-propanol/formic acid/14.8 M NH_{3aq} (79:20:0.12:0.04 [v/v/v/v]) and B = 2-propanol/water/formic acid/ 14.8 M NH_{3aq} (88:10:0.12:0.04 [v/v/v/v]). Total run time was 70 min with a re-equilibration period of 20 min in between runs. HESI settings were as follows: sheath gas (N_2) pressure 35 (arbitrary units), auxiliary gas (N_2) pressure 10 (arbitrary units), auxiliary gas (N_2) T 50 ˚C, sweep gas (N_2) pressure 10 (arbitrary units), spray voltage 4.0 kV (positive ion ESI), capillary temperature 275 °C, S-Lens 70 V. IPLs were analyzed with a mass range of *m/z* 375 to 2000 (resolving power ~~resolution~~ 70,000 ~~parts per million, ppm~~at m/z 200), followed by data dependent MS^2 (resolving power ~~resolution~~ 17,500 ppm at m/z 200)), in which the ten most abundant masses in the mass spectrum (with the exclusion of isotope peaks) were fragmented ~~successively~~ (stepped normalized collision energy 15, 22.5, 30; isolation window 1.0 *m/z*). A dynamic exclusion window of 6 sec was used as well as ~~A~~an inclusion list ~~was used~~ with a mass tolerance of 3 ppm to target specific compounds (Table S1). The Q Exactive Orbitrap MS was calibrated within a mass accuracy range of 1 ppm using the Thermo Scientific Pierce LTQ Velos ESI Positive Ion Calibration Solution (containing a mixture of caffeine, MRFA, Ultramark 1621, and *N*-butylamine in an acetonitrile-methanol-acetic acid solution).

Peak areas for each individual IPL were determined by integrating the combined mass chromatogram (within 3 ppm) of the monoisotopic and first isotope peak of all relevant adducts formed (protonated, ammoniated and/or sodiated adducts may be formed in different proportions depending on the type of IPL). PAF was used as internal standard to continuously monitor MS performance and to asses matrix effects. Reported peak areas have been corrected for these effects. Absolute quantification of IPL GDGTs was not possible due to a lack of standards. Peak areas were not corrected for any possible differences in response factors between the various classes of IPL-crenarchaeol. IPLs with the same headgroup but with the regioisomer of crenarchaeol instead of crenarchaeol as the CL co-elute on the chromatographic system used here and any peak area reported for a crenarchaeol IPL thus represents the sum of both isomers.

To rule out any degradation of the GDGT-IPLs during storage of the sediments at -20°C, the anoxic surface sediment sample at 885 mbsl was also analyzed according to the method previously used by Lengger et al. (2012). The IPL

fraction was separated from the CLs with the use of a silica column and elution with MeOH (Lengger et al., 2012). This IPL fraction was hydrolyzed for 3 h and analyzed by HPLC-APCI/MS (according to Hopmans et al., 2016) and the IPL derived CL-GDGT distribution was compared with previously published data. ~~The total lipid extract from the surface sediment at 885 mbsl was further analyzed by acid hydrolysis to determine the composition and relative abundance of IPL-derived CL (resulting from the acid hydrolysis of IPLs) using the method described by Lengger et al. (2012) and analyzed by HPLC-APCI/MS (according to Hopmans et al., 2016).~~

Nucleic acids extraction, cDNA synthesis and quantitative PCR (QPCR) analyses

Sediment was centrifuged and the excess of water was removed by pipetting before proceeding with the extraction of nucleic acids from the sediment. DNA/RNA of surface (0–0.5 cm) and subsurface (10–12 cm) sediments was extracted with the RNA PowerSoil® Total Isolation Kit plus the DNA elution accessory (Mo Bio Laboratories, Carlsbad, CA). Concentration of DNA and RNA were quantified by Nanodrop (Thermo Scientific, Waltham, MA) and Fluorometric with Quant-iT^{TM} PicoGreen® dsDNA Assay Kit (Life technologies, Netherlands). RNA extracts were treated with DNAse and reverse-transcribed to cDNA as described by Pitcher et al. (2011). Quantification of archaeal 16S rRNA gene copies and *amo*A gene copies were estimated by QPCR by using the following primers; Parch519F and ARC915R (archaeal 16S rRNA gene), CrenAmoAQ-F and CrenAmoAModR (*amo*A gene), as previously described (Pitcher et al., 2011). For details on the QPCR conditions, efficiency and R^2 of the QPCR assays see Table S2.

16S rRNA gene amplicon sequencing, analysis, and phylogeny

PCR reactions were performed with the universal, Bacteria and Archaea, primers S-D-Arch-0159-a-S-15 and S-D-Bact-785-a-A-21 (Klindworth et al., 2013) as previously described in Moore et al. (2015). The archaeal 16S rRNA gene amplicon sequences were analyzed by QIIME v1.9 (Caporaso et al., 2010). Raw sequences were demultiplexed and then quality-filtered with a minimum quality score of 25, length between 250–350, and allowing maximum two errors in the barcode sequence. Taxonomy was assigned based on blast and the SILVA database version 123 (Altschul et al., 1990; Quast et al., 2013). Representative operational taxonomic units (OTUs, clusters of reads with 97% similarity) of archaeal groups were extracted through filter_taxa_from_otu_table.py and filter_fasta.py with QIIME (Caporaso et al., 2010). The phylogenetic affiliation of the partial archaeal 16S rRNA gene sequences was compared to release 123 of the Silva NR SSU Ref database (http://www.arb-silva.de/; Quast et al., 2013) using the ARB software package (Ludwig et al., 2004). Sequences were added to the reference tree supplied by the Silva database using the ARB Parsimony tool. MCG intragroup phylogeny for representative sequences of OTUs affiliated to the MCG lineage was carried out in ARB (Ludwig et al., 2004). Sequences were added by parsimony to a previously-built phylogenetic tree composed of reference sequences of the 17 MCG subgroups known so far (Kubo et al., 2012). Affiliation of any 16S rRNA gene sequences to a given subgroup was done assuming a similarity cutoff of ≥85%.

Cloning, sequencing and phylogeny of the archaeal *amo*A gene

Amplification of the archaeal *amo*A gene was performed as described by Yakimov et al., (2011). PCR reaction mixture was the following (final concentration): Q-solution 1× (PCR additive, Qiagen); PCR buffer 1×; BSA (200 µg ml^{-1}); dNTPs (20 µM); primers (0.2 pmol $µl^{-1}$); $MgCl_2$ (1.5 mM); 1.25 U Taq polymerase (Qiagen, Valencia, CA, USA). PCR conditions for these amplifications were the following: 95°C, 5 min; 35 × [95°C, 1 min; 55°C, 1 min; 72°C, 1 min]; final extension 72°C, 5 min. PCR products were gel purified (QIAquick gel purification kit, Qiagen) and cloned in the TOPO-TA cloning® kit from Invitrogen (Carlsbad, CA, USA) and transformed in *E. coli* TOP10 cells following the manufacturer's recommendations. Recombinant clones plasmid DNAs were purified by Qiagen Miniprep kit and screening by sequencing ($n \geq 30$) using M13R primer by Macrogen Europe Inc. (Amsterdam, The Netherlands). Obtained archaeal *amo*A protein sequences were aligned with already annotated *amo*A sequences by using the Muscle application (Edgar, 2004). Phylogenetic trees were constructed with the Neighbor-Joining method (Saitou and Nei, 1987) and evolutionary distances computed using the Poisson correction method with a bootstrap test of 1,000 replicates.

RESULTS

In this study, we analyzed both IPLs and DNA/RNA extracts from sediments previously collected along the Arabian Sea Murray Ridge within the OMZ (885 mbsl), just below the lower interface (1306 mbsl), and well below the OMZ (2470 and 3003 mbsl). The surface sediment (0-0.5 cm) at 885 mbsl was fully anoxic, however, the surface sediments below the OMZ were partly oxygenated (1306 mbsl), and fully oxygenated at 2470 and 3003 mbsl (Table 1). The subsurface sediments (10-12 cm) were fully anoxic at all stations (Table 1). For more details on the physicochemical conditions in these sediments see Table 1.

Archaeal IPL-GDGTs in the surface and subsurface sediments

A range of IPL-GDGTs (GDGT-0 to 4 and crenarchaeol) with the IPL-types monohexose (MH), dihexose (DH) and hexose-phosphohexose (HPH) was detected in surface and subsurface sediments across the Arabian Sea OMZ (Table 2). For the DH GDGT-0 two structural isomers (type-I with two hexose moieties at both ends of the CL, and type-II with one dihexose moiety; Table 2) were detected and identified based on their mass spectral characteristics (Fig. S2). These isomers were previously also reported in thaumarchaeotal cultures (Elling et al., 2014, 2017). In addition, GDGT-0 with both an ether-bound cyclopentanetetraol moiety and a hexose moiety as head groups was identified (Fig. S2) in some sediments (Table 2). This IPL was previously reported as a glycerol dibiphytanyl nonitol tetraether (GDNT; de Rosa et al. 1983) but was later shown to contain a 2-hydroxymethyl-1-(2,3-dihydroxypropoxy)-2,3,4,5-cyclopentanetetraol moiety by Sugai et al., (1995) on the basis of NMR spectroscopy characterization.

In the surface sediment at 885 mbsl, crenarchaeol IPLs were dominant (44.~~6~~7% of all detected IPL-GDGTs), occurring predominantly with DH as IPL-type (with a hexose head group on both ends; 43.1%; Table 2). IPL-GDGT-2 was the

second most abundant (29.6%), also mainly consisting of the IPL-type DH (29.5%; Table 2). IPL-GDGT-0, -1, -3 and -4 were occurring with relative abundances of 0.3%, 1.7%, 17.8% and 6.1%, respectively (Table 2). Overall, the majority (98.1%; Table 3) of IPL-GDGTs in surface sediment at 885 mbsl with IPL-type DH (all with a hexose molecule on both ends of the CL).

The surface sediment at 1306 mbsl contained mostly IPL-GDGT-0 (37.6% of all detected IPL-GDGTs), almost entirely with the IPL-type HPH (36.6% of the total; Table 2). Slightly less abundant was the IPL-crenarchaeol (35.6%), with the IPL-types HPH (18.7%) and DH type-I (15.5%) in equal amounts and with a minor relative abundance with MH (1.4%). Overall, the IPL-GDGTs in surface sediment at 1306 mbsl mainly contained the IPL-types HPH (55.4%; Table 3) and DH (42.0%~~) (~~; Table 3).

Well below the OMZ, surface sediments from 2470 and 3003 mbsl were both dominated by IPL-GDGT-0 (71.9 and 80.8% of all detected IPL-GDGTs, respectively), predominantly with IPL-type HPH (Table 2; Fig. 1a). The IPL-crenarchaeol had a lower relative abundance (26.6 and 17.6%, respectively) and again was dominated by the member with IPL-type HPH (Table 2). The other IPL-GDGTs occurred in minor quantities (<1%). Overall, IPL-type HPH was, thus, by far the most abundant head group detected in surface sediments at 2470 and 3003 mbsl (97.7% and 97.4%, respectively), in contrast to the other two surface sediments studied (Table 3).

In all subsurface (10-12 cm) sediments (i.e. at 885, 1306, 2470 and 3003 mbsl) the most abundant IPL-GDGTs were DH-crenarchaeol (28.9±3.8%; Table 2) and DH-GDGT-2 (25.5±3.5%; Table 2). DH was also the most commonly observed IPL-type attached to GDGT-3 and GDGT-4 (Table 2). Overall the distributions of the IPL-GDGTs in all subsurface sediments were relatively similar (Fig. 1a) in comparison to the substantial changes observed at the surface (cf. Fig. 1a). Overall, the IPL-type DH was the predominant one detected in subsurface sediment with a relative abundance ranging from 68.8% at 3003 mbsl to 92.9~~0.1~~% at ~~1306~~ 885 mbsl (Table 3). In contrast to all other sediments, in the subsurface sediments at 885 mbsl and 1306 mbsl, two different isomers (Fig. S2) of the DH-GDGT-0 were detected (Table 2). DH type-I (0.9% at 1306 mbsl) is also found in the other surface and subsurface sediments and in combination with other core GDGT structures, whereas the other isomer (DH type-II) only occurs (7.8% at 885 mbsl; 1.8% at 1306 mbsl; Table 2; Fig. S2b). In addition, these subsurface sediments also contain small amounts of GDGT-0 with cyclopentanetetraol and MH head groups (IPL-type HCP; 1.6% at 885 mbsl; 0.4% at 1306 mbsl; Table ~~1~~2; Fig. S2c).

