Limnol. Oceanogr., 59(6), 2014, 1899–1909

doi:10.4319/lo.2014.59.6.1899

Trichodesmium sp. clade distributions in the western North Atlantic Ocean

Mónica Rouco,[1,a] Hannah Joy-Warren,[1] Dennis J. McGillicuddy, Jr.,[2] John B. Waterbury,[1] and Sonya T. Dyhrman [1,3,*]

[1] Department of Biology, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts
[2] Department of Applied Ocean Physics & Engineering, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts
[3] Department of Earth and Environmental Sciences and the Lamont-Doherty Earth Observatory, Columbia University, New York

Abstract

To better understand the factors determining *Trichodesmium* clade (Clade III vs. Clade I) distribution, a quantitative polymerase chain reaction method was developed and applied to a sample set from the western North Atlantic. Clade distributions were compared with a suite of physicochemical parameters on a cruise transect from Woods Hole, Massachusetts, to Barbados during the spring of 2011. Clade I comprised > 90% of the total *Trichodesmium* spp. population. Depth distributions of both clades were similar in waters where there was a shallow (< 20 m) mixed-layer depth (MLD), but segregated in the water column where the MLD was deep (> 80 m). Surface abundances of both clades were higher when the MLD was < 20 m. Temperature and salinity were correlated with both clades in the data set as a whole. However, when stations influenced by a freshwater salinity anomaly were removed, dissolved inorganic phosphorus and total dissolved phosphorus correlated with Clade I cell density only. These correlations underscore the influence of phosphorus in driving *Trichodesmium* sp. dynamics in this region and are suggestive of niche partitioning between the two clades.

Populations of the nitrogen-fixing cyanobacteria *Trichodesmium* are a key component of the phytoplankton community in the North Atlantic Ocean. They provide an important source of new nitrogen to this oligotrophic system, influencing primary productivity and organic matter export to the deep ocean (Capone et al. 2005; Luo et al. 2012). Estimates attribute up to 1.6×10^{12} mol N yr^{-1} in the North Atlantic to the activities of *Trichodesmium* populations alone, which could equal or even exceed the estimated flux of NO_3^- from deep waters (Capone et al. 2005).

With its importance to the marine nitrogen (N) cycle, the *Trichodesmium* genus has been intensively studied. Populations of *Trichodesmium* spp. are known to thrive in stratified, warm waters with high light (Bergman et al. 2013). There are many factors that can influence their growth and N_2 fixation, including geochemical factors such as iron and phosphorus availability, or physical factors such as light, mixed-layer depth or temperature (Bergman et al. 2013 and references therein). However, *Trichodesmium* spp. abundance and contribution to primary productivity varies greatly both spatially and seasonally (Carpenter et al. 2004), making predictive modeling of the distribution and activities of this genus an ongoing challenge (Monteiro et al. 2010; McGillicuddy 2014).

There are many examples of physiological variability within a phytoplankton genus, or even species, driving the distribution and activities of key model systems such as *Prochlorochoccus* sp. (Moore et al. 1998), *Synechococcus* sp. (Palenik 2001), *Ostreococcus* sp. (Rodríguez et al. 2005), and *Emiliania* sp. (Read et al. 2013), among others. Yet the majority of studies investigating the distribution and activities of *Trichodesmium* populations in the North Atlantic Ocean have focused on methods that account for the *Trichodesmium* community as a whole, such as microscope counts of colonies (Luo et al. 2012), Video Plankton Recorder assays of colony abundance (Davis and McGillicuddy 2006), or molecular techniques targeting the *nifH* gene (e.g., Foster et al. 2007). As such, the *Trichodesmium* population is typically only examined at the genus level or by gross differences in colony morphology (the latter of which are known to contain different species [Carpenter et al. 2004; Luo et al. 2012]). Little is known about the environmental factors driving the abundance and horizontal, vertical, and temporal distribution of different *Trichodesmium* species.

Genetic analysis of laboratory cultures (using morphology, pigments, 16S rRNA, ITS, and *hetR*) showed that the genus *Trichodesmium* comprises 6 different species, which group in two major clades (Clade I and III [Hynes et al. 2012]) that are thought to dominate most field populations (Hmelo et al. 2012; Hynes et al. 2012). Clade I includes *T. thiebautii*, *T. tenue*, *T. hildebrandtii*, and *T. spiralis*, and Clade III includes *T. erythraeum* and *T. contortum* (Hynes et al. 2012). Most laboratory studies have focused on *T. erythraeum*, especially the model isolate IMS-101, as the main representative of *Trichodesmium* field populations, and then extrapolated findings to field communities (Bergman et al. 2013). Different *Trichodesmium* species can, however, have distinct physiological capabilities and responses to environmental factors (Dyhrman et al. 2006, 2009; Hynes et al. 2012). For example, studies suggest that *Trichodesmium* sp. responses to increasing CO_2 levels are species-specific (Hutchins et al. 2013).

In this study, a quantitative polymerase chain reaction (q-PCR) method, targeting the *rnpB* gene, was developed and validated in the laboratory, and was then used to

* Corresponding author: sdyhrman@ldeo.columbia.edu

[a] Present address: Lamont-Doherty Earth Observatory, Columbia University, New York

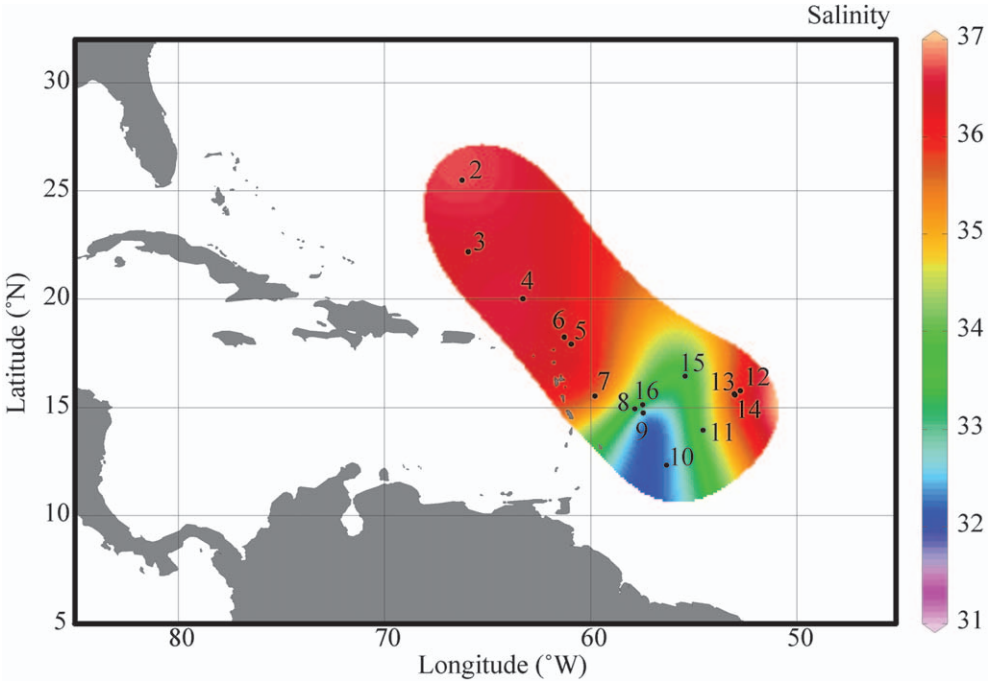

Fig. 1. Surface salinity map and station locations for an April–May 2011 cruise transect in the western North Atlantic (cruise OC471). Data were contoured using Ocean Data View 4.5.6 with the data intensive visualization and analysis (DIVA) grid method (R. Schlitzer; http://odv.awi.de).

investigate *Trichodesmium* Clade I and III abundance and distribution in the western North Atlantic during the spring of 2011. The western North Atlantic is strongly influenced by the Amazon and Orinoco rivers, which discharge low-salinity waters containing comparatively high levels of nutrients (Muller-Karger et al. 1988). The structure, abundance, and physiology of phytoplankton populations in some of these areas are highly affected by the dynamics of this system (Subramaniam et al. 2008). Populations of *Trichodesmium* sp. are often found at high concentrations in the western North Atlantic (Foster et al. 2007), but the environmental factors driving the abundance and distribution of the two major clades are unknown. In order to discern potential factors influencing *Trichodesmium* clade distribution in the western North Atlantic, field samples were collected from depth profiles along a cruise transect from Woods Hole, Massachusetts, to Barbados and clade abundances were compared with a suite of physicochemical parameters.

