2012

Diversity, Distribution, and Expression of Diazotroph nifH Genes in Oxygen-Deficient Waters of the Arabian Sea

Amal Jayakumar

Mamoon M. D. Al-Rshadat
Old Dominion University

Bess B. Ward
Old Dominion University

Margaret R. Mulholland
Old Dominion University, mmulholl@odu.edu

Original Publication Citation

Jayakumar, A., Al-Rshaidat, M. M. D., Ward, B. B., & Mulholland, M. R. (2012). Diversity, distribution, and expression of diazotroph *nifH* genes in oxygen-deficient waters of the Arabian Sea. *FEMS Microbiology Ecology, 82*(3), 597-606. doi:10.1111/j.1574-6941.2012.01430.x

RESEARCH ARTICLE

Diversity, distribution, and expression of diazotroph *nifH* genes in oxygen-deficient waters of the Arabian Sea

Amal Jayakumar[1], Mamoon M.D. Al-Rshaidat[2], Bess B. Ward[1] & Margaret R. Mulholland[2]

[1]Department of Geosciences, Princeton University, Princeton, NJ, USA and [2]Department of Ocean, Earth & Atmospheric Sciences, Old Dominion University, Norfolk, VA, USA

Correspondence: Amal Jayakumar, Department of Geosciences, Princeton University, Princeton, M45 Guyot Hall, Washington Road, NJ 08544, USA.
Tel.: +1 609 258 6294;
fax: +1 609 258 0796;
e-mail: ajayakum@princeton.edu

Present address: Mamoon M.D. Al-Rshaidat, Coastal Environment, College of Marine Sciences, The University of Jordan-Aqaba, Aqaba, Jordan.

Received 13 January 2012; revised 29 May 2012; accepted 7 June 2012.
Final version published online 9 July 2012.

DOI: 10.1111/j.1574-6941.2012.01430.x

Editor: Riks Laanbroek

Keywords
nitrogen fixation; *nifH* gene diversity and expression; oxygen deficient zone; Arabian Sea.

Abstract

The Arabian Sea oxygen minimum zone (OMZ), the largest suboxic region in the world's oceans, is responsible for up to half of the global mesopelagic fixed nitrogen (N) loss from the ocean via denitrification and anammox. Dinitrogen (N_2) fixation is usually attributed to cyanobacteria in the surface ocean. Model prediction and physiological inhibition of N_2 fixation by oxygen, however, suggest that N_2 fixation should be enhanced near the oxygen-deficient zone (ODZ) of the Arabian Sea. N_2 fixation and cyanobacterial *nifH* genes (the gene encoding dinitrogenase reductase) have been reported in surface waters overlying the Arabian Sea ODZ. Here, water samples from depths above and within the Arabian Sea ODZ were examined to explore the distribution, diversity, and expression of *nifH* genes. In surface waters, *nifH* DNA and cDNA sequences related to *Trichodesmium*, a diazotroph known to occur and fix N_2 in the Arabian Sea, were detected. Proteobacterial *nifH* phylotypes (DNA but not cDNA) were also detected in surface waters. Proteobacterial *nifH* DNA and cDNA sequences, as well as *nifH* DNA and cDNA sequences related to strictly anaerobic N-fixers, were obtained from oxygen-deficient depths. This first report of *nifH* gene expression in subsurface low-oxygen waters suggests that there is potential for active N_2 fixation by several phylogenetically and potentially metabolically diverse microorganisms in pelagic OMZs.

Introduction

Productivity in much of the world's ocean is limited by the availability of fixed nitrogen (N). Observations and extrapolations from models suggest either that fixed N losses via denitrification and anammox greatly exceed N inputs to the ocean via dinitrogen (N_2) fixation (Codispoti *et al.*, 2001; Codispoti, 2007), or that these opposing processes are balanced, at least on geological timescales (Gruber & Sarmiento, 1997; Altabet *et al.*, 2002; Deutsch *et al.*, 2007; Altabet, 2007). Although both processes can be variable in time and space, it is hypothesized that the apparent imbalance is owing to an underestimation of N_2 fixation (Codispoti, 2007). To test this hypothesis, the relative distributions of microorganisms that fix N_2 and their activity are being investigated using observations and models.

Denitrification occurs in anoxic and suboxic areas of the world's oceans including marine sediments, anoxic basins and fjords, and pelagic oceanic oxygen minimum zones (OMZs) (Devol, 2008). Here, the term OMZ denotes the oceanographic region, for example the Arabian Sea, where low-oxygen waters are found, while oxygen-deficient zone (ODZ) refers to the depths where oxygen concentrations are low enough to induce anaerobic metabolisms, estimated at 1–2 μM. The Arabian Sea OMZ contains the largest suboxic region in the world's oceans and is responsible for up to half of the global mesopelagic fixed N loss (Naqvi, 2008). While low oxygen is thought to be beneficial for N_2 fixation because oxygen inhibits nitrogenase (the enzyme complex that mediates N_2 fixation), high nitrate concentrations, such as those present in ODZs, are thought to be inhibitory to N_2 fixation in these systems (Postgate, 1998). Nonetheless, even *Trichodesmium*, the best-known marine N-fixing cyanobacterium, fixes N_2 in the presence of dissolved inorganic nitrogen (DIN) (Mulholland *et al.*, 1999, 2001; Holl &

Montoya, 2005). Further, N_2 fixation has recently been documented in areas where substantial DIN concentrations were measured, including coastal waters (Grosse *et al.*, 2010; Mulholland *et al.*, 2012), the Pacific Ocean (Moisander *et al.*, 2010), and the eastern tropical Atlantic Ocean (Voss *et al.*, 2004). In addition, models suggest that N_2 fixation may occur in closer proximity to zones of denitrification than previously thought (Deutsch *et al.*, 2007; Mulholland & Capone, 2009). Denitrification results in the depletion of fixed N relative to phosphorus (P), thereby lowering N : P ratios of dissolved nutrient pools, which may create conditions conducive to N_2 fixation (Deutsch *et al.*, 2007; Capone & Knapp, 2007; Naqvi, 2008). Consistent with the hypothesis that N_2 fixation should occur in proximity to denitrification, low rates of N_2 fixation, as well as gene sequences of *nifH* (the gene encoding a critical component of the nitrogenase enzyme complex) representing several proteobacterial *nifH* phylotypes, were detected in the low-oxygen, nitrate-rich waters of coastal California (Hamersley *et al.*, 2011) and in the OMZ region of the Eastern Tropical South Pacific (Fernandez *et al.*, 2011).

