



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

3 Marc A. Besseling^{1*}, Ellen C. Hopmans¹, Jaap S. Sinninghe Damsté^{1,2}, and Laura Villanueva¹

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

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

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

9 **Abstract.** Benthic Archaea comprise a significant part of the total prokaryotic biomass in marine sediments. Recent
10 genomic surveys suggest they are largely involved in anaerobic processing of organic matter but the distribution and
11 abundance of these archaeal groups is still largely unknown. Archaeal membrane lipids composed of isoprenoid
12 diethers or tetraethers (glycerol dibiphytanyl glycerol tetraether, GDGT) are often used as archaeal biomarkers. Here,
13 we compare the archaeal diversity and intact polar lipid (IPL) composition in both surface (0–0.5 cm) and subsurface
14 (10–12 cm) sediments recovered within, just below, and well below the oxygen minimum zone (OMZ) of the Arabian
15 Sea. Archaeal 16S rRNA gene amplicon sequencing revealed a predominance of Thaumarchaeota (Marine Group I,
16 MG-I) in oxygenated sediments. Quantification of archaeal 16S rRNA and ammonia monooxygenase (*amoA*) of
17 Thaumarchaeota genes and their transcripts indicated the presence of an active *in situ* benthic population, which
18 coincided with a high relative abundance of hexose phosphohexose crenarchaeol, a specific biomarker for living
19 Thaumarchaeota. On the other hand, anoxic surface sediments within the OMZ and all subsurface sediments were
20 dominated by archaea belonging to the Miscellaneous Crenarchaeota Group (MCG), the Thermoplasmatales and
21 archaea of the DPANN superphylum. Members of the MCG were diverse with a dominance of subgroup MCG-12 in
22 anoxic surface sediments. This coincided with a high relative abundance of IPL GDGT-0 with an unknown polar head
23 group. Subsurface anoxic sediments were characterized by higher relative abundance of GDGT-0, 2 and 3 with
24 dihexose IPL-types, as well as GDGT-0 with a cyclopentanetetraol molecule and a hexose, as well as the presence of
25 specific MCG subgroups, suggesting that these groups could be the biological sources of these archaeal lipids.



26 INTRODUCTION

27 Archaea are ubiquitous microorganisms in the marine system (DeLong et al., 1994). They occur in diverse
28 environments, e.g. hydrothermal vents (Stetter et al., 1990), the marine water column (Karner et al., 2001; Massana et
29 al., 2004), in the underlying sediments (Lloyd et al., 2013), and well below the seafloor (Biddle et al., 2006; Lipp et al.,
30 2008), where they are considered key players in diverse biogeochemical processes (Offre et al., 2013, and references
31 cited therein). Specifically marine sediments have been shown to contain a highly diverse archaeal community (Lloyd
32 et al., 2013). The ammonia-oxidizing Thaumarchaeota of the marine group I.1a (further referred to as MG-I) is
33 probably the most widely studied archaeal group in marine sediments. However, in comparison with studies of marine
34 pelagic Thaumarchaeota, the diversity and distribution of benthic Thaumarchaeota is still quite unknown (e.g. Durbin &
35 Teske, 2010; Jorgenson et al., 2012; Learman et al., 2016). Genomic studies have revealed the existence of uncultured
36 archaeal groups other than Thaumarchaeota in marine, predominantly anoxic, sediments such as the Miscellaneous
37 Crenarchaeota Group (MCG; Meng et al., 2014), archaea of the DPANN superphylum (composed of Micrarchaeota,
38 Diapherotrites, Aenigmarchaeota, Nanohaloarchaeota, Parvarchaeota, Nanoarchaeota, Pacearchaeota and
39 Woearchaeota; Castelle et al., 2015; Rinke et al., 2013) and the Marine Benthic Group (MBG) B (Teske & Sørensen,
40 2008), and D (Lloyd et al., 2013). In the case of the archaea belonging to the groups of the MCG and MBG-D,
41 metagenomic studies suggest that they are able to degrade extracellular proteins and aromatic compounds (Lloyd et al.,
42 2013; Meng et al., 2014).

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

55 Archaeal membrane lipids are typically a variation of two main structures, *sn*-2,3-diphytanylglycerol diether (archaeol)
56 with phytanyl (C₂₀) chains in a bilayer structure, and *sn*-2,3-dibiphytanyl diglycerol tetraether (glycerol dibiphytanyl
57 glycerol tetraether, GDGT), in which the two glycerol moieties are connected by two C₄₀ isoprenoid chains, allowing
58 the formation of a monolayer membrane (Koga and Morii, 2007). GDGTs containing 0–4 cyclopentane moieties (Fig.



59 S1) are usually not exclusive to a specific archaeal group (Schouten et al., 2013) with the exception of the GDGT
60 crenarchaeol, containing 4 cyclopentane and one cyclohexane moiety, which is deemed to be exclusive to the
61 Thaumarchaeota phylum (De La Torre et al., 2008; Sinninghe Damsté et al., 2002, 2012). The newly described archaeal
62 groups detected by genetic methods are yet uncultured, therefore, their membrane lipid composition remains unknown.
63 In this study, we determined the archaeal diversity in a marine benthic system along a strong gradient in bottom water
64 oxygen concentrations and compared it with the diversity of archaeal lipids. We aimed to characterize changes in the
65 archaeal benthic community under different physicochemical conditions, as well as to provide clues on the potential
66 archaeal lipid biomarkers produced by uncultured benthic archaea. We analyzed sediments (surface 0–0.5 cm, and
67 subsurface 10–12 cm) of the Murray ridge in the Arabian Sea, which is impinged by one of the strongest present-day
68 oxygen minimum zones (OMZ). Previous studies observed changes in the diversity of archaeal lipids in the same
69 environmental setting in sediments under different oxygen and nutrient concentrations (Lengger et al., 2012; 2014). In
70 our study, we expand the repertory of archaeal lipid diversity previously detected by Lengger et al. (2012; 2014) by re-
71 analyzing the samples with High Resolution Accurate Mass/Mass spectrometry (UHPLC-HRAM MS). In addition, we
72 determined the archaeal diversity by means of 16S rRNA gene amplicon sequencing, as well as the abundance and
73 potential activity of specific archaeal groups by quantitative PCR (QPCR) of 16S rRNA and the metabolic gene coding
74 for the ammonia monooxygenase (*amoA* gene) of Thaumarchaeota.

75 MATERIAL and METHODS

76 Sampling

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

83 Lipid extraction and analysis

84 Total lipids were extracted from surface (upper 0–0.5 cm) and subsurface (10–12 cm) sediments after freeze-drying
85 using a modified Bligh and Dyer method (Bligh and Dyer, 1959) as previously described by Lengger et al. (2014). C₁₆-
86 PAF (1-O-hexadecyl-2-acetyl-sn-glycero-3-phosphocholine) was added to the extracts as an internal standard and the
87 extracts were dried under a stream of nitrogen. The extracts with the added standard were then dissolved by adding
88 solvent (hexane:isopropanol:H₂O 718:271:10 [v/v/v/v]) and filtered through a 0.45 µm, 4 mm-diameter True
89 Regenerated Cellulose syringe filter (Grace Davison, Columbia, MD, USA).



90 IPLs were analyzed according to Sturt et al. (2004) with some modifications. An Ultimate 3000 RS UHPLC, equipped
91 with thermostated auto-injector and column oven, coupled to a Q Exactive Orbitrap MS with Ion Max source with
92 heated electrospray ionization (HESI) probe (Thermo Fisher Scientific, Waltham, MA), was used. Separation was
93 achieved on a YMC-Triart Diol-HILIC column (250 x 2.0 mm, 1.9 μm particles, pore size 12 nm; YMC Co., Ltd,
94 Kyoto, Japan) maintained at 30 °C. The following elution program was used with a flow rate of 0.2 mL min⁻¹: 100% A
95 for 5 min, followed by a linear gradient to 66% A: 34% B in 20 min, maintained for 15 min, followed by a linear
96 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-
97 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}
98 (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
99 settings were as follows: sheath gas (N₂) pressure 35 (arbitrary units), auxiliary gas (N₂) pressure 10 (arbitrary units),
100 auxiliary gas (N₂) T 50 °C, sweep gas (N₂) pressure 10 (arbitrary units), spray voltage 4.0 kV (positive ion ESI),
101 capillary temperature 275 °C, S-Lens 70 V. IPLs were analyzed with a mass range of m/z 375 to 2000 (resolution
102 70,000 parts per million, ppm), followed by data dependent MS² (resolution 17,500 ppm), in which the ten most
103 abundant masses in the mass spectrum (with the exclusion of isotope peaks) were fragmented successively (stepped
104 normalized collision energy 15, 22.5, 30; isolation window 1.0 m/z). An inclusion list was used with a mass tolerance of
105 3 ppm to target specific compounds (Table S1). The Q Exactive Orbitrap MS was calibrated within a mass accuracy
106 range of 1 ppm using the Thermo Scientific Pierce LTQ Velos ESI Positive Ion Calibration Solution (containing a
107 mixture of caffeine, MRFA, Ultramark 1621, and *N*-butylamine in an acetonitrile-methanol-acetic acid solution).
108 The total lipid extract from the surface sediment at 885 mbsl was further analyzed by acid hydrolysis to determine the
109 composition and relative abundance of IPL-derived CL (resulting from the acid hydrolysis of IPLs) using the method
110 described by Lengger et al. (2012) and analyzed by HPLC-APCI/MS (according to Hopmans et al., 2016).

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

112 Sediment was centrifuged and the excess of water was removed by pipetting before proceeding with the extraction of
113 nucleic acids from the sediment. DNA/RNA of surface (0–0.5 cm) and subsurface (10–12 cm) sediments was extracted
114 with the RNA PowerSoil® Total Isolation Kit plus the DNA elution accessory (Mo Bio Laboratories, Carlsbad, CA).
115 Concentration of DNA and RNA were quantified by Nanodrop (Thermo Scientific, Waltham, MA) and Fluorometric
116 with Quant-iT™ PicoGreen® dsDNA Assay Kit (Life technologies, Netherlands). RNA extracts were treated with
117 DNase and reverse-transcribed to cDNA as described by Pitcher et al. (2011). Quantification of archaeal 16S rRNA
118 gene copies and *amoA* gene copies were estimated by QPCR by using the following primers; Parch519F and ARC915R
119 (archaeal 16S rRNA gene), CrenAmoAQ-F and CrenAmoAModR (*amoA* gene), as previously described (Pitcher et al.,
120 2011). For details on the QPCR conditions, efficiency and R² of the QPCR assays see Table S2.

