environmental microbiology

Environmental Microbiology (2011) doi:10.1111/j.1462-2920.2011.02623.x

Environmental factors determining ammonia-oxidizing organism distribution and diversity in marine environments

Nicholas J. Bouskill,[1*†] Damien Eveillard,[2] Diana Chien,[1‡] Amal Jayakumar[1] and Bess B. Ward[1]

[1]*Department of Geosciences, Guyot Hall, Princeton University, Princeton, NJ, USA.*

[2]*Computational Biology Group, LINA, Université of Nantes, CNRS UMR 6241, 2 Rue de la Houssiniere, 44300 Nantes, France.*

Summary

Ammonia-oxidizing bacteria (AOB) and archaea (AOA) play a vital role in bridging the input of fixed nitrogen, through N-fixation and remineralization, to its loss by denitrification and anammox. Yet the major environmental factors determining AOB and AOA population dynamics are little understood, despite both groups having a wide environmental distribution. This study examined the relative abundance of both groups of ammonia-oxidizing organisms (AOO) and the diversity of AOA across large-scale gradients in temperature, salinity and substrate concentration and dissolved oxygen. The relative abundance of AOB and AOA varied across environments, with AOB dominating in the freshwater region of the Chesapeake Bay and AOA more abundant in the water column of the coastal and open ocean. The highest abundance of the AOA *amoA* gene was recorded in the oxygen minimum zones (OMZs) of the Eastern Tropical South Pacific (ETSP) and the Arabian Sea (AS). The ratio of AOA : AOB varied from 0.7 in the Chesapeake Bay to 1600 in the Sargasso Sea. Relative abundance of both groups strongly correlated with ammonium concentrations. AOA diversity, as determined by phylogenetic analysis of clone library sequences and archetype analysis from a functional gene DNA microarray, detected broad phylogenetic differences across the study sites. However, phylogenetic diversity within physicochemically congruent stations was more similar than would be expected by chance. This suggests that the prevailing geochemistry, rather than localized dispersal, is the major driving factor determining OTU distribution.

Received 9 May, 2011; accepted 20 September, 2011. *For correspondence. E-mail njbouskill@lbl.gov; Tel. (+1) 510 486 7490; Fax (+1) 510 486 7152. Present addresses: †Ecology Department, Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94702, USA; ‡Graduate program in Microbiology, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA.

Introduction

Nitrification is an important link within the aquatic nitrogen cycle. Oxidation of ammonium to nitrate determines the distribution of reduced and oxidized forms of nitrogen, maintains the mesopelagic accumulation of nitrate and contributes to the regeneration of NO_3^- in the euphotic zone (Ward *et al*., 1989; Clark *et al*., 2008; Gruber, 2008; Ward, 2008). There is also evidence that the oxidation of ammonia supports chemoautotrophic production within the mesopelagic ocean (Ingalls *et al*., 2006; Hansman *et al*., 2009) and systems of low irradiance (Chen *et al*., 2009), fuelling the growth of heterotrophic organisms. The enzymatic oxidation of ammonia to nitrite is carried out by a few lineages in the *Beta-* and *Gammaproteobacteria* and by members of Group 1.1a and 1.1b of the *Crenarchaeota* (now known as a separate clade, the *Thaumarchaeota*). While the contribution of ammonia-oxidizing bacteria (AOB) to nitrification has long been assumed (see Ward, 2008), ammonia-oxidizing archaea (AOA) were only recently discovered. Following the identification of an ammonia-monooxygenase subunit A (*amoA*) gene on archaeal metagenomic scaffolds (Venter *et al*., 2004; Treusch *et al*., 2005), the genetic capacity of archaea to oxidize ammonia has been demonstrated to be widespread in nature (e.g. Francis *et al*., 2005). Further identification of *amo* genes and active oxidation of ammonia to nitrite has since been verified in pure cultures and enrichments (Konneke *et al*., 2005; de la Torre *et al*., 2008).

Given the potential for nitrogen limitation of primary productivity and the overall importance of nitrification to the global nitrogen (Yool *et al*., 2007) and carbon cycles (Austin *et al*., 2006), a more thorough understanding of the ecology of organisms catalysing this process is required. To this end, a number of culture experiments (Koops and Pommerening-Roser, 2001), *in situ* molecular-based studies (Francis *et al*., 2003; Ward *et al*., 2007; Bouskill *et al*., 2011) and genomic analysis (Arp *et al*., 2007) have done much to determine the

© 2011 Society for Applied Microbiology and Blackwell Publishing Ltd

environmental factors influencing AOB distribution. However, clarification of the ecological range of the AOA is hindered by a lack of representative pure cultures prohibiting ecophysiological examination or genome sequencing, although such studies continue to gather momentum (e.g. Martens-Habbena *et al*., 2009; Walker *et al*., 2010). In light of recent quantitative studies determining AOA to be the numerically dominant nitrifiers within many aquatic environments, in particular the open ocean (e.g. Mincer *et al*., 2007; Agogue *et al*., 2008), a more thorough characterization of the environmental factors regulating the diversity and distribution of AOA is necessary. To further characterize the ecology of ammonia-oxidizing organisms (AOO) this study examines the diversity and distribution of both AOA and AOB in a range of aquatic environments.

Erguder and colleagues (2009) reviewed the growing literature on environmental factors relating to the ecology of AOA and found that different AOA phylotypes (based primarily on *amoA* gene diversity) inhabit a substantial range of environmental conditions in terrestrial and aquatic environments. It has been hypothesized that archaea, as a domain, thrive in areas of chronic energetic stress (Valentine, 2007). This characterization is supported by the relatively limited number of studies comparing the ecology of AOA and AOB, concluding that AOA dominate in marine waters (Francis *et al*., 2007), where substrate concentrations can be an order of magnitude less than within coastal and estuarine zones. A potential rationale for their dominance of the open ocean, recently demonstrated from physiological experiments of *Nitrosopumilus maritimus* (Martens-Habbena *et al*., 2009), could be the higher substrate affinity of AOA over that of AOB. This feature would allow archaea to outcompete bacteria in nutrient-poor environments, consistent with correlations between nitrite production and AOA abundance in marine environments (Wuchter *et al*., 2006; Beman *et al*., 2008).

By comparing the distribution and abundance of AOO across a range of environmental conditions represented by the stations in this study, it is possible to identify the abiotic factors that are important in structuring AOA assemblages. The two Chesapeake Bay stations, CB100 and CB300, span the length of the estuarine gradient from relatively freshwater to nearly seawater salinity respectively. The Sargasso Sea [sampled in 2004 (denoted SS_100m) and in 2007 (BATS)] and North Atlantic (NA) represent contrasting open ocean conditions: the Sargasso Sea is an oligotrophic gyre, whereas NA is seasonally productive. Both the Arabian Sea (AS) and the Eastern Tropical South Pacific (ETSP) are sites of major oxygen minimum zones (OMZs) associated with highly productive surface waters. The core of the OMZs is characterized by anaerobic bacterial metabolisms including denitrification and anammox (Hamersley *et al*., 2007; Ward *et al*., 2009). The strong dissolved oxygen (DO) gradients of the OMZ water column permit examination of the influence of DO on the diversity and distribution of AOO. A second station in the ETSP provided shallow water samples from a continental shelf low-oxygen region.

This study examines the distribution in AOO in relation to oxygen and ammonium concentrations, as well as other important environmental variables, in the upper 250–500 m of the marine water column. Quantitative PCR was used to determine the relative abundance of the AOB and AOA *amoA* genes across different environments. Simultaneously, the 16S rRNA of two *Crenarchaeota* groups containing AOAs, marine group one (MG1) and pSL12 (Barns *et al*., 1996), were also quantified. AOA community composition was examined by sequencing *amoA* clone libraries and DNA hybridization to a custom 70-mer oligonucleotide microarray. The AOA microarray employs archetype probes drawn from 1500 environmental sequences (using data available up to December 2008) as described for previous generations of AOB microarrays (Ward *et al*., 2007; Bouskill *et al*., 2011).

Results

Physicochemical characteristics

Measured values of NH_4^+, NO_2^-, NO_3^-, S and T for all stations are given in Table 1. Trends in water chemistry typical of estuarine gradients were recorded between the oligohaline (CB100) and polyhaline (CB300) stations of the Chesapeake Bay. Salinity increased approximately eightfold from north to south, while average water column temperature decreased. CB100 maintained higher average NH_4^+ and NO_3^- concentrations at both the surface and deep depths than those measured at CB300 (Table 1).

In the open ocean stations (AS, ETSP, SS, NA), NH_4^+ concentrations were consistently below detection limit (~0.1 μM) and salinity was relatively invariant (35.6 ± 0.6), but higher than at CB100 and CB300 (Table 1). Measured concentrations of DO in the OMZ regions show clear gradients between the surface and the oxycline at the upper boundary of the OMZ (Table 1). The AS exhibits both primary (30–82 m) and secondary (250 m) nitrite maxima (PNM and SNM respectively) characteristic of OMZs. In the AS, NO_2^- concentrations can vary from trace levels to higher than 4 μM (USJGOFS, http://www1.whoi.edu/). At the SNM, NO_2^- concentrations have been observed as high as 6.5 μM. However, at the stations sampled for this study only traces of NO_2^- were observed at the PNM and the SNM was approximately 2.4 μM. This NO_2^- concentration is lower than previously

© 2011 Society for Applied Microbiology and Blackwell Publishing Ltd,

© 2011 Society for Applied Microbiology and Blackwell Publishing Ltd, *Environmental Microbiology*

Table 1. Site parameters – including latitude and longitude and individual depths.

