Diazotroph community succession during the VAHINE mesocosm experiment (New Caledonia lagoon)

K. A. Turk-Kubo¹, I. E. Frank¹, M. E. Hogan¹, A. Desnues², S. Bonnet², and J. P. Zehr¹

¹Ocean Sciences Department, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA, 95064, USA
²Mediterranean Institute of Oceanography (MIO) – IRD/CNRS/Aix-Marseille University, IRD Noumea, 101 Promenade R. Laroque, BPA5, 98848 Noumea Cedex, France

Correspondence to: K. A. Turk-Kubo (kturk@ucsc.edu)

Received: 4 March 2015 – Published in Biogeosciences Discuss.: 17 June 2015
Revised: 23 November 2015 – Accepted: 27 November 2015 – Published: 17 December 2015

Abstract. The VAHINE mesocosm experiment, conducted in the low-nutrient low-chlorophyll waters of the Noumea lagoon (coastal New Caledonia) was designed to trace the incorporation of nitrogen (N) fixed by diazotrophs into the food web, using large volume (50 m³) mesocosms. This experiment provided a unique opportunity to study the succession of different N₂-fixing microorganisms (diazotrophs) and calculate in situ net growth and mortality rates in response to fertilization with dissolved inorganic phosphate (DIP) over a 23-day period, using quantitative polymerase chain reaction (qPCR) assays targeting widely distributed marine diazotroph lineages. Inside the mesocosms, the most abundant diazotroph was the heterocyst-forming Richelia associated with Rhizosolenia (Het-1) in the first half of the experiment, while unicellular cyanobacterial Group C (UCYN-C) became abundant during the second half of the experiment. Decreasing DIP concentrations following the fertilization event and increasing temperatures were significantly correlated with increasing abundances of UCYN-C. Maximum net growth rates for UCYN-C were calculated to range between 1.23 ± 0.07 and 2.16 ± 0.07 d⁻¹ in the mesocosms, which are among the highest growth rates reported for diazotrophs. Outside the mesocosms in the New Caledonia lagoon, UCYN-C abundances remained low, despite increasing temperatures, suggesting that the microbial community response to the DIP fertilization created conditions favorable for UCYN-C growth inside the mesocosms. Diazotroph community composition analysis using PCR targeting a component of the nitrogenase gene (nifH) verified that diazotrophs targeted in qPCR assays were collectively among the major lineages in the lagoon and mesocosm samples, with the exception of Crocosphaera-like phylotypes, where sequence types not typically seen in the oligotrophic ocean grew in the mesocosms. Maximum net growth and mortality rates for nine diazotroph phylotypes throughout the 23-day experiment were variable between mesocosms, and repeated fluctuations between periods of net growth and mortality were commonly observed. The field population of diazotrophs in the New Caledonian lagoon waters appeared to be dominated by Het-1 over the course of the study period. However, results from both qPCR and PCR analysis indicated a diverse field population of diazotrophs was present in the lagoon at the time of sampling. Two ecotypes of the Braarudosphaera bigelowii symbiont unicellular group A (UCYN-A) were present simultaneously in the lagoon, with the recently described B. bigelowii/UCYN-A2 association present at higher abundances than the B. bigelowii/UCYN-A1 association.

1 Introduction

Biological nitrogen fixation (BNF), the microbially mediated conversion of dinitrogen (N₂) gas into bioavailable nitrogen (N), is a significant source of new N in oligotrophic oceanic regions where primary productivity is N limited (Gruber and Sarmiento, 1997; Karl et al., 1997), and has the potential to directly impact carbon sequestration (Karl et al., 2012; Karl and Letelier, 2008). BNF has historically been considered an important process in the oligotrophic ocean gyres, but primary productivity in oligotrophic tropical coastal regions can also be N limited (Torréton et al., 2010) and such environ-
mensions have the potential to play a significant role in export production in the world’s oceans due to the transfer of carbon from tidal and wind-generated currents (Hyndes et al., 2014; Gattuso et al., 1998).

The New Caledonian (Noumea) coral lagoon, located off the southwestern coast of New Caledonia (southwestern Pacific), is a tropical low-nutrient low-chlorophyll (LNLC) system and is bounded by one of the world’s largest barrier reefs. Oligotrophic ocean water enters the lagoon from the south over the open shelf, then is driven north by the trade winds and tidal forces and exits through several deep inlets in the intertidal barrier reef that forms the western boundary of the lagoon (Ouillon et al., 2010). Primary productivity is N limited throughout the year (Torréton et al., 2010), giving microorganisms able to fix N2 gas into bioavailable nitrogen (diazotrophs) a competitive edge over non-diazotrophic organisms. High rates of BNF during the austral summer have been reported, in both large size fractions (>10 µm) and small size fractions (<10 µm; Garcia et al., 2007; Biegala and Raimbault, 2008). Garcia et al. (2007) also reported that the percent of N2 fixation measured in the large size fraction had high temporal variability. Large blooms of the most conspicuous and well-studied diazotroph, *Trichodesmium*, have been repeatedly detected in this region using both indirect (via satellite observation; Dupouy et al., 2000, 2011) and direct measurements (Renaud et al., 2005; Masotti et al., 2007; Rodier and Le Borgne, 2008, 2010). Both free-living and organism- and particle-associated unicellular picocyanobacterial diazotrophs are also suspected to be significant contributors to BNF in the lagoon (Garcia et al., 2007; Biegala and Raimbault, 2008), yet the phylogenetic identity of these picocyanobacteria has yet to be determined. Despite evidence that diverse diazotroph communities exist in the Noumea lagoon, there is very little quantitative diazotroph distribution data from this region (Luo et al., 2012), especially for diazotrophs other than *Trichodesmium*.

Marine unicellular cyanobacterial diazotrophs are phylogenetically divided into three groups: the uncultivated unicellular group A (UCYN-A; Zehr et al., 2008, 2001; Tripp et al., 2010), which live in association with strains of the prymnesiophyte *Braarudosphaera bigelowii* (Thompson et al., 2014, 2012; Hagino et al., 2013); the free-living *Crocosphaera* sp. (also referred to as unicellular group B or UCYN-B); and the presumably free-living unicellular group C (UCYN-C), which contains several cultivated cyanobacteria including *Cyanothecce* sp. strain ATCC 51142 (Reddy et al., 1993) and group C TW3 (Taniuchi et al., 2012).

The marine filamentous cyanobacterial diazotrophs include the colonial, non-heterocyst-forming *Trichodesmium*, and the heterocyst-forming symbionts associated with diatoms (DDAs). DDAs form between different strains of *Richelia* sp. associated with diatoms of the genera *Rhizosolenia* (Het-1) and *Hemiaulus* (Het-2; Villareal, 1990, 1992). Het-2 is in an obligate symbiont of *Hemiaulus*, as evidenced by its reduced genome size (Hilton et al., 2013). Although the genome of Het-1 shows evidence of some genome reduction, it remains unclear whether Het-1 is also an obligate symbiont of *Rhizosolenia* (Villareal, 1992; Hilton et al., 2013). The heterocyst-forming *Calothrix* (Het-3) has long been observed living as an epiphyte with *Chaetoceros* (Carpenter and Foster, 2002) but can also grow free from its host (Foster et al., 2010), and has a non-streamlined genome (Hilton et al., 2013).