N_2 fixation is likely an important process in surface waters of the Arabian Sea. Microscopic observations and measurements of *nifH* abundance and diversity suggest that there is a diverse assemblage of diazotrophs in surface waters of the Arabian Sea that includes *Trichodesmium*, unicellular cyanobacteria, *Gammaproteobacteria*, and diatom–diazotroph assemblages (Capone *et al.*, 1998; Mazard *et al.*, 2004; Bird *et al.*, 2005; Parab *et al.*, 2006; Mulholland & Capone, 2009; Padmakumar *et al.*, 2010). Direct measurements of N_2 fixation rates, however, have been reported only once for the Arabian Sea (Capone *et al.*, 1998). To investigate the capacity for N_2 fixation within the Arabian Sea, we examined *nifH* diversity and gene expression within the ODZ as well as in the overlying surface waters.

Materials and methods

Water samples were collected from three stations in the Arabian Sea (Fig. 1) during September and October of 2007 aboard the *R/V Roger Revelle*. Water column nitrate, nitrite, and oxygen concentrations at these stations were reported previously (Ward *et al.*, 2009) and are shown here for one station in Fig. 2. Oxygen concentrations were < 1 μM between 100 and 750 m depth at Station 1 and between 140 and 750 m depth at Station 2. The depleted oxygen layer at Station 3 was much thinner, spanning only 120–260 m in depth. Water samples (up to 12 L) were collected from 2 to 4 depths per station (Table 1): near the surface or in the chlorophyll maximum layer (upper mixed layer), within the oxycline

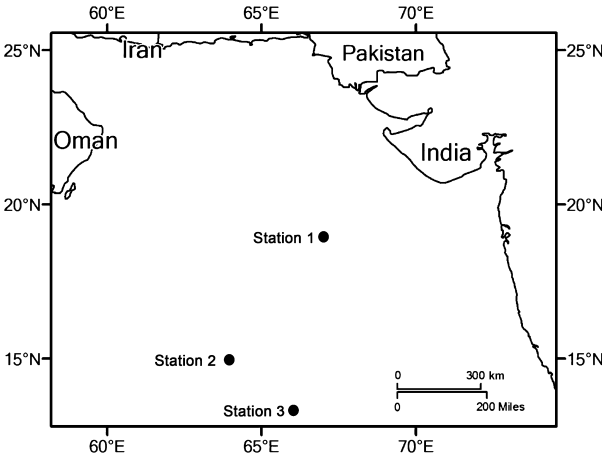

Fig. 1. Map of the Arabian Sea showing the three stations sampled for this study.

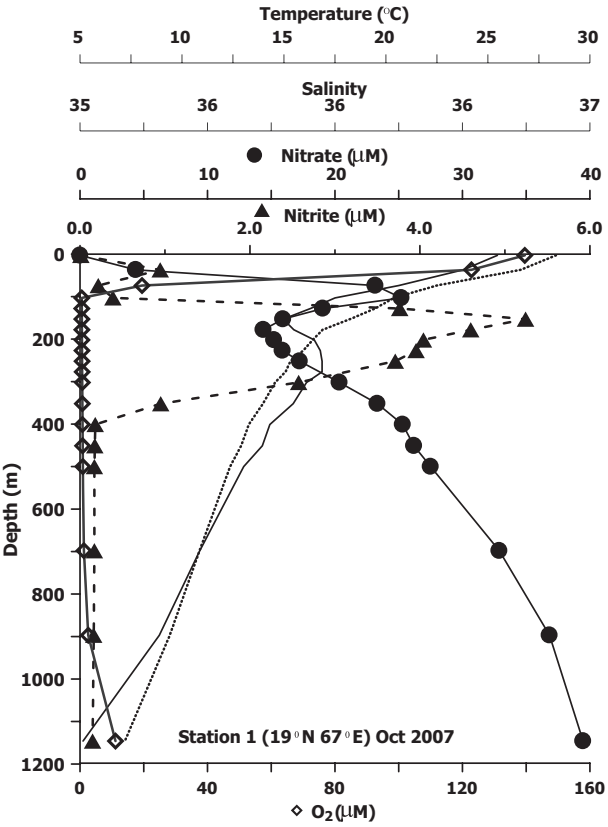

Fig. 2. Vertical distribution of chemical and physical variables at Station 1, showing representative features that characterize the surface, oxycline, and ODZ regions of the water column.

where the oxygen gradient was steepest, and within the ODZ at the depth of the secondary nitrite maximum. Whole water from CTD-mounted Niskin bottles was filtered through Sterivex-GP capsule filters (Millipore), using a peristaltic pump, quick frozen in liquid nitrogen,

Table 1. Stations, depth, and oceanographic features from which DNA and cDNA sequences were obtained

Station	Depth (m)	Feature	DNA clones analyzed	DNA clone ID (number of clones)	cDNA clones analyzed	cDNA clone ID (number of clones)
1 (19°N 67°E)	10	Euphotic zone	17	S1D010M (3)	NA	NA
	60	Oxycline	24	S1D060M (20)	NA	NA
	150	Secondary nitrite maximum (ODZ)	24	S1D150M (23)	47	mRNA S1D150M (25)
	175	Secondary nitrite maximum (ODZ)	23	S1D175M (11)	24	mRNA S1D175M (22)
2 (15°N 64°E)	2.5	Euphotic zone	32	S2D2.5M (0)	NA	NA
	150	Secondary nitrite maximum (ODZ)	32	S2D150M (5)	28	mRNA S2D150M (25)
3 (12°N 64°E)	10	Euphotic zone	32	S3D010M (25)	20	mRNA S3D10M (4)
	65	Chlorophyll Maximum	32	S3D065M (0)	NA	NA
	110	Oxycline	32	S3D110M (5)	28	mRNA S3D110M (23)

Clones analyzed = total number of clones sequenced. (number of clones) = number of clones that were determined by phylogenetic analysis to represent *nifH* genes. NA = cDNA clone libraries were not made from these samples.

and shipped in liquid nitrogen to Princeton University (PU). The samples were stored at −80 °C before being transported to Old Dominion University (ODU) in liquid nitrogen for DNA extraction.

