

Supplementary methods

Study sites and sampling.

During the TRANSAT-1 cruise (September 2002), in total 36 stations were occupied with RV Pelagia following a track from 65°N to 35°N in the eastern basin of the North Atlantic. Quantitative PCR analyses were done on 3 stations: T9, T20 and T28 (Supplementary Fig. 1). Sampling was performed with 22 10-L NOEX (no oxygen exchange) bottles mounted in a frame also holding sensors for conductivity, temperature, depth, chlorophyll fluorescence, and optical backscattering. The specific water masses were identified based on their temperature and salinity characteristics and their oxygen concentrations (Supplementary Table 1). One subsurface sample per station was taken from the lower euphotic layer at around 100 m depth. The waters below the surface thermocline (500 m depth) were formed by North Atlantic Central Water (NACW). Below, the 1000-2000 m depth layer was occupied by either Labrador Seawater (LSW; Sts T20, T28) or Antarctic Intermediate Water (AAIW, St T9). At Station 9, below the AAIW, slope water was present. At around 2000 m depth, the Iceland-Scotland Overflow Water (ISOW) was sampled at Stations T9 and T20. At Station T28, Middle Northeast Atlantic Deep Water (MNEADW) and Lower Northeast Atlantic Deep Water (LNEADW) were sampled at around 3000 m and 4000 m depth, respectively. For each depth, one liter of seawater was filtered onto 0.2 µm polycarbonate filter (Millipore) and the filters were subsequently stored at -80°C until further processing in the lab.

During the ARCHIMEDES-2 cruise (November to December 2006), a total of 20 stations was occupied with RV Pelagia following a track from 26°N to 5°S in the eastern (sub)tropical Atlantic. Quantitative PCR analyses were done at 7 stations: A3, A5, A11, A19, A28, A33 and A48 (Supplementary Fig. 1). Sampling was performed as described above for TRANSAT-1 and according to the water masses (Supplementary Table 1). One subsurface sample per station was taken from the lower euphotic layer at around 100 m depth. The waters below the surface thermocline (250 to 500 m depth) were formed by either North Atlantic Central Water (NACW; Sts. A3, A5 and A48) or South Atlantic Central Water (SACW; Sts. A19, A28, A33). The 700-900 m depth layer was occupied by the Subarctic Intermediate Water (SAIW)

water at the northern stations, or the Antarctic Intermediate Water (AAIW) at the southern stations. Below the AAIW, the Upper Northeast Atlantic Deep Water (UNEADW), Middle Northeast Atlantic Deep Water (MNEADW) and Lower Northeast Atlantic Deep Water (LNEADW) is flowing southwards with cores at around 1800 m, 2500-2750 m and 4000 m depth, respectively. For each depth, ten liters of seawater were filtered through a 0.22 μm Sterivex filter GP unit (Millipore) and 1.8 ml of lysis buffer (40 mM EDTA, 50 mM Tris-HCl, 0.75 M sucrose) was added and the filters subsequently stored at -80°C until analysis.

Molecular analysis

DNA extraction and preparation of the Q-PCR standards. Archaeal 16S rRNA gene fragments were amplified from the plasmid 88exp4 (from the archaeal clones library) and the fosmid HF770_041|11 (obtained from T. Mincer⁴) as 16S rDNA, Marine *Crenarchaeota* Group I (MCGI) and pSL12-like standard, respectively, using the 16S specific primers (Parch519f and ARC915R)³¹. The archaeal *amoA* gene of the standard was amplified from DNA of *Nitrosopumilus maritimus* (obtained from M. Könneke¹²), using specific archaeal *amoA* primers (Supplementary Table 2). The β -proteobacterial *amoA* gene for the standard was amplified using a mix (25% each) of *Nitrosomonas europaea*, *N. eutropha*, *N. marina* and *Nitrospira briensis* using specific β -proteobacterial *amoA* primers (Supplementary Table 2).

Each amplification was performed under the following conditions: 4 min initial denaturation; 35 cycles at, 94°C for 30 s, specific annealing temperature of the primer set for 40 s (Supplementary Table S2), 72°C for 2 min, 80°C for 25 s using for 1 U of Pico Maxx high fidelity DNA polymerase (Stratagene), 10x Pico Maxx PCR buffer, 0.25 mM of each dNTP, 8 μg of BSA, 0.2 μM of primers, 3mM of MgCl_2 and ultra pure sterile water (Sigma). Amplification products were run on an agarose gel (1%), stained with SYBRGold[®] (Invitrogen), bands were isolated and purified using QuickClean 5M gel extraction kit (GenScript). Purified products were quantified using a Nanodrop[®] spectrophotometer and the copy numbers of the 16S rRNA genes of each QPCR standard were calculated directly from the concentration of the purified DNA. Ten-fold serial dilutions ranging from 10^7 to 10^2 copies of a known copy number of the standard were used per QPCR reaction in duplicate to generate an external quantification standard.

DNA extraction and quantitative PCR of the environmental samples.

DNA extraction. Total DNA extraction from marine samples was performed using the UltraClean soil DNA isolation kit (Mobio) for the TRANSAT-1 samples and the UltraClean Mega soil DNA isolation kit (Mobio) followed by a DNA concentration with a Centricon unit (Millipore) for ARCHIMEDES-2 samples.

Q-PCR analysis. All Q-PCR analyses were performed on an iCycler iQ 5 thermocycler (Bio-rad) equipped with i-Cycler iQ software (version 3.1, Bio-Rad). The copy numbers of MCGI 16S rDNA, pSL12-like 16S rDNA, archaeal *amoA* and β -proteobacterial *amoA* in all environmental samples were determined in triplicate on the non-diluted sample and for two different dilutions of the sample (5 times and 25 times diluted). The reaction mixture (20 μ l) contained 1 U of Pico Maxx high fidelity DNA polymerase (Stratagene), 2 μ l of 10x Pico Maxx PCR buffer, 0.25 mM of each dNTP, 8 μ g of BSA, 0.2 μ M of primers, 50 000 times diluted SYBR Green[®] (Invitrogen) (optimized concentration), a final concentration of 10 nM of fluorescein, 3mM of MgCl₂ and ultra pure sterile water (Sigma)³². All reactions were performed in a 96-well QPCR plates (Bio-Rad) with optical tape (Bio-Rad). One μ L of diluted or non-diluted environmental DNA was added to 19 μ L of mix in each well. Accumulation of newly amplified double stranded gene products was followed online as the increase of fluorescence due to the binding of the fluorescent dyes SYBRGreen[®] and fluorescein. Specificity for QPCR reaction was tested on agarose gel electrophoresis and with a melting curve analysis (60°C-94°C) in order to identify unspecific PCR products such as primer dimers or fragments with unexpected fragment lengths. PCR efficiencies and correlation coefficients for standard curves were as follows: for the MCGI 16S rDNA assay, 87.9-105.8% and $r^2 = 0.991-0.999$, for the pSL12-like 16S rDNA assay, 94.7-110.5% and $r^2 = 0.949-0.999$, for the archaeal *amoA* assay, 86.5-97% and $r^2 = 0.983-0.999$ and for the β -proteobacterial *amoA* assay, 77.8-88.3% and $r^2 = 0.989-0.999$. Each gene fragment was detected using a standard for the specific quantification of MCGI 16S rRNA, pSL12-like 16S rRNA, archaeal *amoA* and β -proteobacterial *amoA* genes and primer combinations and annealing temperature as listed in Table 3. Thermocycling was performed as follows: initial denaturation at 95°C for 4 min; amplification: 41 cycles at, 95°C for 30 s, primer annealing temperature for 40 s, and extension at 72°C for 30 s, 80°C for

25 s, with a plate read between each cycle; melting curve 60 – 94°C with a read every 0.5°C held for 1 s between each read.

