Phylogenetic diversity and spatio-temporal distribution of nitrogenase genes (*nifH*) in the northern South China Sea

Liangliang Kong[1], Hongmei Jing[1], Takafumi Kataoka[1], Jun Sun[2], Hongbin Liu[1],*

[1]Division of Life Science, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR
[2]College of Marine Science and Engineering, Tianjin University of Science and Technology, Tianjin, 3000457, PR China

ABSTRACT: We examined the phylogenetic diversity and abundance of diazotrophs along a transect from the Pearl River plume to oceanic waters in the northern South China Sea by constructing a clone library and using the quantitative polymerase chain reaction (qPCR). All the diazotrophic phylogenetic groups recovered, including heterotrophic proteobacteria, phototrophic cyanobacteria and Cluster III diazotrophs, and showed distinct spatio-temporal variation along the transect. At the oceanic sampling stations, gammaproteobacteria formed the dominant diazotrophic group, with cyanobacterial diazotrophs accounting for only a relatively small proportion; at the station influenced by the Pearl River plume all the diazotrophs were non-cyanobacterial, making this station distinct from the other stations. Multidimensional analysis demonstrated that salinity was the key environmental factor determining the spatio-temporal variations of diazotrophic communities along the transect. *Trichodesmium* spp. were the most abundant among the cyanobacterial diazotrophs and were potentially the most important nitrogen fixers at all the oceanic stations. The absence of cyanobacterial diazotrophs in winter, and at the station influenced by the plume, could be due to the relatively high nutrient conditions in these waters.

KEY WORDS: Diazotrophs · *nifH* gene · Phylogenetic diversity · qPCR · South China Sea

INTRODUCTION

Nitrogen fixation, the process of converting atmospheric nitrogen to ammonium, contributes greatly to the new nitrogen input and carbon export in warm tropical and subtropical oceans (Falkowski 1997, Karl et al. 1997, Capone et al. 2005). The fixed nitrogen provides a particularly important source of nitrogen for primary production in the upper layer of water in the open ocean (Dugdale & Goering 1967) where phytoplankton growth is often limited by the bioavailable nitrogen (Vitousek & Howarth 1991). The new nitrogen input enhances carbon fixation and the export of organic matter to the deep ocean (Mahaffey et al. 2005). It has been estimated that nitrogen fixation could fuel up to 50% of primary production in the North Pacific (Karl et al. 1997) and 47% of primary production in the tropical North Atlantic (Carpenter et al. 2004).

Biological nitrogen fixation is catalyzed by a highly conserved enzyme complex which consists of 2 subunits: dinitrogenase, encoded by the *nifDK* genes, and dinitrogenase reductase, encoded by the *nifH* gene (Mehta et al. 2003). The highly conserved sequences in the gene for dinitrogenase reductase reveal the early origin of this enzyme and also provide a good basis for the phylogenetic analysis of different *nifH* phylotypes (Zehr et al. 2003). Molecular analyses based on the conserved region of the *nifH* gene make it possible to prepare clones from low

copy numbers of *nifH* genes in genomic DNA, thus enabling the study of the environmental diazotrophic community (Zehr & McReynolds 1989).

A wide range of diazotrophs have been discovered in aquatic ecosystems, all of which occur in the domains *Bacteria* and *Archaea* (Chien & Zinder 1994, Zehr et al. 2003). *Proteobacteria* and *Cyanobacteria* are the 2 main groups of diazotrophic bacteria inhabiting the upper marine ecosystems. Among them, *Cyanobacteria* are well-studied and the filamentous cyanobacterium *Trichodesmium* spp. is reported to be abundant in oligotrophic tropical and subtropical oceans (Carpenter & Romans 1991, Capone et al. 1997). Oceanic symbiotic *Cyanobacteria*, such as *Richelia* sp., occur as endosymbionts in some genera of diatoms, such as *Rhizosolenia*, *Hemiaulus* and *Chaetoceros*, and are also important nitrogen fixers (Carpenter et al. 1999, Foster & Zehr 2006). More recently, unicellular diazotrophic cyanobacteria, named group A (UCYN-A), group B (UCYN-B) and group C (UCYN-C), have been discovered on the basis of molecular techniques (Zehr et al. 2001, Langlois et al. 2005, Foster et al. 2007), and their contribution to total nitrogen fixation was estimated to be equal to, or even greater than, that of *Trichodesmium* spp. (Falcón et al. 2004, Montoya et al. 2004, Goebel et al. 2007).

Although most cyanobacterial diazotrophic groups can be detected simultaneously in the marine environment (Church et al. 2005a, Foster et al. 2007, Fong et al. 2008), different diazotrophic groups have shown different patterns of distribution in the ocean (Riemann et al. 2010). *Trichodesmium* spp. are widely and abundantly distributed in the warm surface waters of oligotrophic oceans (Capone et al. 1997, Karl et al. 1997), with optimal growth temperatures ranging from 25 to 30°C (Breitbarth 2005). UCYN-A were found to be able to occupy colder and deeper subsurface layers in the South Pacific Ocean (Moisander et al. 2010), consistent with an earlier report that unicellular diazotrophic cyanobacteria were dominant in the deep euphotic layer of the Atlantic Ocean (Langlois et al. 2005). In contrast to these 2 oceanic diazotrophic groups, the symbiotic cyanobacterium *Richelia* sp. is more prevalent in river plumes (Foster et al. 2007, 2009), where nitrogen is limiting but phosphorus and silicate are still available (Grosse et al. 2010).

The South China Sea (SCS), located in the tropical-subtropical western North Pacific, is the largest marginal sea in Asia (Shaw & Chao 1994). The warm and oligotrophic environment makes the SCS an ideal habitat for diazotrophs (Wu et al. 2003,

Chen 2005). Indeed, a recent study on the N budget indicated that up to 20% of settling particles might be contributed by nitrogen fixation in the SCS (Gaye et al. 2009). The trophic condition in the SCS, however, is variable geographically and temporally. The Asian Monsoon influences the circulation pattern and water column stability in the SCS (Wyrtki 1961, Shaw & Chao 1994), and the input from the Pearl and Mekong Rivers affects nutrient availability and nitrogen fixation activity (Cai et al. 2004, Voss et al. 2006). For example, the Mekong River plume can fuel nitrogen fixation in Vietnamese coastal waters, where nitrogen fixation has been reported to supply up to 47% of the N demand of primary production. This figure is higher than is found in other parts of the SCS (Grosse et al. 2010). Furthermore, nitrogen fixation rates measured in the upwelling region off Vietnam showed an approximately 10-fold increase during the monsoon season compared with the intermonsoon season (Voss et al. 2006). Factors mentioned above have increased the heterogeneity of the distribution of diazotrophs in this area. Previous studies have also shown a clear seasonality in the abundance of *Trichodesmium* spp. in the northern SCS, with the highest abundance in summer and autumn, and the lowest in spring (Chen et al. 2003).

So far, studies on the distribution of nitrogen fixers, and nitrogen-fixing activity, have focused on *Trichodesmium* spp. in the northern SCS; however, molecular analyses of the composition and abundance of the diazotrophic community are scarce. The only known studies focusing on the Mekong River plume and the region off the coast of Vietnam have taken place during the intermonsoon period (Moisander et al. 2008, Bombar et al. 2011), which was considered to be less favorable for nitrogen fixation (Voss et al. 2006). The spatial and seasonal changes of diazotrophic communities—especially during the southwest monsoon in summer and the northeast monsoon in winter—are largely unknown. In this study, DNA samples collected in both summer and winter in the northern SCS were examined by the clone library technique and by an SYBR green quantitative polymerase chain reaction (qPCR) assay using specific primer sets; our aims were (1) to investigate the composition of the *nifH*-bearing organisms along a transect from the coastal region (influenced by the Pearl River plume) to the open ocean, (2) to estimate the abundance and distribution of major cyanobacterial diazotrophs, and (3) to compare the structure and distribution of the diazotrophic community in summer and winter.