We also determined the IPL-derived CL-GDGTs in the 885 mbsl surface sediment following the method of Lengger et al. (2012), in order to exclude IPL degradation within the stored samples. The CL-GDGTs composition derived from freshly obtained IPL showed the same distribution (r= 0.99, $\rho < 0.001$) as reported previously (Lengger et al., 2012). ~~p=1.00).~~

Archaeal diversity in the surface and subsurface sediment

Different archaeal groups were detected in surface and subsurface sediment across the Arabian sea OMZ. The surface sediment at 885 mbsl, contained archaeal 16S rRNA gene sequences that were assigned to several archaeal groups (Fig. 1b). The most dominant group was MCG (Total 30.5%, 12.2% attributed to C3; also known as MCG-15, Kubo et al., 2012). Another major group found was the DPANN Woesearchaeota Deep sea Hydrothermal Vent Group 6 (DHVEG-6, 20.3%; Fig. 1b; Castelle et al., 2015). Marine Benthic Group (MBG) -B, -D and -E were also present with 12.2%, 7.7% and 6.9% of the archaeal 16S rRNA gene reads, respectively (Fig. 1b). Sequences affiliated to the Marine Hydrothermal Vent Group (MHVG, 8.1%) of the phylum Euryarchaeota were also detected (Fig. 1b). Other groups, with lower relative abundances, were Thermoplasmatales groups ANT06-05 (5.7%) and F2apm1A36 (3.3%) and the DPANN Aenigmarchaeota (previously named Deep Sea Euryarchaeotic Group, DSEG; 1.6%; Fig. 1b).

Below the OMZ, in partly and fully oxygenated surface sediments at 1306, 2470 and 3003 mbsl (Table 1), the most dominant archaeal group was Thaumarchaeota MG-I with relative abundances of 81.5%, 89.7% and 100%, respectively (Fig. 1b). At 1306 mbsl other archaeal groups, such as MHVG (5.6%), Thermoplasmatales ASC21 (3.2%), DHVEG-6 (2.9%), MBG-B (2.4%) and MCG (1.3%) made up the rest of the archaeal community (Fig. 1b). At 2470 mbsl DHVEG-6 (1.1%) was still detectable besides the MG-I (Fig. 1b).

In the subsurface sediments (10−12 cm), only the DNA extracted from the sediments at 885 and 1306 mbsl gave a positive amplification signal. The archaeal composition of the subsurface (10−12 cm) sediments at 885 mbsl and 1306 mbsl was similar (Fig. 1b; Pearson correlation coefficient of 0.95), with most of the 16S rRNA gene reads classified within the MCG (47.5% and 48.4%, respectively). Other archaeal groups, such as MBG-D (14.4% and 5.7%, respectively), MBG-B (10.1% and 4.4%), the Woesearchaeota (7.8% and 10.4%), were also detected with comparable relative abundances (Fig. 1b). Other archaeal groups such as Thaumarchaeota Terrestrial hot spring, the Euryarchaeota MHVG, MBG-E and the Aenigmarchaeota were detected but at low (< 10%) relative abundance (Fig. 1b). Only minor amount of reads were classified as Thaumarchaeota MG-I (0.5% at 1306 mbsl) (Fig. 1b).

Considering the high relative abundance of the MCG detected in the surface sediment at 885 mbsl, as well as in the subsurface (10−12 cm) sediments at 885 mbsl and 1306 mbsl (between 30.5-48.4% of total archaeal 16S rRNA gene reads detected in those samples), we performed phylogenetic analyses to determine the diversity of subgroups of the MCG within these sediments. A total of 57 representative 16S rRNA gene reads assigned to MCG were extracted from the dataset and incorporated in a MCG phylogenetic tree of Fillol et al. (2015) (Fig. 2). The majority of MCG 16S rRNA gene reads from the 885 mbsl surface sediment (77.3%; Table 4) clustered in subgroup 15. In the 885 mbsl subsurface sediment, the majority of MCG reads clustered within subgroups 8 and 15 (33.6% and 19.6%, respectively; Table 4). In the 1306 mbsl surface sediment there was only a low relative abundance of MCG (Fig. 1b); all MCG archaea detected clustered in subgroup 15 (Table 4). On the other hand, in the 1306 mbsl subsurface sediment the reads clustered in subgroups 15, 2 and 14 (34.3%, 10.9% and 10.9%, respectively; Fig. 2).

As the Thaumarchaeota MGI was dominant in oxygenated sediments at 1306, 2470 and 3003 mbsl (Fig. 1b), we further analyzed the diversity of this group by performing a more detailed phylogeny of the recovered 16S rRNA gene reads attributed to this group. Five OTUs dominated the Thaumarchaeota MG1 (Table 5); we will refer to them as OTU-1 to -5. OTU-1, 2, 3 and 5 were phylogenetically closely related to other known benthic Thaumarchaeota MGI species, such as '*Ca.* Nitrosoarchaeum koreensis MY1' or environmental 16S rRNA gene sequences from marine sediments (Fig.3). On the other hand, OTU-4 clustered with 16S rRNA gene sequences from pelagic Thaumarchaeota MGI species, like *Ca.* Nitrosopelagicus brevis, and also clustered with 16S rRNA sequences recovered from seawater SPM (Fig. 3). OTU-3 was the most abundant OTU in the surface sediment at 1306, 2470, and 3003 mbsl with a relative abundance of 44-68% (Table 5). At 1306 mbsl OTU-4 was the second most abundant (35.1%). This OTU had a much lower relative abundance (1.6% and 0.0%) at 2470 and 3003 mbsl, respectively (Table 5). The relative abundance of OTU-2 increased with increasing sampling station depth (Table 5), OTU-1 and 5 had an abundance <5% in the surface sediments (Table 5).

The diversity of Thaumarchaeota MG1 was further assessed by amplification, cloning and sequencing of the archaeal *amo*A gene. Most of the *amo*A gene sequences from surface (27 out of 29 clones) and subsurface sediment at 885 mbsl (9 out of 10 clones) and just one from the surface sediment from 1306 mbsl (1 out of 58 clones) were closely related with *amo*A gene sequences previously recovered from SPM at 1050 mbsl from this area of the Arabian Sea (Villanueva et al., 2014). Phylogenetically they fall within the 'Water column B, subsurface water' *amo*A clade as defined by Francis et al. (2005) (Fig. 4). At 1306 and 3003 mbsl (surface and subsurface) the majority of recovered *amo*A gene sequences clustered within the 'shallow water/sediment' clade (100 and 98.3%, respectively) and are closely related with *amo*A gene sequences from water column SPM at 170 mbsl (Villanueva et al., 2014) as well as *amo*A gene coding sequences previously detected in sediments (Villanueva et al., 2014; Fig. 4). Of all recovered *amo*A gene sequences from 885 mbsl only a small fraction (8.3%) clustered within the 'shallow water/sediment' clade (Fig. 4).

Abundance and potential activity of archaea in surface and subsurface sediments

The abundance of archaeal 16S rRNA gene copies in the surface sediments of different stations varied slightly: it was lowest at 1306 mbsl (9.8×10^9 copies g^{-1} sediment) and highest at 2470 mbsl (1.5×10^{11}; Fig. 5a). The potential activity, based on the 16S rRNA gene transcripts of the archaeal 16S rRNA gene, was the lowest at 2470 mbsl (5×10^4 transcripts g^{-1} sediment), while a higher potential activity was detected at 885, 1306 and 3003 mbsl ($0.9\text{-}42 \times 10^7$; Fig. 5a). The abundance of archaeal 16S rRNA gene copies in the subsurface sediment varied also within one and a half order of magnitude ($1.1\text{-}54 \times 10^9$; Fig. 5c), with a decrease with increasing water depth. The potential activity showed less variation within the subsurface sediments ($1.2\text{-}22 \times 10^7$ 16S rRNA gene transcripts g^{-1} of sediment; Fig. 5c) than in the surface sediments.

The abundance of Thaumarchaeota was estimated by quantifying the archaeal *amo*A gene copies. The highest abundance of *amo*A gene copies in surface sediment was detected at 2470 mbsl (1.0×10^9 copies g^{-1} sediment), and the

lowest at 885 mbsl (5×10^4; Fig. 5b). *Amo*A gene transcripts in surface sediment were under the detection limit at 885 mbsl but were detected below the OMZ with 4×10^2, 2.3×10^6 and 8×10^3 gene transcripts g^{-1} of sediment at 1306, 2470 and 3003 mbsl, respectively (Fig. 5b). In subsurface sediments, the abundance of *amo*A gene copies was low at 885 and 1306 mbsl (5.4-19 $\times 10^2$ gene transcripts g^{-1} sediment) and higher at 2470 and 3003 mbsl (4.1×10^5, 5.4×10^6, respectively; Fig. 5d). *Amo*A gene transcripts were not detected in the subsurface sediments (Fig. 5d).

DISCUSSION

In this study, we assessed the changes in benthic archaeal diversity and abundance in sediments of the Arabian Sea oxygen minimum zone along a gradient in bottom water oxygen concentrations. The steep Murray Ridge protrudes the OMZ, allowing the study of sediments deposited under varying bottom water oxygen concentrations. All these sediments receive organic matter (OM), the most important fuel for benthic prokaryotic activity in sediments. This OM is produced in a relatively small area of the ocean (i.e. the station within the OMZ, at 885 mbsl, and well below the OMZ, at 3003 mbsl, are only 110 km apart) and, therefore, likely composed of the same primary photosynthate. However, due to differences in the degree of mineralization resulting from different exposure to oxic conditions in the water column, there were differences in OM quality. OM in the sediments within the OMZ has a higher biochemical "quality" based on amino acid composition and intact phytopigments compared to OM in the sediments below the OMZ (Koho et al., 2013). Therefore, changes in the quality and flux of OM received by the different sediment niches could also affect the archaeal community composition as several of the archaeal groups (i.e. MCG and MBG-D) reported here have been suggested to use OM as carbon source in anoxic conditions (Lloyd et al., 2013).

Effect of oxygen availability on archaeal diversity and abundance in the surface sediments

We detected large differences in archaeal diversity between the surface sediment deposited within the OMZ and those deposited below the OMZ. In contrast to the diverse anaerobic archaeal community in the surface of 885 mbsl, in surface sediments at 1306, 2470 and 3003 mbsl, Thaumarchaeota MGI were dominant, representing 80-100% of the archaeal population (Fig. 1). This clear difference in the benthic archaeal population in the surface sediments can be attributed to the oxygen availability as Thaumarchaeota are known to require oxygen for their metabolism (i.e. nitrification; Könneke et al., 2005). In fact, the oxygen penetration depth (OPD) was observed to be 3, 10, and 19 mm in sediments at 1306, 2470, and 3003 mbsl, respectively, while in sediments at 885 mbsl, the OPD was barely 0.1 mm (Table S1; Kraal et al., 2012). The surface (0-5 mm) sediment at 1306 mbsl was not fully oxygenated (OPD of 3 mm), which probably explains the detection in relatively low abundance (ca. 20%) of the anaerobic archaea that thrive in the anoxic sediment from 885 mbsl. The low OPD at 1306 mbsl also explains the low *amo*A gene expression in comparison with the deeper surface sediments (Figs. 5b,d). Overall this indicates the presence of Thaumarchaeota with lower activity in the surface sediments at 1306 mbsl (Fig. 5). Within the Thaumarchaeota MG1 group, we also detected changes in the relative abundance of specific OTUs in the surface sediments at 1306, 2470 and 3003 mbsl (Table 5).

For example, OTU-2 becomes progressively more abundant with increasing water depth, suggesting that this OTU is favored at the higher oxygen concentrations found in the surface sediment at 3003 mbsl. OTU-4 was closely affiliated with '*Ca.* Nitrosopelagicus brevis', a pelagic MG-I member, which indicates that this DNA is most likely derived from the overlaying water column (Table 5), and thus should be considered to represent fossil DNA.

High *amo*A gene abundances were detected in the surface sediment at 2470 and 3003 mbsl, while values in the surface of 885 mbsl were approximately three orders of magnitude less. The lack of oxygen in the surface sediments at 885 mbsl and in the subsurface sediments, as well as undetectable *amo*A gene transcripts at those depths, suggest that in these cases the *amo*A gene DNA signal is fossil. It is well known that under anoxic conditions DNA of marine pelagic microbes may become preserved in sediments even for periods of thousands of years (Boere et al., 2011; Coolen et al., 2004b). The fossil origin of the Thaumarchaeotal *amo*A gene is supported by the phylogenetic affiliation of the *amo*A gene fragments amplified from the 885 mbsl surface sediment, as those sequences were closely related to *amo*A gene sequences detected in the suspended particulate matter in the subsurface water column (Villanueva et al., 2014), thus suggesting that the detected DNA originated from pelagic Thaumarchaeota present in the subsurface water column, as proposed for the presence of OTU-4 16S rRNA gene sequences in the surface sediment (see earlier).

There is a discrepancy between the 16S rRNA gene copy numbers and the *amo*A gene copy numbers within the sediments (Fig. 5). *Amo*A gene copies were consistently lower than the 16S rRNA gene copies, even within sediments that were completely dominated by Thaumarchaeota MG-I. This may be caused by the *amo*A gene primer mismatches and/or the disparity of gene copy numbers within the archaeal genomes (Park et al., 2008).