Although *nifH* genes are regularly recovered from diverse non-cyanobacterial diazotrophs in oligotrophic ocean waters (Halm et al., 2012; Farnelid et al., 2011; Riemann et al., 2010; Langlois et al., 2005; Hewson et al., 2007; Bird and Wyman, 2012; Bonnet et al., 2013; Fong et al., 2008; Turk-Kubo et al., 2014), their activity and relative significance to BNF remains poorly understood (Turk-Kubo et al., 2014). The most widely studied non-cyanobacterial diazotroph, γ-24774A11, is an uncultivated putative gamma proteobacteria most closely related to *Pseudomonas stutzeri* that has been hypothesized to be a potentially important contributor to overall BNF in the South Pacific (Moisander et al., 2014).

Different diazotrophs have potentially different fates in the marine environment. For example, *Trichodesmium* is rarely recovered in sediment traps (Walsby, 1992), but *Trichodesmium*-derived N is efficiently transferred to non-diazotrophic plankton (mainly diatoms and bacteria) at short time scales (48 h) in the surface ocean (Bonnet et al., 2015b). In contrast, blooms of DDAs fuel an important summer export flux at the ALOHA station (Karl et al., 2012). This highlights the importance of characterizing the diazotroph community composition when performing biogeochemical studies on the fate of N2 fixation in the ocean. Such studies are commonly performed on oceanographic cruises, where discrete samples are taken from multiple stations along a cruise track, which passes over many different water masses. This approach is of critical importance to describe the biog graphical distribution of diazotrophs with respect to environmental parameters over large oceanic provinces. However, in order to understand how diazotroph assemblages shift in response to rapid environmental perturbations inside the same water mass, and to track the incorporation of their newly fixed N into the food web, high-frequency sampling of a single water body is required, which is rarely accomplished.

We report here data from the VAHINE mesocosm experiment, detailed by Bonnet (2015), which was a large, multi-institute collaborative project conducted to determine which components of the food web were directly supported by newly fixed N. To answer this question, three large (50 m3) mesocosms were deployed in the New Caledonian (Noumea) lagoon to isolate a part of the water column from physical dispersion without disturbing light penetration and temperature. The same water masses were monitored for 23 days during austral summer conditions. In order to create conditions favorable for diazotrophs, the mesocosms were fertilized with dissolved inorganic phosphorus (DIP) on day 4 of the incubation. This experiment provided unique oppor-
tunities to (1) track rapid diazotroph assemblage shifts using quantitative techniques (quantitative PCR; qPCR) targeting known major marine diazotroph lineages for a long period of time (23 days) in a single water mass; (2) calculate in situ net growth and mortality rates for targeted diazotroph phylotypes in a complex community; (3) determine the abundances of targeted diazotrophs in a coastal LNLC environment, the Noumea lagoon, during the experimental period; and (4) characterize shifts in the diazotroph community composition in both the mesocosms and in the Noumea lagoon during the experimental period using next-generation sequencing of nifH gene amplicons.

2 Methods

2.1 Sampling

Three large volume mesocosms (50 m$^3$), based on the design described in Guieu et al. (2010, 2014), were deployed at the exit of the Noumea lagoon ($22^{\circ}29.073 \text{S}–166^{\circ}26.905 \text{E}$), 28 km off the coast of New Caledonia, on 13 January 2013. Detailed descriptions of the mesocosm site selection, deployment and sampling strategy are provided in Bonnet (2015). Large volume samples (50 L) were collected from 1, 6, and 12 m depths from each mesocosm and outside the mesocosms (hereafter called the Noumea lagoon) once per day at 07:00 using a Teflon® PFA pump and PVC tubing. This daily sample was immediately transferred back to laboratories aboard the R/V Alis, and subsampled for a suite of stock and flux measurements. Samples for DNA analysis were immediately filtered onto 25 mm 0.2 µm Supor® filters (Millipore, Billarica, CA), using gentle peristaltic pumping. All filters were flash frozen in liquid nitrogen, and stored at −80°C until shipment on dry ice from New Caledonia to the University of California, Santa Cruz.

2.2 Determining targeted diazotroph abundances and net growth rates using the quantitative polymerase chain reaction (qPCR)

DNA was extracted using a Qiagen DNeasy Plant kit (Valencia, CA), with modifications to the protocol optimized to recover high-quality DNA from cyanobacteria, including additional cell lysis steps of freeze–thaw cycles, agitation using a bead beater, as well as a proteinase K digestion (Moisander et al., 2008). The quality of DNA extracts was evaluated using a PicoGreen® fluorescence quantitation kit (Molecular Probes, Eugene, OR, USA), according to the manufacturer’s guidelines.

Nine diazotrophic phylotypes were quantified in samples from the mesocosms and the Noumea lagoon, using quantitative PCR (qPCR), using Taqman® assays widely used in the marine environment. This study targeted representatives from known major marine diazotroph lineages: two unicellular diazotrophic symbionts of different *Braarudosphaera bigelovii* strains, UCYN-A1 (Church et al., 2005a), UCYN-A2 (Thompson et al., 2014); two free-living unicellular diazotroph cyanobacterial phylotypes UCYN-B (*Crocospheira* sp.; Moisander et al., 2010), and UCYN-C (Foster et al., 2007); filamentous, colonial, non-heterocyst forming *Trichodesmium* spp. (Church et al., 2005b); a widespread *Rhizosolenia* (Het-1; Church et al., 2005b), *Richelia* associated with *Hemiaulus* (Het-2; Foster et al., 2007), and *Calothrix* associated with *Chaetoceros* (Het-3; Foster et al., 2007); and a widespread *γ*-proteobacterial phylotype *γ*-24774A11 (Moisander et al., 2008).

Recombinant plasmids with the targeted organism nifH fragment were used as qPCR standards, and each 96 well plate was run with a serial dilution ($10^0–10^7$ nifH copies reaction$^{-1}$) of the appropriate standard. All qPCR reactions were set up with 1 µL of DNA extract in a 15 µL volume with the following reagents and final concentrations: 1X Taqman® Master Mix (Applied Biosystems, Carlsbad, CA, USA), 0.4 µM each forward and reverse primers, and 0.2 µM probe (5’-FAM and 3’-TAMRA labeled). Thermocycle parameters were as described in Goebel et al. (2010) for all assays, with the exception of using a 64°C annealing temperature for UCYN-A2. The qPCR reaction efficiencies were as follows: UCYN-A1 – 96.7 ± 2.8 %; UCYN-A2 – 94.9 ± 3.4 %; UCYN-B – 101.3 ± 2.2 %; UCYN-C – 95.4 ± 2.9 %; *γ*-24774A11 – 96.2 ± 3.2 %; *Trichodesmium* – 93.6 ± 1.4 %; Het-1 – 96.1 ± 4.7 %; Het-2 – 98.3 ± 2.0 %; and Het-3 – 97.5 ± 0.9 %. Based on the differences in sample volumes, the limits of detection (LOD) and quantification (LOQ) for all qPCR assays ranged between 11 and 56 nifH copies L$^{-1}$, and between 87 and 444 nifH copies L$^{-1}$, respectively. Samples were determined to be “detected, not quantified” (DNQ) when calculated abundances were greater than the LOD, but less than the LOQ. Abundances are reported as nifH copies L$^{-1}$ rather than cells L$^{-1}$ because there is currently little information about the number of nifH copies per genome in these diazotroph targets.