Methods

Field sample collection—Sampling took place on a cruise transect in the western North Atlantic Ocean aboard the R/V *Oceanus* (cruise OC471) in April and May 2011 (Fig. 1). Samples were collected at approximately the same time of the day from 10 liter bottles mounted on a Rosette sampling device at 15 stations and 5 depths (0, 20, 40, 60, and 80 m) per station. For *Trichodesmium* sampling, 10 liters of seawater was gravity-filtered through 47 mm, 10 μm polycarbonate filters. Both the silicone tubing and the polypropylene in-line filter holders used were washed with 10% bleach and rinsed with distilled water between each use. Filters for q-PCR analysis were stored in 2 mL cryovials, snap-frozen, and stored in liquid nitrogen until future analysis. Filters from duplicate bottles were collected for microscope counts. These filters were fixed with 4% buffered formalin, mounted on microscope slides, and stored at -20°C until microscope counts were performed. In situ profiles of temperature, salinity, photosynthetically active radiation, and density (sigma-theta) were measured by a conductivity, temperature, and depth instrument deployed at each station. Mixed-layer depth (MLD) was determined by the threshold criteria with a potential density difference of 0.125 kg m^{-3} (Monterey and Levitus 1997).

Biogeochemical assays—For nutrient determination (phosphate [PO_4], total dissolved phosphate [TDP], nitrate and nitrite [NO_2 + NO_3], ammonium [NH_4], and total dissolved nitrogen [TDN]), 125 mL of water was filtered through a 0.2 μm, 47 mm polycarbonate filter, and stored frozen (-20°C) in 10% HCl-cleaned bottles until analysis at the Chesapeake Bay Lab at the University of Maryland according to the facility's protocols. Dissolved organic nitrogen (DON) and dissolved organic phosphorus (DOP) were determined by the difference between the TDN and TDP and the respective inorganic pools of N and P.

Microscope counts—Filters from Niskin bottles paired with those for q-PCR were used for microscope counts. Counts of colony abundance were made at nine stations for all depths with a Zeiss epifluorescence microscope by imaging phycoerythrin autofluorescence. Selected stations were also assayed for *Trichodesmium* cell counts. This

Table 1. Primer sequences, product sizes, and efficiencies for the q-PCR reactions targeting the *rnpB* gene. Primer sequences were obtained from Chappell and Webb (2010). Efficiency is calculated as an average of all standard curves ($n = 10$). IMS-101 was used as the Clade III representative and VI-I was used as the Clade I representative.

Trichodesmium strain	Primers 5′-3′	Target size	Efficiency (%)
IMS-101	*Fwd: ACCAACCATTGTTCCTTCG, *Rev: CAAGCCTGCTGGATAACG	199	98.1±4.3
VI-I	*Fwd: GAATCTATGAACGCAACGGAAC, *Rev: ACCAGCAGTGTCGTGAGG	102	98.6±3.2

* Fwd: forward primer; Rev: reverse primer.

involved counting and measuring the length of each filament on the filter also using a Zeiss epifluorescence microscope and phycoerythrin autofluorescence. Images of every filament in the sample were taken using an AxioCam and AxioVision software. Filament counts and lengths were subsequently determined by imaging the entire filter and calculating the total cell number per filter using the following formula: Total cell number per filter = Σ (each filament length divided by average cell length). The average cell length was obtained, as explained below, from measurements of the culture species chosen as representatives of each clade. In addition, total cell number per colony was determined from 14 randomly selected colonies in the sample set. Colonies were defined as a collection of seven or more filaments, and colonies from a variety of sizes (from 9 to 95 filaments per colony) were considered. Every attempt was made to measure every filament from the colony, but these should be considered a conservative cell count, because it was difficult to image and measure all filaments.

Cell cultures—Trichodesmium erythraeum IMS-101 (isolated from the coastal waters off North Carolina in 1992) and *T. thiebuatii* VI-I (isolated at the Bermuda Atlantic Time Series Station in 1998) were selected as the culture representatives for Clade III and Clade I, respectively (Chappell et al. 2012). Additionally, an extra species belonging to Clade III (OC471_st_13.6), isolated from Sta. 13 by JBW, was cultured and used to validate the technique. Cultures were grown in Nalgene polycarbonate flasks in RMP growth media with a Sargasso Sea water base (Hynes et al. 2012) using methods described in Webb et al. (2001). Cultures were maintained at 25°C on a shaker table in a 12 : 12 dark : light cycle (12 h at 56 μmol quanta m^{-2} s^{-1} and 12 h dark). *T. erythraeum* IMS-101 and *T. thiebautii* VI-I cultures were present in the form of free filaments to enable reliable cell counts.

Validation experiments—To test for extraction efficiency, a standard curve was generated for each culture (*T. erythraeum* IMS-101 and *T. thiebautii* VI-I) from direct extractions of individual filters, each with a different cell concentration. The standard curves were then compared with samples diluted from the most concentrated filter, which mimicked the same concentrations as those of the independent filters. Deoxyribonucleic acid (DNA) from the cultured strain OC471_st_13. 6 was also extracted and used as a template for the q-PCR reactions.

Standard curves—Clade-specific standard curves were prepared from dilutions of duplicate filters of 30 mL of *T. erythraeum* IMS-101 and 30 mL of *T. thiebautii* VI-I harvested at a low vacuum and snap-frozen in liquid nitrogen. Cultures were harvested at similar concentrations (2733 cells mL^{-1} and 2221 cells mL^{-1} for IMS-101 and VI-I cultures, respectively). DNA extracts from the duplicate filters were pooled within each species and aliquoted to use for the analysis of the whole data set. All filters were stored in liquid nitrogen until the DNA extraction was performed. For cell counts, 10 0.5 mL samples from the cultures from which standards were taken were collected, and counts were performed using a Sedgewick Rafter slide. Each independent filament in the sampled volume was photographed using a Zeiss Axiocam high-resolution digital camera. Filament and cell length were measured using Axiovision 4.6.3 software. Cell counts per filament were obtained by dividing the total length of a filament by an average empirically determined cell length (6.37 ± 1.7 μm and 6.48 ± 1.35 μm, for *T. erythraeum* IMS-101 and *T. thiebautii* VI-I, respectively). Average cell length was determined by measuring all clearly distinct cells (between 10 and 40 per filament) from 10 randomly selected filaments of each cultured species.

DNA extraction—DNA from cultures and field samples was extracted using Power Plant Pro DNA Isolation Kit (MoBio laboratories, 13400-50), shown to be the most robust DNA extraction kit for *Trichodesmium* samples (Hynes 2009). Filters were transferred to bead tubes provided by the kit and extracted following the manufacturer's instructions. DNA from all samples was eluted in 100 μL of elution solution. DNA samples were kept at −20°C until quantitative PCR (q-PCR) analyses were performed.