The DNA was extracted at ODU from Sterivex filters using the PUREGENE™ Genomic DNA Isolation Kit (Qiagen, Germantown, MD). The RNA was extracted at PU from replicate Sterivex filters using the ALLPrep DNA/RNA Mini Kit (Qiagen). The cDNA was synthesized immediately following purification of RNA using a SuperScript III First Strand Synthesis System (Invitrogen, Carlsbad, CA) following the procedure described by the manufacturer. DNA and cDNA were quantified using PicoGreen dsDNA quantification kit (Invitrogen) according to the manufacturer's specifications. The *nifH* sequences were amplified from environmental DNA and cDNA using Phusion™ High-Fidelity DNA Polymerase (Finnzymes), on an MJ100 Thermal Cycler (MJ Research). A nested reaction was used, as previously described (Zehr *et al.*, 1998), with slight modification: 25 μL PCRs were amplified for 30 cycles (1 min at 98 °C, 1 min at 57 °C, 1 min at 72 °C), first with the outer PCR primers (Zani *et al.*, 2000), followed by amplification with the inner PCR primers (Zehr and McReynolds, 1989). To minimize the possibility of amplifying contaminants (Zehr *et al.*, 2003), negative controls (autoclaved and UV-irradiated water) were run with every PCR experiment, reagents were diluted in freshly autoclaved water, the PCR preparation station was UV irradiated for 1 h before each daily use, and the number of cycles was limited to 30 for each reaction. The PCR reagents were also irradiated where possible. Each reagent was then tested separately for amplification in negative controls.

Amplified fragments were electrophoresed on 1.2% agarose gels, and *nifH* bands were excised and then cleaned using a QIAquick Nucleotide Removal Kit (Qiagen). Cleaned fragments were inserted into a pCR®2.1-TOPO® vector using One Shot® TOP10 Chemically Competent *E. coli*, TOPO TA Cloning® Kit (Invitrogen) according to manufacturer's specification. Clones were picked randomly and amplified using M13 forward (−20) and M13 reverse primers. The products were sequenced at the DNA Analysis Facility at Yale University using Big DyeTM terminator chemistry (Applied Biosystems, Carlsbad, CA). Sequences were edited using FINCHTV ver. 1.4.0 (Geospiza Inc.) and checked for identity using BLAST. Consensus *nifH* sequences (359 bp) were translated to amino acid (aa) sequences (108 aa after trimming the primer region) and aligned using CLUSTALX (Thompson *et al.* 1997) along with published *nifH* sequences from the NCBI database. Neighbor-joining trees were produced from the alignment using distance matrix methods (PAUP 4.0; Sinauer Associates). Bootstrap analysis was used to estimate the reliability of phylogenetic reconstruction (1000 iterations). The *nifH* sequence from *Methanosarcina lacustris* (AAL02156) was used as an outgroup. The new Arabian Sea *nifH* sequences have been deposited in GenBank, DNA sequences accession numbers JF429940–JF429973 and cDNA sequences accession numbers JQ358610–JQ358707.

Results and discussion

Nitrate, nitrite, and oxygen profiles were characteristic of an open ocean OMZ, with a strong nitrate deficit and a clear secondary nitrite maximum within the depth interval where oxygen concentrations were < 1 μM (Fig. 2; Ward *et al.*, 2009). All major nutrients were depleted in surface waters. The stable stratification of warm surface waters likely resulted in little if any communication between the euphotic zone and the depth of the secondary nitrite maximum, with the exception of the downward flux of sinking particulate material. Thus, the dramatically different chemical environments at the

sample depths imply quite different conditions for microbial growth and suggest that the diazotrophic assemblages would likely differ among depths.

A total of 92 DNA clones containing sequences identified as homologous with known *nifH* sequences were retrieved from seven of the nine depths from which clone libraries were made (Table 1). The cDNA clone libraries were made from five depths, and a total of 99 cDNA sequences were obtained (Table 1). Phylogenetic analysis of translated *nifH* DNA and cDNA fragments grouped the sequences into four of the major clusters previously used to characterize *nifH* phylogeny (Chien & Zinder, 1996). Cluster I, including proteobacteria and cyanobacteria, is shown in Fig. 3, and sequences from Clusters II, III, and IV are shown in Fig. 4.

Three sequences from surface waters (10 m) at Station 1 had high identity with *nifH* sequences from *Trichodesmium* spp. (Fig. 3). These sequences clustered with uncultured marine cyanobacterial sequences reported by Mazard *et al.* (2004) from surface waters of the Arabian Sea, as well as with sequences from surface waters in many other locations, including the North Pacific, North Atlantic, and South China Sea.

Although several clones (32 from each depth) of the PCR products obtained with the nested *nifH* primers from the euphotic zone at Station 2 and the chlorophyll maximum at Station 3 (Table 1) were sequenced, no *nifH* DNA sequences were detected among those clones. This suggests that N_2-fixing cyanobacteria were not abundant in these waters. Nevertheless, we did detect *nifH* gene expression: four cDNA sequences from 10 m at Station 3 were almost identical to the *Trichodesmium*-like DNA sequence from 10 m at Station 1.

No sequences related to the recently identified ubiquitous, single-celled N_2-fixing 'Group A' and 'Group B' cyanobacteria (Zehr *et al.*, 2001, 2008) were recovered from any of the DNA or cDNA clone libraries (Fig. 3). Using PCR primers for the 16S rRNA gene designed to be specific for members of the Groups A and B N_2-fixing uncultured cyanobacteria ($UCYN_2$-fix lineage), Mazard *et al.* (2004) detected these relatives of *Crocosphaera* and *Cyanothece* only in the most southern and most northern regions of their N-S transect across the Arabian Sea. However, no sequences related to these groups were detected at their transect stations, which essentially coincide with Stations 1 and 2 in the current study. Mazard *et al.* (2004) suggested that the $UCYN_2$-fix lineage was restricted to very warm (~ 30 °C) oligotrophic waters and that temperature was the factor that limited their distribution. *Trichodesmium*-like DNA sequences, for both *nifH* and 16S rRNA genes, were detected by Mazard *et al.* (2004) at the central Arabian Sea stations on their transect. Thus, although based on a small number of clones, the two studies are consistent in suggesting that *Trichodesmium*-like cyanobacteria are the most important N_2 fixers in the open ocean surface waters of the Arabian Sea. No cyanobacteria-like *nifH* sequences were detected in any of the clone libraries from depths below the surface mixed layer.