***AmoA* gene amplification, sequencing and phylogenetic analysis of archaeal *amoA* sequences.** Archaeal *amoA* sequences were amplified with the primers Arch-*amoA*-for and Arch-*amoA*-rev (Supplementary Table 2) as described in the Q-PCR section and 35 cycles of amplification. AOA *amoA* PCR were purified (QuickClean 5M PCR purification Kit, GenScript) and cloned into the pCR4-TOPO[®] vector for sequencing (Invitrogen) according to manufacturers instructions. Clones were checked for the right insert by PCR and sequencing was performed (Eurofins MWG GmbH, Germany). Sequence data were compiled using ARB software³³ and aligned with complete length sequences of closest relatives obtained from NCBI database. Using ARB, the phylogenetic tree was first generated with the aligned, almost complete length sequences of closest relatives from the NCBI data base using the neighbor-joining method. Then the short-aligned sequences were added to the tree using the ARB parsimony option. To compare the archaeal *amoA*-based richness within each clone library, rarefaction analysis and Chao1 non parametric richness estimations were performed by using DOTUR³⁴. Operational taxonomic units (OTUs) were defined as sequence group in which sequences differed by $\leq 2\%$ (Supplementary Fig. 3 and Table 4).

Archaeal 16S rRNA gene amplification, sequencing and phylogenetic analysis of archaeal 16S rRNA sequences. Archaeal 16S rRNA gene clone libraries were generated from 3 stations: Station T9 at 2071 m, Station A3 at 2 distinct depths (249 m and 2750 m) and one station (ARCHIMEDES-3 cruise, 0.21°N, -18.43°E) close to the station A33 (1°N, -20.60°E) at 7155 m depth by PCR amplification using a forward primer specific for Archaea and a reverse universal primer (Arch21f, 1492r)³⁵ (Supplementary Table 3). The 16S rDNA cloning was performed as described above. Seventy-seven, 90, 97 and 80 archaeal clones from the stations T9 at 2071 m, A3 at 249 m, A3 at 2750 m, and A33 at 7155 m, respectively, were sequenced. Among these clones, 44 (57%), 50 (56%), 83 (86%) and 32 (40%) clones from the stations T9 at 2071 m, A3 at 249 m, A3 at 2750 m, and A33 at 7155 m, respectively, were affiliated to the crenarchaeal group. A DOTUR analysis was performed on the 219 crenarchaeal sequences, 19 different OTUs were defined as sequence groups in which

sequences differed by $\leq 3\%$. The sequences were compared with sequences from public database and analyzed using the ARB phylogenetic package (Supplementary Fig. 4).

Primers specificity.

16S rRNA genes primers. To test whether groups not targeted by our MCGI and pSL12 primers might be responsible for the observed Q-PCR pattern, archaeal 16S rRNA gene fragments were amplified, cloned, screened and sequenced (see above). Using the ARB phylogenetic package³³, the specificity of the primers was tested on our MCGI and pSL12-clade affiliated sequences from the four archaeal clone libraries. The primers used in our quantitative PCR approach for detecting MCGI and the pSL12-like cluster targeted all the obtained sequences (data not shown). The specificity of the primers used by Mincer *et al.*⁴ was also tested in silico on our MCGI sequences. These primers targeted also all the MCGI sequences (data not shown).

***AmoA* genes primers.** To test whether groups of archaeal *amoA* sequences not targeted by the archaeal *amoA* primers used in this study³ might be responsible for the observed Q-PCR pattern, we established two clone libraries with the primers used by Mincer *et al.*⁴ and Francis *et al.*¹¹. These two clone libraries were established from the same bathypelagic sample as used with our primers (Station A33, 2502 m depth).

From the 45 short sequences (only 85 bp) recovered from the clone library with the Mincer *et al.*⁴ set of primers, only 2 sequences fell into cluster A, all the other sequences affiliated with cluster B (data not shown). The relatively short sequences of the *amoA* gene amplified by these primers do not allow phylogenetic analyses. Mincer and coauthors used this set of primers to determine the abundance of archaeal *amoA* genes by Q-PCR. The short fragment amplified by these primers is at the end of the *amoA* gene and is not or not entirely recovered by the primers used by Wuchter *et al.*³ and Francis *et al.*¹¹.

Using the primers of Francis *et al.*, we recovered 38 *amoA* sequences with a length of 595 bp affiliating exclusively to cluster B (Supplementary Fig. 5). Half of these sequences are closely related to the bathypelagic subcluster 2. The sequences belonging to cluster A recovered in our North Atlantic clone library are not detected by this set of primers. Using this set of primers for Q-PCR analysis in North Atlantic

waters would underestimate the number of AOA. Hence, the archaeal *amoA* primers used in our study provide a better coverage of the entire group of marine archaeal *amoA* sequences (clusters A and B, Fig. 2), however, do not recover some subclusters within cluster B which are covered by the primers of Francis *et al.*¹¹ (Supplementary Fig. 5). Recently, Beman *et al.*⁵ designed two new forward primers, Arch-amoAFA and Arch-amoAFB, able to amplify specifically cluster A and B of AOA, respectively. The long and closest neighbor sequences of our North Atlantic sequences (cluster A and B) are recovered in silico by these primers (Fig. 2: EF414272, EU049823, EU049824, EU049846, AB373359, EU049842, AM901414 and AB305361). Thus, we conclude that our sequences would also be recovered by the set of primers of Beman *et al.*⁵.

Supplementary discussion

An archaeal 16S rRNA clone library was also established for one mesopelagic sample (station A3) and 3 bathypelagic samples (stations T9, A3 and A33). In a DNA-based phylogenetic tree, the obtained crenarchaeal sequences fell into the MCGI and pSL12-like groups (Supplementary Fig. 4). For the TRANSAT-1 clone library, more than half of the sequences (57%) fell into one specific OTU (i.e., OTU 5), close to *Nitrosopumilus maritimus*, one of two mesophilic crenarchaeal isolates carrying the *amoA* gene. None of the other sequences (except 1 sequence from the bathypelagic sample of station A3) (Supplementary Fig. 4) belonged to this OTU. The relatively high abundance of this OTU might be responsible for the high abundance of archaeal *amoA* genes in the mesopelagic North Atlantic (Supplementary Table 3). The majority of the ARCHIMEDES bathypelagic sequences (stations A3 and A33) fell into OTU 2, far from the two crenarchaeal isolates harbouring the *amoA* gene (Supplementary Fig. 4). At these two bathypelagic sites, the abundance of archaeal *amoA* gene copy numbers was very low. A specific OTU (i.e., OTU 9), the second most abundant OTU in the bathypelagic samples (i.e., 13% and 25% at station A3 and A33, respectively) is closely related to the sequence AY856365 belonging to a cluster

suggested to carry archaeal nitrogenase (*nif*) genes³⁶. Hence, despite the absence of *amoA* genes in these bathypelagic samples, the archaeal communities in the bathypelagic sub-tropical and equatorial North Atlantic might have another role in the nitrogen cycle, potentially as nitrogen-fixing organisms. For the mesopelagic sample (station A3), where the abundance of *amoA* genes was still relatively high (Supplementary Table 3), 48% of the sequences clustered within the pSL12-like group. This might indicate, as suggested by Mincer *et al.*⁴, that these organisms also contain an *amoA*-like gene. The phylogenetic analyses of the archaeal 16S rRNA genes support the idea that specific water masses harbour specific ecotypes and that the presence of OTU 2, close to *Nitrosopumilus maritimus*, especially in the TRANSAT sample might explain the high abundance of *amoA* genes in this specific water mass.