Quantitative PCR—Primers targeting sequences of the *rnpB* gene of two of the major *Trichodesmium* clades (Clade III and Clade I) were used in this study. The *rnpB* gene encodes a ribonuclease P (RNaseP), which is a type of ribonuclease that cleaves RNA, and it is a common reference gene used in cyanobacterial studies because of its stable expression pattern (Chappell and Webb 2010). The clade-specific primers used here were designed and determined to be clade-specific for use in expression studies (Chappell and Webb 2010) and their sequences are indicated in Table 1. The efficiency of the amplification (E) was determined for each primer set as follows: $E = [10^{(-1/\text{slope})} - 1] \times 100$, and calculated as an average of all the standard curves used during the study ($n = 10$; Table 1). The q-PCR amplification of standards, field samples, and internal controls was performed in triplicate in a 96-well plate format on an iCycler iQ real-time detection system (Bio-Rad, Hercules) using the iQ SYBR®

Green SuperMix (Bio-Rad, Hercules). Internal controls included a negative control with template from the opposite clade to check for cross-clade reactions, and a blank with RNase-free water as a control to check for potential contamination. These controls were always negative. Both internal controls and standards were included in each plate. Each reaction was run in a final volume of 25 μL including 12.5 μL SYBR® Green SuperMix, 200 nmol L^{-1} forward and reverse primers, 2 μL of DNA, and 9.5 μL RNase-free water. Cycling conditions were 50°C for 2 min, 95°C for 10 min; 40 cycles of 95°C for 15 s, 55°C for 1 min with fluorescence read at 55°C, followed by dissociation curve analysis from 60°C to 95°C. Melt curves were used to confirm single amplification products, which was the case for all samples. The fluorescence threshold was set by the analytical software for the iCycler, and the PCR cycle during which this threshold was crossed was designated the threshold cycle (C_T) for each sample. The reported C_T was averaged for each triplicate sample or standard. A standard curve was generated for each plate using the C_T values and the known cell concentration of the standards.

To assure that all data fell within the standard curve range, all field samples were diluted at least 1 : 2. To verify that interference due to inhibitors present in the environment did not occur at our selected dilution, DNA from filters of two depths per station (usually 20 m and 60 m, to assure consistency) was used to test the q-PCR efficiency over a range of additional dilutions. Field samples that yielded a cell abundance of zero were spiked with a known concentration of standard DNA and assayed to confirm that negative results were not related to inhibition. Cell numbers for field samples were determined through interpolation from the standard curve, taking into account the dilution factor. Populations were defined as Clade III or Clade I, depending on whether they amplified with IMS-101 or VI-I primers, respectively.

Sequencing—Three randomly selected field stations were tested for primer specificity. In brief, three q-PCR products from Clade III and three q-PCR products from Clade I reactions were processed for sequencing. The q-PCR products were gel-extracted using the QIAquick gel-extraction kit (Qiagen), and direct-sequenced at the University of Maine (Orono) according to the facility's protocols. Sequence editing and alignments, using the local alignment search tool (BLASTn), were done with Geneious R6.1.3 software. Sequenced q-PCR products were aligned against *rnpB* sequences of both Clade III and Clade I obtained from Chappell and Webb (2010), with GenBank accession numbers: GQ415327-GQ415339. Gene sequences have been deposited to the European Nucleotide Archive, http://www.ebi.ac.uk/ena/data/view/HG974235-HG974240, with the accession numbers HG974235-HG974240.

Statistical analyses—Pair-wise correlation coefficients (Pearson product moment correlation) were calculated between Clade III density or Clade I density and the rest of the metavariables, using Sigma Plot 11.0 (Systat Software).

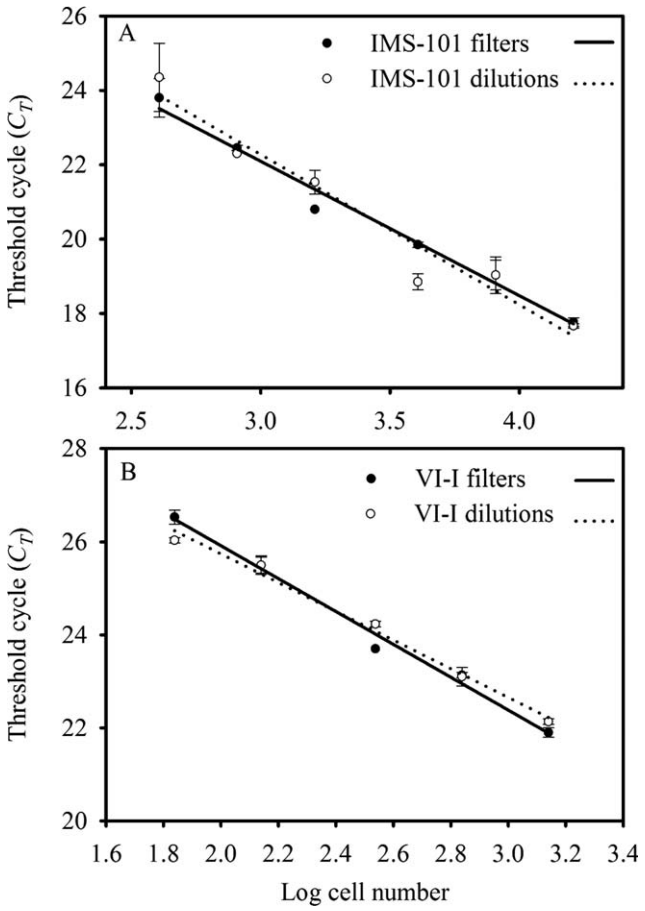

Fig. 2. Extraction efficiency tests comparing the q-PCR signal for samples generated by direct extraction of a filter (filters) or dilution of a concentrated extract to mimic the filter value (dilutions) for (A) *Trichodesmium erythraeum* IMS 101 and (B) *T. thiebautii* VI-I. Error bars indicate standard deviations of triplicate technical replicates.

Results

Validation and quality control—For each culture (*T. erythraeum* IMS-101 and *T. thiebautii* VI-I), the standard curves generated from direct extractions of individual filters yielded average C_T values almost identical to the standard curves created with samples diluted from the most concentrated filter (Fig. 2) with near 100% efficiency (Table 1). In short, whether the concentration in the sample was derived from a direct extraction or diluted to that value, the result was indistinguishable (Fig. 2), indicating no issues with extraction efficiency over the target range. Both duplicate standard curves from the IMS-101 culture yielded C_T values almost identical to both duplicate standard curves from the VI-I cultures, which were obtained from a culture with cell concentrations almost identical to the IMS-101 culture (Fig. 3). None of the blanks (Milli-Q) or negative controls (e.g., PCR reactions in which DNA from one of the clades was combined with primers from the other clade) amplified in any of the plates, demonstrating the specificity of the technique. Field samples that yielded a cell abundance of

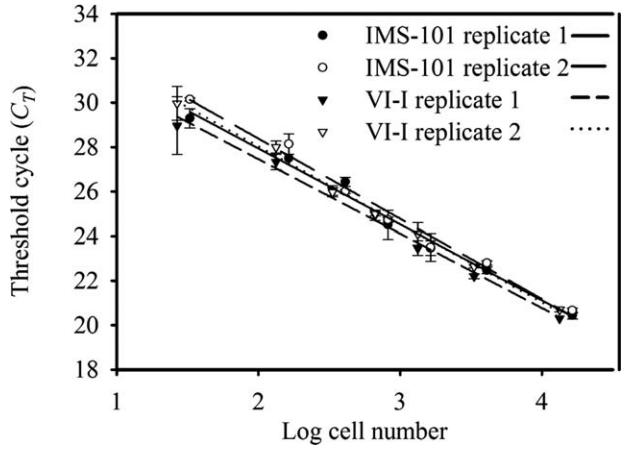

Fig. 3. Standard curves from q-PCR reactions for *Trichodesmium erythraeum* IMS-101 and *T. thiebautii* VI-I. Error bars indicate standard deviations of triplicate samples. Biological replicates are plotted separately.

zero and that were spiked with known additions of either *T. erythraeum* or *T. thiebautii* ($n = 4$) generated expected C_T values, indicating true negatives in these cases. Melt curves confirmed single products in all field samples at a consistent melting point. Lastly, q-PCR efficiencies calculated from dilutions of field samples (the 20 m and 60 m filters per station) fell within the efficiency range of the standards, showing no evidence of PCR inhibitors.