MATERIALS AND METHODS

Study area and sample collection

Samples were collected at 5 stations from the Pearl River plume to the oceanic region of the SCS (22° 00′ N, 114° 00′ E to 18° 01′ N, 118° 00′ E) in August 2009, and along the same transect at 4 stations in January 2010 (Fig. 1). Seawater was collected from 3 to 7 depths within the upper ocean (<150 m) of each station using Niskin bottles (volume 12 l) attached to a conductivity–temperature–depth (CTD) rosette multi-sampler. For collecting DNA samples, 500 to 1000 ml (August 2009) and 1000 to 2000 ml (January 2010) of seawater were filtered through Millipore filters (47 mm, 0.2 µm pore size) using a low vacuum. The filters were flash frozen and stored at −80°C until further analysis. Samples for nutrient analysis were collected from the same depths at all stations, and nitrate (NO_3^-), nitrite (NO_2^-) and soluble reactive phosphate (SRP) were measured onboard with a Technicon AA3 AutoAnalyzer (Bran-Lube) according to classical colorimetric methods. The detection limits for nitrate, nitrite and SRP were 0.1, 0.04 and 0.08 µmol l^{-1}, respectively. The concentration of chlorophyll *a* (chl *a*) was also measured onboard using a Turner Designs fluorometer (model Trilogy 040) after extraction with 90 % acetone (Chen et al. 2009).

DNA extraction, polymerase chain reaction (PCR) amplification of *nifH*, cloning and sequencing

Filters were cut into small pieces and incubated in 700 µl sucrose lysis buffer (50 mM sucrose, 25 mM Tris, 10 mM EDTA) containing RNase A (200 µg ml^{-1}) and lysozyme (1 mg ml^{-1}). Proteins were digested from the lysates by incubating with 1 % sodium dodecyl sulfate (SDS) and proteinase K (200 µg ml^{-1}) at 53°C for 2 h, and removed by phenol:chloroform:isoamyl alcohol (25:24:1) and chloroform:isoamyl alcohol (24:1) containing 5 M NaCl. The genomic DNA was purified by precipitation with 100 % isopropanol followed by washing with 70 % ethanol, then the air-dried genomic DNA pellet was eluted into 50 µl TE buffer. Genomic DNA from 5 surface samples (from Stns A9, A5, A0, A11 and LE09) and 2 samples from the deep chl *a* maximum (DCM) (50 m at Stn A5 and 75 m at Stn A11) in the summer cruise, and samples from 2 surface stations (A9 and A10) in the winter cruise, were selected for the cloning analysis. Fragments of the *nifH* gene were amplified using the nested PCR protocols of Zehr et al. (1998). Nuclease-free water was used for negative controls in each reaction. Positive PCR amplification was confirmed by electrophoresis on a 1.5 % agarose gel, and an amplicon of *nifH* of approximately 360 bp was purified using the Gel Band Purification Kit (GE Healthcare). The purified amplicon was ligated into the cloning vector PCR4.0, by means of the TOPO TA cloning kit (Invitrogen), and then cloned in this vector. Correct insertion was checked by direct PCR amplification of randomly selected clones, using M13 forward and reverse primers, followed by sequencing on a 3730xl DNA Analyzer (Applied Biosystems).

The sequences of *nifH* that we obtained were translated into amino acid sequences using Bioedit (Hall 1999). Operational taxonomic units (OTUs) were identified at a 5 % sequence distance cutoff value generated by DOTUR (Schloss & Handelsman
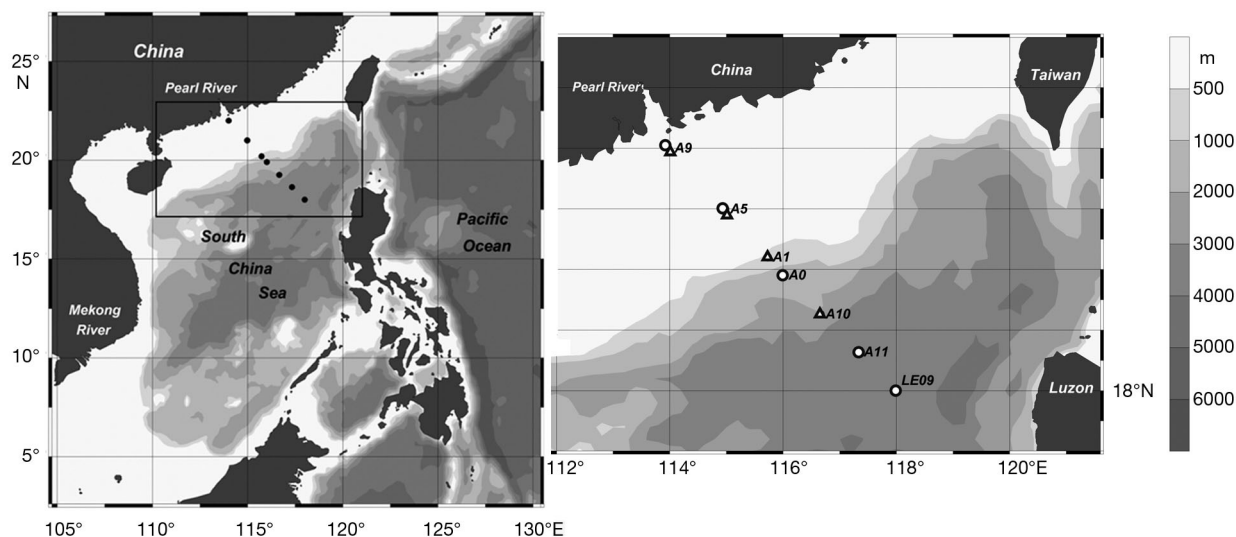

Fig. 1. Sampling stations along a transect in the northern South China Sea. Circles show sampling stations in the summer (August 2009); triangles show sampling stations in the winter (January 2010)

2005). Using the representative sequences of each OTU and the top-hit NifH protein sequences from the GenBank from a BLASTP search, a neighbor-joining tree was constructed using MEGA 4.0 (Tamura et al. 2007) with a bootstrap value of 1000 replicates. All the sequences were deposited in GenBank with the accession numbers HQ455835 to HQ456121.

Quantification of *nifH* genes using quantitative PCR (qPCR) amplification

The abundance of *nifH* genes in 4 major cyanobacterial groups—*Trichodesmium* spp., UCYN-A, UCNY-B and symbiotic heterocystous *Richelia* sp. associated with the diatom *Rhizosolenia* sp. (group Het-1)—were estimated by SYBR green qPCR analysis using previously designed primer sets (Church et al. 2005a,b). After checking the primers' specificity and sensitivity in the SYBR green qPCR assay (Short et al. 2004), 2 new forward primers were designed by the software Primer 3 (Rozen & Skaletsky 2000), targeting UCYN-A and *Crocosphaera watsonii*, respectively (Table 1). The newly designed primer sets were then tested against other closely related targets to make sure that no cross reactivity could occur.