Table S1: Characteristics of the water masses sampled during the TRANSAT-1 and ARCHIMEDES-2 cruise based on CTD^a profiles at the stations (Archaeal *amoA* cloning and archaeal 16S rRNA gene cloning were performed at stations in bold and italic, respectively).

Cruise	Station	Latitude (°N)	Longitude (°E)	Depth (m)	Water mass	Temp	Salinity	Pot temp	Density	CTD oxy	AOU
<i>TRANSAT-1</i>	<i>9</i>	<i>61.63</i>	<i>-20.82</i>	<i>2071</i>	<i>ISOW</i>	<i>2.571</i>	<i>34.983</i>	<i>2.416</i>	<i>27.924</i>	<i>292.3</i>	<i>36.100</i>
TRANSAT-1	9	61.63	-20.82	1633	slope water	3.390	34.952	3.263	27.822	287.4	34.200
TRANSAT-1	9	61.63	-20.82	1014	AAIW	5.644	35.065	5.552	27.660	247.8	56.100
TRANSAT-1	9	61.63	-20.82	498	NACW	7.877	35.196	7.825	27.455	265.9	22.100
TRANSAT-1	9	61.63	-20.82	99	Subsurface	9.100	35.228	9.089	27.283	264.2	15.800
TRANSAT-1	20	55.91	-32.21	2500	ISOW	2.942	34.974	2.739	27.888	284.0	41.795
TRANSAT-1	20	55.91	-32.21	1897	ISOW	3.184	34.972	3.035	27.860	284.9	38.444
TRANSAT-1	20	55.91	-32.21	1147	LSW	3.757	34.915	3.669	27.753	278.8	39.707
TRANSAT-1	20	55.91	-32.21	450	NACW	6.257	35.048	6.216	27.563	237.4	61.839
TRANSAT-1	20	55.91	-32.21	149	Subsurface	7.508	34.938	7.493	27.301	264.2	26.546
TRANSAT-1	28	49.73	-26.13	3878	LNEADW	2.707	34.928	2.359	27.884	253.0	76.054
TRANSAT-1	28	49.73	-26.13	2999	MNEADW	2.873	34.943	2.620	27.874	269.8	56.977
TRANSAT-1	28	49.73	-26.13	1897	LSW	3.233	34.888	3.084	27.788	286.7	36.515
TRANSAT-1	28	49.73	-26.13	597	NACW	5.701	34.946	5.649	27.555	225.6	77.905
TRANSAT-1	28	49.73	-26.13	146	Subsurface	11.214	35.439	11.196	27.082	240.7	26.282
<i>ARCHIMEDES-2</i>	<i>3</i>	<i>24.50</i>	<i>-18.69</i>	<i>2750</i>	<i>MNEADW</i>	<i>2.920</i>	<i>34.952</i>	<i>2.697</i>	<i>27.875</i>	<i>230.9</i>	<i>95.237</i>
ARCHIMEDES-2	3	24.50	-18.69	800	SAIW	8.342	35.188	8.256	27.383	117.1	168.165
ARCHIMEDES-2	3	24.50	-18.69	499	NACW	11.824	35.603	11.758	27.105	145.1	118.393
<i>ARCHIMEDES-2</i>	<i>3</i>	<i>24.50</i>	<i>-18.69</i>	<i>249</i>	<i>NACW</i>	<i>15.902</i>	<i>36.259</i>	<i>15.862</i>	<i>26.749</i>	<i>183.9</i>	<i>57.515</i>
ARCHIMEDES-2	3	24.50	-18.69	99	Subsurface	19.751	36.898	19.732	26.282	216.1	7.355
ARCHIMEDES-2	5	21.32	-20.87	4001	LNEADW	2.368	34.897	2.023	27.888	244.3	87.562
ARCHIMEDES-2	5	21.32	-20.87	2748	MNEADW	2.954	34.956	2.730	27.874	234.3	91.597
ARCHIMEDES-2	5	21.32	-20.87	899	SAIW	7.069	35.045	6.980	27.457	126.8	167.168
ARCHIMEDES-2	5	21.32	-20.87	501	NACW	11.371	35.506	11.306	27.114	109.4	156.863
ARCHIMEDES-2	5	21.32	-20.87	249	NACW	13.821	35.763	13.785	26.823	92.3	160.108
ARCHIMEDES-2	5	21.32	-20.87	100	Subsurface	15.642	35.869	15.627	26.503	56.7	186.391
ARCHIMEDES-2	11	11.98	-20.21	3999	LNEADW	2.365	34.894	2.020	27.886	242.2	89.782
ARCHIMEDES-2	11	11.98	-20.21	2751	MNEADW	2.859	34.938	2.636	27.869	237.5	89.244
ARCHIMEDES-2	11	11.98	-20.21	900	transitional IW	5.993	34.778	5.912	27.389	105.0	196.989
ARCHIMEDES-2	11	11.98	-20.21	499	transitional CW	9.680	35.076	9.622	27.076	44.0	232.912

ARCHIMEDES-2	11	11.98	-20.21	250	transitional CW	12.101	35.207	12.068	26.738	86.5	176.013
ARCHIMEDES-2	11	11.98	-20.21	100	Subsurface	14.206	35.446	14.192	26.492	73.9	177.046
ARCHIMEDES-2	19	4.90	-15.19	4001	LNEADW	2.368	34.892	2.023	27.883	239.1	92.827
ARCHIMEDES-2	19	4.90	-15.19	2751	MNEADW	2.787	34.926	2.566	27.865	236.0	91.279
ARCHIMEDES-2	19	4.90	-15.19	900	AAIW	5.117	34.629	5.041	27.377	127.1	181.609
ARCHIMEDES-2	19	4.90	-15.19	500	SACW	7.938	34.758	7.887	27.101	86.9	201.628
ARCHIMEDES-2	19	4.90	-15.19	251	SACW	13.183	35.314	13.148	26.607	79.0	177.496
ARCHIMEDES-2	19	4.90	-15.19	100	Subsurface	16.219	35.655	16.203	26.205	86.2	154.580
ARCHIMEDES-2	28	-3.39	-14.83	2501	MNEADW	2.948	34.925	2.749	27.848	239.8	85.957
ARCHIMEDES-2	28	-3.39	-14.83	1799	UNEADW	3.781	34.971	3.635	27.801	240.3	78.298
ARCHIMEDES-2	28	-3.39	-14.83	901	AAIW	4.419	34.543	4.349	27.386	149.8	164.233
ARCHIMEDES-2	28	-3.39	-14.83	498	SACW	7.499	34.658	7.450	27.086	108.2	183.404
ARCHIMEDES-2	28	-3.39	-14.83	252	SACW	11.901	35.139	11.868	26.723	94.9	168.898
ARCHIMEDES-2	28	-3.39	-14.83	100	Subsurface	16.398	35.689	16.382	26.190	109.6	130.293
<i>ARCHIMEDES-3</i>	<i>23</i>	<i>-0.21</i>	<i>-18.43</i>	<i>7155</i>	<i>AABW</i>	<i>1.4212</i>	<i>34.752</i>	<i>0.700</i>	<i>27.870</i>	<i>222.4</i>	<i>121.400</i>
ARCHIMEDES-2	33	1.00	-20.60	2502	MNEADW	2.962	34.934	2.762	27.855	245.0	80.673
ARCHIMEDES-2	33	1.00	-20.60	1800	UNEADW	3.900	34.982	3.753	27.798	244.8	72.817
ARCHIMEDES-2	33	1.00	-20.60	749	AAIW	5.307	34.513	5.244	27.260	142.9	164.551
ARCHIMEDES-2	33	1.00	-20.60	498	SACW	7.195	34.637	7.147	27.112	127.5	166.187
ARCHIMEDES-2	33	1.00	-20.60	250	SACW	11.987	35.150	11.955	26.715	103.3	160.023
ARCHIMEDES-2	33	1.00	-20.60	99	Subsurface	16.469	35.666	16.453	26.156	117.0	122.601
ARCHIMEDES-2	48	19.19	-25.24	2500	MNEADW	3.196	34.963	2.993	27.857	234.4	89.284
ARCHIMEDES-2	48	19.19	-25.24	1798	UNEADW	4.091	35.007	3.942	27.798	212.3	103.796
ARCHIMEDES-2	48	19.19	-25.24	901	SAIW	6.707	34.939	6.621	27.423	115.5	181.197
ARCHIMEDES-2	48	19.19	-25.24	499	NACW	10.193	35.245	10.133	27.121	64.0	209.541
ARCHIMEDES-2	48	19.19	-25.24	249	NACW	13.918	35.735	13.881	26.781	83.5	168.521
ARCHIMEDES-2	48	19.19	-25.24	99	Subsurface	18.733	36.508	18.716	26.248	137.5	90.726