Total *Trichodesmium* cell abundance (Clade I + Clade III) assayed by q-PCR was highly correlated with total colony number measured by microscopy in nine randomly picked stations ($r = 0.96$; $p < 0.001$; Fig. 4A). Cell abundances reported by microscope counts were 2836 cells filter^{-1} at Sta. 4 (80 m), 1610 cells filter^{-1} at Sta. 5 (40 m), 7780 cells filter^{-1} at Sta. 10 (60 m), and 3866 cells filter^{-1} at Sta. 15 (80 m; Fig. 4B). Cell abundances reported by q-PCR (Clade III + Clade I estimates), performed on filters from the same stations and depths as were tested by microscopy, were 12,530 cells filter^{-1} at Sta. 4 (80 m); 4915 cells filter^{-1} at Sta. 5 (40 m); 6502 cells filter^{-1} at Sta. 10 (60 m); and 2575 cells filter^{-1} at Sta. 15 (80 m; Fig. 4B). Depending on the size of the colony, total cell-number estimates for a colony varied from 633 cells to 11,385 cells, with the latter representing a conservative estimate of the maximum variability in cell number driven by the presence of a single colony.

All Clade III sequenced products had between 99% and 100% identity to other members of the clade, including *T. erythraeum* GBRTRLI101 (GQ415327), *T. erythraeum* 6-5 (GQ415328), *Trichodesmium* sp. K11-131 (GQ415329), *T. contortum* 21-74 (GQ415330), *T. contortum* 20-70 (GQ415331), and *Trichodesmium* sp. XII-13 (GQ415332). All Clade I sequenced products had 100% identity to other members of the clade, including *T. tenue* Z-1 (GQ415333), *T. tenue* H9-4 (GQ415334), *T. thiebautii* VI-I (GQ415335), *T. thiebautii* II-3 (GQ415336), *T. spiralis* 20-71 (GQ415337), *Trichodesmium* sp. 5-1 JW (GQ415338), and *Trichodesmium* sp. 1-2 (GQ415339). In addition, DNA extracted from the other cultured Clade III strain (OC471_st_13. 6), isolated from Sta. 13 on this transect, amplified with Clade III *rnpB* primers but not with the Clade I primers used in this study.

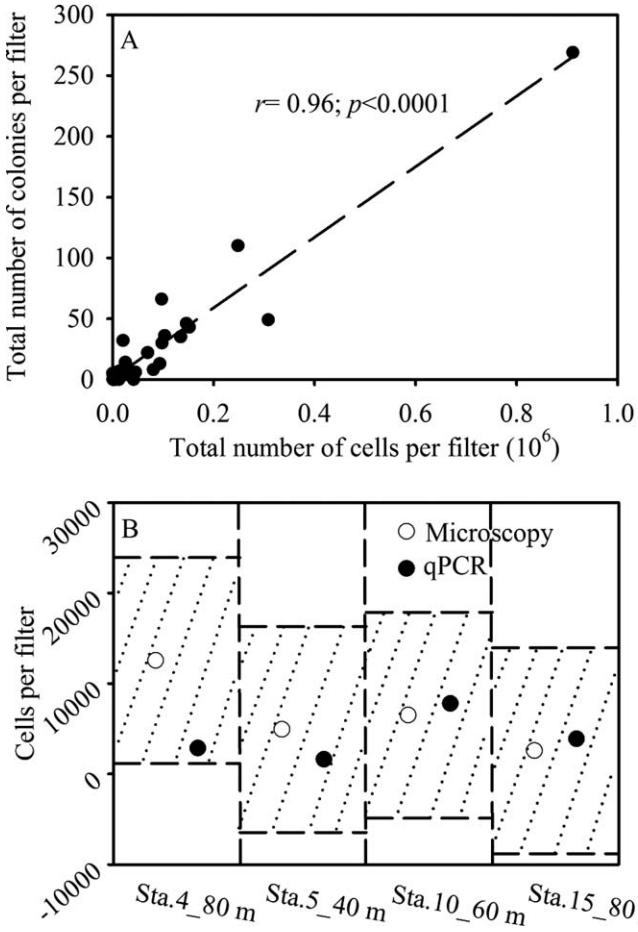

Fig. 4. (A) Relationship between total *Trichodesmium* cell abundance (Clade I + Clade III combined) assayed by q-PCR and total colony number detected by microscopy at nine randomly selected stations. (B) Total number of cells at four selected stations estimated by microscopy and by the q-PCR methodology. The size of the shaded area is calculated independently for each station by the addition and subtraction of 11,385 cells (the colony cell count empirically derived herein) from the q-PCR value. The shaded area includes any possible value that could be expected from the q-PCR filter if a colony is included or excluded in the 10 liter bottle that supplies the q-PCR sample.

Trichodesmium *clade-type abundance and distribution—Trichodesmium* populations were dominated by Clade I cells along the cruise transect (Fig. 5). The average Clade I cell concentration was 7197 ± 17,871 cells L^{-1}, ranging from 85 cells L^{-1} to 113,576 cells L^{-1} (Fig. 5B), whereas the average Clade III cell concentration was 199 ± 550 cells L^{-1}, ranging from 0 cells L^{-1} to 3798 cells L^{-1} (Fig. 5A). Concentrations were highest in the upper 20 m of Sta. 8, 9, 10, and 16, where there was a salinity anomaly (Figs. 1, 5). Surface cell densities ranged from 99 cells L^{-1} to 90,147 cells L^{-1} for Clade I and from 0 cells L^{-1} to 1675 cells L^{-1} for Clade III populations. Near surface (0 and 20 m), concentrations represented maxima in the profiles (Clade

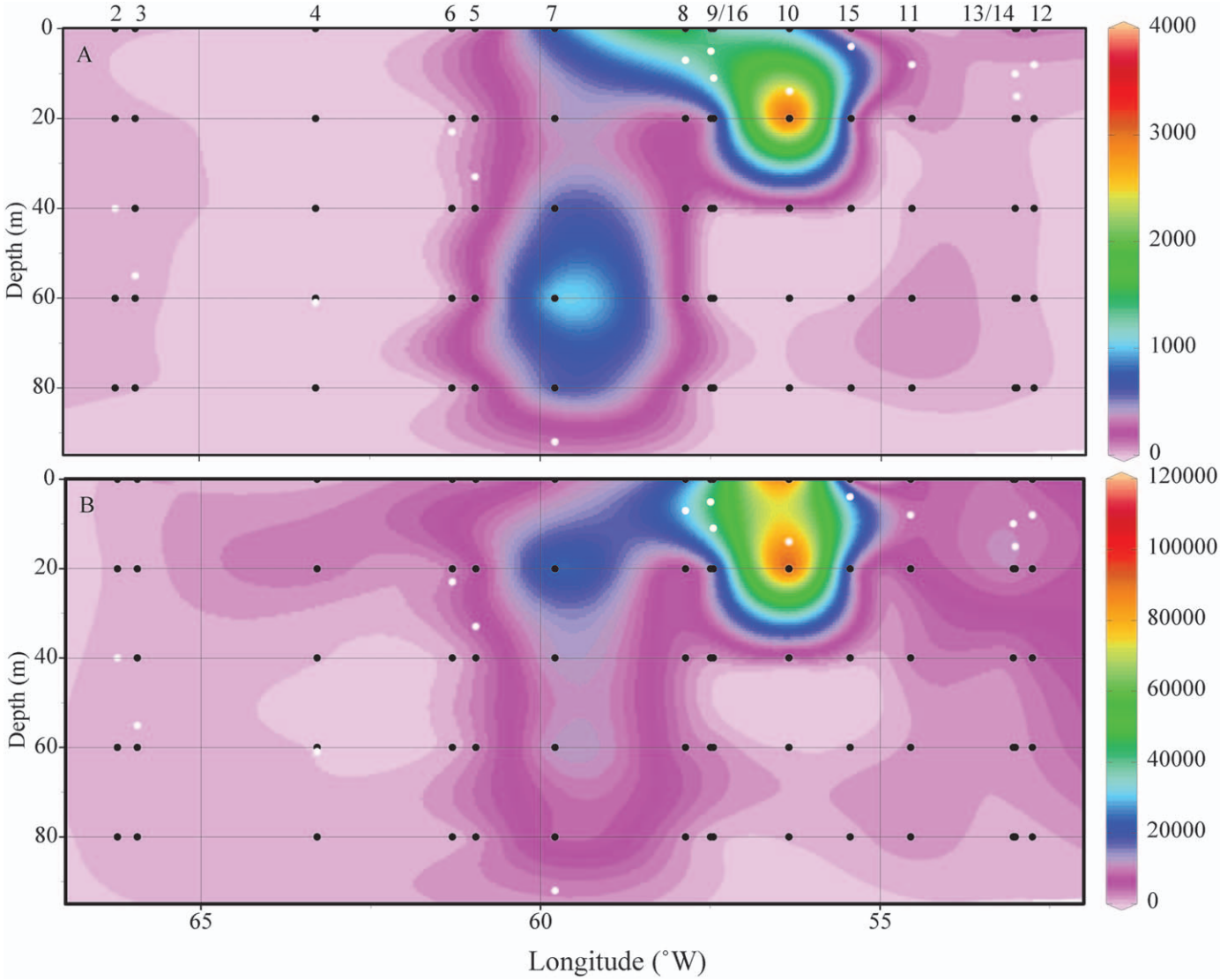

Fig. 5. (A) Clade III and (B) Clade I cell density (cells L^{-1}) in the upper 80 m. Black dots show where actual samples were taken at each station and white dots indicate mixed-layer depth. Station number is indicated above the panel. Data were contoured using Ocean Data View 4.5.6 with the DIVA grid method (R. Schlitzer; http://odv.awi.de). Note that the scales are different in panels A and B. Color bar scale indicates cells L^{-1}.