Location	Station code	Latitude, longitude	Date	Depth (m)	Sample code	Total station depth (m)	Temp. (°C)	Salinity	DO (μM)	NO_2^- (μM)	NO_3^- (μM)	NH_4^+ (μM)	Volume filtered (ml)	Application[a]	References
Chesapeake Bay	CB100_S	39°20′N,	April 2001–2004	2	CB100_S	11	17.4 (7.5)	2.2 (3)	N.M.	1.9 (2.5)	58.7 (34.2)	5.9 (2.8)	~720	1	Ward *et al.* (2007)
	CB100_D	76°13′W		9	CB100_D		17.1 (8)	3.9 (3.3)	N.M.	1.5 (1.7)	52.5 (33.5)	5.3 (3)		1/2	
	CB300_S	37°11′N,	April 2001–2004	2	CB300_S	20	9 (0.3)	18.5 (1.8)	N.M.	0.3 (0.2)	7 (7.2)	0.4 (0.3)	~1330	1/3	
	CB300_D	76°6′W		17	CB300_D		8.6 (0.5)	22 (4.9)	N.M.	0.2 (0.1)	3.6 (3.2)	2.3 (1.7)		1/2/3	
Sargasso Sea	SS	32°46′N,	April 2002	1	SS_1m	4300	20.4	36.4	227.5	0.12	0.15	N.M.	3000	1	
		64°29′W		100	SS_100m		19.5	36	226.7	N.M	0.15	N.M.		1/2/3	
				250	SS_250m		18.2	36	196.8	0.14	0.1	N.M.		1	
Bermuda Atlantic Time Series	BATS	31°29′N,	March 2009	20	BATS_20m	4680[b]	20.1	36.6	226	0.1	0.08	N.M.	3000	3	
		64°6′W		120	BATS_120m		19.7	36.5	221.1	0.14	0.1	N.M.		3	
North Atlantic	NA_01	60°85′N, 26°63′W	May 2008	4.7	NA_2001	1556	8.9	35.3	278.8	N.M	10.8	N.M	2000	3	
	NA_19	61°49′N, 26°34′W		4	NA_2019	1499	9.6	35.3	291.4	N.M	13.4	N.M		3	
	NA_21	61°46′N, 25°95′W		4.5	NA_2021	1403	9.5	35.3	290.8	N.M	8.5	N.M		3	
Eastern tropical South Pacific	Station 9	15°37.92′S,	October 2005	80	ETSP_St.9_80m	981	13.1	34.9	1.8	6.21	8.5	0.6	4300	1/3	Ward *et al.* (2008)
		76°7.75′W		150	ETSP_St.9_150m		12.2	34.9	1.9	4.59	11.8	B.D.		1	
				250	ETSP_St.9_250m		11.4	34.8	2	3.51	26.5	0.11		1/2	
				500	ETSP_St.9_500m		8	34.5	9.3	0.00	28.9	B.D.		1	
Eastern tropical South Pacific	Station 23	13°15.00′S,	October 2005	20	ETSP_St.23_20m	100	13.5	35	7.9	0.90	17.5	0.6	4300	1/3	
		76°29.99′W		60	ETSP_St.23_60m		13.3	35	1.9	2.20	12	0.2	8600	1/3	
				100	ETSP_St.23_100m		13.2	35	2.4	4.72	13.8	0.4		1	
Arabian Sea	AS	15°00.00′N, 64°00.00′E	September 2007	2.5	AS_2.5m	3900	27.5	35.7	152.7	0.02	0.41	0.3	13100	1	Jayakumar *et al.* (2009); Ward *et al.* (2009)
		19°22.98′N, 66°39.35′E		60	AS_St.1_60m	3151	23.3	36.1	22.2	0.24	16.7	B.D.	12300	1/2/3	
		16°44.98′N, 64°30.08′E		80	AS_80m	3151	21.8	36	68.9	0.05	21.85	0.32	14200	1/3	
		15°00.00′N, 64°00.00′W	October 2004	250	AS_St.2_250m	3900	16.7	35.7	3	2.03	18.4	B.D.	12000	1/2/3	

a. Extracted DNA used in: 1 – quantitative PCR; 2 – clone library construction; 3 – microarray hybridization.
b. Steinberg *et al.* (2001).
N.M., not measured; B.D., below detection. The large standard deviations on the values above reflect large seasonal and annual variation. [N.B. Arabian Sea NH_4^+ concentrations were measured at a nearby station (19°N, 67°E)].

reported concentrations for the AS OMZ (Jayakumar *et al.*, 2009). DO concentration in the ETSP decreased below 10 μmol l^{-1} at approximately 50 m, and was below detection (10 μM) down to 400 m depth, where DO began to increase again. Similar to the AS, ETSP NO_2^- concentrations were elevated, but a distinct SNM was difficult to identify under the depth resolution of the sampling.

Quantitative PCR

At each station the abundance of AOB and AOA was measured by quantification of the respective *amoA* genes (Fig. 1). AOB were the dominant AOO at CB100 and the AOA were more abundant at CB300. While the archaeal *amoA* gene was amplified from all stations examined, bacterial *amoA* could not be amplified from all of the marine stations, particularly in the Arabian Sea (amplicons were obtained in only one of seven samples examined), suggesting AOB were either below the detection limit of the assay or have a more restricted distribution. Total AOA *amoA* copy numbers ranged from 5.7×10^2 ($\pm 2.7 \times 10^2$) at CB100 to 1.5×10^6 ($\pm 6 \times 10^5$) copies ml^{-1} water at 100 m at ETSP station 23, and AOA *amoA* copy numbers were generally the highest of all four qPCR assays performed (Fig. 1). When detected, AOB *amoA* gene abundances ranged from ~95 ± 13 in the surface waters of the Arabian Sea to 3.8×10^5 ($\pm 8 \times 10^4$) copies ml^{-1} at 150 m of the OMZ ETSP station (Stn. 9). The ratio of AOA: AOB varied by environment from 0.7 to 1602 at CB100 and the ETSP (at 500 m) respectively. In general, the AOA *amoA* showed significantly higher gene abundances in marine samples compared with AOB *amoA*, except at 150 m in the ETSP, where AOB *amoA* genes were approximately fivefold more abundant than AOA *amoA* genes (Fig. 1).

The abundance of the specific crenarchaeal groups to which AOA are assumed to belong, MG1 and pSL12, varied by environment and did not always correlate with AOA *amoA* gene abundance. There was a strong relationship, however, between AOA *amoA* and MG1 16S rRNA gene abundances when combining data from all stations ($R^2 = 0.62$, $P = 0.02$). No relationship was recorded between the AOA *amoA* gene abundance and pSL12 16S rRNA abundance at any station. However, comparing trends between AOA *amoA* gene abundance and the combined 16S rRNA (MG1 and pSL12) gene abundance data set also found a positive relationship ($R^2 = 0.61$, $P = 0.0001$).

Partial correlation analysis was used to quantify monotonic relationships between AOO abundance and environmental variables across all stations. Positive correlations were recorded between AOB abundance and both NH_4^+ and NO_2^- concentrations when controlling for either DO (NH_4^+: $r = 0.71$, $P = 0.001$; NO_2^-: $r = 0.5$, $P = 0.007$) or salinity (NH_4^+: $r = 0.69$, $P = 0.001$; NO_2^-: $r = 0.45$, $P = 0.02$) (Table 2). AOB abundance was strongly negatively correlated with temperature when taking into account salinity ($r = 0.82$, $P = 0.001$) or DO ($r = 0.85$, $P = 0.001$), and was also negatively correlated with salinity when controlling for DO ($r = 0.5$, $P = 0.01$).

Archaeal nitrifier abundance as inferred from AOA *amoA* gene abundance showed a negative correlation with NH_4^+ concentrations when controlling for temperature ($r = -0.5$, $P = 0.014$). Additionally, when controlling for variability in salinity or temperature, a statistically significant correlation was found with NO_2^- concentrations ($r = 0.4$, $P = 0.04$; $r = 0.55$, $P = 0.003$ respectively). AOA *amoA* gene abundance also showed a negative relationship with DO when controlling for other factors (salinity; $r = -0.5$, $P = 0.006$; Temp; $r = -0.65$, $P = 0.001$), consistent with high AOA *amoA* abundance within the core of the AS and ETSP OMZ. Finally, AOA *amoA* gene abundance showed a significant relationship with salinity when controlling for temperature ($r = 0.62$, $P = 0.004$) or DO ($r = 0.4$, $P = 0.03$) (Table 2).

AOA amoA *phylogeny*

Over 400 AOA *amoA* gene fragments (693 bp) were sequenced from six samples (Table 1) and showed a maximum nucleotide sequence divergence of 22% (±3%). Using an operational taxonomic unit (OTU) level of 3%, 10% or 15% (the array probes are also defined at a 15% divergence), a total of 152, 58 or 31 distinct OTUs were detected in the combined data set. The range of OTUs found within individual clone libraries, at the 10% level, was from 3 in the Sargasso Sea (SS_100m) to 21 at 250 m in the Arabian Sea (Table S1). Rarefaction curves (Fig. S1), also constructed at a 10% cut-off level, and OTU estimator analysis (S_{chao} and S_{ace} in Table S1) demonstrated good library coverage except for CB300 and the AS_St.2_250m. Shannon diversity indices reflected this analysis, with the highest diversity index at AS_St.2_250m and CB300 and the lowest at the SS_100m.

Phylogenetic relationships of representative OTUs, based on translated *amoA* sequences from all libraries, resolved into 5 clusters with > 70% bootstrap support. The amino acid tree (Fig. 2) (also see the Unifrac analysis with nucleic acid sequences, Fig. S2) provides evidence that physicochemically similar stations harbour phylogenetically related AOA populations, regardless of geographic proximity. This indicates that environment type is a good predictor of AOA community structure.