^a: CTD, conductivity, temperature, depth

Water masses: NACW, North Atlantic Central Water ; SACW, South Atlantic Central Water ; transitional CW, transitional Central Water ; LSW, Labrador Sea Water ; ISOW, Iceland-Scotland Overflow Water ; SAIW, Subarctic Intermediate Water ; AAIW, Antarctic Intermediate Water ; transitional IW, transitional intermediate water ; UNEADW, Upper North East Atlantic Deep Water ; MNEADW, Middle North East Atlantic Deep Water ; LNEADW, Lower North East Atlantic Deep Water, AABW, Antarctic Bottom Water.

Parameters: temp, temperature (°C) measured with the CTD rosette ; salinity, salinity (pss) measured with the CTD rosette, pot temp, potential temperature (°C) ; density, potential density anomaly (kg.m^{-3}) ; CTD oxy, oxygen concentration ($\mu\text{mol.kg}^{-1}$) measured with the CTD rosette ; AOU, apparent oxygen utilization ($\mu\text{mol.kg}^{-1}$).

Table S2 – Primers used for detection and quantification of archaeal 16S rRNA and *amoA* genes and bacterial *amoA* genes

Target	Gene	Name	Fragment length (bp)	Annealing temperature (°C)	Sequence (5' to 3')	Reference
Marine <i>Crenarchaeota</i>	16S rRNA	MCGI-391f	122	61.0	AAGGTTARTCCGAGTGRTTTC	3
	16S rRNA	MCGI-554r			TGACCACTTGAGGTGCTG	3
pSL12-like clade	16S rRNA	pSL12_750F	326	65.0	GGTCCRCCAGAACGCGC	4
	16S rRNA	pSL12_876R			GTACTCCCCAGGCGGCAA	4
Archaeal <i>amoA</i>	<i>amoA</i>	Arch- <i>amoA</i> -for	256	58.5	CTGAYTGGGCGTGGACATC	3
	<i>amoA</i>	Arch- <i>amoA</i> -rev			TTCTTCTTTGTTGCCAGTA	3
β- proteobacterial <i>amoA</i>	<i>amoA</i>	<i>amoA</i> -1F	490	61.5	GGGGTTTCTACTGGTGGT	37
	<i>amoA</i>	<i>amoA</i> -r NEW			CCCCTCBGSAAAVCCTTCTTC	38

Table S3: Archaeal 16S rRNA and *amoA* gene, and β -proteobacterial *amoA* gene copy numbers in the samples from the Atlantic Ocean