Preliminary analyses were conducted from three depth profiles on days 19 and 20, and very little vertical stratification was observed for most of the targeted diazotrophs, with abundances for most diazotrophs measured at the same order of magnitude throughout the 15 m mesocosm (Supplement Fig. S1). *Trichodesmium* and Het-3 were the only exceptions to this observation. Based on these findings, qPCR analyses focused on samples taken from the middle of the mesocosm (6 m), collected on odd days from inside the mesocosms, and even days from outside the mesocosms.

Growth and mortality rates were calculated for individual diazotrophs inside the mesocosms when abundances were higher than the LOQ on two consecutive sampling days, as described in Moisander et al. (2012), using the following formula: $k = 2.303 \times \log(n_{1}/(n_{2}-n_{1}))/(t_{2}–t_{1})$, where $n_{x}$ = abundance at time $x$. This assumes that the organisms
were growing exponentially during the experiment, which cannot be easily verified in field populations. These rates implicitly include grazing, mortality and viral lysis, and thus are net growth and mortality rates.

2.3 Determination of diazotroph community composition using high-throughput sequencing of nifH amplicons

In order to evaluate whether the diazotrophs targeted via qPCR assays were representative of the phylotypes present in the lagoon and mesocosms, partial nifH fragments (ca. 360 base pairs) were amplified using a nested PCR assay and universal nifH primers nifH1-4, as described in Turk-Kubo et al. (2014). Two lagoon samples (day 1 and day 22) and three samples from each mesocosm (day 3, day 23, and the day where UCYN-C abundances were beginning to increase, i.e., days 11–15; see discussion below) were chosen for analysis. For each sample, triplicate PCR reactions were pooled. Internal primers were modified 5' common sequence (CS) linkers (CS1_nifH1F: 5'-ACACTGACGA CATG-GTTCTA CATGYGAYCC NAARGCNGA; CS2_nifH2R: 5'-TACGGTAGCA GAGACTTGGT CTADNGCCAT CATTYCNCC) to facilitate library preparation at the DNA Services (DNAS) Facility at the University of Illinois, Chicago, using the targeted amplicons sequencing (TAS) approach described in Green et al. (2015). These libraries were pooled with other libraries to achieve a target depth of ca. 40,000 sequences per sample. Sequencing of paired end reads was performed at the W.M. Keck Center for Comparative and Functional Genomics at the University of Illinois at Urbana-Champaign using Illumina MiSeq technology. De-multiplexed raw paired end reads were merged in the CLC Genomics workbench, and merged reads between 300 and 400 base pairs in length were selected. Quality filtering was performed in QIIME (Caporaso et al., 2010) using the usearch quality filter (usarch_qf) pipeline script, which includes steps for denoising, de novo chimera removal using UCHIME (Edgar et al., 2011) and operational taxonomic unit (OTU) determination using usearch6.1 at 97% nucleotide identity (Edgar, 2010). Representative nucleotide sequences from OTUs with greater than 100 reads (277 out of 2325 OTUs, representing 92% of all sequences that passed the usearch quality filter) were imported into ARB (Ludwig et al., 2004), translated into protein sequences, where non-nifH OTUs or those with frameshifts were discarded. QIIME script exclude_seqs_by_blast.py was used to check for sequences with >92% amino acid identity to known contaminants; none were found. OTUs targeted by each qPCR assay were determined in silico for each group of diazotrophs in ARB by identifying representative sequences that had zero to two mismatches in either primer or the probe binding region, without exceeding a total of four mismatches in all three regions (see Supplement Table S4).

For the characterization of the overall diazotroph community composition in the lagoon and mesocosms, representative sequences from the most highly recovered OTUs (109 OTUs representing 85% of all post quality-filtering sequences) were considered. Translated amino acid sequences were aligned to the existing amino acid alignment in the curated database. Maximum likelihood trees were calculated using translated amino acid sequences from representative sequences and their closest relatives (determined via blastp) in MEGA 6.06 (Tamura et al., 2013), using the JTT-matrix-based model and bootstrapped with 1000 replicate trees. Distribution of read data across samples for each representative sequence was visualized in the Interactive Tree of Life online tool (Letunic and Bork, 2006).

Raw reads (fastq files) were deposited into the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under BioProject ID PRJNA300416 and BioSample accessions SAMN04202524–SAMN04202534.

2.4 PCR amplification of Cyanothece-like organisms

Partial nifH gene fragments (233 base pairs) from Cyanothece-like organisms were PCR amplified using primers designed as part of this study to broadly target the UCYN-C group, including cultivated Cyanothece spp., Gloeocapsa spp., the cyanobionts of diatoms E. turgida, and R. gibba, as well as many closely related uncultivated Cyanothece-like ecotypes. The oligonucleotide primers Cyanothece_nifH_F (5'-CTT AGC AGC AGA ACG GGG AA-3') and Cyanothece_nifH_R (5'-GCA TTG CGA AAC CAC CAC AA-3') were designed using NCBI’s Primer BLAST, screened in silico for cross reactivity to non-target nifH phylotypes in a curated nifH ARB database (Heller et al., 2014) and synthesized by Sigma Oligos (St. Louis, MO, USA).

Duplicate PCR reactions were carried out in 20 µL volumes with 1X Platinum® Taq PCR buffer (Invitrogen, Carlsbad, CA), 3.0 mM MgCl2, 400 µm dNTP mix, 0.2 µM of each forward and reverse primers, 1 U Platinum® Taq polymerase (Invitrogen), and 2 µL of DNA extract. Thermocycle parameters were as follows: the initial denaturation step at 94°C for 5 min was followed by 30 cycles of 94°C for 30 sec, 58°C for 30 sec, and 72°C for 30 sec and a final elongation step at 72°C for 10 min. PCR amplicons were cleaned using the Qiagen Gel Extraction Kit (Qiagen), and sequenced directly using Sanger technology at the UC Berkeley DNA Sequencing Center using Cyanothece_nifH_R to prime the sequencing reaction.

Raw sequences were processed using Sequencer 5.2.4 (Gene Codes Corporation, Ann Arbor, MI) and phylogenetic analyses were conducted in ARB (Ludwig et al., 2004), using a curated database of all nifH sequences available in Genbank (Heller et al., 2014) and in MEGA 6.06 (Tamura et al., 2013).
Sequences were submitted to NCBI’s Genbank database under accession numbers KU160532–KU160537.

2.5 Statistical analyses

Pearson correlation coefficients were calculated for each paired variable, after transformation to assure normality, and samples collected from the incubations and the Noumea lagoon were treated independently. Tables S1 and S2 detail linear associations and transformations required for each variable inside and outside the mesocosms, respectively. Diazotroph abundance and physico-chemical environmental data were analyzed in R (Team, 2012). Biplot form principal component analysis was used to examine the temporal and spatial variation in environmental parameters. Correspondence analysis of the abundance data in each mesocosm separately was performed with the ca package (Nenadic and Greenacre, 2007). Biplots of sampling days (rows) and organisms (columns) were compared to examine reproducibility across mesocosms. Least squares regression models were calculated to fit and predict the effect of temperature, salinity, and PO₄ on diazotroph abundance.

