

1 Diazotroph community succession during the VAHINE 2 mesocosm experiment (New Caledonia Lagoon)

3 Kendra A. Turk-Kubo^{1*}, Ildiko E. Frank¹, Mary E. Hogan¹, Anne Desnues²,
4 Sophie Bonnet², Jonathan P. Zehr¹

5
6 [1] Ocean Sciences Department, University of California, Santa Cruz, 1156 High Street,
7 Santa Cruz, CA, 95064 USA

8 [2] Mediterranean Institute of Oceanography (MIO) – IRD/CNRS/Aix-Marseille
9 University, IRD Noumea, 101 Promenade R. Laroque, BPA5, 98848 Noumea Cedex

10
11 Correspondence to: K.A. Turk-Kubo (kturk@ucsc.edu)

12 13 **Abstract**

14 The VAHINE mesocosm experiment, conducted in the low-nutrient low-chlorophyll
15 waters of the Noumea Lagoon (coastal New Caledonia) was designed to trace the
16 incorporation of nitrogen (N) fixed by diazotrophs into the food web, using large volume
17 (50 m³) mesocosms. This experiment provided a unique opportunity to study the
18 succession of different N₂-fixing microorganisms (diazotrophs) and calculate *in situ* net
19 growth and mortality rates in response to fertilization with dissolved inorganic phosphate
20 (DIP) over a 23-day period, using quantitative polymerase chain reaction (qPCR) assays
21 targeting widely distributed marine diazotroph lineages. Inside the mesocosms, the most
22 abundant diazotroph was the heterocyst-forming *Richelia* associated with *Rhizosolenia*
23 (Het-1) in the first half of the experiment, while unicellular cyanobacterial Group C
24 (UCYN-C) became abundant during the second half of the experiment. Decreasing DIP
25 concentrations following the fertilization event and increasing temperatures were
26 significantly correlated with increasing abundances of UCYN-C. Maximum net growth
27 rates for UCYN-C were calculated to range between 1.23 ± 0.07 and 2.16 ± 0.07 d⁻¹ in
28 the mesocosms, which are among the highest growth rates reported for diazotrophs.

29 Outside the mesocosms in the New Caledonia lagoon, UCYN-C abundances remained
30 low, despite increasing temperatures, suggesting that the microbial community response
31 to the DIP fertilization created conditions favorable for UCYN-C growth inside the

32 mesocosms. Diazotroph community composition analysis using PCR targeting a
33 component of the nitrogenase gene (*nifH*) verified that diazotrophs targeted in qPCR
34 assays were collectively among the major lineages in the lagoon and mesocosm samples,
35 with the exception of *Crocospaera*-like phylotypes, where sequence types not typically
36 seen in the oligotrophic ocean grew in the mesocosms. Maximum net growth and
37 mortality rates for nine diazotroph phylotypes throughout the 23-day experiment were
38 variable between mesocosms, and repeated fluctuations between periods of net growth
39 and mortality were commonly observed. The field population of diazotrophs in the New
40 Caledonian lagoon waters appeared to be dominated by Het-1 over the course of the
41 study period. However, results from both qPCR and PCR analysis indicated a diverse
42 field population of diazotrophs was present in the lagoon at the time of sampling. Two
43 ecotypes of the *Braarudosphaera bigelowii* symbiont unicellular group A (UCYN-A)
44 were present simultaneously in the lagoon, with the recently described *B.*
45 *bigelowii*/UCYN-A2 association present at higher abundances than the *B.*
46 *bigelowii*/UCYN-A1 association.

47

48 **1 Introduction**

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

59 The New Caledonian (Noumea) coral lagoon, located off the southwestern coast
60 of New Caledonia (South Western Pacific), is a tropical low-nutrient low-chlorophyll
61 (LNLC) system and is bounded by one of the world's largest barrier reefs. Oligotrophic
62 ocean water enters the lagoon from the south over the open shelf, then is driven north by

63 the trade winds and tidal forces and exits through several deep inlets in the intertidal
64 barrier reef that forms the western boundary of the lagoon (Ouillon et al., 2010). Primary
65 productivity is N limited throughout the year (Torréton et al., 2010), giving
66 microorganisms able to fix N₂ gas into bioavailable nitrogen (diazotrophs) a competitive
67 edge over non-diazotrophic organisms. High rates of BNF during the austral summer
68 have been reported, in both large size fractions (>10 µm) and small size fractions (<10
69 µm; Garcia et al., 2007; Biegala and Raimbault, 2008). Garcia et al., (2007) also reported
70 that the percent of N₂ fixation measured in the large size fraction had high temporal
71 variability. Large blooms of the most conspicuous and well-studied diazotroph,
72 *Trichodesmium*, have been repeatedly detected in this region using both indirect (via
73 satellite observation; Dupouy et al., 2011; Dupouy et al., 2000) and direct measurements
74 (Renaud et al., 2005; Masotti et al., 2007; Rodier and Le Borgne, 2008, 2010). Both free-
75 living and organism- and particle-associated unicellular picocyanobacterial diazotrophs
76 are also suspected to be significant contributors to BNF in the lagoon (Garcia et al.,
77 2007; Biegala and Raimbault, 2008), yet the phylogenetic identity of these
78 picocyanobacteria has yet to be determined. Despite evidence that diverse diazotroph
79 communities exist in the Noumea lagoon, there is very little quantitative diazotroph
80 distribution data from this region (Luo et al., 2012), especially for diazotrophs other than
81 *Trichodesmium*.

82 Marine unicellular cyanobacterial diazotrophs are phylogenetically divided into
83 three groups: the uncultivated unicellular group A (UCYN-A; Zehr et al., 2001; Zehr et
84 al., 2008; Tripp et al., 2010) which live in association with strains of the prymnesiophyte
85 *Braarudosphaera bigelowii* (Thompson et al., 2014; Thompson et al., 2012; Hagino et al.,
86 2013); the free-living *Crocospaera* sp. (also referred to as unicellular group B or
87 UCYN-B); and the presumably free-living unicellular group C (UCYN-C), which
88 contains several cultivated cyanobacteria including *Cyanothece* sp. strain ATCC 51142
89 (Reddy et al., 1993) and group C TW3 (Taniuchi et al., 2012).

90 The marine filamentous cyanobacterial diazotrophs include the colonial, non-
91 heterocyst-forming *Trichodesmium*, and the heterocyst-forming symbionts associated
92 with diatoms (DDAs). DDAs form between different strains of *Richelia* sp. associated
93 with diatoms of the genera *Rhizosolenia* (Het-1) and *Hemiaulus* (Het-2) (Villareal, 1990,

94 1992). Het-2 is in an obligate symbiont of *Hemiaulus*, as evidenced by its reduced
95 genome size (Hilton et al., 2013). Although the genome of Het-1 shows evidence of
96 some genome reduction, it remains unclear whether Het-1 is also an obligate symbiont of
97 *Rhizosolenia* (Villareal, 1992;Hilton et al., 2013). The heterocyst-forming *Calothrix*
98 (Het-3) has long been observed living as an epiphyte with *Chaetoceros* (Carpenter and
99 Foster, 2002) but can also grow free from its host (Foster et al., 2010), and has a non-
100 streamlined genome (Hilton et al., 2013).

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

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

124 high frequency sampling of a single water body is required, which is rarely
125 accomplished.

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

143

144 **2 Methods**

145 **2.1 Sampling**

146 Three large volume mesocosms (50 m³), based on the design described in Guieu
147 et al. (2010, 2014), were deployed at the exit of the Noumea lagoon (22°29.073 S -
148 166°26.905 E), 28 km off the coast of New Caledonia on January 13, 2013. Detailed
149 descriptions of the mesocosm site selection, deployment and sampling strategy are
150 provided in [Bonnet et al. \(2015a\)](#). Large volume samples (50 L) were collected from 1,
151 6, and 12 m depths from each mesocosm and outside the mesocosms (hereafter called
152 Noumea lagoon) once per day at 7 a.m. using a Teflon® PFA pump and PVC tubing.
153 This daily sample was immediately transferred back to laboratories aboard the R/V Alis,
154 and subsampled for a suite of stock and flux measurements. Samples for DNA analysis

155 were immediately filtered onto 25 mm 0.2 µm Supor® filters (Millipore, Billarica, CA),
156 using gentle peristaltic pumping. All filters were flash frozen in liquid nitrogen, and
157 stored at -80°C until shipment on dry ice from New Caledonia to the University of
158 California, Santa Cruz.

159

160 **2.2 Determining targeted diazotroph abundances and net growth rates** 161 **using the quantitative polymerase chain reaction (qPCR)**

162 DNA was extracted using a Qiagen DNeasy Plant kit (Valencia, CA), with
163 modifications to the protocol optimized to recover high quality DNA from cyanobacteria,
164 including additional cell lysis steps of freeze/thaw cycles, agitation using a bead beater,
165 as well as a proteinase K digestion (Moisander et al., 2008). The quality of DNA extracts
166 was evaluated using a NanoDrop (Thermo Scientific, Waltham, MA), and concentrations
167 were determined using a PicoGreen® dsDNA Quantitation Kit (Molecular Probes,
168 Eugene, OR, USA), according to manufacturer's guidelines.

169 Nine diazotrophic phylotypes were quantified in samples from the mesocosms
170 and the Noumea lagoon, using quantitative PCR (qPCR), using Taqman® assays widely
171 used in the marine environment. This study targeted representatives from known major
172 marine diazotroph lineages: two unicellular diazotrophic symbionts of different
173 *Braarudosphaera bigelowii* strains, UCYN-A1 (Church et al., 2005a), UCYN-A2
174 (Thompson et al., 2014); two free-living unicellular diazotroph cyanobacterial phylotypes
175 UCYN-B (*Crocospaera* sp.; Moisander et al., 2010), and UCYN-C (Foster et al., 2007);
176 filamentous, colonial, non-heterocyst forming *Trichodesmium* spp. (Church et al., 2005a);
177 three diatom-diazotroph associations (DDAs), *Richelia* associated with *Rhizosolenia*
178 (Het-1; Church et al., 2005b), *Richelia* associated with *Hemiaulus* (Het-2; Foster et al.,
179 2007), and *Calothrix* associated with *Chaetoceros* (Het-3; Foster et al., 2007); and a
180 widespread γ-proteobacterial phylotype γ-24774A11 (Moisander et al., 2008).

181 Recombinant plasmids with the targeted organism *nifH* fragment were used as
182 qPCR standards, and each 96 well plate was run with a serial dilution (10^0 - 10^7 *nifH*
183 copies reaction⁻¹) of the appropriate standard. All qPCR reactions were set up with 1 µL
184 of DNA extract in a 15 µL volume with the following reagents and final concentrations:
185 1X Taqman® Master Mix (Applied Biosystems, Carlsbad, CA, USA), 0.4 µM each

186 forward and reverse primers, and 0.2 μM probe (5'-FAM and 3'-TAMRA labeled).
187 Thermocycle parameters were as described in Goebel et al. (2010) for all assays, with the
188 exception of using a 64°C annealing temperature for UCYN-A2. The qPCR reaction
189 efficiencies were as follows: UCYN-A1 – $96.7 \pm 2.8\%$; UCYN-A2 – $94.9 \pm 3.4\%$;
190 UCYN-B – $101.3 \pm 2.2\%$; UCYN-C – $95.4 \pm 2.9\%$; γ -24774A11 – $96.2 \pm 3.2\%$;
191 *Trichodesmium* – $93.6 \pm 1.4\%$; Het-1 – $96.1 \pm 4.7\%$; Het-2 – $98.3 \pm 2.0\%$; and Het-3 –
192 $97.5 \pm 0.9\%$. Based on the differences in sample volumes, the limits of detection (LOD)
193 and quantification (LOQ) for all qPCR assays ranged between 11-56 *nifH* copies L^{-1} , and
194 87-444 *nifH* copies L^{-1} , respectively. Samples were determined to be ‘detected, not
195 quantified’ (DNQ) when calculated abundances were greater than the LOD, but less than
196 the LOQ. Abundances are reported as *nifH* copies L^{-1} , rather than cells L^{-1} because there
197 is currently little information about the number of *nifH* copies per genome in these
198 diazotroph targets.

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

206 Growth and mortality rates were calculated for individual diazotrophs inside the
207 mesocosms when abundances were higher than the LOQ on two consecutive sampling
208 days, as described in Moisaner et al. (2012), using the following formula: $k =$
209 $2.303 * (\log_{10}(N_{t2} - N_{t1})) / (t_2 - t_1)$, where N_x = abundance at time x . This assumes that the
210 organisms were growing exponentially during the experiment, which cannot be easily
211 verified in field populations. These rates implicitly include grazing, mortality and viral
212 lysis, and thus are net growth and mortality rates.