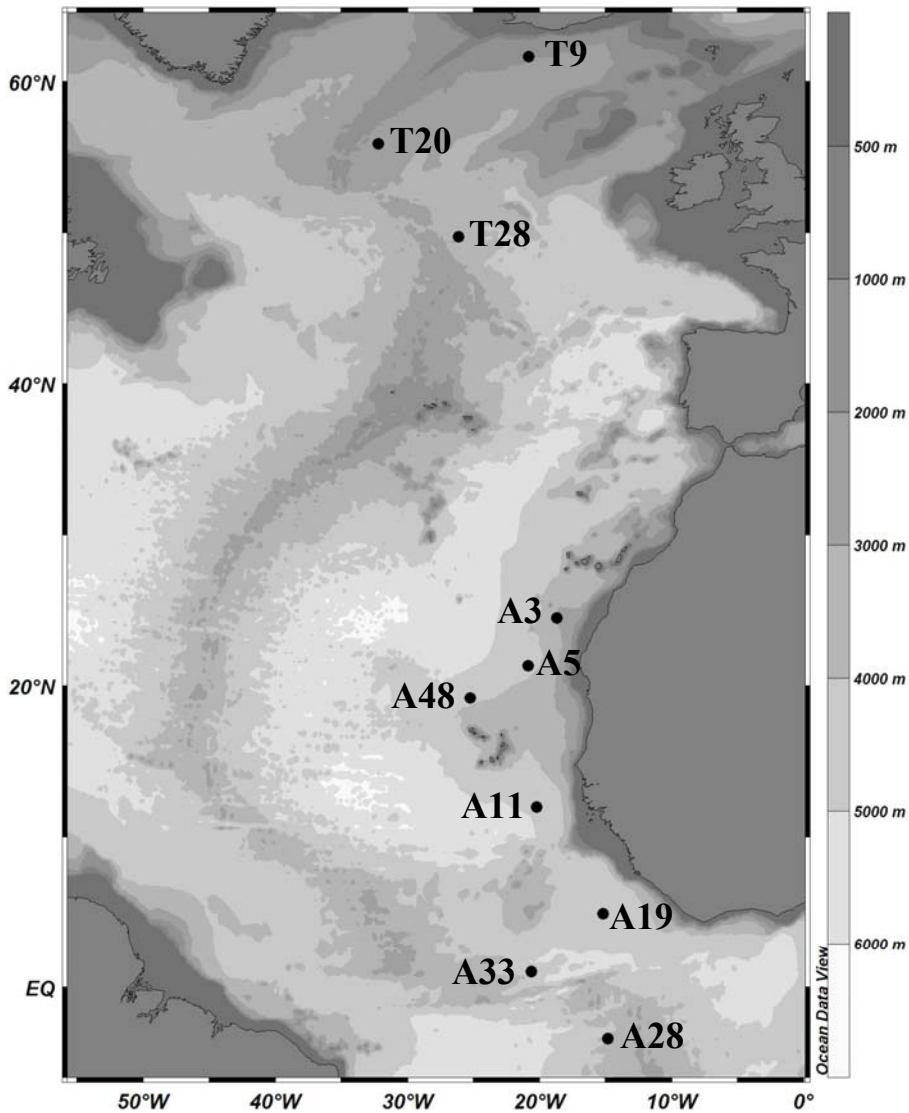
Cruise	Station	Depth (m)	MCGI 16S rDNA copies ml ⁻¹ MW (S.D.)	pSL12-like 16S rDNA copies ml ⁻¹ MW (s.d.)	Archaeal <i>amoA</i> gene copies ml ⁻¹ MW (s.d.)	β -proteobacterial <i>amoA</i> gene copies ml ⁻¹ MW (s.d.)
TRANSAT-1	9	2071	7.84x10 ⁴ (7.50x10 ³)	9.33x10 ² (7.52x10 ¹)	2.36x10 ³ (5.06x10 ²)	2.21x10 ² (5.13x10 ¹)
TRANSAT-1	9	1633	1.39x10 ⁴ (1.83x10 ³)	8.70x10 ² (2.63x10 ²)	4.98x10 ² (2.52x10 ¹)	7.20x10 ¹ (1.88x10 ¹)
TRANSAT-1	9	1014	1.35x10 ⁵ (4.01x10 ⁴)	3.74x10 ³ (1.20x10 ³)	2.39x10 ³ (3.87x10 ²)	6.28x10 ¹ (1.74x10 ¹)
TRANSAT-1	9	498	3.99x10 ⁴ (1.68x10 ⁴)	1.81x10 ³ (3.07x10 ²)	4.68x10 ³ (1.23x10 ³)	1.75x10 ² (5.55x10 ¹)
TRANSAT-1	9	99	6.31x10 ⁴ (1.09x10 ⁴)	1.19x10 ² (1.18x10 ¹)	7.08x10 ³ (1.98x10 ³)	1.99x10 ³ (7.19x10 ²)
TRANSAT-1	20	2500	7.53x10 ² (1.65x10 ²)	9.51x10 ² (8.07x10 ¹)	3.47x10 ² (2.02x10 ¹)	1.97x10 ¹ (2.90x10 ⁰)
TRANSAT-1	20	1897	1.12x10 ³ (1.25x10 ²)	3.69x10 ² (5.48x10 ¹)	4.48x10 ² (1.93x10 ²)	9.98x10 ⁰ (-)
TRANSAT-1	20	1147	1.36x10 ³ (1.71x10 ²)	2.82x10 ² (1.07x10 ²)	1.38x10 ² (4.28x10 ¹)	n.d.
TRANSAT-1	20	450	7.34x10 ³ (3.02x10 ³)	5.89x10 ² (1.20x10 ²)	3.56x10 ³ (1.11x10 ³)	1.60x10 ¹ (-)
TRANSAT-1	20	149	8.87x10 ³ (2.18x10 ³)	3.78x10 ² (7.95x10 ⁰)	2.23x10 ⁴ (3.90x10 ³)	5.93x10 ² (1.14x10 ²)
TRANSAT-1	28	3878	5.10x10 ⁴ (1.10x10 ⁴)	2.72x10 ³ (1.17x10 ³)	8.52x10 ¹ (7.78x10 ⁰)	8.93x10 ⁰ (3.28x10 ⁰)
TRANSAT-1	28	2999	5.36x10 ⁴ (1.44x10 ⁴)	3.53x10 ³ (1.41x10 ³)	1.40x10 ² (8.25x10 ⁰)	3.00x10 ⁰ (4.56x10 ⁻¹)
TRANSAT-1	28	1897	5.00x10 ⁴ (8.02x10 ³)	3.18x10 ³ (-)	7.98x10 ¹ (1.68x10 ¹)	n.d.
TRANSAT-1	28	597	2.13x10 ⁵ (4.69x10 ⁴)	4.52x10 ³ (1.53x10 ³)	5.52x10 ³ (1.07x10 ³)	9.78x10 ² (1.61x10 ¹)
TRANSAT-1	28	146	1.38x10 ⁵ (4.49x10 ⁴)	1.45x10 ³ (6.67x10 ²)	9.52x10 ³ (9.83x10 ²)	1.33x10 ² (2.70x10 ¹)
ARCHIMEDES-2	3	2750	2.30x10 ⁴ (2.03x10 ³)	1.67x10 ² (1.63x10 ¹)	7.80x10 ¹ (2.96x10 ¹)	5.23x10 ⁻² (1.54x10 ⁻²)
ARCHIMEDES-2	3	800	2.21x10 ⁵ (4.93x10 ⁴)	1.41x10 ³ (5.94x10 ¹)	7.98x10 ¹ (2.18x10 ¹)	n.d.
ARCHIMEDES-2	3	499	2.49x10 ⁴ (3.76x10 ³)	1.10x10 ³ (2.62x10 ²)	1.38x10 ² (1.95x10 ¹)	n.d.
ARCHIMEDES-2	3	249	1.24x10 ⁴ (9.64x10 ²)	4.69x10 ³ (9.04x10 ²)	7.48x10 ² (2.51x10 ²)	n.d.
ARCHIMEDES-2	3	99	2.60x10 ³ (4.98x10 ²)	2.30x10 ¹ (1.77x10 ⁰)	2.49x10 ³ (8.78x10 ¹)	8.85x10 ⁻² (4.53x10 ⁻²)
ARCHIMEDES-2	5	4001	1.54x10 ⁴ (2.31x10 ³)	5.79x10 ¹ (-)	2.04x10 ² (4.90x10 ¹)	6.58x10 ⁻¹ (5.23x10 ⁻²)
ARCHIMEDES-2	5	2748	4.01x10 ³ (4.62x10 ²)	6.10x10 ¹ (9.82x10 ⁰)	1.51x10 ¹ (2.29x10 ⁰)	6.02x10 ⁻² (1.10x10 ⁻²)
ARCHIMEDES-2	5	899	1.39x10 ⁴ (2.00x10 ³)	1.32x10 ² (7.71x10 ¹)	7.68x10 ¹ (2.09x10 ¹)	9.70x10 ⁻¹ (7.57x10 ⁻¹)
ARCHIMEDES-2	5	501	7.16x10 ⁴ (1.68x10 ⁴)	2.20x10 ² (1.20x10 ²)	1.59x10 ³ (2.76x10 ²)	9.28x10 ⁻¹ (2.47x10 ⁻²)
ARCHIMEDES-2	5	249	8.37x10 ⁵ (2.29x10 ⁵)	6.16x10 ² (7.07x10 ¹)	4.73x10 ³ (4.68x10 ²)	2.31x10 ⁰ (5.90x10 ⁻¹)
ARCHIMEDES-2	5	100	1.03x10 ⁵ (3.88x10 ⁴)	4.85x10 ² (1.54x10 ²)	2.32x10 ⁴ (2.62x10 ³)	5.94x10 ¹ (6.55x10 ⁰)