Twenty-seven *nifH* DNA sequences from 150 and 175 m at Station 1 and three DNA sequences from 150 m at Station 2 clustered in the alpha-proteobacterial section of Cluster I (Fig. 3). These were most closely related to sequences derived from other marine environments, such as the South China Sea and the English Channel. Of particular interest is the near perfect identity of several clones from 150 and 175 m at Station 1 (S1D150M21 and S1D175M21) with a single sequence (ABW80585) previously reported from water that had been upwelled from oxygen-deficient depths in the Eastern Tropical South Pacific, an environment analogous to the one sampled here (Fernandez *et al.*, 2011). Eighteen clones (including 16 identical sequences; S1D060M24) from Station 1 surface waters (60 m) were also identified as alpha-proteobacterial sequences, but with closest identities to *Sphingomonas azotifigens* (isolated from rice roots) and several uncultured sequences from terrestrial environments.

One group of five *nifH* DNA sequences from 150 m (Station 2) and 175 m (Station 1) was nearly identical to 25 sequences from 10 m at Station 3. All of these were closely related to sequences (ABM67093, associated with aerobic alpha-proteobacteria) retrieved from mangrove sediments (Zhang *et al.*, 2008). No alpha-proteobacterial cDNA sequences were obtained from any of the samples.

Six DNA clones, including representatives from all three stations from oxycline and secondary nitrite maximum depths, were identified as likely beta-proteobacterial *nifH* sequences, with closest identities to cultivated symbiotic or soil bacteria (e.g. *Burkholderia phymatum*, *Sinorhizobium meliloti*, *Cupriavidus* spp., and *Methylocystis rosea*). This cluster included one sequence retrieved as mRNA from the Mediterranean Sea (Man-Aharonovich *et al.*, 2007) that was most closely related to *Burkholderia xenovorans* (Chain *et al.*, 2006). This sequence was thought to represent an active clade of diverse N_2 fixers in the open ocean region of the Mediterranean Sea, where N limitation in ultra-oligotrophic surface waters is thought to control productivity and microbial community composition. By contrast, no beta-proteobacterial cDNA sequences were obtained from any of the samples from the Arabian Sea.

Four DNA sequences, two from 110 m at Station 3 and two from 175 m at Station 1, were identified as likely gamma-proteobacterial *nifH* fragments. These had closest identities to sequences retrieved from aquatic marine

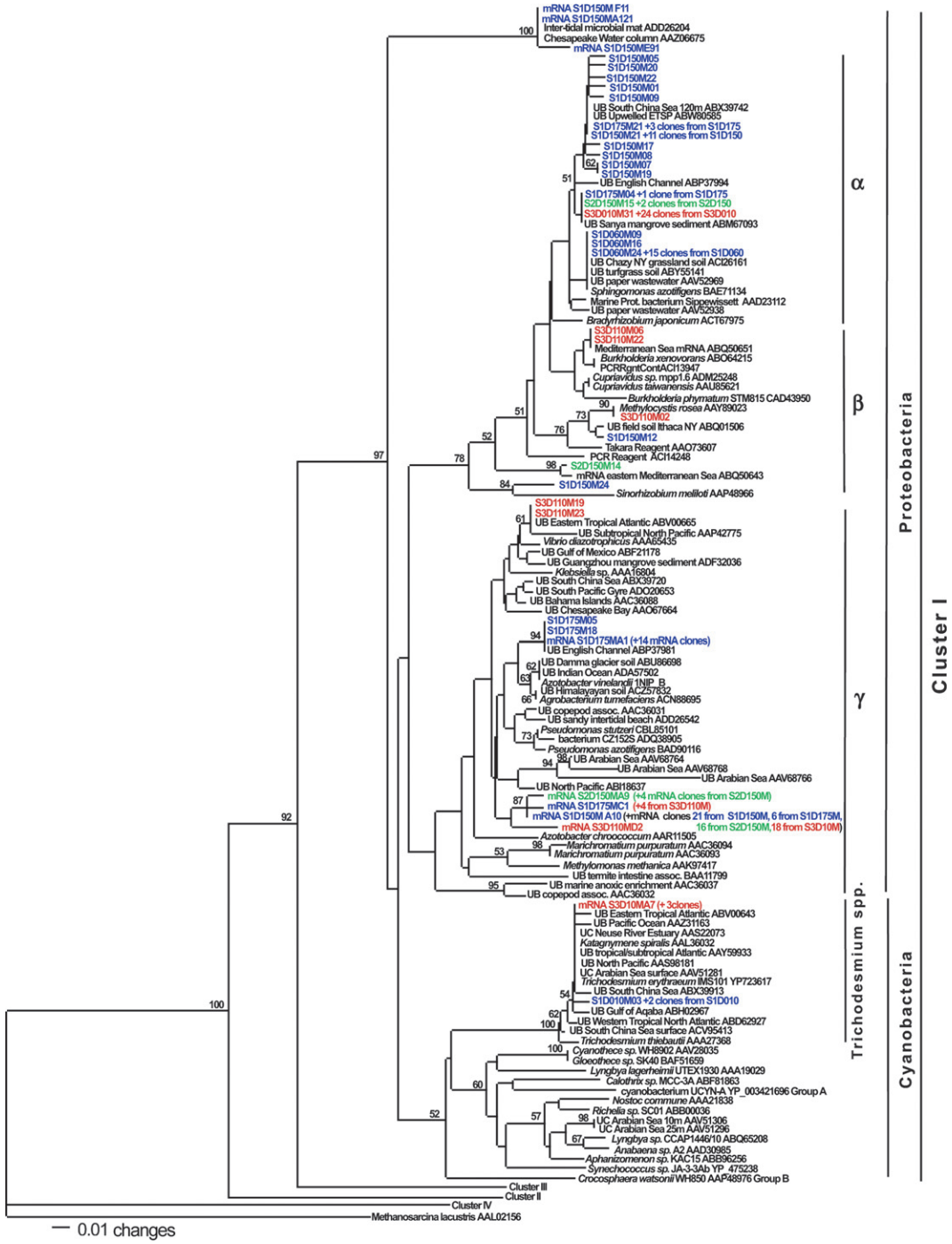

Fig. 3. Neighbor-joining tree of *nifH* amino acid sequences translated from DNA and cDNA sequences for sequences in Cluster I (Chien & Zinder, 1996). Sequences from this study are identified in color (blue = Station 1, green = Station 2, and red = Station 3), and database sequence names are in black, with accession numbers. Bootstrap values > 50% are indicated. For the new sequences, sequence names (e.g. the sequence from Station 1 at the top of the alpha-proteobacterial cluster in the figure, S1D150M05) provide Station number (S1), depth of sample (D150), and clone number (M05). The cDNA sequences names have the prefix mRNA. UB, uncultured bacterium; UC, uncultured cyanobacterium.

environments, including tropical and subtropical surface waters and the English Channel. Fifteen identical cDNA sequences (mRNA S1D175M-A1) were also identical to the DNA sequences from the ODZ at Station 1, suggesting that this gamma-proteobacterial clade is indeed active in N_2 fixation in the ODZ.