213

214 **2.3 Determination of diazotroph community composition using high** 215 **throughput sequencing of *nifH* amplicons**

216 In order to evaluate whether the diazotrophs targeted via qPCR assays were
217 representative of the phylotypes present in the lagoon and mesocosms, partial *nifH*
218 fragments (ca. 360 base pairs) were amplified using a nested PCR assay and universal
219 *nifH* primers *nifH1-4*, as described in Turk-Kubo et al., (2014). Two lagoon samples (day
220 1 and day 22), and three samples from each mesocosm (day 3, day 23, and the day where
221 UCYN-C abundances were beginning to increase, i.e. days 11-15, see discussion below)
222 were chosen for analysis. For each sample, triplicate PCR reactions were pooled. Internal
223 primers were modified 5' common sequence (CS) linkers (CS1_*nifH1F*: 5'-
224 ACACTGACGACATGGTTCTACATGYGAYCCNAARGCNGA, CS2_*nifH2R*: 5'-
225 TACGGTAGCAGAGACTTGGTCTADNGCCATCATYTCNCC) to facilitate library
226 preparation at the DNA Services (DNAS) Facility at the University of Illinois, Chicago,
227 using the targeted amplicons sequencing (TAS) approach described in Green et al.
228 (2015). These libraries were pooled with other libraries to achieve a target depth of ca.
229 40,000 sequences per sample. Sequencing of paired end reads was performed at the W.M.
230 Keck Center for Comparative and Functional Genomics at the University of Illinois at
231 Urbana-Champaign using Illumina MiSeq technology. De-multiplexed raw paired end
232 reads were merged in CLC Genomics workbench, and merged reads between 300-400
233 base pairs in length were selected. Quality filtering was performed in QIIME (Caporaso
234 et al., 2010) using the *usearch* quality filter (*usearch_qf*) pipeline script, which includes
235 steps for denoising, de novo chimera removal using UCHIME (Edgar et al., 2011) and
236 operational taxonomic unit (OTU) determination using *usearch6.1* at 97% nucleotide
237 identity (Edgar 2010). Representative nucleotide sequences from OTUs with greater than
238 100 reads (277 out of 2325 OTUs, representing 92% of all sequences that passed the
239 *usearch* quality filter) were imported into ARB (Ludwig et al., 2004), translated into
240 protein sequences, where non-*nifH* OTUs or those with frameshifts were discarded.
241 QIIME script *exclude_seqs_by_blast.py* was used to check for sequences with >92%
242 amino acid identity to known contaminants; none were found. OTUs targeted by each
243 qPCR assay was determined in silico for each group of diazotrophs in ARB by
244 identifying representative sequences that had 0-2 mismatches in either primer or the
245 probe binding region, without exceeding a total of 4 mismatches in all three regions (see
246 Supplement Table S4).

247 For the characterization of the overall diazotroph community composition in the
248 lagoon and mesocosms, representative sequences from the most highly recovered OTUs
249 (109 OTUs representing 85% of all post quality-filtering sequences) were considered.
250 Translated amino acid sequences were aligned to the existing amino acid alignment in the
251 curated database. Maximum likelihood trees were calculated using translated amino acid
252 sequences from representative sequences and their closest relatives (determined via
253 blastp) in MEGA 6.06 (Tamura et al. 2013), using the JTT matrix based model and
254 bootstrapped with 1000 replicate trees. Distribution of read data across samples for each
255 representative sequence was visualized in the Interactive Tree of Life online tool (Letunic
256 and Bork, 2006).

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

260

261 **2.4 PCR amplification of *Cyanothece*-like organisms**

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

271 Duplicate PCR reactions were carried out in 20 μ L volumes with 1X Platinum®
272 Taq PCR buffer (Invitrogen, Carlsbad, CA), 3.0 mM MgCl₂, 400 μ M dNTP mix, 0.2 μ M
273 of each forward and reverse primers, 1 U Platinum® Taq polymerase (Invitrogen), and 2
274 μ L of DNA extract. Thermocycle parameters were as follows: the initial denaturation
275 step at 94°C for 5 min was followed by 30 cycles of 94°C for 30 sec, 58°C for 30 sec, and
276 72°C for 30 sec and a final elongation step at 72°C for 10 min. PCR amplicons were
277 cleaned using the QiaQuick Gel Extraction Kit (Qiagen), and sequenced directly using

278 Sanger technology at the UC Berkeley DNA Sequencing Center using
279 *Cyanothece_nifH_R* to prime the sequencing reaction.

280 Raw sequences were processed using Sequencher 5.2.4 (Gene Codes Corporation,
281 Ann Arbor, MI) and phylogenetic analyses were conducted in ARB (Ludwig et al.,
282 2004), using a curated database of all *nifH* sequences available in Genbank (Heller et al.,
283 2014) and in MEGA 6.06 (Tamura et al., 2011). [Sequences were submitted to NCBI's](#)
284 [Genbank database under accession numbers KU160532-KU160537.](#)

285

286 **2.5 Statistical analyses**

287 Pearson correlation coefficients were calculated for each paired variable, after
288 transformation to assure normality, and samples collected from the incubations and the
289 Noumea lagoon were treated independently. Supplement Tables S1 and S2 detail linear
290 associations and transformations required for each variable inside and outside the
291 mesocosms, respectively. Diazotroph abundance and physico-chemical environmental
292 data was analyzed in R (Team, 2012). Biplot from principal component analysis was used
293 to examine the temporal and spatial variation in environmental parameters.

294 Correspondence analysis of the abundance data in each mesocosm separately was
295 performed with the *ca* package (Nenadic and Greenacre, 2007). Biplots of sampling days
296 (rows) and organisms (columns) were compared to examine reproducibility across
297 mesocosms. Least squares regression models were calculated to fit and predict the effect
298 of temperature, salinity, and PO₄ on diazotroph abundance.

299

300 **3 Results and Discussion**

301 **3.1 Diazotroph community structure in VAHINE mesocosms experiment** 302 **and in the Noumea Lagoon**

303 In order to 1) characterize the diazotroph community composition in the
304 mesocosms and the lagoon using a qualitative approach that is complementary to qPCR
305 and 2) evaluate whether the qPCR assays used in this study, which are widely used in
306 studies of the oligotrophic ocean, target the major lineages of diazotrophs in the Noumea
307 Lagoon, fragments of the *nifH* gene were amplified from a subset of samples using a

308 well-established nested PCR approach (Zehr and Turner, 2001), and sequenced using
309 Illumina MiSeq technology.

310 In total, 636,848 paired end reads were recovered from 11 samples, and 544,209
311 passed the quality filtering steps described above (Supplemental Table S4). Clustering at
312 97% nucleotide identity yielded 2,325 OTUs with greater than 4 reads, and 334 OTUs
313 with greater than 100 reads. 277 of these OTUs (representing 479,402 reads or 88% of all
314 reads that passed quality filtering; Supplemental Table S4) remained after removing non-
315 *nifH* reads or those with frameshifts, and were used in downstream analyses. A majority
316 of these 277 OTUs (152 OTUs, representing 311,550 reads, or 65% of the reads selected
317 for analysis) were affiliated with *nifH* cluster 1B Cyanobacteria. Reads affiliated with
318 *nifH* cluster 1G, which is composed primarily of γ -proteobacterial phylotypes, was the
319 second most highly recovered group (88 OTUs, representing 120,586 reads, or 25.2% of
320 the reads selected for analysis). Reads that were closely related to *nifH* cluster 1J/1K,
321 comprised primarily of α - and β -proteobacteria, were also recovered, but only comprised
322 8.4% of the reads selected for analysis (19 OTUs, 40,032 reads). Cluster 3 and cluster 1O
323 affiliated reads were recovered, but together accounted for less than 2% of the reads
324 selected for analysis (Table 1).

325 The two OTUs with the highest relative abundance (OTU1890 and OTU2317),
326 accounted for 31% of the reads selected for analysis and were closely related to the
327 prymnesiophyte symbiont UCYN-A2 (Supplemental Table S5). Both OTUs were
328 present in the lagoon in day 1 and day 22 samples, and were recovered at high relative
329 abundances from all three mesocosms throughout the experiment (Fig. 1). The third most
330 highly recovered OTU (OTU1) was a γ -proteobacteria closely related to γ -24774A11, a
331 heterotrophic diazotroph with widespread occurrence (Moisander et al., 2014; Langlois et
332 al., 2015), that is also preferentially amplified by the *nifH* primers used (Turk et al.,
333 2011). This sequence type was present in the lagoon samples, and had high relative
334 abundances in all three mesocosms in midpoint samplings (days 11, 13, and 15 for M1,
335 M2, and M3, respectively; Fig. 1). OTU2280 (cluster 1J/1K) was the OTU with the
336 fourth highest relative abundance. It does not have high sequence similarity to any
337 uncultivated or cultivated organisms, with the closest relative, an uncultivated
338 rhizosphere isolate (Genbank accession no. KC667160), sharing only 86% nucleotide

339 sequence similarity. This is true for all but one of the 1J/1K OTUs, OTU119, which is
340 closely related (98% nucleotide identity) to an environmental sequence recovered from
341 Heron Reef (Genbank accession no. EF175779).

342 Also among OTUs that had high recovery were UCYN-A1 (OTU2008,
343 OTU1754, and OTU2), other UCYN-A2 OTUs (OTU2325, OTU1548, and OTU664),
344 *Trichodesmium* (OTU5 and OTU35), UCYN-C (OTU12) and two OTUs affiliated with
345 cluster 1G (OTU2218 and OTU2199) (Fig.1 and Supplemental Table S5). 1G OTUs are
346 recovered from the lagoon sample at day 22, and have high relative abundances in all
347 three mesocosms by the end of the experiment. These sequence types are not closely
348 related to γ -24774A11, thus are not quantified using qPCR assays in the study, and their
349 quantitative importance in the mesocosm environment cannot be determined based on
350 this qualitative measure. It is important to note that high relative abundances in PCR-
351 based libraries is not indicative of high abundances, as often these sequence types
352 dominate PCR libraries yet are present at low abundances in the environment (Hewson et
353 al., 2007, Bonnet et al., 2013, Turk-Kubo et al., 2014, Shizoaki et al., 2014, Bentzon-Tilia
354 et al., 2015), presumably as a result of preferential amplification (Turk et al., 2011).

355 A majority of the OTUs with high relative abundances (159 out of 277,
356 representing 380,556 out of 479,402 reads) affiliated with the following lineages targeted
357 by qPCR assays used in this study: UCYN-A1, UCYN-A2, UCYN-B, UCYN-C,
358 *Trichodesmium*, and *Richelia* associated with *Rhizosolenia* (Het-1), and γ -24774A11. To
359 determine whether the qPCR assays used would target the diazotrophs present,
360 representative sequences were identified that contained between 0-2 mismatches in the
361 primer and probe binding regions, without exceeding a total of 4 mismatches for all three
362 regions. For UCYN-A, *Trichodesmium*, Het-1, and γ -24774A11 lineages, nearly all of
363 the sequence types present met these criteria (between 97-100%), thus would be
364 quantified in the qPCR assays (Table 1). Only 26% of the recovered *Cyanothece*-like
365 sequences would successfully be targeted by the UCYN-C assay, however, this coverage
366 increases to 85% when including two of the most highly recovered OTUs that have a
367 third mismatch at the 5' end of the probe binding region, thus are still likely to be
368 quantified. Together with the results of UCYN-C specific PCRs (see section 3.3 below),

369 these results indicate that qPCR assays targeting UCYN-C likely quantified most of the
370 phylotypes present, but absolute abundances may be underestimated.

371 A group of 32 OTUs clustered within the *Chroococcales* and were closely related
372 to both *Crocospaera watsonii* 8501 (93-97% amino acid identity) and *Cyanothece* sp.
373 WH8904 (93-97% amino acid identity). These OTUs were not targeted by the UCYN-B
374 qPCR assay, with 4-8 total mismatches in the primer/probe binding regions. So despite
375 being considered putative *Crocospaera* sp. for the purpose of this study, it is possible
376 that they are actually from a *Cyanothece*-like organism (not targeted by the UCYN-C
377 qPCR assay) or another member of the *Chroococcales*. These OTUs were most highly
378 recovered at the final sampling day (day 23) in M2 and M3, but were not recovered in the
379 lagoon libraries.

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

389

390 **3.2 Most abundant diazotrophs in VAHINE mesocosms experiment shift** 391 **from *Richelia* to *Cyanothece*-like ecotypes**

392 During the VAHINE mesocosm experiment, three periods could be roughly
393 defined based on the identity of the most abundant diazotroph and biogeochemical
394 parameters (described in detail in Berthelot et al., 2015). A diazotroph community similar
395 to the Noumea lagoon field community and low DIP concentrations characterized the
396 period prior to the DIP spike (hereafter referred to as P0). The second period, days 5-14
397 (hereafter referred as P1), was characterized by high abundances of Het-1, high DIP
398 availability, as well as moderate N₂ fixation rates (10.1±1.3 nmol N L⁻¹ d⁻¹; Bonnet et al.,
399 2015b). In the second half of the experiment, days 15-23 (hereafter referred as P2), the

400 UCYN-C population grew to high abundances, and was characterized by low DIP
401 availability, and high N₂ fixation rates (27.3±1.0 nmol N L⁻¹ d⁻¹; Bonnet et al., 2015b),
402 with rates reaching > 60 nmol N L⁻¹ d⁻¹, ranking among the highest reported in marine
403 waters (Luo et al., 2012).

404 At the initiation of the mesocosm experiment, the most abundant diazotrophs in
405 the Noumea lagoon were Het-1 (3.1 x 10⁴ *nifH* copies L⁻¹), and *Richelia* associated with
406 *Hemiaulus* (Het-2; 1.2x10⁴ *nifH* copies L⁻¹), as well as the *B. bigelowii*-associated
407 UCYN-A2 (1.5 x 10⁴ *nifH* copies L⁻¹) and UCYN-A1 (5.6 x 10³ *nifH* copies L⁻¹).
408 *Trichodesmium* and the uncultivated unicellular cyanobacterial group C (UCYN-C) were
409 present but at lower abundances, 2.8x10³ *nifH* copies L⁻¹ and 7.8x10² *nifH* copies L⁻¹,
410 respectively. γ -24774A11 and *Calothrix* associated with *Chaetoceros* (Het-3) were also
411 present but at abundances too low to quantify (DNQ), and no *Crocospaera* sp. (UCYN-
412 B) could be detected (Supplement Table S3).

