Sources of iron and phosphate affect the distribution of diazotrophs in the North Atlantic

Citation: Ratten, Jenni-Marie, Julie LaRoche, Dhwani K. Desai, Rachel U. Shelley, William M. Landing, Ed Boyle, Gregory A. Cutter, and Rebecca J. Langlois. "Sources of Iron and Phosphate Affect the Distribution of Diazotrophs in the North Atlantic." Deep Sea Research Part II: Topical Studies in Oceanography 116 (June 2015): 332–341.

As Published: https://doi.org/10.1016/j.dsr2.2014.11.012

Publisher: Elsevier

Persistent URL: http://hdl.handle.net/1721.1/120778

Version: Original manuscript: author's manuscript prior to formal peer review

Sources of iron and phosphate affect the distribution of diazotrophs in the North Atlantic

Jenni-Marie Ratten[a]

Julie LaRoche[b] (corresponding author)

Dhwani Desai[c]

Rachel U. Shelley[d]

William M. Landing[e]

Ed Boyle[f]

Gregory A. Cutter[g]

Rebecca J. Langlois[h] (corresponding author)

[a] Dalhousie University, Biology Department, Halifax, NS B3H 4R2, CANADA, jenni.ratten@dal.ca

[b] Dalhousie University, Biology Department, 1355 Oxford St, Halifax, NS B3H 4R2, CANADA, Julie.laroche@dal.ca, 01-902-494-4249

[c] Dalhousie University, Biology Department, Halifax, NS B3H 4R2, CANADA, dhwani.desai@dal.ca

[d] Florida State University, Earth, Ocean and Atmospheric Science, Tallahassee, FL 32306, USA, rachel.shelley@univ-brest.fr

[e] Florida State University, Earth, Ocean and Atmospheric Science, Tallahassee, FL 32306, USA, wlanding@fsu.edu

21 [f] Massachusetts Institute of Technology, Department of Earth, Atmospheric, and
22 Planetary Sciences, Cambridge, MA 02139, USA, eaboyle@mit.edu
23 [g] Old Dominion University, Ocean, Earth & Atmospheric Sciences, Norfolk, VA 23529,
24 USA, gcutter@odu.edu
25 [h] Dalhousie University, Oceanography Department, 1355 Oxford St, Halifax, NS B3H
26 4R2, CANADA, rebecca.langlois@dal.ca, 01-902-494-4377
27

Abstract

Dinitrogen fixation supplies nutrient-depleted oceanic surface waters with new biologically-available fixed nitrogen. Diazotrophs are the only organisms that can fix dinitrogen, but the factors controlling their distribution patterns in the ocean are not well understood. In this study, the relative abundances of eight diazotrophic phylotypes in the subtropical North Atlantic Ocean were determined by quantitative PCR (qPCR) of the *nifH* gene using TaqMan probes. A total of 152 samples were collected at 27 stations during two GEOTRACES cruises (USGT10 and USGT11) from Lisbon, Portugal to Mindelo, Cape Verde Islands (USGT10), and from Woods Hole, MA, USA via the Bermuda Time Series (BATS) to Praia, Cape Verde Islands (USGT11). Seven of the eight diazotrophic phylotypes tested were detected. These included free-living and symbiotic cyanobacteria (unicellular groups (UCYN) A, B and C, *Trichodesmium,* the diatom-associated cyanobacteria *Rhizoselinia-Richelia* and *Hemiaulus-Richelia*) and a γ-proteobacteria (Gamma A, AY896371), The *nifH* gene abundances were analyzed in the context of a large set of hydrographic parameters, macronutrient and trace metal concentrations measured in parallel with DNA sampling using the PRIMER-E software to determine the environmental variables that most influenced the abundances and distribution of the diazotrophic phylotypes. We observed a geographic segregation of diazotrophic phylotypes between east and west, with UCYN A, UCYN B and UCYN C and the *Rhizosolenia-Richelia* symbiont associated with the eastern North Atlantic (east of 45°W), whereas *Trichodesmium* and Gamma A were detected across the basin. *Hemiaulus-Richelia* symbionts were primarily found in temperate waters near the North American coast. The highest diazotrophic phylotype abundance and diversity were

associated with temperatures greater than 22 °C in the surface mixed layer and a high supply of iron and phosphorus from North African aeolian mineral dust deposition and from remineralized nutrients upwelled at the edge of the oxygen minimum zone off the northwestern coast of Africa.