In the anoxic surface sediment at 885 mbsl (within the OMZ), we detected a highly diverse archaeal population composed of MCG, Thermoplasmatales, MBG-B, -D and -E, Woesearchaeota, and MHVG. Archaeal groups such as MCG and MBG-B and E have been previously described in anoxic marine sediments, where they have been suggested to be involved in anaerobic OM degradation (e.g. Biddle et al., 2006; Inagaki et al., 2003; Castelle et al., 2015). Members of the DPANN Woesearchaeota were only present in the surface sediment at 885 mbsl but not in the subsurface anoxic sediments at 885 and 1306 mbsl, suggesting that their presence here is not solely dependent on the absence of oxygen but possibly also on the OM composition and availability in surface and subsurface sediments. Alternatively, the Woesearchaeota 16S rRNA gene signal could also originate from the water column and deposited in the surface sediment at 885 mbsl as fossil DNA as observed for the case of Thaumarchaeota as mentioned above.

Archaeal community composition in the anoxic subsurface sediments

The archaeal diversity in the subsurface sediment (10–12 cm) from both 885 and 1306 mbsl (i.e. dominated by MCG, MBG-B, -D and –E) is similar to that observed in the surface sediment at 885 mbsl. This supports that oxygen availability is an important factor for determining the diversification of archaeal groups (Fig. 1b). MCG, one of the dominant archaeal groups in these sediments, showed substantial differences in the distribution of its subgroups (Table 4). All subsurface sediments had a high intra-group diversity of MCG. This is in contrast with the surface sediment at

885 and 1306 mbsl where a high relative abundance of the subgroup MCG-15 is detected (Table 4). A recent survey of the ecological niches and substrate preferences of the MCG in estuarine sediments based on genomic data pointed to MCG-6 archaea as degraders of complex extracellular carbohydrate polymers (plant-derived), while subgroups 1, 7, 15 and 17 have mainly the potential to degrade detrital proteins (Lazar et al., 2016). Lazar et al. (2016) also described the presence of aminopeptidases coded in the genome bin of MCG-15, suggesting that this subgroup could be specialized in degradation of extracellular peptides in comparison with the other MCG subgroups, which would be restricted to the use of amino acid and oligopeptides. Considering the dominance of the MCG-15 subgroup in the surface sediments analyzed in this study (Table 4), we hypothesize that the proteinaceous OM deposited in the surface sediment, which mainly originates from photosynthate, is still quite undegraded. This would favor the MCG-15 in this niche, fueling its metabolism by the degradation of peptides extracellularly, while in subsurface sediments, other MCG groups such as 2, 8 and 14 would be more favored.

The archaeal 16S rRNA gene abundance in the subsurface sediments progressively declined with increasing water depth, while the potential activity was similar. This can be due to the expected decrease in the flux of OM being delivered to these anoxic sediments layers attributed to higher degradation of OM in oxygenated bottom waters and the progressively larger oxic zone in the sediments (Lengger et al., 2012; Nierop et al., 2017). This results in lower organic carbon concentrations and a decreased biochemical quality of the OM (Koho et al., 2013; Nierop et al., 2017) to sustain the heterotrophic archaeal population inhabiting the anoxic subsurface sediments. Also the presence or lack of macrofauna in the analyzed sediments would have an effect on the OM composition, sediments within the OMZ are less prone to bioturbation which most likely resulted in higher OM preservation (Koho et al., 2013). Differences in the OM biochemical composition can influence the microbial community composition as was shown recently for North Sea sediments (Oni et al., 2015).

Benthic archaea as potential sources for archaeal IPLs

Archaeal lipids in surface and deeper sediments of the Murray Ridge (Lengger et al., 2012, 2014), as well as in the overlaying water column (Pitcher et al., 2011; Schouten et al., 2012), have been studied earlier. The study by Lengger et al. (2012) was limited to the determination of MH-, DH- and HPH-crenarchaeol with HPLC/ ESI-MS^2 using a specific selected reaction monitoring method (SRM; Pitcher et al., 2011). A follow-up study of Lengger et al. (2014) reported MH-, DH- and HPH-IPLs with multiple CL-GDGTs. In our study, we expanded the screening for IPLs carrying different polar head groups in combination with multiple CLs using high resolution accurate mass/mass spectrometry (see Table S1). By applying this method, we were able to target a broader range of IPLs in these sediments. ~~unraveled the unknown diversity of IPL-GDGTs in the sediments under study, which~~ This allows a more direct comparison with the archaeal diversity detected by gene-based methods. Note that the study of Lengger et al. (2014) used a different sampling resolution (surface sediment used was 0-2 compared to our 0-0.5 cm) and our results can, therefore, not be directly compared.

The fully oxygenated surface sediments showed a dominance of GDGT-0 and crenarchaeol mostly with HPH as IPL-type (Table 2). This is the expected IPL-GDGT signature of Thaumarchaeota as previously observed in pure cultures (Elling et al., 2015, 2017; Pitcher et al., 2010; Qin et al., 2015; Schouten et al., 2008; Sinninghe Damsté et al., 2012). The predominance of the HPH IPL-type in surface (0-2 cm) sediments was previously interpreted as an indication of the presence of an active Thaumarchaeotal population synthesizing membrane lipids *in situ* (Lengger et al., 2012, 2014), taking into account the labile nature of sedimentary ~~HPH IPLs~~phospholipids (Harvey et al., 1986; Schouten et al., 2010; Xie et al., 2013)~~(Harvey et al., 1986; Schouten, et al., 2010)~~. This hypothesis is strongly supported by our data because (i) the archaeal community in the oxygenated surface (0-0.5 cm) sediments is dominated by Thaumarchaeota (Fig. 1) and (ii) the high abundance of thaumarchaeotal *amo*A gene copies and gene transcripts detected in the oxygenated surface sediments. On the other hand, in the anoxic surface sediment at 885 mbsl, crenarchaeol was predominantly present with DH as the predominant IPL-type (Table 2). This is considered to be a fossil signal of Thaumarchaeota deposited from the water column due to a higher preservation potential of glycolipid head groups (as present in DH) as previously suggested (Lengger et al., 2012, 2014). However, ~~(~~Logemann et al.~~,~~ (2011) showed in a 100 day degradation experiment that IPL GDGTs (ether-bound lipids) were hardly degraded in anoxic sediments and, hence, the differences in degradation rates between phospholipid versus glycolipid GDGTs still need to be determined, especially on longer time scales that apply to sediments. Nevertheless, the presence of *amo*A gene sequences in the 885 mbsl surface sediment, which are derived from the deeper water column, as well as the much lower *amo*A gene abundance and lack of *amo*A gene expression (Fig. 5b, d) supports the contention that the crenarchaeol IPLs in the surface sediment at 885 mbsl are predominantly fossil since evidence for active Thaumarchaeota is lacking.

The low relative abundance of GDGT-0 IPLs in the surface sediment at 885 mbsl (Table 2) is remarkable. Only MH-GDGT-0 was detected in low relative abundance (0.3 %), whereas any other of the IPL-types with GDGT-0 as CL that were screened for in our study (Table S2; Fig. 1b) was absent. In contrast, Lengger et al. (2012) reported a significant amount of IPL-derived CL-GDGT-0 (i.e. 18.5% of total CL GDGTs) when the head groups of the IPLs are released by acid hydrolyses and the remaining CLs were analyzed in a surface (0-0.5 cm) sediment from the same site. We re-analyzed the IPL-derived CL-GDGT composition in the surface (0-0.5 cm) sediment at 88 mbsl and recovered an identical CL-GDGT distribution as reported by Lengger et al. (2014). The discrepancy between CL and IPL distribution may partly be explained by the underestimation of MH IPLs by our method. However, the difference in response factor for the different IPL types is not sufficient to explain this discrepancy. Therefore, we speculate it is due to the presence of an IPL-type with unknown head groups not included in our analytical window. This unknown IPL GDGT-0 may originate from any of the archaeal groups present in the surface sediment at 885 mbsl, such as MCG, Thermoplasmatales, MBG-B, MBG-E and Euryarchaeota MHVG. Woesearchaeota also occur relatively abundant in the surface sediments at 885 mbsl (Fig. 1) but recent studies suggest that their small genomes lack the genes coding for

414 the enzymes of the GDGT biosynthetic pathway (Jahn et al., 2004; Podar et al., 2013; Villanueva et al., 2017; Waters et
415 al., 2003). Therefore, they are not expected to contribute to the IPL-GDGT pool. Ruling out the Woesearchaeota as a
416 possible source of IPL-GDGTs, the IPL GDGT-0 with unknown polar head group(s) in the surface sediment at 885
417 mbsl may be attributed to the MCG, which make up 30.5% of the archaeal 16S rRNA gene reads in this sediment. Most
418 of these MCG archaea fall into subgroup MCG-15 (Table 4). Previous studies proposed butanetriol dibiphytanyl
419 glycerol tetraethers (BDGTs) as putative biomarker of the MCG based on the correlation between the presence of these
420 components and MCG in estuarine sediments (Meador et al., 2014). However, we did not detect any IPL BDGTs in the
421 sediments analyzed in our study. Buckles et al. (2013) suggested that members of the MCG and Crenarchaeota group
422 1.2 could be the biological source of IPL GDGT-0 found in the anoxic hypolimnion of a tropical lake. Considering
423 these evidences, it is possible that the unknown IPL GDGT-0 present in the surface sediment at 885 mbsl could be a
424 biomarker for MCG.
425 In subsurface sediments, the IPL GDGT distribution was remarkably different from that detected in the surface
426 oxygenated sediment as higher relative abundances of GDGT-1, 2, 3 and 4 were detected in detriment of GDGT-0,
427 similar to the distribution detected in the surface sediments at 885 mbsl. This may represent new archaeal production in
428 the anoxic sediments, selective preservation of archaeal lipids produced in the water column and surface sediments, or
429 both. The HCP GDGT-0 and two isomers of the DH GDGT-0 (Fig. S2) were detected in the subsurface sediments at
430 885 and 1306 mbsl but not in those from deeper waters (Table 2). Since these IPLs were not detected in the surface
431 sediments, it is likely that they are produced *in situ*. Unfortunately, we only obtained information on the archaeal
432 community composition of the subsurface sediments at shallow water depth, so we cannot compare these with the
433 subsurface sediments from deeper waters that lack these DH moieties, which could have provided a clue towards the
434 archaeal source of these IPLs. An IPL composed of GDGT-0 with a cyclopentanetetraol head group has been
435 previously detected in cultures of the hyperthermophilic crenarchaeal *Sulfolobales* (Langworthy et al., 1974; Sturt et al.,
436 2004). However, members of the *Sulfolobales* were not detected in our 16S rRNA gene amplicon sequencing data. We
437 also detected a high relative abundance of MCG (up to 48.4% relative abundance) in the subsurface sediment at 885 and
438 1306 mbsl (Fig. 1). The diversity of the MCG population in the subsurface sediments was much higher in comparison
439 with the diversity in surface sediments at 885 mbsl as sequences closely related to the MCG subgroups, 2, 8, 10, 14, 5b,
440 15, and 17 were detected both in the 885 mbsl and 1306 mbsl subsurface sediments (Fig. 2). This presence of these
441 different MCG members, likely caused by niche differentiation (see before), may be the origin of the unusual DH-
442 GDGT-0 isomer and the HCP-GDGT-0 IPL that we detected within the subsurface sediments at 885 and 1306 mbsl.

443 **CONCLUSIONS**

444 By using a combined 16S rRNA gene amplicon sequencing and IPL analysis with high-resolution accurate mass/mass
445 spectrometry we have unraveled the high diversity of benthic archaea harbored ~~specially~~ in oxygenated and anoxic

sediments of the Arabian Sea, as well as ~~increasing~~ widening ~~the repertoire~~our detection window of archaeal intact polar lipid~~s detected~~. DNA-based analyses revealed a dominance of active benthic *in situ* Thaumarchaeota in those sediment where oxygen was present, which coincided with high relative abundance of the HPH-crenarchaeol previously suggested to be a marker of living Thaumarchaeota ~~(Pitcher et al., 2011a)~~. In the anoxic marine sediments analyzed here, members of the MCG, DPANN and Euryarchaeota Thermoplasmatales dominated. We also observed a high diversity within the MCG with a more diverse population in subsurface sediments. ~~Surface anoxic sediments with an important MCG population coincided with a high abundance of an IPL GDGT-0 with a yet-unknown polar head group. This result, together with previous studies, suggests that this IPL GDGT-0 could be a potential biomarker for MCG. Besides, s~~Subsurface anoxic sediments had a high relative abundance of IPL GDGT-1, 2, and 3 with DH headgroups, which could either be attributed to fossil signal due to the more recalcitrant nature of the glycosidic bonds ~~(Schouten et al., 2010)~~ or being IPLs synthesized by the archaeal groups detected in those sediments. In addition, ~~unusual headgroups were also detected, as an~~ IPL GDGT-0 was also detected with a hexose head group on both ends of the core lipid, two hexoses on one end, and a cyclopentanetetraol molecule bound to the core lipid and a hexose attached to it. ~~Possibly ruling out the involvement of archaeal m~~Members of the DPANN could possibly be ruled out of ~~in~~ making those lipids due to the lack of lipid biosynthetic pathway. ~~(Podar et al., 2013; Waters et al., 2003), d~~Dominant archaeal members in those sediments such as the MCG and Thermoplasmatales, could be potential biological sources of these IPLs. To conclude, this combined approach has shed light on the possible biological sources of specific archaeal IPLs and also detected a highly diverse ~~er diversity of~~ benthic archaeal community. ~~than previously thought.~~

Acknowledgments

Elda Panoto is thanked for assistance with molecular analyses. We would like to thank the captain and crew of the RV *Pelagia* as well as the cruise leader, technicians and scientists participating in cruise 64PE301. This PASOM cruise was funded by the Earth and Life Science and Research Council (ALW) with financial aid from the Netherlands Organization for Scientific Research (NWO) (grant 817.01.015) to Prof. G.J. Reichart (PI). NIOZ is acknowledged for the studentship of MAB. This research was further supported by the NESSC and SIAM Gravitation Grants (024.002.001 and 024.002.002) from the Dutch Ministry of Education, Culture and Science (OCW) to JSSD.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J.: Basic local alignment search tool., J. Mol. Biol., 215(3), 403–10, doi:10.1016/S0022-2836(05)80360-2, 1990.