413 During P0, evidence of shifts from the original Noumea lagoon diazotroph
414 assemblage could be observed in all three mesocosms (Fig. 2b-e). In M1, M2 and M3,
415 Het-1 remained the most abundant diazotroph, and the relative proportions of Het-2 and
416 UCYN-A1 were similar to those measured outside the mesocosms. However, UCYN-A2
417 abundances decreased and *Trichodesmium* abundances increased over this three-day
418 period with respect to their abundances in the lagoon. These changes were most
419 pronounced in M3. As in the lagoon, UCYN-C and γ -24774A11 were present at low
420 abundances (ca. 10² *nifH* copies L⁻¹) in the day 3 sampling.

421 Diazotroph assemblages within each mesocosm remained relatively consistent
422 throughout the first thirteen days of the experiment, with Het-1 the most abundant
423 diazotroph (Fig. 2b-d). However, a significant shift in community composition began
424 about halfway through the 23-day experiment, evidenced by the increasing abundances of
425 UCYN-C and Het-3, and decreasing abundances of Het-1 and UCYN-A1 (Fig. 2b-d and
426 3). Increasing UCYN-C abundances were first seen in M1 (day 11), then M2 (day 13),
427 followed by M3 (day 15), which reflected the time in each mesocosm when the DIP
428 turnover time dropped below 1 d⁻¹, signaling DIP limitation (Berthelot et al.,
429 2015; Moutin et al., 2005). In all three mesocosms, which acted as biological replicates,
430 UCYN-C abundances were not only strongly correlated with the decreasing DIP

431 concentrations ($p=0.004$, $r=-0.51$), but also increasing temperatures in the mesocosms
432 ($p=0.000$, $r=+0.86$) and could be weakly correlated with increasing salinity
433 ($p=0.000$, $r=+0.56$), decreasing Het-1 ($p=0.002$, $r=-0.54$), Het-2 ($p=0.01$, $r=-0.37$) and
434 UCYN-A1 ($p=0.000$, $r=-0.64$) abundances. A similar correlation between UCYN-C
435 abundance and both increasing salinity and temperature was observed outside in the
436 Noumea Lagoon (although UCYN-C abundances in the lagoon remained low; see section
437 3.4). Salinity in each of the three mesocosms and in the lagoon did increase gradually
438 during the experimental period, but were elevated inside the mesocosms in P2, likely due
439 to evaporative loss (Bonnet et al., 2015a). Together, this data suggests that temperatures
440 below 25.6°C are not optimal temperatures for this group, and that it may tolerate slightly
441 elevated salinities better than other diazotrophs present. This is consistent with the
442 warmer temperatures required for growth for the UCYN-C isolate TW3 (Taniuchi et al.,
443 2012), and the occurrence of the UCYN-C group in low latitude, warm waters ($> 28^{\circ}\text{C}$)
444 in the Tropical North Atlantic (Langlois et al., 2008). However, changes in temperature
445 and salinity alone cannot explain the high abundances of UCYN-C in these mesocosms,
446 as similar increases were recorded in the lagoon without a corresponding increase in
447 UCYN-C abundances, and other diazotrophs also can grow optimally at these
448 temperatures (e.g. *Crocospaera* sp. (Webb et al., 2009) and *Trichodesmium* sp.
449 (Breitbarth et al., 2007)).

450 Het-3 was undetected in all mesocosms until day 7, and then remained at low
451 abundances (usually below quantitation limits) until day fifteen of the incubation. After
452 day 15, Het-3 abundances were of similar order of magnitude to their abundances in the
453 Noumea Lagoon (ca. 10^2 - 10^3 *nifH* copies L^{-1}). Increases in Het-3 abundances after day
454 15 could be strongly correlated to increases in temperature ($p=0.000$, $r=+0.82$) as well as
455 salinity ($p=0.000$, $r=+0.78$), total chl a ($p=0.000$, $r=+0.71$), and UCYN-C abundances
456 ($p=0.000$, $r=+0.77$). Het-3 abundances could also be weakly correlated to increases in
457 bulk N_2 fixation rates ($p=0.01$, $r=+0.49$), and decreases in PO_4 concentrations ($p=0.003$,
458 $r=-0.54$), UCYN-A1 ($p=0.000$, $r=-0.64$) and Het-1 ($p=0.03$, $r=-0.38$) abundances. Het-3
459 abundances were never greater than 5.8×10^3 *nifH* copies L^{-1} , however their distribution
460 throughout the water column (Supplement Fig. S1) and their recovery in the sediment
461 traps (Bonnet et al., 2015b) suggests that these diazotrophs are sinking out of the water

462 column, and could possibly play a role in supplying fixed N to sediments in this shallow
463 coastal system.

464 Both UCYN-A2 and UCYN-A1 abundances peaked early in the mesocosms, at
465 days 5 and 7, respectively. UCYN-A2 abundances were not strongly correlated to any
466 environmental factors, and were only weakly correlated to bulk N₂ fixation rates
467 ($p=0.008$, $r=+0.49$) and N₂ fixation rates associated with the <10 μm size fraction
468 ($p=0.045$, $r=+0.38$). However, decreasing UCYN-A1 abundances could be strongly
469 correlated to temperature increases ($p=0.000$, $r=-0.75$), and decreasing PO₄
470 concentrations ($p=0.000$, $r=+0.69$). This implies that the *B. bigelowii*/UCYN-A1
471 association benefitted either directly or indirectly from the day 4 introduction of PO₄ in
472 the early part of the experiment. There is evidence that UCYN-A1 *nifH* transcription (a
473 proxy for active N₂ fixation) is limited by the availability of inorganic P (Turk-Kubo et
474 al., 2012). UCYN-A1 is unable to directly utilize components of the dissolved organic
475 phosphate (DOP) pool, such as phosphoesters and phosphonates (Tripp et al., 2010), but
476 nothing is known about the capability of the *B. bigelowii* host to utilize DOP substrates.
477 The correlation between UCYN-A1 abundances (thus assumed correlation between the
478 symbiosis as a whole) and inorganic P concentrations in the VAHINE mesocosms
479 provides further evidence that inorganic P may be the preferred P source for the *B.*
480 *bigelowii* host.

481 Multiple regression models underscore the importance of a small number of
482 environmental factors in the abundances of three of the diazotrophs targeted. The log
483 abundances of UCYN-A1, UCYN-C and Het-3 in all three mesocosms and in the lagoon
484 can be modeled well using only the temperature, salinity, and PO₄ data, with R² values
485 between 0.77-0.85, and R²-cv between 0.60-0.76 (Table 2), when the sample location
486 (M1-M3 and outside) are included as a variable. Although the three-predictor model
487 provides the highest quality fit to the data, the goodness of prediction values (R²-cv) are
488 comparable between models that use all three predictors versus models using just T alone
489 in the case of UCYN-A1 and UCYN-C, and just PO₄ alone for all three diazotrophs
490 (Table 2). These results from the linear regression model imply that the most important
491 environmental factor best correlated with the dramatic increase in UCYN-C and Het-3
492 abundances was the decreasing concentration of PO₄.

493 Considering that UCYN-C maintained low population abundances in the Noumea
494 Lagoon during the time of the mesocosm experiment, [despite changes in temperature and](#)
495 [salinity that mirror changes in the mesocosms \(Bonnet et al., 2015a\)](#), it follows that
496 UCYN-C may have benefitted either indirectly from the DIP fertilization, or directly
497 from a physical aspect of the mesocosms themselves. A biofilm had accumulated on the
498 sides of the mesocosm bags, and although this was not sampled for molecular analysis, it
499 is possible that the UCYN-C ecotype was a component of this biofilm community, thus
500 dependent on a physical environment not representative of water column conditions.
501 [Another explanation may be that the decreased turbulence in the mesocosm environment](#)
502 [created favorable conditions for this ecotype.](#) Finally, it is possible that the inverse
503 correlation between UCYN-C abundances and DIP concentrations may be a result of
504 UCYN-C being able to outcompete other diazotrophs for organic P substrates, under low
505 DIP conditions. *Cyanothece* sp. strains PCC 8801 and 8802 have genes used in
506 phosphonate metabolism (Bandyopadhyay et al., 2011), which strongly implies that some
507 strains are able to use this organic P substrate to meet cellular P requirements. The strain
508 most closely related to the UCYN-C ecotype, *Cyanothece* sp. CCY0100, also has genes
509 for phosphonate metabolism and transport (JGI website). It is evident that some
510 component of the microbial community in the mesocosms was utilizing DOP, [as DOP](#)
511 [stocks were drawn down to levels seen in the lagoon \(\$0.105 \pm 0.011 \mu\text{mol L}^{-1}\$ \)](#) in the
512 second half of the experiment (Berthelot et al., 2015), however, UCYN-C abundances
513 could not be correlated to this drawdown (data not shown). *Trichodesmium erythraeum*
514 IMS101 and *Calothrix rhizosoleniae* SC01 (Het-3) are the only two other diazotrophs
515 targeted in this study known to possess the metabolic capability to utilize phosphonates
516 (Dyhrman et al., 2006; Hilton et al., 2013). *Richelia* associated with *Hemiaulus* (Het-2) do
517 not have any genes for the metabolism of phosphonates (Hilton et al., 2013), but it
518 remains unclear whether the symbiont of *Rhizosolenia* (Het-1) is able to use
519 phosphonates, as four genes related to phosphonate metabolism were identified in the
520 genome (Hilton, 2014). However, as with UCYN-C, no correlation between DOP
521 concentration and *Trichodesmium*, Het-1 or Het-3 abundances were seen (data not
522 shown). [This may be in part due to shortcomings in our understanding of the chemical](#)

523 composition of the DOP pool (Dyhrman et al., 2007), and which organic P compounds
524 are bioavailable to which organisms.

525

526 **3.3 Phylogenetic identity of diazotrophs targeted with UCYN-C qPCR assay**

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

545 It is important to note that the sequences amplified using this *Cyanothece*-specific
546 PCR are targeted by the UCYN-C qPCR assay, and the phylotypes with 3 mismatches in
547 the probe binding region were not recovered with this assay. This may be due to
548 differences primer efficiency, the number of amplification steps, and/or depth of
549 sequencing, but these results lend further support that important UCYN-C phylotypes
550 were quantified with qPCR assays.

551

552 **3.4 *In situ* net growth and mortality rates**

553 The VAHINE mesocosm experiment provided a rare opportunity to repeatedly
554 sample the same water mass for an extended period of time, thus the ability to empirically
555 determine *in situ* net growth or mortality rates for individual diazotroph phylotypes,
556 based on the change in *nifH* gene copies L⁻¹ between sampling days. Growth (and
557 mortality) rates are critical input parameters for mathematical models of oceanic N
558 budgets. Culture-based rates or estimates are often employed because there have been so
559 few measurements of *in situ* rates from natural populations of diazotrophs, and some
560 species such as UCYN-A remain uncultivated. However, due to the lack of competition
561 and grazing in culture, these rates may be overestimated compared to *in situ* rates.

562 Surprisingly, maximum net growth rates for each diazotroph phylotype were among
563 the highest reported for oceanic diazotrophic organisms (even when considering culture-
564 based growth rates), yet showed substantial variability between mesocosms, both in terms
565 of the absolute rates and the patterns of growth and mortality across time (Table 3). For
566 example, maximum growth rates for UCYN-A2 ranged between 0.52 and 1.71 d⁻¹, but
567 occurred early in the experiment in M3 (day 5), in the middle of the experiment in M2
568 (day 11) and at the end of the experiment in M1 (day 20). The two phylotypes with
569 consistent timing of maximum growth rates across mesocosms were UCYN-C and Het-3.
570 UCYN-C had the highest maximum growth rates of all phylotypes, which ranged
571 between 1.23 - 2.16 d⁻¹ and occurred within a four-day period (day 11-15) in all
572 mesocosms. Het-3, which was virtually absent for the first half of the experiment in all
573 three mesocosms, had maximum net growth rates between 0.57 - 1.09 day⁻¹, that occurred
574 within a five day period (day 15 – 20) in all mesocosms.

575 Moisander et al. (2012) reported maximum net growth rates from nutrient amendment
576 experiments conducted in the South Western Pacific close to New Caledonia for UCYN-
577 A1 (0.19 d⁻¹), UCYN-B (0.61 d⁻¹) and γ -24774A11 (0.52 d⁻¹). The maximum net growth
578 rates calculated for these phylotypes during the VAHINE project were considerably
579 higher, at 0.73 d⁻¹, 1.38 d⁻¹, and 1.07 d⁻¹ for UCYN-A1, UCYN-B and γ -24774A11,
580 respectively. These results are unexpected considering that the rates determined by
581 Moisander et al. (2012) were from a series of nutrient amendment incubations in 4.5-L
582 bottles, where presumably favorable conditions for the growth of diazotrophs were
583 present.

584 The maximum net growth rates determined *in situ* for UCYN-A2 and UCYN-C were
585 among the highest measured at 1.71 d^{-1} , and 2.16 d^{-1} , respectively, and represent the only
586 reported growth rates for these uncultivated diazotrophs. Very little is known about the
587 newly described uncultivated *B. bigelowii*/UCYN-A2 association, but the difference
588 between UCYN-A1 and UCYN-A2 net growth rates and their patterns within the same
589 mesocosm indicate that the growth rates of each association are likely dependent upon
590 different environmental variables. It also seems plausible, due to the [difference in size](#)
591 [between the two *B. bigelowii* hosts \(Thompson et al., 2014\), that they are grazed by](#)
592 [different protists.](#)

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

604 Maximum net growth rates for the filamentous diazotrophs were also generally higher
605 than previously reported growth rates. *Trichodesmium* sp. maximum net growth rates for
606 were as high as 1.46 ± 0.05 (M1) and $1.55 \pm 0.02 \text{ d}^{-1}$ (M3), and microscopic analyses
607 indicated (data not shown) that both *T. erythraeum* and *T. thiebautii* were present in the
608 mesocosms, the former being the dominant *Trichodesmium* species. These calculated
609 rates are much higher than the specific growth rate previously reported for cultures of *T.*
610 *erythraeum* ($0.29 \pm 0.04 \text{ d}^{-1}$; Hutchins et al., 2007) and net growth rates for *T. thiebautii*
611 populations in the Noumea lagoon ($0.11 - 0.38 \text{ d}^{-1}$; Rodier and Le Borgne, 2008).
612 Maximum growth rates for the DDAs Het-1, Het-2, and Het-3 were 1.28 d^{-1} , 2.24 d^{-1} , and
613 1.09 d^{-1} , respectively. Although these are the first reported growth rates for these DDAs
614 determined using *in situ* cellular abundances, Foster et al. (2011) determined maximum

615 net growth rates to be much lower when using nanoSIMS-based techniques to quantify
616 ¹⁵N incorporation in biomass, which is to be expected if the DDAs are using other N
617 sources (eg. dissolved organic N and/or recycled N) in addition to fixing N.

