

## ORIGINAL ARTICLE

# Heterotrophic organisms dominate nitrogen fixation in the South Pacific Gyre

Hannah Halm<sup>1,5</sup>, Phyllis Lam<sup>1</sup>, Timothy G Ferdelman<sup>1</sup>, Gaute Lavik<sup>1</sup>, Thorsten Dittmar<sup>2</sup>, Julie LaRoche<sup>3</sup>, Steven D'Hondt<sup>4</sup> and Marcel MM Kuypers<sup>1</sup>

<sup>1</sup>Max Planck Institute for Marine Microbiology, Bremen, Germany; <sup>2</sup>Max Planck Research Group for Marine Geochemistry, University of Oldenburg, Oldenburg, Germany; <sup>3</sup>IFM-GEOMAR, Kiel, Germany and <sup>4</sup>Graduate School of Oceanography, University of Rhode Island, Narragansett Bay Campus, Narragansett, RI, USA

Oceanic subtropical gyres are considered biological deserts because of the extremely low availability of nutrients and thus minimum productivities. The major source of nutrient nitrogen in these ecosystems is N<sub>2</sub>-fixation. The South Pacific Gyre (SPG) is the largest ocean gyre in the world, but measurements of N<sub>2</sub>-fixation therein, or identification of microorganisms involved, are scarce. In the 2006/2007 austral summer, we investigated nitrogen and carbon assimilation at 11 stations throughout the SPG. In the ultra-oligotrophic waters of the SPG, the chlorophyll maxima reached as deep as 200 m. Surface primary production seemed limited by nitrogen, as dissolved inorganic carbon uptake was stimulated upon additions of <sup>15</sup>N-labeled ammonium and leucine in our incubation experiments. N<sub>2</sub>-fixation was detectable throughout the upper 200 m at most stations, with rates ranging from 0.001 to 0.19 nmol N h<sup>-1</sup>. N<sub>2</sub>-fixation in the SPG may account for the production of 8–20% of global oceanic new nitrogen. Interestingly, comparable <sup>15</sup>N<sub>2</sub>-fixation rates were measured under light and dark conditions. Meanwhile, phylogenetic analyses for the functional gene biomarker *nifH* and its transcripts could not detect any common photoautotrophic diazotrophs, such as, *Trichodesmium*, but a prevalence of  $\gamma$ -proteobacteria and the unicellular photoheterotrophic Group A cyanobacteria. The dominance of these likely heterotrophic diazotrophs was further verified by quantitative PCR. Hence, our combined results show that the ultra-oligotrophic SPG harbors a hitherto unknown heterotrophic diazotrophic community, clearly distinct from other oceanic gyres previously visited.

*The ISME Journal* (2012) 6, 1238–1249; doi:10.1038/ismej.2011.182; published online 15 December 2011

**Subject Category:** microbial ecology and functional diversity of natural habitats

**Keywords:** diazotrophs;  $\gamma$ -proteobacteria; *nifH*; (photo)heterotrophic nitrogen fixation; UCYN-A; ultra-oligotrophy

## Introduction

Nitrogen fixation is the primary source of fixed nitrogen in the open ocean, and much of it is believed to occur in oligotrophic subtropical gyres (Sohm *et al.*, 2011). It is a key process influencing global oceanic productivity and associated carbon fluxes. These open-ocean ecosystems cover 40% of the earth's surface, and their nitrogen-fixing communities have a substantial role in global oceanic carbon and nitrogen budgets (Karl *et al.*, 2002). The largest of all gyre systems is the South Pacific Gyre (SPG). Owing to its extreme remoteness from any continents, surface water of the SPG is the most oligotrophic in global oceans, with the lowest sea

surface chlorophyll *a* (chl *a*) concentrations and the clearest natural water in the world (McClain *et al.*, 2004; Morel *et al.*, 2007). In contrast to the subtropical gyres of the northern hemisphere, investigations of N-cycling in the southern hemisphere are scarce. Very few nitrogen fixation rates have been reported so far for the central SPG (Moutin *et al.*, 2008; Raimbault and Garcia, 2008). Even less is known about the organisms responsible for N<sub>2</sub>-fixation therein. One recent study targeted the biomarker nitrogenase gene, *nifH*, but could not detect any of the common diazotrophic cyanobacteria like *Trichodesmium* or other large Group B cyanobacteria (such as *Crocospaera*) at the central gyre stations, in contrast to the gyre edge or other ocean gyre systems (Bonnet *et al.*, 2008). Rather, only low numbers of the unicellular Group A cyanobacteria (UCYN-A) (~300 copies per l) were enumerated (Bonnet *et al.*, 2008). However, the sparseness of data available precluded any detailed characterization of the diazotrophic community within this enormous, largely unexplored ultra-oligotrophic biome.

Correspondence: H Halm, IFM-GEOMAR, Düsternbrooker Weg 20, 24105 Kiel, Germany.

E-mail: hhalm@ifm-geomar.de

<sup>5</sup>Current address: IFM-GEOMAR, Düsternbrooker Weg 20, 24105 Kiel, Germany.

Received 22 May 2011; revised 3 November 2011; accepted 3 November 2011; published online 15 December 2011

Here, we quantify rates of  $N_2$ -fixation and investigate the community structure responsible for  $N_2$ -fixation in the heart of this most oligotrophic gyre. In December 2006 to January 2007, we surveyed: (1) a zonal transect from the western edge ( $23^\circ 51' S$   $165^\circ 39' W$ ) of the gyre to the center ( $27^\circ 44' S$   $117^\circ 37' W$ ) and (2) a zonal transect along the southern edge of the gyre and across the south Pacific subtropical front (from  $38^\circ 04' S$   $133^\circ 06' W$  to  $45^\circ 58' S$   $163^\circ 11' W$ ). Rates of  $N_2$ -fixation were experimentally determined, in parallel with experiments to determine rates of ammonium and leucine uptakes, and the corresponding rates of dissolved inorganic carbon (DIC) fixation. Organisms responsible for active  $N_2$ -fixation were revealed by phylogenetic analyses of the biomarker gene *nifH* and their expressed transcripts. Various groups of  $N_2$ -fixing organisms were further quantified with newly designed and already available primers and probes targeting *nifH*.

## Materials and methods

### *Water-column sampling and hydrochemical measurements*

Sampling was conducted onboard the *R/V Roger Revelle* (Knox-02RR) in austral summer (December 2006/January 2007) from Apia, Samoa to Dunedin, New Zealand (see Figures 1a and b). Water samples were taken in the upper 200 m from a CTD-rosette system (Sea-Bird, Bellevue, WA, USA) fitted with twenty-four 2.5 l Niskin bottles and a fluorometer (Seapoint, Exeter, NH, USA) for chl *a* measurements. However, chl *a* was not measured at the first station because of technical problems with the sensor. Water samples were taken from three to five depths, from the surface down to the chlorophyll maxima. Ammonium was measured on board with the fluorescence method (Holmes *et al.*, 1999). Another 50 ml of seawater subsample was immediately frozen for later, shore-based phosphate, nitrate and nitrite analyses using a nutrient autoanalyser (Alpha Laval, Traacs 800, Bran Luebbe GmbH, Norderstedt, Germany). DIC was measured according to the flow-injection technique (Hall and Aller, 1992). Dissolved organic carbon (DOC) and total nitrogen (TN) were measured using a total organic carbon analyzer (TOC-VCPH, Shimdazu, Duisburg, Germany) equipped with a TNM-1 Modul for TN analyses and an ASI-V autosampler. Dissolved organic nitrogen was calculated by subtracting total inorganic N (ammonium and nitrate) from TN. The samples were not filtered to avoid the contamination of C and N through filtering. Concentrations of particulate organic carbon (POC) and particulate organic nitrogen (PON) in oligotrophic regions are negligible. The analysis of DOC and TN was validated on a daily basis with consensus deep-sea reference material (CRM Program, <http://yyy.rsmas.miami.edu/groups/biogeochem/CRM.html>) provided by the University of Miami.