Cluster 1 includes sequences found primarily from the deeper sampling depths of the AS (AS_St.2_250m) and ETSP (ETSP_St.9_250m). Phylogenetic similarity between the two OMZs is further supported by Unifrac

© 2011 Society for Applied Microbiology and Blackwell Publishing Ltd, *Environmental Microbiology*

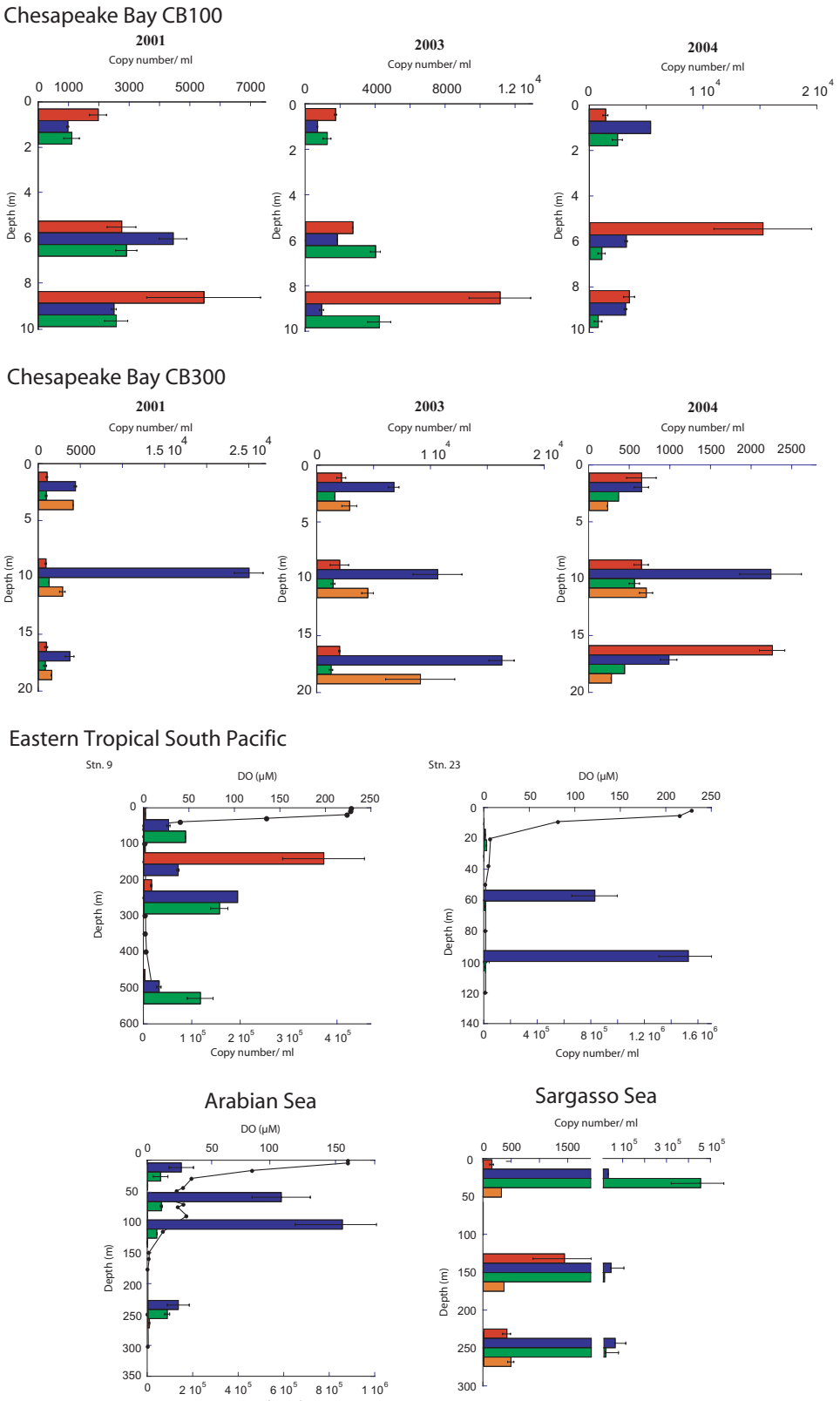


Fig. 1. Quantitative PCR analysis in depth profiles at several stations. Bars represent the mean of triplicate reactions, while the error bars are the standard deviations around the mean. Red – AOB *amoA*, blue – AOA *amoA*, green – MG1 16S rRNA, orange – pSL12 16S rRNA.

© 2011 Society for Applied Microbiology and Blackwell Publishing Ltd, *Environmental Microbiology*

Table 2. Results of the partial correlation analysis relating AOA *amoA* copy numbers to environmental variables.

	Correlation with						
Controlling for	NH_4^+	NO_2^-	NO_3^-	DO	Temp.	Sal.	Depth
	AOB						
Salinity	0.69*	0.45*	–0.1	–0.17	–0.81**	–	N/A
Temperature	0.35	0.28	–0.15	–0.24	–	0.31	0.27
Dissolved oxygen	0.71*	0.5*	0.32	–	–0.85**	–0.48	0.19
	AOA						
Salinity	0.04	0.39	0.38	–0.5*	–0.23	–	0.46*
Temperature	–0.46*	0.55*	–0.26	–0.65**	–	0.62**	0.61**
Dissolved oxygen	–0.38	–0.04	–0.31	–	0.15	0.42*	0.30

Statistically significant relationships are asterisked (* $\leq$ 0.05, ** $\leq$ 0.001).

analysis and construction of DNA-based phylogenetic trees (Fig. S2). This cluster includes several relatively divergent clades that are distinct from all the others, possibly indicating AOA populations adapted to low DO.

Cluster 2 (Fig. 2) contains clones and related environmental sequences predominantly representing an open-ocean deep-water ecotype, similar to cluster B described by Beman and colleagues (2008). Sequences from the AS_St.2_250m and ETSP_St.9_250m, but none from the other samples, fell into cluster 2. AS_St.2_250m sequences were distributed throughout cluster 2, while the ETSP_St.9_250m sequences formed two distinct subclusters. A number of AS_St.2_250m sequences were phylogenetically similar to published sequences from the bathypelagic realm of the North Atlantic Ocean (Agogue *et al*., 2008) and from mesopelagic depths (770 m) at station ALOHA, Hawaii and Monterey Bay, CA (Mincer *et al*., 2007) (Fig. 2 and Fig. S2). AS_St.2_250m sequences were also phylogenetically affiliated with sequences previously reported from the Peruvian OMZ (Lam *et al*., 2009). A small number of AS_St.2_250m sequences clustered with other *amoA* sequences from other open ocean stations within cluster 3, including the AS_St.1_60m and SS_100m stations. Amino acid sequences from AS_St.1_60m were distributed throughout clusters 3 and 4 and showed similarity to sequences from a eutrophic coastal region of the South China Sea (Dang *et al*., 2008), the Gulf of Mexico and from coral symbionts (Siboni *et al*., 2008).

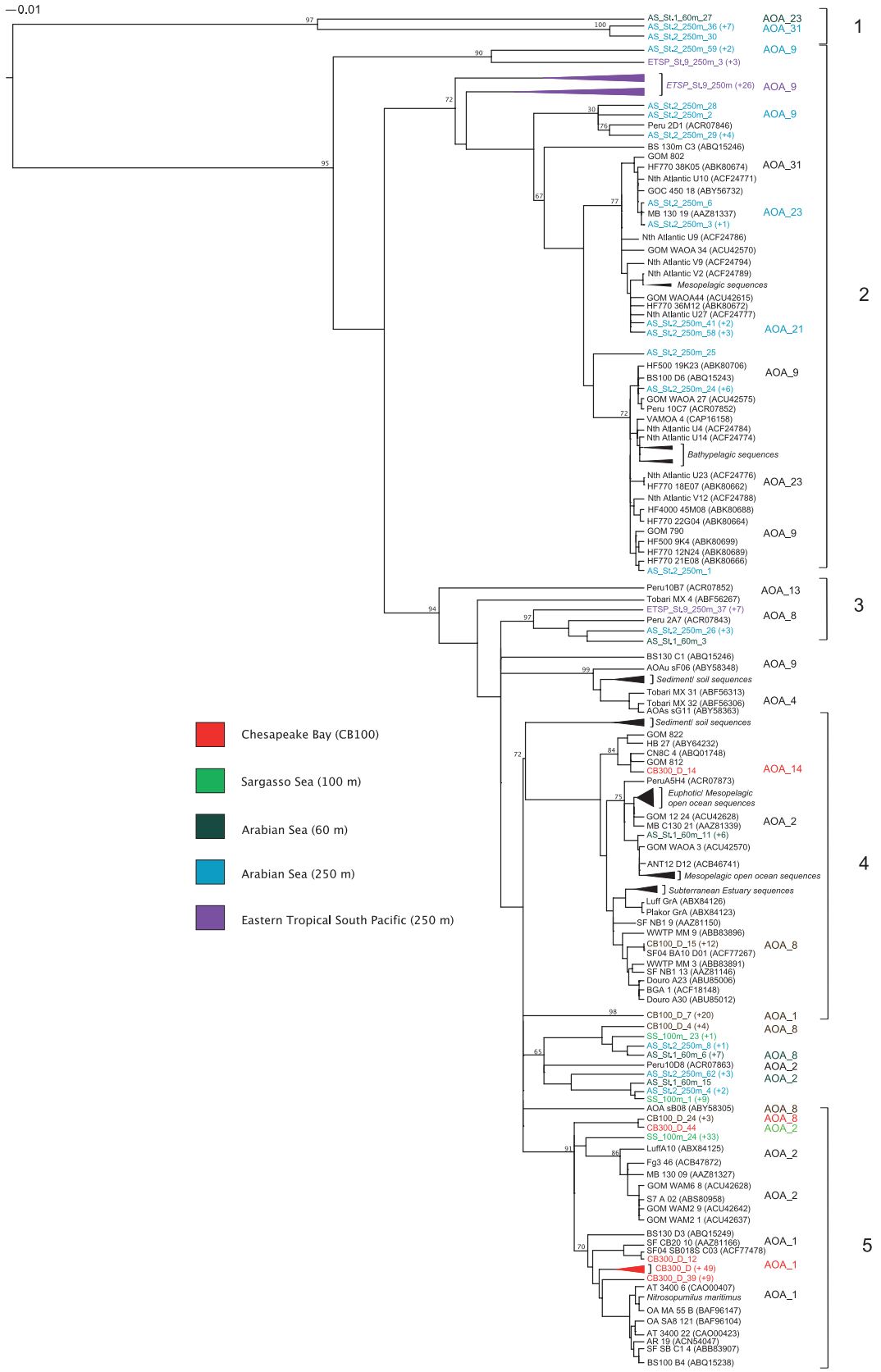
Sargasso Sea AOA communities were characterized by low diversity – the 50 SS *amoA* amplicons sequenced in this study distributed into three OTUs. Two of these OTUs clustered with open ocean sequences from Arabian Sea. However, the bulk of sequences from SS_100m fell into one OTU that was phylogenetically related to sequences amplified from the tissue of marine sponges in cluster 5 (Steger *et al*., 2008).

Sequences cloned from the Chesapeake Bay water column fell into clusters 4 and 5. The closest relative of the largest number of CB100 sequences were sequences from the Douro estuary, San Francisco Bay and a number of sequences from wastewater and groundwater treatment works, possibly representing organisms adapted to high nutrient concentrations. CB100 sequences were also related to those from CB300. Sequences from CB300 formed a unique cluster related to an environmental sequence from an aquarium biofiltration system (Urakawa *et al*., 2008). CB300 sequences were also closely related to *N. maritimus* (the majority of sequences were between 97% and 99% similar), which was isolated from a marine aquarium (Konneke *et al*., 2005). In the case of the Chesapeake Bay, the Unifrac environmental matrix and DNA tree (Fig. S2) support the overarching conclusions of the phylogenetic analysis that geochemically similar environments harbour phylogenetically similar *amoA* sequences (Fig. 2). AOA assemblages from the oligo- and polyhaline environments of the Chesapeake Bay were more similar to each other than to sequences retrieved from the oceanic marine environments.