III: 424 ± 581 cells L^{-1} and 351 ± 968 cells L^{-1} for 0 m and 20 m, respectively; Clade I: 15,903 ± 24,476 cells L^{-1} and 14,316 ± 28,424 cells L^{-1} for 0 m and 20 m, respectively) with the exception of a Clade III maximum (1268 cells L^{-1}) at 60 m for Sta. 7 (Figs. 5, 6).

Although there was some depth segregation between clades at Sta. 7, the clades followed similar trends in the data set as a whole (Fig. 6). A significant positive correlation was found between Clade III and Clade I cell concentrations along the cruise transect (Table 2; $r = 0.84$, $p < 0.001$), but the strength of this correlation decreased when stations highly influenced by the riverine plume were removed (Table 2; $r = 0.63$, $p < 0.001$).

Trichodesmium *clade-type abundance and distribution as a function of physicochemical parameters*—The mixed-layer depth (MLD) along the cruise transect ranged from 23 m to 61 m between Sta. 2 and 6, and deepened to 92 m at Sta. 7. Sta. 7 was the only station where relatively high cell concentrations of both clades were observed at all depths. The MLD was very shallow, ranging from 4 m to 15 m, from Sta. 8 to 16. With the exception of Sta. 7, surface abundances of both clades were higher when the MLD was < 20 m (Fig. 7).

The average surface temperature in the system was 26.47 ± 1.19°C, with surface temperature values ranging between 24.6°C to 28.8°C, with the highest surface temperature at Sta. 10 (28.75°C). A wide range in surface salinity was observed along the cruise transect, with values as high as 36.7 at Sta. 2 and as low as 31.2 at Sta. 16, representing the strong influence from the Amazon river plume along this transect (Fig. 1). None of the dissolved nutrient concentrations (NH_4, NO_2+NO_3, TDN, TDP, PO_4, DOP, Si) were correlated with salinity. Both Clade III and Clade I

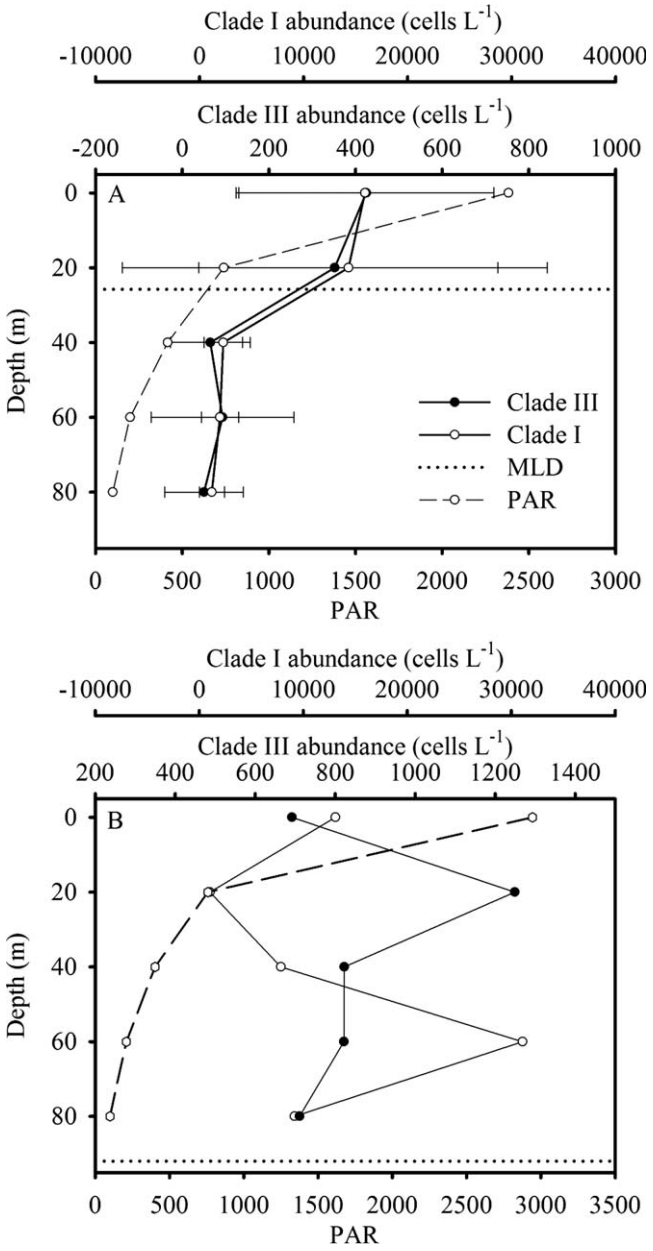

Fig. 6. (A) Mean vertical distributions of Clade III and Clade I cell density (cells L^{-1}) and mean photosynthetically active radiation (PAR) across the entire transect. (B) Vertical distribution of Clade I and Clade III cell density (cells L^{-1}) and PAR at Sta. 7. The dotted line corresponds to the mean mixed-layer depth (MLD) across (A) the entire transect, and (B) the MLD at Sta. 7. Note the subsurface peak in Clade III cell abundance at 60 m. Error bars represent 95% confidence intervals. Note that the 95% confidence intervals are high at the surface because they include values from the entire transect, which vary from 85 cells L^{-1} to 113,576 cells L^{-1} for Clade I populations and from 0 cells L^{-1} to 3798 cells L^{-1} for Clade III populations.

populations showed the highest abundances at stations most strongly influenced by the river plume (Fig. 1). A significant correlation between the abundance of each clade with salinity (Table 2; $r = -0.64$, $p < 0.001$ and $r = -0.74$, $p < 0.001$, for Clade III and Clade I, respectively) and temperature (Table 2; $r = 0.43$, $p < 0.001$ and $r = 0.49$, $p < 0.001$, for Clade III and Clade I, respectively) was apparent when all stations were considered. When the stations highly influenced by the riverine plume were excluded (surface salinity < 34.16), a significant correlation was still observed between the cell abundance of both clades with salinity and temperature (Table 2). However, with the low-salinity stations removed, there was a significant correlation between Clade I abundance and TDP (Table 2; $r = 0.49$; $p < 0.001$), PO_4 (Table 2; $r = 0.30$; $p = 0.046$) and DOP (Table 2; $r = 0.42$; $p = 0.004$) that was not present for Clade III.

Discussion

Despite many years of study, and the potential for substantial heterogeneity, the vertical and horizontal distribution of particular *Trichodesmium* clades remains largely unexplored. Herein, a q-PCR assay for the two dominant *Trichodesmium* clades was developed and applied to investigate their abundance, vertical and horizontal distribution, and variation with various physical and chemical parameters across a cruise transect in the western North Atlantic Ocean.

q-PCR assay—The q-PCR technique used here was extensively validated to assure as reliable results as possible, including checking extraction efficiency, and screening field samples from each station for PCR inhibition. Extraction efficiency was robust over a wide range of samples from all stations and controls. C_T values from all of the field samples fell within the range assayed by the standard curves, and standards and field samples always amplified with near 100% efficiency. Clade I and Clade III q-PCR products, amplified from randomly selected field samples, yielded sequences between 99% and 100% identical to other known members of their same clade, indicating primer specificity. In addition, only primers for the Clade III amplified DNA extracted from a *T. erythraeum* culture derived from Sta. 13 (OC471_st_13. 6 strain), which was identified according to morphology and pigment composition.