Biddle, J. F., Lipp, J. S., Lever, M. a, Lloyd, K. G., Sørensen, K. B., Anderson, R., Fredricks, H. F., Elvert, M., Kelly, T. J., Schrag, D. P., Sogin, M. L., Brenchley, J. E., Teske, A., House, C. H. and Hinrichs, K.-U.: Heterotrophic Archaea dominate sedimentary subsurface ecosystems off Peru., Proc. Natl. Acad. Sci. U. S. A., 103(10), 3846–3851, doi:10.1073/pnas.0600035103, 2006.

Bligh, E. G., Dyer, W. J.: A rapid method of total lipid extraction and Purification, Can. J. Biochem. Physiol., 37, 911–

917, doi:10.1139/o59-099., 1959.

Boere, A. C., Rijpstra, W. I. C., de Lange, G. J., Sinninghe Damsté, J. S. and Coolen, M. J. L.: Preservation potential of ancient plankton DNA in Pleistocene marine sediments, Geobiology, 9, 377–393, doi:10.1111/j.1472-4669.2011.00290.x, 2011.

Buckles, L. K., Villanueva, L., Weijers, J. W. H., Verschuren, D. and Sinninghe Damsté, J. S.: Linking isoprenoidal GDGT membrane lipid distributions with gene abundances of ammonia-oxidizing Thaumarchaeota and uncultured crenarchaeotal groups in the water column of a tropical lake (Lake Challa, East Africa)., Environ. Microbiol., 2, doi:10.1111/1462-2920.12118, 2013.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Peña, A. G., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights, D., Koenig, J. E., Ley, R. E., Lozupone, C. A., Mcdonald, D., Muegge, B. D., Pirrung, M., Reeder, J., Sevinsky, J. R., Turnbaugh, P. J., Walters, W. A., Widmann, J., Yatsunenko, T., Zaneveld, J. and Knight, R.: correspondEnce QIIME allows analysis of high- throughput community sequencing data Intensity normalization improves color calling in SOLiD sequencing, Nat. Publ. Gr., 7(5), 335–336, doi:10.1038/nmeth0510-335, 2010.

Castelle, C. J., Wrighton, K. C., Thomas, B. C., Hug, L. A., Brown, C. T., Wilkins, M. J., Frischkorn, K. R., Tringe, S. G., Singh, A., Markillie, L. M., Taylor, R. C., Williams, K. H. and Banfield, J. F.: Genomic Expansion of Domain Archaea Highlights Roles for Organisms from New Phyla in Anaerobic Carbon Cycling, Curr. Biol., 25, 1–12, doi:10.1016/j.cub.2015.01.014, 2015.

Coolen, M. J. L., Muyzer, G., Rijpstra, W. I. C., Schouten, S., Volkman, J. K. and Sinninghe Damsté, J. S.: Combined DNA and lipid analyses of sediments reveal changes in Holocene haptophyte and diatom populations in an Antarctic lake, Earth Planet. Sci. Lett., 223(1–2), 225–239, doi:10.1016/j.epsl.2004.04.014, 2004a.

Coolen, M. J. L., Hopmans, E. C., Rijpstra, W. I. C., Muyzer, G., Schouten, S., Volkman, J. K. and Sinninghe Damsté, J. S.: Evolution of the methane cycle in Ace Lake (Antarctica) during the Holocene: response of methanogens and methanotrophs to environmental change, Org. Geochem., 35(10), 1151–1167, doi:10.1016/j.orggeochem.2004.06.009, 2004b.

Delong, E. F. and Pace, N. R.: Environmental Diversity of Bacteria and Archaea, Soc. Syst. Biol., 50(4), 470–478, doi:10.1080/10635150118513, 2001.

DeLong, E. F., Ying Wu, K., Prézelin, B. B. and Jovine, R. V. M.: High abundance of Archaea in Antarctic marine picoplankton, Nature, 371, 695–697, doi:10.1038/371695a0, 1994.

Durbin, A. M. and Teske, A.: Sediment-associated microdiversity within the Marine Group I Crenarchaeota, Environ. Microbiol. Rep., 2(5), 693–703, doi:10.1111/j.1758-2229.2010.00163.x, 2010.

Edgar, R. C.: MUSCLE: Multiple sequence alignment with high accuracy and high throughput, Nucleic Acids Res., 32(5), 1792–1797, doi:10.1093/nar/gkh340, 2004.

Elling, F. J., Konneke, M., Lipp, J. S., Becker, K. W., Gagen, E. J. and Hinrichs, K. U.: Effects of growth phase on the membrane lipid composition of the thaumarchaeon Nitrosopumilus maritimus and their implications for archaeal lipid distributions in the marine environment, Geochim. Cosmochim. Acta, 141, 579–597, doi:10.1016/j.gca.2014.07.005, 2014.

Elling, F. J., Könneke, M., Mußmann, M., Greve, A. and Hinrichs, K.-U.: Influence of temperature, pH, and salinity on membrane lipid composition and TEX86 of marine planktonic thaumarchaeal isolates, Geochim. Cosmochim. Acta, 171, 238–255, doi:10.1016/j.gca.2015.09.004, 2015.

Elling, F. J., Könneke, M., Nicol, G. W., Stieglmeier, M., Bayer, B., Spieck, E., de la Torre, J. R., Becker, K. W., Thomm, M., Prosser, J. I., Herndl, G. J., Schleper, C. and Hinrichs, K. U.: Chemotaxonomic characterisation of the thaumarchaeal lipidome, Environ. Microbiol., 19(7), 2681–2700, doi:10.1111/1462-2920.13759, 2017.

Fillol, M., Auguet, J.-C., Casamayor, E. O. and Borrego, C. M.: Insights in the ecology and evolutionary history of the Miscellaneous Crenarchaeotic Group lineage., ISME J., 1–13, doi:10.1038/ismej.2015.143, 2015.

Harvey, H. R., Fallon, R. D. and Patton, J. S.: The effect of organic matter and oxygen on the degradation of bacterial membrane lipids in marine sediments, Geochim. Cosmochim. Acta, 50, 795–804, doi:10.1016/0016-7037(86)90355-8, 1986.

Hopmans, E. C., Schouten, S. and Sinninghe, J. S.: The effect of improved chromatography on GDGT-based palaeoproxies, Org. Geochem., 93, 1–6, doi:10.1016/j.orggeochem.2015.12.006, 2016.

Inagaki, F., Suzuki, M., Takai, K., Oida, H., Sakamoto, T., Aoki, K., Nealson, K. H. and Horikoshi, K.: Microbial Communities Associated with Geological Horizons in Coastal Subseafloor Sediments from the Sea of Okhotsk Microbial Communities Associated with Geological Horizons in Coastal Subseafloor Sediments from the Sea of Okhotsk, Appl. Environ. Microbiol., 69(12), 7224–7235, doi:10.1128/AEM.69.12.7224, 2003.

Ingalls, A. E., Huguet, C. and Truxal, L. T.: Distribution of intact and core membrane lipids of archaeal glycerol dialkyl glycerol tetraethers among size-fractionated particulate organic matter in Hood Canal, Puget Sound, Appl. Environ. Microbiol., 78(5), 1480–1490, doi:10.1128/AEM.07016-11, 2012.

Jahn, U., Summons, R., Sturt, H., Grosjean, E. and Huber, H.: Composition of the lipids of Nanoarchaeum equitans and their origin from its host Ignicoccus sp. strain KIN4/I, Arch. Microbiol., 182(5), 404–413, doi:10.1007/s00203-004-0725-x, 2004.

Jorgenson, S. L., Hannisdal, B., Lanzén, A., Baumberger, T., Flesland, K., Fonseca, R., Øvreås, L., Steen, I. H., Thorseth, I. H., Pedersen, R. B. and Schleper, C.: Correlating microbial community pro fi les with geochemical data in highly strati fi ed sediments from the Arctic Mid-Ocean Ridge, Proc. Natl. Acad. Sci., 109(42), E2846–E2855, doi:10.1594/PANGAEA.786687, 2012.

Karner, M. B., DeLong, E. F. and Karl, D. M.: Archaeal dominance in the mesopelagiczone of the Pacific Ocean, Nature, 409(6819), 507–510, doi:10.1038/35054051, 2001.

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M. and Glockner, F. O.: Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies, Nucleic Acids Res., 41(1), 1–11, doi:10.1093/nar/gks808, 2013.

Koga, Y. and Morii, H.: Biosynthesis of ether-type polar lipids in archaea and evolutionary considerations., Microbiol. Mol. Biol. Rev., 71(1), 97–120, doi:10.1128/MMBR.00033-06, 2007.

Koho, K. A., Nierop, K. G. J., Moodley, L., Middelburg, J. J., Pozzato, L., Soetaert, K., Van Der Plicht, J. and Reichart, G. J.: Microbial bioavailability regulates organic matter preservation in marine sediments, Biogeosciences, 10(2), 1131–1141, doi:10.5194/bg-10-1131-2013, 2013.

Könneke, M., Bernhard, A. E., de la Torre, J. R., Walker, C. B., Waterbury, J. B. and Stahl, D. A.: Isolation of an autotrophic ammonia-oxidizing marine archaeon, Nature, 437(22), 543–546, doi:10.1038/nature03911, 2005.

Kraal, P., Slomp, C. P., Reed, D. C., Reichart, G.-J. and Poulton, S. W.: Sedimentary phosphorus and iron cycling in and below the oxygen minimum zone of the northern Arabian Sea, Biogeosciences, 9(7), 2603–2624, doi:10.5194/bg-9-2603-2012, 2012.

Kubo, K., Lloyd, K. G., F Biddle, J., Amann, R., Teske, A. and Knittel, K.: Archaea of the Miscellaneous Crenarchaeotal Group are abundant, diverse and widespread in marine sediments, ISME J., 6(10), 1949–1965, doi:10.1038/ismej.2012.37, 2012.

De La Torre, J. R., Walker, C. B., Ingalls, A. E., Könneke, M. and Stahl, D. A.: Cultivation of a thermophilic ammonia oxidizing archaeon synthesizing crenarchaeol, Environ. Microbiol., 10(3), 810–818, doi:10.1111/j.1462-2920.2007.01506.x, 2008.

Langworthy, T. A., Mayberry, W. R. and Smith, P. F.: Long-chain glycerol diether and polyol dialkyl glycerol triether lipids of Sulfolobus acidocaldarius, J. Bacteriol., 119(1), 106–116, 1974.

Lazar, C. S., Baker, B. J., Seitz, K., Hyde, A. S., Dick, G. J., Hinrichs, K. U. and Teske, A. P.: Genomic evidence for distinct carbon substrate preferences and ecological niches of Bathyarchaeota in estuarine sediments, Environ. Microbiol., 18(4), 1200–1211, doi:10.1111/1462-2920.13142, 2016.

Learman, D. R., Henson, M. W., Thrash, J. C., Temperton, B., Brannock, P. M., Santos, S. R., Mahon, A. R. and Halanych, K. M.: Biogeochemical and microbial variation across 5500 km of Antarctic surface sediment implicates organic matter as a driver of benthic community structure, Front. Microbiol., 7, 1–11, doi:10.3389/fmicb.2016.00284, 2016.

Lengger, S. K., Hopmans, E. C., Reichart, G.-J., Nierop, K. G. J., Sinninghe Damsté, J. S. and Schouten, S.: Intact polar and core glycerol dibiphytanyl glycerol tetraether lipids in the Arabian Sea oxygen minimum zone. Part II: Selective preservation and degradation in sediments and consequences for the TEX86, Geochim. Cosmochim. Acta, 98, 244–258, doi:10.1016/j.gca.2012.05.003, 2012.

Lengger, S. K., Hopmans, E. C., Sinninghe Damsté, J. S. and Schouten, S.: Impact of sedimentary degradation and deep water column production on GDGT abundance and distribution in surface sediments in the Arabian Sea: Implications for the TEX86 paleothermometer, Geochim. Cosmochim. Acta, 142, 386–399, doi:10.1016/j.gca.2014.07.013, 2014.