Keywords

Diazotrophs

North Atlantic

Iron

Aerosol

nifH

GEOTRACES

qPCR

Abbreviations

Al	Aluminum
Anammox	anaerobic ammonium oxidation
ANOSIM	Analysis of Similarities
BATS	Bermuda Atlantic Time Series
BNF	Biological Nitrogen Fixation

CVOO	Cape Verde Ocean Observatory
Fe	iron
Gamma A	*γ-proteobacteria A*, AY896371
Het 1	*Rhizosolenia-Richelia* symbiont
Het 2	*Hemiaulus-Richelia* symbiont
MAR	mid-Atlantic ridge
OMZ	oxygen minimum zone
PCA	Principle Component Analysis
qPCR	quantitative PCR
SIMPER	Similarity Percentages
SML	surface mixed layer
UCYN A	Unicellular cyanobacteria Group A (related to *Candidatus Atelocyanobacterium thalassa*)
UCYN B	Unicellular cyanobacteria Group B (related to *Crocosphaera*)
UCYN C	Unicellular cyanobacteria Group C (related to *Cyanothece*)

1. Introduction

Dinitrogen fixation is an important source of biologically-available nitrogen in the marine environment, as fixed forms of nitrogen are scarce in most open ocean surface waters. This is particularly true in the oligotrophic subtropical gyres (Vitousek and Howarth

1991, Karl et al. 2002). Dinitrogen fixation is carried out by a specific group of bacteria and Archaea called diazotrophs. Until recently it was believed that the majority of dinitrogen fixation in the ocean was performed by the large, surface bloom-forming *Trichodesmium*, a non-heterocystous, filamentous cyanobacterium, and by symbiotic associations between diatoms and the diazotroph *Richelia* sp. (Foster et al. 2007). Phylogenetic studies using *nifH*, the gene encoding for the iron protein subunit of the nitrogenase enzyme (Zehr et al. 1998, Langlois et al. 2005, Turk et al. 2011) have revealed a much more diverse diazotrophic flora that includes unicellular and symbiotic cyanobacteria, heterotrophic bacteria and Archaea, all potentially contributing significantly to global oceanic dinitrogen fixation. High-throughput next generation sequencing studies have further enriched our knowledge of diazotroph phylogenetic diversity, and have identified the presence of unexplored heterotrophic diazotroph groups throughout the world's oceans (Farnelid et al. 2011).

Although the abundance of the diazotrophs *Trichodesmium* and *Richelia* can be determined by microscopy counts, many of the diazotrophic unicellular cyanobacteria and heterotrophic bacteria cannot be visually identified with certainty in marine microbial communities. Microscopic images of the elusive UCYN A, one of the most widely distributed diazotrophic cyanobacteria, have been obtained only recently (Krupke et al. 2013, Thompson et al. 2013). To date, most oceanic heterotrophic diazotrophs are known only by their *nifH* sequences. To further complicate the matter, the abundance of diazotrophs is generally several orders of magnitude lower than the dominant phytoplankton and bacterioplankton (e.g., *Prochlorococcus* and *Pelagibacter*). This presents a challenge for detection and cultivation techniques. Quantitative-PCR (q-

PCR) and TaqMan probes have been used to circumvent some of these difficulties (Langlois et al. 2008), allowing the quantitative detection of diverse phylogenetic clades defined by specific *nifH* sequences. This approach has already yielded valuable information on the *nifH* phylotype distribution and abundance in the Pacific (Goebel et al. 2007, Church et al. 2008, Moisander et al. 2010) and Atlantic Oceans (Langlois et al. 2008, Turk et al. 2011).