618 For many diazotrophs, the pattern of net growth and mortality rates indicated a very
619 dynamic process that appeared specific to each mesocosm. Most diazotrophs experienced
620 multiple transitions between net growth and net mortality within a single mesocosm
621 throughout the 23-day incubation (Table 3). For example, in M1, net growth rates for
622 *Trichodesmium* were observed approximately every other sampling period (on days 9, 13,
623 15, 19, and 23), and in some intervening periods experienced mortality rates similar in
624 magnitude to growth rates (e.g. day 7 and 11). In contrast, net growth occurred on days 7,
625 11, 13, 20, and 23 in M3, and one of the largest net mortality rates was measured at day 9
626 (-2.16 +/- 0.18 d⁻¹). There are no aspects of the experimental design that can be invoked
627 to explain this variability; in fact biogeochemical parameters and picoplankton population
628 dynamics were well replicated in all three mesocosms (Bonnet et al., 2015a). Therefore
629 the dynamic nature of diazotroph growth and mortality rates in each mesocosm most
630 likely results from a combination of grazing pressure and viral lysis, which can be
631 expected to reflect natural variations in the grazers and virus present.

632

633 **3.5 Symbiotic *Richelia* and UCYN-A ecotypes are abundant in the Noumea** 634 **lagoon during the VAHINE mesocosms experiment**

635 The VAHINE mesocosm experiment provided an additional opportunity to
636 characterize the abundances of targeted diazotrophs in the Noumea lagoon using
637 quantitative techniques. As described above, the diazotroph assemblage at the onset of
638 the VAHINE mesocosms experiment was dominated by Het-1, Het-2, UCYN-A2 and
639 UCYN-A1 (see section 3.1). Het-1 remained the most abundant of the targeted
640 diazotrophs for most of the sampling days ($3.1 \times 10^4 - 5.5 \times 10^5$ *nifH* copies L⁻¹) during the
641 period of study (Fig. 2e, Supplement Table S3). Het-2 abundances were consistently
642 lower than Het-1, and variable throughout the period of study, ranging between $4.0 \times 10^3 -$
643 1.6×10^4 *nifH* copies L⁻¹, and reaching peak abundances on day 10 and again on day 18
644 (Supplement Table S3). Bonnet et al. (2015a) discusses the BNF rates in detail, however,

645 it should be noted that in the lagoon only Het-2 abundances had significantly strong
646 positive correlation with bulk N₂ fixation rates ($p=0.01$, $r=+0.78$).

647 Although *Richelia* living in association with *Rhizosolenia* (Het-1) and *Hemiaulus*
648 (Het-2) are known to be widely distributed and potentially significant N₂-fixers in
649 oligotrophic oceans (Carpenter et al., 1999;Subramaniam et al., 2008;Karl et al., 2012),
650 reports of their presence in coastal waters in the Southwest Pacific are rare, and previous
651 studies during the austral summer in the Noumea Lagoon reported very low abundances
652 of heterocystous symbionts (Rodier and Le Borgne, 2010;Biegala and Raimbault, 2008).
653 This is the first report on quantitative abundances of *Richelia* in the Noumea lagoon,
654 which are comparable to abundances reported from the Tropical North Atlantic (Foster et
655 al., 2007;Goebel et al., 2010) and the North Pacific Subtropical Gyre (Church et al.,
656 2008;Foster and Zehr, 2006), where the DDAs are thought to have a significant impact on
657 C sequestration due to their productivity and rapid sinking (Subramaniam et al.,
658 2008;Karl et al., 2012;Villareal et al., 2012), and may play a role in supporting the
659 benthic community (Houlbreque and Ferrier - Pagès, 2009).

660 UCYN-A2 was the second most abundant member of the **targeted** diazotrophic
661 **assemblage** in the lagoon for the first 10 days of the experiment (Fig. 2e), and was
662 present at abundances as high as 1.1×10^5 *nifH* copies L⁻¹ (day 6; Supplement Table S3).
663 UCYN-A2 abundances declined steadily throughout the period of the 23-day experiment
664 ($p=0.04$, $r=-0.63$), and had a significantly strong negative correlation to sea temperature
665 ($p=0.003$, $r=-0.82$) and salinity ($p=0.03$, $r=-0.89$), both of which were increasing in
666 Noumea lagoon throughout the VAHINE deployment (Bonnet et al., 2015b). UCYN-A1
667 abundances were consistently lower than those of UCYN-A2, ranging between 2.8×10^3 –
668 6.4×10^4 *nifH* copies L⁻¹ with peaks in abundance at day 6 and again at day 22
669 (Supplement Table S3). Interestingly, UCYN-A1 abundances had significant positive
670 correlation to total chl *a* in the > 10 μm size fraction ($p=0.03$, $r=+0.72$). The significant
671 inverse relationship between UCYN-A2 abundances and water temperatures in the
672 Noumea lagoon, suggests that the *B. bigelowii*/UCYN-A2 association may thrive at lower
673 temperatures than other diazotrophs (e.g. *Trichodesmium*), similar to the *B.*
674 *bigelowii*/UCYN-A1 association (Moisander et al., 2010).

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

685 *Trichodesmium* spp. have been routinely observed in the Noumea Lagoon using
686 satellite observations (e.g. Dupouy et al., 2000; Dupouy et al., 2011) and direct field
687 measurements (Renaud et al., 2005; Masotti et al., 2007; Rodier and Le Borgne, 2010,
688 2008). Conspicuous blooms that occur in the austral spring and summer months have
689 been associated with warm sea surface temperatures (>26°C) and recent NO_3^- and soluble
690 reactive P (SRP) enrichments in the lagoon (Rodier and Le Borgne, 2008). Although
691 *Trichodesmium* was present at relatively low abundances during the first 8 days of the
692 experiment ($3.4 \times 10^2 - 6.5 \times 10^3$ *nifH* copies L^{-1} ; Supplement Table S3), abundances
693 increased steadily throughout the experiment ($p=0.01$, $r=+0.73$). *Trichodesmium* reached
694 peak abundances of 1.4×10^5 *nifH* copies L^{-1} on the final sampling day (day 22; Fig. 2e).
695 The only environmental parameter with which changes in *Trichodesmium* abundances
696 could be strongly correlated was NH_4^+ ($p=0.02$, $r=+0.81$).

697 The UCYN-C group comprised a minor part of the Noumea lagoon diazotroph
698 community, and was detected at abundances between $3.2 \times 10^2 - 4.8 \times 10^3$ *nifH* copies L^{-1}
699 (Supplement Table S3) during the experimental period. However, UCYN-C abundances
700 did show a linear increase over the duration of the experiment ($p=0.01$, $r=+0.74$), and
701 significant positive correlations to changes in salinity ($p=0.02$, $r=+0.80$), and NH_4^+
702 ($p=0.03$, $r=+0.76$). The unicellular UCYN-C group was first described when *nifH*
703 sequences with 90% nucleotide identity to the cultivated *Cyanothece* sp. PCC 51142
704 were recovered from surface waters in the Tropical North Atlantic (Langlois et al.,
705 2005; Foster et al., 2007). This phylotype has rarely been quantified in the oligotrophic

706 ocean, and when present, abundances are low (Foster et al., 2007;Langlois et al.,
707 2008;Needoba et al., 2007;Goebel et al., 2010;Ratten et al., 2014), therefore very little is
708 known about its distribution or importance. The qPCR assay used in this study (Foster et
709 al., 2007) amplifies not only the uncultivated Atlantic phylotype, but also many
710 cultivated *Cyanothece*- and *Gloeothece*-like isolates, which are expected to be present in
711 an environment like the Noumea Lagoon, as *Cyanothece*-and *Gloeothece*-like organisms
712 have been reported in shallow marine sediments (Hunter et al., 2006;Bauer et al., 2008)
713 and intertidal sands (Reddy et al., 1993;Ohki et al., 2008).

714 γ -24774A11 was also consistently detected at low abundances (ca. 10^2 - 10^3 *nifH*
715 copies L⁻¹; Supplement Table S3), which increased throughout the duration of the
716 experiment ($p=0.03$, $r=+0.65$). Despite being the most widely studied marine
717 heterotrophic diazotroph, very little is known about this phylotype. Recent studies
718 suggest that this γ -proteobacterial diazotroph is likely to be free-living and may be able to
719 sustain in a broad range of environmental conditions, as evidenced by low but uniform
720 abundances throughout the photic zone in the South Pacific (Moisander et al., 2014).

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

727 *Crocospaera* sp. (UCYN-B), previously reported to be members of the
728 unicellular diazotroph community in the Noumea Lagoon (Biegala and Raimbault, 2008),
729 were not initially detected, but were present at low abundances that increased over the 23
730 day period ($p=0.04$, $r=+0.63$) to abundances as high as 2.7×10^3 *nifH* copies L⁻¹, and had a
731 significant strong correlation with total chl *a* ($p=0.02$, $r=+0.81$). It must be noted that the
732 *Crocospaera*-like sequences that were recovered from the mesocosm libraries would not
733 likely amplify efficiently with the qPCR assay used. Therefore, it is not clear whether
734 these abundances included a *Crocospaera* population not recovered in the PCR-based
735 *nifH* libraries, or were in part the result of cross-reactivity with the *Crocospaera*-like

736 population that were amplified from mesocosm samples during P1 and P2, and possibly
737 present in the lagoon as well.

738

739 **4 Conclusions**

740 The VAHINE mesocosm experiment was conducted to trace the incorporation of N₂
741 fixed by diazotrophs into the food web (Bonnet et al., 2015a). Despite lack of sampling
742 replication within each mesocosm, consistent patterns in both the relative abundances of
743 targeted diazotrophic phylotypes, as well as shifts in diazotrophic community
744 composition, were reproducibly seen in all three mesocosms. Although the timing of the
745 increases were specific to an individual mesocosm, UCYN-C and Het-3 abundances
746 increased over time, while UCYN-A1 and Het-1 abundances decreased over time in all
747 three mesocosms (Fig. 3). The experimental conditions selected for the growth of UCYN-
748 C during P2, an ecotype never before quantified at high abundances in the marine water
749 column, which enabled the calculation of growth rates of this uncultivated ecotype, and
750 provided insight into its dynamics with respect to environmental parameters. Although
751 the data strongly suggests that the drawdown of DIP provided an environment favorable
752 for high UCYN-C growth rates, further studies are required to better understand the
753 environmental conditions that stimulated this bloom, and whether such blooms are seen
754 in the Noumea Lagoon itself.

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

764 The analysis of the diazotroph assemblage outside the mesocosms represents both the
765 first quantitative data on targeted diazotroph phylotypes, as well as the first *nifH*-based
766 diversity libraries on the populations in the Noumea Lagoon. Not previously considered

767 to be a significant diazotrophs in this region, DDAs must now be considered a potentially
768 important contributor to BNF in these waters, especially in the austral summer. Although
769 the presence of UCYN-A in the lagoon has long been suspected due to the relative
770 importance of daytime nitrogen fixation in the small size fraction (Biegala and
771 Raimbault, 2008), we report the first quantitative data on UCYN-A abundances in the
772 Noumea lagoon. Furthermore, the co-occurrence of two UCYN-A ecotypes revealed in
773 this study, provides important insight into the overlap in environmental niches for these
774 two ecotypes.

775

776 **Author contributions**

777 SB designed and executed the experiments, and SB and AD sampled the experiments for
778 molecular analyses. MH extracted DNA samples, and KT conducted all qPCR and PCR
779 analysis, and analyzed the data. IF performed statistical analyses. KT and JZ prepared
780 the manuscript with input from all co-authors.

781

782 **Acknowledgements**

783 Funding for this research was provided by the Agence Nationale de la Recherche (ANR
784 starting grant VAHINE ANR-13-JS06-0002), INSU-LEFE-CYBER program, GOPS and
785 IRD. The authors thank the captain and crew of the R/V Alis. We also thank the SEOH
786 divers service from Noumea, as well as the technical support and the divers of the IRD
787 research center of Noumea. We gratefully acknowledge C. Guieu, J-M. Grisoni and F.
788 Louis for the mesocosms design and their useful advice. [We would also like to thank five](#)
789 [anonymous reviewers for their time and critiques](#). The molecular component of this work
790 was funded by a Gordon and Betty Moore Foundation Marine Investigator award (JPZ)
791 and the National Science Foundation Science Center for Microbial Oceanography
792 Research and Education (C-MORE, Grant #EF-0424599).