3 Results and discussion

3.1 Diazotroph community structure in the VAHINE mesocosms experiment and in the Noumea lagoon

In order to (1) characterize the diazotroph community composition in the mesocosms and the lagoon using a qualitative approach that is complementary to qPCR and (2) evaluate whether the qPCR assays used in this study, which are widely used in studies of the oligotrophic ocean, target the major lineages of diazotrophs in the Noumea lagoon, fragments of the \( \text{nifH} \) gene were amplified from a subset of samples using a well-established nested PCR approach (Zehr and Turner, 2012; Table S4). Clustering at 97 % nucleotide identity yielded 2325 OTUs with greater than 4 reads, and 334 OTUs with greater than 100 reads. 277 of these OTUs (representing 479 402 reads or 88 % of all reads that passed quality filtering steps) were affiliated with \( \text{nifH} \) gene reads that passed the quality filtering steps (first column), and the number of OTUs affiliated with each group that are successfully targeted by qPCR assays used in this study. Rows in bold are pooled results for the entire \( \text{nifH} \) subgroup or group. Partial \( \text{nifH} \) sequences were classified according to the convention defined in Zehr et al. (2003). OTU – operational taxonomic unit.

Table 1. In silico qPCR coverage analysis. Taxonomic assignment for all \( \text{nifH} \) amplicons reads that passed the quality filtering steps (first column), and the number of OTUs affiliated with each group that are successfully targeted by qPCR assays used in this study. Rows in bold are pooled results for the entire \( \text{nifH} \) subgroup or group. Partial \( \text{nifH} \) sequences were classified according to the convention defined in Zehr et al. (2003). OTU – operational taxonomic unit.

<table>
<thead>
<tr>
<th>OTU</th>
<th>No. of OTUs</th>
<th>% OTUs</th>
</tr>
</thead>
<tbody>
<tr>
<td>UCYN-A</td>
<td>68 (260221)</td>
<td>88 % (99 %)</td>
</tr>
<tr>
<td>UCYN-B</td>
<td>32 (15917)</td>
<td>0 % (0 %)</td>
</tr>
<tr>
<td>UCYN-C</td>
<td>18 (11324)</td>
<td>50 % (26 %)*</td>
</tr>
<tr>
<td>Tricho.</td>
<td>18 (20034)</td>
<td>78 % (97 %)</td>
</tr>
<tr>
<td>Het-1</td>
<td>2 (1738)</td>
<td>100 % (100 %)</td>
</tr>
<tr>
<td>Het-2</td>
<td>0 (0)</td>
<td>0 % (0 %)</td>
</tr>
<tr>
<td>Het-3</td>
<td>0 (0)</td>
<td>0 % (0 %)</td>
</tr>
<tr>
<td>Other</td>
<td>14 (2316)</td>
<td>na</td>
</tr>
<tr>
<td>1G</td>
<td>88 (120586)</td>
<td>82 % (99 %)</td>
</tr>
<tr>
<td>( \gamma ) – 24774A11</td>
<td>22 (51594)</td>
<td>82 % (99 %)</td>
</tr>
<tr>
<td>Other</td>
<td>66 (60992)</td>
<td>na</td>
</tr>
<tr>
<td>1J/1K</td>
<td>19 (40032)</td>
<td>100 % (100 %)</td>
</tr>
<tr>
<td>3</td>
<td>16 (6825)</td>
<td>100 % (100 %)</td>
</tr>
<tr>
<td>1O</td>
<td>2 (409)</td>
<td>na</td>
</tr>
</tbody>
</table>

* 61 % (85 %) when a third mismatch at the 5’ end (probe) is allowed.

In total, 636 848 paired end reads were recovered from 11 samples, and 544 209 passed the quality filtering steps described above (Table S4). Clustering at 97 % nucleotide identity yielded 2325 OTUs with greater than 4 reads, and 334 OTUs with greater than 100 reads. 277 of these OTUs (representing 479 402 reads or 88 % of all reads that passed quality filtering; Table S4) remained after removing non-\( \text{nifH} \) reads or those with mismatches, and were used in downstream analyses. A majority of these 277 OTUs (152 OTUs, representing 311 550 reads, or 65 % of the reads selected for analysis) were affiliated with \( \text{nifH} \) cluster 1B Cyanobacteria. Reads affiliated with \( \text{nifH} \) cluster 1G, which is composed primarily of \( \gamma \)-proteobacterial phylotypes, was the second most highly recovered group (88 OTUs, representing 120 586 reads, or 25.2 % of the reads selected for analysis). Reads that were closely related to \( \text{nifH} \) cluster 1J/1K, comprised primarily of \( \alpha \) - and \( \beta \)-proteobacteria, were also recovered, but only comprised 8.4 % of the reads selected for analysis (19 OTUs, 40 032 reads). Cluster 3 and cluster 1O affiliated reads were recovered, but together accounted for less than 2 % of the reads selected for analysis (Table 1).

The two OTUs with the highest relative abundance (OTU1890 and OTU2317), accounted for 31 % of the reads selected for analysis and were closely related to the prymnesiophyte symbiont UCYN-A2 (Table S5). Both OTUs were present in the lagoon in day 1 and day 22 samples, and were recovered at high relative abundances from all three mesocosms throughout the experiment (Fig. 1). The third most highly recovered OTU (OTU1) was a \( \gamma \)-proteobacteria closely related to \( \gamma \)-24774A11, a heterotrophic diazotroph with widespread occurrence (Moisander et al., 2014; Langlois et al., 2015), that is also preferentially amplified by the \( \text{nifH} \) primers used (Turk et al., 2011). This sequence type was present in the lagoon samples, and had high relative abundances in all three mesocosms in midpoint samplings (days 11, 13, and 15 for M1, M2, and M3, respectively; Fig. 1). OTU2280 (cluster 1J/1K) was the OTU with the fourth highest relative abundance. It does not have high sequence similarity to any uncultivated or cultivated organisms, with the closest relative, an uncultivated rhizosphere isolate (Genbank accession no. KC667160), sharing only 86 % nucleotide sequence similarity. This is true for all but one of the 1J/1K OTUs, OTU119, which is closely related (98 % nucleotide identity) to an environmental sequence recovered from Heron Reef (Genbank accession no. EF175779).

Also among OTUs that had high recovery were UCYN-A1 (OTU2008, OTU1754, and OTU2), other UCYN-A2...
Figure 1. Maximum likelihood tree calculated using partial nifH amino acid sequences recovered from the Noumea lagoon (NL) and mesocosms (M1, M2, and M3). Relative abundances of nifH reads associated with each operational taxonomic unit (OTU) are indicated for each sample by shaded boxes, with intense shading indicating high relative abundances, and light shading indicating low relative abundances. Trees were bootstrapped using 1000 replicate trees, and nodes with values > 50 are displayed. Branch lengths were inferred using the JTT model, and the scale bar indicates the number of substitutions per site. OTUs that are targeted by qPCR assays used in this study are marked with a black diamond (♦), and two UCYN-C sequences that are likely to amplify are marked with a circle (⊙). nifH cluster designations according to the convention in Zehr et al. (2003) are notated at the right. d1 – day 1; d11 – day 11; d13 – day 13; d15 – day 15; d22 – day 22; d23 – day 23.
OTUs (OTU2325, OTU1548, and OTU664), *Trichodesmium* (OTU5 and OTU35), UCYN-C (OTU12) and two OTUs affiliated with cluster 1G (OTU2218 and OTU2199) (Fig. 1 and Table S5). 1G OTUs are recovered from the lagoon sample at day 22, and have high relative abundances in all three mesocosms by the end of the experiment. These sequence types are not closely related to γ-24774A11, and thus are not quantified using qPCR assays in the study, and their quantitative importance in the mesocosm environment cannot be determined based on this qualitative measure. It is important to note that high relative abundances in PCR-based libraries are not indicative of high abundances, as often these sequence types dominate PCR libraries yet are present at low abundances in the environment (Hewson et al., 2007; Bonnet et al., 2013; Turk-Kubo et al., 2014; Shizoaki et al., 2014; Bentzon-Tilia et al., 2015), presumably as a result of preferential amplification (Turk et al., 2011).