Marine diazotrophs play a critical role in the oceanic nitrogen cycle, because they currently provide the most significant source of fixed nitrogen to the ocean through biological nitrogen fixation (BNF) of dinitrogen gas (Duce et al. 2008). Over geological time scales, the magnitude of the global oceanic fixed nitrogen inventory has been determined by the balance between BNF and the combined nitrogen loss processes of denitrification (Altabet 2006, Codispoti 2006) and anaerobic ammonia oxidation (anammox)

Diazotroph distribution has been utilized to estimate dinitrogen fixation rates and model the factors controlling BNF. However, the oceans remain vastly under sampled with respect to diazotroph abundance, distribution and community structure (Luo et al. 2012, Fernández et al. 2013), making it problematic to validate model-based predictions concerning the fate of dinitrogen fixation in a changing ocean (Goebel et al. 2007, Monteiro et al. 2010, Sohm et al. 2011). It is therefore important to understand what environmental factors control the distribution and activity of marine diazotrophs.

Environmental parameters such as temperature, availability of phosphate, water column stability, upward diffusive fluxes of nutrients, light, and input of iron (Fe) via atmospheric mineral dust deposition (Fernández et al. 2013) have all been proposed as factors

controlling the distribution of diazotrophs. Although detected in almost every oceanic environment, diazotrophs are most abundant in the warm tropical and subtropical oceans where fixed nitrogen is depleted in surface waters (Langlois et al. 2008, Church et al. 2008, Stal 2009, Moisander et al. 2010). Diazotrophs are not limited by fixed nitrogen availability, but both phosphorus and dissolved Fe availability have been implicated in the control of the geographical distribution of diazotrophs and BNF (Falkowski 1997, Karl et al. 2002, Mills et al 2004, Moore et al. 2009). In the oligotrophic subtropical North Atlantic gyre, mineral dust deposition is the most significant source of dissolved Fe to the surface of the ocean (Gao et al. 2001, Jickells et al. 2005). However in the eastern tropical Atlantic, between the Cape Verde Islands and the northwest African coast, upwelled regenerated nutrients from the sub-surface oxygen minimum zone are another potential source of macro- (N, P, Si) and micro-nutrients (e.g. Fe, Co) to the surface layers (Rijkenberg et al. 2012, Fitzsimmons et al. 2013).

We used qPCR and eight phylotype-specific TaqMan probes and primer sets, representing the most commonly occurring marine diazotrophs in the Atlantic Ocean (Langlois, et al. 2008) to estimate *nifH* abundances in an East-West transect across the subtropical North Atlantic Ocean. We compared the distribution and relative abundance of *nifH* phylotypes with hydrographic parameters, macronutrients and trace metal distributions from the surface down to 400 m as well as aerosol aluminum (Al) and Fe concentrations. This was possible through coordinated sampling of nucleic acids, a suite of trace metals dissolved in the water column and aerosols during the 2010 and 2011 US GEOTRACES research cruises.

2. Materials and methods

2.1 Cruise track and sample collection

Samples for measuring *nifH* gene abundance were collected during two GEOTRACES cruises (USGT10 and USGT11) that took place in the subtropical North Atlantic Ocean from October 16th till November 2nd 2010 and from November 7th till December 10th 2011, respectively (figure 1). The cruise track (figure 1) included stations at the Bermuda Atlantic Time-series (BATS) site, Cape Verde Ocean Observatory (CVOO) site and the mid-Atlantic ridge (MAR). Seawater samples were collected from the conventional CTD/rosette at six depths per station ranging from 2-1000 m. Immediately after collection 1-2 L of seawater were vacuum filtered onto 0.22 µm Durapore filters (Millipore) to collect the natural microbial communities. The filters were flash frozen using liquid nitrogen and stored at -80 °C until analysis in the laboratory. In total, 152 samples were collected from 27 stations with an average of 6 depths per station ranging between 2 and 1000 m. Up to three samples were collected in the surface mixed layer (SML) at all of the stations sampled. A broad suite of trace metals and other macronutrients were sampled during these two US GEOTRACES cruises (Deep-Sea Research special issue), enabling the analysis of the nucleic acid-derived *nifH* abundance measurements within the context of a large database of chemical and hydrographic parameters.