121 **16S rRNA gene amplicon sequencing, analysis, and phylogeny**



122 PCR reactions were performed with the universal, Bacteria and Archaea, primers S-D-Arch-0159-a-S-15 and S-D-Bact-
123 785-a-A-21 (Klindworth et al., 2013) as previously described in Moore et al. (2015). The archaeal 16S rRNA gene
124 amplicon sequences were analyzed by QIIME v1.9 (Caporaso et al., 2010). Raw sequences were demultiplexed and
125 then quality-filtered with a minimum quality score of 25, length between 250–350, and allowing maximum two errors
126 in the barcode sequence. Taxonomy was assigned based on blast and the SILVA database version 123 (Altschul et al.,
127 1990; Quast et al., 2013). Representative operational taxonomic units (OTUs, clusters of reads with 97% similarity) of
128 archaeal groups were extracted through filter_taxa_from_otu_table.py and filter_fasta.py with QIIME (Caporaso et al.,
129 2010). The phylogenetic affiliation of the partial archaeal 16S rRNA gene sequences was compared to release 123 of
130 the Silva NR SSU Ref database (<http://www.arb-silva.de/>; Quast et al., 2013) using the ARB software package (Ludwig
131 et al., 2004). Sequences were added to the reference tree supplied by the Silva database using the ARB Parsimony tool.
132 MCG intragroup phylogeny for representative sequences of OTUs affiliated to the MCG lineage was carried out in
133 ARB (Ludwig et al., 2004). Sequences were added by parsimony to a previously-built phylogenetic tree composed of
134 reference sequences of the 17 MCG subgroups known so far (Kubo et al., 2012). Affiliation of any 16S rRNA gene
135 sequences to a given subgroup was done assuming a similarity cutoff of $\geq 85\%$.

136 Cloning, sequencing and phylogeny of the archaeal *amoA* gene

137 Amplification of the archaeal *amoA* gene was performed as described by Yakimov et al., (2011). PCR reaction mixture
138 was the following (final concentration): Q-solution 1× (PCR additive, Qiagen); PCR buffer 1×; BSA (200 $\mu\text{g ml}^{-1}$);
139 dNTPs (20 μM); primers (0.2 $\text{pmol } \mu\text{l}^{-1}$); MgCl_2 (1.5 mM); 1.25 U Taq polymerase (Qiagen, Valencia, CA, USA). PCR
140 conditions for these amplifications were the following: 95°C, 5 min; 35 × [95°C, 1 min; 55°C, 1 min; 72°C, 1 min];
141 final extension 72°C, 5 min. PCR products were gel purified (QIAquick gel purification kit, Qiagen) and cloned in the
142 TOPO-TA cloning® kit from Invitrogen (Carlsbad, CA, USA) and transformed in *E. coli* TOP10 cells following the
143 manufacturer's recommendations. Recombinant clones plasmid DNAs were purified by Qiagen Miniprep kit and
144 screening by sequencing ($n \geq 30$) using M13R primer by MacroGen Europe Inc. (Amsterdam, The Netherlands).
145 Obtained archaeal *amoA* protein sequences were aligned with already annotated *amoA* sequences by using the Muscle
146 application (Edgar, 2004). Phylogenetic trees were constructed with the Neighbor-Joining method (Saitou and Nei,
147 1987) and evolutionary distances computed using the Poisson correction method with a bootstrap test of 1,000
148 replicates.

149 RESULTS

150 In this study, we analyzed both IPLs and DNA/RNA extracts from sediments previously collected along the Arabian
151 Sea Murray Ridge within the OMZ (885 mbsl), just below the lower interface (1306 mbsl), and well below the OMZ
152 (2470 and 3003 mbsl). The surface sediment (0-0.5 cm) at 885 mbsl was fully anoxic, however, the surface sediments



153 below the OMZ were partly oxygenated (1306 mbsl), and fully oxygenated at 2470 and 3003 mbsl (Table 1). The
154 subsurface sediments (10-12 cm) were fully anoxic at all stations (Table 1). For more details on the physicochemical
155 conditions in these sediments see Table 1.

156 **Archaeal IPL-GDGTs in the surface and subsurface sediments**

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

166 In the surface sediment at 885 mbsl, crenarchaeol IPLs were dominant (44.6% of all detected IPL-GDGTs), occurring
167 predominantly with DH as IPL-type (with a hexose head group on both ends; 43.1%; Table 2). IPL-GDGT-2 was the
168 second most abundant (29.6%), also mainly consisting of the IPL-type DH (29.5%; Table 2). IPL-GDGT-0, -1, -3 and -
169 4 were occurring with relative abundances of 0.3%, 1.7%, 17.8% and 6.1%, respectively (Table 2). Overall, the
170 majority (98.1%; Table 3) of IPL-GDGTs in surface sediment at 885 mbsl with IPL-type DH (all with a hexose
171 molecule on both ends of the CL).

172 The surface sediment at 1306 mbsl contained mostly IPL-GDGT-0 (37.6%), almost entirely with the IPL-type HPH
173 (36.6%; Table 2). Slightly less abundant was the IPL-crenarchaeol (35.6%), with the IPL-types HPH (18.7%) and DH
174 type-I (15.5%) in equal amounts and with a minor relative abundance with MH (1.4%). Overall, the IPL-GDGTs in
175 surface sediment at 1306 mbsl mainly contained the IPL-types HPH (55.4%; Table 3) and DH (42.0%) (Table 3).

176 Well below the OMZ, surface sediments from 2470 and 3003 mbsl were both dominated by IPL-GDGT-0 (71.9 and
177 80.8%, respectively), predominantly with IPL-type HPH (Table 2; Fig. 1a). The IPL-crenarchaeol had a lower relative
178 abundance (26.6 and 17.6%, respectively) and again was dominated by the member with IPL-type HPH (Table 2). The
179 other IPL-GDGTs occurred in minor quantities (<1%). Overall, IPL-type HPH was, thus, by far the most abundant head
180 group detected in surface sediments at 2470 and 3003 mbsl (97.7% and 97.4%, respectively), in contrast to the other
181 two surface sediments studied (Table 3).

182 In all subsurface (10-12 cm) sediments (i.e. at 885, 1306, 2470 and 3003 mbsl) the most abundant IPL-GDGTs were
183 DH-crenarchaeol (28.9±3.8%; Table 2) and DH-GDGT-2 (25.5±3.5%; Table 2). DH was also the most commonly
184 observed IPL-type attached to GDGT-3 and GDGT-4 (Table 2). Overall the distributions of the IPL-GDGTs in all
185 subsurface sediments were relatively similar (Fig. 1a) in comparison to the substantial changes observed at the surface



186 (cf. Fig. 1a). Overall, the IPL-type DH was the predominant one detected in subsurface sediment with a relative
187 abundance ranging from 68.8% at 3003 mbsl to 90.1% at 1306 mbsl (Table 3). In contrast to all other sediments, in the
188 subsurface sediments at 885 mbsl and 1306 mbsl, two different isomers (Fig. S2) of the DH-GDGT-0 were detected
189 (Table 2). DH type-I (0.9% at 1306 mbsl) is also found in the other surface and subsurface sediments and in
190 combination with other core GDGT structures, whereas the other isomer (DH type-II) only occurs (7.8% at 885 mbsl;
191 1.8% at 1306 mbsl; Table 2; Fig. S2b). In addition, these subsurface sediments also contain small amounts of GDGT-0
192 with cyclopentanetetraol and MH head groups (IPL-type HCP; 1.6% at 885 mbsl; 0.4% at 1306 mbsl; Table 1; Fig.
193 S2c).

194 IPL derived CL-GDGTs from the 885 mbsl surface sediment were re-analyzed (after Lengger et al., 2012), in order to
195 exclude degradation within the stored samples, but the IPL derived CL-GDGTs composition showed the same
196 distribution as reported previously (Lengger et al., 2012) ($p=1.00$).

197 **Archaeal diversity in the surface and subsurface sediment**

198 Different archaeal groups were detected in surface and subsurface sediment across the Arabian sea OMZ. The surface
199 sediment at 885 mbsl, contained archaeal 16S rRNA gene sequences that were assigned to several archaeal groups (Fig.
200 1b). The most dominant group was MCG (Total 30.5%, 12.2% attributed to C3; also known as MCG-15, Kubo et al.,
201 2012). Another major group found was the DPANN Woesearchaeota Deep sea Hydrothermal Vent Group 6 (DHVEG-
202 6, 20.3%; Fig. 1b; Castelle et al., 2015). Marine Benthic Group (MBG) -B, -D and -E were also present with 12.2%,
203 7.7% and 6.9% of the archaeal 16S rRNA gene reads, respectively (Fig. 1b). Sequences affiliated to the Marine
204 Hydrothermal Vent Group (MHVG, 8.1%) of the phylum Euryarchaeota were also detected (Fig. 1b). Other groups,
205 with lower relative abundances, were Thermoplasmatales groups ANT06-05 (5.7%) and F2apm1A36 (3.3%) and the
206 DPANN Aenigmarchaeota (previously named Deep Sea Euryarchaeotic Group, DSEG; 1.6%; Fig. 1b).

207 Below the OMZ, in partly and fully oxygenated surface sediments at 1306, 2470 and 3003 mbsl (Table 1), the most
208 dominant archaeal group was Thaumarchaeota MG-I with relative abundances of 81.5%, 89.7% and 100%, respectively
209 (Fig. 1b). At 1306 mbsl other archaeal groups, such as MHVG (5.6%), Thermoplasmatales ASC21 (3.2%), DHVEG-6
210 (2.9%), MBG-B (2.4%) and MCG (1.3%) made up the rest of the archaeal community (Fig. 1b). At 2470 mbsl
211 DHVEG-6 (1.1%) was still detectable besides the MG-I (Fig. 1b).

212 In the subsurface sediments (10–12 cm), only the DNA extracted from the sediments at 885 and 1306 mbsl gave a
213 positive amplification signal. The archaeal composition of the subsurface (10–12 cm) sediments at 885 mbsl and 1306
214 mbsl was similar (Fig. 1b; Pearson correlation coefficient of 0.95), with most of the 16S rRNA gene reads classified
215 within the MCG (47.5% and 48.4%, respectively). Other archaeal groups, such as MBG-D (14.4% and 5.7%,
216 respectively), MBG-B (10.1% and 4.4%), the Woesearchaeota (7.8% and 10.4%), were also detected with comparable
217 relative abundances (Fig. 1b). Other archaeal groups such as Thaumarchaeota Terrestrial hot spring, the Euryarchaeota



218 MHVG, MBG-E and the Aenigmarchaeota were detected but at low (< 10%) relative abundance (Fig. 1b). Only minor
219 amount of reads were classified as Thaumarchaeota MG-I (0.5% at 1306 mbsl) (Fig. 1b).