The Unifrac environmental matrix showed that sequences in clusters 4 and 5 were closely associated with those found under similar environmental conditions (Fig. S2) and are likely affiliated with the 'shallow-water' ecotype A (Beman *et al*., 2008). Of the DNA sequences considered in this case, the AS_St.1_60m sequences stand apart as the most divergent in the environmental matrix and separate into discrete mono-environmental clusters in the phylogenetic trees (Fig. S2). The SS

Fig. 2. Maximum-likelihood phylogenetic tree constructed from translated DNA sequences from AOA *amoA* clone library analyses of various environments. Bootstrap values were calculated over 100 replicates. CB100 and CB300 have been abbreviated as CB1 and CB3, respectively, in the sequence names in the tree.

© 2011 Society for Applied Microbiology and Blackwell Publishing Ltd, *Environmental Microbiology*



© 2011 Society for Applied Microbiology and Blackwell Publishing Ltd, *Environmental Microbiology*

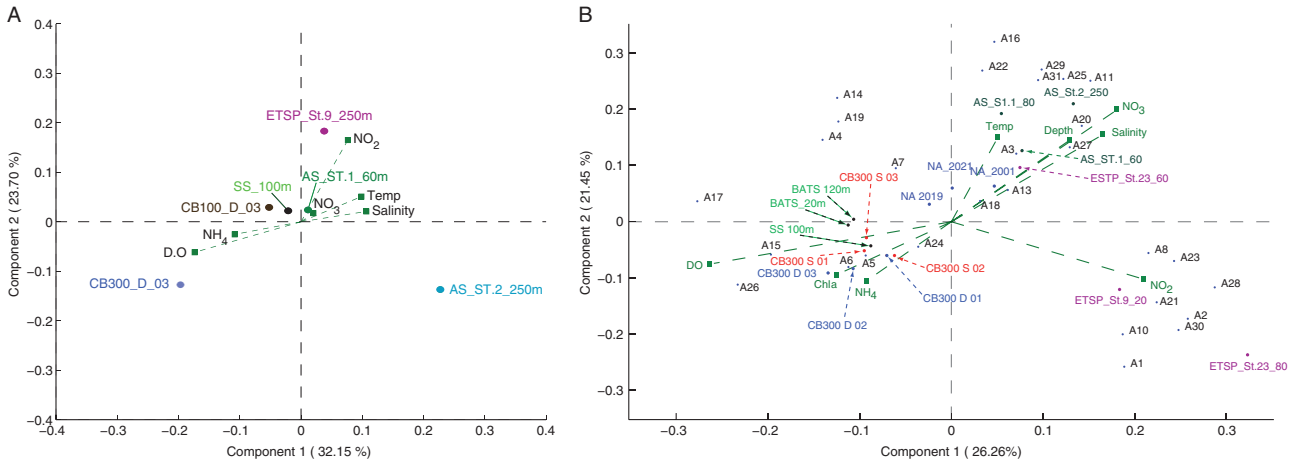


Fig. 3. Principal components analysis based on a correlation matrix combining pre-correlated physicochemical and biological factors. Data for the plots are taken from (A) clone library data (OTU cut-off at 15%) and (B) array data. The same geographic locations are grouped by colour. Physicochemical abbreviations: Temp, temperature; NH_4, ammonium; NO_2, nitrite; NO_3, nitrate; D.O, dissolved oxygen.

sequences show phylogenetic similarity with sequences from the Gulf of California and Monterey Bay sampled at similar depths.

Principal components analysis (PCA), correlating both physicochemical and clone library matrices, found that ammonium and DO concentrations positively influenced the genotype-environment relationship at the polyhaline Chesapeake Bay station (CB300) (Fig. 3A). AOA communities at AS_St.1_60m and at ETSP_St.9_250m were positively ordinated with NO_2^- concentrations, T and S, but less so with nitrate. The PCA did not reproduce the patterns of phylogenetic similarity shown in the tree; for example, there was a clear differentiation between the AS and ETSP stations at 250 m in the PCA, although the sequences from these two samples formed related clusters (Fig. 2). In addition to sequences that overlap two stations (at a 10% OTU cut-off level), the PCA also takes into account the relative abundance of those sequences and thus complements the phylogenetic analysis. According to this analysis, the AOA assemblage structure of several stations (CB100, SS_100m and AS_St.2_250m) was unrelated to the environmental variables considered.

AOA archetypes microarray analysis

Classification of community composition using Shannon evenness detected variation in the overall contribution of individual archetypes to the total hybridization signal among environments. Evenness was highest in water column samples from the surface waters of the North Atlantic (0.89 where a value = 1 would mean total signal intensity was split equally across all probes), the AS_St.1_60m (0.86) and the deeper samples at both the AS (AS_St.2_250m) and the ETSP (ETSP_St.9_250m) (0.85). Evenness was lowest at BATS_120m (0.69), consistent with the same measurement estimated from the clone libraries at SS_100m (Fig. 4).

The most significant relative fluorescence ratio (RFR) signals across all samples were attributed to archetypes AOA_9 and AOA_12, followed by archetypes AOA_4, AOA_26 and AOA_17 (Fig. 4). AOA_9 represents a number of uncultivated environmental sequences derived primarily from the marine water column, notably the Monterey Bay, Gulf of California and off Hawaii at station ALOHA. This archetype was a significant component of all samples but particularly high within the coastal ETSP station (at 20 m), and the Arabian Sea (at 80 m and 250 m). The AOA_12 archetype was a significant component of CB300 samples in all three years, but was a smaller component of open ocean sites at the deep ETSP station and NA_01. This is unsurprising as the archetype was compiled from sequences derived primarily from representatives of Tobari sediments, a hypernutrified Mexican estuary and also soil clones. Archetype AOA_17 representing sequences from terrestrial and aquatic sediments, specifically, the San Francisco Bay Estuary and the deep sea at the Western Pacific Continental Margin, was a large component of CB300 and BATS samples. A number of archetype probes produced low signals in all the samples. These archetypes include AOA_1, which includes *N. maritimus* and the largest number of sequenced *amoA* genes deposited into GenBank. AOA_1 was detected at CB300 and at the ETSP stations, but at very low levels.

A PCA based on the array signals and environmental data (Fig. 3B) found a similar overall distribution of estuarine and open ocean stations as that detected in the clone library analysis. CB300 samples, both surface AND deep,

© 2011 Society for Applied Microbiology and Blackwell Publishing Ltd, *Environmental Microbiology*

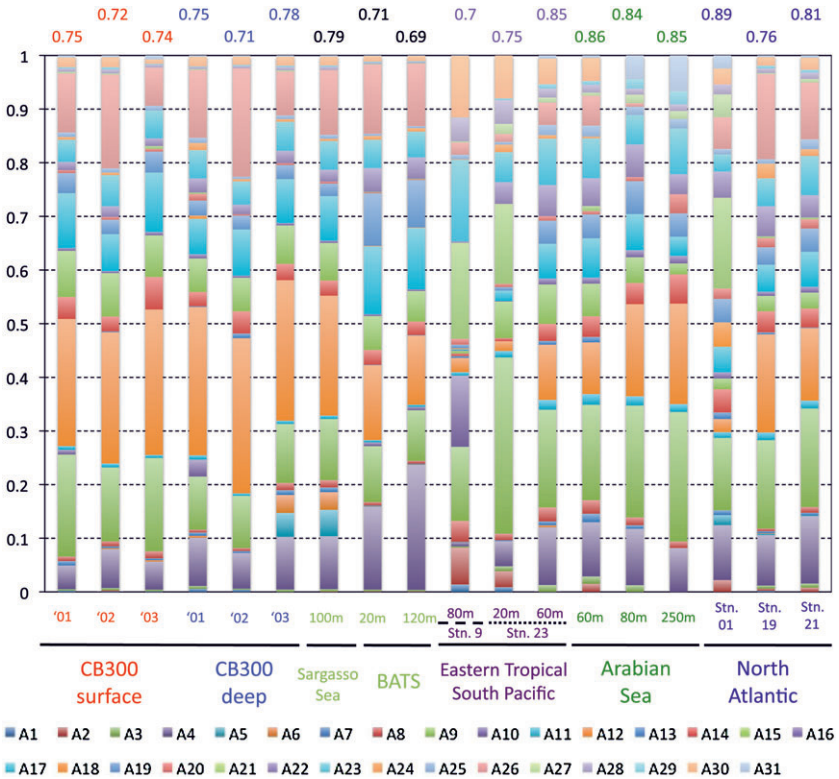


Fig. 4. Composite bar plots showing RFR for all archetype probes from the 20 station analysed in this study. The stations are grouped by geographical location. The corresponding evenness indices (*J*) are given above the individual sample bars.

generally formed tight intra-sample clusters, yet also show inter-sample clustering, with both depths ordinating closely together, indicating phylogenetic overlap between the two stations. This clustering was mainly due to the higher average contribution of several archetypes (AOA_5, _6, _24 and _26) and the low hybridization signal for a number of archetypes common to the other samples (e.g. AOA_3, _20, _27). ETSP_St.9_80m was the most dissimilar sample mainly due to the contribution to the RFR from probes not commonly found in other samples (notably AOA_1, _2, _10, _21, _23 and _30).

In general, stations that clustered together by ordination had similar geochemistry (e.g. NA_19 and NA_21), most notably the tight cluster of the AS stations and ETSP_St.23_60m that are all taken from water column depths of very low DO (between 1.8 and 68 μmol l^{-1}). This is consistent with the clustering of sequences from low oxygen depth in Fig. 2. However, the two samples from the ETSP OMZ (ETSP_St.9_80m and ETSP_St.23.20m) showed no relationship with DO concentrations, but were ordinated with nitrite concentration, which usually co-varies with DO in OMZs. These AOA assemblage–environment relationships also noted that the surface and deep depths of CB300 harboured assemblages positively related to dissolved oxygen, ammonium and chlorophyll-*a* concentrations but negatively ordinated with temperature, salinity and nitrate concentrations (Fig. 3B). The cluster of OMZ samples (excluding ETSP_St.9_80m and ETSP_St.23.20m) showed the opposite relationship with all these variables.