793

794 **References**

795 Bandyopadhyay, A., Elvitigala, T., Welsh, E., Stöckel, J., Liberton, M., Min, H., Sherman, L. A.,
796 and Pakrasi, H. B.: Novel metabolic attributes of the genus *Cyanothece*, comprising a group of
797 unicellular nitrogen-fixing cyanobacteria, MBio, 2, e00214-00211, 2011.
798

799 Bauer, K., Díez, B., Lugomela, C., Seppälä, S., Borg, A. J., and Bergman, B.: Variability in
800 benthic diazotrophy and cyanobacterial diversity in a tropical intertidal lagoon, *FEMS Microbiol*
801 *Ecol*, 63, 205-221, 2008.
802
803 Berthelot, H., Moutin, T., L'Helguen, S., Leblanc, K., Hélias, S., Grosso, O., Leblond, N.,
804 Charrière, B., and Bonnet, S.: N₂ fixation and DON as significant N sources for primary and
805 export productions during the VAHINE mesocosms experiment (New Caledonia lagoon),
806 *Biogeosciences*, 2015.
807
808 [Bentzon-Tilia, M., Traving, S. J., Mantikci, M., Knudsen-Leerbeck, H., Hansen, J. L., Markager,](#)
809 [S., and Riemann, L.: Significant N₂ fixation by heterotrophs, photoheterotrophs and heterocystous](#)
810 [cyanobacteria in two temperate estuaries, *The ISME journal*, 9\(2\), 273-285, 2015.](#)
811
812 Biegala, I. C., and Raimbault, P.: High abundance of diazotrophic picocyanobacteria (< 3 μm) in
813 a Southwest Pacific coral lagoon, *Aquat Microb Ecol*, 51, 45-53, 10.3354/ame01185, 2008.
814
815 Bird, C., and Wyman, M.: Transcriptionally active heterotrophic diazotrophs are widespread in
816 the upper water column of the Arabian Sea, *FEMS Microbiol Ecol*, 84, 189-200, 10.1111/1574-
817 6941.12049, 2012.
818
819 Bonnet, S., Dekaezemacker, J., Turk-Kubo, K. A., Moutin, T., Hamersley, R. M., Grosso, O.,
820 Zehr, J. P., and Capone, D. G.: Aphotic N₂ Fixation in the Eastern Tropical South Pacific Ocean,
821 *PLoS ONE*, 8, e81265, [doi:10.1371/journal.pone.0081265](https://doi.org/10.1371/journal.pone.0081265), 2013.
822
823 Bonnet, S.: Introduction to the project VAHINE: VARIability of vertical and tropHic transfer of
824 fixed N₂ in the south wEst Pacific, *Biogeosciences*, [in preparation](#), 2015a.
825
826 [Bonnet, S., Berthelot, H., Turk-Kubo, K. Fawcett, F. Rahav, E., Berman-Frank, I., and l'Helguen,](#)
827 [S. Dynamics of N₂ fixation and fate of diazotroph-derived nitrogen during the VAHINE](#)
828 [mesocosm experiment \(New Caledonia\), *in preparation* 2015b.](#)
829
830 [Bonnet, S., Berthelot, H., Turk-Kubo, K., Cornet-Bartaux, V., Fawcett, S. E., Berman-Frank, I.,](#)
831 [Barani, A., Dekaezemacker, J., Benavides, M., Charriere, B., and Capone, D. G.: Diazotroph](#)
832 [derived nitrogen supports diatoms growth in the South West Pacific: a quantitative study using](#)
833 [nanoSIMS, *Limnology and Oceanography*, *under revision*.](#)
834
835 Breitbarth, E., Oschlies, A., and LaRoche, J.: Physiological constraints on the global distribution
836 of *Trichodesmium* - effect of temperature on diazotrophy, *Biogeosciences*, 4, 53-61,
837 [doi:10.5194/bg-4-53-2007](https://doi.org/10.5194/bg-4-53-2007), 2007.
838
839 [Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K.,](#)
840 [Fierer, N., Pena, A. G., Goodrich, J. K., and Gordon, J. I.: QIIME allows analysis of high-](#)
841 [throughput community sequencing data, *Nature Methods*, 7, 335-336, \[doi:10.1038/nmeth.f.303\]\(https://doi.org/10.1038/nmeth.f.303\),](#)
842 [2010.](#)
843
844 Carpenter, E. J., Montoya, J. P., Burns, J., Mulholland, M. R., Subramaniam, A., and Capone, D.
845 G.: Extensive bloom of a N₂-fixing diatom/cyanobacterial association in the tropical Atlantic
846 Ocean, *Mar Ecol Prog Ser*, 185, 273-283, 1999.
847

848 Carpenter, E. J., and Foster, R. A.: Marine cyanobacterial symbioses, in: *Cyanobacteria in*
849 *Symbiosis*, edited by: Rai, A. N., Bergman, B., and Rasmussen, U., Kluwer Academic Publishers,
850 Netherlands, 11-17, 2002.

851

852 Church, M., Jenkins, B., Karl, D., and Zehr, J.: Vertical distributions of nitrogen-fixing
853 phylotypes at Stn ALOHA in the oligotrophic North Pacific, *Aquat Microb Ecol*,
854 [doi:10.3354/ame038003](https://doi.org/10.3354/ame038003), 2005a.

855

856 Church, M., Short, C., Jenkins, B., Karl, D., and Zehr, J.: Temporal Patterns of Nitrogenase Gene
857 (*nifH*) Expression in the Oligotrophic North Pacific Ocean, *Appl Environ Microbiol*, 2005b.

858

859 Church, M. J., Bjorkman, K. M., Karl, D. M., Saito, M. A., and Zehr, J. P.: Regional distributions
860 of nitrogen-fixing bacteria in the Pacific Ocean, *Limnol Oceanogr*, 53, 63-77, 2008.

861

862 Dupouy, C., Neveux, J., Subramaniam, A., Mulholland, M. R., Montoya, J. P., Campbell, L.,
863 Carpenter, E. J., and Capone, D. G.: Satellite captures *Trichodesmium* blooms in the southwestern
864 tropical Pacific, *EOS, Transactions American Geophysical Union*, 81, 13-16, 2000.

865

866 [Dupouy, C., Benielli-Gary, D., Neveux, J., Dandonneau, Y., and Westberry, T.: A new algorithm](#)
867 [for detecting *Trichodesmium* surface blooms in the South Western Tropical Pacific,](#)
868 [Biogeosciences, 8, 1-17, doi:10.5194/bg-8-1-2011, 2011.](#)

869

870 Dyhrman, S. T., Chappell, P. D., Haley, S. T., Moffett, J. W., Orchard, E. D., Waterbury, J. B.,
871 and Webb, E. A.: Phosphonate utilization by the globally important marine diazotroph
872 *Trichodesmium*, *Nature*, 439, 68-71, 2006.

873

874 [Dyhrman, S. T., Ammerman, J. W., and Van Mooy, B. A. S.: Microbes and the Marine](#)
875 [Phosphorus Cycle, *Oceanography*, 20, 110-116, 2007.](#)

876

877 [Edgar, R. C.: Search and clustering orders of magnitude faster than BLAST, *Bioinformatics*, 26,](#)
878 [2460-2461,1367-4803, 2010.](#)

879

880 [Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., and Knight, R.: UCHIME improves](#)
881 [sensitivity and speed of chimera detection, *Bioinformatics*, 27, 2194-2200, 2011.](#)

882

883 Farnelid, H., Andersson, A. F., Bertilsson, S., Al-Soud, W. A., Hansen, L. H., Sorensen, S.,
884 Steward, G. F., Hagstrom, A., and Riemann, L.: Nitrogenase gene amplicons from global marine
885 surface waters are dominated by genes of non-cyanobacteria, *PLoS ONE*, 6, e19223,
886 [10.1371/journal.pone.0019223](https://doi.org/10.1371/journal.pone.0019223), 2011.

887

888 Fong, A. A., Karl, D., Lukas, R., Letelier, R. M., Zehr, J. P., and Church, M. J.: Nitrogen fixation
889 in an anticyclonic eddy in the oligotrophic North Pacific Ocean., *ISME Journal*, 2, 663-676, 2008.

890

891 Foster, R. A., and Zehr, J. P.: Characterization of diatom–cyanobacteria symbioses on the basis of
892 *nifH*, *hetR* and 16S rRNA sequences, *Environ. Microbiol.*, 8, doi: 1913-1925, [10.1111/j.1462-](https://doi.org/10.1111/j.1462-2920.2006.01068.x)
893 [2920.2006.01068.x](https://doi.org/10.1111/j.1462-2920.2006.01068.x), 2006.

894

895 Foster, R. A., Subramaniam, A., Mahaffey, C., Carpenter, E. J., Capone, D. G., and Zehr, J. P.:
896 Influence of the Amazon River plume on distributions of free-living and symbiotic cyanobacteria
897 in the western tropical north Atlantic Ocean, *Limnol Oceanogr*, 52, 517-532, 2007.

898

899 Foster, R. A., Goebel, N. L., and Zehr, J. P.: Isolation of *Calothrix rhizosoleniae*
900 (CYANOBACTERIA) strain SC01 from *Chaetoceros* (BACILLARIOPHYTA) spp. diatoms of
901 the subtropical north Pacific Ocean, *J Phycol*, 46, 1028-1037, doi:10.1111/j.1529-
902 8817.2010.00885.x, 2010.
903
904 Foster, R. A., Kuypers, M. M. M., Vagner, T., Paerl, R. W., Musat, N., and Zehr, J. P.: Nitrogen
905 fixation and transfer in open ocean diatom–cyanobacterial symbioses, *ISME J*, 5, 1484-1493,
906 doi:10.1038/ismej.2011.26, 2011.
907
908 Garcia, N., Raimbault, P., and Sandroni, V.: Seasonal nitrogen fixation and primary production in
909 the Southwest Pacific: nanoplankton diazotrophy and transfer of nitrogen to picoplankton
910 organisms, *Mar Ecol Prog Ser*, 343, 25-33, 2007.
911
912 Gattuso, J.-P., Frankignoulle, M., and Wollast, R.: Carbon and carbonate metabolism in coastal
913 aquatic ecosystems, *Annu Rev Ecol Syst*, 405-434, doi:10.1146/annurev.ecolsys.29.1.405, 1998.
914
915 Goebel, N. L., Turk, K. A., Achilles, K. M., Paerl, R. W., Hewson, I., Morrison, A. E., Montoya,
916 J. P., Edwards, C. A., and Zehr, J. P.: Abundance and distribution of major groups of diazotrophic
917 cyanobacteria and their potential contribution to N₂ fixation in the tropical Atlantic Ocean,
918 *Environ. Microbiol.*, 12, 3272-3289, doi:10.1111/j.1462-2920.2010.02303.x, 2010.
919
920 Gómez, F., Furuya, K., and Takeda, S.: Distribution of the cyanobacterium *Richelia*
921 *intracellularis* as an epiphyte of the diatom *Chaetoceros compressus* in the western Pacific
922 Ocean, *J Plankton Res*, 27, 323-330, 2005.
923
924 [Green, S. J., Venkatramanan, R., and Naqib, A.: Deconstructing the Polymerase Chain Reaction:
925 Understanding and Correcting Bias Associated with Primer Degeneracies and Primer-Template
926 Mismatches, *PLOS One*, doi: 10.1371/journal.pone.0128122, 2015.](#)
927
928 Gruber, N., and Sarmiento, J. L.: Global patterns of marine nitrogen fixation and denitrification,
929 *Global Biogeochem. Cy.*, 11, 235-266, 1997.
930
931 Hagino, K., Onuma, R., Kawachi, M., and Horiguchi, T.: Discovery of an endosymbiotic
932 nitrogen-fixing cyanobacterium UCYN-A in *Braarudosphaera bigelowii* (Prymnesiophyceae),
933 *PLoS One*, 8, e81749, doi:10.1371/journal.pone.0081749, 2013.
934
935 Halm, H., Lam, P., Ferdelman, T. G., Lavik, G., Dittmar, T., LaRoche, J., D'Hondt, S., and
936 Kuypers, M. M.: Heterotrophic organisms dominate nitrogen fixation in the South Pacific Gyre,
937 *ISME J*, 6, 1238-1249, doi:10.1038/ismej.2011.182, 2012.
938
939 Heller, P., Tripp, H. J., Turk-Kubo, K., and Zehr, J. P.: ARBitrator: a software pipeline for on-
940 demand retrieval of auto-curated *nifH* sequences from GenBank, *Bioinformatics*,
941 doi:10.1093/bioinformatics/btu417, 2014.
942
943 Hewson, I., Moisander, P. H., Achilles, K. M., Carlson, C. A., Jenkins, B. D., Mondragon, E. A.,
944 Morrison, A. E., and Zehr, J. P.: Characteristics of diazotrophs in surface to abyssopelagic waters
945 of the Sargasso Sea, *Aquat Microb Ecol*, 46, 15-30, 2007.
946
947 Hilton, J.: Ecology and evolution of diatom-associated cyanobacteria through genetic analyses,
948 Ph.D., Ocean Sciences Department, University of California, Santa Cruz, Santa Cruz, 2014.
949