Standard curves were determined by analyzing 10-fold serial dilutions of linear plasmids with the target *nifH* inserts, ranging from 5×10^1 to 5×10^6 gene copies, and were constructed with a linear regression of threshold cycle number (C_T) values plotted against the initial gene copy number on a log scale. qPCR was performed in triplicate in a final volume of 15 µl reaction mixture with 2 µl of extracted DNA from environmental samples, 1× Faststart Universal SYBR Green Master (ROX) (Roche Applied Science) and 300 nM each of the forward and reverse primers on a Mx3005P qPCR System (Stratagene); the PCR conditions were 95ºC for 10 min, followed by 45 cycles of 95ºC for 15 s and 60ºC for 60 s. Amplification speci-

ficity was determined by a gradual increase in temperature from 60 to 95ºC for each post-amplification melting curve. A sample inhibition test was conducted with the addition of 2 µl of randomly selected samples to the plasmid reaction, and high amplification efficiency of ≥98.5 % was obtained in all samples (Short et al. 2004, Foster et al. 2007). The gene copy number was calculated from the C_T value applied to the regression formula generated from the standard curve.

Statistical analysis

In order to reveal patterns among different stations regarding the distribution of diazotrophic assemblages and related environmental variables, detrended correspondence analysis (DCA) was carried out using CANOCO V4.5 (Biometrics-Plant Research International) to determine whether linear or unimodal species models were more suitable for our data. The length of the first DCA axis is 2.246 for diazotrophic assemblages; therefore, canonical correspondence analysis (CCA), which assumes unimodal distributions of OTUs along environmental gradients, was performed in order to reveal the relationships between community structures and environmental variables. Temperature, salinity, chl *a*, nutricline depth and phosphate concentration were included as explanatory variables in CCA, and biplot scaling was used. The effects of high collinearity among those factors were removed by eliminating variables with a variance inflation factor (VIF) >20, one at a time, beginning with the variable with the highest VIF. Forward selection was then used to determine the minimum set of environmental variables that could explain the largest amount of variance in the community. The statistical significance of an explanatory variable added in the course of forward selection was tested with the Monte Carlo permutation test (999 permutations, p ≤ 0.05).

Table 1. Primer sets specific to different cyanobacterial diazotrophic groups used for quantitative polymerase chain reaction (qPCR) analysis

Target	Forward primer (5′–3′)	Reverse primer (5′–3′)	Source
Trichodesmium	GACGAAGTATTGAAGCCAGGTTTC	CGGCCAGCGCAACCTA	Church et al. (2005a)
UCYN-A[a]	GGAACTGTAGAAGATATTGAACTTGA	ACCACGACCAGCACATCCA	Church et al. (2005a), present study
Crocosphaera[b]	AGTGTGTAGAATCTGGTGGTCCT	TCTTCTAGGAAGTTGATGGAGGTGAT	Church et al. (2005a), present study
Richelia (Het-1)	CGGTTTCCGTGGTGTACGTT	AATACCACGACCCGCACAAC	Church et al. (2005b)

[a]The forward primer for UCYN-A was designed to target a sequence in A11-DCM-07 (HQ456051)
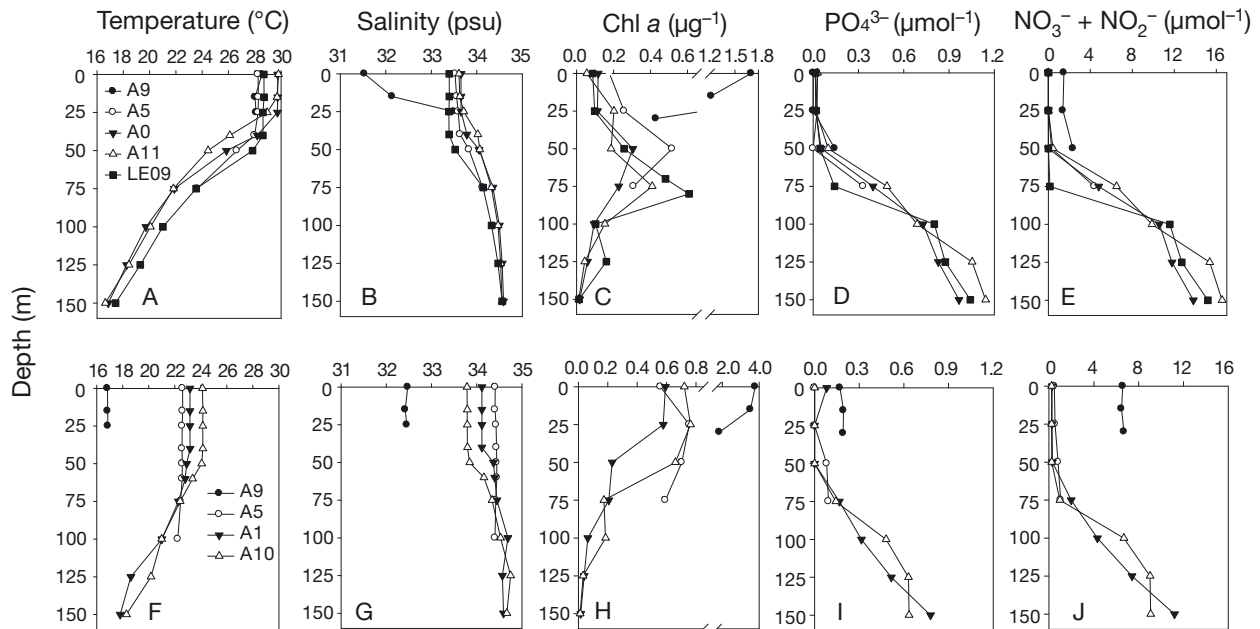[b]The forward primer for *Crocosphaera* was designed to target a sequence in LE09-S-54 (HQ456027)

Fig. 2. Depth profiles of environmental factors at the sampling stations in summer (upper panel) and winter (lower panel): (A,F) temperature, (B,G) salinity, (C,H) chlorophyll *a* (chl *a*), (D,I) soluble reactive phosphate (PO_4^{3-}), and (E,J) nitrate plus nitrite ($NO_3^- + NO_2^-$). Note that different *x*-axis scales were used for chl *a* in C and H