2.2 DNA extraction and qPCR

In the laboratory, liquid nitrogen-frozen filters were crushed with plastic homogenizers and incubated for five min with a 5 mg mL^{-1} lysozyme in TE buffer solution. DNA was

183 extracted using the AllPrep RNA/DNA Mini Kit (Qiagen) following the manufacturer's
184 protocol, except that DNA was eluted twice with 40 µl TE buffer and incubated for five
185 minutes before centrifuging. DNA was stored in small aliquots to avoid freeze/thaw
186 cycles. DNA concentrations were determined using the Quanti-iT PicoGreen dsDNA
187 reagent (Molecular Probes, Life Sciences). The abundances of eight *nifH* phylotypes
188 were determined by qPCR using the specific TaqMan probes and primers for Het 1
189 (*Rhizosolenia-Richelia* symbionts; Church et al. 2005) and Het 2 (*Hemiaulus-Richelia*
190 symbiont; Foster et al. 2007), *Trichodesmium*, UCYN A (*Candidatus*
191 *Atelocyanobacterium thalassa*), UCYN B (*Crocosphaera*), UCYN C (*Cyanothece*),
192 Gamma A (*gamma-proteobacteriaA*) and Cluster III (Langlois et al. 2008). Universal
193 TaqMan mastermix and concentrations of primers, probes and BSA were as in Langlois
194 et al. (2008) in a reaction volume of 25 µl, which included either 5 µl of plasmid
195 standard, DNA sample or PCR water as template. Plasmid standards, samples and no-
196 template controls were run in duplicate on the Roche LightCycler 480 using clear 384-
197 well plates. Samples were amplified using the following program: 95 °C for 10 min, 45
198 cycles of [95 °C for 15 sec, 60 °C for 1 min]. Data was collected at 60 °C. A ramp of
199 1.6 °C sec^{-1} was used at each step. Amplification curves were analyzed using LinReg
200 (version 2013.0) (Ramakers et al. 2003). All qPCR reactions amplified with efficiencies
201 greater than 97%. Average primer efficiencies (Langlois et al. 2012) were 97% for
202 *Rhizosolenia-Richelia* symbiont, *Hemiaulus-Richelia* symbiont and UCYN A, 91% for
203 *Trichodesmium*, UCYN B and UCYN C, 92% for Gamma A and 95% for Cluster III. As it
204 is not yet known how the *nifH* copy numbers relate to diazotroph biomass or cell
205 density, the qPCR results are reported throughout the manuscript as *nifH* copies mL^{-1}

and represent the number of *nifH* copies detected in environmental DNA samples in a known volume of seawater. All phylotypes except Cluster III were detected. Hence Cluster III is not included in the analysis.

2.3 PRIMER-E analysis

2.3.1 Preparing the matrices

The *nifH* gene abundances of the seven detected diazotrophic phylotypes (*Rhizosolenia-Richelia* symbiont, *Hemiaulus-Richelia* symbiont, *Trichodesmium*, UCYN A, UCYN B, UCYN C and Gamma A) and their corresponding environmental variables were analyzed in all 152 samples or a subset of the surface mixed layer samples (SML) only (52 samples) using the PRIMER-E software V.6 (Clarke and Gorley, 2006). Environmental metadata was obtained from BCO-DMO (Biological and Chemical Oceanography Data Management Office; links to all data sets are listed in Supplemental Table 2). The dataset was first divided into a community data matrix (containing gene abundances of the *nifH* phylotypes) and a corresponding matrix of environmental measurements (including dissolved metal concentrations, nutrients, organic material and physical parameters).

Missing values in one or more environmental variables resulted in the deletion of the entire sample from the data base and hence, such samples were not included in the subsequent statistical analysis. A correlation matrix (draftsmans plot) was generated for each pair of environmental variables. Only one variable was retained from pairs with a correlation > 0.9. This resulted in a final dataset comprised of 64 samples divided into a

subset of samples collected within the SML (37 samples) and a subset of samples originating from below the SML.