### *$^{15}N$ and $^{13}C$ incubation experiments*

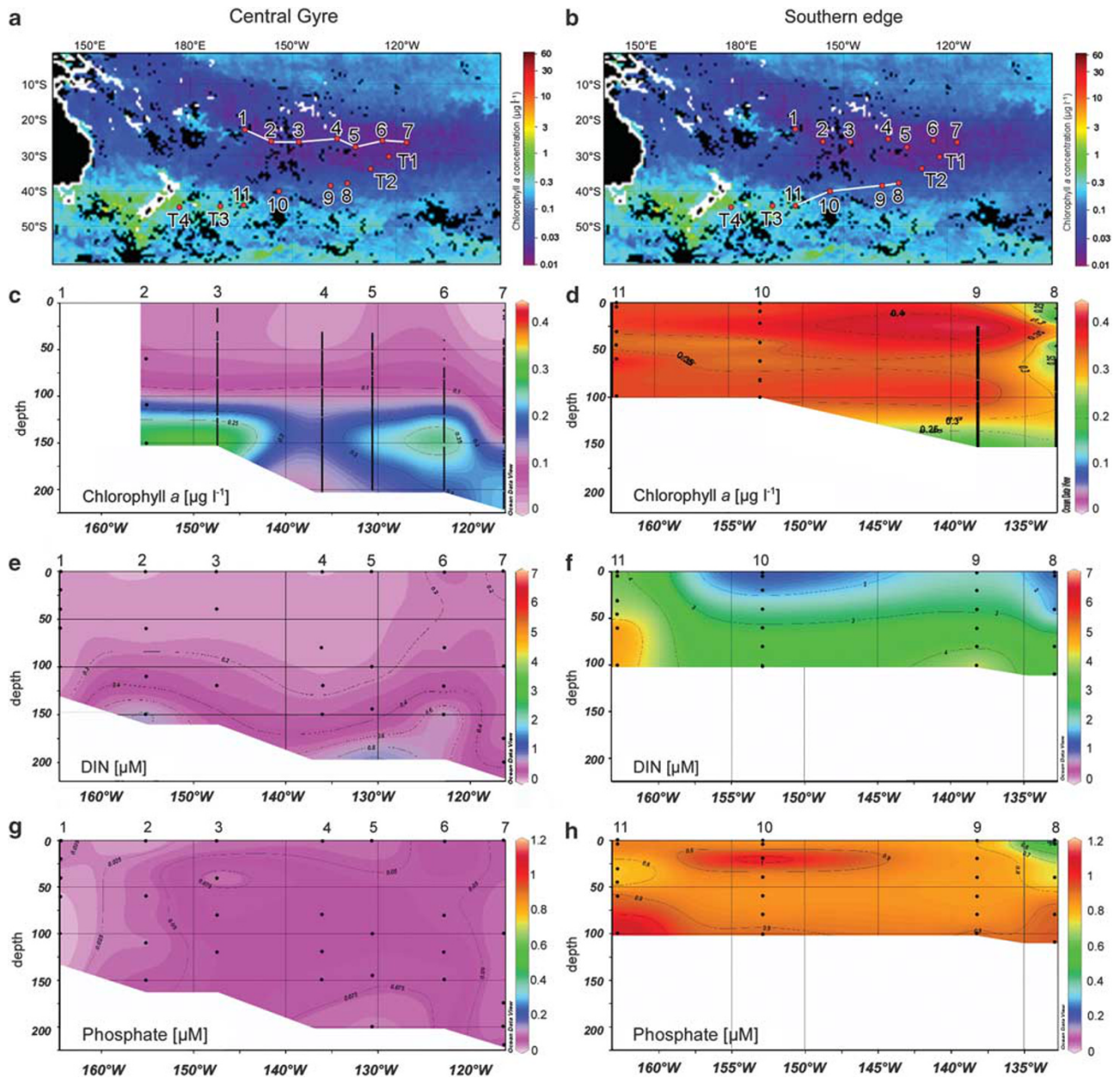
Nitrogen fixation and bicarbonate uptake rates were measured simultaneously by conducting incubation experiments with  $^{15}N$ -dinitrogen gas and  $DI^{13}C$  as described in the study by Montoya *et al.* (1996) with the following modifications. Seawater from each sampled depth was filled into 1 l gas-tight glass bottles without headspace. To each bottle, 4 ml of  $^{15}N_2$ -gas and  $^{13}C$ -sodium hydrogen carbonate (200  $\mu M$  end concentration) was added with gas-tight syringes.

Ammonium and leucine assimilation rates were determined in parallel on water samples from the surface and chl *a* maxima. Seawater subsamples were filled into 250 ml glass bottles in triplicate and incubated with either  $^{15}N$ -ammonium (2  $\mu M$  end concentration) or  $^{15}N$ -leucine (19.5  $\mu M$  end concentration), whereas  $^{13}C$ -sodium bicarbonate (200  $\mu M$  end concentration) was added to both experimental series at the same time.

All experimental samples were incubated in a water bath with surface water continuously flowing through on deck for 5–8 h. If the CTD was run during the evening, water subsamples for the incubations were stored at room temperature in the dark until the next morning. Bottles for the lower depths were covered using a black gauze material (mesh size = 1 mm<sup>2</sup>) to reduce light intensity. For each experimental series parallel dark incubations were also made by covering subsampled bottles with aluminium foil and placed inside a dark opaque plastic bag. Experiments were terminated by filtering the incubated water on to combusted glass-fiber filters (GF/F Millipore, Volketswil, Switzerland) and frozen at  $-20^\circ C$  until further analyses in a shore-based laboratory. For all stations, surface water was filtered for natural  $^{15}N$ - and  $^{13}C$ -stable isotopic background measurements of particulate organic carbon and nitrogen. In a shore-based laboratory, GF/F filters were freeze dried and acidified with hydrochloric acid (36%) in a dessicator to remove inorganic carbon in the samples. They were then introduced into a Thermo Flash EA 1112 elemental analyzer coupled to an isotopic ratio mass spectrometer (Thermo Delta Plus XP, Thermo Fisher Scientific, Schwerte, Germany) to determine C- and N-stable isotopic compositions. Carbon and nitrogen assimilation rates were then calculated as the increase in  $^{13}C/^{12}C$  and  $^{15}N/^{14}N$  isotopic ratios with time, and are presented here in units of  $nM h^{-1}$ . At station 11, experiments were performed for only one depth 2 days after sampling because of stormy weather.

### *Molecular ecological analyses*

To determine total microbial abundance, 15 or 30 ml of seawater subsample was fixed with 1% (final concentration) of paraformaldehyde, filtered on a polycarbonate membrane filter (0.2 mm pore size, GTTP, Millipore), and frozen at  $-20^\circ C$  until further processing. A section of each filter was stained with DAPI (1  $\mu M$ ) and cells were counted under an epifluorescence microscope (Axiophot, Zeiss, Oberkochen, Germany).



**Figure 1** Map showing the locations of the sampling stations for the northern transect in the center of the gyre (a) and at the southern edge (b), both superimposed on a SeaWiFS surface chlorophyll *a* composite image (first week of January 2007); and the corresponding vertical distribution of chlorophyll *a* concentrations (c, d), dissolved inorganic nitrogen (DIN) (e, f) and phosphate concentrations (g, h).

For DNA extraction, between 2.5 and 5 l of seawater subsamples was filtered through polycarbonate membrane filters (0.2 μm pore size, GTTP, Millipore). For RNA extraction, 10 l of seawater from the surface and chlorophyll maximum was filtered through a Sterivex filter (0.22 μm pore size; Millipore) and stabilized with RNAlater (Ambion, Life Technologies, Darmstadt, Germany). They were frozen at  $-80^{\circ}\text{C}$  until further processing in a shore-based laboratory.

DNA was extracted from stations 6, 7, 8 and 9 from the surface and chlorophyll maxima, and additionally from 40 m at station 9 and from 60 m at station 10, following the protocol by Zhou *et al.*

(1996). Thereafter, DNA was further purified using Wizard DNA Clean-Up System (Promega, Mannheim, Germany). Clone libraries were constructed for the *nifH* genes amplified from total DNA using *nifH* primers designed by Zani *et al.* (2000) (Supplementary Table S1) and the nested PCR approach as previously described in the study by Langlois *et al.* (2005), but with 45 and 40 cycles for the first and second PCRs, respectively.

RNA was extracted from surface water samples and chlorophyll maxima of stations 4, 6, 7, 8, 9 and 10, using the Totally RNA kit (Applied Biosystems, Darmstadt, Germany) with previous cell-lysis and proteolysis procedures within the filter cartridges as

described previously by Somerville *et al.* (1989) and modified in the study by Lam *et al.* (2007). To examine the expression of the nitrogenase gene *nifH*, first-strand complementary DNA (cDNA) was obtained from RNA extracts after reverse transcription using the SuperScript III First-Strand Synthesis SuperMix Kit (Invitrogen, Groningen, The Netherlands) with the gene-specific primer, *nifH3* (Zani *et al.*, 2000; Langlois *et al.*, 2008; Halm *et al.*, 2009), and the reaction was then treated with 2 Units RNaseH (Ambion). *nifH* cDNA fragments were subsequently amplified from the first-strand reaction (1–20 ng  $\mu\text{l}^{-1}$ ) with AmpliTaq-Gold polymerase (Applied Biosystems) and the primers and the nested PCR method as described previously (Langlois *et al.*, 2005), but with 40 and 35 cycles for the first and second PCRs, respectively.

