Nitrogen (N) is an essential nutrient in the ocean and limits biological productivity in most marine ecosystems. At the global scale, biological N fixation is the largest source of N to the ocean, whereas anaerobic microbial processes are responsible for N losses (1, 2). However, biological N fixation and anaerobic N losses are ultimately connected by nitrification, the microbial oxidation of ammonia to nitrite and nitrate, which occurs in a wide variety of environments and plays a central role in the global nitrogen cycle. Catalyzed by the enzyme ammonia monooxygenase, the ability to oxidize ammonia was previously thought to be restricted to a few groups within the β- and γ-Proteobacteria. However, recent metagenomic studies have revealed the existence of unique ammonia monooxygenase α-subunit (amoA) genes derived from uncultivated, nonextremophilic Crenarchaeota. Here, we report molecular evidence for the widespread presence of ammonia-oxidizing archaea (AOA) in marine water columns and sediments. Using PCR primers designed to specifically target archaeal amoA, we find AOA to be pervasive in areas of the ocean that are critical for the global nitrogen cycle, including the base of the euphotic zone, suboxic water columns, and estuarine and coastal sediments. Diverse and distinct AOA communities are associated with each of these habitats, with little overlap between water columns and sediments. Within marine sediments, most AOA sequences are unique to individual sampling locations, whereas a small number of sequences are evidently cosmopolitan in distribution. Considering the abundance of nonextremophilic archaea in the ocean, our results suggest that AOA may play a significant, but previously unrecognized, role in the global nitrogen cycle.

Crenarchaeota | nitrification | ammonia monooxygenase

Nitrification, the microbial oxidation of ammonia to nitrite and nitrate, occurs in a wide variety of environments and plays a central role in the global nitrogen cycle. Catalyzed by the enzyme ammonia monooxygenase, the ability to oxidize ammonia was previously thought to be restricted to a few groups within the β- and γ-Proteobacteria. However, recent metagenomic studies have revealed the existence of unique ammonia monooxygenase α-subunit (amoA) genes derived from uncultivated, nonextremophilic Crenarchaeota. Here, we report molecular evidence for the widespread presence of ammonia-oxidizing archaea (AOA) in marine water columns and sediments. Using PCR primers designed to specifically target archaeal amoA, we find AOA to be pervasive in areas of the ocean that are critical for the global nitrogen cycle, including the base of the euphotic zone, suboxic water columns, and estuarine and coastal sediments. Diverse and distinct AOA communities are associated with each of these habitats, with little overlap between water columns and sediments. Within marine sediments, most AOA sequences are unique to individual sampling locations, whereas a small number of sequences are evidently cosmopolitan in distribution. Considering the abundance of nonextremophilic archaea in the ocean, our results suggest that AOA may play a significant, but previously unrecognized, role in the global nitrogen cycle.

amoa encodes the catalytic α-subunit of ammonia monooxygenase, the enzyme responsible for catalyzing the rate-limiting step in bacterial ammonia oxidation. Because amoA is well conserved and required by all AOB, this gene has been used extensively as a molecular marker for cultivation-independent studies of AOB communities. The possibility that some archaea carry an amoA was confirmed by Schleper et al. (19), who found a Sargasso Sea-like amoA homolog on the same 43-kb metagenomic fragment as a 16S rRNA gene from soil Crenarchaeota, suggesting that archaea capable of ammonia oxidation may be present in soils. A definitive link between this novel amoA and archaeal ammonia oxidation was only recently established by cultivation of an ammonia-oxidizing member of the marine group 1 Crenarchaeota, the first cultivated representative from this dominant archaeal lineage (20). How widespread ammonia-oxidizing archaea (AOA) are in the marine environment, and their relative importance in the global N cycle, remain unknown.

In this study, we use unique PCR primers specifically targeting the archaeal amoA to document the distribution and diversity of AOA in water columns and sediments of the ocean.

Abbreviations: amoA, ammonia monooxygenase α-subunit; AOA, ammonia-oxidizing archaea; AOB, ammonia-oxidizing bacteria; ETNP, Eastern Tropical North Pacific; OMZ, oxygen minimum zone; OTU, operational taxonomic unit.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. DQ14825–DQ14848 and DQ148573–DQ148905).