The environmental variables used for the principal component analysis were expressed on broadly different scales precluding a direct comparison without biasing of the results. In order to derive meaningful distances between samples using Euclidian distances, we first square-root transformed or log transformed the variables that covered several orders of magnitude to bring them within a common numerical range. This generated values all ranging within 4 orders of magnitude, allowing variables to be compared without biases. Each variable of the environmental matrix was then normalized by subtracting their mean and dividing by their standard deviation prior to further analysis. The *nifH* phylotypes abundance data was log-transformed and compared using Bray-Curtis Similarities, a similarity (or distance) measure that ignores joint absences of variables between samples.

2.3.2 Multivariate analysis pipeline

For all phylotype matrices a BEST (Bio-Env + Stepwise) test was carried out. This test determines which environmental variables best explain the microbial community composition. The comparison was carried out between the transformed environmental matrix and the Bray-Curtis similarities of the phylotype data set. A combination of variables which showed the maximum correlation with the phylotype distribution was identified for further analysis (Supplemental Table 1), using LINKTREE (linkage tree), a program that describes the best way to split the samples into groups based on a

250 threshold value for each environmental variable (for example group1 > 0.4 µM PO_4^{3-} >
251 group 2).

252 Principle Component Analysis (PCA) was performed with the environmental matrix of
253 the SML samples. The first three components of the PCA captured 85.2% (PC1 =
254 23.4%, PC2 = 51.3%, PC3 = 10.5%) of the variance.

255 Based on the clustering obtained in the PCA plots and on the results of the
256 BEST/LINKTREE test and utilizing information about known drivers of diversity, the
257 samples were categorized into groups. These categories included geographical location
258 (east or west of 45°W, north or south of 30°N), nutrient concentration (high, low), trace
259 metal concentrations in the water column and in aerosols (high, low), dust origin
260 (European, North African, Saharan, Marine, North American) and rain (present, absent).
261 For variables with continuous data, high and low concentrations were defined by a
262 threshold derived from the published peer-reviewed literature. If no definition could be
263 found in the literature, the variables were categorized based on an evaluation of the
264 present dataset (LINKTREE analysis) and literature values (Table 1).

265 An ANOSIM (analysis of similarity) test was utilized to compare the diazotrophic
266 communities within these predefined groups and to determine whether the distribution
267 of the *nifH* phylotypes were significantly different between the predefined groups for
268 each environmental variable (Table 1). The Bray-Curtis similarities of the log-
269 transformed diazotroph community data (based on *nifH* counts of the various phylotypes
270 at each site) was used for the ANOSIM.

271 For each categorization that showed a positive ANOSIM test ($p < 0.05$), the
272 discriminating phylotypes in the groups of this categorization (e.g. high and low aerosol

load) were identified using the SIMPER (Similarity Percentages) routine (Table 2; the three most contributing phylotypes are highlighted in bold and italics)

3.0 Results

3.1 Distribution of seven diazotrophic phylotypes during the US Atlantic GEOTRACES cruises

Although *nifH* phylotypes were detected throughout the water column, they were most abundant in the surface mixed layer (SML; Figure 2). Along the east-west transect, the sum of all *nifH* phylotypes (*nifH* copies mL^{-1}) in the SML was highest close to the CVOO (Cape Verde Ocean Observatory) on the eastern side of the transect (> 100 *nifH* copies mL^{-1}), dropping off around the MAR and rising again on the western side of the basin to > 100 *nifH* copies mL^{-1} (Figure 3a). The Shannon diversity index (a measure of abundance and evenness of species present) showed a similar trend, with high diversity of *nifH* phylotypes in the eastern basin and lower diversity observed in the center of the gyre (Figure 3b). Although the diversity and abundance of *nifH* phylotypes varied across the Atlantic basin, temperature and N* (N* = 0.87 (N - 16P + 2.9 µM); Gruber and Sarmiento 1997) remained relatively constant within the gyre, at 25.2 °C and 0, dropping at the continental shelf edge to 19 °C and -0.8, respectively (Figure 3 c and d). Major nutrient concentrations (N, P, Si) concentrations were low in the SML and relatively constant throughout the transect; except for a tendency towards depletion of nitrate relative to phosphate in the western Atlantic and at CVOO. Phosphate concentrations were low across the gyre, averaging 7.5 nM. However, phosphate

concentrations rose to 16 – 74 nM near the American coast and to 21 nM east of CVOO (Figure 3e).