RESULTS

Hydrographic conditions

The study area stretched from river-influenced shelf waters to an oligotrophic oceanic basin. Vertical profiles of temperature, salinity, chl *a* and nutrients in both summer and winter are shown in Fig. 2. Both chl *a* and the concentrations of nitrite plus nitrate in the surface water were higher in the winter than in the summer. Phosphate concentration was higher in the summer than in winter—when it became undetectable. Clear DCM was detected at depths of 50 to 75 m in the summer, but it was not obvious in the winter. Generally, the plume influenced Stn A9, which is located close to the Pearl River estuary; the water here was characterized by lower salinity and much higher nutrient and chl *a* concentrations as compared with the other stations. The concentrations of surface chl *a* decreased along the transect from the Pearl River plume to the open ocean.
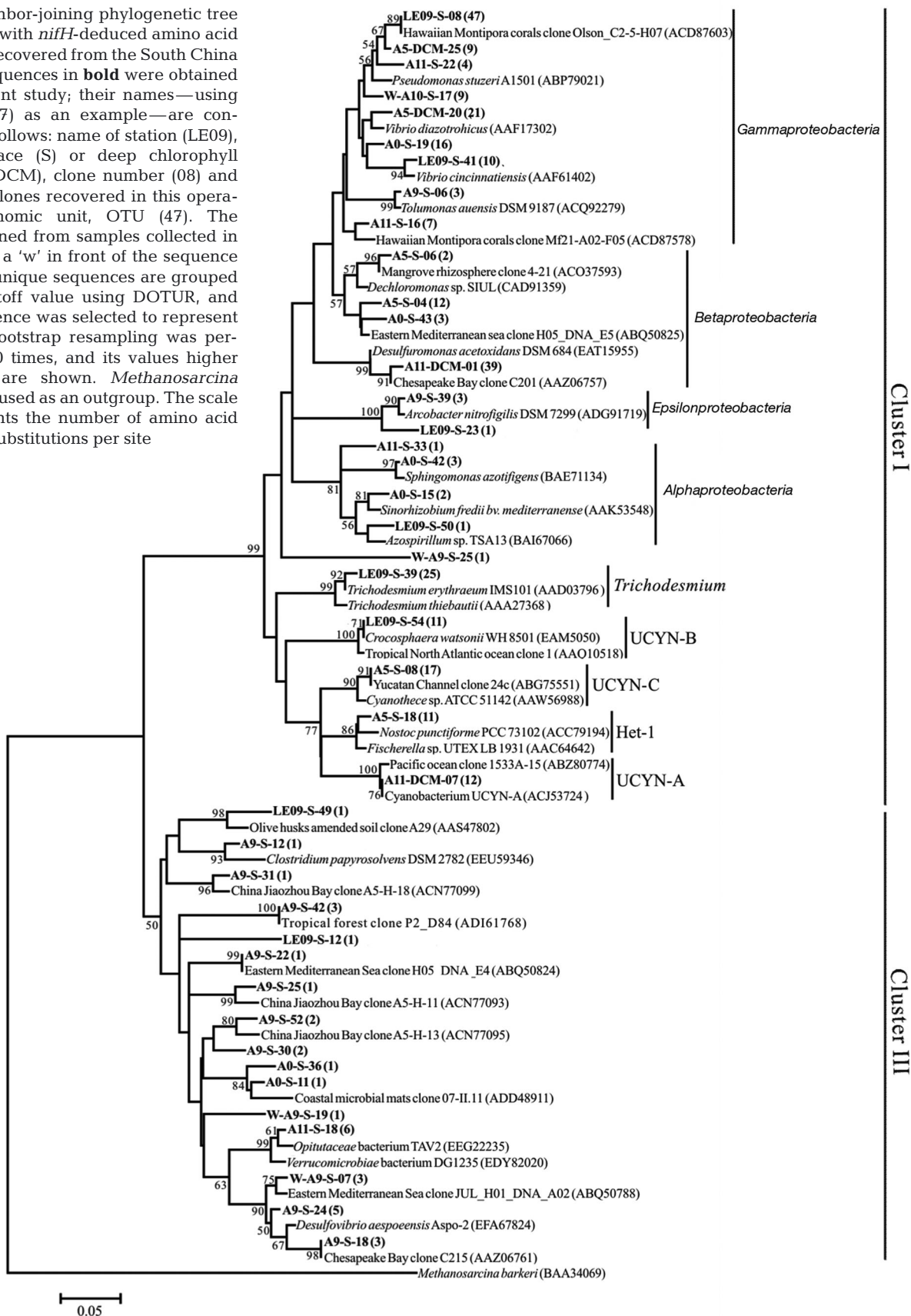
nifH phylogeny and the composition of the community

A total of 9 clone libraries were constructed in our study, with 7 samples from summer waters (5 surface and 2 DCM samples) and 2 surface samples from winter waters. Overall, 303 *nifH* sequences were successfully recovered with 63.3 to 100% sequence similarities and they were classified into 41 distinct OTUs based on a 5% amino acid sequence cutoff value. According to the *nifH* gene affiliation described by Chien & Zinder (1994), all of the sequences were grouped into *nifH* Cluster I (i.e. *Cyanobacteria* and *Proteobacteria*) and Cluster III (i.e. diverse anaerobic bacteria, such as *Clostridium*, *Deltaproteobacteria* and sulfate reducers).

Of the 303 sequences recovered, 270 and 33 sequences were distributed in Clusters I and III, respectively (Fig. 3). In Cluster I we identified 4 proteobacterial *nifH* groups and 5 cyanobacterial *nifH* groups. A total of 194 *nifH* sequences (20 OTUs) affiliated to *Proteobacteria*, with ~75.1 to 100% homology, were distributed into groups of *Gamma-*, *Beta-*, *Alpha-* and *Epsilonproteobacteria*. Among these proteobacterial *nifH* sequences, *Gammaproteobacteria* was the largest group (126 sequences, 9 OTUs), followed by *Betaproteobacteria* (56 sequences, 4 OTUs). *Trichodesmium* was the most abundant cyanobacterial group (25 sequences); sequences belonging to UCYN-A, UCYN-B and UCYN-C were also recovered in our study. Eleven *nifH* sequences obtained from surface water in Stns A5 and A0 were highly similar to the symbiotic cyanobacterial group

Fig. 3. Neighbor-joining phylogenetic tree constructed with *nifH*-deduced amino acid sequences recovered from the South China Sea. The sequences in **bold** were obtained in the present study; their names—using LE09-S-08(47) as an example—are constructed as follows: name of station (LE09), depth (surface (S) or deep chlorophyll maximum (DCM), clone number (08) and number of clones recovered in this operational taxonomic unit, OTU (47). The clones obtained from samples collected in winter have a 'w' in front of the sequence name. The unique sequences are grouped at a 5% cutoff value using DOTUR, and only 1 sequence was selected to represent the OTU. Bootstrap resampling was performed 1000 times, and its values higher than 50% are shown. *Methanosarcina barkeri* was used as an outgroup. The scale bar represents the number of amino acid substitutions per site
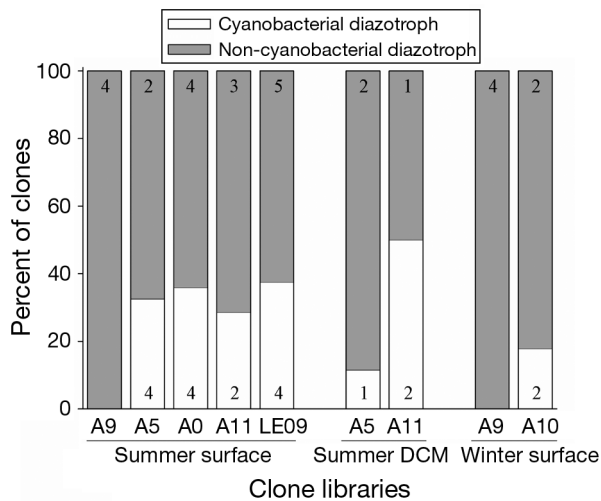
Fig. 4. Proportion of cyanobacterial and non-cyanobacterial diazotrophs detected in all clone libraries in the northern South China Sea. Numbers in the gray and white bars indicate the number of phylogenetic groups identified in each clone library. The classification of the phylogenetic groups corresponds to that of the sub-clusters in Fig. 3. All sequences recovered in Cluster III were considered as belonging to 1 group of non-cyanobacterial diazotrophs
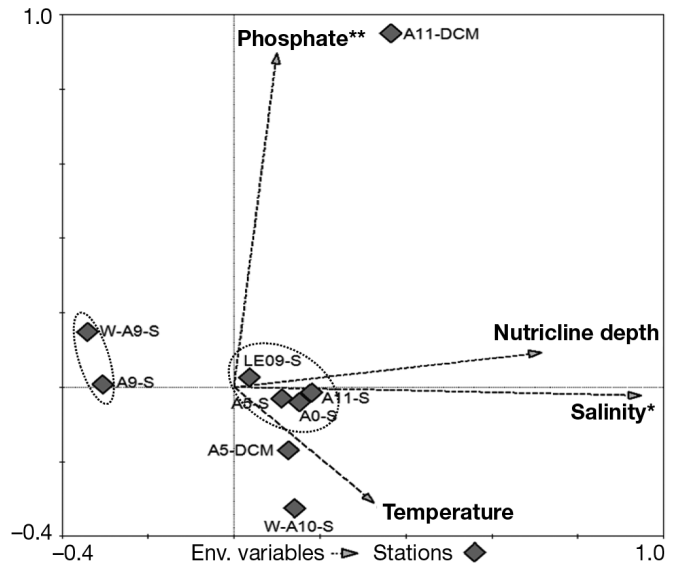

Fig. 5. Canonical correspondence analysis (CCA) biplot based on the phylogenetic composition of diazotrophs at different stations with the biotic and abiotic data set as explanatory variables. *Significant at p ≤ 0.05; **significant at p ≤ 0.01, as determined by 999 Monte Carlo permutations

Het-1. The 33 sequences in Cluster III formed 16 OTUs and only 3 of them were similar to cultured organisms, with a sequence identity higher than 95%; most of the Cluster III *nifH* sequences have their closest matches to uncultured environmental clones that were originally obtained from aquatic or terrestrial ecosystems.