Lincoln, S. a, Wai, B., Eppley, J. M., Church, M. J., Summons, R. E. and DeLong, E. F.: Planktonic Euryarchaeota are a significant source of archaeal tetraether lipids in the ocean., Proc. Natl. Acad. Sci. U. S. A., 111(27), 9858–63, doi:10.1073/pnas.1409439111, 2014a.

Lincoln, S. a, Wai, B., Eppley, J. M., Curch, M. J., Summons, R. E. and Delong, E. F.: Reply to Schouten et al .: Marine Group II planktonic Euryarchaeota are significant contributors to tetraether lipids in the ocean, Proc. Natl. Acad. Sci., 111(41), 4286, doi:10.1073/pnas.1416736111, 2014b.

Lipp, J. S. and Hinrichs, K.-U.: Structural diversity and fate of intact polar lipids in marine sediments, Geochim.

Cosmochim. Acta, 73(22), 6816–6833, doi:10.1016/j.gca.2009.08.003, 2009.

Lipp, J. S., Morono, Y., Inagaki, F. and Hinrichs, K.-U.: Significant contribution of Archaea to extant biomass in marine subsurface sediments., Nature, 454(August), 991–994, doi:10.1038/nature07174, 2008.

Lloyd, K. G., Schreiber, L., Petersen, D. G., Kjeldsen, K. U., Lever, M. a, Steen, A. D., Stepanauskas, R., Richter, M., Kleindienst, S., Lenk, S., Schramm, A. and Jørgensen, B. B.: Predominant archaea in marine sediments degrade detrital proteins., Nature, 496(7444), 215–8, doi:10.1038/nature12033, 2013.

Logemann, J., Graue, J., Köster, J., Engelen, B., Rullkötter, J. and Cypionka, H.: A laboratory experiment of intact polar lipid degradation in sandy sediments, Biogeosciences, 8(9), 2547–2560, doi:10.5194/bg-8-2547-2011, 2011.

Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, Buchner, A., Lai, T., Steppi, S., Jobb, G., Förster, W., Brettske, I., Gerber, S., Ginhart, A. W., Gross, O., Grumann, S., Hermann, S., Jost, R., König, A., Liss, T., Lüssmann, R., May, M., Nonhoff, B., Reichel, B., Strehlow, R., Stamatakis, A., Stuckmann, N., Vilbig, A., Lenke, M., Ludwig, T., Bode, A. and Schleifer, K.-H.: ARB: a software environment for sequence data., Nucleic Acids Res., 32(4), 1363–71, doi:10.1093/nar/gkh293, 2004.

Massana, R., Castresana, J., Balague, V., Guillou, L., Romari, K., Groisillier, A., Valentin, K. and Pedros-Alio, C.: Phylogenetic and Ecological Analysis of Novel Marine Stramenopiles, Appl. Environ. Microbiol., 70(6), 3528–3534, doi:10.1128/AEM.70.6.3528, 2004.

Meador, T. B., Bowles, M., Lazar, C. S., Zhu, C., Teske, A. and Hinrichs, K.-U.: The archaeal lipidome in estuarine sediment dominated by members of the Miscellaneous Crenarchaeotal Group, Environ. Microbiol., 17(7), 2441–2458, doi:10.1111/1462-2920.12716, 2015.

Meng, J., Xu, J., Qin, D., He, Y., Xiao, X. and Wang, F.: Genetic and functional properties of uncultivated MCG archaea assessed by metagenome and gene expression analyses, ISME J., 8(3), 650–659, doi:10.1038/ismej.2013.174, 2014.

Moore, E. K., Villanueva, L., Hopmans, E. C., Rijpstra, W. I. C., Mets, A., Dedysh, S. N. and Sinninghe Damsté, J. S.: Abundant trimethylornithine lipids and specific gene sequences are indicative of planctomycete importance at the oxic/anoxic interface in sphagnum-dominated northern wetlands, Appl. Environ. Microbiol., 81(18), 6333–6344, doi:10.1128/AEM.00324-15, 2015.

Nierop, K. G. J., Reichart, G., Veld, H. and Sinninghe Damsté, J. S.: The influence of oxygen exposure time on the composition of macromolecular organic matter as revealed by surface sediments on the Murray Ridge (Arabian Sea), Geochim. Cosmochim. Acta, 206, 40–56, 2017.

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

Oni, O. E., Schmidt, F., Miyatake, T., Kasten, S., Witt, M., Hinrichs, K.-U. and Friedrich, M. W.: Microbial Communities and Organic Matter Composition in Surface and Subsurface Sediments of the Helgoland Mud Area, North Sea, Front. Microbiol., 6(November), 1–16, doi:10.3389/fmicb.2015.01290, 2015.

Park, S.-J., Park, B.-J. and Rhee, S.-K.: Comparative analysis of archaeal 16S rRNA and amoA genes to estimate the abundance and diversity of ammonia-oxidizing archaea in marine sediments, Extremophiles, 12(4), 605–615, doi:10.1007/s00792-008-0165-7, 2008.

Pitcher, A., Rychlik, N., Hopmans, E. C., Spieck, E., Rijpstra, W. I. C., Ossebaar, J., Schouten, S., Wagner, M. and Sinninghe Damsté, J. S.: Crenarchaeol dominates the membrane lipids of Candidatus Nitrososphaera gargensis, a thermophilic group I.1b Archaeon, ISME J., 4(4), 542–52, doi:10.1038/ismej.2009.138, 2010.

Pitcher, A., Villanueva, L., Hopmans, E. C., Schouten, S., Reichart, G.-J. and Sinninghe Damsté, J. S.: Niche segregation of ammonia-oxidizing archaea and anammox bacteria in the Arabian Sea oxygen minimum zone., ISME J., 5(12), 1896–904, doi:10.1038/ismej.2011.60, 2011.

Podar, M., Makarova, K. S., Graham, D. E., Wolf, Y. I., Koonin, E. V and Reysenbach, A.-L.: Insights into archaeal evolution and symbiosis from the genomes of a nanoarchaeon and its inferred crenarchaeal host from Obsidian Pool, Yellowstone National Park., Biol. Direct, 8(1), 9, doi:10.1186/1745-6150-8-9, 2013.

Qin, W., Carlson, L. T., Armbrust, E. V., Devol, A. H., Moffett, J. W., Stahl, D. A. and Ingalls, A. E.: Confounding effects of oxygen and temperature on the TEX_{86} signature of marine Thaumarchaeota, Proc. Natl. Acad. Sci., 112(35), 10979–10984, doi:10.1073/pnas.1501568112, 2015.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. and Glockner, F. O.: The SILVA ribosomal RNA gene database project: improved data processing and web-based tools, Nucleic Acids Res., 41(November 2012), D590–D596, doi:10.1093/nar/gks1219, 2013.

Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N. N., Anderson, I. J., Cheng, J.-F., Darling, A. E., Malfatti, S., Swan, B. K., Gies, E. a, Dodsworth, J. a, Hedlund, B. P., Tsiamis, G., Sievert, S. M., Liu, W.-T., Eisen, J. a, Hallam, S. J.,

Kyrpides, N. C., Stepanauskas, R., Rubin, E. M., Hugenholtz, P. and Woyke, T.: Insights into the phylogeny and coding potential of microbial dark matter., Nature, 499(7459), 431–437, doi:10.1038/nature12352, 2013.

De Rosa, M. and Gambacorta, A.: The lipids of archaebacteria, The lipids of archaeabacteria, 27, 153–175, 1988.

Rossel, P. E., Lipp, J. S., Fredricks, H. F., Arnds, J., Boetius, A., Elvert, M. and Hinrichs, K. U.: Intact polar lipids of anaerobic methanotrophic archaea and associated bacteria, Org. Geochem., 39(8), 992–999, doi:10.1016/j.orggeochem.2008.02.021, 2008.

Saitou N, N. M.: The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees', Mol. Biol. Evol., 4(4), 406–425, doi:citeulike-article-id:93683, 1987.

Schleper, C., Jurgens, G. and Jonuscheit, M.: Genomic studies of uncultivated archaea, Nat. Rev. Microbiol., 3(6), 479–488, doi:10.1038/nrmicro1159, 2005.

Schouten, S., Hopmans, E. C., Baas, M., Boumann, H., Standfest, S., Könneke, M., Stahl, D. a. and Sinninghe Damsté, J. S.: Intact membrane lipids of "Candidatus Nitrosopumilus maritimus," a cultivated representative of the cosmopolitan mesophilic group I crenarchaeota, Appl. Environ. Microbiol., 74(8), 2433–2440, doi:10.1128/AEM.01709-07, 2008.

Schouten, S., Middelburg, J. J., Hopmans, E. C. and Sinninghe Damsté, J. S.: Fossilization and degradation of intact polar lipids in deep subsurface sediments: A theoretical approach, Geochim. Cosmochim. Acta, 74(13), 3806–3814, doi:10.1016/j.gca.2010.03.029, 2010.

Schouten, S., Pitcher, A., Hopmans, E. C., Villanueva, L., van Bleijswijk, J. and Sinninghe Damsté, J. S.: Intact polar and core glycerol dibiphytanyl glycerol tetraether lipids in the Arabian Sea oxygen minimum zone: I. Selective preservation and degradation in the water column and consequences for the TEX86, Geochim. Cosmochim. Acta, 98, 228–243, doi:10.1016/j.gca.2012.05.002, 2012.

Schouten, S., Hopmans, E. C. and Sinninghe Damsté, J. S.: The organic geochemistry of glycerol dialkyl glycerol tetraether lipids: A review, Org. Geochem., 54, 19–61, doi:10.1016/j.orggeochem.2012.09.006, 2013.

Schouten, S., Villanueva, L., Hopmans, E. C., van der Meer, M. T. J. and Sinninghe Damsté, J. S.: Are Marine Group II Euryarchaeota significant contributors to tetraether lipids in the ocean?, Proc. Natl. Acad. Sci. U. S. A., 111(41), E4285, doi:10.1073/pnas.1416176111, 2014.

Schubotz, F., Wakeham, S. G., Lipp, J. S., Fredricks, H. F. and Hinrichs, K.-U.: Detection of microbial biomass by intact polar membrane lipid analysis in the water column and surface sediments of the Black Sea, Environ. Microbiol., 11(10), 2720–2734, doi:10.1111/j.1462-2920.2009.01999.x, 2009.

Sinninghe Damsté, J. S., Schouten, S., Hopmans, E. C., van Duin, A. C. T. and Geenevasen, J. A. J.: Crenarchaeol: the characteristic core glycerol dibiphytanyl glycerol tetraether membrane lipid of cosmopolitan pelagic crenarchaeota, J. Lipid Res., 43, 1641–1651, doi:10.1194/jlr.M200148-JLR200, 2002.

Sinninghe Damsté, J. S., Rijpstra, W. I. C., Hopmans, E. C., Jung, M.-Y., Kim, J.-G., Rhee, S.-K., Stieglmeier, M. and Schleper, C.: Intact Polar and Core Glycerol Dibiphytanyl Glycerol Tetraether Lipids of Group I.1a and I.1b Thaumarchaeota in Soil, Appl. Environ. Microbiol., 78(19), 6866–6874, doi:10.1128/AEM.01681-12, 2012.

Spang, A., Caceres, E. F. and Ettema, T. J. G.: Genomic exploration of the diversity, ecology, and evolution of the archaeal domain of life, Science (80-.)., 357(6351), doi:10.1126/science.aaf3883, 2017.

Stetter, K. O., Fiala, G., Huber, G., Huber, R. and Segerer, a: Hyperthermophilic microorganisms, FEMS Microbiol. Rev., 75, 117–124, doi:10.1111/j.1574-6968.1990.tb04089.x, 1990.

Sturt, H. F., Summons, R. E., Smith, K., Elvert, M. and Hinrichs, K.-U.: Intact polar membrane lipids in prokaryotes and sediments deciphered by high-performance liquid chromatography/electrospray ionization multistage mass spectrometry—new biomarkers for biogeochemistry and microbial ecology, Rapid Commun. Mass Spectrom., 18(6), 617–628, doi:10.1002/rcm.1378, 2004.

Sugai, A., Sakuma, R., Fukuda, I., Kurosawa, N. and Itoh, Y. H.: The Structure of the Core Polyol of the Ether Lipids from Sulfolobus acidocaldarius, Lipids, 30(4), 339–344, doi:10.1007/BF02536042, 1995.

Teske, A.: Marine deep sediment microbial communities., 2013.

Teske, A. and Sørensen, K. B.: Uncultured archaea in deep marine subsurface sediments: have we caught them all?, ISME J., 2(1), 3–18, doi:10.1038/ismej.2007.90, 2008.

Villanueva, L., Schouten, S. and Sinninghe Damsté, J. S.: Depth-related distribution of a key gene of the tetraether lipid biosynthetic pathway in marine Thaumarchaeota., Environ. Microbiol., doi:10.1111/1462-2920.12508, 2014.

Villanueva, L., Schouten, S. and Sinninghe Damsté, J. S.: Phylogenomic analysis of lipid biosynthetic genes of Archaea shed light on the "lipid divide," Environ. Microbiol., 19(1), 54–69, doi:10.1111/1462-2920.13361, 2017.