Determination of colony abundance is one of the standard metrics used for the study of *Trichodesmium* abundance in field communities (Luo et al. 2012). Total *Trichodesmium* cell abundance (Clade I and Clade III combined) assayed by q-PCR correlated with total colony abundance assayed by microscopy, indicating that the q-PCR method is a robust assay for the determination of *Trichodesmium* abundance. Direct comparisons of cell counts for this genus are challenging because *Trichodesmium* is filamentous and forms colonies; thus, cell counts are a derived value and have to be calculated from filament lengths and cell length. Regardless, similar values were observed for two of four comparison samples (Sta. 10 and 15). Two of the samples (Sta. 4 and 5) showed greater variability, with q-PCR overestimating the cell count relative to the microscope values. There are several possibilities that could explain discrepancies between these two methodologies. Microscope counts were done using an

Table 2. Pearson correlation coefficients (coef.) between Clade III (ELD) or Clade I (TLD) and assayed physical and chemical parameters. Significant correlations are indicated in bold.

	TLD	NH_4	NO_2+NO_3	TDN*	TDP*	PO_4	DOP*	Si	PAR*	Sal*	Temp*
All stations†											
ELD											
Pearson coef.	0.84	−0.06	−0.08	0.13	0.10	−0.01	0.14	−0.02	0.16	−0.64	0.43
p-value	**<0.001**	0.60	0.49	0.27	0.38	0.90	0.24	0.90	0.18	**<0.001**	**<0.001**
TLD											
Pearson coef.		−0.07	−0.08	0.10	0.15	−0.03	0.22	0.02	0.22	−0.74	0.49
p-value		0.57	0.48	0.38	0.20	0.79	0.061	0.84	0.06	**<0.001**	**<0.001**
Low riverine influence†											
ELD											
Pearson coef.	0.63	−0.16	−0.11	0.24	0.12	0.27	0.06	−0.09	0.03	−0.66	0.42
p-value	**<0.001**	0.30	0.47	0.12	0.44	0.08	0.71	0.56	0.86	**<0.001**	**<0.001**
TLD											
Pearson coef.		−0.17	−0.13	0.16	0.49	0.30	0.42	−0.09	0.09	−0.74	0.61
p-value		0.26	0.42	0.30	**<0.001**	**0.046**	**0.004**	0.55	0.56	**<0.001**	**<0.001**

* TDN: total dissolved nitrogen; TDP: total dissolved phosphorus; PAR: photosynthetically active radiation; DOP: dissolved organic phosphorus; Sal: salinity; Temp: temperature.

† All stations: Analyses include data from all stations ($n = 74$). Low riverine influence: Analyses exclude data from stations highly influenced by the riverine plume (e.g., surface salinity < 34.16; $n = 44$).

empirically derived cell length from the culture representatives, because it is not possible to accurately measure cell divisions in each of the field filaments. Variations in this number have been observed for *Trichodesmium* species other than *T. erythraeum* and *T. thiebautii* (Hynes et al. 2012), and could be a source of error. However, in order to explain the values observed here, the cell length in *Trichodesmium* populations would need to be 3–4 times lower than those of our culture representatives, and this has not been observed for any *Trichodesmium* isolates to date (Hynes et al. 2012). Variability in the number of *rnpB* copies per cell could also explain discrepancies between methodologies if more than one *rnpB* copy per *Trichodesmium* cell is present. This is a potential source of error with all q-PCR assays in the field, which is a common and broadly applied tool for the determination of microbe abundance in field populations (Church et al. 2005; Dyhrman et al. 2010). The gene is present as a single copy in the *T. erythraeum* IMS-101 genome, and appears to exist as a single copy in *T. thiebautii* VI-I as well. The latter strain was isolated from the study region; and taken together, this evidence suggests that copy number may be consistent across field populations. Two copies of the *rnpB* gene could also be present in a cell if it had fully replicated its DNA without having completed division. Again, this caveat applies to any q-PCR methodology in the field. In order to observe a two-fold overestimate of the number of cells with the q-PCR relative to microscopy, all the cells must be synchronously dividing and captured at this point at the time of sampling, which is not likely to happen in natural populations. Also, both of the standard curves were prepared with exponentially growing cultures; and thus, the effect of growth rate and cell cycle should be accounted for to some extent in this study. Finally, discrepancies between methodologies could be derived from the variable distribution of *Trichodesmium* filaments or colonies in the 10 liter sample volume. *Trichodesmium* populations are dilute, with patchy distributions of colonies that have upward of 11,000 cells each on this cruise transect. The inclusion or exclusion of one single colony in any of the duplicate 10 liter samples used individually for q-PCR or microscopy could account for the observed differences in this study. Despite the variability linked to estimates of cell counts, q-PCR shows excellent agreement with colony number across the data set as a whole.

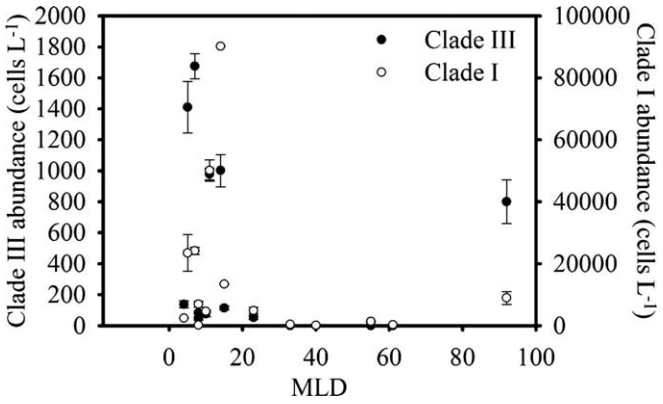

Fig. 7. Surface (0 m) Clade III and Clade I cell density (cells L^{-1}) as a function of mixed-layer depth (MLD). The highest abundances are associated with MLD < 20 m, with the exception of an anomaly found at Sta. 7 (92 m).

Trichodesmium *clade abundance*—Total surface *Trichodesmium* cell densities along the cruise transect (0 to 93,945 cells L^{-1}), obtained by adding the Clade III and Clade I counts, fell within the range found by other studies in the area (Luo et al. 2012). This suggests that the majority of the *Trichodesmium* in this region is falling into one of the two clades assayed herein. Along the cruise transect, Clade I

populations were the most abundant, with cell densities between 10 and 90 times higher than the Clade III communities. Previous studies of clade-specific iron-limitation responses on a cruise transect in the North Atlantic found that Clade I populations in surface waters were the most biologically active group, representing between 99.8% and 100% of all the cDNA quantified in the samples (Chappell et al. 2012). These results were corroborated by an alternative clade-specific q-PCR analysis of bulk water samples in the same stations (Hynes 2009). Several surface colonies collected in a 16S rDNA survey at the Bermuda Atlantic Time Series were also found to be mostly related (> 99% identity) to *T. thiebautii* (Hmelo et al. 2012).

T. erythraeum strain IMS-101 is by far the most commonly studied isolate of *Trichodesmium* (Bergman et al. 2013) because it has been in culture for many years and was the source strain for the only *Trichodesmium* genome sequence that is currently available (http://genome.jgi-psf.org/trier/trier.home.html). The dominance of Clade I populations in the western North Atlantic is striking; and if this is a consistent pattern, it calls into question efforts in predictive modeling based on strain IMS-101 alone. A more comprehensive seasonal and geographic survey is warranted to assess whether the dominance of the Clade I populations is common.