A majority of the OTUs with high relative abundances (159 out of 277, representing 380 536 out of 479 402 reads) is affiliated with the following lineages targeted by qPCR assays used in this study: UCYN-A1, UCYN-A2, UCYN-B, UCYN-C, *Trichodesmium*, and *Richelia* associated with *Rhizosolenia* (Het-1), and γ-24774A11. To determine whether the qPCR assays used would target the diazotrophs present, representative sequences were identified that contained between zero and two mismatches in the primer and probe binding regions, without exceeding a total of four mismatches for all three regions. For UCYN-A, *Trichodesmium*, Het-1, and γ-24774A11 lineages, nearly all of the sequence types present met these criteria (between 97 and 100 %), and thus would be quantified in the qPCR assays (Table 1). Only 26 % of the recovered *Cyanothece*-like sequences would successfully be targeted by the UCYN-C assay; however, this coverage increases to 85 % when including two of the most highly recovered OTUs that have a third mismatch at the 5’ end of the probe binding region, and thus are still likely to be quantified. Together with the results of UCYN-C specific PCRs (see Sect. 3.3 below), these results indicate that qPCR assays targeting UCYN-C likely quantified most of the phylotypes present, but absolute abundances may be underestimated.

A group of 32 OTUs clustered within the Chroococcales and were closely related to both *Crocosphaera watsonii* 8501 (93–97 % amino acid identity) and *Cyanothece* sp. WH8904 (93–97 % amino acid identity). These OTUs were not targeted by the UCYN-B qPCR assay, with four to eight total mismatches in the primer/probe binding regions. Despite being considered putative *Crocosphaera* sp. for the purpose of this study, it is possible that they are actually from a *Cyanothece*-like organism (not targeted by the UCYN-C qPCR assay) or another member of the Chroococcales. These OTUs were most highly recovered at the final sampling day (day 23) in M2 and M3, but were not recovered in the lagoon libraries.

The heterocyst-forming symbionts of diatoms are conspicuously underrepresented in these libraries, with only two Het-1 OTUs recovered, despite being one of the most abundant diazotrophs measured in the lagoon and mesocosms (see discussion below). Furthermore, there were no *Richelia* associated with *Hemiaulus* (Het-2) or *Calothrix* associated with *Chaetoceros* (Het-3) sequences recovered despite being detected using qPCR assays. These findings lend further support to previous reports that this degenerate PCR assay does not work well for these phylotypes (Foster and Zehr, 2006). This underscores a major limitation of characterizing community composition using solely qualitative PCR techniques.

### 3.2 Most abundant diazotrophs in VAHINE mesocosms

Experiment shift from *Richelia* to *Cyanothece*-like ecotypes

During the VAHINE mesocosm experiment, three periods could be roughly defined based on the identity of the most abundant diazotroph and biogeochemical parameters (described in detail in Berthelot et al., 2015). A diazotroph community similar to the Noumea lagoon field community and low DIP concentrations characterized the period prior to the DIP spike (hereafter referred to as P0). The second period, days 5–14 (hereafter referred as P1), was characterized by high abundances of Het-1, high DIP availability, as well as moderate N2 fixation rates (10.1 ± 1.3 nmol N L−1 d−1; Bonnet et al., 2015a). In the second half of the experiment, days 15–23 (hereafter referred as P2), the UCYN-C population grew to high abundances, and was characterized by low DIP availability, and high N2 fixation rates (27.3 ± 1.0 nmol N L−1 d−1; Bonnet et al., 2015a), with rates reaching > 60 nmol N L−1 d−1, ranking among the highest reported in marine waters (Luo et al., 2012).

At the initiation of the mesocosm experiment, the most abundant diazotrophs in the Noumea lagoon were Het-1 (3.1 × 104 nifH copies L−1), and *Richelia* associated with *Hemiaulus* (Het-2; 1.2 × 104 nifH copies L−1), as well as the *B. bigelowii*-associated UCYN-A2 (1.5 × 104 nifH copies L−1) and UCYN-A1 (5.6 × 103 nifH copies L−1). *Trichodesmium* and the uncultivated unicellular cyanobacterial group C (UCYN-C) were present but at lower abundances, 2.8 × 103 nifH copies L−1 and 7.8 × 102 nifH copies L−1, respectively. γ-24774A11 and *Calothrix* associated with *Chaetoceros* (Het-3) were also present but at abundances too low to quantify (DNQ), and no *Crocosphaera* sp. (UCYN-B) could be detected (Table S3).

During P0, evidence of shifts from the original Noumea lagoon diazotroph assemblage could be observed in all three mesocosms (Fig. 2b–e). In M1, M2 and M3, Het-1 remained the most abundant diazotroph, and the relative proportions of Het-2 and UCYN-A1 were similar to those measured outside the mesocosms. However, UCYN-A2 abundances decreased and *Trichodesmium* abundances increased over this 3-day period with respect to their abundances in the lagoon. These changes were most pronounced in M3. As in the l-
given, UCYN-C and γ-24774A11 were present at low abundances (ca. 10^2 nifH copies L^{-1}) in the day 3 sampling.

Diazotroph assemblages within each mesocosm remained relatively consistent throughout the first thirteen days of the experiment, with Het-1 the most abundant diazotroph (Fig. 2b–d). However, a significant shift in community composition began about halfway through the 23-day experiment, evidenced by the increasing abundances of UCYN-C and Het-3, and decreasing abundances of Het-1 and UCYN-A1 (Figs. 2b–d and 3). Increasing UCYN-C abundances were first seen in M1 (day 11), then M2 (day 13), followed by M3 (day 15), which reflected the time in each mesocosm when the DIP turnover time dropped below 1 d^{-1}, signaling DIP limitation (Berthelot et al., 2015; Moutin et al., 2005). In all three mesocosms, which acted as biological replicates, UCYN-C abundances were not only strongly correlated with the decreasing DIP concentrations (p = 0.004, r = −0.51), but also increasing temperatures in the mesocosms (p = 0.000, r = +0.86) and could be weakly correlated with increasing salinity (p = 0.000, r = +0.56), decreasing Het-1 (p = 0.002, r = −0.54), Het-2 (p = 0.01, r = −0.37) and UCYN-A1 (p = 0.000, r = −0.64) abundances. A similar correlation between UCYN-C abundance and both increasing salinity and temperature was observed outside in the Noumea lagoon (although UCYN-C abundances in the lagoon remained low; see Sect. 3.4). Salinity in each of the three mesocosms and in the lagoon did increase gradually during the experimental period, but were elevated inside the mesocosms in P2, likely due to evaporative loss (Bonnet, 2015). Together, this suggests that temperatures below 25.6 °C are not optimal temperatures for this group, and that it may tolerate slightly elevated salinities better than other diazotrophs present. This is consistent with the warmer temperatures required for growth for the UCYN-C isolate TW3 (Taniuchi et al., 2012), and the occurrence of the UCYN-C group in low latitude, warm waters (> 28 °C) in the tropical North Atlantic (Langlois et al., 2008). However, changes in temperature and salinity alone cannot explain the high abundances of UCYN-C in these mesocosms, as similar increases were recorded in the lagoon without a corresponding increase in UCYN-C abundances, and other diazotrophs also can grow optimally at these temperatures (e.g., Crocosphaera sp. (Webb et al., 2009) and Trichodesmium sp.; Breitharth et al., 2007).