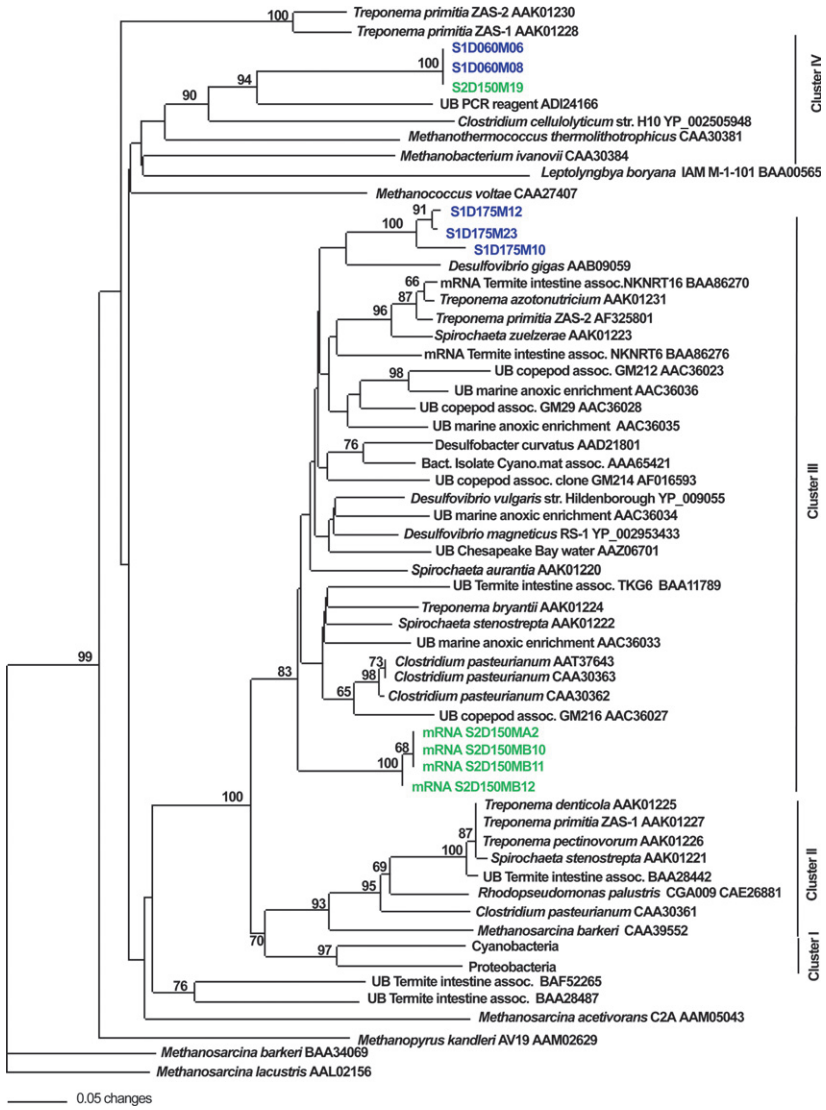

Fig. 4. Neighbor-joining tree of *nifH* amino acid sequences translated from DNA and cDNA sequences for sequences in Clusters II, III, and IV (Chien & Zinder, 1996). Sequences from this study are identified as described for Fig. 3.

The gamma-proteobacterial subcluster also contains sequences obtained from zooplankton guts (Braun *et al.*, 1999) and from soils and marine sediments. Several sequences in this cluster were identified as belonging to heterotrophic bacterial symbionts of dinoflagellates in surface waters of the southern Indian Ocean (Farnelid *et al.*, 2010). The largest group of expressed gamma-proteobacterial *nifH* genes that was detected, however, was not closely related to any of the published DNA sequences (Fig. 3). These cDNA sequences (73 in total) comprised a separate subcluster and were retrieved from all three stations at both oxycline and ODZ depths.

A cluster of anomalous cDNA sequences (top of tree, Fig. 3) were clearly associated with Cluster I, but did not fall within the cyanobacteria or any specific proteobacterial group. The translated sequences were unambiguously recognized as *nifH* sequences, but their affiliation is unknown. They were not associated with DNA sequences from this study but were identical to DNA sequences from unpublished studies, retrieved from Chesapeake Bay water column (accession no. AAZ06675) and marine microbial mats (accession no. ADD26204) from sandy intertidal beach on a Dutch barrier island.

Only six of the *nifH* DNA sequences reported here were identified phylogenetically as falling outside of Cluster I (Fig. 4). Three DNA sequences from 175 m at Station 1 grouped in Cluster III with heterotrophic bacterial genera typical of anaerobic environments, such as

sulfate reducers and free living Spirochetes from freshwater sediments (Lilburn *et al.*, 2001). A group of four cDNA sequences from the ODZ at Station 2 grouped into Cluster III, but were not identical to any of the DNA sequences. Cluster III *nifH* sequences typically derive from anaerobic gut symbionts and anoxic marine sediments (Chien & Zinder, 1996). Although the ODZ is not a sulfidic environment, the total absence of oxygen may be conducive to both free living and symbiotic N_2 fixers, at least some of which were expressing the *nifH* gene when sampled. Two DNA sequences from 60 m at Station 1 and one sequence from 150 m at Station 2 were only distantly related to other *nifH* sequences, but were identified with strong bootstrap support within Cluster IV. These sequences had greatest identity at the protein level (although still < 70% identity) with a sequence previously reported as a contaminant in PCR reagents from an unpublished study investigating *nifH* genes in surface ocean waters (accession no. ADI24166).