ARCHIMEDES-2	11	3999	1.44×10^4 (2.78×10^3)	4.38×10^2 (2.07×10^2)	1.03×10^1 (1.73×10^0)	1.41×10^0 (7.43×10^{-1})
ARCHIMEDES-2	11	2751	2.54×10^4 (1.15×10^3)	1.92×10^2 (1.04×10^1)	3.40×10^1 (5.97×10^0)	2.56×10^0 (1.32×10^0)
ARCHIMEDES-2	11	900	1.21×10^5 (4.37×10^4)	2.13×10^2 (6.17×10^1)	3.75×10^1 (1.08×10^1)	1.33×10^1 (-)
ARCHIMEDES-2	11	499	1.50×10^5 (3.05×10^4)	4.95×10^2 (2.84×10^1)	1.92×10^2 (3.50×10^1)	5.30×10^1 (8.84×10^0)
ARCHIMEDES-2	11	250	4.02×10^5 (7.04×10^4)	2.29×10^3 (3.64×10^2)	2.57×10^3 (6.73×10^2)	1.13×10^2 (3.84×10^1)
ARCHIMEDES-2	11	100	2.35×10^5 (3.53×10^4)	2.87×10^2 (5.75×10^1)	2.53×10^4 (7.42×10^2)	1.87×10^1 (-)
ARCHIMEDES-2	19	4001	2.86×10^3 (6.52×10^2)	5.73×10^2 (1.48×10^2)	3.80×10^0 (2.28×10^0)	n.d.
ARCHIMEDES-2	19	2751	3.23×10^4 (6.54×10^3)	2.85×10^2 (1.90×10^1)	6.05×10^0 (4.95×10^{-1})	4.33×10^{-2} (-)
ARCHIMEDES-2	19	900	1.60×10^4 (4.98×10^3)	6.75×10^2 (-)	1.84×10^1 (4.62×10^0)	n.d.
ARCHIMEDES-2	19	500	2.85×10^4 (6.88×10^3)	4.80×10^2 (2.09×10^2)	1.41×10^1 (-)	n.d.
ARCHIMEDES-2	19	251	1.27×10^5 (5.84×10^4)	1.57×10^3 (7.26×10^2)	8.15×10^2 (1.15×10^2)	n.d.
ARCHIMEDES-2	19	100	7.09×10^4 (3.77×10^4)	1.15×10^2 (1.18×10^1)	3.32×10^4 (1.05×10^4)	n.d.
ARCHIMEDES-2	28	2501	3.31×10^4 (3.71×10^3)	3.67×10^3 (1.46×10^3)	5.54×10^1 (7.27×10^0)	1.52×10^{-1} (-)
ARCHIMEDES-2	28	1799	5.19×10^2 (9.93×10^1)	2.87×10^3 (5.08×10^2)	6.46×10^0 (1.20×10^0)	n.d.
ARCHIMEDES-2	28	901	8.54×10^4 (2.30×10^3)	1.12×10^3 (5.30×10^2)	1.50×10^1 (4.22×10^0)	1.27×10^{-1} (4.54×10^{-2})
ARCHIMEDES-2	28	498	3.83×10^4 (2.07×10^4)	1.83×10^3 (5.90×10^2)	1.70×10^2 (1.27×10^1)	4.36×10^{-2} (2.09×10^{-2})
ARCHIMEDES-2	28	252	4.71×10^5 (8.41×10^4)	2.08×10^4 (2.25×10^3)	2.39×10^3 (4.87×10^2)	2.34×10^{-1} (7.09×10^{-2})
ARCHIMEDES-2	28	100	1.28×10^4 (4.17×10^3)	1.38×10^2 (9.85×10^0)	1.20×10^4 (8.52×10^2)	4.31×10^{-1} (1.67×10^{-1})
ARCHIMEDES-2	33	2502	5.35×10^4 (1.36×10^4)	2.90×10^3 (6.65×10^2)	6.47×10^1 (1.88×10^1)	n.d.
ARCHIMEDES-2	33	1800	2.70×10^4 (2.12×10^3)	1.53×10^3 (2.92×10^2)	3.53×10^1 (3.69×10^0)	n.d.
ARCHIMEDES-2	33	749	4.51×10^5 (1.88×10^5)	3.61×10^3 (2.02×10^2)	1.29×10^4 (1.91×10^3)	3.95×10^{-2} (-)
ARCHIMEDES-2	33	498	3.24×10^5 (4.54×10^4)	1.27×10^3 (2.9×10^2)	1.39×10^2 (3.09×10^1)	3.89×10^{-2} (-)
ARCHIMEDES-2	33	250	3.36×10^5 (9.88×10^4)	4.84×10^3 (1.04×10^3)	2.68×10^3 (1.24×10^2)	9.82×10^{-2} (4.36×10^{-2})
ARCHIMEDES-2	33	99	3.10×10^4 (3.67×10^3)	1.56×10^3 (3.36×10^2)	2.50×10^4 (1.20×10^3)	3.19×10^{-2} (-)
ARCHIMEDES-2	48	2500	1.69×10^3 (1.18×10^2)	2.63×10^2 (8.87×10^1)	6.65×10^0 (7.07×10^{-2})	1.62×10^{-1} (-)
ARCHIMEDES-2	48	1798	7.07×10^3 (1.90×10^3)	1.27×10^2 (4.36×10^1)	6.78×10^0 (1.31×10^0)	n.d.
ARCHIMEDES-2	48	901	5.04×10^3 (2.14×10^3)	7.18×10^1 (1.06×10^0)	1.73×10^0 (1.16×10^0)	1.96×10^{-1} (-)
ARCHIMEDES-2	48	499	2.43×10^5 (7.20×10^4)	7.48×10^2 (3.10×10^2)	5.33×10^2 (1.08×10^2)	1.81×10^{-1} (4.08×10^{-2})
ARCHIMEDES-2	48	249	1.83×10^4 (5.20×10^3)	5.93×10^2 (1.88×10^2)	3.24×10^2 (8.53×10^1)	1.21×10^{-1} (-)

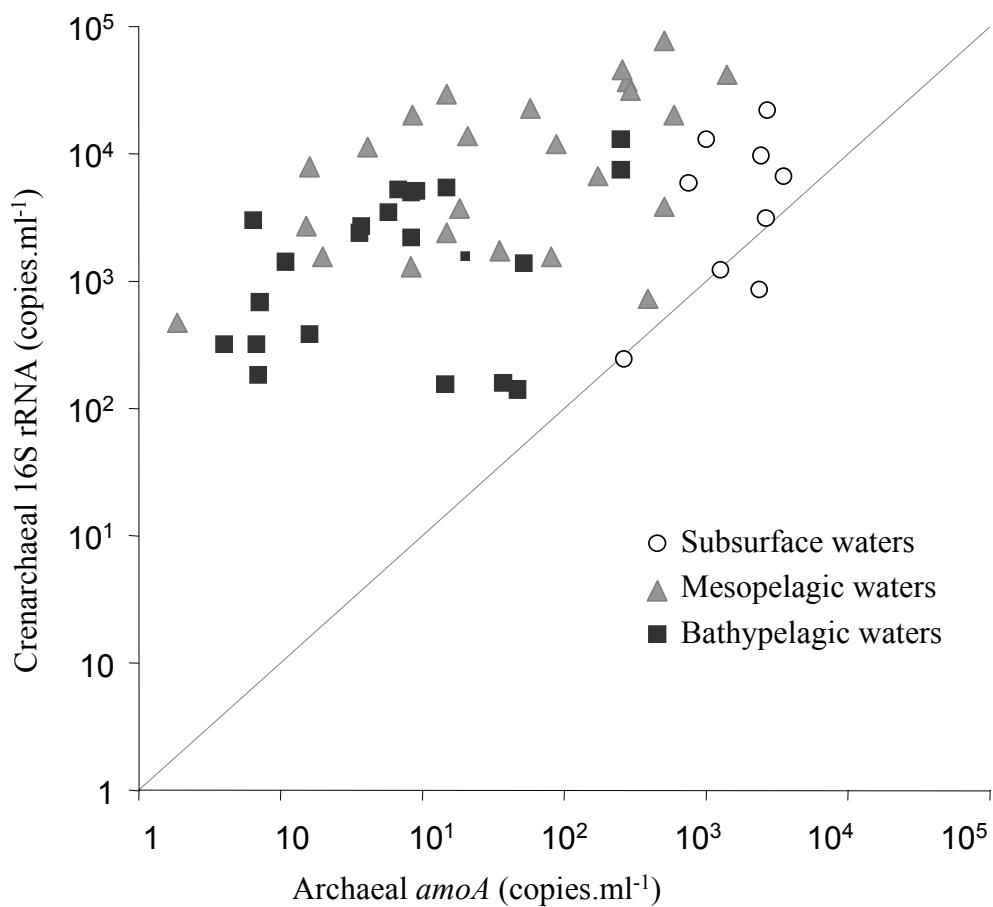
s.d.: standard deviation ; MW: marine water ; n.d.: not detected (i.e., no specific product formed upon Q-PCR or below detection limit).