The non-cyanobacterial *nifH* sequences were dominant in all clone libraries, especially in the river-plume-influenced Stn A9, in which cyanobacterial *nifH* sequences were not detected. However, compared with the DCM and winter waters, cyanobacterial *nifH* sequences were more prevalent in the surface waters in summer, accounting for 28.6 to 37.5% of the total sequences obtained. Besides the relatively prevailing cyanobacterial *nifH* sequences in the surface oceanic waters in summer, these summer oceanic surface *nifH* communities were more diverse, containing more diazotrophic groups (5–8 groups, see Fig. 4) than other samples (3–4 groups).

Statistical analysis

After excluding factors with VIFs >20, 4 variables—temperature, salinity, phosphate concentration and nutricline depth—were used for CCA. Close correlations were found between salinity and nutricline depth (r = 0.70) and between phosphate concen-

tration and temperature (r = −0.62). Forward selection analysis with significance tests of Monte Carlo permutations indicated that both salinity ($\lambda_A = 0.27$, p = 0.034) and phosphate concentration ($\lambda_A = 0.21$, p = 0.002) contributed significantly to the total variance and were closely associated with the first and the second axes, respectively. The first axis explained 51.4% of the total variance and, together, the first and second axes explained 91.7%. Biplot scaling of CCA based on the canonical axes 1 and 2 demonstrated a clear inter-sample relationship: plume-influenced Stn A9 in both summer and winter, characterized by low salinity, were plotted together on the opposite direction from salinity. The DCM at Stn A11 had a much higher phosphate concentration and therefore plotted distantly from other stations; the rest of the oceanic stations along the transect had similar salinity, temperature and phosphate concentrations and thus distributed relatively closely (Fig. 5).

Quantification of cyanobacterial *nifH* phylotypes

The abundance of 4 *nifH* phylotypes (*Trichodesmium*, UCYN-A, UCYN-B and Het-1) at all stations was determined using qPCR. We focused mainly on the cyanobacterial diazotrophs because they are one of the major sources of primary production in the ocean. Among 5 cyanobacterial *nifH* phy-

Table 2. Cross reactivity of different primer sets used in our study as indicated by the threshold cycle number (C_T) values obtained with target and non-target templates. Numbers in parentheses indicate mismatches between template and forward or reverse primer. ND = not detected

Template	Accession number	Primer set			
		Trichodesmium	UCYN-A	UCYN-B	Het-1
Trichodesmium	HQ456014	15.7 (0, 0)	ND (7, 4)	27.5 (1, 3)	35.7 (7, 3)
UCYN-A	HQ456051	31.6 (9, 4)	15.2 (0, 0)	37.1 (6, 5)	37.8 (14, 2)
UCYN-B	HQ456027	34.2 (10, 3)	ND (6, 1)	16.1 (0, 0)	35.3 (12, 1)
Het-1	HQ455894	34.3 (7, 4)	ND (7, 2)	34.9 (4, 5)	15.8 (0, 0)
UCYN-C	HQ455886	ND (9, 4)	39.2 (4, 0)	33.2 (4, 3)	32.8 (11, 2)
Gammaproteobacteria	HQ455968	31.7 (8, 5)	ND (5, 2)	35.8 (4, 6)	35.6 (12, 3)
Non-template controls		31.5	ND	ND	34.6

Table 3. Sensitivity and accuracy of the standard curves determined by primer pairs for different targets

Targets	Sensitivity				Accuracy (copy no. × 10^3)	
	Slope	R^2	Y-intercept[a]	Linearity range	Added	Detected
Trichodesmium	−3.39	0.99	34.81	10^1–10^6	5.00	4.80
UCYN-A	−3.58	0.99	35.83	10^1–10^6	5.00	5.11
UCYN-B	−3.58	0.99	37.10	10^1–10^6	5.00	4.91
Richelia (Het-1)	−3.49	0.99	35.63	10^1–10^6	5.00	5.75

[a]Y-intercept indicates the theoretical C_T value generated by quantitative polymerase chain reaction (qPCR) with 1 gene copy

lotypes detected in this study, primers targeting UCYN-C for the SYBR green qPCR assay failed to work. Therefore, the estimation of UCYN-C abundance was not included in this study. The specificity of each primer set and the cross reactivity were tested before quantifying the abundance of 4 *nifH* phylotypes in all samples (Table 2). For the primer sets specific to *Trichodesmium* spp. and group Het-1, the C_T values obtained for the non-target templates were similar or slightly higher than for the no-template controls (NTC). Gel electrophoresis and melting curve analyses showed that primer-dimers or hairpin structures might cause fluorescent signals (data not shown). A low cross reactivity was detected between the UCYN-B specific primer set and the most closely related *Trichodesmium* sequence, but the C_T value of 27.5 is 11.4 cycles more than that of the target, indicating a low amplification efficiency of 3 to 4 orders of magnitude. Therefore, it will not significantly affect the estimation of the abundance of the target *nifH* gene. No significant cross reactivity was detected for the UCYN-A primers. Four independent standard curves were constructed, with qPCR efficiency ranging from 90.2 to 97.2% and high coefficients ($R^2 \geq 0.99$, Table 3).

Distinct horizontal and vertical distribution of *nifH* gene abundance in the northern SCS was uncovered