Proteobacterial *nifH* sequences have been reported as contaminants of PCR reagents (Zehr *et al.*, 2003), and such contaminants probably occur much more frequently than reported. It is therefore important to consider whether the Cluster IV and proteobacterial *nifH* sequences reported here do indeed represent native N_2-fixing microorganisms in the Arabian Sea, or whether they might be due to contamination of reagents. The possibility of contamination was minimized by rigorous inclusion of negative controls in the PCR analysis and a relatively low number of amplification cycles (as recommended by Zehr *et al.*, 2003), but even failure to amplify negative controls does not guarantee that sequences obtained in samples were derived from the sample alone. Amplification of trace contaminants may be enhanced in the presence of sample DNA, such that greater sensitivity is obtained in the sample than in the negative control reactions. Use of 30 cycles was intended to minimize this problem without compromising sensitivity. The proteobacterial sequences reported here are not related to others previously reported as contaminants, but this is no guarantee because almost any microbe could be present as a trace contaminant. We suggest that the sequences affiliated with sequences derived from known anaerobic species and environments may be less likely to derive from contaminants, as the reagents were not anoxic and full oxygenation of the reaction conditions likely prevailed prior to analysis.

It would be most compelling in support of active N_2 fixation if the same sequences were detected in both DNA and RNA. Some examples of identical DNA and cDNA sequences were found, but more often, the cDNA sequences were not very closely related to DNA sequences from the same samples. Detection of cDNA *nifH* sequences implies *nifH* gene expression in the cyanobacteria (*Trichodesmium* like), the *Gammaproteobacteria*, and among the sequences associated with strict anaerobes in Cluster III. The dominant cDNA clade in the *Gammaproteobacteria* contained numerous sequences from all three stations at oxycline and ODZ depths. While these sequences were not closely related to DNA sequences obtained from the same samples, the fact that they were retrieved from multiple depths, excluding the surface waters, suggests that they are characteristic of the subsurface low-oxygen waters of the Arabian Sea. That these sequences were not detected in the DNA implies that cells containing them were not abundant relative to other potential diazotrophs. Their expression, however, was apparently favored by the conditions in the ODZ at the time of sampling. Such disparity between diversity and abundance of sequences in DNA vs. cDNA clone libraries is not unusual (Short and Zehr, 2007), and even PCR and qPCR analyses of the same sample can yield quite different abundances of *nifH* sequences in clone libraries (Turk *et al.*, 2011). Thus, PCR bias could explain the differential representation in our DNA and cDNA libraries. Other explanations, for example dependent upon transient responses to environmental conditions, are possible and could be tested with further experimental work.

We conclude from this discussion of the contamination issue that the *nifH* DNA sequences reported here in the gamma-proteobacterial and alpha-proteobacterial regions of Cluster I likely represent microorganisms present at the depths of the oxycline and secondary nitrite maximum of the ODZ, based on their affinity with many other sequences from aquatic marine environments. Although we did not detect gene expression in the alphaproteobacterial clade, the DNA sequences were abundant and phylogenetically consistent with *nifH* detection from marine environments. The small number of sequences representing beta-proteobacterial *nifH* sequences is somewhat more suspect, given that they are most closely related to sequences previously reported from soil and wastewater treatment sites. Nonetheless, several sequences related to *Burkholderia* sp. were reported as dinoflagellate symbionts (Farnelid *et al.*, 2010) and identified as actively expressed in surface waters of the Mediterranean (Man-Aharonovich *et al.*, 2007), suggesting that members of this cluster are indeed natural inhabitants of the marine environment. Further exploration of the activity and regulation of these *nifH* genes are required to determine whether they contribute to N_2 fixation in the ODZ.

Direct measurements of N_2 fixation by *Trichodesmium* (Capone *et al.*, 1998; Mulholland & Capone, 2009) and indirect evidence from geochemical tracers and models (Deutsch *et al.*, 2007) all indicate that N_2 fixation occurs in surface waters of the Arabian Sea. Isotope mass

balances derived from stable isotope measurements of nitrate in the surface waters of the Arabian Sea indicate that a significant fraction of primary production depends on N_2 fixation (Brandes *et al.*, 1998). The occurrence of N_2 fixation in surface waters of the Arabian Sea is supported by the detection of both 16S rRNA and *nifH* genes representing N_2-fixing cyanobacteria (Mazard *et al.*, 2004; Bird *et al.*, 2005), as well as the detection of *Trichodesmium*-like *nifH* DNA and cDNA sequences in the present study.

A recent geochemical model suggested that denitrification and N_2 fixation are closely coupled in space and time in large regions of the ocean (Deutsch *et al.*, 2007), including in the Arabian Sea (Mulholland & Capone, 2009). Thus, evidence of N_2 fixation in both surface and subsurface waters presented here is consistent with recent models of N budgets in this region and with recent findings in the Eastern Tropical South Pacific (Fernandez *et al.*, 2011). The *nifH* sequences reported here are the first found in the ODZ of the Arabian Sea and suggest that the genetic potential for N_2 fixation occurs at the same sites where fixed N loss occurs by denitrification and anammox.

The *nifH* sequences retrieved from three stations in the open ocean Arabian Sea suggest that a diverse diazotrophic community including *Alpha-*, *Beta-*, and *Gammaproteobacteria* is present in the ODZ. Detection of numerous cDNA sequences from these stations implies that at least some of the *nifH* containing organisms are active in N_2 fixation. Differences in identity between DNA and cDNA may be due to clone library biases or transient environmental conditions that induced expression by low abundance members of the assemblage. While the presence of *nifH* genes in the ODZ suggests the capacity for N_2 fixation within and adjacent to oxygen-deficient waters, active N_2 fixation has not yet been measured in subsurface waters or within the ODZ. Molecular studies examining the distribution and abundance of *nifH* genes have recently shown the presence of *nifH* genes in a wide range of environments not previously thought to harbor N_2 fixers: archaeal N_2 fixers near hydrothermal vent systems (Mehta *et al.*, 2005; Mehta & Baross, 2006), heterotrophic bacteria associated with flagellates in surface waters (Farnelid *et al.*, 2010), and cyanobacteria in nitrate replete surface waters (Voss *et al.*, 2004; Grosse *et al.*, 2010; Moisander *et al.*, 2010). The diversity of diazotrophs and their geographical range is clearly broader than previously appreciated.