950 Hilton, J. A., Foster, R. A., Tripp, H. J., Carter, B. J., Zehr, J. P., and Villareal, T. A.: Genomic
951 deletions disrupt nitrogen metabolism pathways of a cyanobacterial diatom symbiont, *Nat*
952 *Commun*, 4, 1767, doi:10.1038/ncomms2748, 2013.
953
954 Houlbreque, F., and Ferrier-Pagès, C.: Heterotrophy in tropical scleractinian corals, *Biological*
955 *Reviews*, 84, 1-17, doi:10.1111/j.1469-185X.2008.00058.x, 2009.
956
957 Hunter, E. M., Mills, H. J., and Kostka, J. E.: Microbial community diversity associated with
958 carbon and nitrogen cycling in permeable shelf sediments, *Appl Environ Microbiol*, 72, 5689-
959 5701, 2006.
960
961 Hutchins, D. A., Fu, F. X., Zhang, Y., Warner, M. E., Feng, Y., Portune, K., Bernhardt, P. W.,
962 and Mulholland, M. R.: CO₂ control of *Trichodesmium* N₂ fixation, photosynthesis, growth rates,
963 and elemental ratios: Implications for past, present, and future ocean biogeochemistry, *Limnol*
964 *Oceanogr*, 52, 1293-1304, 2007.
965
966 Hyndes, G. A., Nagelkerken, I., McLeod, R. J., Connolly, R. M., Lavery, P. S., and Vanderklift,
967 M. A.: Mechanisms and ecological role of carbon transfer within coastal seascapes, *Biological*
968 *Reviews*, 89, 232-254, 2014.
969
970 Karl, D., Letelier, R., Tupas, L., Dore, J., Christian, J., and Hebel, D.: The role of nitrogen
971 fixation in biogeochemical cycling in the subtropical North Pacific Ocean, *Nature*, 388, 533-538,
972 1997.
973
974 Karl, D. M., and Letelier, R. M.: Nitrogen fixation-enhanced carbon sequestration in low nitrate,
975 low chlorophyll seascapes, *Mar Ecol Prog Ser*, 364, 257-268, doi:10.3354/meps07547, 2008.
976
977 Karl, D. M., Church, M. J., Dore, J. E., Letelier, R. M., and Mahaffey, C.: Predictable and
978 efficient carbon sequestration in the North Pacific Ocean supported by symbiotic nitrogen
979 fixation, *Proceedings of the National Academy of Sciences*, 109, 1842-1849,
980 doi:10.1073/pnas.1120312109, 2012.
981
982 Langlois, R. J., LaRoche, J., and Raab, P. A.: Diazotrophic diversity and distribution in the
983 tropical and subtropical Atlantic Ocean, *Appl Environ Microbiol*, 71, 7910-7919, 2005b.
984
985 Langlois, R. J., Hummer, D., and LaRoche, J.: Abundances and distributions of the dominant
986 *nifH* phylotypes in the Northern Atlantic Ocean, *Appl Environ Microbiol*, 74, 1922-1931, 2008.
987
988 Langlois, R., Großkopf, T., Mills, M., Takeda, S., and LaRoche, J.: Widespread Distribution and
989 Expression of Gamma A (UMB), an Uncultured, Diazotrophic, γ -Proteobacterial *nifH* Phylotype,
990 *PLoS ONE*, 10, e0128912, 0121932-0126203, 2015.
991
992 Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, Buchner, A., Lai, T.,
993 Steppi, S., Jobb, G., Förster, W., Brettske, I., Gerber, S., Ginhart, A. W., Gross, O., Grumann, S.,
994 Hermann, S., Jost, R., König, A., Liss, T., Lüßmann, R., May, M., Nonhoff, B., Reichel, B.,
995 Strehlow, R., Stamatakis, A., Stuckmann, N., Vilbig, A., Lenke, M., Ludwig, T., Bode, A., and
996 Schleifer, K. H.: ARB: a software environment for sequence data, *Nucleic Acids Res*, 32, 1363-
997 1371, 2004.
998
999 Luo, Y. W., Doney, S. C., Anderson, L. A., Benavides, M., Berman-Frank, I., Bode, A., Bonnet,
1000 S., Boström, K. H., Böttjer, D., Capone, D. G., Carpenter, E. J., Chen, Y. L., Church, M. J., Dore,

1001 J. E., Falcón, L. I., Fernández, A., Foster, R. A., Furuya, K., Gómez, F., Gundersen, K., Hynes, A.
1002 M., Karl, D. M., Kitajima, S., Langlois, R. J., LaRoche, J., Letelier, R. M., Marañón, E.,
1003 McGillicuddy Jr, D. J., Moisander, P. H., Moore, C. M., Mouriño-Carballido, B., Mulholland, M.
1004 R., Needoba, J. A., Orcutt, K. M., Poulton, A. J., Rahav, E., Raimbault, P., Rees, A. P., Riemann,
1005 L., Shiozaki, T., Subramaniam, A., Tyrrell, T., Turk-Kubo, K. A., Varela, M., Villareal, T. A.,
1006 Webb, E. A., White, A. E., Wu, J., and Zehr, J. P.: Database of diazotrophs in global ocean:
1007 abundance, biomass and nitrogen fixation rates, *Earth Syst. Sci. Data*, 4, 47-73, doi:10.5194/essd-
1008 4-47-2012, 2012.
1009
1010 Masotti, I., Ruiz-Pino, D., and Le Bouteiller, A.: Photosynthetic characteristics of *Trichodesmium*
1011 in the southwest Pacific Ocean: importance and significance, *Marine Ecology-Progress Series*,
1012 338, 47-59, 2007.
1013
1014 Moisander, P. H., Beinart, R. A., Voss, M., and Zehr, J. P.: Diversity and abundance of
1015 diazotrophic microorganisms in the South China Sea during intermonsoon., *ISME J*, 2, 954-967,
1016 2008.
1017
1018 Moisander, P. H., Beinart, R. A., Hewson, I., White, A. E., Johnson, K. S., Carlson, C. A.,
1019 Montoya, J. P., and Zehr, J. P.: Unicellular cyanobacterial distributions broaden the oceanic N₂
1020 fixation domain, *Science*, 327, doi:1512-1514, 10.1126/science.1185468, 2010.
1021
1022 Moisander, P. H., Zhang, R., Boyle, E. A., Hewson, I., Montoya, J. P., and Zehr, J. P.: Analogous
1023 nutrient limitations in unicellular diazotrophs and *Prochlorococcus* in the South Pacific Ocean,
1024 *ISME J*, 6, 733-744, 2012.
1025
1026 Moisander, P. H., Serros, T., Paerl, R. W., Beinart, R. A., and Zehr, J. P.: Gammaproteobacterial
1027 diazotrophs and *nifH* gene expression in surface waters of the South Pacific Ocean, *ISME J*,
1028 doi:10.1038/ismej.2014.49, 2014.
1029
1030 Moutin, T., Van Den Broeck, N., Beker, B., Dupouy, C., Rimmelin, P., and Le Bouteiller, A.:
1031 Phosphate availability controls *Trichodesmium* spp. biomass in the SW Pacific Ocean, *Marine*
1032 *Ecology-Progress Series*, 297, 15-21, 2005.
1033
1034 Needoba, J. A., Foster, R. A., Sakamoto, C., Zehr, J. P., and Johnson, K. S.: Nitrogen fixation by
1035 unicellular diazotrophic cyanobacteria in the temperate oligotrophic North Pacific Ocean, *Limnol*
1036 *Oceanogr*, 52, 1317-1327, 2007.
1037
1038 Nenadic, O., and Greenacre, M.: Correspondence analysis in R, with two-and three-dimensional
1039 graphics: The ca package, 2007.
1040
1041 Ohki, K., Kamiya, M., Honda, D., Kumazawa, S., and Ho, K. K.: MORPHOLOGICAL AND
1042 PHYLOGENETIC STUDIES ON UNICELLULAR DIAZOTROPHIC CYANOBACTERIA
1043 (CYANOPHYTES) ISOLATED FROM THE COASTAL WATERS AROUND SINGAPORE, *J*
1044 *Phycol*, 44, 142-151, 2008.
1045
1046 Ouillon, S., Douillet, P., Lefebvre, J.-P., Le Gendre, R., Jouon, A., Bonneton, P., Fernandez, J.-
1047 M., Chevillon, C., Magand, O., and Lefèvre, J.: Circulation and suspended sediment transport in a
1048 coral reef lagoon: The south-west lagoon of New Caledonia, *Mar Pollut Bull*, 61, 269-296, 2010.
1049

1050 Ratten, J., LaRoche, J., Desai, D. K., Shelley, R., Landing, W. M., Boyle, E. A., Cutter, G. A.,
1051 and Langlois, R.: Source of iron and phosphate affect the distribution of diazotrophs in the North
1052 Atlantic, *Deep Sea Res II*, doi:10.1016/j.dsr2.2014.11.012, 2014.

1053

1054 Reddy, K., Haskell, J., Sherman, D., and Sherman, L.: Unicellular, aerobic nitrogen-fixing
1055 cyanobacteria of the genus *Cyanothece*, *J Bacteriol*, 175, 1284, 1993.

1056

1057 Renaud, F., Pringault, O., and Rochelle-Newall, E.: Effects of the colonial cyanobacterium
1058 *Trichodesmium* spp. on bacterial activity, *Aquat Microb Ecol*, 41, 261-270, 2005.

1059

1060 Riemann, L., Farnelid, H., and Steward, G. F.: Nitrogenase genes in non-cyanobacterial plankton:
1061 prevalence, diversity and regulation in marine waters, *Aquat Microb Ecol*, 61, 225-237,
1062 doi:10.3354/Ame01431, 2010.

1063

1064 Rodier, M., and Le Borgne, R.: Population dynamics and environmental conditions affecting
1065 *Trichodesmium* spp.(filamentous cyanobacteria) blooms in the south-west lagoon of New
1066 Caledonia, *J Exp Mar Biol Ecol*, 358, 20-32, 2008.

1067

1068 Rodier, M., and Le Borgne, R.: Population and trophic dynamics of *Trichodesmium thiebautii* in
1069 the SE lagoon of New Caledonia. Comparison with *T. erythraeum* in the SW lagoon, *Mar Pollut*
1070 *Bull*, 61, 349-359, 2010.

1071

1072 [Shiozaki, T., Ijichi, M., Kodama, T., Takeda, S., and Furuya, K.: Heterotrophic bacteria as major
1073 nitrogen fixers in the euphotic zone of the Indian Ocean, *Global Biogeochemical Cycles*, 28\(10\),
1074 1096-1110, 2014.](#)

1075

1076 Subramaniam, A., Yager, P. L., Carpenter, E. J., Mahaffey, C., Bjorkman, K., Cooley, S., Kustka,
1077 A. B., Montoya, J. P., Sanudo-Wilhelmy, S. A., Shipe, R., and Capone, D. G.: Amazon River
1078 enhances diazotrophy and carbon sequestration in the tropical North Atlantic Ocean, *Proc Natl*
1079 *Acad Sci USA*, 105, 10460-10465, doi:10.1073/pnas.0710279105, 2008.

1080

1081 Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S.: MEGA5: molecular
1082 evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum
1083 parsimony methods, *Mol Biol Evol*, 28, 2731-2739, doi:10.1093/molbev/msr121, 2011.

1084

1085 Taniuchi, Y., Chen, Y.-I. L., Chen, H.-Y., Tsai, M.-L., and Ohki, K.: Isolation and
1086 characterization of the unicellular diazotrophic cyanobacterium Group C TW3 from the tropical
1087 western Pacific Ocean, *Environ. Microbiol.*, 14, 641-654, doi:10.1111/j.1462-2920.2011.02606.x,
1088 2012.

1089

1090 Team, R. D. C.: R: A language and environment for statistical computing, Foundation for
1091 Statistical Computing, <http://www.R-project.org/>, 2012.

1092

1093 Thompson, A., Carter, B. J., Turk-Kubo, K., Malfatti, F., Azam, F., and Zehr, J. P.: Genetic
1094 diversity of the unicellular nitrogen-fixing cyanobacteria UCYN-A and its prymnesiophyte host,
1095 *Environ Microbiol*, 16, 3238-3249, doi:10.1111/1462-2920.12490, 2014.

1096

1097 Thompson, A. W., Foster, R. A., Krupke, A., Carter, B. J., Musat, N., Vaultot, D., Kuypers, M. M.
1098 M., and Zehr, J. P.: Unicellular Cyanobacterium Symbiotic with a Single-Celled Eukaryotic Alga,
1099 *Science*, 337, 1546-1550, doi:10.1126/science.1222700, 2012.

1100

1101 Torréron, J.-P., Rochelle-Newall, E., Pringault, O., Jacquet, S., Faure, V., and Briand, E.:
1102 Variability of primary and bacterial production in a coral reef lagoon (New Caledonia), Mar
1103 Pollut Bull, 61, 335-348, 2010.
1104
1105 Tripp, H. J., Bench, S. R., Turk, K. A., Foster, R. A., Desany, B. A., Niazi, F., Affourtit, J. P., and
1106 Zehr, J. P.: Metabolic streamlining in an open ocean nitrogen-fixing cyanobacterium, Nature,
1107 464, 90-94, doi:10.1038/nature08786, 2010a.
1108
1109 Turk, K., Rees, A. P., Zehr, J. P., Pereira, N., Swift, P., Shelley, R., Lohan, M., Woodward, E. M.
1110 S., and Gilbert, J.: Nitrogen fixation and nitrogenase (*nifH*) expression in tropical waters of the
1111 eastern North Atlantic, ISME J, 5, 1201-1212, 10.1038/ismej.2010.205, 2011.
1112
1113 Turk-Kubo, K. A., Achilles, K. M., Serros, T. R., Ochiai, M., Montoya, J. P., and Zehr, J. P.:
1114 Nitrogenase (*nifH*) gene expression in diazotrophic cyanobacteria in the Tropical North Atlantic
1115 in response to nutrient amendments, Frontiers in Aquatic Microbiology, 3, 386,
1116 doi:10.3389/fmicb.2012.00386, 2012.
1117
1118 Turk-Kubo, K. A., Karamchandani, M., Capone, D. G., and Zehr, J. P.: The paradox of marine
1119 heterotrophic nitrogen fixation: abundances of heterotrophic diazotrophs do not account for
1120 nitrogen fixation rates in the Eastern Tropical South Pacific, Environ Microbiol, 16, 3095-3114,
1121 doi:10.1111/1462-2920.12346, 2014.
1122
1123 Villareal, T. A.: Laboratory culture and preliminary characterization of the nitrogen-fixing
1124 *Rhizosolenia-Richelia* symbiosis, Mar Ecol, 11, 117-132, 1990.
1125
1126 Villareal, T. A.: Marine nitrogen-fixing diatom - cyanobacteria symbioses, in: Marine Pelagic
1127 Cyanobacteria: *Trichodesmium* and other Diazotrophs, edited by: Carpenter, E. J., Capone, D. G.,
1128 and Rueter, J. G., Series C: Mathematical and Physical Sciences, Kluwer Academic Publishers,
1129 The Netherlands, 163-175, 1992.
1130
1131 Villareal, T. A., Brown, C. G., Brzezinski, M. A., Krause, J. W., and Wilson, C.: Summer diatom
1132 blooms in the North Pacific subtropical gyre: 2008–2009, PLoS ONE, 7, e33109,
1133 doi:10.1371/journal.pone.0033109, 2012.
1134
1135 Vu, T. T., Stolyar, S. M., Pinchuk, G. E., Hill, E. A., Kucek, L. A., Brown, R. N., Lipton, M. S.,
1136 Osterman, A., Fredrickson, J. K., and Konopka, A. E.: Genome-scale modeling of light-driven
1137 reductant partitioning and carbon fluxes in diazotrophic unicellular cyanobacterium *Cyanothece*
1138 sp. ATCC 51142, Plos Computational Biology, 8, e1002460, doi:10.1371/journal.pcbi.1002460,
1139 2012.
1140
1141 Walsby, A.: The gas vesicles and buoyancy of *Trichodesmium*, in: Marine Pelagic Cyanobacteria:
1142 *Trichodesmium* and Other Diazotrophs, Springer, 141-161, 1992.
1143
1144 Webb, E. A., Ehrenreich, I. M., Brown, S. L., Valois, F. W., and Waterbury, J. B.: Phenotypic
1145 and genotypic characterization of multiple strains of the diazotrophic cyanobacterium,
1146 *Crocospaera watsonii*, isolated from the open ocean., Environ. Microbiol., 11, 338-348, 2009.
1147
1148 Zehr, J. P., Waterbury, J. B., Turner, P. J., Montoya, J. P., Omoregie, E., Steward, G. F., Hansen,
1149 A., and Karl, D. M.: Unicellular cyanobacteria fix N₂ in the subtropical North Pacific Ocean,
1150 Nature, 412, 635-638, 2001.
1151