Discussion

Quantitative distribution of AOO

Quantification of AOO *amoA* abundances in the water column of the Chesapeake Bay estuary over a 3-year period (April 2001, 2003 and 2004) recorded a transition in dominant ammonia-oxidizing populations between stations CB100 and CB300. AOB were the dominant AOO in the freshwater eutrophic station (AOA : AOB = 0.6 ± 0.5), while AOA dominated the more saline, mesotrophic water column towards the mouth of the Bay (AOA : AOB = 6.4 ± 8.4). At first glance, a number of previous studies quantifying estuarine populations of AOA and AOB simultaneously would seem to contradict the present results. AOA were more abundant in the freshwater regions of the San Francisco Bay Estuary (Mosier and Francis, 2008), the Fitzroy River Estuary (Abell *et al.*, 2010) and Huntington Beach subterranean Estuary (Santoro *et al.*, 2008). If the relative distributions of AOA and AOB are considered in relation to ammonium concentrations rather than salinity, however, agreement between this and previous studies (e.g. Mosier and Francis, 2008; Santoro *et al.*,

© 2011 Society for Applied Microbiology and Blackwell Publishing Ltd, *Environmental Microbiology*

2008) is found. The contrasting statistical relationships between the different AOO and ammonium concentrations suggest that AOB are associated with high substrate concentrations, and might be considered copiotrophic when compared with oligotrophic AOA. This relationship also extends to artificial systems such as wastewater or groundwater treatment plants (Urakawa *et al*., 2008; Wells *et al*., 2009) and soil microcosms (Verhamme *et al*., 2011), where AOB are often found to be the dominant AOO. Furthermore, Reed and colleagues (2010) demonstrated that increasing available substrate in an AOA *amoA*-dominated aquifer stimulated AOB communities increasing their abundance relative to AOA. In further comparison with previous studies, AOB abundance at station CB100 was higher than water column values (~10^2 copies ml^{-1}) reported for a freshwater lake in Northwestern China (Jiang *et al*., 2009), using identical primer pairs. AOA abundances at CB300 were in the lower range of values reported in physicochemically similar water columns (e.g. Urakawa *et al*., 2010). More recently, Hollibaugh and colleagues (2011), using different qPCR primers, found that AOA (~10^4 copies ng^{-1} DNA) vastly outnumbered AOB in a polyhaline estuary in the South Eastern United States, despite high ammonium concentrations.

Quantification of AOB and AOA *amoA* abundances at a range of marine stations found the crenarchaeal *amoA* gene to numerically dominate open ocean AOO communities. In contrast, the AOB *amoA* was either undetected or recorded in low abundance. The one exception to this was in the low-oxygen waters of the deep ETSP stations. The dominance of AOA in the marine environment appears to be a general feature of AOO population distributions (e.g. Beman *et al*., 2008; Santoro *et al*., 2010). Reasons for such contrasting niche differentiation between AOB and AOA are beginning to accumulate from genomic and physiological data of cultured representatives (e.g. Arp *et al*., 2007; Martens-Habbena *et al*., 2009; Walker *et al*., 2010). Physiological features that allow archaea to persist under low-nutrient conditions characteristic of most oceanic systems include the reduction in futile ion cycling from isoprenoidal ether-linked lipid membranes (Valentine, 2007), a remarkably high specific affinity for ammonium (some 200 times higher than cultivated AOB, Martens-Habbena *et al*., 2009) and the biosynthesis of unique compatible solutes (Walker *et al*., 2010). Such features likely permit AOA to effectively outcompete AOB for available NH_4^+ throughout the water column.

The highest abundance of AOA *amoA* genes was recorded at the lowest DO in the ETSP and in the oxycline of the AS. The magnitude of the highest *amoA* copy numbers recorded in the AS (8.6×10^5 copies ml^{-1}) and ETSP (1.5×10^6 copies ml^{-1}) is higher than that reported for the Gulf of California OMZ (Beman *et al*., 2008), but on the order of that previously reported within the OMZ associated with the Peruvian and Chilean upwelling system (Lam *et al*., 2009; Molina *et al*., 2010).

The presence of high AOA and AOB abundance in low-oxygen waters of two geographically distinct oceanic regions is intriguing. Bacterial ammonia oxidization is energetically favourable under oxic conditions. *In situ* analysis of oxygen concentrations using sensitive amperometric sensor for ultra-low O_2 concentrations (Revsbech *et al*., 2009) reveals the OMZ to be essentially anoxic (< 10 nM O_2). While the presence of the *amoA* gene in this study is not evidence of AOO activity, both the process of ammonia oxidation (Lipshultz *et al*., 1990) and the activity of AOA (Stewart *et al*., 2011) have been demonstrated within the upper boundary of the Peruvian and Chilean OMZ respectively. Nevertheless, and despite the fact this study has not expressly accounted for the role of physical processes (e.g. particle-attached transport or horizontal mixing), it would seem unlikely that a group whose primary metabolism is aerobic could maintain the inferred population numbers in a permanently anoxic environment.

At the present time, little is understood about the metabolic capability of AOA inhabiting the OMZ. AOB, on the other hand, can oxidize ammonium under microaerophilic conditions, and some groups may favour low oxygen concentrations. It is also possible that these phylogenetically diverse AOA maintain an active anaerobic metabolism. Bacterial ammonium oxidizers can generate ATP and grow, albeit slowly, under anaerobic conditions via nitrifier denitrification (Schmidt and Bock, 1997; Schmidt, 2008). The recent publication of the *N. maritimus* genome (Walker *et al*., 2010) revealed no apparent anaerobic metabolism, and the presence of a nitrite reductase (Hallam *et al*., 2006) in the *Crenarchaea* appears to be a pathway towards intracellular nitrite detoxification rather than a respiratory function (Bartossek *et al*., 2010). However, many of the AOA *amoA* sequences from the OMZ are phylogenetically diverse when compared with previously published sequences and seemingly unique to the OMZ. These sequences may therefore represent organisms adapted to survive under anaerobic conditions. However, given the range of *amoA* : 16S rRNA ratios found in this study (between 1.5 and 38), it is possible that the abundances reported here are overestimates of the true population number. Ratios greater than 1 appear to be a feature of marine AOA communities across the California Current (Santoro *et al*., 2010), the Gulf of California (Beman *et al*., 2008) and ALOHA (Mincer *et al*., 2007).

AOA genotypic diversity

Previous large-scale studies of AOA *amoA* (Francis *et al*., 2005) and the archaeal 16S rRNA (Auguet *et al*., 2009)

© 2011 Society for Applied Microbiology and Blackwell Publishing Ltd, *Environmental Microbiology*

have recorded close evolutionary relatedness between organisms inhabiting physicochemically similar environments, independent of geographical distance between the environments. This was also recorded in the diversity component of the present study. In the Chesapeake Bay two distinct AOA communities appear to inhabit the two stations. A similar phylogenetic division has been previously recorded across estuarine gradients for AOB, with estuarine populations characterized by *Nitrosomonas* lineages in freshwater areas with increasing *Nitrosospira* dominance towards polyhaline areas (Francis *et al*., 2003; Ward *et al*., 2007). AOA *amoA* sequences from CB100, the freshwater end of the bay, were phylogenetically affiliated with sequences from environments of high ammonium concentrations such as wastewater and groundwater treatment plants (Park *et al*., 2006; van der Wielen *et al*., 2009), indicating an AOA genotype adapted to eutrophic conditions. A large number of sequences from CB300 were closely related to the *N. maritimus* sequence, the only group of sequences in this study that appeared to be closely related to the cultivated strain.

The phylogenetically distant cluster found at the AS and ETSP OMZ might represent AOA adapted to survival under anaerobic conditions, for which the presence of the *amoA* gene represents either a secondary metabolism or an ancestral remnant no longer active. However, this study also observed considerable diversity of the OMZ samples. Sequences from these stations were related to water column sequences from both oxic and anoxic environments, including the oxic marine water column of the Gulf of California (Mincer *et al*., 2007; Beman *et al*., 2008) and the Peruvian OMZ (Lam *et al*., 2009; Molina *et al*., 2010).

AOA genotypic distribution assessed with an amoA *microarray*

The higher throughput and greater depth of clade detection afforded by the microarrays allowed a broader investigation of the relationship between environmental factors and AOA functional diversity than was possible with clone libraries from a smaller selection of depths. While the percentage of a community detected by the microarray is dependent on the development of an adequate sequence database, the application of an *amoA* microarray in this study achieved greater coverage of the AOA community than that detected by a clone library. Greater coverage has also been demonstrated for 16S rRNA arrays compared with 16S rRNA clone libraries (e.g. DeSantis *et al*., 2007). The AOA communities reconstructed using the array are consistent with diversity patterns observed in the clone library data.

The AOA archetypes could be separated into three main categories: (i) ubiquitous archetypes important in all samples, (ii) archetypes significant only at specific stations and (iii) a small number of archetypes rarely contributing to signal intensity. Cosmopolitan distribution of *amoA* sequences has been recorded previously (see Francis *et al*., 2005) and in the present study, three *amoA* archetypes fell into this first category and contributed significant signal hybridization throughout the data set: archetypes AOA_4, AOA_9 and AOA_12. Sequences contributing to archetype AOA_9 were marine in origin and included representatives from low-oxygen environments, which might explain the ubiquitous nature of this archetype within the stations sampled here (Fig. 2). However, sequences making up archetypes AOA_4 and AOA_12 were mainly from a recognized soil/sediment clade (e.g. Francis *et al*., 2005) rather than water column samples.

Sequences matching common archetypes (e.g. AOA_9) were distributed throughout the phylogenetic tree (Fig. 2). While this phylogeny is based on entire sequenced fragments (635 bp), the archetypes are defined on the basis of a highly variable 70-mer region within the amplified fragment. Thus, the archetype algorithm maximizes the discrimination among probes, but may not group sequences coherently within the overall phylogeny. The fact that sequences matching AOA_12 were not found in the clone library suggests that sequences homologous to this archetype region are poorly amplified by the PCR primers used for cloning.

Nonetheless, environmentally similar samples ordinated together on the basis of relatively higher RFR of certain archetypes (Fig. 3B), which, taking into account the phylogeny of the sequences making up the archetype, can be considered to be characteristic of the samples. For example, the major probes separating the Chesapeake Bay stations from the marine stations (e.g. AOA_5, _6, _14 and _26) represent sequences that were originally sampled from the estuarine environment. Similarly, the probes that differentiate two of the ETSP samples (ETSP_St.9_80 and ETSP_St.23_250m) contain phylogenetic sequences from the marine environment. Interestingly, archetype AOA_1 was rarely detected in the array signal. While this archetype represents the greatest number of published sequences as well as *N. maritimus*, its low signal strength in these samples covering a very wide geographic range suggests that *N. maritimus* may not be representative of oceanic AOA clades. The same conclusion can be drawn from the phylogeny described above (Fig. 2), as most of the AOA *amoA* sequences retrieved, with the exception of a number of sequences from CB300, also were not closely related to *N. maritimus*.