For both DNA and cDNA, amplicons of the predicted 359 bp size were cloned using the TOPO TA cloning kit for sequencing (pCR4 vector, Invitrogen). Plasmids of positive clones were sequenced using the BigDye terminator kit 3.0 (Applied Biosystems) and an ABI PRISM 3130  $\times$  L genetic analyzer (Applied Biosystems). From all clone libraries, only sequences with a length of 359 bp and of good quality without chimera were used. The partial sequences were first aligned and compared using the BioEdit software (Ibis Biosciences, Carlsbad, CA, USA) (Hall, 1999), and grouped based on sequence similarity. A total of 198 sequences in total were then used for final amino acid-based alignments along with sequences obtained from public databases, and phylogenetic analyses were performed using the ARB software package (TU Munich, Munich, Germany) (Ludwig *et al.*, 2004). Phylogenetic *nifH* trees were constructed by performing neighbor-joining, maximum likelihood and maximum parsimony analyses based on amino acids and using different sets of filters to exclude the influence of highly variable positions, and without any filters. Topologies derived by these different approaches were compared with each other to construct a consensus tree. The *nifH* sequences were deposited in GenBank under accession numbers HM210309–HM210414.

#### Quantitative PCR

The TaqMan fluorogenic PCR method was used to quantify the *nifH* genes and the reverse-transcribed cDNA as described in the studies by Langlois *et al.* (2008) and Lam *et al.* (2007) with the following modifications. On the basis of *nifH* sequence information from the samples, primers and TaqMan probes were designed to target six prominent phylotypes using the software Allele ID 7.7 (Premier Biosoft, Palo Alto, CA, USA) (Supplementary Table S1). The specificities of their sequences were verified with our ARB database, and were also checked against the NCBI database with BLAST searches. Quantitative PCR (qPCR) standards for the different phylotypes were obtained by extracting

plasmids from the correct clones determined in the above phylogenetic analyses. Plasmid extraction and purification were performed using the Qiagen Plasmid Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. PCR products amplified from correct purified plasmids with specific primers for the different groups were taken as standards. Concentrations of standards were measured using a DNA 1000 kit run on a 2100 Bioanalyzer (Agilent Technologies, Böblingen, Germany). The newly designed qPCR primers and (FAM and BHQ dual-labeled) probe sets (Biomers, Ulm, Germany) were tested for specificity and optimal primer and probe concentrations, using the various representative plasmids as positive and negative controls. No unspecific amplification was observed.

The qPCRs for both DNA and cDNA with the newly developed primer and probe sets were run on a iQ5 Multicolor Real-Time PCR detection system (Bio-Rad, Munich, Germany). Each 25  $\mu\text{l}$  qPCR reaction contained 12.5  $\mu\text{l}$  of 2  $\times$  TaqMan Master Mix (Applied Biosystems), 0.5 pmol  $\mu\text{l}^{-1}$  (Group A) or 0.8 pmol  $\mu\text{l}^{-1}$  (all other phylotypes) each of the forward and reverse primers, 0.2 pmol  $\mu\text{l}^{-1}$  of TaqMan probe, 9.3 or 8.7  $\mu\text{l}$  PCR water, as well as 2  $\mu\text{l}$  of either standard or environmental sample. The optimal DNA dilution for qPCR was tested to be 1:25, which was used in subsequent qPCRs. For cDNA, 2  $\mu\text{l}$  of 2.5  $\times$  diluted reverse-transcribed first-strand reaction was used. The cycling parameters were 2 min at 50  $^{\circ}\text{C}$ , 10 min at 95  $^{\circ}\text{C}$ , followed by 40 cycles of 15 s at 95  $^{\circ}\text{C}$  and 1 min at 60  $^{\circ}\text{C}$ . All samples were run in triplicate together with no-template controls. Standards were serially diluted ( $10^7$ – $10^1$  copies) and also run in triplicate. The efficiency of all primers was between 78 and 99.8% for all analyses. The iQ Optical System Software (version 1.0, Bio-Rad) was used to analyze the qPCR data.

The qPCR for *Trichodesmium*, UCYN-A, *Crocospaera* (group B), *Cyanothece* (group C) and  $\gamma$ -proteobacteria from the Atlantic Ocean (AO) and cluster III, were performed with primer and probe sets from Langlois *et al.* (2008), and run with an ABI Prism 7000 (Applied Biosystems) as described previously (Langlois *et al.*, 2008), using 5  $\mu\text{l}$  samples or standards in each reaction. Standards were serially diluted ( $10^7$ – $10^1$  copies). All samples, standards and no-template controls were run in duplicate for each primer/probe set. No-template controls were undetectable after 40 cycles. The ABI 7000 system SDS software (version 1.2.3; with RQ study application) was used to analyze the qPCR data.

## Results and Discussion

### *Nutrient distributions and primary productivity in an ultra-oligotrophic gyre system*

The ultra-oligotrophic status of the SPG reaches its zenith in the northern transect that cuts through the

center of the gyre (Figure 1a). A weak thermocline and consequently a weak pycnocline (Supplementary Figures S1a and m) were present throughout the central gyre (stations 1–7) at water depths between 50 and 100 m. Chl *a* concentrations were close to the detection limit of the sensor in the upper 60 m of the water column (Figure 1c). Chl *a* maxima were observed below the pycnocline and along generally the same isopycnal surface ( $\sigma_t$  between 26 and 26.5 kg m<sup>-3</sup>; Supplementary Figure S1a). In terms of depth, these deep chlorophyll maxima were situated at a minimum of 150 m at station 2 ( $\leq 0.39 \mu\text{g chl } a \text{ per l}$ ), or as deep as 200 m at station 7 ( $\leq 0.28 \mu\text{g chl } a \text{ per l}$ ) (Figure 1c). Concentrations of dissolved inorganic nitrogen (DIN = ammonium + nitrite + nitrate) and phosphate were close to detection limits from the surface to 100 m, and increased slightly with depth to 0.8 and  $\leq 0.1 \mu\text{M}$ , respectively, towards the deep chlorophyll maxima (Figures 1e and g).

Along the southern transect, the pycnocline shoaled towards the south-west (Supplementary Figure S1b). At stations 8, 9 and 10, two chlorophyll maxima were measured, the upper one at 40–50 m and a lower maximum with slightly higher concentrations at 100 m depth (Figure 1d). Surface concentrations of DIN (1–6  $\mu\text{M}$ , mean  $\sim 3.1 \mu\text{M}$ ) and phosphate (0.84  $\mu\text{M}$ ) were considerably higher than those in the central gyre (Figures 1f and h). A notably lower phosphate concentration (0.11  $\mu\text{M}$ ) was measured near the surface at station 8, which coincided with a local minimum in nitrate, but a maximum in ammonium (Supplementary Figures S1f and h) and a minimum in surface chl *a*. Although nutrient concentrations were significantly higher along the southern edge of the SPG, this region is still considered highly oligotrophic compared with the rest of the world's oceans.

Low rates of primary production in the center of the gyre ( $\leq 2.6 \text{ nM C h}^{-1}$ ) were determined through non-stimulated uptake of <sup>13</sup>C-bicarbonate under light conditions (Figure 2a), and were consistent with the low abundance of phytoplankton as reflected by the vanishingly small chl *a* content ( $< 0.4 \mu\text{g Chl } a \text{ per l}$ ) of the water sampled. Primary productivity was  $\sim 10$ -fold higher at sites along the southern rim of the gyre (stations 8–11) than along the northern transect (Figure 2a). Although  $\sim 3$ -fold higher rates ( $\sim 6.9 \text{ nM C h}^{-1}$ ) were measured for the north-eastern part of the SPG in another study (Raimbault and Garcia, 2008), all rates reported thus far are very low relative to more mesotrophic oceanic regions like the equatorial west Pacific ( $120 \text{ nM C h}^{-1}$ ) (Raimbault and Garcia, 2008). In comparison, incubations amended with <sup>15</sup>NH<sub>4</sub><sup>+</sup> and <sup>15</sup>N-leucine significantly enhanced <sup>13</sup>C-bicarbonate uptake in the center of the gyre (up to  $10.2 \text{ nM C h}^{-1}$ ; *t*-test,  $P < 0.05$ ), and such enhancement was generally light dependent (Figure 2, Supplementary Figure S4), indicating that primary production in the center of the SPG was nitrogen limited. Meanwhile, the rates of all three incubations were