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

230 As the Thaumarchaeota MGI was dominant in oxygenated sediments at 1306, 2470 and 3003 mbsl (Fig. 1b), we further
231 analyzed the diversity of this group by performing a more detailed phylogeny of the recovered 16S rRNA gene reads
232 attributed to this group. Five OTUs dominated the Thaumarchaeota MGI (Table 5); we will refer to them as OTU-1 to -
233 5. OTU-1, 2, 3 and 5 were phylogenetically closely related to other known benthic Thaumarchaeota MGI species, such
234 as ‘*Ca. Nitrosoarchaeum koreensis* MY1’ or environmental 16S rRNA gene sequences from marine sediments (Fig. 3).
235 On the other hand, OTU-4 clustered with 16S rRNA gene sequences from pelagic Thaumarchaeota MGI species, like
236 *Ca. Nitrosopelagicus brevis*, and also clustered with 16S rRNA sequences recovered from seawater SPM (Fig. 3).
237 OTU-3 was the most abundant OTU in the surface sediment at 1306, 2470, and 3003 mbsl with a relative abundance of
238 44–68% (Table 5). At 1306 mbsl OTU-4 was the second most abundant (35.1%). This OTU had a much lower relative
239 abundance (1.6% and 0.0%) at 2470 and 3003 mbsl, respectively (Table 5). The relative abundance of OTU-2 increased
240 with increasing sampling station depth (Table 5), OTU-1 and 5 had an abundance <5% in the surface sediments (Table
241 5).

242 The diversity of Thaumarchaeota MGI was further assessed by amplification, cloning and sequencing of the archaeal
243 *amoA* gene. Most of the *amoA* gene sequences from surface (27 out of 29 clones) and subsurface sediment at 885 mbsl
244 (9 out of 10 clones) and just one from the surface sediment from 1306 mbsl (1 out of 58 clones) were closely related
245 with *amoA* gene sequences previously recovered from SPM at 1050 mbsl from this area of the Arabian Sea (Villanueva
246 et al., 2014). Phylogenetically they fall within the ‘Water column B, subsurface water’ *amoA* clade as defined by
247 Francis et al. (2005) (Fig. 4). At 1306 and 3003 mbsl (surface and subsurface) the majority of recovered *amoA* gene
248 sequences clustered within the ‘shallow water/sediment’ clade (100 and 98.3%, respectively) and are closely related
249 with *amoA* gene sequences from water column SPM at 170 mbsl (Villanueva et al., 2014) as well as *amoA* gene coding



250 sequences previously detected in sediments (Villanueva et al., 2014; Fig. 4). Of all recovered *amoA* gene sequences
251 from 885 mbsl only a small fraction (8.3%) clustered within the ‘shallow water/sediment’ clade (Fig. 4).

252 **Abundance and potential activity of archaea in surface and subsurface sediments**

253 The abundance of archaeal 16S rRNA gene copies in the surface sediments of different stations varied slightly: it was
254 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
255 activity, based on the 16S rRNA gene transcripts of the archaeal 16S rRNA gene, was the lowest at 2470 mbsl (5×10^4
256 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.
257 5a). The abundance of archaeal 16S rRNA gene copies in the subsurface sediment varied also within one and a half
258 order of magnitude ($1.1\text{--}54 \times 10^9$; Fig. 5c), with a decrease with increasing water depth. The potential activity showed
259 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
260 the surface sediments.

261 The abundance of Thaumarchaeota was estimated by quantifying the archaeal *amoA* gene copies. The highest
262 abundance of *amoA* gene copies in surface sediment was detected at 2470 mbsl (1.0×10^9 copies g^{-1} sediment), and the
263 lowest at 885 mbsl (5×10^4 ; Fig. 5b). *AmoA* gene transcripts in surface sediment were under the detection limit at 885
264 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,
265 2470 and 3003 mbsl, respectively (Fig. 5b). In subsurface sediments, the abundance of *amoA* gene copies was low at
266 885 and 1306 mbsl ($5.4\text{--}19 \times 10^2$ gene transcripts g^{-1} sediment) and higher at 2470 and 3003 mbsl (4.1×10^5 , 5.4×10^6 ,
267 respectively; Fig. 5d). *AmoA* gene transcripts were not detected in the subsurface sediments (Fig. 5d).

268 **DISCUSSION**

269 In this study, we assessed the changes in benthic archaeal diversity and abundance in sediments of the Arabian Sea
270 oxygen minimum zone along a gradient in bottom water oxygen concentrations. The steep Murray ridge protrudes the
271 OMZ, allowing the study of sediments deposited under varying bottom water oxygen concentrations but receiving
272 organic matter (OM), the most important fuel for benthic prokaryotic activity in sediments, produced in a relatively
273 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
274 110 km apart) and, therefore, likely composed of the same primary photosynthate. However, there were differences in
275 OM quality, as OM within the OMZ had a higher biochemical “quality” based on amino acid composition and intact
276 phytopigments compared to OM below the OMZ (Koho et al., 2013). Therefore, changes in the quality and flux of OM
277 received by the different sediment niches could also affect the archaeal community composition as several of the
278 archaeal groups (i.e. MCG and MBG-D) reported here have been suggested to use OM as carbon source in anoxic
279 conditions (Lloyd et al., 2013).

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



281 We detected large difference in archaeal diversity between the surface sediment deposited within the OMZ and those
282 deposited below the OMZ. In contrast to the diverse anaerobic archaeal community in the surface of 885 mbsl, in
283 surface sediments at 1306, 2470 and 3003 mbsl, Thaumarchaeota MG1 were dominant, representing 80-100% of the
284 archaeal population (Fig. 1). This clear difference in the benthic archaeal population in the surface sediments can be
285 attributed to the oxygen availability as Thaumarchaeota are known to require oxygen for their metabolism (i.e.
286 nitrification; Könneke et al., 2005). In fact, the oxygen penetration depth (OPD) was observed to be 3, 10, and 19 mm
287 in sediments at 1306, 2470, and 3003 mbsl, respectively, while in sediments at 885 mbsl, the OPD was barely 0.1 mm
288 (Table S1; Kraal et al., 2012). The surface (0-5 mm) sediment at 1306 mbsl was not fully oxygenated (OPD of 3 mm),
289 which probably explains the detection in a relatively low abundance (ca. 20%) of the anaerobic archaea that thrive in
290 the anoxic sediment from 885 mbsl. The low OPD at 1306 mbsl also explains the low *amoA* gene expression in
291 comparison with the deeper surface sediments (Fig. 5b, d). Overall this indicates the presence of Thaumarchaeota with
292 lower activity at 1306 mbsl (Fig. 5). Within the Thaumarchaeota MG1 group, we also detected changes in the relative
293 abundance of specific OTUs in the surface sediments at 1306, 2470 and 3003 mbsl (Table 5). For example, OTU-2
294 becomes progressively more abundant with increasing water depth, suggesting that this OTU is favored at the higher
295 oxygen concentrations found in the surface sediment at 3003 mbsl. OTU-4 was closely affiliated with '*Ca.*
296 *Nitrosopelagicus brevis*', a pelagic MG-I member, which indicates that this DNA is most likely derived from the
297 overlaying water column (Table 5), and thus should be considered to represent fossil DNA.

298 High *amoA* gene abundances were detected in the surface sediment at 2470 and 3003 mbsl, while values in the surface
299 of 885 mbsl were approximately three orders of magnitude less. The lack of oxygen in the surface sediments at 885
300 mbsl and in the subsurface sediments, as well as undetectable *amoA* gene transcripts at those depths, suggest that in
301 these cases the *amoA* gene DNA signal is fossil. It is well known that under anoxic conditions DNA of marine pelagic
302 microbes may become preserved in sediments even for periods of thousands of years (Boere et al., 2011; Coolen et al.,
303 2004b). The fossil origin of the Thaumarchaeotal *amoA* gene is supported by the phylogenetic affiliation of the *amoA*
304 gene fragments amplified from the 885 mbsl surface sediment, as those sequences were closely related to *amoA* gene
305 sequences detected in the suspended particulate matter in the subsurface water column (Villanueva et al., 2014), thus
306 suggesting that the detected DNA originated from pelagic Thaumarchaeota present in the subsurface water column, as
307 proposed for the presence of OTU-4 16S rRNA gene sequences in the surface sediment (see earlier).

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

312 In the anoxic surface sediment at 885 mbsl (within the OMZ), we detected a highly diverse archaeal population
313 composed of MCG, Thermoplasmatales, MBG-B, -D and -E, Woesearchaeota, and MHVG. Archaeal groups such as



314 MCG and MBG-B and E have been previously described in anoxic marine sediments, where they have been suggested
315 to be involved in anaerobic OM degradation (e.g. Biddle et al., 2006; Inagaki et al., 2003; Castelle et al., 2015).
316 Members of the DPANN Woese archaeota were only present in the surface sediment at 885 mbsl but not in the
317 subsurface anoxic sediments at 885 and 1306 mbsl, suggesting that their presence here is not solely dependent on the
318 absence of oxygen but possibly also of the OM composition and availability in surface and subsurface sediments.
319 Alternatively, the Woese archaeota 16S rRNA gene signal could also originate from the water column and deposited in
320 the surface sediment at 885 mbsl as fossil DNA as observed for the case of Thaumarchaeota as mentioned above.

321 **Archaeal community composition in the anoxic subsurface sediments**

322 The archaeal diversity in the subsurface sediment (10–12 cm) in both 885 and 1306 mbsl (i.e. dominated by MCG,
323 MBG-B, -D and -E) is similar to that observed in the surface sediment at 885 mbsl. This supports that oxygen
324 availability is an important factor for determining the diversification of archaeal groups (Fig. 1b). MCG, one of the
325 dominant archaeal groups in these sediments showed substantial differences in the distribution of its subgroups. A high
326 relative abundance of the subgroup MCG-15 was noted in surface sediments, while the subsurface sediments had a
327 higher intra-group diversity. A recent survey of the ecological niches and substrate preferences of the MCG in estuarine
328 sediments based on genomic data pointed to MCG-6 archaea as degraders of complex extracellular carbohydrate
329 polymers (plant-derived) as detrital proteins, while subgroups 1, 7, 15 and 17 have mainly the potential to degrade
330 detrital proteins (Lazar et al., 2016). Lazar et al. (2016) also described the presence of aminopeptidases coded in the
331 genome bin of MCG-15, suggesting that this subgroup could be specialized in degradation of extracellular peptides in
332 comparison with the other MCG subgroups, which would be restricted to the use of amino acid and oligopeptides.
333 Considering the dominance of the MCG-15 subgroup in the surface sediments analyzed in this study, we can
334 hypothesize that the protein content in the OM deposited in the surface sediment, mainly originated from photosynthate,
335 could be still quite undegraded and therefore the MCG-15 would be favored in this niche and fueled its metabolism by
336 the degradation of peptides extracellularly, while in subsurface sediments, other MCG groups such as 2, 8 and 14 would
337 be more favored.