© 2005 by The National Academy of Sciences of the USA.
Materials and Methods

Site Descriptions, Sample Collection, and DNA Extraction. Water samples were collected from three depths ($r_{p} = 15.7, 15.8, and 15.9$) bracketing the secondary nitrite maximum in the suboxic zone of the Black Sea (April 2003) and two stations (designated C1 and M2) in Monterey Bay (April 2005) by using Niskin bottles mounted on a hydrography wire. Water samples were either filtered directly onto Sterivex capsules (0.2 μm filter, Millipore) with positive pressure (Black Sea) or a peristaltic pump was used (Monterey Bay). After filtration, capsules were capped, packed in polyvinyl bags, frozen in liquid nitrogen, and stored at −80°C until DNA extraction. Briefly, DNA was extracted from Sterivex capsules by using 1.6 ml of lysis buffer (20 mM EDTA/400 mM NaCl/750 mM sucrose/50 mM Tris), 200 μl of 10% SDS, and 100 μl of proteinase K (10 mg/ml), incubated at 55°C for 2 h. After the addition of 1 ml of 100% ethanol, lysates were purified by using DNeasy columns (Qiagen, Valencia, CA).

Ten liters of water was collected in November 2003 from the secondary nitrite maximum (200-m depth) within the OMZ off the Pacific coast of Mexico, filtered onto a SpiralCap capsule, frozen, and stored below −70°C. For lysis, 12 ml of Tris-EDTA/1% SDS and 750 μl of proteinase K (10 mg/ml) was added to the capsules, followed by incubation at 55°C for 2 h. The lysate was removed, placed on ice for 90 min, and centrifuged for 40 min at 17,000 × g to remove cell debris. DNA was ethanol-precipitated overnight, pelleted at 32,000 × g for 30 min, resuspended in 300 μl of sterile water, and further purified by using Qiagen DNeasy columns.

Surface sediments were collected from San Francisco Bay and Bahía del Tóbari, Mexico (Gulf of California), using box cores deployed off small vessels in July 2004 and October 2004, respectively. Sediments were collected by hand from three sites in the Elkhorn Slough estuary (from head to mouth: Hudson’s Landing, Hummingbird Island, and Vierra Marsh) in October 2004 and from the surf zone at the mouth of the unconfined Huntington Beach aquifer in July 2004. Soil aggregates were collected in June 2004 from the saturated C horizon at a noncontaminated “background” site near Oak Ridge, TN. Cores were collected from each environment by using cut-off 5–cc syringes, frozen on dry ice, and stored at −80°C. DNA was extracted from ~0.25 g of sediment (0–0.5 cm depth interval) or soil by using the FastDNA SPIN kit for soil (Obiogene, Carlsbad, CA).

PCR Primer Design, Amplification, and Cloning of Archaeal amoA Fragments. PCR primers were designed based on alignments of archaeal amoA genes and deduced amino acid sequences from the Sargasso Sea (GenBank accession no. AACY01435967) and German soil (GenBank accession no. AJ627422). Gene fragments (635 bp) were amplified by using the PCR primers Arch-amoAF (5’-STAAATGGTCGTGCGAGTACACAGC-3’) and Arch-amoAR (5’-CGGGCCATCCATCTCTAATG-3’) with the following protocol: 95°C for 5 min; 30 cycles consisting of 94°C for 45 s, 53°C for 60 s, and 72°C for 60 s; and 72°C for 15 min. Triplicate PCRs were pooled (to minimize PCR bias), gel-purified, and cloned with the TOPO-TA cloning kit (Invitrogen). White colonies were transferred to 96-well plates containing LB broth (with 50 μg/ml kanamycin), grown overnight at 37°C, and PCR-screened directly for the presence of inserts by using T7 and M13R vector primers.