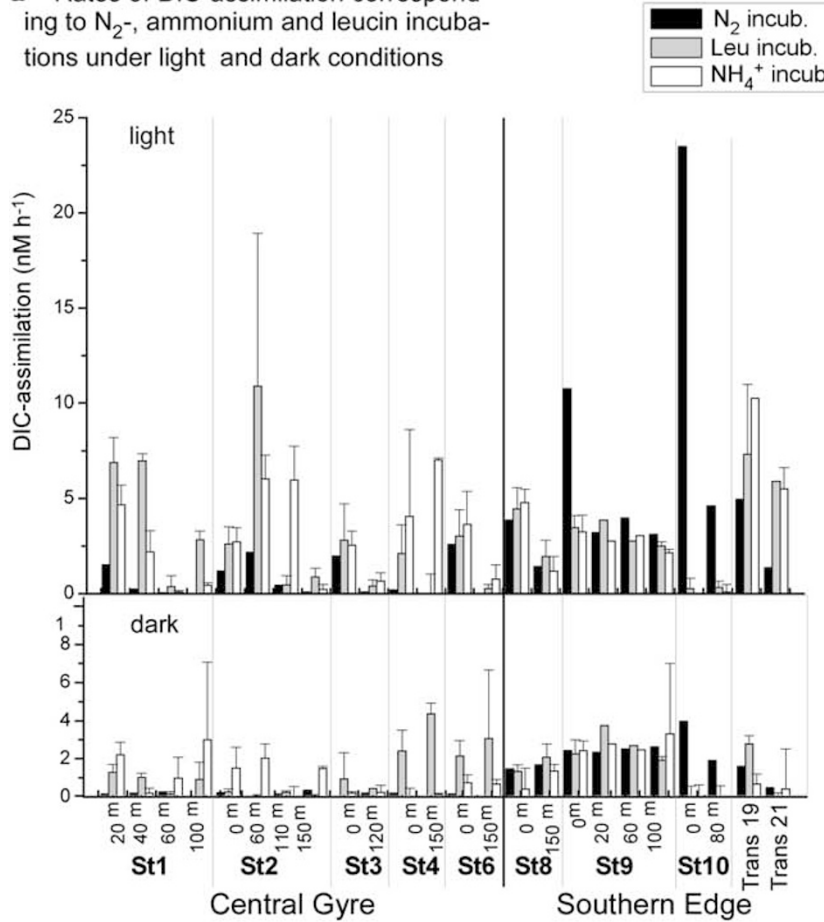
comparable along the southern transect (*t*-test,  $P > 0.05$ ) (Figure 2a), suggesting that primary production therein was more likely limited by factors other than nitrogen.

#### Nitrogen fixation in the SPG

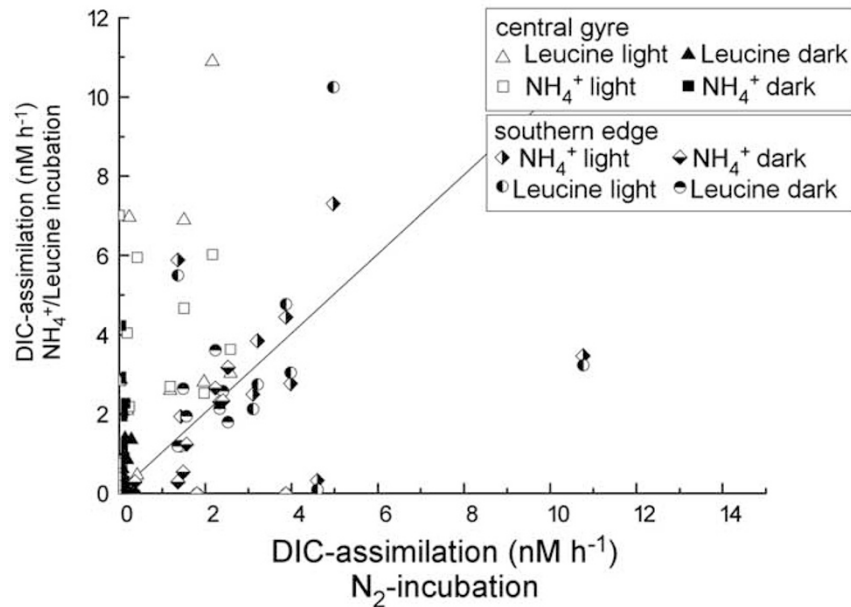
Substantial light and dark N<sub>2</sub>-fixation rates were detected at all but one station (station 11; Supplementary Figure S2, Supplementary Table S2). Within the central gyre, rates ranged between 0.001 and 0.04 nM N h<sup>-1</sup> (Supplementary Figure S2, Supplementary Table S2), and a very high N<sub>2</sub>-fixation rate of 0.19 nM N h<sup>-1</sup> was measured at 100 m at station 3. Apart from this one sample, the rates of the southern transect were comparable (0.012–0.035 nM N h<sup>-1</sup>). Over the entire span of sample locations and depths, no significant correlation was discernible between N<sub>2</sub>-fixation rates and chlorophyll concentrations or total microbial abundance. Intriguingly, similar <sup>15</sup>N<sub>2</sub>-fixation rates were measured in the parallel light and dark incubations at most stations, indicating that N<sub>2</sub>-fixation activity was most likely independent of light.

The mean daily N<sub>2</sub>-fixation rate (calculated as the sum of 12 h rates measured in the dark and in the light) at the surface of the central gyre stations (stations 2–7) was  $0.4 \pm 0.3 \text{ nM N day}^{-1}$ , whereas at the southern edge (stations 8–10), the daily rate was  $1.5 \pm 0.3 \text{ nM N day}^{-1}$ . Raimbault and Garcia (2008) found the highest N<sub>2</sub>-fixation rates ( $1.8 \text{ nM N day}^{-1}$ ) in the upper 50 m of the water column in the eastern and central SPG, but these rates declined rapidly below 50 m. In contrast, our measured nitrogen-fixation rates were similar throughout the surface mixed layer in the center of the gyre during December 2006/January 2007 (Figure 3, Supplementary Figure S3). The SPG usually reaches its maximum expansion and thus the minimum chl *a* concentrations in September, although the magnitude of seasonal variation is not as large as that in other subtropical gyres (McClain *et al.*, 2004; Morel *et al.*, 2007). Hence, the fact that N<sub>2</sub>-fixation was found to be important even during our sampling in the austral summer, strongly implies that N<sub>2</sub>-fixation is a significant source of fixed nitrogen to this ultra-oligotrophic system throughout the year. The N<sub>2</sub>-fixation rates in the center of the SPG (up to 0.19 nM N h<sup>-1</sup>) and at its southern edge (0.01–0.035 nM N h<sup>-1</sup>) were comparable with those rates generally obtained from the North Pacific Gyre (0.01–0.15 nM N h<sup>-1</sup>) (Montoya *et al.*, 2004)—except for the higher rates (0.36 nM N h<sup>-1</sup>) observed during a *Trichodesmium* bloom in an anti-cyclonic downwelling eddy (Fong *et al.*, 2008). The N<sub>2</sub>-fixation rates in the SPG were also substantially lower than those in the tropical Atlantic (up to 3.1 nM N h<sup>-1</sup>; Voss *et al.* (2004)). Diazotrophs have higher requirements for iron than non-diazotrophic primary producers. Hence, the low iron inputs to the SPG, as a result of the remoteness of the SPG to any continents, may ultimately limit N<sub>2</sub>-fixation