Het-3 was undetected in all mesocosms until day 7, and then remained at low abundances (usually below quantitation limits) until day fifteen of the incubation. After day 15, Het-3 abundances were on a similar order of magnitude to their abundances in the Noumea lagoon (ca. 10^2–10^3 nifH copies L^{-1}). Increases in Het-3 abundances after day 15 could be strongly correlated to increases in temperature (p = 0.000, r = +0.82) as well as salinity (p = 0.000, r = +0.78), total chl a (p = 0.000, r = +0.71), and UCYN-C abundances (p = 0.000, r = +0.77). Het-3 abundances could also be weakly corre-
Figure 2. Abundances of targeted diazotrophs at the 6 m depth during the VAHINE mesocosm experiment and in the Noumea lagoon (NL) during the experimental period. Het-1 is the most abundant diazotroph in P0 and P1, while UCYN-C abundances become abundant during P2 in all three mesocosms, M1 (b), M2 (c), and M3 (d). In the NL (e), increasing abundances of UCYN-C during the experimental period are not observed. DIP concentrations (a) decreased steadily following the spike at day 4 (data from Bonnet, 2015).

2015), it follows that UCYN-C may have benefitted either indirectly from the DIP fertilization, or directly from a physical aspect of the mesocosms themselves. A biofilm had accumulated on the sides of the mesocosm bags, and although this was not sampled for molecular analysis, it is possible that the UCYN-C ecotype was a component of this biofilm community, and thus dependent on a physical environment not representative of water column conditions. Another explanation may be that the decreased turbulence in the mesocosm environment created favorable conditions for this ecotype. Finally, it is possible that the inverse correlation between UCYN-C abundances and DIP concentrations may be a result of UCYN-C being able to outcompete other diazotrophs for organic P substrates, under low DIP conditions. *Cyanobacterium* sp. strains PCC 8801 and 8802 have genes used in phosphonate metabolism (Bandyopadhyay et al., 2011), which strongly implies that some strains are able to use this organic P substrate to meet cellular P requirements. The strain most closely related to the UCYN-C ecotype, *Cyanobacterium* sp. CCY0100, also has genes for phosphonate metabolism and transport (JGI website). It is evident that some component of the microbial community in the
mesocosms was utilizing DOP, as DOP stocks were drawn down to levels seen in the lagoon (0.105 ± 0.011 μmol L⁻¹) in the second half of the experiment (Berthelot et al., 2015), however, UCYN-C abundances could not be correlated to this drawdown (data not shown). Trichodesmium erythraeum IMS101 and Calothrix rhizosoleniae SC01 (Het-3) are the only two other diazotrophs targeted in this study known to possess the metabolic capability to utilize phosphonates (Dyhrman et al., 2006; Hilton et al., 2013). Richelia associated with Hemiaulus (Het-2) do not have any genes for the metabolism of phosphonates (Hilton et al., 2013), but it remains unclear whether the symbiont of Rhizosolenia (Het-1) is able to use phosphonates, as four genes related to phosphonate metabolism were identified in the genome (Hilton, 2014). However, as with UCYN-C, no correlations between DOP concentration and Trichodesmium, Het-1 or Het-3 abundances were seen (data not shown). This may be in part due to shortcomings in our understanding of the chemical composition of the DOP pool (Dyhrman et al., 2007), and which organic P compounds are bioavailable to which organisms.

3.3 Phylogenetic identity of diazotrophs targeted with UCYN-C qPCR assay

The UCYN-C qPCR assay used in this study was designed to target a nifH sequence type recovered from Amazon-influenced waters in the tropical North Pacific (Foster et al., 2007). In addition to uncultivated sequences of marine origin, this assay also targets cultivated members of the UCYN-C group including Cyanothece sp. strains AT51142, AT51472, CCY 0110, as well as some freshwater cyanobacterial symbionts of diatoms (Fig. 4). Due to the importance of this group in the mesocosm experiment, and the uncertainty about exactly which organism(s) were targeted with the qPCR assay, PCR amplification of a partial nifH fragment was used to specifically characterize the identity of the Cyanothece-like organisms. Two closely related sequence types were recovered from 6 m samples on days 15 and 20, represented by 60341CB and 60343CB (Fig. 4), which shared 95 % nucleotide similarity. The Cyanothece-like sequences recovered from the mesocosms were most closely related to Cyanothece sp. CCY0100 (92–93 % nucleotide identity; Fig. 4), which was isolated from coastal waters in Chwaka Bay, Zanzibar. They shared only 90–91 % nucleotide identity with the sequences used to design the oceanic UCYN-C primer of Foster et al. (2007) and the UCYN-C isolate TW3 (Taniuchi et al., 2012). Thus the ecotypes that reached such high abundances in the mesocosms were more closely related to an ecotype reported in a coastal environment than those recovered from open water regimes.

It is important to note that the sequences amplified using this Cyanothece-specific PCR are targeted by the UCYN-C qPCR assay, and the phylogenotypes with three mismatches in the probe binding region were not recovered with this assay. This may be due to differences in primer efficiency, the number of amplification steps, and/or depth of sequencing, but these results lend further support to the assertion that important UCYN-C phylogenotypes were quantified with qPCR assays.

3.4 In situ net growth and mortality rates

The VAHINE mesocosm experiment provided a rare opportunity to repeatedly sample the same water mass for an extended period of time, and thus the ability to empirically determine in situ net growth or mortality rates for individ-
Surprisingly, maximum net growth rates for each diazotroph phylotype were among the highest reported for UCYN-A2 and UCYN-C were among the highest measured at 1.71, and 2.16 d\(^{-1}\), respectively, and represent the only reported growth rates for these uncultivated diazotrophs. Very little is known about the newly described uncultivated \textit{B. bigelowii}/UCYN-A2 association, but the difference between UCYN-A1 and UCYN-A2 net growth rates and their patterns within the same mesocosm indicate that the growth rates of each association are likely dependent upon different environmental variables. It also seems plausible, due to the difference in size between the two \textit{B. bigelowii} hosts (Thompson et al., 2014), that they are grazed by different protists.

Experimental growth rates as high as 1.92 d\(^{-1}\) have been reported for batch cultures of \textit{Cyanotherce} sp. ATCC 51142 (Vu et al., 2012), and Taniuchi et al. (2012) measured rates up to 0.85 d\(^{-1}\) in the presumably oligotrophic UCYN-C TW3 isolated from the Kuroshio Current. Thus, rates calculated for UCYN-C in the VAHINE mesocosms experiment are similar in magnitude. However, such high rates in culture are a direct result of the lack of competition with other organisms for nutrients important for growth and the lack of grazing pressure. Despite high growth rates in culture, TW3 is found at low abundances in the Kuroshio Current, presumably due to these factors (Taniuchi et al., 2012). Although the factors behind such high growth rates for UCYN-C in the mesocosms are not clear, it is possible that UCYN-C was free from significant grazing in these experiments.