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



346 biochemical composition can influence the microbial community composition as was shown recently for North Sea
347 sediments (Oni et al., 2015).

348 **Benthic archaea as potential sources for archaeal IPLs**

349 Previous studies have evaluated the archaeal core lipid composition derived from IPLs and CLs as well as estimating
350 the different IPLs of crenarchaeol attributed to Thaumarchaeota in the same sediments studied here (Lengger et al.,
351 2012, 2014), as well as in the overlaying water column (Pitcher et al., 2011; Schouten et al., 2012). The studies by
352 Lengger et al. (2012, 2014) were limited to the determination of MH-, DH- and HPH-crenarchaeol with HPLC/ESI-
353 MS² using a specific selected reaction monitoring method (SRM; Pitcher et al., 2011). In our study we expanded the
354 screening for IPLs carrying different polar head groups in combination with multiple CLs using high resolution
355 accurate mass/mass spectrometry (see Table S1). By applying this method, we unraveled the unknown diversity of IPL-
356 GDGTs in the sediments under study, which allows a more direct comparison with the archaeal diversity detected by
357 gene-based methods.

358 Fully oxygenated surface sediments showed a dominance of GDGT-0 and crenarchaeol mostly with HPH as IPL-type
359 (Table 1). This is the expected IPL-GDGT signature of Thaumarchaeota as previously observed in pure cultures
360 (Pitcher et al., 2010; Qin et al., 2015; Schouten et al., 2008; Sinninghe Damsté et al., 2012). The predominance of the
361 HPH IPL-type in these sediments was previously interpreted by Lengger et al. (2012) as an indication of the presence of
362 an active Thaumarchaeota population synthesizing membrane lipids *in situ* (Lengger et al., 2012 & 2014), taking into
363 account the labile nature of HPH IPLs (Harvey et al., 1986; Schouten, et al., 2010). This hypothesis is strongly
364 supported by (i) the fact that the archaeal community is dominated by Thaumarchaeota (Fig. 1) and (ii) the high
365 abundance of thaumarchaeotal *amoA* gene copies and gene transcripts detected in oxygenated sediments in our study.
366 On the other hand, in the anoxic surface sediment at 885 mbsl, crenarchaeol was predominantly present with DH as the
367 IPL-type (Table 1). This was considered to be a fossil signal of Thaumarchaeota deposited from the water column due
368 to a higher preservation potential of glycolipid head groups (as present in DH) as previously suggested (Lengger et al.,
369 2012 & 2014). The presence of *amoA* gene sequences from the deep water column in the 885 mbsl surface sediment as
370 well as the much lower *amoA* gene abundance and lack of *amoA* gene expression (Fig. 5b, d) supports the contention
371 that the crenarchaeol IPLs are fossil.

372 We detected an increase in the ratio IPL-GDGT-0/IPL-crenarchaeol within the oxygenated surface sediments with
373 increasing water depth. This trend may be explained to be due to decreasing bottom water temperatures with increasing
374 water depth, which results in a higher ratio of GDGT-0/crenarchaeol (Schouten et al., 2002), and would thus point to a
375 population of Thaumarchaeota actively adjusting their membrane lipid composition according to environmental
376 conditions.

377 A striking result is the low relative abundance of GDGT-0 IPLs in the surface sediment at 885 mbsl (Table 1). Only
378 MH-GDGT-0 was detected in low relative abundance (0.3 %), whereas GDGT-0 with any other of the IPL-types that



379 were screened for in our study (Table S2; Fig. 1b) was absent. In contrast, Lengger et al. (2012) reported a significant
380 amount of IPL-derived CL-GDGT-0 (when the head groups are removed by acid hydrolyses and the remaining core
381 lipid is analyzed) in this surface sediment. This discrepancy must be due to the presence of an IPL-type with unknown
382 head groups not included in our analytical window. To test this hypothesis, we re-analyzed the IPL-derived CL-GDGT
383 composition in the surface sediment at 885 mbsl and we recovered an identical CL-GDGT distribution as reported by
384 Lengger et al. (2014), confirming the presence of a GDGT-0 of unknown IPL-type. This IPL could originate from any
385 of the archaeal groups present in the surface sediment at 885 mbsl, such as MCG, Thermoplasmatales, MBG-B, MBG-
386 E and Euryarchaeota MHVG. Woese archaeota also occur relatively abundant in the surface sediments at 885 mbsl (Fig.
387 1) but recent studies suggest that their small genomes lack the genes coding for the enzymes of the GDGT biosynthetic
388 pathway (Jahn et al., 2004; Podar et al., 2013; Villanueva et al., 2017; Waters et al., 2003). Therefore, they are not
389 expected to contribute to the IPL-GDGT pool. Ruling out the Woese archaeota as a possible source of IPL-GDGTs, the
390 IPL GDGT-0 with unknown polar head group(s) in the surface sediment at 885 mbsl may be attributed to the MCG,
391 which make up to 30.5% of the archaeal 16S rRNA gene reads in this sample. Most of these MCG archaea fall into
392 subgroup MCG-15 (Table 4). Previous studies proposed butanetriol dibiphytanyl glycerol tetraethers (BDGTs) as
393 putative biomarker of the MCG based on the correlation between this compound and the presence of MCG in estuarine
394 sediment (Meador et al., 2014). However, we did not detect any IPL BDGTs in the sediments analyzed in our study.
395 Buckles et al. (2013) also suggested that members of the MCG and Crenarchaeota group 1.2 could be the biological
396 source of IPL GDGT-0 found in the anoxic hypolimnion of a tropical lake. Considering these evidences, it is possible
397 that the unknown IPL GDGT-0 present in the surface sediment at 885 mbsl could be a biomarker for MCG. We also
398 detected a high relative abundance of MCG (up to 48.4% relative abundance) in surface and subsurface sediment at 885
399 mbsl and subsurface sediment at 1306 mbsl (Fig. 1). However, the diversity of the MCG population was higher in the
400 subsurface sediments in comparison with diversity in surface sediments at 885 mbsl as sequences closely related to the
401 MCG subgroups, 2, 8, 10, 14, 5b, 15, and 17 were detected both in the 885 mbsl and 1306 mbsl subsurface sediment
402 samples (Fig. 2). This niche differentiation of MCG members may also be the origin of the different DH-GDGT-0
403 moieties (and the one GDGT-0 with IPL-type HCP) that we detected within the subsurface sediments (Table 1). There
404 is also a high abundance of IPL-GDGT-0 present within SPM in the OMZ (Schouten et al., 2012). We, therefore,
405 cannot exclude the possibility that the detected IPL derived CL-GDGT-0 is derived from archaea in the water column.
406 In subsurface sediments, the IPL GDGT distribution was remarkably different from that detected in the surface
407 oxygenated sediment as higher relative abundances of GDGT-1, 2, 3 and 4 were detected in detriment of GDGT-0,
408 similar to the distribution detected in the surface sediments at 885 mbsl. It remains unclear whether this represents new
409 archaeal production in the anoxic sediments or selective preservation of archaeal lipids produced in the water column.
410 At 885 mbsl and at 1306 mbsl, in the subsurface sediment we also found two variants of the IPL-type DH GDGT-0 and
411 the IPL-type HCP (Fig. S2). GDGT-0 with the HCP IPL-type and the GDGT-0 with a DH headgroup attached were not



412 detected in any other sediments or in combination with other GDGTs. Since the GDGT-0 were not detected in the
413 surface sediments it is likely that they are produced *in situ*. Unfortunately, we only obtained archaeal community
414 compositions at those two depths so we cannot compare these with the samples missing these moieties. GDGT-0 with a
415 cyclopentanetetraol head group has been previously detected in cultures of the hyperthermophilic crenarchaeal
416 *Sulfolobales* (Langworthy et al., 1974; Sturt et al., 2004). However, the source(s) of the DH-GDGT-0 and the GDGT-0
417 with a cyclopentanetetraol head group remains unknown.

418 CONCLUSIONS

419 By using a combined 16S rRNA gene amplicon sequencing and IPL analysis with high-resolution accurate mass/mass
420 spectrometry we have unraveled the high diversity of benthic archaea harbored specially in anoxic sediments, as well as
421 increasing the repertoire of archaeal intact polar lipid detected. DNA-based analyses revealed a dominance of active
422 benthic *in situ* Thaumarchaeota in those sediment where oxygen was present, which coincided with high relative
423 abundance of the HPH-crenarchaeol previously suggested to be a marker of living Thaumarchaeota (Pitcher et al.,
424 2011a). In the anoxic marine sediments analyzed here, members of the MCG, DPANN and Euryarchaeota
425 Thermoplasmatales dominated. We also observed a high diversity within the MCG with a more diverse population in
426 subsurface sediments. Surface anoxic sediments with an important MGC population coincided with a high abundance
427 of an IPL GDGT-0 with a yet-unknown polar head group. This result, together with previous studies, suggests that this
428 IPL GDGT-0 could be a potential biomarker for MGC. Besides, subsurface anoxic sediments had a high relative
429 abundance of IPL GDGT-1, 2, and 3 with DH headgroups, which could be attributed to fossil signal due to the more
430 recalcitrant nature of the glycosidic bonds (Schouten et al., 2010) or being IPLs synthesized by the archaeal groups
431 detected in those sediments. In addition, unusual headgroups were also detected, as an IPL GDGT-0 with a hexose head
432 group on both ends of the core lipid, two hexoses on one end, and a cyclopentanetetraol molecule bound to the core
433 lipid and a hexose attached to it. Possibly ruling out the involvement of archaeal members of the DPANN in making
434 those lipids due to the lack of lipid biosynthetic pathway (Podar et al., 2013; Waters et al., 2003), dominant archaeal
435 members in those sediments such as the MCG and Thermoplasmatales, could be potential biological sources of these
436 IPLs. To conclude, this combined approach has shed light on the possible biological sources of specific archaeal IPLs
437 and also detected a higher diversity of benthic archaeal than previously thought.