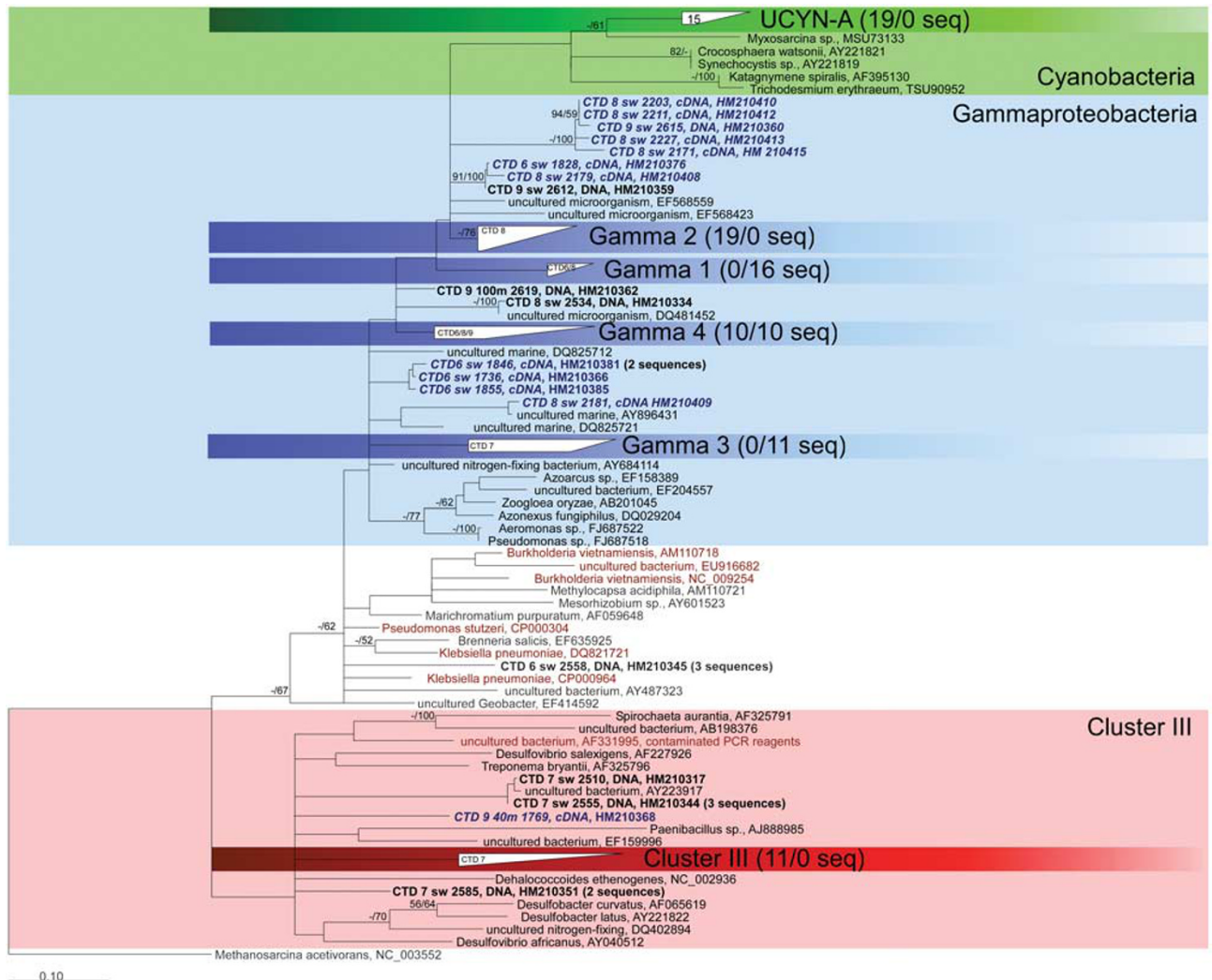
**a** Rates of DIC-assimilation corresponding to  $N_2$ -, ammonium and leucine incubations under light and dark conditions



**b** DIC-assimilation of  $N_2$  incubation vs. DIC-assimilation Leucine or ammonium incubations



**Figure 2** Assimilation of <sup>13</sup>C-labeled DIC in incubations with  $N_2$ ,  $NH_4^+$  and leucine under light and dark conditions (a); relationship between non-stimulated ( $N_2$  incubation) and stimulated ( $NH_4^+$  and leucine incubations) assimilation of DIC (b). The straight line indicates the 1:1 relationship should DIC assimilation be the same in stimulated and non-stimulated incubations.



**Figure 3** Phylogenetic reconstruction based on *NifH* amino-acid sequences translated from DNA/cDNA sequences retrieved in this study. Clusters I, III and IV have been defined by Zehr *et al.* (2003). Branching patterns that were not supported by all three treeing algorithms used are shown as multifurcations. Bootstrapped values (%) above 50, out of 100 resampling in maximum parsimony and maximum likelihood analyses are shown on branches. The scale bar represents 10% estimated sequence divergence. Numbers after group names indicate the numbers of DNA/cDNA sequences retrieved for the subgroup in this study. Sequences in brown indicate likely contaminated PCR products previously reported, but are mostly distant from the sequences retrieved in this study. The CTD numbers within highlighted clusters denote the stations from which sequences originated. Please see Supplementary Figure S5 for detailed views on individual subclusters.

(Wagener *et al.*, 2008). A recent study (Mohr *et al.*, 2010) showed that the method used here would likely have underestimated  $N_2$ -fixation rates in the SPG.

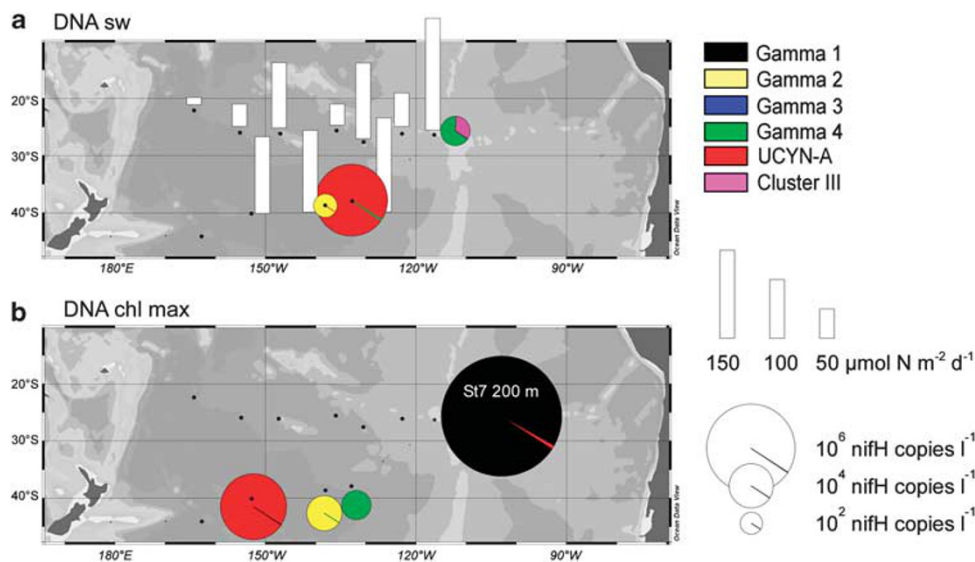
Absolute  $N_2$ -fixation rates in the SPG may appear low relative to other ocean gyre ecosystems, but measurable rates of  $N_2$ -fixation down to the depth of the chlorophyll maximum (up to 200 m) and the vast expanse of the SPG, substantially boost basin-wide estimates of  $N_2$ -fixation rates for the SPG. Depth-integrated rates (Table 1) were generally lower for the center of the gyre ( $12\text{--}190\ \mu\text{mol N m}^{-2}\ \text{day}^{-1}$ , mean  $93.8 \pm 61.3\ \mu\text{mol N m}^{-2}\ \text{day}^{-1}$ ) than for the southern transect ( $103\text{--}160\ \mu\text{mol N m}^{-2}\ \text{day}^{-1}$ , mean  $134.3 \pm 29\ \mu\text{mol N m}^{-2}\ \text{day}^{-1}$ ). These rates were comparable to rates previously reported for the central gyre ( $48\text{--}135\ \mu\text{mol N m}^{-2}\ \text{day}^{-1}$ ) (Moutin *et al.*, 2008), and were not significantly different

from the rates ( $30\text{--}60\ \mu\text{mol N m}^{-2}\ \text{day}^{-1}$ ) estimated by Raimbault and Garcia (2008), or those modeled for the South Pacific ( $55\ \mu\text{mol N m}^{-2}\ \text{day}^{-1}$ ) (Deutsch *et al.*, 2007). Similar rate estimates have also been given for the North Pacific (Zehr *et al.*, 2001; Montoya *et al.*, 2004). Assuming little seasonal variability, our crude estimates indicate that the SPG may account for up to  $15\ \text{Tg a}^{-1}$  of oceanic nitrogen input, equivalent to approximately 11–13% of the global pelagic  $N_2$ -fixation ( $117\text{--}135\ \text{Tg N a}^{-1}$  (Codispoti *et al.*, 2001; Gruber, 2004; Codispoti, 2007)) or approximately 5–6.5% of total new nitrogen input into the world's ocean ( $231\text{--}294\ \text{Tg N a}^{-1}$  (Codispoti *et al.*, 2001; Gruber, 2004)). If we take the methodical underestimation shown by Mohr *et al.* (2010) into consideration, this estimate of oceanic N-input in the SPG is likely to be even higher.

**Table 1** Integrated rates of nitrogen fixation and corresponding C-uptake for the euphotic zone at stations with more than one incubation depth

Station	Latitude	Longitude	Integrated rates		Range of uptake C:N values	Average of uptake C:N values
			$N_2$ fixation ( $\mu\text{mol m}^{-2} \text{ day}^{-1}$ )	C uptake ( $\mu\text{mol m}^{-2} \text{ day}^{-1}$ )		
1	23°51'S	165°39'W	12.4	645.1	3–205	86 ± 106
2	26°03'S	156°54'W	38.0	2186.2	29–196	105 ± 71
3	27°57'S	148°35'W	111.3	2069.0	0.3–175	75 ± 71
4	26°29'S	137°56'W	36.0	395.2	11–17	14 ± 3
5	28°27'S	131°23'W	129.0	2191.5	8.5–53	26 ± 20
6	27°55'S	123°10'W	57.6	2080.9	11–124	41 ± 60
7	27°44'S	117°37'W	190.9	2787.4	7–54	18 ± 20
8	38°04'S	133°06'W	160.3	6597.7	26–76	55 ± 25
9	39°19'S	139°48'W	139.6	8720.4	50–198	90 ± 71
10	41°51'S	153°06'W	103.0	11 059.9	80–264	176 ± 100

Rates were integrated from the surface to the depth of chlorophyll maximum (between 150 and 200 m in the center and 100 m at the southern transect), assuming 12 h each of day-time and night-time for the corresponding light- and dark- rates measured. C:N values were calculated by using uptake rates.