Waters, E., Hohn, M. J., Ahel, I., Graham, D. E., Adams, M. D., Barnstead, M., Beeson, K. Y., Bibbs, L., Bolanos, R.,

693 Keller, M., Kretz, K., Lin, X., Mathur, E., Ni, J., Podar, M., Richardson, T., Sutton, G. G., Simon, M., Soll, D., Stetter,
694 K. O., Short, J. M. and Noordewier, M.: The genome of Nanoarchaeum equitans: insights into early archaeal evolution
695 and derived parasitism., Proc. Natl. Acad. Sci. U. S. A., 100(22), 12984–8, doi:10.1073/pnas.1735403100, 2003.

696 Xie, S., Lipp, J. S., Wegener, G., Ferdelman, T. G. and Hinrichs, K.-U.: Turnover of microbial lipids in the deep
697 biosphere and growth of benthic archaeal populations, Proc. Natl. Acad. Sci., 110(15), 6010–6014,
698 doi:10.1073/pnas.1218569110, 2013.

699 Yakimov, M. M., Cono, V. La, Smedile, F., DeLuca, T. H., Juárez, S., Ciordia, S., Fernández, M., Albar, J. P., Ferrer,
700 M., Golyshin, P. N. and Giuliano, L.: Contribution of crenarchaeal autotrophic ammonia oxidizers to the dark primary
701 production in Tyrrhenian deep waters (Central Mediterranean Sea)., ISME J., 5(6), 945–61,
702 doi:10.1038/ismej.2010.197, 2011.

703

704

Figure legends

Fig. 1. (**A**) Relative abundances of the IPL-GDGTs (sum of the IPL-types MH, DH and HPH) for the different core GDGTs in the surface (0-0.5 cm) and subsurface sediments (10-12 cm) and **(B)** the archaeal community composition as revealed by 16S rRNA gene reads (with average abundance above of > 1%) in the surface sediments at 885, 1306, 2470, and 3003 mbsl and in the subsurface sediments at 885 and 1306 mbsl.

Fig. 2. Maximum likelihood phylogenetic tree of the archaeal groups MCG+C3 (modified from Fillol et al., 2015). Extracted OTUs from the Arabian Sea sediments assigned as MCG were inserted in the tree. The number of detected reads per OTU per samples are indicated. Per MCG subgroup the relative abundance is given as detected at the different stations and sediments depths, this is also noted in Table 4. Scale bar represents a 2% sequence dissimilarity.

Fig. 3. Maximum likelihood phylogenetic tree of MG-I OTUs recovered within the sediment based on the 16S rRNA gene (colored in blue). Sequences from cultured representatives of Thaumarchaeota MG-I are indicated in red. Environmental sequences of MG-I members are indicated in black with their origin specified. The relative abundances of the various OTUs are listed in Table 4. Scale bar represents a 2% sequence dissimilarity.

Fig. 4. Maximum likelihood phylogenetic tree of *amo*A gene coding sequences recovered from surface (S; 0-0.5 cm) and subsurface (SS; 10-12 cm) sediments (colored in blue) at 885 mbsl, 1306 mbsl and 3003 mbsl (155 clones). *Amo*A gene coding sequences recovered from SPM (colored in orange) at 170 mbsl (28 clones), SPM at 1050 (25 clones) reported by Villanueva et al. (2014). ** indicates *amo*A gene sequences recovered from surface sediments at 3003 mbsl previously reported in Villanueva et al., (2015). Scale bar represents a 2% sequence dissimilarity.

Fig. 5. Abundance of Thaumarchaeotal 16S rRNA (**A,C**) and *amo*A (**B,D**) gene fragment copies per gram of dry weight in the surface sediment (0-0.5 cm) (**A,B**) and the subsurface sediment (10-12 cm) (**C,D**). Black bars indicate the amount of DNA 16S rRNA or *amo*A gene fragment copies and the gray bars indicate the RNA (gene transcripts) of 16S rRNA or *amo*A gene fragment copies. Error bars indicate standard deviation based on $n = 3$ experimental replicates.

728 **Table 1. Bottom water temperature and oxygen concentration, oxygen penetration depth in the sediment, and**
729 **TOC content and pore water composition of the surface (0-0.5 cm) sediment[a]**

Station (mbsl)	T (°C)	BWO ($\mu mol \cdot L^{-1}$)	OPD (mm)	TOC (wt %)	NH_4^+ (µM)	NO_2^- (µM)	NO_3^- (µM)	HPO_4^{2-} (µM)
885	10	2.0	0.1	5.6 (± 0.2)	2	1.2	1.3	9.2
1306	6.7	14.3	2.9	2.9 (± 0.1)	2.6*	0.1*	36.2*	5.6
2470	2.1	63.8	9.8	0.8 (± 0.1)	-[b]	-	-	-
3003	1.4	82.9	19	0.7 (± 0.1)	55.6	8.3	46.2	3.8

[a] Data from Kraal et al. (2012) and Lengger et al. (2014)

[b] no data available

730

731

Table 2. Total IPL abundance and ~~H~~heatmap[a] of the relative abundance (%) of the detected IPLs and sum (not color coded) per IPL-GDGT ~~. Response areas of the summed IPLs are given per sample~~in the sediments studied.

~~Sample~~Sediment	Depth (mbsl)	GDGT-0 MH	GDGT-0 DH I[b]	GDGT-0 DH II[b]	GDGT-0 HCP[c]	GDGT-0 HPH	GDGT-0 Sum	GDGT-1 MH	GDGT-1 DH I[b]	GDGT-1 HPH	GDGT-1 Sum	GDGT-2 MH	GDGT-2 DH I[b]	GDGT-2 HPH	GDGT-2 Sum
Surface (0-0.5 cm)	885	0.3	ND[d]	ND	ND	ND	0.3	0.1	1.6	ND	1.7	0.1	29.5	ND	29.6
	1306	1.1	ND	ND	ND	36.6	37.6	0.1	1.5	0.2	1.7	ND	15.4	ND	15.4
	2470	0.2	0.1	ND	ND	71.5	71.9	0.0	0.1	0.4	0.5	ND	0.8	ND	0.8
	3003	0.5	0.1	ND	ND	80.3	80.8	ND	0.2	ND	0.2	ND	0.9	ND	0.9
Subsurface (10-12 cm)	885	0.3	ND	7.8	1.6	2.1	11.9	0.1	1.7	0.1	1.9	0.2	27.0	ND	27.1
	1306	2.2	0.9	1.8	0.4	2.1	7.4	0.2	6.7	ND	6.9	0.1	29.7	ND	29.7
	2470	4.3	2.7	ND	ND	18.6	25.6	0.1	5.8	ND	5.9	ND	23.2	ND	23.2
	3003	9.1	3.4	ND	ND	13.0	25.5	0.2	4.3	ND	4.6	ND	21.9	ND	21.9

~~Sample~~Sediment	Depth (mbsl)	GDGT-3 MH	GDGT-3 DH I[b]	GDGT-3 HPH	GDGT-3 Sum	GDGT-4 MH	GDGT-4 DH I[b]	GDGT-4 HPH	GDGT-4 Sum	Crenarchaeol MH	Crenarchaeol DH I[b]	Crenarchaeol HPH	Crenarchaeol Sum	IPL a~~A~~bundance [au . g sed dw^{-1}][e]
Surface (0-0.5 cm)	885	ND	17.8	ND	17.8	ND	6.1	ND	6.1	1.3	43.1	0.3	44.6	2.7E+09
	1306	0.0	6.9	ND	6.9	ND	2.7	ND	2.7	1.4	15.5	18.7	35.6	1.2E+10
	2470	ND	0.2	ND	0.2	ND	0.0	ND	0.0	0.2	0.6	25.8	26.6	2.2E+09
	3003	ND	0.4	ND	0.4	ND	0.0	ND	0.0	0.4	0.2	17.1	17.6	1.3E+10
Subsurface (10-12 cm)	885	0.1	15.9	ND	15.9	ND	9.4	ND	9.4	1.1	31.1	1.5	33.8	2.0E+09
	1306	0.0	14.5	ND	14.5	ND	6.1	ND	6.1	2.7	32.4	0.4	35.5	2.2E+09
	2470	ND	9.6	ND	9.6	ND	2.9	ND	2.9	3.5	28.3	1.0	32.8	7.8E+08
	3003	ND	9.7	ND	9.7	ND	5.6	ND	5.6	8.2	23.9	0.6	32.7	1.6E+09

[a] Green colors indicate a low relative abundance, red colors indicate a high relative abundance

[b] DH isomers were detected as a GDGT with a glycosidically-bound hexose moieity on both ends of the core (I) and with one glycosidically-bound dihexose moiety on one end (II).

[c] HCP is an IPL-type with an ether-bound cyclopentanetetraol moiety on one end and an hexose moiety on the other (previously reported as GDNT; e.g. De Rosa and Gambacorta, 1988; Sturt et al., 2004).

[d] ND = not detected

[e] Response area of summed IPLs given in au (arbitrary units) per gram of dry weight (dw) sediment.

Table 3. Relative abundance of IPL-GDGTs grouped by polar head group[a].

Sample	Depth (mbsl)	MH	DH	HCP	HPH
Surface (0-0.5 cm)	885	1.7%	98.1%	0.0%	0.3%
	1306	2.6%	42.0%	0.0%	55.4%
	2470	0.5%	1.8%	0.0%	97.7%
	3003	0.8%	1.8%	0.0%	97.4%
~~Deep~~Subsurface (10-12 cm)	885	1.8%	92.9%	1.6%	3.7%
	1306	5.2%	91.9%	0.4%	2.5%
	2470	7.9%	72.6%	0.0%	19.6%
	3003	17.6%	68.8%	0.0%	13.6%

[a]Polar head group types detected: MH = monohexose, DH = dihexose, both isomers combined, HCP = monohexose and cyclopentanetetraol, HPH = monohexose and phosphohexose.

[b] ND = not detected

747 **Table 4. Relative abundance (in %) of -MCG- and C3--assigned 16S rRNA gene reads relative to total archaeal**
748 **reads and distribution (in %) of various subgroups for a station within and a station just below the OMZ**

749

Subgroup	885 mbsl		1306 mbsl	
	Surface (0—0.5 cm)	Subsurface (10-12 cm)	Surface (0-0.5 cm)	Subsurface (10-12 cm)
Total	30.5	47.5	1.3	48.8
1		4.6		8.6
2		9.7		10.9
3		<1		2.3
4		<1		
5b		<1		
8	2.3	33.6		10.3
10		13.4		4.0
12	13.6	7.7		8.0
13		1.2		2.3
14	2.3	3.1		10.9
15	77.3	19.6	100	34.3
17	4.5	5.7		8.6

750

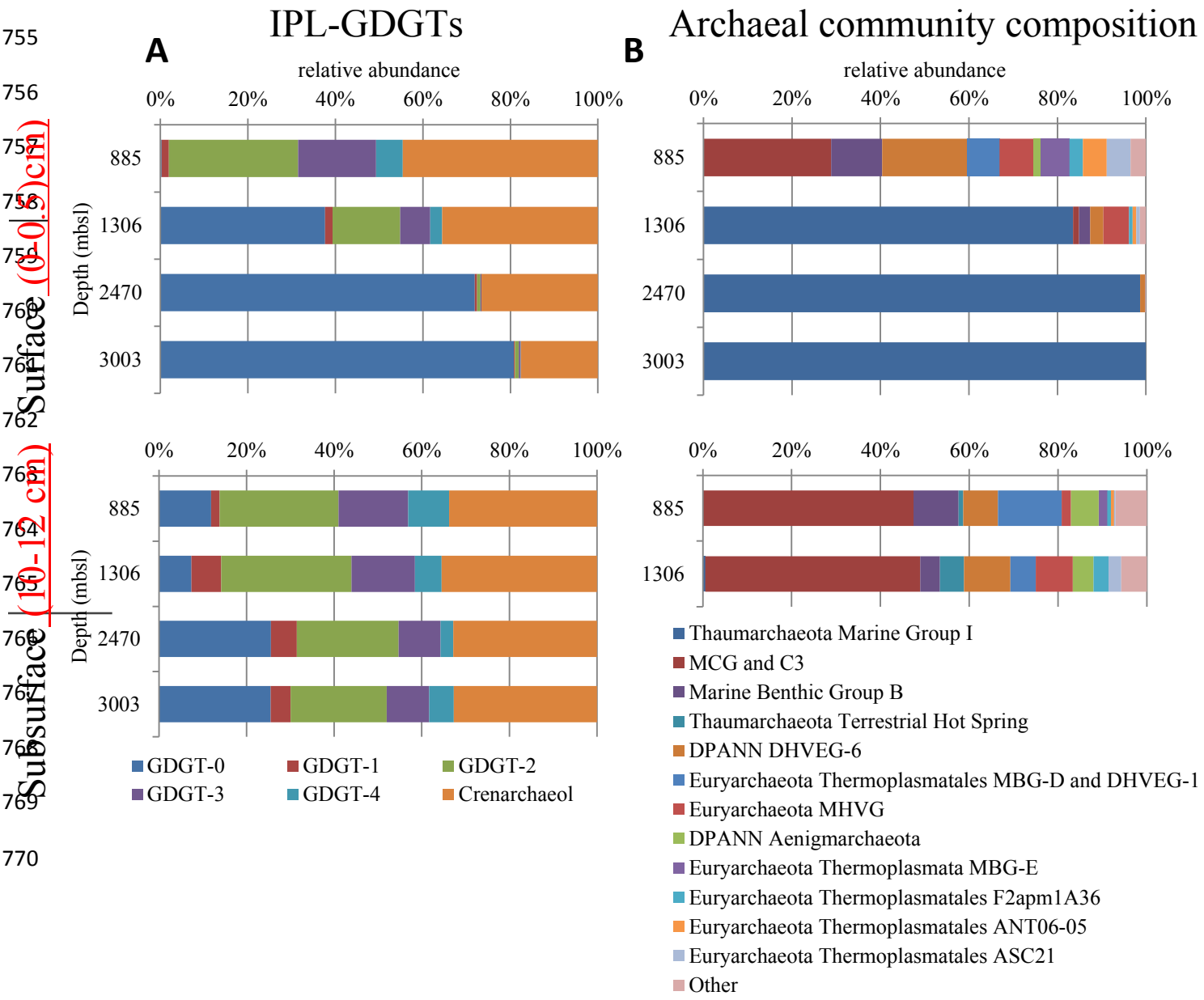
751 **Table 5.** Total Thaumarchaeota MG-I 16S rRNA gene reads and distribution per OTU (%) in surface sediments.