The PCA combining both biological (RFR) and physicochemical variables (Fig. 3B) showed slightly different clustering between stations than that recorded by correspondence analysis (CA) of the biological data alone

© 2011 Society for Applied Microbiology and Blackwell Publishing Ltd,

(Fig. S3; also see *Supporting information* for further discussion of the CA data). This could be expected from the use of two different techniques to display the data (PCA versus CA) and through the correlation of a matrix combining all variables. In the PCA, the CB samples were positively related with DO, ammonium and chlorophyll, and negatively with nitrate and temperature, while several OMZ samples clustered together and had the opposite relationships with all of these variables. The only OMZ samples not included in this OMZ cluster were positively ordinated with nitrite, another important characteristic of OMZ waters. While this analysis does not point to a single environmental variable that controls AOA community composition, it does show that communities are not randomly assembled and suggests regional scale controls on community composition. In the current analysis, the first two principal components accounted for 40% of data set variability, indicating that not all the physicochemical factors important in structuring AOA communities are accounted for in this study.

Conclusion

In the present study, a combination of molecular techniques and multivariate statistics has demonstrated that AOA assemblages are more likely to be phylogenetically related when sampled from environmental similar conditions, regardless of geographical proximity. This relationship could be attributed to a few environmental parameters (notably concentrations of ammonium and dissolved oxygen). The current study extends understanding of the important environmental factors structuring AOO assemblages under natural conditions; however, further work is required to understand the temporal variability, response to perturbation of AOA communities and the extent of adaptive capacity of seemingly environmentally constrained populations. These research avenues are important for determining how resilient nitrifying assemblages might be under anticipated anthropogenic forcing of the marine nitrogen cycle (Duce *et al.*, 2008).

Experimental procedures

Station description, sample collection and DNA extraction

A description of the sampling locations, water column depths and physicochemical conditions of the different stations is given in Table 1 (for a more detailed description of individual sampling efforts see references in Table 1; furthermore, a link to an online map of sample location is provided in *Supporting information*). Briefly, for all samples, water was filtered onto 0.2 μm Sterivex filters (volumes given in Table 1), which were drained capped and frozen in liquid nitrogen. DNA was extracted from thawed filters using the Puregene DNA kit (Gentra, Minneapolis, MN). Subsets of the sample extracts were subjected to the three different kinds of analyses described below (see applications column of Table 1 for a summary).

Quantitative PCR

DNA concentrations were quantified by Nanodrop ND-1000 spectrophotometer (Nanodrop technologies, Wilmington, DE) in duplicate. Real-time PCR amplification of the archaeal and bacterial *amoA*, MG1 and pSL12 16S rRNA genes were run on the Stratagene MX-3000 (Stratagene, LaJolla, CA) using primers and annealing temperatures listed in Table S2. All reactions were performed in either duplicate or triplicate on 96-well reaction plates alongside the relevant triplicate amplified standards (described below). Reaction mixtures contained 12.5 μl (1× final concentration) of Ex-taq Premix perfect SYBR Green Real-time master mix (Takara, Madison, WI), 0.4 mmol l^{-1} each primer, 2 μl of DNA template (12 ng) made up to 20 μl with ultra-pure sterile water. All reactions were carried out under the following conditions: initial denaturation at 95°C for 30 s, followed by 40 cycles of denaturation at 95°C for 5 s, annealing at the gene-specific temperature (Table S2) for 20 s and extension at 72°C for 30 s. Following amplification, products were denatured to produce a melt curve to assess the stringency of product formation. Products were also visualized on a 2% agarose gel to evaluate amplification specificity.

Standards for qPCR were generated from cloned PCR products of the AOA *amoA*, MG1 and pSL12 16S rRNA genes. Standards for AOB qPCR experiments were serial dilutions of *amoA* amplified from DNA extracted from *Nitrosomonas cryotolerans* cultures, made up to the same concentrations as the plasmid standards. All plasmids were based on clones amplified from DNA extracted from the middle depth of the SS in April 2002 and cloned as described below. The standards were 10-fold serial dilutions of the plasmid from 1.0×10^{-2} to 1.0×10^{-8} ($R^2 = 0.999$ for all standard curves). Copy number ng^{-1} DNA was calculated as described previously (Bouskill *et al.*, 2007) and expressed finally as copy number ml^{-1}. Samples with a cycle threshold value (C_T) equal to, or above that of the negative control were treated as below the detection limit of the assay.

amoA *PCR, library construction and sequencing*

A 635 bp fragment of the AOA ammonia monooxygenase (*amoA*) gene was amplified using the *amoA*-F/R primer set (Table S2) using 2 U of the PHUSION polymerase (Finnzymes, NEB). Triplicate PCR reactions were carried out under the following conditions; 98°C for 30 s, 30 cycles of 98°C (5 s), 53°C (10 s) and 72°C (15 s) with a final extension for 7 min at 72°C. Pooled PCR products were cloned into vectors as described previously (Jayakumar *et al.*, 2004) and inserts were sequenced in a forward and reverse direction on the ABI 310 genetic analyser (Applied biosystems) using the M13F and R primers. Raw sequences were downloaded to Sequencher v.4 and contigs generated automatically.

Nucleic acid sequences were translated into amino acid sequences (http://ca.expasy.org/) and aligned alongside

© 2011 Society for Applied Microbiology and Blackwell Publishing Ltd, *Environmental Microbiology*

environmental *amoA* sequences from GenBank using the CLUSTAL X program (Thompson *et al*., 1997). Bootstrapped neighbour-joining phylogenetic trees (using a Jukes-Cantor distance matrix) were created using Paup* (v. 3.1). OTU determination, library richness estimates (Chao 1, ACE) and rarefaction curves were generated using DOTUR (Schloss and Handelsman, 2005). Shannon diversity and evenness indices were calculated using the VEGAN program in R (Oksanen *et al*., 2008; R Development Core Team, 2008).

To detect potential patterns of phylogenetic relatedness under similar geochemical conditions, *amoA* sequences from GenBank searches and the current sequences were compiled and each individual sequence tagged as belonging to a specific environment. Sequences were included in the analysis according to the following criteria: (i) high-quality data: sequences were > 500 bp, containing no ambiguous bases and could be translated into an amino acid sequence, (ii) one representative OTU sequence per 10% OTU cut-off cluster per study was used in building the final tree rather than all sequences from that study, and (iii) each sequence was obtained from studies that adequately reported important geochemical factors and initially grouped into ocean or estuarine stations. Sequences were aligned and a neighbour-joining tree constructed as above and uploaded to Unifrac (Lozupone *et al*., 2006). The aggregate phylogenetic similarity between sequences from different environments was assessed by generating an environmental matrix. The newick file from this matrix was used to construct a dendrogram representing the phylogenetic relatedness between stations. All *amoA* sequences from this study are available in GenBank under Accession Nos JF969736–JF970026.

AOA microarray analysis

The array was developed following the archetype array approach outlined and employed previously (e.g. Ward *et al*., 2007; Bulow *et al*., 2008) with 90-mer oligonucleotide probes that included an AOA *amoA*-specific 70-mer region and 20-mer control region (5′-GTACTACTAGCCTAGGCTAG-3′) bound onto a glass slide. The design and spotting of the *amoA* probes has been described previously (Ward *et al*., 2007; Ward, 2008). The 31 AOA archetype probes represent 1500 sequences taken from clone libraries built from terrestrial, aquatic and geothermal environments and the sequences from the current study. The probes and their characteristics are listed in Table S3.

Target preparation microarray hybridization and data analysis

Hybridization targets were prepared from isothermally amplified whole-genome products. Briefly, DNA extracts were initially digested with a ChoI restriction endonuclease (Roche, Indianapolis, IN) and the product used as a template in a whole-genome amplification (WGA) at 37°C using random octamers and a Phi29 polymerase (Genomiphi, V2, GE Healthcare, Piscataway, NJ). WGA products were labelled with an amino-allyl-dUTP (Ambion, Austin, TX) during linear amplification using a Klenow enzyme (Invitrogen, Carlsbad, CA). The reaction contained 3.6 mM d(AGC)TP, 0.4 mM dTTP and 4.4 mM dUaa and the *amoA* fragment was isothermally amplified for 1.5 h. The Klenow product was purified and conjugated with Cy3 as described previously (Ward *et al*., 2007).

Cy3-labelled PCR product (200 ng) was combined with 2× hybridization buffer (1× final concentration; Agilent, Santa Clara, CA), 0.25 pmol of a Cy5-labelled complementary 20-mer standard oligonucleotide and incubated at 95°C for 5 min before being cooled to room temperature. Samples were hybridized to triplicate arrays by overnight incubation at 65°C and washed as described previously (Ward *et al*., 2007). The arrays were scanned with an Agilent laser scanner (Agilent Technologies, Palo Alto, CA) and analysed using the Gene Pix Pro 6.0 software (Molecular Devices, Sunnyvale, CA). Various steps employed to filter and quantify the signal hybridization data have been described previously (Taroncher-Oldenburg *et al*., 2003). Signal hybridization intensities were standardized to the total fluorescence across the AOA probe set to give a RFR allowing comparison between different arrays (Ward *et al*., 2007). All of the original array files are available at GEO (Gene Expression Omnibus; http://www.ncbi.nlm.nih.gov/projects/geo/) at NCBI (National Center for Biotechnology Information) under GEO Accession No. GPL13883.