**Figure 4** Distribution of depth-integrated  $N_2$ -fixation rates and the corresponding diazotrophic community structure in the SPG. The sizes of white bars indicate the rates of nitrogen fixation integrated over the euphotic zone. The pies show the contribution of various diazotrophic groups to total *nifH* gene copies at the surface (a) and at the chlorophyll maximum (b), as determined by quantitative PCR and shown in a logarithmic scale.

#### Unique, heterotrophic, diazotrophic communities in the SPG?

Gene abundance and active transcription of the nitrogenase gene (*nifH*), the functional gene biomarker for  $N_2$ -fixation, detected in parallel samples, yields important insights into the community structure of the  $N_2$ -fixation in the SPG. Sequence analyses of the *nifH* genes and their expressed transcripts revealed a diazotrophic community consisting mainly of heterotrophic  $\gamma$ -proteobacteria that fell into the conventionally defined cluster I (Zehr *et al.*, 2003), and the unicellular cyanobacteria Group A (UCYN-A), along with few sequences affiliated with the so-called cluster III (Figure 3, Supplementary Table S3). Sequences related to the common photoautotrophic diazotrophs *Trichodesmium* and *Crocospaera* were not detected in any

of our clone libraries (Figure 4, Supplementary Figure S5). Their absence at both gene (DNA) and transcript (mRNA) levels was further confirmed by negative real-time PCR results using common primers for these groups (Langlois *et al.*, 2008). *NifH* genes of *Cyanothece* and green sulfur bacteria were also not detectable.

A total of 26 cyanobacterial UCYN-A sequences were obtained from station 7 through cDNA (as reversely transcribed from the transcripts mRNA) (Supplementary Table S3). They are up to 100% (94–100%) similar to sequences from the North Pacific Gyre and <90% similar to the closest cultured relatives *Synechococcus* and *Cyanothece* (Figure 4). UCYN-A was consistently found to be active in all samples examined in both the gyre center and the southern gyre edge (*nifH*



cDNA:  $0.4\text{--}8.4 \times 10^3$  copies per l), whereas high *nifH* gene abundance ( $\leq 2.56 \times 10^5$  *nifH* copies per l) was also detected (Table 2, Figure 4), far exceeding the only other report for the central SPG (184 gene copies per l) (Bonnet *et al.*, 2008). This phylogenetic group has previously been described for the North Pacific Gyre (Church *et al.*, 2005a,b; Zehr *et al.*, 2008), the Atlantic (Langlois *et al.*, 2005, 2008) and the Red Sea (Foster *et al.*, 2009), in comparable abundance as found in this study. These small unicellular cyanobacteria have been shown to express *nifH* mainly during the day (Zehr *et al.* 2007). A recent study showed that UCYN-A lack the entire photosystem II apparatus and are therefore unable to photosynthesize, rendering them most likely photoheterotrophs (Zehr *et al.*, 2008; Tripp *et al.*, 2010).

The vast majority of the obtained *nifH* gene DNA and cDNA sequences belonged to  $\gamma$ -proteobacteria, as seen in clone libraries (Supplementary Table S3), whereas the qPCR results showed Group A to be widely distributed, but dominant in numbers only at two stations (Figures 3 and 4, Table 2). Various groups of  $\gamma$ -proteobacteria make up higher numbers at other stations. Sequences of  $\gamma$ -proteobacteria fell into several subgroups, named Gamma 1–4. The subgroup Gamma 1 consisted of 31 cDNA sequences from the surface water at station 6, although one cDNA sequence from station 8 and one DNA sequence from station 9 might also be closely affiliated (Figure 3). Using specifically designed qPCR primers and probe, this subgroup was enumerated to be particularly abundant ( $7.7 \times 10^6$

gene copies per l) at the deep chlorophyll maximum right at the gyre center (station 7), with their expression also detectable (Table 2). Gamma 2 sequences were detected from surface water at station 8 (29 sequences; Figure 3, Supplementary Figure S5). By qPCR, they were also found to be abundant and active at stations 9 and 10 along the southern transect (Figure 3, Table 2), but their potential significance in the gyre center could not be fully assessed because of the high detection limit of this primer-probe set (Table 2). In all, 18 of the retrieved cDNA sequences belonged to Gamma 3, and they were obtained mainly from the surface water of station 7. This subgroup was detected as both DNA and cDNA throughout the gyre, but was not sufficiently high to be reliably quantified. The subgroup Gamma 4, which was related to *Vibrio* ( $\leq 95\%$  similarity) and a number of sequences from the Atlantic and Pacific, consisted of 13 DNA sequences and 10 cDNA sequences. They were relatively ubiquitous in the SPG compared with other  $\gamma$ -proteobacterial subgroups, although not necessarily as numerous (up to  $1.6 \times 10^3$  gene copies per l) (Table 2). In addition, primers and probes previously designed to target  $\gamma$ -proteobacterial *nifH* from the Atlantic Ocean revealed extremely high expression in the center of the gyre (station 6,  $3.25 \times 10^6$  cDNA per l; Table 2). These organisms are related to the subgroup Gamma 3 (94% similarity to CTD7 sw 1891, HM210395).

Finally, a group of 38 sequences belong to cluster III, with the closest *Dehalococcoides*

**Table 2** Abundance (mean  $\pm$  s.d. real-time PCR runs) of various phylotypes of nitrogenase gene, *nifH*, as gene copies (DNA) and their expressed transcripts quantified after reverse transcription (cDNA) at different sampling stations

Station	Depth (m)		Gamma 1 $10^3$ copies per l	Gamma 2 $10^3$ copies per l	Gamma 3 $10^3$ copies per l	Gamma 4 $10^3$ copies per l	AO $10^3$ copies per l	Cluster III $10^3$ copies per l	UCYN-A $10^3$ copies per l
4	150	cDNA	$0.8 \pm 0.0$		0				
6	0	DNA		+	0	0	0	0	0
		cDNA	+		0	0	$324.7 \pm 32.48$	0	$8.4 \pm 2.5$
6	150	DNA	0	+	0	*		0	0
		cDNA	0	+	+	+	0		
7	0	DNA	0	+	0	$0.61 \pm 0.2$	0	$0.3 \pm 0.1^*$	0
		cDNA			+	$0.05 \pm 0.00$	0	+	$3.5 \pm 0.5^*$
7	220	DNA	$7736 \pm 852$	+	0	+		0	$81.8 \pm 40.9$
	200	cDNA	$0.1 \pm 0.0$	+	0	0	0	$1.9 \pm 2.5$	$3.2 \pm 0.9$
8	0	DNA	0	+	0	$1.56 \pm 0.09^*$		0	$255.6 \pm 147.8$
		cDNA	0	+	0	+	0	$2.4 \pm 1.7$	$1.2 \pm 0.6$
8	110	DNA	0	+	+	$1.17 \pm 0.02$			
		cDNA			0	0	0		
9	0	DNA	0	$217 \pm 98$	0	0	0	0	$0.3 \pm 0.1$
		cDNA		+	+	+	0	+	
9	100	DNA	$1.2 \pm 0.5$	$1986 \pm 2350$	0	$1.15 \pm 0.35^*$	0		
		cDNA	0	$74 \pm 34$	+	$0.06 \pm 0.03^*$		0	$1.6 \pm 0.3$
10	0	cDNA	0	$148 \pm 112$	0	+	0		$0.4 \pm 0.0$
10	60	DNA	0		0	0	0	$0.3 \pm 0.0$	$144.7 \pm 51.3$
Detection limit			<0.100	<1.500	<0.100	<0.050	<0.010	<0.100	<0.100

Abbreviations: AO, gamma-proteobacterial group first detected in the Atlantic Ocean; cDNA, complementary DNA; UCYN-A, unicellular Group A cyanobacteria.

Asterisks (\*) denote clone library construction for the specified samples, and the corresponding phylotypes to which the subsequently retrieved sequences belong whereas '+' indicates detectable but not quantifiable phylotypes.