438 Acknowledgments

439 Elda Panoto is thanked for assistance with molecular analyses. We would like to thank the captain and crew of the RV
440 *Pelagia* as well as the cruise leader, technicians and scientists participating in cruise 64PE301. This PASOM cruise was
441 funded by the Earth and Life Science and Research Council (ALW) with financial aid from the Netherlands
442 Organization for Scientific Research (NWO) (grant 817.01.015) to Prof. G.J. Reichart (PI). NIOZ is acknowledged for



443 the studentship of MAB. This research was further supported by the NESSC and SIAM Gravitation Grants
444 (024.002.001 and 024.002.002) from the Dutch Ministry of Education, Culture and Science (OCW) to JSSD.

445 References

- 446 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J.: Basic local alignment search tool., *J. Mol. Biol.*,
447 215(3), 403–10, doi:10.1016/S0022-2836(05)80360-2, 1990.
- 448 Biddle, J. F., Lipp, J. S., Lever, M. a, Lloyd, K. G., Sørensen, K. B., Anderson, R., Fredricks, H. F., Elvert, M., Kelly,
449 T. J., Schrag, D. P., Sogin, M. L., Brenchley, J. E., Teske, A., House, C. H. and Hinrichs, K.-U.: Heterotrophic Archaea
450 dominate sedimentary subsurface ecosystems off Peru., *Proc. Natl. Acad. Sci. U. S. A.*, 103(10), 3846–3851,
451 doi:10.1073/pnas.0600035103, 2006.
- 452 Bligh, E. G., Dyer, W. J.: A rapid method of total lipid extraction and Purification, *Can. J. Biochem. Physiol.*, 37, 911–
453 917, doi:10.1139/o59-099., 1959.
- 454 Boere, A. C., Rijpstra, W. I. C., de Lange, G. J., Sinninghe Damsté, J. S. and Coolen, M. J. L.: Preservation potential of
455 ancient plankton DNA in Pleistocene marine sediments, *Geobiology*, 9, 377–393, doi:10.1111/j.1472-
456 4669.2011.00290.x, 2011.
- 457 Buckles, L. K., Villanueva, L., Weijers, J. W. H., Verschuren, D. and Damsté, J. S. S.: Linking isoprenoidal GDGT
458 membrane lipid distributions with gene abundances of ammonia-oxidizing Thaumarchaeota and uncultured
459 crenarchaeotal groups in the water column of a tropical lake (Lake Challa, East Africa), *Environ. Microbiol.*, 2,
460 doi:10.1111/1462-2920.12118, 2013.
- 461 Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Peña, A. G.,
462 Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights, D., Koenig, J. E., Ley, R. E., Lozupone, C. A.,
463 Mcdonald, D., Muegge, B. D., Pirrung, M., Reeder, J., Sevinsky, J. R., Turnbaugh, P. J., Walters, W. A., Widmann, J.,
464 Yatsunencko, T., Zaneveld, J. and Knight, R.: correspondEce QIIME allows analysis of high- throughput community
465 sequencing data Intensity normalization improves color calling in SOLiD sequencing, *Nat. Publ. Gr.*, 7(5), 335–336,
466 doi:10.1038/nmeth0510-335, 2010.
- 467 Castelle, C. J., Wrighton, K. C., Thomas, B. C., Hug, L. A., Brown, C. T., Wilkins, M. J., Frischkorn, K. R., Tringe, S.
468 G., Singh, A., Markillie, L. M., Taylor, R. C., Williams, K. H. and Banfield, J. F.: Genomic Expansion of Domain
469 Archaea Highlights Roles for Organisms from New Phyla in Anaerobic Carbon Cycling, *Curr. Biol.*, 25, 1–12,
470 doi:10.1016/j.cub.2015.01.014, 2015.
- 471 Coolen, M. J. L., Muyzer, G., Rijpstra, W. I. C., Schouten, S., Volkman, J. K. and Sinninghe Damsté, J. S.: Combined
472 DNA and lipid analyses of sediments reveal changes in Holocene haptophyte and diatom populations in an Antarctic
473 lake, *Earth Planet. Sci. Lett.*, 223(1–2), 225–239, doi:10.1016/j.epsl.2004.04.014, 2004a.
- 474 Coolen, M. J. L., Hopmans, E. C., Rijpstra, W. I. C., Muyzer, G., Schouten, S., Volkman, J. K. and Sinninghe Damsté,
475 J. S.: Evolution of the methane cycle in Ace Lake (Antarctica) during the Holocene: response of methanogens and
476 methanotrophs to environmental change, *Org. Geochem.*, 35(10), 1151–1167, doi:10.1016/j.orggeochem.2004.06.009,
477 2004b.
- 478 DeLong, E. F., Ying Wu, K., Prézelin, B. B. and Jovine, R. V. M.: High abundance of Archaea in Antarctic marine
479 picoplankton, *Nature*, 371, 695–697, doi:10.1038/371695a0, 1994.
- 480 Durbin, A. M. and Teske, A.: Sediment-associated microdiversity within the Marine Group I Crenarchaeota, *Environ.*
481 *Microbiol. Rep.*, 2(5), 693–703, doi:10.1111/j.1758-2229.2010.00163.x, 2010.
- 482 Edgar, R. C.: MUSCLE: Multiple sequence alignment with high accuracy and high throughput, *Nucleic Acids Res.*,
483 32(5), 1792–1797, doi:10.1093/nar/gkh340, 2004.
- 484 Fillol, M., Auguet, J.-C., Casamayor, E. O. and Borrego, C. M.: Insights in the ecology and evolutionary history of the
485 Miscellaneous Crenarchaeotic Group lineage., *ISME J.*, 1–13, doi:10.1038/ismej.2015.143, 2015.
- 486 Harvey, H. R., Fallon, R. D. and Patton, J. S.: The effect of organic matter and oxygen on the degradation of bacterial
487 membrane lipids in marine sediments, *Geochim. Cosmochim. Acta*, 50, 795–804, doi:10.1016/0016-7037(86)90355-8,
488 1986.
- 489 Hopmans, E. C., Schouten, S. and Sinninghe, J. S.: The effect of improved chromatography on GDGT-based
490 palaeoproxies, *Org. Geochem.*, 93, 1–6, doi:10.1016/j.orggeochem.2015.12.006, 2016.
- 491 Inagaki, F., Suzuki, M., Takai, K., Oida, H., Sakamoto, T., Aoki, K., Neelson, K. H. and Horikoshi, K.: Microbial
492 Communities Associated with Geological Horizons in Coastal Subseafloor Sediments from the Sea of Okhotsk
493 Microbial Communities Associated with Geological Horizons in Coastal Subseafloor Sediments from the Sea of



- 494 Okhotsk, *Appl. Environ. Microbiol.*, 69(12), 7224–7235, doi:10.1128/AEM.69.12.7224, 2003.
- 495 Ingalls, A. E., Huguet, C. and Truxal, L. T.: Distribution of intact and core membrane lipids of archaeal glycerol dialkyl
496 glycerol tetraethers among size-fractionated particulate organic matter in Hood Canal, Puget Sound, *Appl. Environ.*
497 *Microbiol.*, 78(5), 1480–1490, doi:10.1128/AEM.07016-11, 2012.
- 498 Jahn, U., Summons, R., Sturt, H., Grosjean, E. and Huber, H.: Composition of the lipids of *Nanoarchaeum equitans* and
499 their origin from its host *Ignicoccus* sp. strain KIN4/I, *Arch. Microbiol.*, 182(5), 404–413, doi:10.1007/s00203-004-
500 0725-x, 2004.
- 501 Jorgenson, S. L., Hannisdal, B., Lanzén, A., Baumberger, T., Flesland, K., Fonseca, R., Øvreås, L., Steen, I. H.,
502 Thorseth, I. H., Pedersen, R. B. and Schleper, C.: Correlating microbial community profiles with geochemical data in
503 highly stratified sediments from the Arctic Mid-Ocean Ridge, *Proc. Natl. Acad. Sci.*, 109(42), E2846–E2855,
504 doi:10.1594/PANGAEA.786687, 2012.
- 505 Karner, M. B., DeLong, E. F. and Karl, D. M.: Archaeal dominance in the mesopelagic zone of the Pacific Ocean,
506 *Nature*, 409(6819), 507–510, doi:10.1038/35054051, 2001.
- 507 Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M. and Glockner, F. O.: Evaluation of general
508 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies, *Nucleic*
509 *Acids Res.*, 41(1), 1–11, doi:10.1093/nar/gks808, 2013.
- 510 Koga, Y. and Morii, H.: Biosynthesis of ether-type polar lipids in archaea and evolutionary considerations., *Microbiol.*
511 *Mol. Biol. Rev.*, 71(1), 97–120, doi:10.1128/MMBR.00033-06, 2007.
- 512 Koho, K. A., Nierop, K. G. J., Moodley, L., Middelburg, J. J., Pozzato, L., Soetaert, K., Van Der Plicht, J. and Reichart,
513 G. J.: Microbial bioavailability regulates organic matter preservation in marine sediments, *Biogeosciences*, 10(2),
514 1131–1141, doi:10.5194/bg-10-1131-2013, 2013.
- 515 Könneke, M., Bernhard, A. E., de la Torre, J. R., Walker, C. B., Waterbury, J. B. and Stahl, D. A.: Isolation of an
516 autotrophic ammonia-oxidizing marine archaeon, *Nature*, 437(22), 543–546, doi:10.1038/nature03911, 2005.
- 517 Kraal, P., Slomp, C. P., Reed, D. C., Reichart, G.-J. and Poulton, S. W.: Sedimentary phosphorus and iron cycling in
518 and below the oxygen minimum zone of the northern Arabian Sea, *Biogeosciences*, 9(7), 2603–2624, doi:10.5194/bg-9-
519 2603-2012, 2012.
- 520 Kubo, K., Lloyd, K. G., F Biddle, J., Amann, R., Teske, A. and Knittel, K.: Archaea of the Miscellaneous
521 Crenarchaeotal Group are abundant, diverse and widespread in marine sediments, *ISME J.*, 6(10), 1949–1965,
522 doi:10.1038/ismej.2012.37, 2012.
- 523 De La Torre, J. R., Walker, C. B., Ingalls, A. E., Könneke, M. and Stahl, D. A.: Cultivation of a thermophilic ammonia
524 oxidizing archaeon synthesizing crenarchaeol, *Environ. Microbiol.*, 10(3), 810–818, doi:10.1111/j.1462-
525 2920.2007.01506.x, 2008.
- 526 Langworthy, T. A., Mayberry, W. R. and Smith, P. F.: Long-chain glycerol diether and polyol dialkyl glycerol triether
527 lipids of *Sulfolobus acidocaldarius*, *J. Bacteriol.*, 119(1), 106–116, 1974.
- 528 Lazar, C. S., Baker, B. J., Seitz, K., Hyde, A. S., Dick, G. J., Hinrichs, K. U. and Teske, A. P.: Genomic evidence for
529 distinct carbon substrate preferences and ecological niches of Bathyarchaeota in estuarine sediments, *Environ.*
530 *Microbiol.*, 18(4), 1200–1211, doi:10.1111/1462-2920.13142, 2016.
- 531 Learman, D. R., Henson, M. W., Thrash, J. C., Temperton, B., Brannock, P. M., Santos, S. R., Mahon, A. R. and
532 Halanych, K. M.: Biogeochemical and microbial variation across 5500 km of Antarctic surface sediment implicates
533 organic matter as a driver of benthic community structure, *Front. Microbiol.*, 7, 1–11, doi:10.3389/fmicb.2016.00284,
534 2016.
- 535 Lengger, S. K., Hopmans, E. C., Reichart, G.-J., Nierop, K. G. J., Sinninghe Damsté, J. S. and Schouten, S.: Intact polar
536 and core glycerol dibiphytanyl glycerol tetraether lipids in the Arabian Sea oxygen minimum zone. Part II: Selective
537 preservation and degradation in sediments and consequences for the TEX86, *Geochim. Cosmochim. Acta*, 98, 244–258,
538 doi:10.1016/j.gca.2012.05.003, 2012.
- 539 Lengger, S. K., Hopmans, E. C., Sinninghe Damsté, J. S. and Schouten, S.: Impact of sedimentary degradation and deep
540 water column production on GDGT abundance and distribution in surface sediments in the Arabian Sea: Implications
541 for the TEX86 paleothermometer, *Geochim. Cosmochim. Acta*, 142, 386–399, doi:10.1016/j.gca.2014.07.013, 2014.
- 542 Lipp, J. S. and Hinrichs, K.-U.: Structural diversity and fate of intact polar lipids in marine sediments, *Geochim.*
543 *Cosmochim. Acta*, 73(22), 6816–6833, doi:10.1016/j.gca.2009.08.003, 2009.
- 544 Lipp, J. S., Morono, Y., Inagaki, F. and Hinrichs, K.-U.: Significant contribution of Archaea to extant biomass in
545 marine subsurface sediments., *Nature*, 454(August), 991–994, doi:10.1038/nature07174, 2008.
- 546 Lloyd, K. G., Schreiber, L., Petersen, D. G., Kjeldsen, K. U., Lever, M. a, Steen, A. D., Stepanauskas, R., Richter, M.,