	Depth (mbsl)			
	885	1306	2470	3003
Total reads	0	915	1341	1305
OTU ID #1	n.a.[a]	4.3	2.5	3.0
OTU ID #2	n.a.	3.9	8.1	13.6
OTU ID #3	n.a.	43.6	67.6	61.8
OTU ID #4	n.a.	35.1	1.6	0
OTU ID #5	n.a.	3.3	4.7	2.1

752 [a] n.a. = not applicable

753

Fig. 1.

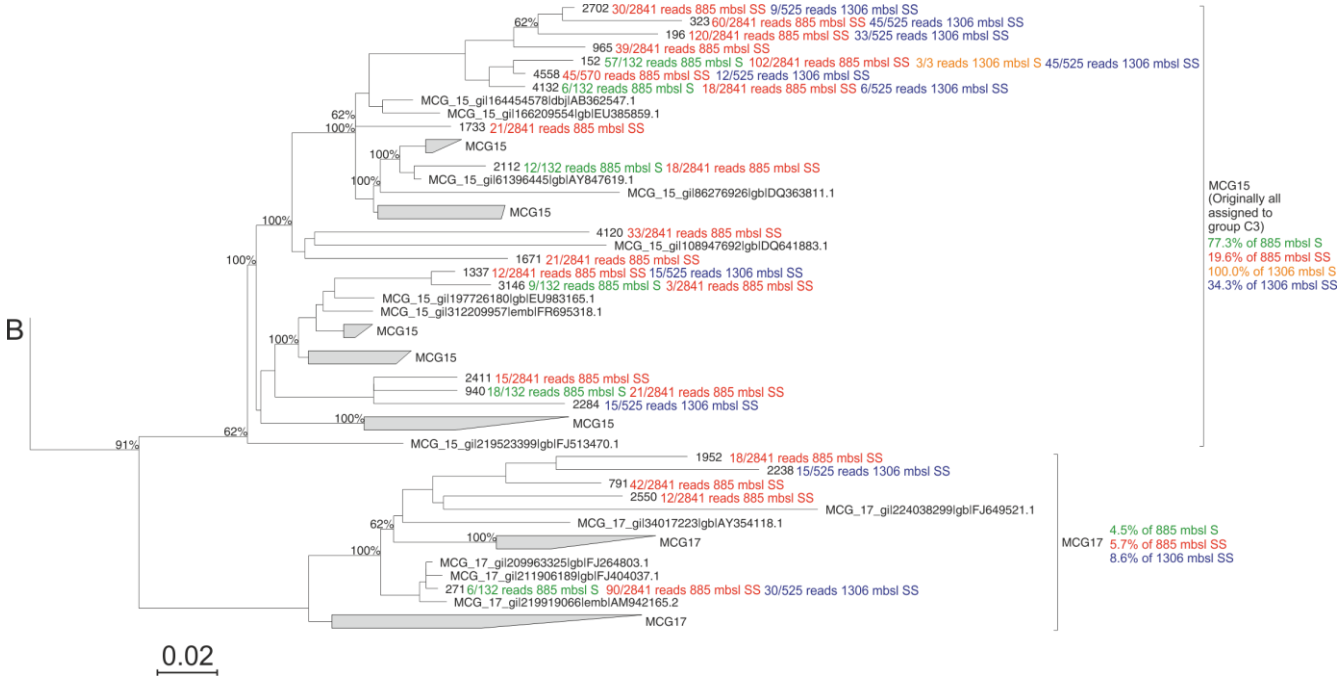

IPL-GDGTs
Archaeal community composition
A
B
relative abundance
0%
20%
40%
60%
80%
100%
Depth (mbsl)
885
1306
2470
3003
Surface (0-0.8 cm)
Subsurface (10-12 cm)
GDGT-0
GDGT-1
GDGT-2
GDGT-3
GDGT-4
Crenarchaeol
Thaumarchaeota Marine Group I
MCG and C3
Marine Benthic Group B
Thaumarchaeota Terrestrial Hot Spring
DPANN DHVEG-6
Euryarchaeota Thermoplasmatales MBG-D and DHVEG-1
Euryarchaeota MHVG
DPANN Aenigmarchaeota
Euryarchaeota Thermoplasmata MBG-E
Euryarchaeota Thermoplasmatales F2apm1A36
Euryarchaeota Thermoplasmatales ANT06-05
Euryarchaeota Thermoplasmatales ASC21
Other

771 **Fig. 2.**


MCG4
MCG4 <1% of 885 mbsl SS
2231 15/2841 reads 885 mbsl SS
MCG9up
MCG9
176 162/2841 reads 885 mbsl SS 27/525 reads 1306 mbsl SS
MCG_2_DQ301965
MCG_2_AB109879
MCG2
9.7% of 885 mbsl SS
10.9% of 1306 mbsl SS
3481 12/2841 reads 885 mbsl SS
2889 12/2841 reads 885 mbsl SS
309 42/2841 reads 885 mbsl SS 3/525 reads 1306 mbsl SS
4457 48/2841 reads 885 mbsl SS 27/525 reads 1306 mbsl SS
MCG11
MCG12
121 18/132 reads 885 mbsl S 219/2841 reads 885 mbsl SS 42/525 reads 1306 mbsl SS
MCG_12_gil224038293|gb|FJ649515.1
MCG_12_gil262293344|gb|GQ994171.1
MCG_12_gil259123950|gb|GQ848385.1
MCG12
13.6% of 885 mbsl S
7.7% of 885 mbsl SS
8.0% of 1306 mbsl SS
MCG8
MCG_8_gil125381510|gb|EF367607.1
83 3/132 reads 885 mbsl S 330/2841 reads 885 mbsl SS 33/525 reads 1306 mbsl SS
MCG_8_gil224038298|gb|FJ649520.1
MCG8
MCG8
MCG_8_gil108947665|gb|DQ641856.1
43 576/2841 reads 885 mbsl SS 18/525 reads 1306 mbsl SS
MCG_8_AY396665
MCG_8_gil48527083|gb|AY625574.1
MCG_8_gil39843283|gb|AY396657.1
MCG_8_gil61105732|gb|AY940581.1
MCG_8_gil290563556|gb|GU475232.1
MCG_8_gil312209975|emb|FR695336.1
MCG8
MCG8
2.3% of 885 mbsl S
33.6% of 885 mbsl SS
10.3% of 1306 mbsl SS
3585 48/2841 reads 885 mbsl SS 3/525 reads 1306 mbsl SS
MCG_8_gil269799371|gb|FJ853479.1
MCG_8_gil229562917|gb|FJ901676.1
MCG_8_gil220937143|gb|FJ484638.1
MCG3
1344 15/2841 reads 885 mbsl SS 12/525 reads 1306 mbsl SS
MCG3
<1% of 885 mbsl SS
2.3% of 1306 mbsl SS
628 54/525 reads 1306 mbsl SS
MCG_14_gil189909385|gb|EU729296.1
2780 12/2841 reads 885 mbsl SS
MCG_14_gil229563149|gb|FJ901908.1
1255 3/132 reads 885 mbsl S 24/2841 reads 885 mbsl SS 3/525 reads 1306 mbsl SS
MCG_14_gil206149491|gb|FJ175640.1
1985 18/2841 reads 885 mbsl SS
2790 12/2841 reads 885 mbsl SS
2556 12/2841 reads 885 mbsl SS
MCG_14_gil120564674|gb|EF125496.1
MCG14
MCG_14_gil224925070|gb|FJ746533.1
MCG14
2.3% of 885 mbsl S
3.1% of 885 mbsl SS
10.9% of 1306 mbsl SS
MCG6
MCG5b
1925 18/2841 reads 885 mbsl SS
MCG5b
<1% of 885 mbsl SS
MCG5a
MCG5a
2189 15/2841 reads 885 mbsl SS
MCG_1_HM483675
916 30/2841 reads 885 mbsl SS 9/525 reads 1306 mbsl SS
2881 3/2841 reads 885 mbsl SS 9/525 reads 1306 mbsl SS
MCG_1_FR695325
2371 12/2841 reads 885 mbsl SS
MCG_1_GU972250
1600 27/2841 reads 885 mbsl SS
1137 6/2841 reads 885 mbsl SS 27/525 reads 1306 mbsl SS
MCG1
MCG_1_FJ902682
MCG1
4.6% of 885 mbsl SS
8.6% of 1306 mbsl SS
881 39/2841 reads 885 mbsl SS
MCG7
245 129/2841 reads 885 mbsl SS
4182 15/2841 reads 885 mbsl SS
1894 18/2841 reads 885 mbsl SS
2002 24/2841 reads 885 mbsl SS
1238 33/2841 reads 885 mbsl SS
1821 21/2841 reads 885 mbsl SS
MCG_10_gil152031995|gb|EF680211.1
1400 6/2841 reads 885 mbsl SS 21/525 reads 1306 mbsl SS
550 60/2841 reads 885 mbsl SS
MCG_10_gil307712883|gb|HM244153.1
832 42/2841 reads 885 mbsl SS
MCG_10_gil164373782|gb|EU329775.1
MCG_10_gil46804828|dbj|AB178094.1
MCG_10_gil50830558|gb|AY592495.1
MCG_10_gil220938012|gb|FJ485507.1
MCG_10_gil229562713|gb|FJ901472.1
MCG_10_gil39843293|gb|AY396667.1
1112 33/2841 reads 885 mbsl SS
MCG_10_gil291065908|gb|GU567003.1
MCG10
13.4% of 885 mbsl SS
4.0% of 1306 mbsl SS
MCG16
MCG13
2376 9/2841 reads 885 mbsl SS 6/525 reads 1306 mbsl SS
MCG13
MCG13
MCG_13_gil83722730|gb|DQ302019.1
1223 24/2841 reads 885 mbsl SS 6/525 reads 1306 mbsl SS
MCG_13_gil259123994|gb|GQ848429.1
MCG13
1.2% of 885 mbsl SS
2.3% of 1306 mbsl SS
A
0.02