1152 Zehr, J. P., and Turner, P. J.: Nitrogen fixation: nitrogenase genes and gene expression, in:
1153 *Methods in Microbiology: Marine Microbiology*, edited by: Paul, J. H., Academic Press, Ltd.,
1154 London, 271-286, 2001.
1155
1156 Zehr, J., Jenkins, B., Short, S., and Steward, G.: Nitrogenase gene diversity and microbial
1157 community structure: a cross - system comparison, *Environ. Microbiol.*, 5, 539-554, 2003.
1158
1159 Zehr, J. P., Bench, S. R., Carter, B. J., Hewson, I., Niazi, F., Shi, T., Tripp, H. J., and Affourtit, J.
1160 P.: Globally distributed uncultivated oceanic N₂-fixing cyanobacteria lack oxygenic photosystem
1161 II, *Science*, 322, 1110-1112, 2008.
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186

1187 **Tables**

1188

1189

1190

	targeted by qPCR		
	no. OTUs (no. sequences)	no. OTUs (no. sequences)	% OTUs (% sequences)
1B	152 (311550)		
UCYN-A	68 (260221)	60 (258005)	88% (99%)
UCYN-B	32 (15917)	0 (0)	0% (0%)
UCYN-C	18 (11324)	9 (2915)	50% (26%) ☉
<i>Tricho.</i>	18 (20034)	14 (19386)	78% (97%)
Het-1	2 (1738)	2 (1738)	100% (100%)
Het-2	0 (0)	0 (0)	0% (0%)
Het-3	0 (0)	0 (0)	0% (0%)
other	14 (2316)	na	
1G	88 (120586)		
γ-24774A11	22 (51594)	18 (50833)	82% (99%)
other	66 (68992)	na	
1J/1K	19 (40032)		
3	16 (6825)		
1O	2 (409)		

☉ 61% (85%) when a third mismatch at the 5' end (probe) is allowed

1191

Table 1. In silico qPCR coverage analysis. Taxonomic assignment for all *nifH*

1192

amplicons reads that passed the quality filtering steps (first column), and the number of

1193

OTUs affiliated with each group that are successfully targeted by qPCR assays used in

1194

this study. Partial *nifH* sequences were classified according to the convention defined in

1195

Zehr et al. (2003). OTU – operational taxonomic unit.

1196

1197

1198

1199

1200

1201

1202

1203

1204

1205

	Temp., Sal. & PO ₄		Temp. & PO ₄		Temp.		PO ₄	
	R ²	R ² -cv	R ²	R ² -cv	R ²	R ² -cv	R ²	R ² -cv
UCYN-A1	0.85	0.76	0.84	0.76	0.70	0.59	0.73	0.63
UCYN-C	0.84	0.72	0.79	0.68	0.71	0.60	0.75	0.66
Het-3	0.77	0.60	0.70	0.57	0.55	0.39	0.70	0.60

1206

1207 **Table 2. Results from multiple regression models predicting diazotroph abundances**

1208 **from environmental parameters.** Abundances from both the Noumea lagoon and the
1209 VAHINE mesocosms experiments are included in the linear regression models. R²-cv is
1210 the cross-validated fit of the model, and when similar to the R² value, indicates a high
1211 predictive ability for future datasets.

1212

1213

1214

1215

1216

1217

1218

1219

1220

1221

1222

1223

1224

1225

1226

1227

Net growth rate / net mortality rate (d ⁻¹)										
time	UCYN-A1	UCYN-A2	UCYN-B	UCYN-C	γ-24774A11	Tricho.	Het-1	Het-2	Het-3	
M1	5	0.54 ± 0.05	0.26 ± 0.04	na	-0.04 ± 0.53	0.27 ± 0.13	0.15 ± 0.04	0.07 ± 0.05	-0.55 ± 0.18	na
	7	0.08 ± 0.07	-0.45 ± 0.04	na	na	-0.08 ± 0.11	-0.94 ± 0.05	0.03 ± 0.03	0.3 ± 0.17	na
	9	-0.81 ± 0.07	-0.48 ± 0.04	-0.41 ± 0.16	na	-0.24 ± 0.06	1.46 ± 0.05	0.34 ± 0.04	0.55 ± 0.03	na
	11	-0.47 ± 0.01	-0.02 ± 0.11	-0.65 ± 0.16	0.46 ± 0.23	0.19 ± 0.03	-1.12 ± 0.03	-0.15 ± 0.09	0.01 ± 0.04	na
	13	0.73 ± 0.04	-0.16 ± 0.15	0.85 ± 0.03	1.23 ± 0.07	-0.15 ± 0.22	0.39 ± 0.03	-0.23 ± 0.09	-0.21 ± 0.03	na
	15	-0.61 ± 0.04	0.18 ± 0.11	na	0.52 ± 0.03	0.27 ± 0.22	0.16 ± 0.04	-1.39 ± 0.04	-0.41 ± 0.02	0.21 ± 0.1
	18	-0.23 ± 0.08	-0.03 ± 0.08	na	0.06 ± 0.02	0.2 ± 0.02	-0.62 ± 0.08	0.7 ± 0.01	-0.03 ± 0.02	0.15 ± 0.16
	19	0.55 ± 0.08	0.31 ± 0.08	0.1 ± 0.11	0.29 ± 0.02	0.03	0.71 ± 0.08	0.7 ± 0.02	0.49 ± 0	-0.13 ± 0.22
	20	0.6 ± 0.05	0.77 ± 0.04	0.65 ± 0.13	-0.24 ± 0.07	0.2 ± 0.03	0.01 ± 0.01	-0.35 ± 0.02	-0.37 ± 0.02	0.57 ± 0.16
	23	0.47 ± 0.04	0.02 ± 0.06	0.55 ± 0.08	0.22 ± 0.07	-0.8 ± 0.07	0.74 ± 0.02	0.06 ± 0.01	0.01 ± 0.06	-0.1 ± 0.04
M2	5	--	--	--	--	--	--	--	--	--
	7	na	-1.63 ± 0.22	na	na	na	na	na	na	na
	9	-0.06 ± 0.03	0.52 ± 0.26	-0.36 ± 0.13	0.51 ± 0.24	-0.91 ± 0.04	0.03 ± 0.01	-0.53 ± 0.01	-0.06 ± 0.01	na
	11	-0.15 ± 0.03	0.29 ± 0.14	na	0.58 ± 0.24	0.31 ± 0.09	0.04 ± 0.02	0.56 ± 0.07	0.1 ± 0.01	na
	13	-0.47 ± 0.01	0.38 ± 0.1	na	0.79 ± 0.14	0.18 ± 0.09	-0.5 ± 0.06	-0.26 ± 0.07	0.44 ± 0.03	na
	15	-0.5 ± 0.07	0.33 ± 0.11	0.18 ± 0.07	2.16 ± 0.07	0.43 ± 0.06	0.22 ± 0.05	-1.29 ± 0.03	-0.63 ± 0.03	na
	18	-0.08 ± 0.11	na	0.3 ± 0.17	-0.27 ± 0.03	0.04 ± 0.07	0.14 ± 0.04	0.35 ± 0.02	-0.04 ± 0.01	0.76 ± 0.15
	19	--	--	--	--	--	--	--	--	--
	20	na	-0.46 ± 0.14	na	na	na	na	na	na	na
	23	0.43 ± 0.05	na	0.36 ± 0.07	0.05 ± 0.09	1.07 ± 0.07	0.09 ± 0.04	0.00 ± 0.07	0.04 ± 0.09	0.04 ± 0.09
M3	5	0.13 ± 0.02	1.71 ± 0.32	0.7 ± 0.48	0.96 ± 0.43	na	0.17 ± 0.03	0.64 ± 0.03	-0.35 ± 0.04	na
	7	0.29 ± 0.02	-1.64 ± 0.12	-0.42 ± 0.15	-0.38 ± 0.09	-0.39 ± 0.09	0.45 ± 0.06	-0.08 ± 0.05	0.54 ± 0.04	na
	9	-0.38 ± 0.03	-0.03 ± 0.1	-0.14 ± 0.3	0.13	0.29 ± 0.12	-2.16 ± 0.18	-0.63 ± 0.05	-0.72 ± 0.03	na
	11	-0.4 ± 0.03	0.18 ± 0.05	0.68 ± 0.27	0.1 ± 0.15	0.06 ± 0.14	1.2 ± 0.18	1.17 ± 0.06	0.84 ± 0.03	na
	13	-0.43 ± 0.07	0.16 ± 0.1	na	-0.02 ± 0.09	0.5 ± 0.09	0.8 ± 0.05	0.06 ± 0.06	-0.03 ± 0.03	na
	15	-0.48 ± 0.11	0.42 ± 0.11	na	1.91 ± 0.02	-0.95 ± 0.1	-0.4 ± 0.01	-1.16 ± 0.03	-0.66 ± 0.05	na
	18	-0.22 ± 0.21	-0.03 ± 0.07	0.21 ± 0.01	0.2 ± 0.04	-0.56 ± 0.11	-0.35 ± 0.01	-0.48 ± 0.01	-0.69 ± 0.07	0.06 ± 0.05
	19	-0.3 ± 0.19	-0.36 ± 0.14	-0.99 ± 0.06	-0.44 ± 0.06	1.03 ± 0.1	0.05 ± 0.01	-0.16 ± 0.11	2.24 ± 0.06	1.09 ± 0.06
	20	0.55 ± 0.12	1.25 ± 0.13	1.38 ± 0.06	0.28 ± 0.07	na	1.55 ± 0.02	1.28 ± 0.11	0.45 ± 0.07	-1.12 ± 0.05
	23	0.45 ± 0.14	-0.04 ± 0.08	0.07 ± 0.02	-0.02 ± 0.06	na	0.72 ± 0.03	-0.36 ± 0.02	-0.14 ± 0.08	0.77 ± 0.11

1228

1229 **Table 3. Diazotroph net growth and mortality rates (d⁻¹) during VAHINE mesocosms**

1230 **experiment.** Rates in bold are the maximum rates measured in each mesocosm. ‘-’

1231 denotes periods where no rates could be calculated due to missing data, and ‘na’ denotes

1232 missing rates due to abundances being ‘detected, not quantified’ (DNQ) or undetected

1233 (UD). Standard error (±) reported for each growth rate is derived from qPCR

1234 measurements for replicate amplifications.