(Fig. 6). All of the cyanobacterial *nifH* genes were detected with up to 10^6 gene copies l^{-1} in summer, the highest abundance of the *nifH* gene usually being detected in the upper water layers (0 to 50 m); cyanobacterial *nifH* genes were undetectable in winter and at the station influenced by the Pearl River plume. *Trichodesmium* was the most abundant cyanobacterial diazotroph found at all oceanic stations. One or 2 peaks of *Trichodesmium nifH* gene abundance were exhibited through the water columns at all the oceanic stations. The first peak usually appeared at 25 or 50 m, and the second peak existed at 100 or 150 m at Stns A0 and LE09, with the gene abundance up to 1.1×10^6 copies l^{-1}. At Stns A5, A11 and LE09 the UCYN-A *nifH* genes were detected only in the surface and subsurface (<50 m) layers with a range of 2.4×10^2 to 1.6×10^4 copies l^{-1}, which is 1 or 2 orders of magnitude lower than the copy number of *Trichodesmium* spp. The highest abundance of the UCYN-A *nifH* gene (1.3×10^5 gene copies l^{-1}) was detected at a depth of 50 m at Stn A0, where this gene was abundant throughout the water column from the surface to 125 m. The abundance of UCYN-B and Het-1 *nifH* gene sequences was similar to that of UCYN-A in the surface layers; these gene sequences became undetectable below 25 m at most of the stations. In winter, only UCYN-B
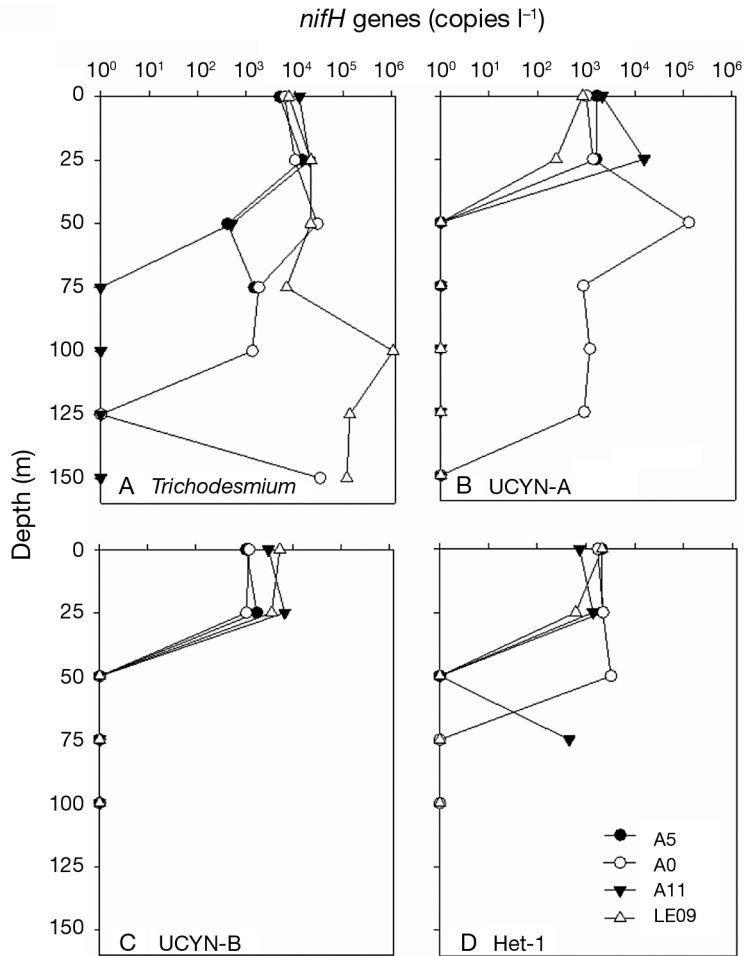
Fig. 6. Vertical distribution of the abundance of *nifH* genes at all the oceanic stations in summer. One gene copy was used to represent where the *nifH* gene was below detection. Quantitative polymerase chain reaction (qPCR) analysis of samples from the plume-influenced station A9, and of samples collected in winter, did not generate positive signals

nifH sequences were detected sporadically in the surface layer at Stn A10 (6.5×10^3 gene copies l^{-1}).

DISCUSSION

The SCS is an oligotrophic nitrogen-limited marginal sea with an N:P ratio of less than 3.1 in the surface layer of the basin (Chen 2005). The water column in summer is oligotrophic and well stratified with a clear DCM and depleted inorganic nitrogen in the surface, while a strong and cold northeast monsoon in winter enhances vertical mixing that brings abundant inorganic nutrients to the upper layers. The depleted inorganic N with detectable phosphate in the summer and the opposite condition in winter observed in this study indicates that the oceanic sea-

water in the SCS is more N-limited in summer than in winter, which would be reflected in a difference in the composition and abundance of the diazotrophs. In addition, the Pearl River plume is rather constrained to the nearshore region and quickly turns eastwards with the cyclonic circulation during the southwest monsoon in summer, although the Pearl River discharge has a much higher nitrate concentration and an N:P ratio of 100 to 600 (Cai et al. 2004). In our study, Stn A9 is the only station influenced by the Pearl River plume, which resulted in a diazotrophic community distinct from those at other stations.

The dominance of gammaproteobacteria in our study is not consistent with previous studies, which showed that cyanobacterial diazotrophs, especially *Trichodesmium* spp., are dominant in the oligotrophic Pacific and Atlantic (Foster et al. 2007, Fong et al. 2008, Langlois et al. 2008) as well as in the southern SCS (Moisander et al. 2008, Bombar et al. 2011). However, several recent studies in the eastern North Atlantic and worldwide oceans have suggested that non-cyanobacterial diazotrophs dominate at the ocean surface (Farnelid et al. 2011, Turk et al. 2011). A possible reason why this has been reported is the preferential amplification of gammaproteobacterial sequences in clone libraries (Turk et al. 2011). Comparatively, alphaproteobacteria accounted for a relatively small portion in all the diazotrophic communities in our study, in contrast to the finding that alphaproteobacteria formed the predominant group in the diazotrophic communities off the Vietnamese coast in the southern SCS (Moisander et al. 2008). This difference could be the result of geographical variation, although alphaproteobacteria and gammaproteobacteria are generally the 2 main heterotrophic groups in oligotrophic marine waters (Zehr et al. 1998, Falcón et al. 2004). Cyanobacterial *nifH* gene sequences showed much higher similarities to the sequences previously obtained in the global oceans, reflecting their ubiquitous distribution.

Cluster III diazotrophs are anaerobic bacteria periodically detected in the ocean's surface (Church et al. 2005a, Foster et al. 2007, Man-Aharonovich et al. 2007). Most (67%) of the Cluster III *nifH* sequences recovered from our study were from the plume-influ-

enced Stn A9, implying that this group might be terrestrial nitrogen fixers being carried out to sea by the freshwater discharge. These anaerobic bacteria could be associated with particles, such as marine snow and fecal pellets, to achieve anoxic microenvironments (Steward et al. 2004). It is interesting to note that 6 sequences were affiliated with *Verrucomicrobiae* bacterium DG1235 and *Opitutaceae* bacterium TAV2, belonging to the phylum *Verrucomicrobia*. Recently, the first report on the physiology of nitrogen fixation by the strain *Methylacidiphilum fumariolicum* SolV within this phylum revealed that the activity of nitrogenase was not inhibited under ammonia concentrations up to 94 mM (Khadem et al. 2010). To the best of our knowledge, we report here for the first time that *Verrucomicrobia nifH* genes were detected from an oceanic environment.

Multidimensional analysis demonstrated that salinity was the major environmental parameter explaining the variability in diazotrophic communities among stations. Other environmental factors, such as concentrations of inorganic nutrients, may have played a more important role in determining the composition and distribution of diazotrophic communities, but, because they were mostly below the limit of detection, the importance of the impact of such factors cannot be quantified. Significantly negative relationships between salinity and nutrient concentrations or N:P ratio were suggested in the Mekong River plume in the southern part of the SCS (Bombar et al. 2011). Samples from Stn A9, located in the Pearl River plume, contained only non-cyanobacterial diazotrophs; this may be a consequence of a high nitrate concentration associated with the low salinity resulting from the direct effect of the fresh water discharge. Among surface samples from oceanic stations, the composition of the community was found to show a relatively high level of similarity, but in these samples the composition of the community was distinct from that found in the plume-influenced Stn A9—in both summer and winter. Generally, the spatial variation in the composition of diazotrophic communities was more obvious than the seasonal variation.