Statistical analysis

Univariate statistical analysis was performed using the statistics toolbox (v. 7.2) in MATLAB (v. 7.6). QPCR data normality was tested using a D'Agostino-Pearson K2 test and non-normally distributed data were log transformed. A two-way ANOVA was used to test for significant differences between DNA copy numbers of the same gene at different water column depths. Additionally, ANOVAs also compared bacterial *amoA* copy numbers with archaeal *amoA* copy numbers at the same depth and archaeal *amoA* with crenarchaeal and pSL12 16S rRNA copy numbers where appropriate. Correlations between *amoA* abundance (of either AOA or AOB) and environmental variables common to all stations [ammonium (NH_4^+), nitrite (NO_2^-), nitrate (NO_3^-), temperature (T), salinity (S)] were assessed using partial correlation analyses to control for different variables. At station CB100 and CB300, particulate carbon (PC), particulate nitrogen (PN), PC : PN ratio, urea, total dissolved nitrogen (TDN), dissolved organic carbon (DOC), dissolved organic nitrogen (DON) were also included in the above analysis. While DO concentrations were added only for stations ETSP and AS. PCA correlating environmental and biological (sequence and array hybridization data) matrices was used to assess potential interactions between AOA community assemblages and environmental factors common to all assessed stations (NH_4^+, NO_2^-, NO_3^-, Chlorophyll-*a*, DO, T, S and depth) as described previously (Bouskill *et al*., 2011).

Acknowledgements

We thank M.J. Perry, T. Rynearson and N. Briggs, who collected the North Atlantic samples and provided the accom-

© 2011 Society for Applied Microbiology and Blackwell Publishing Ltd, *Environmental Microbiology*

panying metadata for those samples. The microarray was printed by Donna Storton (Princeton University) and we gratefully acknowledge the Princeton University Microarray Facility. This work was supported by NSF grants to B.B.W.

References

Abell, G.C., Revill, A.T., Smith, C., Bissett, A.P., Volkman, J.K., and Robert, S.S. (2010) Archaeal ammonia oxidizers and *nirS*-type denitrifiers dominate sediment nitrifying and denitrifying populations in a subtropical macrotidal estuary. *ISME J* **4:** 286–300.

Agogue, H., Brink, M., Dinasquet, J., and Herndl, G.J. (2008) Major gradients in putatively nitrifying and non-nitrifying Archaea in the deep North Atlantic. *Nature* **456:** 788–791.

Arp, D.J., Chain, P.S., and Klotz, M.G. (2007) The impact of genome analyses on our understanding of ammonia-oxidizing bacteria. *Annu Rev Microbiol* **61:** 503–528.

Auguet, J.C., Barberan, A., and Casamayor, E.O. (2009) Global ecological patterns in uncultured Archaea. *ISME J* **4:** 182–190.

Austin, A.T., Sala, O.E., and Jackson, R.B. (2006) Inhibition of nitrification alters carbon turnover in the Patagonian Steppe. *Ecosystems* **9:** 1257–1265.

Barns, S.M., Delwiche, C.F., Palmer, J.D., and Pace, N.R. (1996) Perspective on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proc Natl Acad Sci USA* **93:** 9188–9193.

Bartossek, R., Nicol, G.W., Lanzen, A., Klenk, H.P., and Schleper, C. (2010) Homologues of nitrite reductases in ammonia-oxidizing archaea: diversity and genomic context. *Environ Microbiol* **12:** 1075–1088.

Beman, J.M., Popp, B.N., and Francis, C.A. (2008) Molecular and biogeochemical evidence for ammonia oxidation by marine Crenarchaeota in the Gulf of California. *ISME J* **2:** 429–441.

Bouskill, N.J., Barnhart, E.P., Galloway, T.S., Handy, R.D., and Ford, T.E. (2007) Quantification of changing *Pseudomonas aeruginosa soda, htpX* and *Mt* gene abundance in response to trace metal toxicity: a potential *in situ* biomarker of environmental health. *FEMS Microbiol Ecol* **60:** 276–286.

Bouskill, N.J., Eveillard, D., O'Mullan, G.D., Jackson, G.A., and Ward, B.B. (2011) Seasonal and annual reoccurrence in betaproteobacterial ammonia-oxidizing bacterial population structure. *Environ Microbiol* **13:** 872–886.

Bulow, S., Francis, C.A., Jackson, G.A., and Ward, B.B. (2008) Sediment denitrifier community composition and *nirS* gene expression investigated with functional gene microarrays. *Environ Microbiol* **10:** 3057–3069.

Chen, Y., Liqin, W., Boden, R., Hillebrand, A., Kumaresan, D., Moussard, H., *et al.* (2009) Life without light: microbial diversity and evidence of sulfur- and ammonium-based chemolithotrophy in Movile Cave. *ISME J* **3:** 1093–1104.

Clark, D.R., Rees, A.P., and Joint, I. (2008) Ammonium regeneration and nitrification rates in the oligotrophic Atlantic Ocean; implications for new production estimates. *Limnol Oceanogr* **53:** 52–62.

Dang, H., Zhang, X., Sun, J., Li, T., Zhang, Z., and Yang, G. (2008) Diversity and spatial distribution of sediment ammonia-oxidizing Crenarchaeota in response to estuarine and environmental gradients in Changjiang Estuary and East China Sea. *Microbiology* **154:** 2084–2095.

DeSantis, T.Z., Brodie, E.L., Moberg, J.P., Zubieta, I.X., Piceno, Y.M., and Andersen, G.L. (2007) High-density universal 16S rRNA microarray analysis reveals broader diversity than typical clone library when sampling the environment. *Microb Ecol* **53:** 371–383.

Duce, R.A., LaRoche, J., Altieri, K., Arrigo, K.R., Baker, A.R., Capone, D.G., *et al.* (2008) Impacts of atmospheric nitrogen on the open ocean. *Science* **320:** 893–897.

Erguder, T.H., Boon, N., Wittebolle, L., Marzorati, M., and Verstraete, W. (2009) Environmental factors shaping the ecological niches of ammonia oxidizing archaea. *FEMS Microbiol Ecol* **33:** 855–869.

Francis, C.A., O'Mullan, G.D., and Ward, B.B. (2003) Diversity of ammonia monooxygenase (*amoA*) genes across environmental gradients in Chesapeake Bay sediments. *Geobiology* **1:** 129–140.

Francis, C.A., Roberts, K.J., Beman, J.M., Santoro, A.E., and Oakley, B.B. (2005) Ubiquity and diversity of ammonia-oxidizing archaea in water columns and sediments of the ocean. *Proc Natl Acad Sci USA* **102:** 14683–14688.

Francis, C.A., Beman, M.J., and Kuypers, M.M.M. (2007) New processes and players in the nitrogen cycle: the microbial ecology of anaerobic and archaeal ammonia oxidation. *ISME J* **1:** 19–27.

Gruber, N. (2008) The marine nitrogen cycle: overview and challenges. In *Nitrogen in the Marine Environment*. Capone, D.G., Bronk, D.A., Mulholland, M.R., and Carpenter, E.J. (eds). San Diego, CA, USA: Elsevier, pp. 1–50.

Hallam, S.J., Mincer, T.J., Schleper, C., Preston, C.M., Roberts, K., Richardson, P.M., and DeLong, E.F. (2006) Pathways of carbon assimilation and ammonia oxidation suggested by environmental genomic analysis of marine Crenarchaeota. *PLoS Biol* **4:** e95.

Hamersley, M.R., Lavik, G., Woebken, D., Rattray, J.E., Lam, P., Hopmans, E.C., *et al.* (2007) Anaerobic ammonium oxidation contributes significantly to nitrogen loss from the Peruvian oxygen minimum zone. *Limnol Oceanogr* **52:** 923–933.

Hansman, R., Griffin, S., Watson, J.T., Druffel, E.R.M., Ingalls, A.E., Pearson, A., and Aluwihare, L.I. (2009) The radiocarbon signature of microorganisms in the mesopelagic ocean. *Proc Natl Acad Sci USA* **106:** 6513–6518.

Hollibaugh, J.T., Gifford, S., Sharma, S., Bano, N., and Moran, M.A. (2011) Metatranscriptomic analysis of ammonia-oxidizing organisms in an estuarine bacterioplankton assemblage. *ISME J* **5:** 866–878.

Ingalls, A., Shah, S.R., Hansman, R.L., Aluwihare, L.I., Santos, G.M., Druffel, E.R.M., and Pearson, A. (2006) Quantifying archaeal community autotrophy in the mesopelagic ocean using natural radiocarbon. *Proc Natl Acad Sci USA* **103:** 6442–6447.

Jayakumar, A., O'Mullan, G.D., Naqvi, S.W., and Ward, B.B. (2009) Denitrifying bacterial community composition changes associated with stages of denitrification in oxygen minimum zones. *Microb Ecol* **58:** 350–362.

Jayakumar, D.A., Francis, C.A., Naqvi, S.W.A., and Ward, B.B. (2004) Diversity of nitrite reductase genes (*nirS*) in the denitrifying water column of the coastal Arabian Sea. *Aquat Microbiol Ecol* **34:** 69–78.

Jiang, H., Dong, H., Yu, B., Lv, G., Deng, S., Berzins, N., and Dai, M. (2009) Diversity and abundance of ammonia-oxidizing archaea and bacteria in Qinghai Lake, Northwestern China. *Geomicrobiol J* **26:** 199–211.

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

Koops, H.-P., and Pommerening-Roser, A. (2001) Distribution and ecophysiology of the nitrifying bacteria emphasizing cultured species. *FEMS Microbiol Ecol* **37:** 1–9.

Lam, P., Lavik, G., Jensen, M.M., van de Vossenberg, J., Schmid, M., Woebken, D., *et al*. (2009) Revising the nitrogen cycle in the Peruvian oxygen minimum zone. *Proc Natl Acad Sci USA* **106:** 4752–4727.

Lipshultz, F., Wofsy, S.C., Ward, B.B., Codispoti, L.A., Freidrich, G., and Elkins, J.W. (1990) Bacterial transformations of inorganic nitrogen in the oxygen-deficient waters of the Eastern Tropical South Pacific ocean. *Deep Sea Res* **37:** 1513–1541.

Lozupone, C., Hamady, M., and Knight, R. (2006) Unifrac – an online tool for comparing microbial community diversity in a phylogenetic context. *BMC Bioinformatics* **7:** 371–377.

Martens-Habbena, W., Berube, P., Urakawa, H., de la Torre, J.R., and Stahl, D.A. (2009) Ammonia oxidation kinetics determine niche separation of nitrifying archaea and bacteria. *Nature* **461:** 976–979.

Mincer, T.J., Church, M.J., Taylor, L.T., Preston, C., Karl, D.M., and DeLong, E.F. (2007) Quantitative distribution of presumptive archaeal and bacterial nitrifiers in Monterey Bay and the North Pacific Subtropical Gyre. *Environ Microbiol* **9:** 1162–1175.