1235

1236

1237

1238

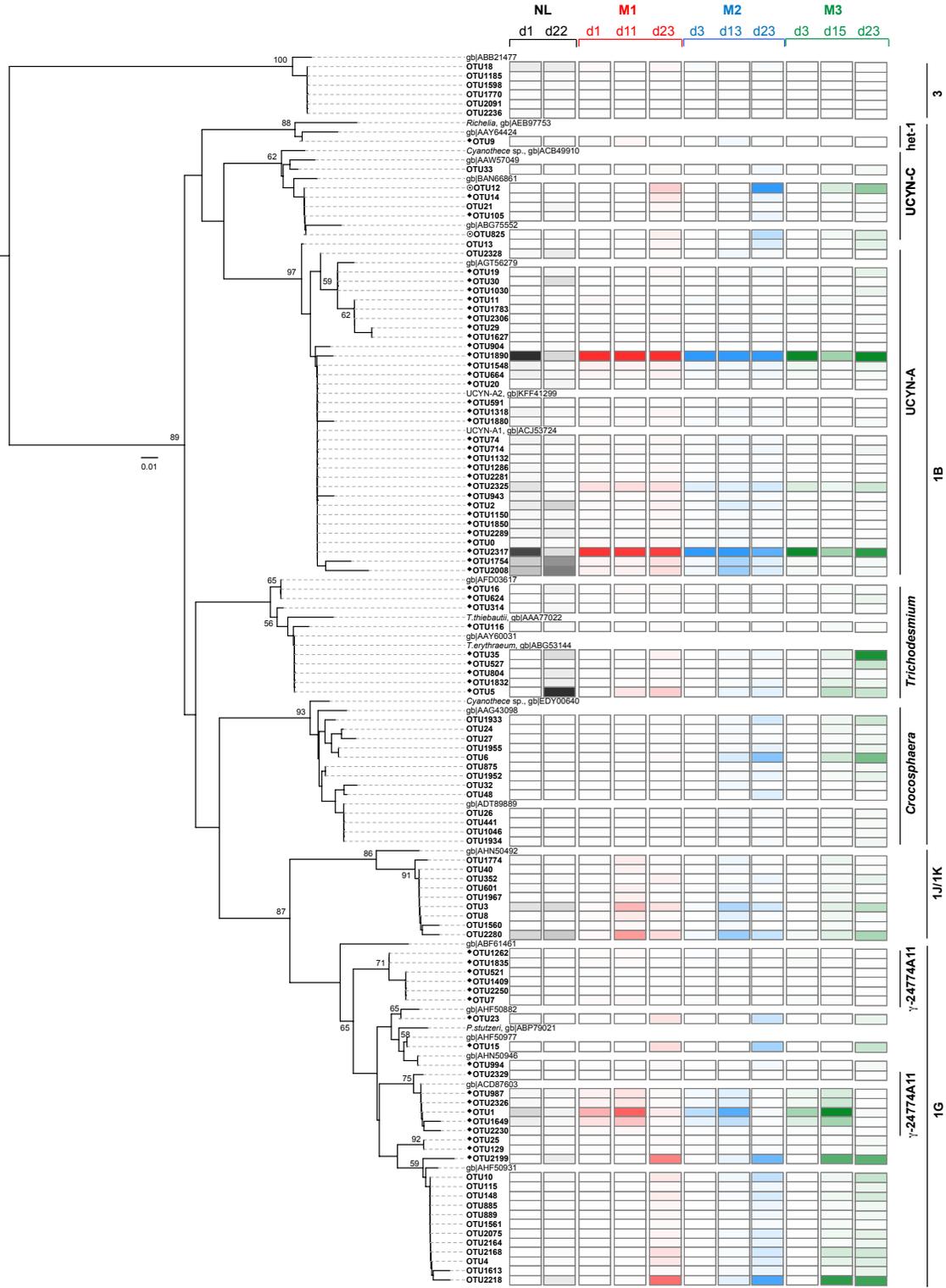
1239

1240

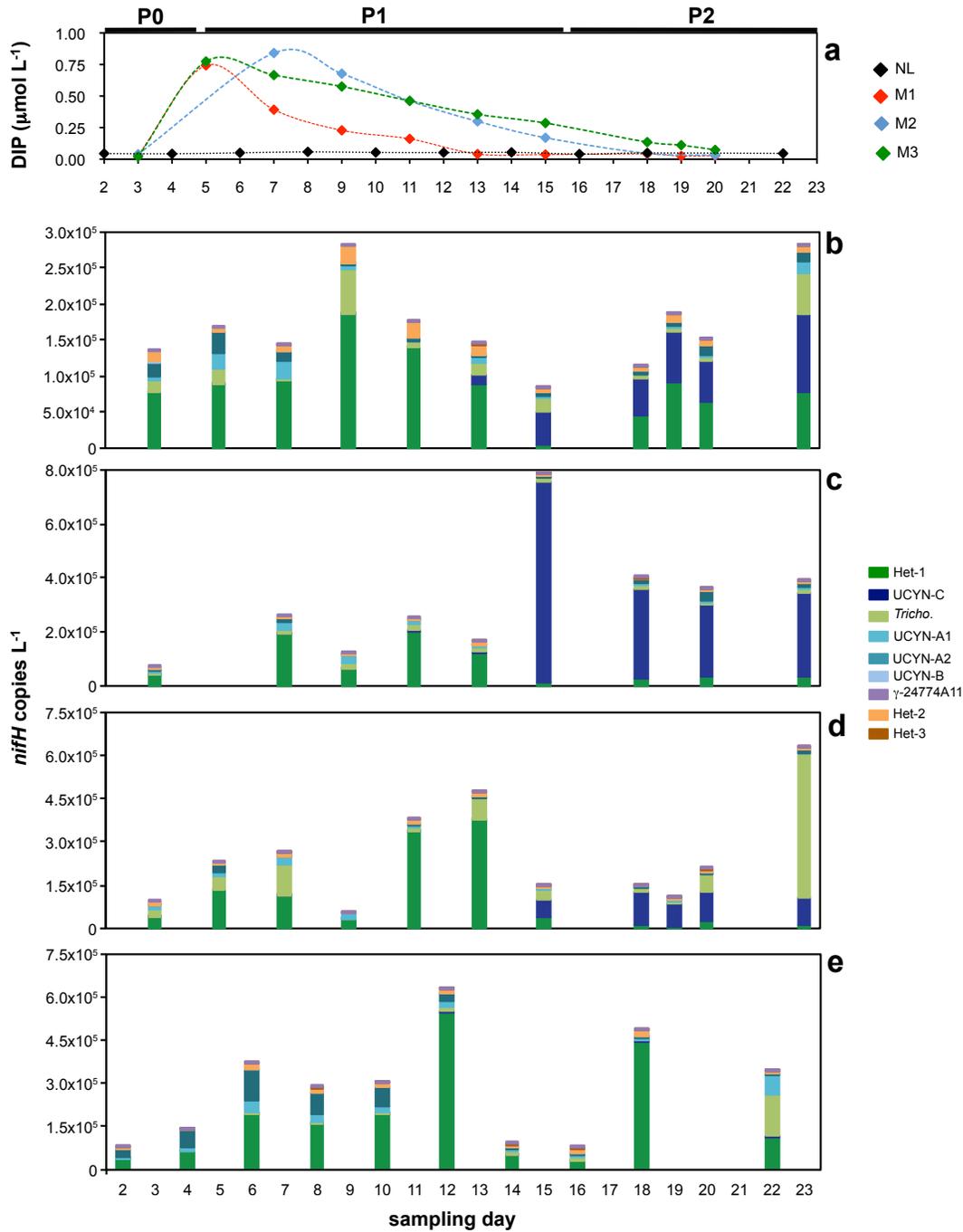
1241

1242

1243



1246 **Figure 1. Maximum likelihood tree calculated using partial *nifH* amino acid sequences**
1247 ***recovered from the Noumea Lagoon (NL) and mesocosms (M1, M2, and M3).*** Relative
1248 abundances of *nifH* reads associated with each operational taxonomic unit (OTU) are
1249 indicated for each sample by shaded boxes, with intense shading indicating high relative
1250 abundances, and light shading indicating low relative abundances. Trees were
1251 bootstrapped using 1000 replicate trees, and nodes with values >50 are displayed. Branch
1252 lengths were inferred using the JTT model, and the scale bar indicates the number of
1253 substitutions per site. OTUs that are targeted by qPCR assays used in this study are
1254 marked with a black diamond (◆), and two UCYN-C sequences that are likely to amplify
1255 are marked with a circle (⊙). *nifH* cluster designations according to the convention in
1256 Zehr et al. (2003) are notated at the right. d1 – day 1; d11 – day 11, d13 – day 13, d15 –
1257 day 15, d22 – day 22, d23 – day 23.
1258
1259
1260



1261

1262 **Figure 2.** *Abundances of targeted diazotrophs at the 6m depth during the VAHINE*
 1263 *mesocosm experiment and in the Noumea lagoon (NL) during the experimental period.*

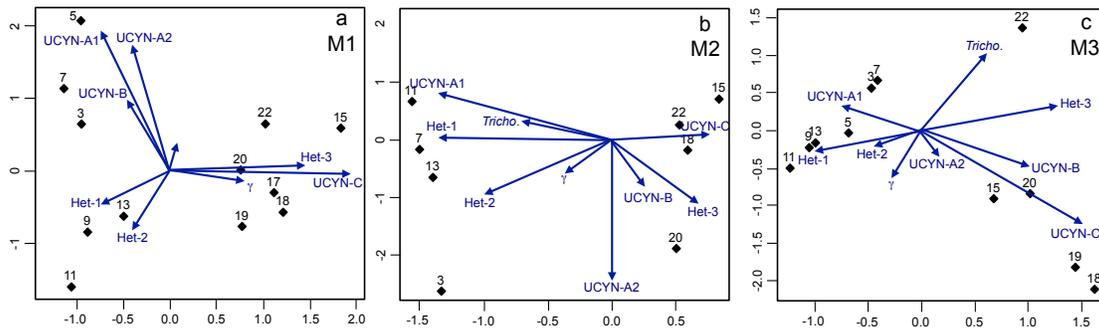
1264 A succession from a Het-1 dominated community to a UCYN-C dominated community is

1265 seen in all three mesocosms, M1 (b), M2 (c), and M3 (d), during P2. In the NL, growth of UCYN-C is not

1266 observed. DIP concentrations (a) decreased steadily following the spike at day 4 (data from in Bonnet et al., 2015a).

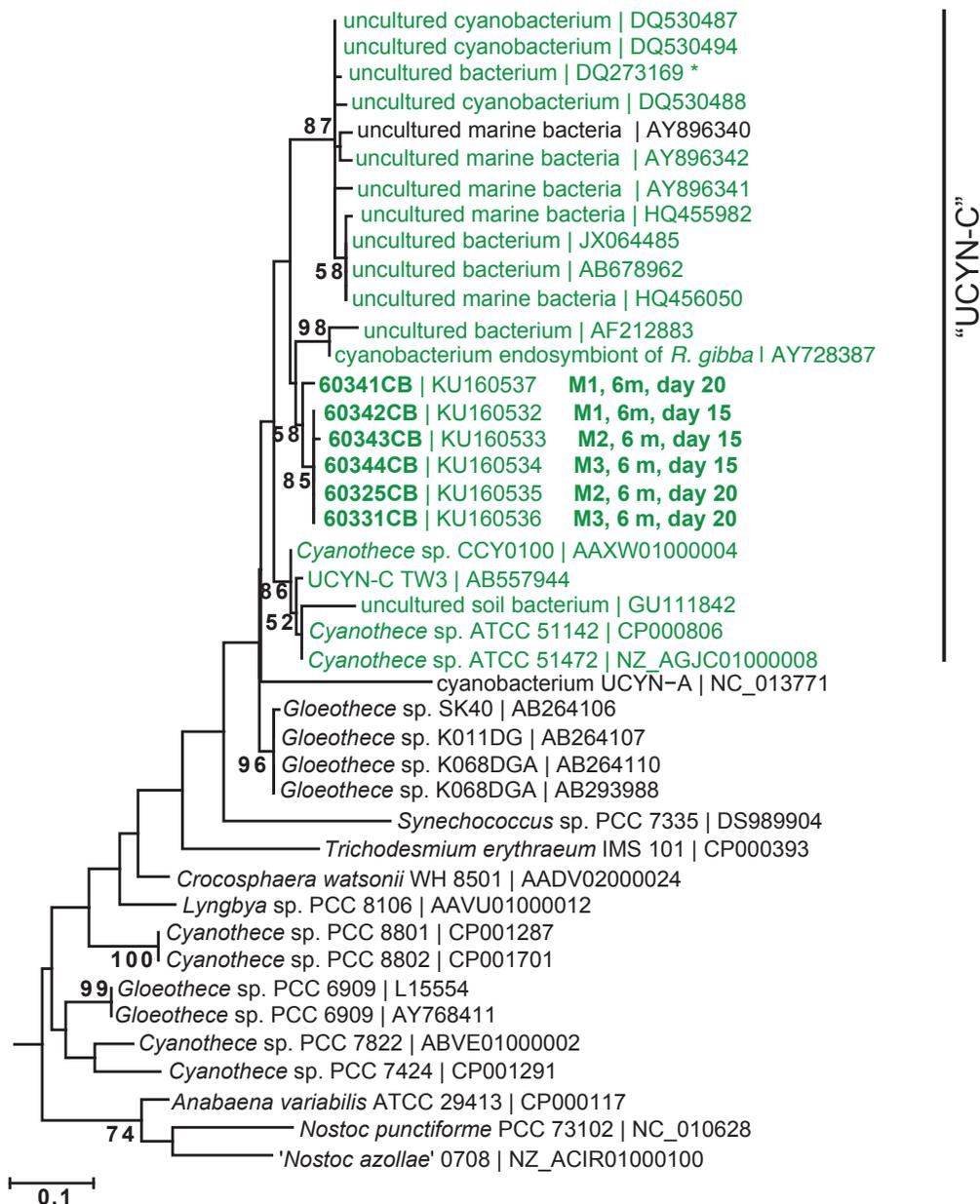
1267

1268
1269



1270
1271
1272
1273
1274
1275
1276
1277
1278

Figure 3. Correspondence analysis biplot of diazotroph abundances for each mesocosm. The horizontal axis is representative of time, evidenced by the progression of time points projected onto the x-axis. Variances covered by the two axes are 61%+18% in M1, 88%+5% in M2, and 56%+30% in M3.



1279

1280 **Figure 4. Maximum likelihood tree of Cyanothece-like diazotrophs based on partial**

1281 ***nifH* nucleotide sequences.** Sequences that are targeted by the UCYN-C qPCR assay

1282 (Foster et al., 2007) with no greater than 2 mismatches in each primer and probe

1283 sequence are green, and the original sequence used for sequence design is marked with an

1284 asterisk (*). *Cyanothece*-like sequences recovered from the VAHINE mesocosms are

1285 bold. Bootstrap trees were calculated using 1000 replicate trees, and nodes with values

1286 >50 are displayed. Branch lengths were inferred using the Tamura-Nei model, and the
1287 scale bar indicates the number of substitutions per site.