Figure 4 shows the *nifH* gene copy numbers (copies mL^{-1}) for the seven detected *nifH* phylotypes. The most commonly detected phylotypes were *Trichodesmium* and UCYN A, reaching maximum abundances of 391 and 105 *nifH* copies mL^{-1} at individual stations, respectively (Figure 4 a and b). *Trichodesmium* was the most frequently detected phylotype throughout the transect with abundances greater than 150 *nifH* copies mL^{-1} detected at six stations (Figure 1) and contributed the most to the overall abundance of the sum of the *nifH* genes (Figure 3a, Figure 4a). *Rhizosolenia-Richelia* symbiont and UCYN C were the least abundant phylotypes with maximum abundances of 4 and 6 copies mL^{-1}, respectively (Figure 4d). Although far less abundant than *Trichodesmium*, *Rhizosolenia-Richelia* symbiont and Gamma A phylotype distributions paralleled the variation in the *Trichodesmium* phylotype distribution (Figure 4c, d). The unicellular cyanobacterial phylotypes UCYN A, B, and C were most abundant on the eastern side of the transect in the region between 30°W – 25°W, directly west of CVOO (Figure 4 c-d and supplemental figure 2f-h), reaching significant *nifH* abundances below the SML also (SMLD 62-85 m). In contrast, the *Hemiaulus- Richelia* symbiont phylotype was mainly found in colder waters (19 – 25 °C) near the American coast (Figure 4c).

We investigated the correlation between mineral aerosol concentrations (and therefore implied deposition) and *nifH* abundance in surface waters (Figure 4e). The high aerosol concentrations of Al and Fe observed on the eastern side of the transect (both elements exceeding 1000 ng m^{-3} at most sample points with maxima of 7620 and 5760 ng m^{-3} respectively at CVOO) coincided with the highest abundance of *nifH* copies mL^{-1}

(Figures 3a and 4e). Even though it has been shown that the North African dust samples have low fractional Fe solubility compared to aerosols originating from North America, the very high amount of North African dust that is transported to the eastern North Atlantic implies that the flux of soluble aerosol Fe would be higher at CVOO (Shelley et al. 2014). A rapid decrease in aerosol Al and Fe loadings coincided with the decrease in diversity and *nifH* abundance at 40°W.

3.2 Multivariate Statistical Analysis

The BEST test applied to the entire dataset provided an initial identification of the environmental variables most relevant in determining the observed diazotrophic community composition. In total 26 environmental variables were tested including dissolved inorganic nutrients, and hydrographic parameters (Supplemental Table 2). The results of this analysis demonstrated the importance of the SML as a determining factor for the community composition (p = 0.01; Supplemental Table 1). We therefore carried out the BEST analysis on two subsets composed of samples present above and below the SML referred to as SML or deep samples, respectively. The SML subset was significantly correlated with environmental variables (p = 0.01; Supplemental Table 1). In contrast the deep sample subset showed no significant correlation, likely due to the very low *nifH* phylotype abundances measured in most deep samples (p = 0.31; Figure 2), and was not analyzed further.

For the SML temperature, phosphate and other environmental variables, identified from the BEST analysis, were used in a PCA (Figure 5). None of the available dissolved trace metal data that were measured during the cruise showed significant influence on

the distribution of the *nifH* phylotypes. This included the biologically-relevant cofactors cobalt, vanadium, copper and zinc.

The PCA conducted with the environmental variables identified with BEST resulted in three significantly different (ANOSIM, $p < 0.01$) clusters of samples. The most important determining factors were east-west segregation, mineral dust concentration and nutrient upwelling of P and Fe.