Sequencing and Phylogenetic, Rarefaction, and Statistical Analyses. Sequencing of T7/M13 PCR products was performed by using vector primers on Applied Biosystems 3730xl capillary sequencers. Nucleotide sequences were assembled and edited by using SEQUENCER, version 4.1 (Gene Codes, Ann Arbor, MI). A 606-bp region of 377 archaeal amoA sequences was chosen for phylogenetic analysis. Nucleotide and amino acid alignments (based on 202 residues) were generated by using MACCLADE (http://macclade.org). Neighbor-joining phylogenetic trees were constructed based on alignments of DNA sequences (using Jukes–Cantor corrected distances) and amino acid sequences within PAUP*4.0b10. Bootstrap analysis was used to estimate the reliability of phylogenetic reconstructions (1,000 replicates). To compare the archaeal amoA-based richness within each clone library, rarefaction analysis and Chao1 nonparametric richness estimations were performed by using DOTUR (21). Operational taxonomic units (OTUs) were defined as sequence groups in which sequences differed by ≤2% or 5%.

Results and Discussion

Ubiquity of AOA. Nonextremophilic Crenarchaeota have been previously detected in a wide variety of environments and are widespread in the ocean (12, 13); however, their role in the marine C or N cycle has remained a mystery. Our results suggest that archaea capable of ammonia oxidation are ubiquitous in marine water columns and sediments. An extensive PCR survey of archaeal amoA genes using our primers resulted in amplification of 635-bp fragments (corresponding to essentially the entire gene) from multiple sites or depths within eight different geographic locations: the water columns of Monterey Bay (base of the euphotic zone), the Eastern Tropical North Pacific (ETNP; OMZ), and the Black Sea (suboxic zone); coastal and estuarine sediments from San Francisco Bay, Elkhorn Slough, and Huntington Beach, CA, and Bahía del Tóbari, Mexico, and for purposes of comparison, a saturated (suboxic) soil sample from Oak Ridge, TN. In total, we produced 15 clone libraries and a database of 377 archaeal amoA sequences from water columns, sediments, and soil (Fig. 1).

Primers were designed based on the novel archaeal amoA sequences from German soil and the Sargasso Sea water column (18, 19). These two genes shared 76% DNA identity, but 85% identity and 91% similarity at the amino acid level. Because these archaeal amoA genes are distant homologs of known bacterial amoA genes from β- and γ-Proteobacteria (19), it is not possible to recover them from environmental DNA samples with bacterial amoA (or pmoA) primers. Although designed based on a limited number of available sequences, our primers recovered 377 archaeal amoA sequences ranging from 69% to 96% identity to the German soil sequence and from 69% to 99% identity to the Sargasso Sea sequence. Based on this extensive database, it may be possible to design new, more degenerate primers that capture even greater archaeal amoA diversity.

AOA Diversity. Overall, we recovered a total of 138 unique OTUs (based on a 2% cutoff) from 377 individual archaeal amoA sequences. Both observed and extrapolated AOA richness varied widely within water column and sediment libraries (Fig. 2), with Chao1 richness estimates ranging from 5 to 37 OTUs for water column libraries and 11 to 48 OTUs for sediment libraries (see Table 1, which is published as supporting information on the PNAS web site). These richness levels are comparable to, or higher than, those observed for bacterial amoA clones obtained from water columns and sediments (22, 23). Based on phylogenetic analysis of PCR clone libraries, individual sample locations often had unique archaeal amoA sequences associated with them (Fig. 1); however, clades of these unique AOA sequences clustered according to habitat. For example, all sequences from North San Francisco Bay (NB-1), our only freshwater site, fell exclusively into one coherent cluster. Similarly, all amoA sequences recovered from Tennessee soil fell into one large, phylogenetically distinct cluster (see soil/sediment cluster in Fig. 1). The cluster also included the amoA sequence from soil metagenomic clone 549, which is derived from a member of the most widely distributed Crenarchaeota lineage (group 1b).
Fig. 1. Phylogenetic relationships among archaeal amoA sequences from water columns, sediments, and soil. Sequences are color-coded according to location; symbols are used to distinguish different sites or depths within a given location. One representative of sequence groups ≥ 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 points. Brackets highlight phylogenetic clusters referred to in the text. HI, Hummingbird Island; HL, Hudson’s Landing; VM, Vierra Marsh.