Maximum net growth rates for the filamentous diazotrophs were also generally higher than previously reported growth rates. \textit{Trichodesmium} sp. maximum net growth rates for were as high as 1.46 ± 0.05 (M1) and 1.55 ± 0.02 d\(^{-1}\) (M3), and microscopic analyses indicated (data not shown) that both \textit{T. erythraeum} and \textit{T. thiebautii} were present in the
mesocosms, the former being the dominant Trichodesmium species. These calculated rates are much higher than the specific growth rate previously reported for cultures of T. erythraeum (0.29 ± 0.04 d⁻¹; Hutchins et al., 2007) and net growth rates for T. thiebautii populations in the Noumea lagoon (0.11–0.38 d⁻¹; Rodier and Le Borgne, 2008). Maximum growth rates for the DDAs Het-1, Het-2, and Het-3 were 1.28, 2.24, and 1.09 d⁻¹, respectively. Although these are the first reported growth rates for these DDAs determined using in situ cellular abundances, Foster et al. (2011) determined maximum net growth rates to be much lower when using nanoSIMS-based techniques to quantify ¹⁵N incorporation in biomass, which is to be expected if the DDAs are using other N sources (e.g., dissolved organic N and/or recycled N) in addition to fixing N.

For many diazotrophs, the pattern of net growth and mortality rates indicated a very dynamic process that appeared specific to each mesocosm. Most diazotrophs experienced multiple transitions between net growth and net mortality within a single mesocosm throughout the 23-day incubation (Table 3). For example, in M1, net growth rates for Trichodesmium were observed approximately every other sampling period (on days 9, 13, 15, 19, and 23), and in some intervening periods experienced mortality rates similar in magnitude to growth rates (e.g., days 7 and 11). In contrast, net growth occurred on days 7, 11, 13, 20, and 23 in M3, and one of the largest net mortality rates was measured at day 9 (−2.16 ± 0.18 d⁻¹). There are no aspects of the experimental design that can be invoked to explain this variability; in fact biogeochemical parameters and picoplankton population dynamics were well replicated in all three mesocosms (Bonnet, 2015). Therefore the dynamic nature of diazotroph growth and mortality rates in each mesocosm most likely results from a combination of grazing pressure and viral lysis, which can be expected to reflect natural variations in the grazers and virus present.

### 3.5 Symbiotic Richelia and UCYN-A ecotypes are abundant in the Noumea lagoon during the VAHINE mesocosms experiment

The VAHINE mesocosm experiment provided an additional opportunity to characterize the abundances of targeted diazotrophs in the Noumea lagoon using quantitativa
tive techniques. As described above, the diazotroph assemblage at the onset of the VAHINE mesocosm experiment was dominated by Het-1, Het-2, UCYN-A2 and UCYN-A1 (see Sect. 3.1). Het-1 remained the most abundant of the targeted diazotrophs for most of the sampling days ($3.1 \times 10^4$--$5.5 \times 10^5$ nifH copies L$^{-1}$) during the period of study (Fig. 2e, Table S3). Het-2 abundances were consistently lower than Het-1, and variable throughout the period of study, ranging between $4.0 \times 10^3$ and $1.6 \times 10^5$ nifH copies L$^{-1}$, and reaching peak abundances on day 10 and again on day 18 (Table S3). Bonnet et al. (2015a) discusses the BNF rates in detail; however, it should be noted that in the lagoon, only Het-2 abundances had a significantly strong positive correlation with bulk N$_2$ fixation rates ($p = 0.01$, $r = +0.78$).

Although Richelia living in association with Rhizosolenia (Het-1) and Hemiaulus (Het-2) are known to be widely distributed and potentially significant N$_2$ fixers in oligotrophic oceans (Carpenter et al., 1999; Subramaniam et al., 2008; Karl et al., 2012), reports of their presence in coastal waters in the southwestern Pacific are rare, and previous studies during the austral summer in the Noumea lagoon reported very low abundances of heterocystous symbionts (Rodier and Le Borgne, 2010; Biegala and Raimbault, 2008). This is the first report on quantitative abundances of Richelia in the Noumea lagoon, which are comparable to abundances reported from the tropical North Atlantic (Foster et al., 2007; Goebel et al., 2010) and the North Pacific Subtropical Gyre (Church et al., 2008; Foster and Zehr, 2006), where the DDAs are thought to have a significant impact on C sequestration due to their productivity and rapid sinking (Subramaniam et al., 2008; Karl et al., 2012; Villareal et al., 2012), and may play a role in supporting the benthic community (Houlbreque and Ferrier-Pagès, 2009).

UCYN-A2 was the second most abundant member of the targeted diazotroph assemblage in the lagoon for the first 10 days of the experiment (Fig. 2e), and was present at abundances as high as $1.1 \times 10^5$ nifH copies L$^{-1}$ (day 6; Table S3). UCYN-A2 abundances declined steadily throughout the period of the 23-day experiment ($p = 0.04$, $r = -0.63$), and had a significantly strong negative correlation with sea temperature ($p = 0.003$, $r = -0.82$) and salinity ($p = 0.03$, $r = -0.89$), both of which were increasing in the Noumea lagoon throughout the VAHINE deployment (Bonnet, 2015). UCYN-A1 abundances were consistently lower than those of UCYN-A2, ranging between $2.8 \times 10^3$ and $6.4 \times 10^4$ nifH copies L$^{-1}$ with peaks in abundance at day 6 and again at day 22 (Table S3). Interestingly, UCYN-A1 abundances had a significant positive correlation with total chl $a$ in the $>10 \mu$m size fraction ($p = 0.03$, $r = +0.72$). The significant inverse relationship between UCYN-A2 abundances and water temperatures in the Noumea lagoon suggests that the B. bigelowii/UCYN-A2 association may thrive at lower temperatures than other diazotrophs (e.g., Trichodesmium), similar to the B. bigelowii/UCYN-A1 association (Moisander et al., 2010).

The coexistence of the B. bigelowii/UCYN-A1 and B. bigelowii/UCYN-A2 ecotypes in the Noumea lagoon, both present at reasonably high abundances, indicates that the ecological niches of these cryptic symbioses overlap. Very little information currently exists about the ratio of UCYN-A1 to UCYN-A2 in tropical oligotrophic oceanic regimes, but UCYN-A1 is typically found at higher relative abundances (as inferred from clone-library based studies; Thompson et al., 2014) in oligotrophic regions. The Noumea lagoon is the first location where the co-occurrence of each ecotype has been verified using quantitative techniques, but it must be noted that the difference in abundances may be a result of the number of UCYN-A2 cells associated with each B. bigelowii host, which has been reported to be as high as 11 : 1 (Thompson et al., 2014).

Trichodesmium spp. have been routinely observed in the Noumea lagoon using satellite observations (e.g., Dupouy et al., 2000, 2011) and direct field measurements (Renaud et al., 2005; Masotti et al., 2007; Rodier and Le Borgne, 2008, 2010). Conspicuous blooms that occur in the austral spring and summer months have been associated with warm sea surface temperatures ($>26 ^\circ C$) and recent NO$_3^-$ and soluble reactive P (SRP) enrichments in the lagoon (Rodier and Le Borgne, 2008). Although Trichodesmium was present at relatively low abundances during the first 8 days of the experiment ($3.4 \times 10^2$--$6.5 \times 10^3$ nifH copies L$^{-1}$; Table S3), abundances increased steadily throughout the experiment ($p = 0.01$, $r = +0.73$). Trichodesmium reached peak abundances of $1.4 \times 10^5$ nifH copies L$^{-1}$ on the final sampling day (day 22; Fig. 2e). The only environmental parameter with which changes in Trichodesmium abundances could be strongly correlated was NH$_4^+$ ($p = 0.02$, $r = +0.81$).