Trichodesmium *clade distribution*—Both clades were distributed throughout the upper 80 m of the water column. The highest surface abundances of both *Trichodesmium* clades were associated with shallow MLD, because it has been previously observed for the *Trichodesmium* genus in the eastern North Atlantic (Tyrrell et al. 2003) and central North Atlantic Ocean (Sañudo-Wilhelmy et al. 2001). Similarly, studies in the North Pacific showed a drastic decrease in *Trichodesmium* spp. populations below the MLD (Church et al. 2005).

Cell abundance of both clades had similar near-surface maxima at all stations, with the exception of Sta. 7 where there was an unusually deep mixed layer associated with an anticyclonic mesoscale eddy (Olson 2014). The depth segregation observed at this station hints at the possibility of niche partitioning between clades. Sta. 7 had the deepest mixed layer (92 m) observed along this cruise transect, which might allow for potential segregation of species in Clade III as a function of light or other physical and chemical factors. Subsurface peaks of Clade III populations were also found by a similar study in the North Atlantic, where MLDs were generally deeper than in this study (Hynes 2009). Clade III populations might represent two different species, or even ecotypes, with unique light-adaptation strategies. This might be indicated by differences in pigment characteristics (such as the Phycourobilin : Phycoerythrobilin ratio) that are observed between species of the two clades (Chappell et al. 2012). This characteristic has been previously observed in another abundant cyanobacterial genus in the oligotrophic regions, *Prochlorococcus* sp., whose isolates occupy two different niches based on light intensity and have differences in pigment ratios (Moore et al. 1998). Further investigation is required to understand whether light, or other factors that vary with depth, might drive partitioning between and within *Trichodesmium* clades.

The western North Atlantic is highly influenced by fluvial inputs from the Amazon River (Muller-Karger et al. 1988). Riverine water with high nutrient concentrations and low salinities can be transported hundreds of kilometers away from the Brazilian coast, creating a unique environment that can affect the structure and abundance of the phytoplankton communities in the area (Carpenter et al. 1999; Subramaniam et al. 2008). A strong correlation was found between the abundance of both clades and salinity and temperature, and the highest concentration of both Clade III and Clade I cells were found in stations where salinities were between 31.23 and 33.94 (e.g., 0 m at Sta. 8, 9, 10, and 16). This indicates that factors influenced by the presence of the Amazon river plume, such as iron (Tovar-Sanchez et al. 2006), nutrient concentration, or stratification, are potentially driving the growth and cell abundance of the *Trichodesmium* spp. population as a whole, consistent with previous studies that also observed an increase in cell densities of *Trichodesmium* spp. communities in this area (Carpenter et al. 1999; Foster et al. 2007; Subramaniam et al. 2008). None of the dissolved nutrient concentrations measured here (NH_4, NO_2+NO_3, TDN, TDP, PO_4, DOP, Si) correlated with salinity, indicating that other resources, such as iron, might be driving the observed distributions.

To better resolve potential physicochemical drivers of *Trichodesmium* spp. distribution outside of the riverine influence, stations within the riverine plume (surface salinity < 34.16) were removed from the data set, and correlation analyses were performed again. In the smaller data set, cell abundances of both clades still correlated with salinity and temperature, which underscores the association of these two factors with processes driving the distribution and growth of populations of both clades.

In the reduced data set, Clade I positively correlated with various phosphorus pools, including TDP, PO_4, and DOP. This pattern highlights the importance of phosphorus in driving *Trichodesmium* sp. growth and N_2 fixation in this low-phosphorus system. These results are also suggestive of a potential role for phosphorus in driving niche partitioning between clades. Relatively little is known about clade-specific responses to phosphorus availability. Common phosphorus-regulated genes related to phosphate scavenging and DOP hydrolysis are common in species from both clades (Orchard et al. 2009). Although the genes required for phosphonate transport and metabolism have been observed in the genomes of species from the two clades (*T. erythraeum*, *T. thiebautii*, *T. tenue*, and *T. spiralis* [Dyhrman et al. 2006]), only *T. erythraeum* seems to produce phosphonate compounds (Dyhrman et al. 2009). It is possible that the apparent ability to produce phosphonates is a factor in the lack of correlation observed between Clade III and the phosphorus pools. Investigation of the abundance and distribution of both clades in areas with higher phosphorus availability, such as the North Pacific, could help to further elucidate the role of phosphorus in defining niche partitioning between the two *Trichodesmium* clades, as well as the factors controlling the distribution and abundance of Clade III. In addition, it is important to

consider the role of the heterotrophic bacterial epibionts associated with *Trichodesmium* colonies in interpretation of these correlations. Cell to cell communication in the form of quorum sensing can dramatically influence the activity of biogeochemically relevant enzymes, such as alkaline phosphatase, involved in hydrolyzing DOP within *Trichodesmium* colonies (Van Mooy et al. 2012). Differences in the community structure of epibionts between *Trichodesmium* clades might then play an important role in defining niche spaces, because the presence of one epibiont group vs. another might determine inter-clade differences in the utilization of certain phosphorus compounds. Further laboratory and field studies are required in order to elucidate the potential influence of phosphorus on the definition of niche spaces within the *Trichodesmium* genus.

There are a number of potential factors that could influence *Trichodesmium* clade distribution that were not assayed here, such as iron concentration and epibiont diversity. Iron availability is well-known to be an important factor controlling *Trichodesmium* growth (Chappell and Webb 2010; Chappell et al. 2012), and could play a role in determining the distributions of different clades in this system. However, previous studies saw no evidence for iron stress in either clade in samples from this region (Chappell et al. 2012), indicating that this factor may not be a consistent driver of clade niche differentiation in the study area. As highlighted above, epibiont diversity could be substantial (Hmelo et al. 2012) and has been shown to influence the activity of enzymes in the colony (Van Mooy et al. 2012). The extent to which epibiont diversity or activity can influence *Trichodesmium* physiology and distribution should be examined in the future. Lastly, mesoscale eddies can affect physicochemical regimes, and recent studies have demonstrated how these features affect the distribution and N_2 fixation of *Trichodesmium* spp. populations as a whole, mainly as a consequence of changes in nutrient regimes or eddy-driven transport (Davis and McGillicuddy 2006; Agawin et al. 2013). Further investigation could focus on defining the influence of mesoscale physical variability on the abundance and distribution of the two clades.

To summarize, Clade I populations are by far the most abundant clade observed in this study of the western North Atlantic. The data herein suggest possible niche partitioning between clades with depth, and as a function of phosphorus, but greater seasonal and spatial resolution of *Trichodesmium* clade distribution is required to fully resolve niche segregation in this genus. The q-PCR method developed and validated here will facilitate further inquiry into the distribution and possible niche partitioning within this genus in this and other systems. In addition, as more *Trichodesmium* isolates and sequence information become available, a higher resolution approach at the species level may be possible to further understand the eco-physiology of this genus.

Acknowledgments

We thank A. Heithoff, L. Wurch, and E. Olson, for participating in the sample collection; and the captain and crew of the R/V *Oceanus* for their help at sea. We also thank S. Haley and K. Frischkorn for valuable technical assistance, and the two anonymous reviewers that provided comments and feedback on earlier versions of this manuscript.

This work was supported by the National Science Foundation Biological Oceanography Program (Ocean Sciences-0925284). M. Rouco was additionally supported by the Ramon Areces Foundation and The Center for Microbial Oceanography: Research and Education.