- 547 Kleindienst, S., Lenk, S., Schramm, A. and Jørgensen, B. B.: Predominant archaea in marine sediments degrade detrital
548 proteins., *Nature*, 496(7444), 215–8, doi:10.1038/nature12033, 2013.
- 549 Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, Buchner, A., Lai, T., Steppi, S., Jobb, G.,
550 Förster, W., Brettske, I., Gerber, S., Ginhart, A. W., Gross, O., Grumann, S., Hermann, S., Jost, R., König, A., Liss, T.,
551 Lüssmann, R., May, M., Nonhoff, B., Reichel, B., Strehlow, R., Stamatakis, A., Stuckmann, N., Vilbig, A., Lenke, M.,
552 Ludwig, T., Bode, A. and Schleifer, K.-H.: ARB: a software environment for sequence data., *Nucleic Acids Res.*, 32(4),
553 1363–71, doi:10.1093/nar/gkh293, 2004.
- 554 Massana, R., Castresana, J., Balague, V., Guillou, L., Romari, K., Groisillier, A., Valentin, K. and Pedros-Alio, C.:
555 Phylogenetic and Ecological Analysis of Novel Marine Stramenopiles, *Appl. Environ. Microbiol.*, 70(6), 3528–3534,
556 doi:10.1128/AEM.70.6.3528, 2004.
- 557 Meador, T. B., Bowles, M., Lazar, C. S., Zhu, C., Teske, A. and Hinrichs, K.-U.: The archaeal lipidome in estuarine
558 sediment dominated by members of the Miscellaneous Crenarchaeotal Group, *Environ. Microbiol.*, 17(7), 2441–2458,
559 doi:10.1111/1462-2920.12716, 2015.
- 560 Meng, J., Xu, J., Qin, D., He, Y., Xiao, X. and Wang, F.: Genetic and functional properties of uncultivated MCG
561 archaea assessed by metagenome and gene expression analyses, *ISME J.*, 8(3), 650–659, doi:10.1038/ismej.2013.174,
562 2014.
- 563 Moore, E. K., Villanueva, L., Hopmans, E. C., Rijpstra, W. I. C., Mets, A., Dedysh, S. N. and Sinninghe Damsté, J. S.:
564 Abundant trimethylornithine lipids and specific gene sequences are indicative of planctomycete importance at the
565 oxic/anoxic interface in sphagnum-dominated northern wetlands, *Appl. Environ. Microbiol.*, 81(18), 6333–6344,
566 doi:10.1128/AEM.00324-15, 2015.
- 567 Nierop, K. G. J., Reichart, G., Veld, H. and Sinninghe Damsté, J. S.: The influence of oxygen exposure time on the
568 composition of macromolecular organic matter as revealed by surface sediments on the Murray Ridge (Arabian Sea),
569 *Geochim. Cosmochim. Acta*, 206, 40–56, 2017.
- 570 Offre, P., Spang, A. and Schleper, C.: Archaea in Biogeochemical Cycles, *Annu. Rev. Microbiol.*, 437–457,
571 doi:10.1146/annurev-micro-092412-155614, 2013.
- 572 Oni, O. E., Schmidt, F., Miyatake, T., Kasten, S., Witt, M., Hinrichs, K.-U. and Friedrich, M. W.: Microbial
573 Communities and Organic Matter Composition in Surface and Subsurface Sediments of the Helgoland Mud Area,
574 North Sea, *Front. Microbiol.*, 6(November), 1–16, doi:10.3389/fmicb.2015.01290, 2015.
- 575 Park, S.-J., Park, B.-J. and Rhee, S.-K.: Comparative analysis of archaeal 16S rRNA and amoA genes to estimate the
576 abundance and diversity of ammonia-oxidizing archaea in marine sediments, *Extremophiles*, 12(4), 605–615,
577 doi:10.1007/s00792-008-0165-7, 2008.
- 578 Pitcher, A., Rychlik, N., Hopmans, E. C., Spieck, E., Rijpstra, W. I. C., Ossebaar, J., Schouten, S., Wagner, M. and
579 Damsté, J. S. S.: Crenarchaeol dominates the membrane lipids of *Candidatus Nitrososphaera gargensis*, a thermophilic
580 group I.1b Archaeon, *ISME J.*, 4(4), 542–52, doi:10.1038/ismej.2009.138, 2010.
- 581 Pitcher, A., Villanueva, L., Hopmans, E. C., Schouten, S., Reichart, G.-J. and Sinninghe Damsté, J. S.: Niche
582 segregation of ammonia-oxidizing archaea and anammox bacteria in the Arabian Sea oxygen minimum zone., *ISME J.*,
583 5(12), 1896–904, doi:10.1038/ismej.2011.60, 2011.
- 584 Podar, M., Makarova, K. S., Graham, D. E., Wolf, Y. I., Koonin, E. V and Reysenbach, A.-L.: Insights into archaeal
585 evolution and symbiosis from the genomes of a nanoarchaeon and its inferred crenarchaeal host from Obsidian Pool,
586 Yellowstone National Park., *Biol. Direct*, 8(1), 9, doi:10.1186/1745-6150-8-9, 2013.
- 587 Qin, W., Carlson, L. T., Armbrust, E. V., Devol, A. H., Moffett, J. W., Stahl, D. A. and Ingalls, A. E.: Confounding
588 effects of oxygen and temperature on the TEX₈₆ signature of marine Thaumarchaeota, *Proc. Natl. Acad. Sci.*, 112(35),
589 10979–10984, doi:10.1073/pnas.1501568112, 2015.
- 590 Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. and Glockner, F. O.: The SILVA
591 ribosomal RNA gene database project: improved data processing and web-based tools, *Nucleic Acids Res.*,
592 41(November 2012), D590–D596, doi:10.1093/nar/gks1219, 2013.
- 593 Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N. N., Anderson, I. J., Cheng, J.-F., Darling, A. E., Malfatti, S., Swan,
594 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.,
595 Kyrpides, N. C., Stepanauskas, R., Rubin, E. M., Hugenholtz, P. and Woyke, T.: Insights into the phylogeny and coding
596 potential of microbial dark matter., *Nature*, 499(7459), 431–437, doi:10.1038/nature12352, 2013.
- 597 De Rosa, M. and Gambacorta, A.: The lipids of archaeobacteria, *The lipids of archaeobacteria*, 27, 153–175, 1988.
- 598 Rossel, P. E., Lipp, J. S., Fredricks, H. F., Arnds, J., Boetius, A., Elvert, M. and Hinrichs, K. U.: Intact polar lipids of
599 anaerobic methanotrophic archaea and associated bacteria, *Org. Geochem.*, 39(8), 992–999,
600 doi:10.1016/j.orggeochem.2008.02.021, 2008.