As already pointed out by Man-Aharonovich et al. (2007), the functional *nifH* communities could sometimes be totally different from what was expected from the analysis of DNA samples. Although cyanobacterial diazotrophs were not the predominant groups recovered from our clone libraries, they could nevertheless be important nitrogen fixers in the SCS, albeit with apparent variations in spatio-temporal

distribution (Chen et al. 2003, Wu et al. 2003). The nitrogen fixation rate was also measured on the same cruise (but at different stations); we found the surface nitrogen fixation rate to be 0.047 to 0.43 nmol N l^{-1} h^{-1} in the summer and 0.011 to 0.087 nmol N l^{-1} h^{-1} in winter (J. Sun unpubl. data). These rates are similar to those measured in the tropical eastern North Atlantic (0.1 to 0.5 nmol N l^{-1} h^{-1}) (Turk et al. 2011), but higher than those measured at Stn ALOHA (0.01 to 0.15 nmol N l^{-1} h^{-1}) (Montoya et al. 2004) or in the North Pacific (0.014 to 0.053 nmol N l^{-1} h^{-1}) (Zehr et al. 2007). They are, however, much lower than the nitrogen fixation rate (5.76 to 22.77 nmol N l^{-1} h^{-1}) recorded in the transitional region between the Mekong River plume and the adjacent SCS waters (Grosse et al. 2010). In addition, the nitrogen fixation rate obtained in the northern SCS in summer showed a relatively higher value in shelf waters (near Stn A5) than in the oceanic region, with the major contribution (>50%) being from large-sized diazotrophs (>10 μm), presumably *Trichodesmium* spp. In agreement with the distribution of the *nifH* gene in the present study, the nitrogen fixation rate in the study area was low in winter when compared with summer, and no spatial pattern was recorded.

Among diazotrophic cyanobacteria, *Trichodesmium* spp. are the dominant group in our study area in summer. We estimated that the average abundance of *Trichodesmium* spp. in the upper water layer (<25 m) is ~10^4 cells l^{-1} by assuming that 1 *Trichodesmium* cell contains only 1 genome (Moisander et al. 2008) and a typical trichome contains ~100 cells (Capone et al. 1997). Our estimation is slightly higher than that of the previous observation in the northern SCS, with an average density of 77 trichomes l^{-1} at the surface, but much lower than values reported off the Vietnamese coast (Moisander et al. 2008), the North Pacific (Church et al. 2005b) and the Amazon River plume in the Atlantic (Foster et al. 2007). The occurrence of a second peak at the deep layer (~100 m), which also appears in the northern SCS (Chen et al. 2003) and North Pacific (Marumo & Asaoka 1974), is probably the result of the downward flux of *Trichodesmium* spp. *Trichodesmium* colonies can sink to a depth of 100 to 200 m by synthesizing or consuming cellular carbohydrates (Chen et al. 2003) and the presence of concentrated trichomes in the deep layer was also confirmed by our fluorescence microscopy observation (data not shown).

The highest abundance of UCYN-A *nifH* gene sequences was detected at a depth of 50 m at Stn A0; this is so far the highest abundance of UCYN-A detected in the SCS and is comparable to the obser-

vations in the North Pacific (Church et al. 2005a, 2008, Fong et al. 2008) and the Amazon River plume (Foster et al. 2007). This group of diazotrophs was found to be more widely distributed in the tropical and subtropical ocean, including nutrient-replete waters (Moisander et al. 2010), such as the upwelling region in the Eastern Equatorial Atlantic (Foster et al. 2009), but their *in situ* N-fixing function is still not well understood. By comparison, UCYN-B preferentially inhabited the oligotrophic surface waters; their distribution in our study was constrained to the upper water layers (<25 m) with an increasing trend towards the open ocean. Symbiotic cyanobacteria, such as *Rhizosolenia–Richelia* and *Hemiaulus–Richelia*, have been reported to be abundant in the Amazon plume (Foster et al. 2007) and Mekong River plume (Moisander et al. 2008, Bombar et al. 2011). However, only the symbiotic group Het-1 *nifH* sequence (*Rhizosolenia–Richelia* symbioses) was found in our study, and the lack of group Het-2 (*Hemiaulus–Richelia* symbioses) could be the result of a bias caused by our small sampling volume (1 to 2 l), although the latter group has been found to be more abundant than the former in the southern SCS (Moisander et al. 2008). In addition, in winter, cyanobacterial diazotrophs were below the detection limit by qPCR; this is also in accordance with the relatively low winter nitrogen fixation rate (J. Sun et al. unpubl. data). The absence of cyanobacterial diazotrophs in the plume-influenced station, and in winter samples, confirms that the elevated N would create an environment that is not conducive to cyanobacterial diazotrophs and their distribution is confined to areas with limited bioavailable N (Capone et al. 1997, Church et al. 2008).

CONCLUSIONS

This is the first study on the diversity, distribution and abundance of *nifH* genes in the northern SCS. A highly diversified diazotrophic community was revealed—including heterotrophic proteobacteria, phototrophic cyanobacteria and Cluster III diazotrophs—and the community structures varied both spatially and temporally. Samples from the plume-influenced station contained only non-cyanobacterial diazotrophs; this is in contrast to the samples from the other oceanic stations. Salinity was found to be the most important environmental factor determining the spatio-temporal changes in diazotrophic communities along the transect from the river plume to the oligotrophic open ocean. In addition, *Trichodesmium*

spp. were the most abundant cyanobacterial diazotrophs—and potentially the most important nitrogen fixers in the shelf and oceanic waters. The absence of cyanobacterial diazotrophs in winter, and in the plume-influenced area, could be due to the relatively high nutrient conditions in these areas.

LITERATURE CITED

Bombar D, Moisander PH, Dippner JW, Foster RA, Voss M, Karfeld B, Zehr JP (2011) Distribution of diazotrophic microorganisms and *nifH* gene expression in the Mekong River plume during intermonsoon. Mar Ecol Prog Ser 424:39–52

Breitbarth E (2005) Ecophysiology of the cyanobacterium *Trichodesmium*. PhD dissertation, Christian Albrechts University of Kiel, Kiel

Cai W, Dai M, Wang Y, Zhai W and others (2004) The biogeochemistry of inorganic carbon and nutrients in the Pearl River estuary and the adjacent Northern South China Sea. Cont Shelf Res 24:1301–1319

Capone DG, Zehr JP, Paerl HW, Bergman B, Carpenter EJ (1997) *Trichodesmium*, a globally significant marine cyanobacterium. Science 276:1221–1229

Capone DG, Burns JA, Montoya JP, Subramaniam A and others (2005) Nitrogen fixation by *Trichodesmium* spp. An important source of new nitrogen to the tropical and subtropical North Atlantic Ocean. Global Biogeochem Cycles 19, GB2024, doi:10.1029/2004GB002331

Carpenter EJ, Romans K (1991) Major role of the cyanobacterium *Trichodesmium* in nutrient cycling in the North Atlantic Ocean. Science 254:1356–1358

Carpenter EJ, Montoya JP, Burns J, Mulholland MR, Subramaniam A, Capone DG (1999) Extensive bloom of a N$_2$-fixing diatom/cyanobacterial association in the tropical Atlantic Ocean. Mar Ecol Prog Ser 185:273–283

Carpenter EJ, Subramaniam A, Capone DG (2004) Biomass and primary productivity of the cyanobacterium *Trichodesmium* spp. in the tropical N Atlantic Ocean. Deep-Sea Res I 51:173–203

Chen B, Liu H, Landry MR, Dai M, Huang B, Sun J (2009) Close coupling between phytoplankton growth and microzooplankton grazing in the western South China Sea. Limnol Oceanogr 54:1084–1097

Chen YlL, Chen HY, Lin YH (2003) Distribution and downward flux of *Trichodesmium* in the South China Sea as influenced by the transport from the Kuroshio Current. Mar Ecol Prog Ser 259:47–57

Chen YlL (2005) Spatial and seasonal variations of nitrate-based new production and primary production in the South China Sea. Deep-Sea Res I 52:319–340

Chien YT, Zinder S (1994) Cloning, DNA sequencing, and characterization of a *nifD*-homologous gene from the

archaeon *Methanosarcina barkeri* 227 which resembles *nifD1* from the eubacterium *Clostridium pasteurianum.* J Bacteriol 176:6590–6598