The UCYN-C group comprised a minor part of the Noumea lagoon diazotroph community, and was detected at abundances between $3.2 \times 10^2$ and $4.8 \times 10^3$ nifH copies L$^{-1}$ (Table S3) during the experimental period. However, UCYN-C abundances did show a linear increase over the duration of the experiment ($p = 0.01$, $r = +0.74$), and significant positive correlations with changes in salinity ($p = 0.02$, $r = +0.80$), and NH$_4^+$ ($p = 0.03$, $r = +0.76$). The unicellular UCYN-C group was first described when nifH sequences with 90% nucleotide identity to the cultivated Cyanothecae sp. PCC 51142 were recovered from surface waters in the tropical North Atlantic (Langlois et al., 2005; Foster et al., 2007). This phylotype has rarely been quantified in the oligotrophic ocean, and when present, abundances are low (Foster et al., 2007; Langlois et al., 2008; Needoba et al., 2007; Goebel et al., 2010; Ratten et al., 2014); therefore, very little is known about its distribution or importance. The qPCR assay used in this study (Foster et al., 2007) amplifies not only the uncultivated Atlantic phylotype, but also many cultivated Cyanothecae- and Gloeotheca-like isolates, which are ex-
pected to be present in an environment like the Noumea lagoon, as *Cyanobacterium* and *Gloeothece*-like organisms have been reported in shallow marine sediments (Hunter et al., 2006; Bauer et al., 2008) and intertidal sands (Reddy et al., 1993; Ohki et al., 2008).

\[ \gamma = 2477.4A11 \] was also consistently detected at low abundances (ca. 10^2–10^3 \( \text{nifH} \) copies \( \text{L}^{-1} \); Table S3), which increased throughout the duration of the experiment (\( p = 0.03, r = +0.65 \)). Despite being the most widely studied marine heterotrophic diazotroph, very little is known about this phylotype. Recent studies suggest that this \( \gamma \)-proteobacterial diazotroph is likely to be free-living and may be able to survive in a broad range of environmental conditions, as evidenced by low but uniform abundances throughout the photic zone in the South Pacific (Moisander et al., 2014).

The epiphytic *Chaetoceros* symbiont, *Calothrix* (Het-3), was also consistently detected at low abundances (ca. 10^2–10^3 \( \text{nifH} \) copies \( \text{L}^{-1} \); Table S3). *Chaetoceros* can often be found as part of the neritic diatom assemblage; therefore, this association is generally observed in coastal or coastal transition zone regions (Gómez et al., 2005). This is the first report of Het-3 in the coastal oligotrophic waters surrounding New Caledonia.

*Crocospheara* sp. (UCYN-B), previously reported to be members of the unicellular diazotrophic community in the Noumea lagoon (Biegala and Raimbault, 2008), were not initially detected, but were present at low abundances that increased over the 23-day period (\( p = 0.04, r = +0.63 \)) to abundances as high as 2.7 × 10^3 \( \text{nifH} \) copies \( \text{L}^{-1} \), and had a significant strong correlation with total chl \( a \) (\( p = 0.02, r = +0.81 \)). It must be noted that the *Crocospheara*-like sequences that were recovered from the mesocosm libraries would not likely amplify efficiently with the qPCR assay used. Therefore, it is not clear whether these abundances included a *Crocospheara* population not recovered in the PCR-based \( \text{nifH} \) libraries, or were in part the result of cross-reactivity with the *Crocospheara*-like population that were amplified from mesocosm samples during P1 and P2, and possibly present in the lagoon as well.

4 Conclusions

The VAHINE mesocosm experiment was conducted to trace the incorporation of \( \text{N}_2 \) fixed by diazotrophs into the food web (Bonnet, 2015). Despite a lack of sampling replication within each mesocosm, consistent patterns in both the relative abundances of targeted diazotrophic phylotypes, as well as shifts in diazotrophic community composition, were reproducibly seen in all three mesocosms. Although the timing of the increases was specific to an individual mesocosm, UCYN-C and Het-3 abundances increased over time, while UCYN-A1 and Het-1 abundances decreased over time in all three mesocosms (Fig. 3). The experimental conditions selected for the growth of UCYN-C during P2, an ecotype never before quantified at high abundances in the marine water column, which enabled the calculation of growth rates of this uncultivated ecotype, and provided insight into its dynamics with respect to environmental parameters. Although the data strongly suggest that the drawdown of DIP provided an environment favorable for high UCYN-C growth rates, further studies are required to better understand the environmental conditions that stimulated this bloom, and whether such blooms are seen in the Noumea lagoon itself.

The experimental setup provided a rare opportunity to calculate in situ net growth rates for natural populations of diazotrophs, including the uncultivated UCYN-A. This study provided the first growth rates for the UCYN-C phylotype and for UCYN-A2, both of which were surprisingly high, implying not only favorable conditions, but also a lack of grazing pressure. Maximum net growth rates were high for all diazotroph ecotypes, but most also experienced intermittent periods of growth and mortality within the 23-day experiment, which was also an unexpected finding. Along with net rates recently reported by Moisander et al. (2012), we anticipate that these data will be important for future modeling efforts.

The analysis of the diazotrophic assemblage outside the mesocosms represents both the first quantitative data on targeted diazotroph phylotypes, as well as the first \( \text{nifH} \)-based diversity libraries on the populations in the Noumea lagoon. Not previously considered to be a significant diazotrophs in this region, DDAs must now be considered a potentially important contributor to BNF in these waters, especially in the austral summer. Although the presence of UCYN-A in the lagoon has long been suspected due to the relative importance of daytime nitrogen fixation in the small size fraction (Biegala and Raimbault, 2008), we report the first quantitative data on UCYN-A abundances in the Noumea lagoon. Furthermore, the co-occurrence of the two UCYN-A ecotypes revealed in this study provides an important insight into the overlap in environmental niches for these two ecotypes.

The Supplement related to this article is available online at doi:10.5194/bg-12-7435-2015-supplement.

Author contributions. S. Bonnet designed and executed the experiments, and S. Bonnet and A. Desnues sampled the experiments for molecular analyses. M. E. Hogan extracted DNA samples, and K. A. Turk-Kubo conducted all qPCR and PCR analysis and analyzed the data. I. E. Frank performed statistical analyses. K. A. Turk-Kubo and J. P. Zehr prepared the manuscript with input from all co-authors.

Acknowledgements. Funding for this research was provided by the Agence Nationale de la Recherche (ANR starting grant VAHINE
Edited by: D. G. Capone

References


Bonnet, S.: Introduction to the project VAHINE: Variability of vertical and tropHic transfer of fixed N2 in the south west Pacific Ocean, Biogeosciences Discuss., submitted, 2015.


Farrellid, H., Andersson, A. F., Bertilsson, S., Al-Soud, W. A., Hansen, L. H., Sorensen, S., Steward, G. F., Hagstrom, A., and Riemann, L.: Nitrogenase gene amplicons from global marine surface waters are dominated by genes of non-cyanobacteria,


Hilton, J.: Ecology and evolution of diatom-associated cyanobacteria through genetic analyses, Ph.D., Ocean Sciences Department, University of California, Santa Cruz, Santa Cruz, 2014.


K. A. Turk-Kubo et al.: Diazotroph community succession during the VAHINE mesocosm experiment

7451