References

Agawin, N., A. Tovar-Sanchez, K. Knoth de Zarruk, C. Duarte, and S. Agustí. 2013. Variability in the abundance of *Trichodesmium* and nitrogen fixation activities in the subtropical NE Atlantic. J. Plankton Res. **35:** 1126–1140, doi:10.1093/plankt/fbt059

Bergman, B., G. Sandh, S. Lin, J. Larsson, and E. J. Carpenter. 2013. *Trichodesmium*—a widespread marine cyanobacterium with unusual nitrogen fixation properties. FEMS Microbiol. Rev. **37:** 286–302, doi:10.1111/j.1574-6976.2012.00352.x

Capone, D. G., and others. 2005. Nitrogen fixation by *Trichodesmium* spp.: An important source of new nitrogen to the tropical and subtropical North Atlantic Ocean. Glob. Biogeochem. Cycles **19:** GB2024, doi:10.1029/2004GB002331

Carpenter, E. J., J. Montoya, J. Burns, M. R. Mulholland, A. Subramaniam, and D. G. Capone. 1999. Extensive bloom of a N_2-fixing diatom/cyanobacterial association in the tropical Atlantic Ocean. Mar. Ecol. Prog. Ser. **185:** 273–283, doi:10.3354/meps185273

———, A. Subramaniam, and D. G. Capone. 2004. Biomass and primary productivity of the cyanobacterium *Trichodesmium* spp. in the tropical N Atlantic Ocean. Deep Sea Res. Part I **51:** 173–203, doi:10.1016/j.dsr.2003.10.006

Chappell, P. D., J. W. Moffett, A. M. Hynes, and E. A. Webb. 2012. Molecular evidence of iron limitation and availability in the global diazotroph *Trichodesmium*. ISME J. **6:** 1728–1739, doi:10.1038/ismej.2012.13

———, and E. A. Webb. 2010. A molecular assessment of the iron stress response in the two phylogenetic clades of *Trichodesmium*. Environ. Microbiol. **12:** 13–27, doi:10.1111/j.1462-2920.2009.02026.x

Church, M., B. Jenkins, D. Karl, and J. Zehr. 2005. Vertical distributions of nitrogen-fixing phylotypes at Stn Aloha in the oligotrophic North Pacific Ocean. Aquat. Microb. Ecol. **38:** 3–14, doi:10.3354/ame038003

Davis, C. S., and D. J. McGillicuddy. 2006. Transatlantic abundance of the N_2-fixing colonial cyanobacterium *Trichodesmium*. Science **312:** 1517–1520, doi:10.1126/science.1123570

Dyhrman, S. T., C. R. Benitez-Nelson, E. D. Orchard, S. T. Haley, and P. J. Pellechia. 2009. A microbial source of phosphonates in oligotrophic marine systems. Nat. Geosci. **2:** 696–699, doi:10.1038/ngeo639

———, P. D. Chappell, S. T. Haley, J. W. Moffett, E. D. Orchard, J. B. Waterbury, and E. A. Webb. 2006. Phosphonate utilization by the globally important marine diazotroph *Trichodesmium*. Nature **439:** 68–71, doi:10.1038/nature04203

———, S. T. Haley, J. A. Borchert, B. Lona, N. Kollars, and D. L. Erdner. 2010. Parallel analyses of *Alexandrium catenella* cell concentrations and shellfish toxicity in the Puget Sound. Appl. Environ. Microbiol. **76:** 4647–4654, doi:10.1128/AEM.03095-09

Foster, R., A. Subramaniam, C. Mahaffey, E. Carpenter, D. G. Capone, and J. Zehr. 2007. Influence of the Amazon River plume on distributions of free-living and symbiotic cyano-

bacteria in the western tropical North Atlantic Ocean. Limnol. Oceanogr. **52:** 517–532, doi:10.4319/lo.2007.52.2.0517

Hmelo, L., B. Van Mooy, and T. Mincer. 2012. Characterization of bacterial epibionts on the cyanobacterium *Trichodesmium*. Aquat. Microb. Ecol. **67:** 1–14, doi:10.3354/ame01571

Hutchins, D. A., F. Fei-Xue, E. A. Webb, N. Walworth, and A. Tagliabue. 2013. Taxon-specific response of marine nitrogen fixers to elevated carbon dioxide concentrations. Nat. Geosci. **6:** 790–795, doi:10.1038/ngeo1858

Hynes, A. M. 2009. Diversity of the marine cyanobacterium *Trichodesmium*: Characterization of the Woods Hole culture collection and quantification of field populations. Ph.D. thesis. Massachusetts Institute of Technology and Woods Hole Oceanographic Institution.

———, E. A. Webb, S. C. Doney, and J. B. Waterbury. 2012. Comparison of cultured *Trichodesmium* (Cyanophyceae) with species characterized from the field. J. Phycol. **48:** 196–210, doi:10.1111/j.1529-8817.2011.01096.x

Luo, Y.-W., and others. 2012. Database of diazotrophs in global ocean: Abundances, biomass and nitrogen fixation rates. Earth Syst. Sci. Data **4:** 47–73, doi:10.5194/essd-4-47-2012

McGillicuddy Jr., D. 2014. Do *Trichodesmium* spp. populations in the North Atlantic export most of the nitrogen they fix? Glob. Biogeochem. Cycles **28:** GB0046, doi:10.1002/2013GB004652

Monteiro, F. M., M. J. Follows, and S. Dutkiewicz. 2010. Distribution of diverse nitrogen fixers in the global ocean. Glob. Biogeochem. Cycles **24:** GB3017, doi:10.1029/2009GB003731

Monterey, G., and S. Levitus. 1997. Seasonal variability of mixed layer depth for the world ocean. NOAA Atlas NESDIS 14. Washington, D.C.: U.S. Government Printing Office.

Moore, L., G. Rocap, and S. Chisholm. 1998. Physiology and molecular phylogeny of coexisting *Prochlorococcus* ecotypes. Nature **576:** 220–223.

Muller-Karger, F., C. McClain, and P. Richardson. 1988. The dispersal of the Amazon's water. Nature **333:** 56–59, doi:10.1038/333056a0

Olson, E. 2014. Investigating the role of *Trichodesmium* spp. in the oceanic nitrogen cycle through observations and models. Ph.D. thesis. Massachusetts Institute of Technology and Woods Hole Oceanographic Institution.

Orchard, E., E. Webb, and S. Dyhrman. 2009. Molecular analysis of the phosphorus starvation response in *Trichodesmium* spp. Environ. Microbiol. **11:** 2400–2411, doi:10.1111/j.1462-2920.2009.01968.x

Palenik, B. 2001. Chromatic adaptation in marine *Synechococcus* strains. Appl. Environ. Microbiol. **67:** 991–994, doi:10.1128/AEM.67.2.991-994.2001

Read, B., and others. 2013. Pan genome of the phytoplankton *Emiliania* underpins its global distribution. Nature **499:** 209–213, doi:10.1038/nature12221

Rodríguez, F., E. Derelle, L. Guillou, F. Le Gall, D. Vaulot, and H. Moreau. 2005. Ecotype diversity in the marine picoeukaryote *Ostreococcus* (Chlorophyta, Prasinophyceae). Environ. Microbiol. **7:** 853–859, doi:10.1111/j.1462-2920.2005.00758.x

Sañudo-Wilhelmy, S. A., and others. 2001. Phosphorus limitation of nitrogen fixation by *Trichodesmium* in the central Atlantic Ocean. Nature **411:** 66–69, doi:10.1038/35075041

Subramaniam, A., and others. 2008. Amazon River enhances diazotrophy and carbon sequestration in the tropical North Atlantic Ocean. PNAS **105:** 10460–10465, doi:10.1073/pnas.0710279105

Tovar-Sanchez, A., and others. 2006. Effects of dust deposition and river discharges on trace metal composition of *Trichodesmium* spp. in the tropical and subtropical North Atlantic Ocean. Limnol. Oceanogr. **51:** 1755–1761, doi:10.4319/lo.2006.51.4.1755

Tyrrell, T., E. Marañón, A. J. Poulton, A. R. Bowie, D. S. Harbour, and E. M. S. Woodward. 2003. Large-scale latitudinal distribution of *Trichodesmium* spp. in the Atlantic Ocean. J. Plankton Res. **25:** 405–416, doi:10.1093/plankt/25.4.405

Van Mooy, B., and others. 2012. Quorum sensing control of phosphorus acquisition in *Trichodesmium* consortia. ISME J. **6:** 422–429, doi:10.1038/ismej.2011.115

Webb, E., J. Moffett, and J. Waterbury. 2001. Iron stress in open-ocean cyanobacteria (*Synechococcus, Trichodesmium, and Crocosphaera* spp.): Identification of the IdiA Protein. Appl. Environ. Microbiol. **67:** 5444–5452, doi:10.1128/AEM.67.12.5444-5452.2001

Associate editor: Robert R. Bidigare

Received: 27 February 2014
Accepted: 09 July 2014
Amended: 19 July 2014