(82% similarity). This includes DNA sequences from the surface water at station 7 and a single cDNA sequence from 40 m depth at station 9. The expression of these *nifH* was confirmed for stations 7 and 8 (80–2400 *nifH* copies per l). Taken together, despite the fact that the relative importance of the various subgroups differed with respect to sampling sites and depths, in general higher *nifH* gene abundance of these (photo)heterotrophic diazotrophs seem to be associated with higher  $N_2$ -fixation rates (Supplementary Figure S3a), while no particular relationships can be discerned regarding *nifH* transcription (Supplementary Figure S3b).

Both sequence analyses and qPCR revealed a diazotrophic community dominated by a diverse cluster of  $\gamma$ -proteobacteria in the SPG, even exceeding the abundance of photoheterotrophic UCYN-A in a number of sites examined. The presence of  $\gamma$ -proteobacterial diazotrophs in oligotrophic oceans has been reported by other studies (Church *et al.*, 2005a; Zehr *et al.*, 2007; Langlois *et al.*, 2008; Foster *et al.*, 2009; Turk *et al.* 2011), yet photoautotrophic cyanobacteria usually predominated in those communities, whereas  $\gamma$ -proteobacteria were only of minor importance. A diazotrophic marine community dominated by highly diverse  $\gamma$ -proteobacteria has not yet been described to the best of our knowledge.  $\gamma$ -Proteobacteria are often found in the bacterioplankton of the photic zone, and they are often linked to the excretion of organic carbon from phototrophs (Zehr *et al.*, 2007; Moisaner *et al.*, 2008). Although we cannot fully discount the presence of *Trichodesmium* at other times of the year, as they are known to occur as episodic blooms in the North Pacific,  $N_2$ -fixation primarily driven by heterotrophic  $\gamma$ -proteobacteria and photoheterotrophic UCYN-A cyanobacteria throughout the SPG would explain our measured rates of  $N_2$ -fixation, especially those in the absence of detectable chlorophyll.  $N_2$ -fixation would therefore function independent of light level in the SPG.

#### Functioning of a heterotrophic diazotrophic community

Results from additional incubation experiments with  $^{15}NH_4^-$  and  $^{13}C$ -bicarbonate further indicated that photoautotrophic diazotrophs were unlikely to be the dominant primary producers in the SPG. Ammonium assimilation rates ( $0.4$ – $1.7$   $nM\ h^{-1}$ ) were detected to be the highest in surface water, and were generally higher under light conditions. These increases in ammonium assimilation were coupled with enhanced DIC uptake under light conditions (Figure 2b, Supplementary Figure S4), suggesting that the main photoautotrophs in these waters relied more directly on external ammonium sources rather than performing  $N_2$ -fixation themselves, and were thus likely nitrogen limited. These observations are consistent with the absence of detectable photoautotrophic diazotrophs through our qualitative and

quantitative *nifH* analyses. In agreement, the measured  $N_2$ -fixation rates were similar under light and dark conditions, whereas autotrophic  $N_2$ -fixing cyanobacteria are known to exhibit distinct diurnal patterns of  $N_2$ -fixation activity (*Trichodesmium* in the daytime or *Crocospaera* at night) (Bergman *et al.*, 1997; Zehr *et al.*, 2007)—unless both day-time and night-time phylotypes were present with comparable although temporally alternate activities. However, this latter hypothesis was found unlikely based on the *nifH* analyses.

A diazotrophic community dominated by photoheterotrophic or heterotrophic bacteria would mean that an external source of organic carbon is required, such as provided by primary producers in the same water. Heterotrophic and photoheterotrophic nitrogen-fixing organisms likely take up organic carbon in the form of small molecules (Kirchman *et al.*, 2000), which may not be limiting for heterotrophs in the SPG. In the gyre center, we measured  $40$ – $60$   $\mu M$  of DOC (Supplementary Table S5) and up to  $7$   $\mu M$  dissolved organic nitrogen (Supplementary Figures S1i and j). Phytoplankton in oligotrophic areas tend to produce more DOC than they can use, because of the limitation of mineral nutrients (Biersmith and Benner, 1998). Although chl *a* concentrations mostly bordered the detection limit above the deep chlorophyll maximum in the gyre center, very low levels of photoautotrophic production remained detectable in these waters as shown by the almost-strictly light-dependent DIC assimilation (Supplementary Figure S2). Owing to the strong vertical stratification and the subsequent lack of ventilation within subtropical gyres, DOC accumulates in these waters to attain some of the highest concentrations in the world's surface ocean (Hansell *et al.*, 2004). Taken together, phytoplankton overproduction and accumulation of DOC could probably explain the prevalence of diverse heterotrophic and photoheterotrophic  $N_2$ -fixing bacteria that are not limited in nitrogen but can use the abundant DOC, as long as they are not limited by other factors. Phytoplankton in the SPG, although present only in low abundance, might indirectly benefit from fixed nitrogen excreted by these  $N_2$ -fixing heterotrophs and nitrogen released upon their cell lyses. Further characterization of the DOC, especially on the lability, would help confirm this possibility. Alternatively, some phytoplankton might directly feed upon these heterotrophic diazotrophs to obtain limiting nutrients, such as nitrogen. Recently, high bacterivory by the smallest phytoplankton has been shown for the North Atlantic Ocean (Zubkov and Tarran, 2008). At the southern edge of the gyre, the slightly higher nutrient availability seems to have also stimulated lithoautotrophic growth more than photosynthesis alone, as shown by the substantial DIC assimilation in the dark. These potential chemolithoautotrophs may include archaeal ammonia oxidizers, which have been found even in the euphotic zone of the North Pacific Subtropical Gyre (Church *et al.*, 2009).

Like photoautotrophs, part of carbon fixed by chemolithoautotrophs may then be channeled to the (photo)heterotrophic diazotrophs, whereas the former could benefit from the subsequently fixed nitrogen in return. Alternatively, there could also be some dark CO<sub>2</sub> fixation due to carboxylation or other anaplerotic reactions in growing autotrophic and heterotrophic microorganisms (Werkman and Wood, 1942; Dijkhuizen and Harder, 1985).

#### *Is the SPG ecosystem different from the other ocean gyres?*

The dominance of heterotrophic diazotrophs suggests that the SPG microbial ecosystem is distinctly different from previously described gyre systems, in which photoautotrophic cyanobacteria are generally known to be the principal N<sub>2</sub> fixers (Sohm *et al.*, 2011). We hypothesize a close and probably mutualistic link between (photo)heterotrophic nitrogen fixation and the release of DOC by phytoplankton within the SPG. Heterotrophic and photoheterotrophic diazotrophs acquire new nitrogen from the atmosphere and use the excess DOC released by phytoplankton, the nitrogen limitation of which is in turn lessened by the new nitrogen inputs through (photo)heterotrophic diazotrophs. Such a community may be representative of the ultra-oligotrophic communities that occur in remote regions of large ocean gyres exemplified by the SPG. In spite of its vast size, the SPG has scarcely been investigated with modern biogeochemical and molecular ecological tools. These ultra-oligotrophic conditions might have been more prevalent in ancient oceans, at a time when open-ocean areas and the cohering gyres were much larger than at present (Scortese and Golonka, 1992).

Some studies suggest that the oligotrophic nature of ecosystems associated with the subtropical gyres are currently intensifying, as productivity shifts landward towards continental margins (Polovina *et al.*, 2008). Conversely, anthropogenic increases in atmospheric deposition of nitrogen (Duce *et al.*, 2008) may alleviate the ultra-oligotrophic conditions of the SPG and shift the heterotrophic diazotrophic community towards an autotrophic one dominated by, for example, *Trichodesmium*, such as in the North Pacific. Anthropogenic atmospheric nitrogen deposition has already affected other gyre systems like the Atlantic (Duce *et al.*, 2008), which might also have been characterized by heterotrophic diazotrophs in the not so distant past. The SPG may serve as one of the last, and perhaps disappearing, oceanic regions to explore the function of ultra-oligotrophic microbial communities.

#### Acknowledgements

We sincerely thank the captain, crew and scientific party of the *R/V Roger Revelle Expedition Knox02RR*, in particular, Jim Dorrance and David C Smith for assistance

with sampling. Daniela Franzke, Thomas Max, Michael Seidel, Matthias Friebe, Sarah Sokoll and Gabriele Klockgether are thanked for laboratory and technical support. This study was financed by the US National Science Foundation (grant no. OCE-0527167), the DFG-supported project SFB754 (<http://www.sfb754.de>) and the Max-Planck Society.