- 601 Saitou N, N. M.: The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees', *Mol. Biol.*
602 *Evol.*, 4(4), 406–425, doi:citeulike-article-id:93683, 1987.
- 603 Schouten, S., Hopmans, E. C., Schefuß, E. and Sinninghe Damsté, J. S.: Distributional variations in marine
604 crenarchaeotal membrane lipids: a new tool for reconstructing ancient sea water temperatures?, *Earth Planet. Sci. Lett.*,
605 204, 265–274, 2002.
- 606 Schouten, S., Hopmans, E. C., Baas, M., Boumann, H., Standfest, S., Könneke, M., Stahl, D. a. and Sinninghe Damsté,
607 J. S.: Intact membrane lipids of “*Candidatus Nitrosopumilus maritimus*,” a cultivated representative of the cosmopolitan
608 mesophilic group I crenarchaeota, *Appl. Environ. Microbiol.*, 74(8), 2433–2440, doi:10.1128/AEM.01709-07, 2008.
- 609 Schouten, S., Middelburg, J. J., Hopmans, E. C. and Sinninghe Damsté, J. S.: Fossilization and degradation of intact
610 polar lipids in deep subsurface sediments: A theoretical approach, *Geochim. Cosmochim. Acta*, 74(13), 3806–3814,
611 doi:10.1016/j.gca.2010.03.029, 2010.
- 612 Schouten, S., Pitcher, A., Hopmans, E. C., Villanueva, L., van Bleijswijk, J. and Sinninghe Damsté, J. S.: Intact polar
613 and core glycerol dibiphytanyl glycerol tetraether lipids in the Arabian Sea oxygen minimum zone: I. Selective
614 preservation and degradation in the water column and consequences for the TEX₈₆, *Geochim. Cosmochim. Acta*, 98,
615 228–243, doi:10.1016/j.gca.2012.05.002, 2012.
- 616 Schouten, S., Hopmans, E. C. and Sinninghe Damsté, J. S.: The organic geochemistry of glycerol dialkyl glycerol
617 tetraether lipids: A review, *Org. Geochem.*, 54, 19–61, doi:10.1016/j.orggeochem.2012.09.006, 2013.
- 618 Schubotz, F., Wakeham, S. G., Lipp, J. S., Fredricks, H. F. and Hinrichs, K.-U.: Detection of microbial biomass by
619 intact polar membrane lipid analysis in the water column and surface sediments of the Black Sea, *Environ. Microbiol.*,
620 11(10), 2720–2734, doi:10.1111/j.1462-2920.2009.01999.x, 2009.
- 621 Sinninghe Damsté, J. S., Schouten, S., Hopmans, E. C., van Duin, A. C. T. and Geenevasen, J. A. J.: Crenarchaeol: the
622 characteristic core glycerol dibiphytanyl glycerol tetraether membrane lipid of cosmopolitan pelagic crenarchaeota, *J.*
623 *Lipid Res.*, 43, 1641–1651, doi:10.1194/jlr.M200148-JLR200, 2002.
- 624 Sinninghe Damsté, J. S., Rijpstra, W. I. C., Hopmans, E. C., Jung, M.-Y., Kim, J.-G., Rhee, S.-K., Stieglmeier, M. and
625 Schleper, C.: Intact Polar and Core Glycerol Dibiphytanyl Glycerol Tetraether Lipids of Group I.1a and I.1b
626 Thaumarchaeota in Soil, *Appl. Environ. Microbiol.*, 78(19), 6866–6874, doi:10.1128/AEM.01681-12, 2012.
- 627 Stetter, K. O., Fiala, G., Huber, G., Huber, R. and Seegerer, a: Hyperthermophilic microorganisms, *FEMS Microbiol.*
628 *Rev.*, 75, 117–124, doi:10.1111/j.1574-6968.1990.tb04089.x, 1990.
- 629 Sturt, H. F., Summons, R. E., Smith, K., Elvert, M. and Hinrichs, K.-U.: Intact polar membrane lipids in prokaryotes
630 and sediments deciphered by high-performance liquid chromatography/electrospray ionization multistage mass
631 spectrometry—new biomarkers for biogeochemistry and microbial ecology, *Rapid Commun. Mass Spectrom.*, 18(6),
632 617–628, doi:10.1002/rcm.1378, 2004.
- 633 Sugai, A., Sakuma, R., Fukuda, I., Kurosawa, N. and Itoh, Y. H.: The Structure of the Core Polyol of the Ether Lipids
634 from *Sulfolobus acidocaldarius*, *Lipids*, 30(4), 339–344, doi:10.1007/BF02536042, 1995.
- 635 Teske, A. and Sørensen, K. B.: Uncultured archaea in deep marine subsurface sediments: have we caught them all?,
636 *ISME J.*, 2(1), 3–18, doi:10.1038/ismej.2007.90, 2008.
- 637 Villanueva, L., Schouten, S. and Sinninghe Damsté, J. S.: Depth-related distribution of a key gene of the tetraether lipid
638 biosynthetic pathway in marine Thaumarchaeota, *Environ. Microbiol.*, doi:10.1111/1462-2920.12508, 2014.
- 639 Villanueva, L., Schouten, S. and Damsté, J. S. S.: Phylogenomic analysis of lipid biosynthetic genes of Archaea shed
640 light on the “lipid divide,” *Environ. Microbiol.*, 19(1), 54–69, doi:10.1111/1462-2920.13361, 2017.
- 641 Waters, E., Hohn, M. J., Ahel, I., Graham, D. E., Adams, M. D., Barnstead, M., Beeson, K. Y., Bibbs, L., Bolanos, R.,
642 Keller, M., Kretz, K., Lin, X., Mathur, E., Ni, J., Podar, M., Richardson, T., Sutton, G. G., Simon, M., Soll, D., Stetter,
643 K. O., Short, J. M. and Noordevier, M.: The genome of *Nanoarchaeum equitans*: insights into early archaeal evolution
644 and derived parasitism., *Proc. Natl. Acad. Sci. U. S. A.*, 100(22), 12984–8, doi:10.1073/pnas.1735403100, 2003.
- 645 Yakimov, M. M., Cono, V. La, Smedile, F., DeLuca, T. H., Juárez, S., Ciordia, S., Fernández, M., Albar, J. P., Ferrer,
646 M., Golyshin, P. N. and Giuliano, L.: Contribution of crenarchaeal autotrophic ammonia oxidizers to the dark primary
647 production in Tyrrhenian deep waters (Central Mediterranean Sea), *ISME J.*, 5(6), 945–61,
648 doi:10.1038/ismej.2010.197, 2011.
- 649
- 650

651 **Figure legends**

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

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

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

664 **Fig. 4.** Maximum likelihood phylogenetic tree of *amoA* gene coding sequences recovered from surface (S) and
665 subsurface (SS) sediments (colored in blue) at 885 mbsl, 1306 mbsl and 3003 mbsl (155 clones). *AmoA* gene coding
666 sequences recovered from SPM (colored in orange) at 170 mbsl (28 clones), SPM at 1050 (25 clones) reported by
667 Villanueva et al. (2014). ** indicates *amoA* gene sequences recovered from surface sediments at 3003 mbsl previously
668 reported in Villanueva et al., (2015). Scale bar represents a 2% sequence dissimilarity.

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

673



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

Station (mbsl)	T (°C)	BWO ($\mu\text{mol}\cdot\text{L}^{-1}$)	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

676

677


 678 **Table 2. Heatmap^a of the relative abundance (%) of the detected IPLs and sum (not color coded) per IPL-**
 679 **GDGT.**

Sample	Depth (mbsl)	GDGT-0						GDGT-1				GDGT-2						
		MH	I ^b		DH	HCP ^c	HPH	Sum	MH	I ^b		DH	HPH	Sum	MH	I ^b		DH
Surface	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	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			

680

Sample	Depth (mbsl)	GDGT-3				GDGT-4				Crenarchaeol							
		MH	I ^b		DH	HPH	Sum	MH	I ^b		DH	HPH	Sum	MH	I ^b		DH
Surface	885	ND	17.8	ND	17.8	ND	6.1	ND	6.1	1.3	43.1	0.3	44.6				
	1306	0.0	6.9	ND	6.9	ND	2.7	ND	2.7	1.4	15.5	18.7	35.6				
	2470	ND	0.2	ND	0.2	ND	0.0	ND	0.0	0.2	0.6	25.8	26.6				
	3003	ND	0.4	ND	0.4	ND	0.0	ND	0.0	0.4	0.2	17.1	17.6				
Subsurface	885	0.1	15.9	ND	15.9	ND	9.4	ND	9.4	1.1	31.1	1.5	33.8				
	1306	0.0	14.5	ND	14.5	ND	6.1	ND	6.1	2.7	32.4	0.4	35.5				
	2470	ND	9.6	ND	9.6	ND	2.9	ND	2.9	3.5	28.3	1.0	32.8				
	3003	ND	9.7	ND	9.7	ND	5.6	ND	5.6	8.2	23.9	0.6	32.7				

681 ^a Green colors indicate a low relative abundance, red colors indicate a high relative abundance682 ^b DH isomers were detected as a GDGT with a glycosidically-bound hexose moiety on both ends of the core (I) and

683 with one glycosidically-bound dihexose moiety on one end (II).

684 ^c HCP is an IPL-type with an ether-bound cyclopentanetetraol moiety on one end and an hexose moiety on the other

685 (previously reported as GDNT; e.g. De Rosa and Gambacorta, 1988; Sturt et al., 2004).

686 ^d ND = not detected

687

688 **Table 3. Relative abundance of IPL-GDGTs grouped by polar head group^a.**

Sediment	Depth (mbsl)	MH	DH	HCP	HPH
Surface	885	1.7%	98.1%	ND ^b	0.3%
	1306	2.6%	42.0%	ND	55.4%
	2470	0.5%	1.8%	ND	97.7%
	3003	0.8%	1.8%	ND	97.4%
Deep	885	1.8%	92.9%	1.6%	3.7%
	1306	5.2%	91.9%	0.4%	2.5%
	2470	7.9%	72.6%	ND	19.6%
	3003	17.6%	68.8%	ND	13.6%

690 Polar head group types detected: MH = monohexose, DH = dihexose, both isomers combined, HCP = monohexose and

691 cyclopentanetetraol, HPH = monohexose and phosphohexose.