Table S4 – Observed and estimated richness in archaeal *amoA* gene libraries

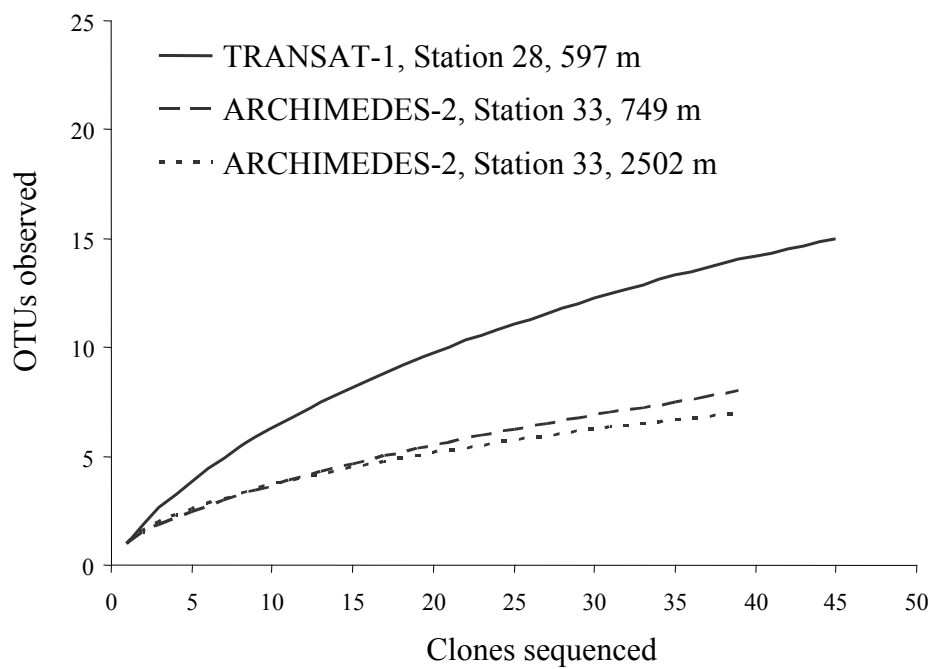
Station	Depth	Water Mass	No. of clones sequenced	No. of OTUs (2%)	Chao1 (2%)
TRANSAT-1, station 28	597 m	NACW	45	15	20
ARCHIMEDES-2, station 33	749 m	AAIW	39	8	18
ARCHIMEDES-2, station 33	2502 m	MNEADW	39	7	8.5



Supplementary Figure S1: Location of the stations where samples for quantitative PCR were taken during the TRANSAT-1 (T) and ARCHIMEDES-2 (A) cruise in the North Atlantic.

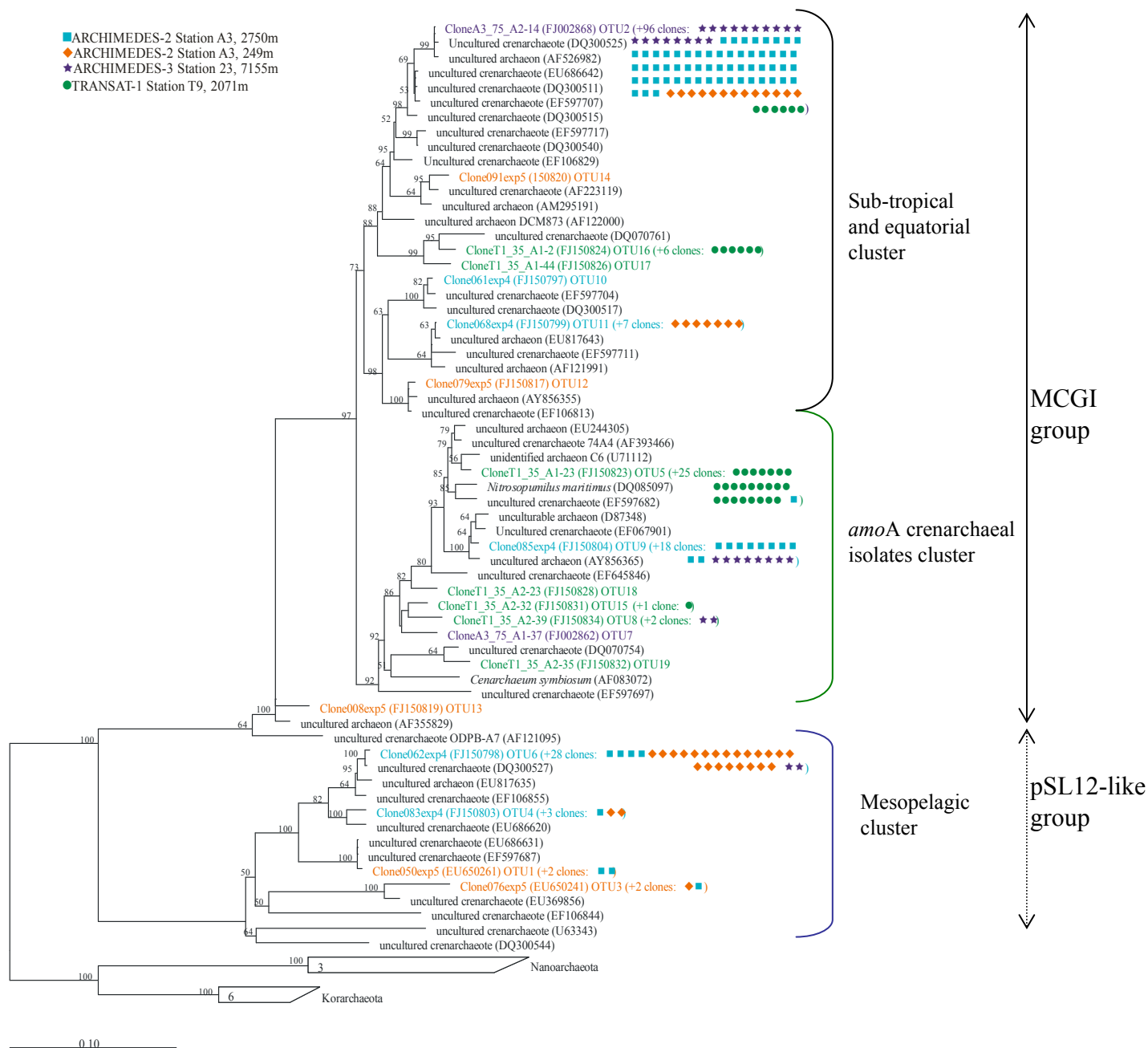


Supplementary Figure S2: Relationship between archaeal *amoA* gene copies and crenarchaeal 16S rRNA gene copies across all the stations and depths ($n=55$).