Francis et al. PNAS October 11, 2005 vol. 102 no. 41 14685

ENVIRONMENTAL SCIENCES
known in soils (19, 24). Interestingly, 141 of 142 water column sequences from the Black Sea, Monterey Bay, and the ETNP, despite widely different geographical locations and biogeochemical characteristics, fell into three distinct clusters (A, B, and C; Fig. 1). The largest of these clusters (cluster A) contains >70% of the water column sequences, including the Sargasso Sea sequence. It is tempting to speculate that these water column sequences correspond to strictly “planktonic” AOA; however, several sequences from Elkhorn Slough sediments also fell into these clades and were >99% identical to Black Sea and Monterey Bay water column sequences (Fig. 1). A number of sedimentary sequences also fell into the soil/sediment cluster. Based on Fig. 1, it is evident that while all but one of the water column sequences fell into the water column clusters, and all soil sequences fell into the large soil/sediment cluster, sequences within these clusters are not found exclusively in those environments. This finding in part reflects the wide range of archael amoA sequences recovered from sediments, with sequences falling in numerous clusters throughout the phylogenetic tree (Fig. 1).

AOA Diversity in the Water Column: The Black Sea, Monterey Bay, and ETNP. Based on the identification of an archael amoA gene in a Sargasso Sea surface water genomic library, we anticipated that similar genes would be present in other marine water column environments where nitrification plays a significant role. The Black Sea is the world’s largest chemically stratified anoxic marine basin and encompasses a pronounced suboxic (O2 <5 μM) zone in which nitrogen cycling (e.g., denitrification, anammox) is extremely important (25, 26). Archael amoA was amplified from multiple depths within the suboxic zone of the Black Sea. Interestingly, this gene was present at densities (σ2) similar to those where anaerobic ammonia oxidation has been reported (26). The majority (>88%) of the archael amoA sequences obtained from the Black Sea suboxic zone fell into cluster A and showed 90–92% nucleotide identity to the Sargasso Sea gene sequence. Within this cluster, there were four predominant sequence types differing by ~1–2% at the DNA level (visible in an expanded tree, Fig. 3, which is published as supporting information on the PNAS web site), primarily at the third position. The remaining Black Sea sequences fell into clusters B and C. Overall, the Black Sea clone libraries were among the least diverse in this study (Fig. 2), and we appear to have recovered most or all of the OTUs predicted by the Chao1 richness estimator (Table 1).

The majority of sequences from the upper water column of Monterey Bay fell into cluster A along with the closely related Black Sea sequences (Fig. 1). Although archael amoA was successfully amplified from all depths analyzed between 30 and 80 m, for brevity we report here only sequence data derived from 30 m at a coastal site (C1) and 40 m at an offshore site (M2). The coastal station (C1) was used in part to determine the extent of overlap between planktonic AOA communities in Monterey Bay and benthic communities in Elkhorn Slough. (Because Elkhorn Slough is a small estuary that is tidally flushed by Monterey Bay seawater, Monterey Bay represents a potential source and sink for nutrients and microorganisms to the slough.) Near the offshore site (M2), nitrification rates (27) and AOB diversity (22) have recently been examined. Although there was some compositional overlap between the coastal (C1) and offshore (M2) archael amoA clone libraries, the coastal library was more diverse (Fig. 2). One subcluster within cluster A is comprised of seven C1 sequences (and one ETNP sequence) that share 93–99% identity with the Sargasso Sea sequence at the DNA level. Translation of these sequences revealed that all are identical to the Sargasso sequence at the amino acid level. These results indicate that the “microdiversity” within these coastal Monterey Bay sequences corresponds entirely to synonymous changes, with no effect on archael ammonia monooxygenase function. Additionally, most of the microdiverse (1–2% divergent) clusters of Black Sea and Monterey Bay sequences in cluster A are identical at the amino acid level. Given the differences between the suboxic zone of the Black Sea and the oxygenated euphotic zone of Monterey Bay, the compositional overlap between the archael amoA sequences from these two environments is surprising and suggests that planktonic AOA can tolerate a wide range of O2 levels.