Church MJ, Jenkins BD, Karl DM, Zehr JP (2005a) Vertical distributions of nitrogen-fixing phylotypes at Stn ALOHA in the oligotrophic North Pacific Ocean. Aquat Microb Ecol 38:3–14

Church MJ, Short CM, Jenkins BD, Karl DM, Zehr JP (2005b) Temporal patterns of nitrogenase gene (*nifH*) expression in the oligotrophic North Pacific Ocean. Appl Environ Microbiol 71:5362–5370

Church MJ, Björkman KM, Karl DM, Saito MA, Zehr JP (2008) Regional distributions of nitrogen-fixing bacteria in the Pacific Ocean. Limnol Oceanogr 53:63–77

Dugdale RC, Goering JJ (1967) Uptake of new and regenerated forms of nitrogen in primary productivity. Limnol Oceanogr 12:196–206

Falcón LI, Carpenter EJ, Cipriano F, Bergman B, Capone DG (2004) N_2 Fixation by unicellular bacterioplankton from the Atlantic and Pacific Oceans: phylogeny and *in situ* rates. Appl Environ Microbiol 70:765–770

Falkowski PG (1997) Evolution of the nitrogen cycle and its influence on the biological sequestration of CO_2 in the ocean. Nature 387:272–275

Farnelid H, Andersson AF, Bertilsson S, Al-Soud WA and others (2011) Nitrogenase gene amplicons from global marine surface waters are dominated by genes of non-cyanobacteria. PLoS ONE 6:e19223

Fong AA, Karl DM, Lukas R, Letelier RM, Zehr JP, Church MJ (2008) Nitrogen fixation in an anticyclonic eddy in the oligotrophic North Pacific Ocean. ISME J 2:663–676

Foster RA, Zehr JP (2006) Characterization of diatom–cyanobacteria symbioses on the basis of *nifH*, *hetR* and 16S rRNA sequences. Environ Microbiol 8:1913–1925

Foster RA, Subramaniam A, Mahaffey C, Carpenter EJ, Capone DG, Zehr JP (2007) Influence of the Amazon River plume on distributions of free-living and symbiotic cyanobacteria in the western tropical north Atlantic Ocean. Limnol Oceanogr 52:517–532

Foster RA, Subramaniam A, Zehr JP (2009) Distribution and activity of diazotrophs in the Eastern Equatorial Atlantic. Environ Microbiol 11:741–750

Gaye B, Wiesner MG, Lahajnar N (2009) Nitrogen sources in the South China Sea, as discerned from stable nitrogen isotopic ratios in rivers, sinking particles, and sediments. Mar Chem 114:72–85

Goebel NL, Edwards CA, Church MJ, Zehr JP (2007) Modeled contributions of three types of diazotrophs to nitrogen fixation at Stn ALOHA. ISME J 1:606–619

Grosse J, Bombar D, Doan HN, Nguyen LN, 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

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98

Karl D, Letelier R, Tupas L, Dore J, Christian J, Hebel D (1997) The role of nitrogen fixation in biogeochemical cycling in the subtropical North Pacific Ocean. Nature 388:533–538

Khadem AF, Pol A, Jetten MSM, Op den Camp HJM (2010) Nitrogen fixation by the verrucomicrobial methanotroph '*Methylacidiphilum fumariolicum*' SolV. Microbiology 156:1052–1059

Langlois RJ, LaRoche J, Raab PA (2005) Diazotrophic diversity and distribution in the tropical and subtropical Atlantic Ocean. Appl Environ Microbiol 71:7910–7919

Langlois RJ, Hummer D, LaRoche J (2008) Abundances and distributions of the dominant *nifH* phylotypes in the Northern Atlantic Ocean. Appl Environ Microbiol 74: 1922–1931

Mahaffey C, Michaels AF, Capone DG (2005) The conundrum of marine N_2 fixation. Am J Sci 305:546–595

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

Marumo R, Asaoka O (1974) Distribution of pelagic blue-green algae in the North Pacific Ocean. J Oceanogr Soc Jpn 30:77–85

Mehta MP, Butterfield DA, Baross JA (2003) Phylogenetic diversity of nitrogenase (*nifH*) genes in deep-sea and hydrothermal vent environments of the Juan de Fuca Ridge. Appl Environ Microbiol 69:960–970

Moisander PH, Beinart RA, Voss M, Zehr JP (2008) Diversity and abundance of diazotrophic microorganisms in the South China Sea during intermonsoon. ISME J 2: 954–967

Moisander PH, Beinart RA, Hewson I, White AE and others (2010) Unicellular cyanobacterial distributions broaden the oceanic N_2 fixation domain. Science 327:1512–1514

Montoya JP, Holl CM, Zehr JP, Hansen A, Villareal TA, Capone DG (2004) High rates of N_2 fixation by unicellular diazotrophs in the oligotrophic Pacific Ocean. Nature 430:1027–1032

Riemann L, Farnelid H, Steward GF (2010) Nitrogenase genes in non-cyanobacterial plankton prevalence, diversity and regulation in marine waters. Aquat Microb Ecol 61:235–247

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics methods and protocols: methods in molecular biology. Humana Press, Totowa, NJ, p 365–386

Schloss PD, Handelsman J (2005) Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. Appl Environ Microbiol 71:1501–1506

Shaw P, Chao S (1994) Surface circulation in the South China Sea. Deep-Sea Res I 41:1663–1683

Short SM, Jenkins BD, Zehr JP (2004) Spatial and temporal distribution of two diazotrophic bacteria in the Chesapeake Bay. Appl Environ Microbiol 70:2186–2192

Steward GF, Zehr JP, Jellison R, Montoya JP, Hollibaugh JT (2004) Vertical distribution of nitrogen-fixing phylotypes in a meromictic, hypersaline lake. Microb Ecol 47:30–40

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599

Turk KA, Rees AP, Zehr JP, Perira N and others (2011) Nitrogen fixation and nitrogenase (*nifH*) expression in tropical waters of the eastern North Atlantic. ISME J 5:1201–1212

Vitousek PM, Howarth RW (1991) Nitrogen limitation on land and in the sea—How can it occur? Biogeochemistry 13:87–115

Voss M, Bombar D, Loick N, Dippner JW (2006) Riverine influence on nitrogen fixation in the upwelling region off Vietnam, South China Sea. Geophys Res Lett 33, L07604, doi:10.1029/2005GL025569

Wu J, Chung SW, Wen LS, Liu KK, Chen YlL, Chen HY, Karl DM (2003) Dissolved inorganic phosphorus, dissolved iron, and *Trichodesmium* in the oligotrophic South China Sea. Global Biogeochem Cycles 17, 1008, doi:10.1029/2002GB001924

Wyrtki K (1961) Physical oceanography of the southeast Asian waters. NAGA Report vol 2, Scientific results of marine investigations of the South China Sea and the Gulf of Thailand. Scripps Institution of Oceanography, La Jolla, CA

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 JP, Waterbury JB, Turner PJ, Montoya JP and others (2001) Unicellular cyanobacteria fix N_2 in the subtropical North Pacific Ocean. Nature 412:635–638

Zehr JP, Jenkins BD, Short SM, Steward GF (2003) Nitrogenase gene diversity and microbial community structure: a cross-system comparison. Environ Microbiol 5:539–554

Zehr JP, Montoya JP, Jenkins BD, Hewson I and others (2007) Experiments linking nitrogenase gene expression to nitrogen fixation in the North Pacific subtropical gyre. Limnol Oceanogr 52:169–183