#### References

- Bergman B, Gallon JR, Rai AN, Stal LJ. (1997). N<sub>2</sub> fixation by non-heterocystous cyanobacteria. *FEMS Microbiol Rev* **19**: 139–185.
- Biersmith A, Benner R. (1998). Carbohydrates in phytoplankton and freshly produced dissolved organic matter. *Marine Chem* **63**: 131–144.
- Bonnet S, Guieu C, Bruyant F, Prasil O, Van Wambeke F, Raimbault P *et al.* (2008). Nutrient limitation of primary productivity in the Southeast Pacific (BIO-SOPE cruise). *Biogeosciences* **5**: 215–225.
- 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, Scott 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, Wai B, Karl DM, DeLong EF. (2009). Abundances of crenarchaeal *amoA* genes and transcripts in the Pacific Ocean. *Environ Microbiol* **12**: 679–688.
- Codispoti L, Brandes JA, Christensen JP, Devol AH, Naqvi S, Pearl HW *et al.* (2001). The oceanic fixed nitrogen and nitrous oxide budgets: moving targets as we enter the anthropocene? *Scientia Marina* **65**: 85–105.
- Codispoti LA. (2007). An oceanic fixed nitrogen sink exceeding 400 Tg N a<sup>-1</sup> vs the concept of homeostasis in the fixed-nitrogen inventory. *Biogeosciences* **4**: 233–253.
- 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.
- Dijkhuizen L, Harder W. (1985). Microbial metabolism of carbon dioxide. In: Dalton H (eds). *Comprehensive Biotechnology*, vol. 1. Pergamon Press: Oxford, pp 409–423.
- Duce RA, LaRoche J, Altieri K, Arrigo KR, Baker AR, Capone DG *et al.* (2008). Impacts of atmospheric anthropogenic nitrogen on the open ocean. *Science* **320**: 893–897.
- 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, Paytan A, Zehr JP. (2009). Seasonality of N<sub>2</sub> fixation and *nifH* gene diversity in the Gulf of Aqaba (Red Sea). *Limnol Oceanogr* **54**: 219–233.
- Gruber N. (2004). The dynamics of the marine nitrogen cycle and its influence on atmospheric CO<sub>2</sub> variations. In: Follows M, Oguz T (eds). *The Ocean Carbon Cycle and Climate, NATO ASI Series*. Kluwer Academic: Dordrecht, pp 97–148.
- Hall PO, Aller RC. (1992). Rapid, small-volume, flow-injection analysis for sigma-CO<sub>2</sub> and NH<sub>4</sub><sup>+</sup> in marine and fresh-waters. *Limnol Oceanogr* **37**: 1113–1119.

- 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.
- Halm H, Musat N, Lam P, Langlois RJ, Musat F, Peduzzi S *et al.* (2009). Co-occurrence of denitrification and nitrogen fixation in a meromictic lake, Lake Cadagno (Switzerland). *Environ Microbiol* **11**: 1945–1956.
- Hansell DA, Bates NR, Olson DB. (2004). Excess nitrate and nitrogen fixation in the North Atlantic Ocean. *Marine Chem* **84**: 243–265.
- Holmes RM, Aminot A, K erouel R, Hooker BA, Peterson BJ. (1999). A simple and precise method for measuring ammonium in marine and freshwater ecosystems. *Can J Fish Aquat Sci* **56**: 1801–1808.
- Karl D, Michales A, Bergman B, Capone D, Carpenter E, Letelier R *et al.* (2002). Dinitrogen fixation in the world’s oceans. *Biogeochemistry* **57/58**: 47–98.
- Kirchman DL, Meon B, Cottrell MT, Hutchins DA, Weeks D, Bruland KW. (2000). Carbon versus iron limitation of bacterial growth in the California upwelling regime. *Limnol Oceanogr* **45**: 1681–1688.
- Lam P, Jensen MM, Lavik G, McGinnis DF, M uller B, Schubert CJ *et al.* (2007). Linking crenarchaeal and bacterial nitrification to anammox in the Black Sea. *Proc Natl Acad Sci USA* **104**: 7104–7109.
- Langlois RJ, H ummer D, LaRoche J. (2008). Abundance and distributions of the dominant *nifH* phylotypes in the Northern Atlantic Ocean. *Appl Environ Microbiol* **74**: 1922–1931.
- Langlois RJ, LaRoche J, Raab PA. (2005). Diazotrophic diversity and distribution in the tropical and subtropical Atlantic Ocean. *Appl Environ Microbiol* **71**: 7910–7919.
- Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar *et al.* (2004). ARB: a software environment for sequence data. *Nucleic Acids Res* **32**: 1363–1371.
- McClain CR, Signorini SR, Christian JR. (2004). Subtropical gyre variability observed by ocean-color satellites. *Deep-Sea Research Part II-Topical Studies in Oceanography* **51**: 281–301.
- Mohr W, Gro kopf T, Wallace DWR, LaRoche J. (2010). Methodical underestimation of oceanic nitrogen fixation rates. *Plos One* **5**: e12583.
- 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**: 996–996.
- Montoya JP, Voss V, K ahler P, Capone DG. (1996). A simple, high-precision, high-sensitivity tracer assay for N<sub>2</sub> fixation. *Appl Environ Microbiol* **62**: 968–993.
- Montoya JP, Holl CM, Zehr JP, Hansen A, Villareal TA, Capone DG. (2004). High rates of N<sub>2</sub> fixation by unicellular diazotrophs in the oligotrophic Pacific Ocean. *Nature* **430**: 1027–1031.
- Morel A, Gentili B, Claustre H, Babin M, Bricaud A, Ras A *et al.* (2007). Optical properties of the “clearest” natural waters. *Limnol Oceanogr* **52**: 217–229.
- Moutin T, Karl DM, Dunhamel S, Rimmelin P, Raimbault P, VanMoory BAS *et al.* (2008). Phosphate availability and the ultimate control of new nitrogen input by nitrogen fixation in the tropical Pacific Ocean. *Biogeochemistry* **5**: 95–109.
- Polovina JJ, Howell EA, Abecassis M. (2008). Ocean’s least productive waters are expanding. *Geophys Res Lett* **35**: 2–6.
- Raimbault P, Garcia N. (2008). Evidence for efficient regenerated production and dinitrogen fixation in nitrogen-deficient waters of the South Pacific Ocean: impact on new and export production estimates. *Biogeochemistry* **5**: 323–338.
- Scortese CR, Golonka J. (1992). *Paleogeographic Atlas*. University of Texas at Arlington: USA.
- Sohm JA, Webb EA, Capone DG. (2011). Emerging patterns of marine nitrogen fixation. *Nat Rev Microbiol* **9**: 499–508.
- Somerville CC, Knight IT, Straube WL, Colwell RR. (1989). Simple, rapid method for direct isolation of nucleic acids from aquatic environments. *Appl Environ Microbiol* **55**: 548–554.
- Tripp HJ, Bench SR, Turk KA, Foster RA, Desany BA, Niazi F, *et al.*, Zehr JP (2010). Metabolic streamlining in an open-ocean nitrogen-fixing cyanobacterium. *Nature* **464**: 90–94.
- Turk KA, Rees AP, Zehr JP, Pereira N, Swif P, Shelley R *et al.* (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 M, Peeken I. (2004). Patterns of nitrogen fixation along 10°N in the tropical Atlantic. *Geophys Res Lett* **31**: doi:10.1029/2004GL020127.
- Werkman CH, Wood HG. (1942). Heterotrophic assimilation of carbon dioxide. *Adv Enzymol* **2**: 135–182.
- Wagener T, Guieu C, Losno R, Bonnet S, Mahowald N. (2008). Revisiting atmospheric dust export to the Southern Hemisphere Ocean: biogeochemical implications. *Global Biogeochem Cycles* **22**: doi:10.1029/2007GB002984.
- 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, Bench SR, Carter BJ, Hewson I, Niazi F, Shi T *et al.* (2008). Globally distributed uncultivated oceanic N<sub>2</sub>-fixing Cyanobacteria lack oxygenic photosystem II. *Science* **322**: 1110–1112.
- 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, Mondragon E, Short CM *et al.* (2007). Experiments linking nitrogenase gene expression to nitrogen fixation in the North Pacific subtropical gyre. *Limnol Oceanogr* **52**: 169–183.
- Zehr JP, Waterbury JB, Turner PJ, Montoya JP, Omoregie E, Steward GF *et al.* (2001). Unicellular cyanobacteria fix N<sub>2</sub> in the subtropical North Pacific Ocean. *Nature* **412**: 635–638.
- Zhou J, Bruns MA, Tiedje JM. (1996). DNA recovery from soils of diverse composition. *Appl Environ Microbiol* **62**: 316–322.
- Zubkov MV, Tarran GA. (2008). High bacterivory by the smallest phytoplankton in the North Atlantic Ocean. *Nature* **455**: 224–248.

Supplementary Information accompanies the paper on The ISME Journal website (<http://www.nature.com/ismej>)