Through their opposing impacts on regional nitrogen cycling, two aspects of the Arabian Sea underscore this region's biogeochemical complexity. The first is the extensive mid-depth ODZ within which denitrification and anammox occur, comprising a major loss term for fixed N in the global fixed N budget. Second, the demonstrated presence of *Trichodesmium* spp. and other diazotrophs in surface and subsurface waters suggests that N_2 fixation may counter the ODZ-related losses of fixed N, at least to some degree (Mulholland & Capone, 2009). Geochemical evidence and rate estimates suggest that rates of N_2 fixation in the Arabian Sea could be high (Somasundar *et al.*, 1990; Gruber & Sarmiento, 1997; Brandes *et al.*, 1998; Capone *et al.*, 1998; Bange *et al.*, 2000; Deutsch *et al.*, 2007). Models estimate the annual input of new N via N_2 fixation to be 3.3 Tg N $year^{-1}$ in the Arabian Sea (Bange *et al.*, 2000, 2005). New evidence for the presence of diazotrophic organisms in oxygen-deficient waters supports the need for direct measurement of N_2 fixation rates, as well as gene expression, to assess the net contribution of this process to the Arabian Sea's nitrogen and carbon budgets.

Acknowledgements

This work was supported by grants from the National Science Foundation to AJ, MRM, and BBW. We thank the captain and crew of the *R/V Roger Revelle*, Wajih Naqvi, and the scientists at the National Institute of Oceanography who participated in the cruise.

References

Altabet MA (2007) Constraints on oceanic N balance/imbalance from sedimentary N-15 records. *Biogeosciences* **4**: 75–86.

Altabet MA, Higginson MJ & Murray DA (2002) The effect of millennial-scale changes in Arabian Sea denitrification on atmospheric CO2. *Nature* **415**: 159–162.

Bange H, Rixen T, Johansen A, Siefert R, Ramesh R, Ittekkot V, Hoffmann M & Andreae M (2000) A revised nitrogen budget for the Arabian Sea. *Global Biogeochem Cycles* **14**: 1283–1297.

Bange H, Naqvi S & Codispoti L (2005) The nitrogen cycle in the Arabian Sea. *Prog Oceanogr* **65**: 145–158.

Bird C, Martinez J, O'Donnell A & Wyman M (2005) Spatial distribution and transcriptional activity of an uncultured clade of planktonic diazotrophic gamma-proteobacteria in the Arabian Sea. *Appl Environ Microbiol* **71**: 2079–2085.

Brandes JA, Devol AH, Yoshinari T, Jayakumar D & Naqvi SWA (1998) Isotopic composition of nitrate in the central Arabian Sea and eastern tropical North Pacific: a tracer for mixing and nitrogen cycle. *Limnol Oceanogr* **43**: 1680–1689.

Braun S, Proctor L, Zani S, Mellon M & Zehr J (1999) Molecular evidence for zooplankton-associated nitrogen-fixing anaerobes based on amplification of the nifH gene. *FEMS Microbiol Ecol* **28**: 273–279.

Capone DG & Knapp AN (2007) A marine nitrogen cycle fix? *Nature* **445**: 159–160.

Capone DG, Subramaniam A, Montoya JP, Voss M, Humborg C, Johansen A, Siefert J & Carpenter EJ (1998) An extensive bloom of the N2-fixing cyanobacterium *Trichodesmium erythraeum* in the central Arabian Sea. *Mar Ecol Prog Ser* **172**: 281–292.

Chain PS, Denef VJ, Konstantinidis KT *et al.* (2006) *Burkholderia xenovorans* LB400 harbors a multi-replicon, 9.73-Mbp genome shaped for versatility. *P Natl Acad Sci USA* **103**: 15280–15287.

Chien YT & Zinder SH (1996) Cloning, functional organization, transcript studies, and phylogenetic analysis of the complete nitrogenase structural genes (nifHDK2) and associated genes in the archaeon *Methanosarcina barkeri* 227. *J Bacteriol* **178**: 143–148.

Codispoti LA (2007) An oceanic fixed nitrogen sink exceeding 400 Tg Na-1 vs the concept of homeostasis in the fixed-nitrogen inventory. *Biogeosciences* **4**: 233–253.

Codispoti LA, Brandes JA, Christensen JP, Devol AH, Naqvi SWA, Paerl HW & Yoshinari T (2001) The oceanic fixed nitrogen and nitrous oxide budgets: moving targets as we enter the anthropocene? *Sci Mar* **65**: 85–105.

Deutsch C, Sarmiento JL, Sigman DM, Gruber N & Dunne JP (2007) Spatial coupling of nitrogen inputs and losses in the ocean. *Nature* **445**: 163–167.

Devol AH (2008) Denitrification including Anammox. *Nitrogen in the Marine Environment* (Capone DG, Bronk DA, Mulholland MR & Carpenter EJ, eds), pp. 263–301. Elsevier, Amsterdam.

Farnelid H, Tarangkoon W, Hansen G, Hansen PJ & FRiemann L (2010) Putative N2-fixing heterotrophic bacteria associated with dinoflagellate-Cyanobacteria consortia in the low-nitrogen Indian Ocean. *Aquat Microb Ecol* **61**: 105–117.

Fernandez C, Farias L & Ulloa O (2011) Nitrogen fixation in denitrified marine waters. *PLoS ONE* **6**: e20539.

Grosse J, Bombar D, Doan HN, Nguyen LH & Voss M (2010) The Mekong River plume fuels nitrogen fixation and determines phytoplankton species distribution in the South China Sea during low- and high-discharge season. *Limnol Oceanogr* **55**: 1668–1680.

Gruber N & Sarmiento JL (1997) Global patterns of marine nitrogen fixation and denitrification. *Global Biogeochem Cycles* **11**: 235–266.

Hamersley MR, Turk KA, Leinweber A, Gruber N, Zehr JP, Gunderson T & Capone DG (2011) Nitrogen fixation within the water column associated with two hypoxic basins in the Southern California Bight. *Aquat Microb Ecol* **63**: 193–205.

Holl CM & Montoya JP (2005) Interactions between nitrate uptake and nitrogen fixation in continuous cultures of the marine diazotroph, *Trichodesmium*. *J Phycol* **41**: 1178–1183.