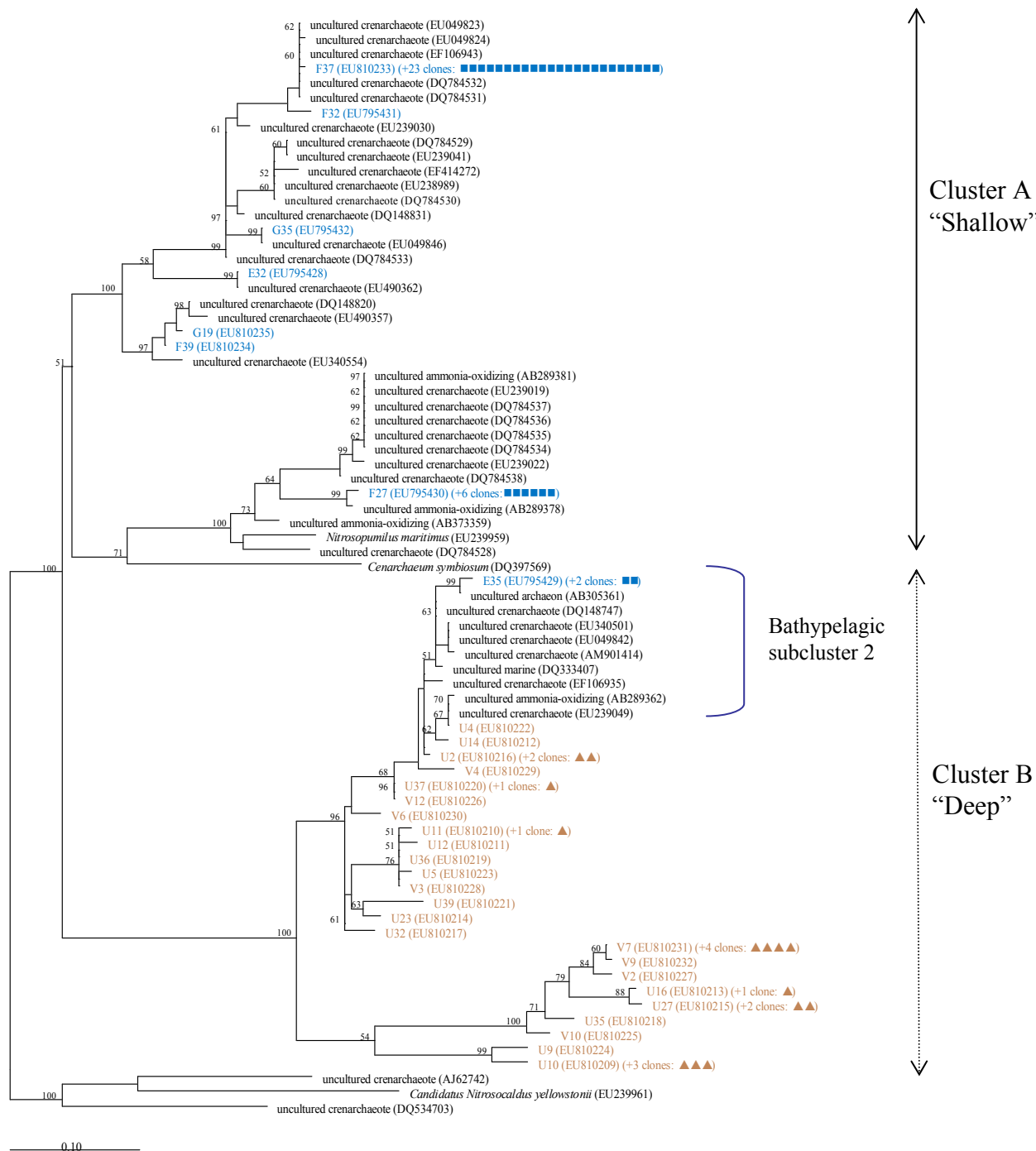


Supplementary Figure S3: Rarefaction curves indicating archaeal *amoA* richness within clone libraries. OTUs were defined as group of sequences differing by $\leq 2\%$ at the DNA level.

■ ARCHIMEDES-2 Station A3, 2750m
 ◆ ARCHIMEDES-2 Station A3, 249m
 ★ ARCHIMEDES-3 Station 23, 7155m
 ● TRANSAT-1 Station T9, 2071m



Supplementary Figure S4: Phylogenetic analyses of crenarchaeal 16s rRNA recovered from the North Atlantic Ocean at different stations and depths. Neighbor-joining bootstrap tree of entire 16S rDNA sequences (including the ones of this study) to which we added partial sequence using the ARB parsimony tool. Only one representative of sequence group $\geq 97\%$ identical is shown; additional symbols show the total number of clones represented by a sequence. Database sequences are given in black. Bootstrap values ($>50\%$) are indicated at branch point.



Supplementary Figure S5: Phylogenetic tree of archaeal *amoA* sequences recovered from the bathypelagic zone at Station A33 with the set of primers used in this study (in blue) and the primers of Francis *et al.* 11 (in brown). Neighbor-joining bootstrap tree of long *amoA*-like nucleotide sequences (including the sequences recovered with the primers of Francis *et al.* 11 EU810209-EU810232) to which we added our partial sequences (217 bp, EU795424-EU795460, EU810233-EU810235) and the sequences of Wuchter *et al.* 3 (217 bp, DQ784528-DQ784537) using the ARB parsimony tool. One representative of sequence group $\geq 99\%$ identical is shown; additional symbols show the total number of clones represented by a sequence. Database sequences are shown in black. Bootstrap values ($>50\%$) are indicated at branch point.

Supplementary notes (references)

31. Coolen, M.J.L. *et al.* Evolution of the methane cycle in Ace Lake (Antarctica) during the Holocene: Response of methanogens and methanotrophs to environmental change. *Org. Geochem.* **35**, 1151-1167 (2004).
32. Coolen, M.J.L. *et al.* Ancient DNA derived from alkenone-biosynthesizing haptophytes and other algae in Holocene sediments from the Black Sea. *Paleoceanography* **21**, PA1005, 10.1029/2005PA001188 (2006).
33. Ludwig, W. *et al.* ARB: a software environment for sequence data. *Nucleic Acids Res.* **32**, 1363-1371 (2004).
34. Schloss, P.D. & Handelsman, J. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl. Environ. Microbiol.* **71**, 1501-1506 (2005).
35. Delong, E.F. Archaea in coastal marine environments. *Proc. Natl. Acad. Sci. USA* **89**, 5685-5689 (1992).
36. Mehta, M.P., Huber, J.A. & Baross, J.A. Incidence of novel and potentially archaeal nitrogenase genes in the deep Northeast Pacific Ocean. *Environ. Microbiol.* **7**, 1525-1534 (2005).
37. Rotthauwe, J.H., Witzel, K.P. & Liesack, W. The ammonia monooxygenase structural gene *amoA* as a functional marker: Molecular fine-scale analysis of natural ammonia-oxidizing populations. *Appl. Environ. Microbiol.* **63**, 4704-4712 (1997).
38. Hornek, R. *et al.* Primers containing universal bases reduce multiple *amoA* gene specific DGGE band patterns when analysing the diversity of beta-ammonia oxidizers in the environment. *J. Microbiol. Methods* **66**, 147-155 (2006).