This conclusion is consistent with our results from the ETNP, one of the largest pelagic OMZs in the ocean, where archael amoA was amplified successfully from a 200-m depth under low oxygen conditions (O2 = 3.1 μM) at the secondary nitrite maximum (NO2 = 2.5 μM). The ETNP and coastal Monterey Bay sequences were the most diverse of the six water column samples in the study (Fig. 2), with sequences from the ETNP falling into water column clusters A and B and one sediment cluster (see below). In contrast to the Black Sea and Monterey Bay libraries, the majority (21 of 26) of the ETNP sequences fell into cluster B, which included only nine sequences from the other water column libraries. Overall, we appear to have identified several key phylogenetic groups/clusters of planktonic AOA,
based on the six archaeal amoA clone libraries generated from these three geographically and biogeochemically distinct water column environments.

**Significance of AOA in the Water Column.** Previous studies have shown marine crenarchaeal abundance in the water column to increase after blooms in surface phytoplankton and to be negatively correlated with chlorophyll concentrations (28). Archaea also dominate the microbial community in regions of the ocean where nitrification is known to be important (9) and, in the Santa Barbara Channel, crenarchaeal abundance and NO2 - concentrations were found to be significantly correlated over time (29). Many of the highest spatial and temporal patterns observed for marine Crenarchaeota as a group are therefore similar to patterns expected for nitrifying microorganisms. We suggest here that a significant portion of the archaeal community may be oxidizing remineralized NH4+ (i.e., performing the first step of chemooautotrophic nitrification), and that some of these broader archaeal patterns in the water column may be driven in part by AOA.

In the Black Sea, for example, our AOA data are consistent with earlier work focused on marine Archaea as a group. Roughly 10% of the archaeal 16S rRNA clones previously reported from the chemocline (i.e., suboxic zone) of the Black Sea were most closely related to marine group 1 Crenarchaeota (30). Diversity was extremely low and only two 16S rRNA sequences (sharing 97% identity) were identified. In the present study, we recovered distinct phylogenetic clusters from the suboxic zone, but found archaeal amoA diversity to be quite low compared with other water column environments.

In addition, the presence of AOA in the suboxic water columns of the Black Sea and the ETNP is consistent with a peak in the concentration of crenarchaeol (the distinct membrane lipid biomarker for the Crenarchaeota) previously reported in the OMZ of the northwestern Arabian Sea (33). Lipid biomarker for the Crenarchaeota (31, 32) previously compared with other water column environments.

involved in this N2O production pathway (35), and some AOA and denitrification represent major sources of N2O in the ocean where nitrification is known to be important (9) and, in the suboxic zone, but found archaeal amoA diversity to be quite low compared with other water column environments.

AOA Diversity in Sediments: Elkhorn Slough, Bahía del Tóbari, San Francisco Bay, and Huntington Beach. Sediment DNA extracts from four geographically distinct locations were PCR-screened to determine the distribution and diversity of archaeal amoA sequences in coastal and estuarine sediments. Unique AOA communities were associated with each geographic location, and within these locations, different clusters were often specifically associated with individual sample sites. In Elkhorn Slough, different levels of richness (Fig. 2) and surprisingly little compositional overlap were found among clone libraries generated from three sites that span the estuarine gradient from head to mouth of the slough. For each slough site, at least one distinct clade of sequences was site-specific, with no overlap with sequences from other sites (Fig. 1). Similar shifts in AOA community composition and richness were observed in Bahía del Tóbari, a small estuary located along the west coast of Mexico in the Gulf of California. Multiple sequences from the interior of Tóbari (site 4) fell into the soil/sediment cluster, yet no sequences from the mouth of the bay (site 6) fell into this cluster (Fig. 1); in addition, the archaeal amoA library from the interior site was among the least diverse of the sediment libraries analyzed in this study, whereas the mouth site was among the most diverse (Fig. 2 and Table 1).