Lilburn T, Kim K, Ostrom N, Byzek K, Leadbetter J & Breznak J (2001) Nitrogen fixation by symbiotic and free-living spirochetes. *Science* **292**: 2495–2498.

Man-Aharonovich D, Kress N, Bar Zeev E, Berman-Frank I & Beja O (2007) Molecular ecology of nifH genes and transcripts in the eastern Mediterranean Sea. *Environ Microbiol* **9**: 2354–2363.

Mazard S, Fuller NJ, Orcutt KM, Brindle O & Scanlan DJ (2004) PCR analysis of the distribution of unicellular cyanobacterial dizaotrophs in the Arabian Sea. *Appl Environ Microbiol* **7**: 1525–1534.

Mehta MP & Baross JA (2006) Nitrogen fixation at 92°C by a hydrothermal vent Archaeon. *Science* **314**: 1783–1786.

Mehta MP, Huber JA & Baross JA (2005) Incidence of novel and potentially archaeal nitrogenase genes in the deep Northeast Pacific Ocean. *Environ Microbiol* **7**: 1525–1534.

Moisander PH, Beinart RA, Hewson I, White AE, Johnson KS, Carlson CA, Montoya JP & Zehr JP (2010) Unicellular cyanobacterial distributions broaden the oceanic N2 fixation domain. *Science* **327**: 1512–1514.

Mulholland MR & Capone DG (2009) Dinitrogen fixation in the Indian Ocean. *Indian Ocean Biogeochemical Processes and Ecological Variability* (Wiggert JD, Hood RR, Naqvi SWA, Brink KH & Smith SL, eds), pp. 167–186. American Geophysical Union, Washington, DC.

Mulholland MR, Ohki K & Capone DG (1999) Nitrogen utilization and metabolism relative to patterns of N2 fixation cultures of *Trichodesmium* NIBB1067. *J Phycol* **35**: 966–988.

Mulholland MR, Ohki K & Capone DG (2001) Nutrient controls on nitrogen uptake and metabolism by natural populations and cultures of *Trichodesmium* (Cyanobacteria). *J Phycol* **37**: 1001–1009.

Mulholland MR, Bernhardt PW, Blanco-Garcia JL, Mannino A, Hyde K, Mondragon E, Turk K, Moisander PH & Zehr JP (2012) Rates of dinitrogen fixation and the abundance of diazotrophs in North American coastal waters between Cape Hatteras and Georges Bank. *Limnol Oceanogr* **57**: 1067–1083.

Naqvi SWA (2008) The Indian Ocean. *Nitrogen in the Marine Environment* (Capone DG, Bronk DA, Mulholland MR & Carpenter EJ, eds), pp. 631–681. Elsevier, Amsterdam.

Padmakumar KB, Menon NR & Sanjeevan VN (2010) Occurrence of endosymbiont *Richelia intracellularis* (Cyanophyta) within the diatom *Rhizosolenia hebetata* in the Northern Arabian Sea. *Int J Biodivers Conserv* **2**: 70–74.

Parab SG, Matondkar SGP, Gomes HdR & Goes JI (2006) Monsoon driven changes in phytoplankton populations in the eastern Arabian Sea as revealed by microscopy and HPLC pigment analysis. *Cont Shelf Res* **26**: 2538–2558.

Postgate JR (1998) *The Fundamentals of Nitrogen Fixation*. Cambridge University Press, Cambridge.

Short SM & Zehr JP (2007) Nitrogenase gene expression in the Chesapeake Bay Estuary. *Environ Microbiol* **9**: 1591–1596.

Somasundar K, Rajendran A, Dileep Kumar M & Sen Gupta R (1990) Carbon and nitrogen budgets of the Arabian Sea. *Mar Chem* **30**: 363–377.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F & Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **24**: 4876–4882.

Turk KA, Rees AP, Zehr JP, Pereira N, Swift P, Shelley R, Lohan M, Woodward EMS & Gilbert J (2011) Nitrogen fixation and nitrogenase (nifH) expression in tropical waters of the eastern North Atlantic. *ISME J* **5**: 1201–1212.

Voss M, Croot P, Lochte K, Mills MM & Peeken I (2004) Patterns of nitrogen fixation along 10oN in the tropical Atlantic. *Geophys Res Lett* **31**: L23S09.

Ward BB, Devol AH, Rich JJ, Chang BX, Bulow SE, Naik H, Pratihary A & Jayakumar A (2009) Denitrification as the dominant nitrogen loss process in the Arabian Sea. *Nature* **461**: 78–82.

Zani S, Mellon MT, Collier JL & Zehr JP (2000) Expression of *nifH* genes in natural microbial assemblages in Lake George, New York, detected by reverse transcriptase PCR. *Appl Environ Microbiol* **66**: 3119–3124.

Zehr JP & McReynolds LA (1989) Use of degenerate oligonucleotides for amplification of the nifH gene from the marine cyanobacterium *Trichodesmium thiebautii*. *Appl Environ Microbiol* **55**: 2522–2526.

Zehr JP, Mellon MT & Zani S (1998) New nitrogen-fixing microorganisms detected in oligotrophic oceans by amplification of nitrogenase (nifH) genes. *Appl Environ Microbiol* **64**: 3444–3450.

Zehr J, Waterbury J, Turner P, Montoya J, Omoregie E, Steward G, Hansen A & Karl D (2001) Unicellular cyanobacteria fix N-2 in the subtropical North Pacific Ocean. *Nature* **412**: 635–638.

Zehr J, Crumbliss L, Church M, Omoregie E & Jenkins B (2003) Nitrogenase genes in PCR and RT-PCR reagents: implications for studies of diversity of functional genes. *Biotechniques* **35**: 996–1005.

Zehr JP, Bench SR, Carter BJ, Hewson I, Niazi F, Shi T, Tripp HJ & Affourtit JP (2008) Globally distributed uncultivated oceanic N_2-fixing Cyanobacteria lack oxygenic photosystem II. *Science* **322**: 1110–1112.

Zhang Y, Dong J, Yang Z, Zhang S & Wang Y (2008) Phylogenetic diversity of nitrogen-fixing bacteria in mangrove sediments assessed by PCR-denaturing gradient gel electrophoresis. *Arch Microbiol* **190**: 19–28.