Molina, V., Belmar, L., and Ulloa, O. (2010) High diversity of ammonia-oxidizing archaea in permanent and seasonal oxygen-deficient waters of the eastern South Pacific. *Environ Microbiol* **12:** 2450–2465.

Mosier, A.C., and Francis, C.A. (2008) Relative abundance and diversity of ammonia-oxidizing archaea and bacteria in the San Francisco Bay estuary. *Environ Microbiol* **10:** 3002–3016.

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Simpson, G.L., Solymos, P., *et al*. (2008) VEGAN: community ecology package. R package version 1.15-0 [WWW document]. URL: http://cran.r-project.org/, http://vegan.r-forge.org/.

Park, H-D., Wells, G.F., Bae, H., Criddle, C.S., and Francis, C.A. (2006) Occurrence of ammonia-oxidizing archaea in wastewater treatment plant bioreactors. *Appl Environ Microbiol* **72:** 5643–5647.

R Development Core Team (2008) R; a language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. ISBN 3-900052-07-0 [WWW document] URL: http://www.R-project.org.

Reed, D.W., Smith, J.M., Francis, C.A., and Fujita, Y. (2010) Responses of ammonia-oxidizing bacteria and archaeal populations to organic nitrogen amendments in low-nutrient groundwater. *Appl Environ Microbiol* **76:** 2517–2523.

Revsbech, N.P., Larsen, L.H., Gundersen, J., Dalsgaard, T., Ulloa, O., and Thamdrup, B. (2009) Determination of ultra-low oxygen concentrations in oxygen minimum zones by the STOX sensor. *Limnol Oceanogr Methods* **7:** 371–381.

Santoro, A.E., Francis, C.A., de Sieyes, N.R., and Boehm, A.B. (2008) Shifts in the relative abundance of ammonia-oxidizing bacteria and archaea across physicochemical gradients in a subterranean estuary. *Environ Microbiol* **10:** 1068–1079.

Santoro, A.E., Casciotti, K.L., and Francis, C.A. (2010) Activity, abundance and diversity of nitrifying archaea and bacteria in the central California Current. *Environ Microbiol* **12:** 1989–2006.

Schloss, P.D., and Handelsman, J. (2005) Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl Environ Microbiol* **71:** 1501–1503.

Schmidt, I. (2008) Nitric oxide: interaction with the ammonia monooxygenase and regulation of metabolic activities in ammonia oxidizers. In *Nitrogen in the Marine Environment*. Capone, D.G., Bronk, D.A., Mulholland, M.R., and Carpenter, E.J. (eds). San Diego, CA, USA: Elsevier, pp. 121–135.

Schmidt, I., and Bock, E. (1997) Anaerobic ammonia oxidation with nitrogen dioxide by *Nitrosomonas eutropha*. *Antonie Van Leeuwenhoek* **167:** 106–111.

Siboni, N., Ben-Dov, E., Sivan, K., and Kushmaro, A. (2008) Global distribution and diversity of coral-associated Archaea and their possible role in the coral holobiont nitrogen cycle. *Environ Microbiol* **10:** 2979–2990.

Steger, D., Ettinger-Epstein, P., Whalan, S., Hentschel, U., de Nys, R., Wagner, M., and Taylor, M.W. (2008) Diversity and mode of transmission of ammonia-oxidizing archaea in marine sponges. *Environ Microbiol* **10:** 1087–1094.

Steinberg, D.K., Carlson, C.A., Bates, N.R., Johnson, R.J., Michaels, A.F., and Knap, A.H. (2001) Overview of the US JGOFS Bermuda Atlantic Time-series Study (BATS): a decade-scale look at ocean biology and biogeochemistry. *Deep Sea Res* **48:** 1405–1447.

Stewart, F.J., Ulloa, O., and DeLong, E.F. (2011) Microbial metatranscriptomics in a permanent marine oxygen minimum zone. *Environ Microbiol*. doi:10.1111/j.1462-2920.2010.02400.x.

Taroncher-Oldenburg, G., Griner, E.M., Francis, C.A., and Ward, B.B. (2003) Oligonucleotide microarray for the study of functional gene diversity in the nitrogen cycle in the environment. *Appl Environ Microbiol* **69:** 1159–1171.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. (1997) The CLUSTAL_X windows interface; flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25:** 4876–4882.

de la Torre, J.R., Walker, C.B., Ingalls, A.E., Konneke, M., and Stahl, D.A. (2008) Cultivation of a thermophilic ammonia oxidizing archaeon synthesizing crenarchaeol. *Environ Microbiol* **10:** 810–818.

Treusch, A.H., Leininger, S., Kletzin, A., Schuster, S.C., Klenk, H.P., and Schleper, C. (2005) Novel genes for nitrite reductase and Amo-related proteins indicate a role of uncultivated mesophilic crenarchaeota in nitrogen cycling. *Environ Microbiol* **7:** 1985–1995.

Urakawa, H., Tajima, Y., Numata, Y., and Tsuneda, S. (2008) Low temperature decreases the phylogenetic diversity of ammonia-oxidizing archaea and bacteria in aquarium biofiltration systems. *Appl Environ Microbiol* **74:** 894–900.

Urakawa, H., Martens-Habbena, W., and Stahl, D.A. (2010)

© 2011 Society for Applied Microbiology and Blackwell Publishing Ltd, *Environmental Microbiology*

High abundance of ammonia-oxidizing Archaea in coastal waters, determined using a modified DNA extraction method. *Appl Environ Microbiol* **76:** 2129–2135.
Valentine, D.L. (2007) Adaptations to energy stress dictate the ecology and evolution of archaea. *Nat Rev Microbiol* **5:** 316–323.
Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., *et al*. (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304:** 66–74.
Verhamme, D.T., Prosser, J.I., and Nicol, G.W. (2011) Ammonia concentration determines differential growth of ammonia-oxidizing archaea and bacteria in soil microcosms. *ISME J* **5:** 1067–1071.
Walker, C.B., de la Torre, J.R., Klotz, M.G., Urakawa, H., Pinel, N., Arp, D.J., *et al*. (2010) *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proc Natl Acad Sci USA* **107:** 8818–8823.
Ward, B.B. (2008) Nitrification in the marine environment. In *Nitrogen in the Marine Environment*. Capone, D.G., Bronk, D.A., Mulholland, M.R., and Carpenter, E.J. (eds). San Diego, CA, USA: Elsevier, pp. 199–262.
Ward, B.B., Kilpatrick, K.A., Renger, E.H., and Eppley, R.W. (1989) Biological nitrogen cycling in the nitracline. *Limnol Oceanogr* **34:** 493–513.
Ward, B.B., Eveillard, D., Kirshtein, J.D., Nelson, J.D., Voytek, M.A., and Jackson, G.A. (2007) Ammonia-oxidizing bacterial community composition in estuarine and oceanic environments assessed using a functional gene microarray. *Environ Microbiol* **9:** 2522–2538.
Ward, B.B., Tuit, C.B., Jayakumar, A., Rich, J.J., Moffett, J., and Naqvi, S.W.A. (2008) Organic carbon, and not copper, controls denitrification in oxygen minimum zones of the ocean. *Deep Sea Res Part I Oceanogr Res Pap* **55:** 1672–1683.
Ward, B.B., Devol, A., Rich, J.J., Chang, B., Bulow, S.E., Naik, H., *et al*. (2009) Denitrification as the dominant nitrogen loss process in the Arabian Sea. *Nature* **461:** 78–81.
Wells, G.F., Park, H.D., Yeung, C.H., Eggleston, B., Francis, C.A., and Criddle, C.S. (2009) Ammonia-oxidizing communities in a highly aerated full-scale activated sludge bioreactor: betaproteobacterial dynamics and low relative abundance of Crenarchaea. *Environ Microbiol* **11:** 2310–2328.
van der Wielen, P.W.J.J., Voost, S., and van der Kooij, D. (2009) Ammonia-oxidizing bacteria and archaea in groundwater treatment and drinking water distribution systems. *Appl Environ Microbiol* **75:** 4687–4695.
Wuchter, C., Abbas, B., Coolen, M.J., Herfort, L., van Bleijswijk, J., Timmers, P., *et al*. (2006) Archaeal nitrification in the ocean. *Proc Natl Acad Sci USA* **103:** 12317–12322.
Yool, A., Martin, A.P., Fernandez, C., and Clark, D.R. (2007) The significance of nitrification for oceanic new production. *Nature* **447:** 999–1002.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Rarefaction curves for clone libraries (CB100 – yellow; CB300 – turquoise; SS – blue; AS_PNM – green; AS_SNM – purple; ETSP – red).
Fig. S2. Unifrac analysis of site relatedness based on *amoA* DNA sequences. Sequences are uploaded to the Unifrac software (Lozupone *et al*., 2006) and manually designated to belong to a given environment. On the left is the Unifrac output detailing the phylogenetic relationship between groups of sequences grouped by environment, values are the branch lengths that gives the extent of divergence between different clusters. On the right are circular trees created using only the sequences that fall within a given cluster. Neighbour-joining trees are built using PHYLIP and bootstrapped 100 times. The trees are uploaded to iTol (Letunic and Bork, 2006) for manual annotation. Abbreviations: O, Ocean; E, Estuarine; FW, Freshwater. SMT, Southern Marianas Trench; ETSP, Eastern Tropical South Pacific; AS, Arabian Sea; HI, Hawaii; GOM, Gulf of Mexico; Sarg., Sargasso; GOC, Gulf of California; Mont., Monterey; ST, sediment; M, Marine. To reduce complexity, only the sequences from the present study are coloured in the circular tress.
Fig. S3. Correspondence analysis using individual RFR values from each site. The archetype probes appear in red with the prefix A and archetype number Sites are colour-coded and labelled. CB300_S and CB300_D are the surface and deep depths respectively. Clusters discussed in the text are colour-coded cluster 1 – black, cluster 2 – blue and cluster 3 – green.
Table S1. Phylogenetic statistics associated with the clone library study. *H* – Shannon diversity; *D* – Simpson diversity; *J* – Shannon evenness. S_{CHAO1} and S_{ACE} – CHAO and ACE statistical estimators.
Table S2. Primers used in the study.
Table S3. Probe designations and the original environment type for the cloned sequences the probe was compiled from. Environmental designation: AB, aquarium biofiltration; C, coral; E, estuary; GM, geothermal mine; HS, hydrothermal spring; M, marine; S, sediment; TS, terrestrial soil.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

© 2011 Society for Applied Microbiology and Blackwell Publishing Ltd, *Environmental Microbiology*