The factors that influence AOA diversity and community structure in Elkhorn Slough and Bahía del Tóbari are currently unknown. Although salinity has shown to be a key factor influencing the distribution and diversity of AOB in estuarine systems, based on bacterial 16S rRNA (37) and amoA genes (33), differences between the sites within Elkhorn Slough and Bahía del Tóbari are probably not salinity-driven: the slough experiences daily tidal flushing by Monterey Bay seawater, and Tóbari receives only episodic inputs of fresh water via agricultural runoff (38). However, differences in AOA community composition in San Francisco Bay, the largest estuary on the west coast of the United States, are likely associated with salinity. As alluded to earlier, sequences from the 30.5-psu salinity units (psu) Central San Francisco Bay (CB-20) site were distributed throughout a number of different regions of the tree, whereas all of the low-salinity (0.5 psu) North San Francisco Bay (NB-1) sequences fell exclusively into one distinct phylogenetic cluster (Fig. 1). This clustering suggests that the North San Francisco Bay archaeal amoA sequences may represent a unique, low-salinity AOA type. Many of the Central San Francisco Bay (CB-20) sequences fell into a well supported cluster with clones from two well flushed Elkhorn Slough sites (Hummingbird Island and Vierra Marsh). These sequences share 93–100% identity at the DNA level and may represent a ubiquitous AOA group in coastal estuarine sediments of this geographic region.

As the primary substrate required for nitrification, NH4+ might also be expected to influence AOA community structure. Coastal permeable sediments from Huntington Beach are characterized by consistently elevated NH4+ concentrations caused by hydrologic connection with groundwater (39), which is highly enriched in NH4+ (groundwater [NH4+] > 150 μM). The archaeal amoA library from this site was the most diverse in this study (Fig. 2), yet many of the Huntington Beach sequences fell into a distinct cluster with a single sequence from the most NH4+-rich site sampled in Elkhorn Slough (Hudson’s Landing).

The four sedimentary environments included in this study were selected in part because of highly variable physical and biogeochemical conditions found within and between them. Despite these pronounced differences, one broad phylogenetic cluster (located directly above cluster A) included sequences from all four of the marine sedimentary environments we sampled (Fig. 1). This cluster also included the only sequence from the ETNP that did not fall into any of the water column clusters. Given the differences between these five environments, this cluster may represent a cosmopolitan group of marine AOA. Preliminary characterization of ammonia-oxidizing enrichment cultures further supports this conclusion, as several archaeal amoA sequences recovered from both Elkhorn Slough sediment enrichments and Monterey Bay seawater enrichments fall into this cluster (K.J.R. and C.A.F., unpublished data).

**Significance of AOA in Marine Sediments.** Although phylogenetically diverse crenarchaeal 16S rRNA phylotypes have been previously identified in sediments of freshwater lakes (40, 41), estuaries (42), continental margins (43), and the deep sea (44), the biogeochemical role of Crenarchaeota in these sedimentary environments is unknown. Clearly, the ubiquity and diversity of AOA in coastal and estuarine sediments suggests that they may be pivotal in benthic N cycling. In coastal and estuarine ecosystems, nitrification is known to link the mineralization of organic nitrogen to the loss of fixed
nitrogen by sedimentary denitrification. A common characteristic of all of the sediment samples analyzed in this study, as well as the suboxic water column and soil samples, is that denitrification functional genes (i.e., nitrite reductase) have also been successfully amplified from each of them. These results suggest that AOA may be intimately associated with denitrifiers and, like AOB, play an important role in coupled nitrification–denitrification. Interestingly, bacterial amoA could not be amplified from some of the samples analyzed in this study (e.g., Tòbari samples), which may indicate that archaeal amoA is more widespread than bacterial amoA, and that AOA are actually more widespread than AOB in sedimentary environments. Future studies focused on quantifying the relative abundance and expression of bacterial versus archaeal amoA genes under nitrifying conditions will be imperative for understanding the relative contributions of what are apparently two functionally equivalent, but evolutionarily distinct, groups of microorganisms.

We thank Amal Jayakumar for collecting the ETNP water sample and geochemical data; Scott Wankel for collecting sediment samples during the San Francisco Bay RMP sediment cruise; and Brendan Bohannan, John Spear, and Gregory O’Mullan for helpful comments on the manuscript. This work was supported in part by National Science Foundation Grant MCB-0433804 (to C.A.F.), a National Science Foundation Graduate Research Fellowship (to A.E.S.), and a National Science Foundation Dissertation Improvement Grant (to J.M.B.). B.B.O. was supported by National Science Foundation Microbial Observatories Grant MCB-0132101 and the National Aeronautics and Space Administration Astrobiology Institute.