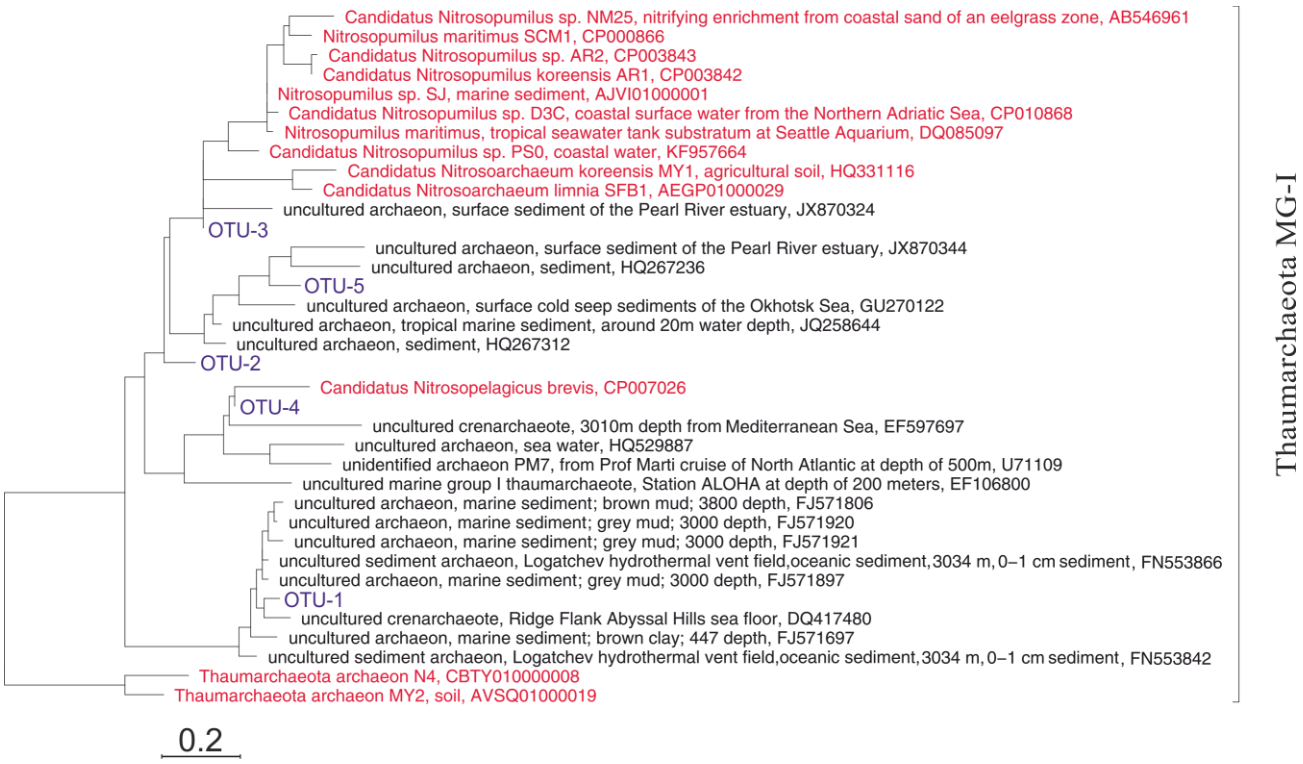
772

773


B
2702 30/2841 reads 885 mbsl SS 9/525 reads 1306 mbsl SS
323 60/2841 reads 885 mbsl SS 45/525 reads 1306 mbsl SS
196 120/2841 reads 885 mbsl SS 33/525 reads 1306 mbsl SS
965 39/2841 reads 885 mbsl SS
152 57/132 reads 885 mbsl S 102/2841 reads 885 mbsl SS 3/3 reads 1306 mbsl S 45/525 reads 1306 mbsl SS
4558 45/570 reads 885 mbsl SS 12/525 reads 1306 mbsl SS
4132 6/132 reads 885 mbsl S 18/2841 reads 885 mbsl SS 6/525 reads 1306 mbsl SS
MCG_15_gil164454578ldbjlAB362547.1
MCG_15_gil166209554lgblEU385859.1
1733 21/2841 reads 885 mbsl SS
MCG15
2112 12/132 reads 885 mbsl S 18/2841 reads 885 mbsl SS
MCG_15_gil61396445lgblAY847619.1
MCG_15_gil86276926lgblDQ363811.1
MCG15
4120 33/2841 reads 885 mbsl SS
MCG_15_gil108947692lgblDQ641883.1
1671 21/2841 reads 885 mbsl SS
1337 12/2841 reads 885 mbsl SS 15/525 reads 1306 mbsl SS
3146 9/132 reads 885 mbsl S 3/2841 reads 885 mbsl SS
MCG_15_gil197726180lgblEU983165.1
MCG_15_gil312209957lemblFR695318.1
MCG15
MCG15
2411 15/2841 reads 885 mbsl SS
940 18/132 reads 885 mbsl S 21/2841 reads 885 mbsl SS
2284 15/525 reads 1306 mbsl SS
MCG15
MCG_15_gil219523399lgblFJ513470.1
MCG15 (Originally all assigned to group C3)
77.3% of 885 mbsl S
19.6% of 885 mbsl SS
100.0% of 1306 mbsl S
34.3% of 1306 mbsl SS
1952 18/2841 reads 885 mbsl SS
2238 15/525 reads 1306 mbsl SS
791 42/2841 reads 885 mbsl SS
2550 12/2841 reads 885 mbsl SS
MCG_17_gil224038299lgblFJ649521.1
MCG_17_gil34017223lgblAY354118.1
MCG17
MCG_17_gil209963325lgblFJ264803.1
MCG_17_gil211906189lgblFJ404037.1
271 6/132 reads 885 mbsl S 90/2841 reads 885 mbsl SS 30/525 reads 1306 mbsl SS
MCG_17_gil219919066lemblAM942165.2
MCG17
MCG17 4.5% of 885 mbsl S
5.7% of 885 mbsl SS
8.6% of 1306 mbsl SS
0.02

774

775 **Fig. 3.**

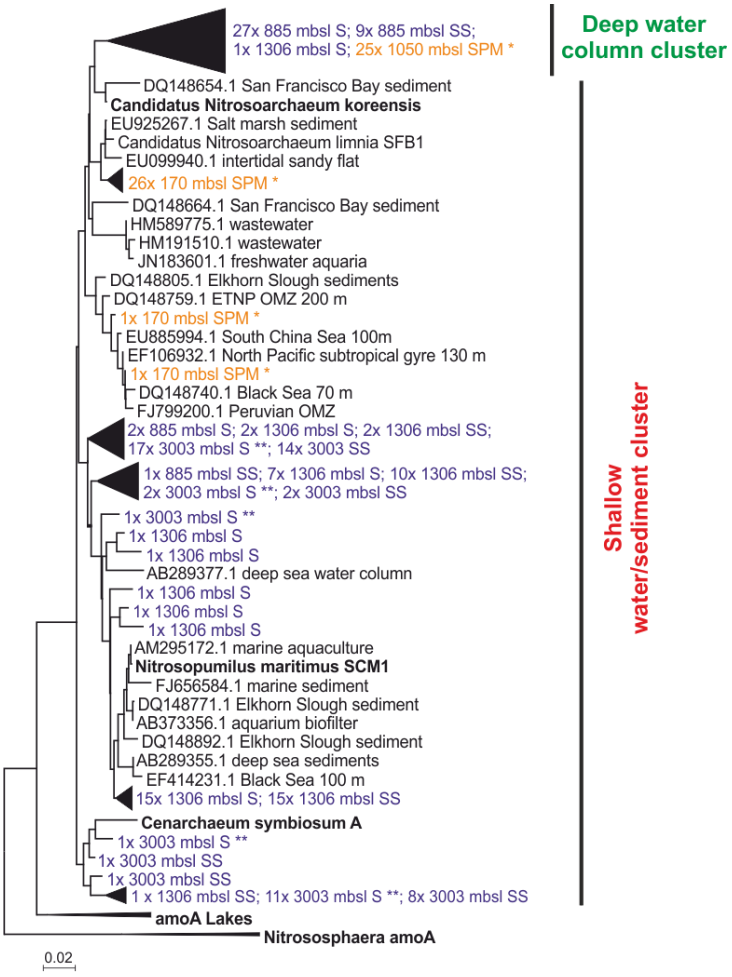


776
777

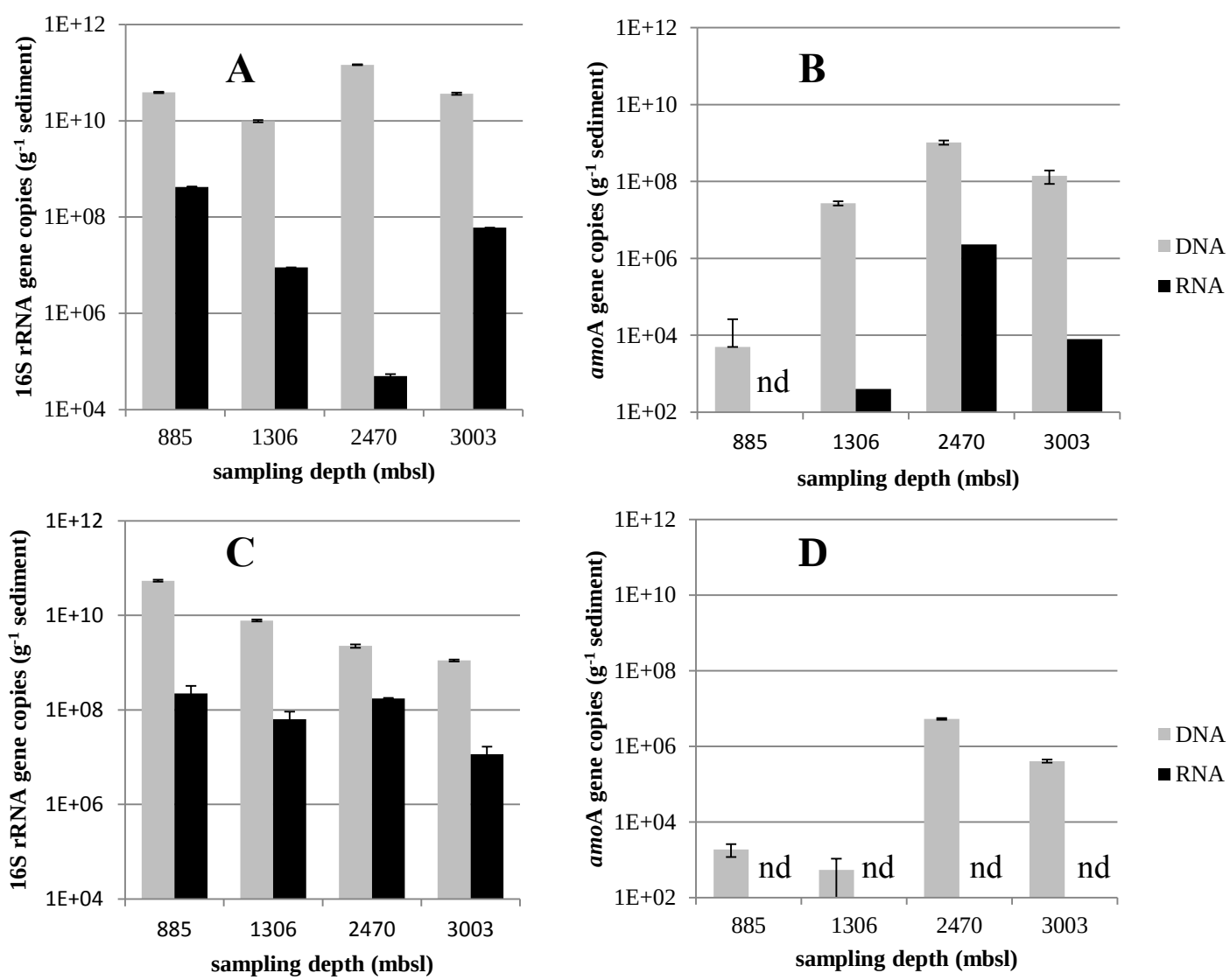
778

779 **Fig. 4.**

780


27x 885 mbsl S; 9x 885 mbsl SS;
1x 1306 mbsl S; 25x 1050 mbsl SPM *
Deep water column cluster
DQ148654.1 San Francisco Bay sediment
Candidatus Nitrosoarchaeum koreensis
EU925267.1 Salt marsh sediment
Candidatus Nitrosoarchaeum limnia SFB1
EU099940.1 intertidal sandy flat
26x 170 mbsl SPM *
DQ148664.1 San Francisco Bay sediment
HM589775.1 wastewater
HM191510.1 wastewater
JN183601.1 freshwater aquaria
DQ148805.1 Elkhorn Slough sediments
DQ148759.1 ETNP OMZ 200 m
1x 170 mbsl SPM *
EU885994.1 South China Sea 100m
EF106932.1 North Pacific subtropical gyre 130 m
1x 170 mbsl SPM *
DQ148740.1 Black Sea 70 m
FJ799200.1 Peruvian OMZ
2x 885 mbsl S; 2x 1306 mbsl S; 2x 1306 mbsl SS;
17x 3003 mbsl S **; 14x 3003 SS
1x 885 mbsl SS; 7x 1306 mbsl S; 10x 1306 mbsl SS;
2x 3003 mbsl S **; 2x 3003 mbsl SS
1x 3003 mbsl S **
1x 1306 mbsl S
1x 1306 mbsl S
AB289377.1 deep sea water column
1x 1306 mbsl S
1x 1306 mbsl S
1x 1306 mbsl S
AM295172.1 marine aquaculture
Nitrosopumilus maritimus SCM1
FJ656584.1 marine sediment
DQ148771.1 Elkhorn Slough sediment
AB373356.1 aquarium biofilter
DQ148892.1 Elkhorn Slough sediment
AB289355.1 deep sea sediments
EF414231.1 Black Sea 100 m
15x 1306 mbsl S; 15x 1306 mbsl SS
Cenarchaeum symbiosum A
1x 3003 mbsl S **
1x 3003 mbsl SS
1x 3003 mbsl SS
1 x 1306 mbsl SS; 11x 3003 mbsl S **; 8x 3003 mbsl SS
amoA Lakes
Nitrososphaera amoA
0.02
Shallow water/sediment cluster

781 **Fig. 5.**


A
B
C
D
16S rRNA gene copies (g-1 sediment)
amoA gene copies (g-1 sediment)
1E+12
1E+10
1E+08
1E+06
1E+04
1E+02
885
1306
2470
3003
sampling depth (mbsl)
nd
DNA
RNA