692 ^bND = not detected

693



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

696

Subgroup	885 mbsl		1306 mbsl	
	surface	subsurface	surface	subsurface
	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

697



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

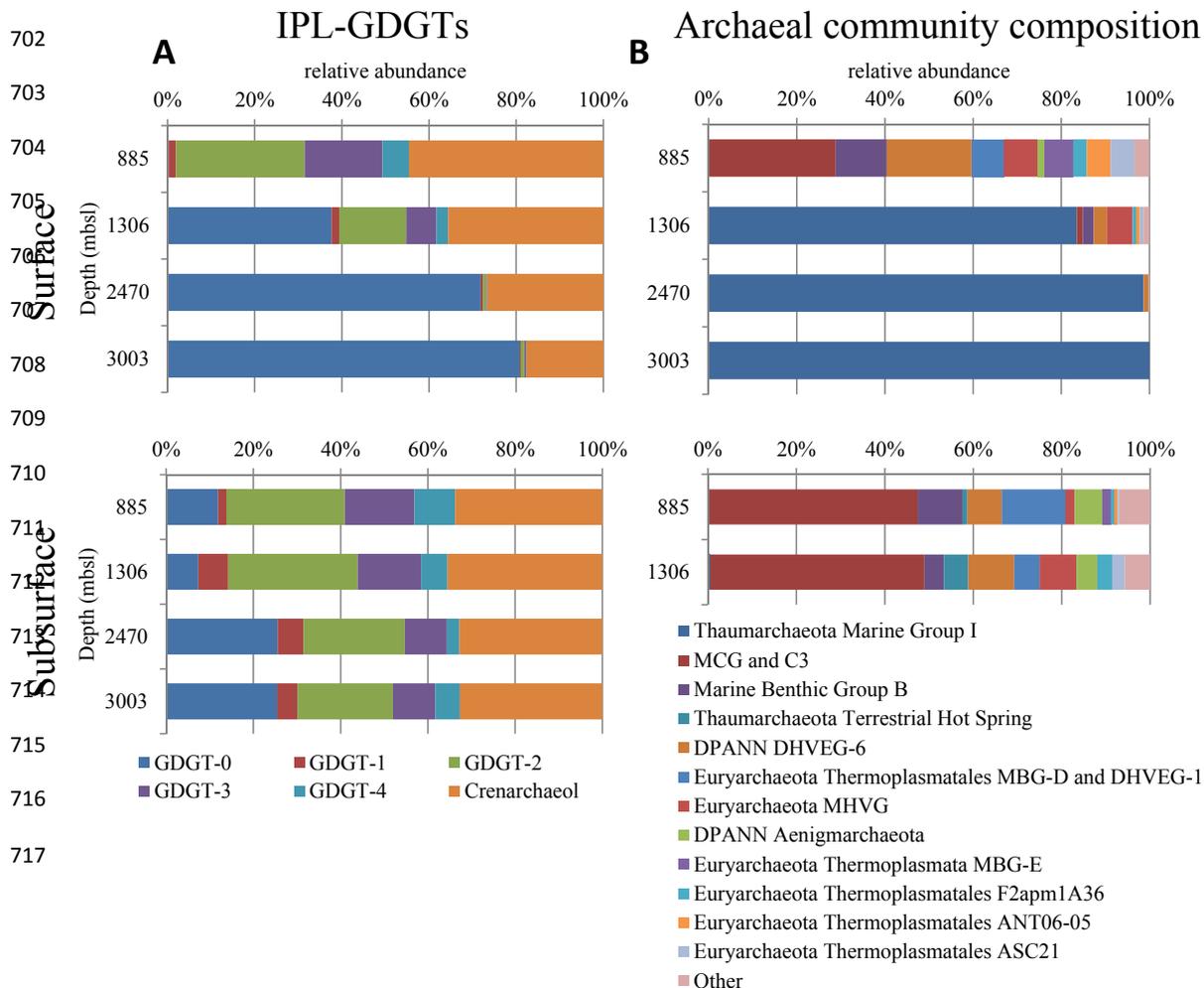
	Depth (mbsl)			
	885	1306	2470	3003
Total reads	0	915	1341	1305
OTU ID #1	n.a. ^a	4.3	2.5	3.0
OTU ID #2	n.a.	3.9	8.1	13.6
OTU ID #3	n.a.	43.6	67.6	61.8
OTU ID #4	n.a.	35.1	1.6	0
OTU ID #5	n.a.	3.3	4.7	2.1

699 ^a n.a. = not applicable

700

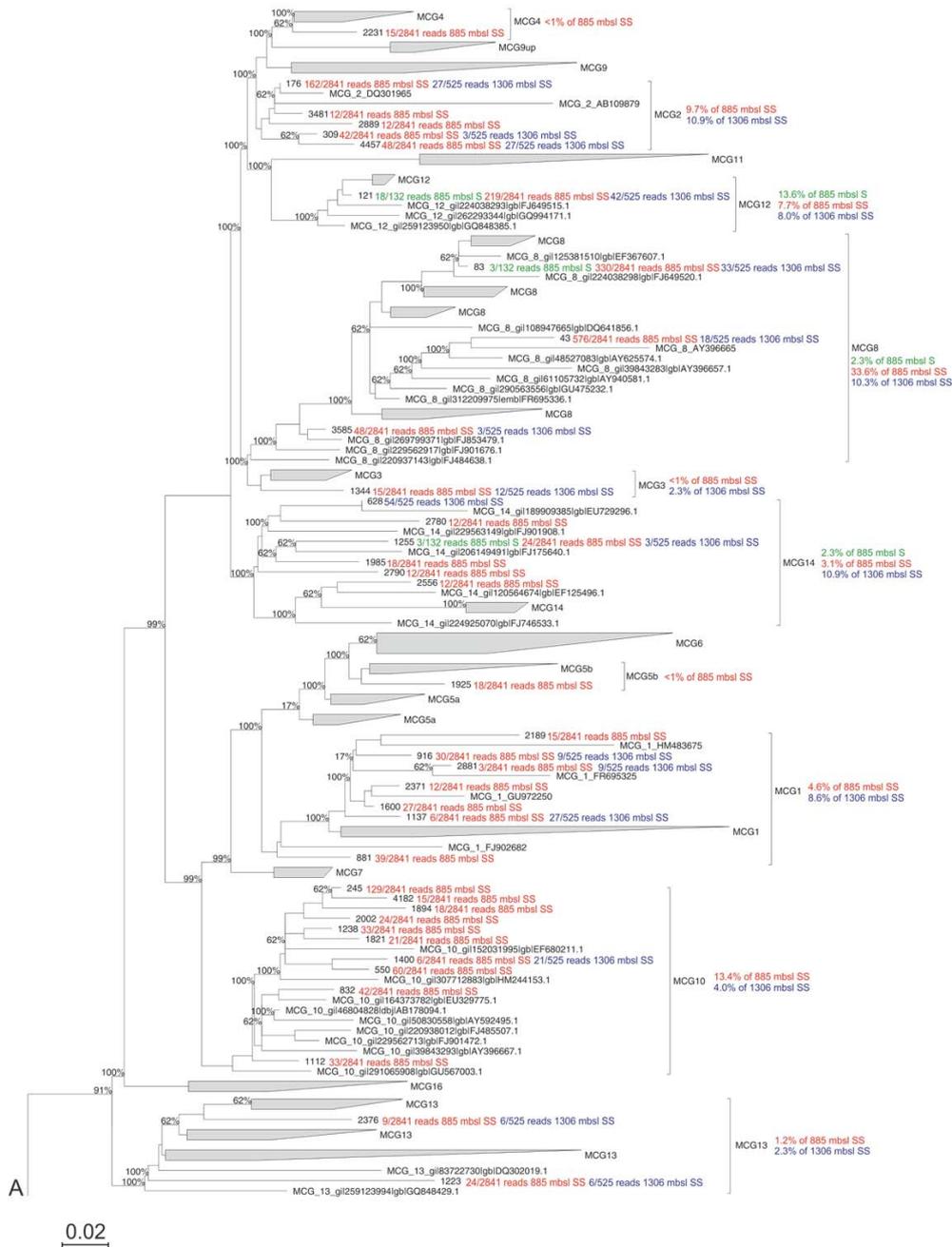


701 **Fig. 1.**



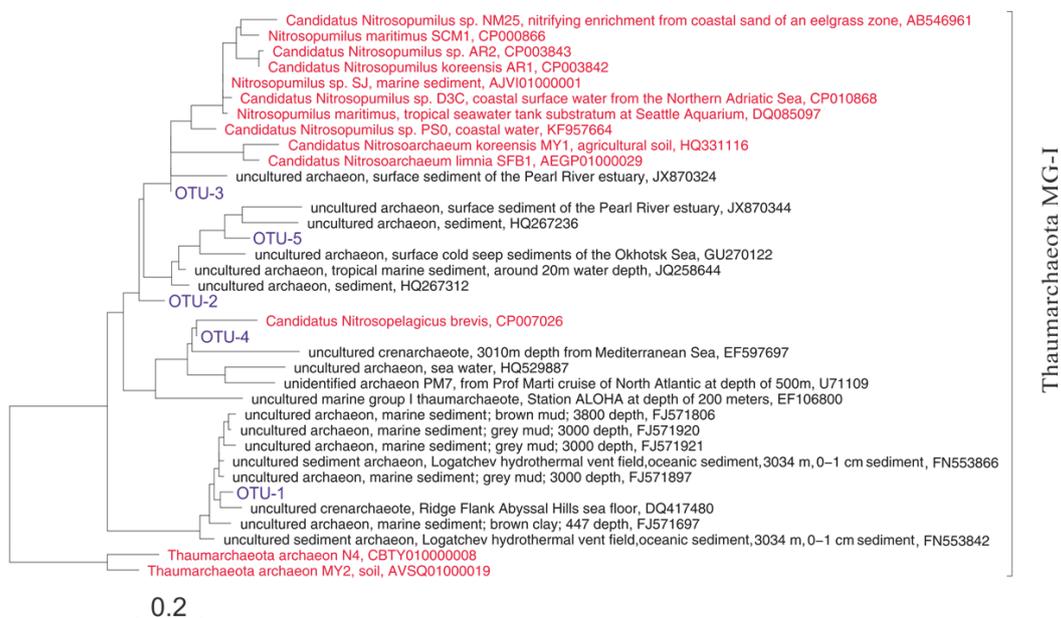


718 Fig. 2.





722 **Fig. 3.**



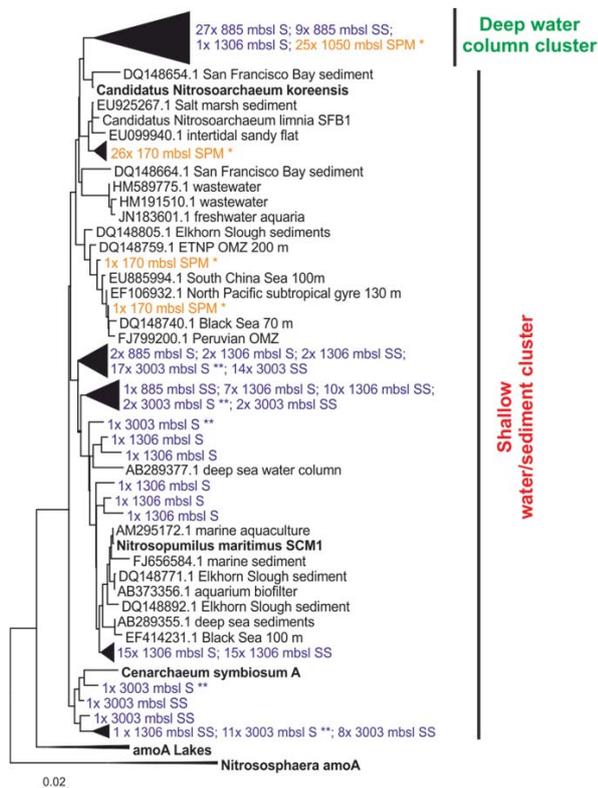
723
724

725



726 Fig. 4.

727





728 Fig. 5.