The two largest clusters were composed of samples collected east and west of 40°W both characterized by high water temperatures and low macronutrient concentrations (Figure 5 a). The eastern cluster was dominated by high aerosol Fe concentrations of African origin. The large western cluster was dominated by aerosols originating from marine and North American sources containing lower Fe concentrations (Figure 5 b and c).

Two overlapping western clusters composed of samples collected near the North American coast, in waters with high nutrients and low temperature were not significantly different from each other, as determined by hierarchical clustering analysis (results not shown). The samples from the North American coast were subjected to low aeolian Fe originating from North America rather than North Africa. The original PCA plot was overlaid with *nifH* phylotype abundances (supplemental figure 1). Some of the phylotypes were distributed evenly across the clusters (*Trichodesmium* and *Rhizosolenia-Richelia* symbionts), whereas the high abundances of UCYN A, UCYN B, UCYN C and Gamma A *nifH* copies coincided with the North African high aerosol cluster in the east of the basin; *Hemiaulus-Richelia nifH* copies were associated with the

higher macro-nutrient and lower temperature conditions near the North American coast (Supplemental Figure 1).

To further analyze whether these differential patterns of phylotype distribution had a statistical significance, we performed ANOSIM tests on the community matrix (i.e. the *nifH* phylotype abundances). The samples were divided into groups or categories as explained in the methods. Categorizations which showed a positive ANOSIM test ($p < 0.05$) were temperature longitude and latitude, PO_4^{3-} (which correlated with NO_3^- and SiO_2), aerosol loadings and aerosol origin (Table 1). Dissolved Al contributed to clustering in the PCA, but was not significant in the ANOSIM. Dissolved Fe concentrations were not significant even though Fe has been previously identified as a driving factor for diazotrophic distribution (Table 1).

The phylotypes which contributed to significant differences between the groups were identified with the SIMPER routine and are listed in Table 2. *Trichodesmium* and Gamma A, which were both distributed throughout the transect, were not generally discriminating taxa, except in the case of the aerosol origin categories where higher *Trichodesmium* abundances were found in association with conditions where the air-mass back trajectories indicated that the aerosols did not have an obvious continental source (Shelley et al. 2014).

4.0 Discussion

4.1 The diazotroph community structure along an East-West transect in the North Atlantic Ocean

It is now well established that the marine diazotrophic community is much more diverse than previously thought (Zehr et al. 1998). However, data on the large-scale distribution and structure of the diazotrophic communities across oceanic basins are sparse, as easily seen from the compilation of the available observations of diazotrophs and their distribution on global maps of the world's oceans (Luo et al. 2012). Vast regions of the oceans remain undersampled spatially and temporally with respect to diazotrophic phylotypes and abundances. In particular, the lack of observations is most noticeable in colder waters outside of the tropical oceans (Luo et al. 2012), probably because until recently marine BNF has been mainly investigated in warm waters where large blooms of *Trichodesmium* are easily noticed (Capone et al. 1997). In more recent years, *nifH*-based phylogenetic studies have firmly established the widespread distribution of diazotrophic microorganisms other than *Trichodesmium*, extending the distribution range of diazotrophs to the global oceans (Farnelid et al. 2011) and in particular to more temperate oceanic regions (Needoba et al. 2007, Blais et al. 2012).

Our study has focused on the detection of seven *nifH* phylotypes that have been previously identified as dominant in the SML (surface mixed layer) of the North Atlantic. The east-west transect we present here spans from 17°N to 40°N, latitudes that are under sampled with respect to diazotrophs. Previous east-west transects crossed the Atlantic ocean at latitudes of 10°N and 0 – 20°N, omitting the North Atlantic gyre (Langlois et al. 2008, Goebel et al. 2010). Building on the work of Langlois et al. (2008)

and Goebel et al. (2010) who presented results on some, but not all, of the same phylotypes discussed in our study, general similarities can be drawn between distribution patterns observed for these three transects. All show differential distributions of diazotrophs with UCYN A predominantly detected east of 40°W and *Hemiaulus-Richelia* symbiont mainly in the western Atlantic. Similarly to Langlois et al. (2008), the UCYN B and UCYN C phylotypes were also detected at higher abundances east of 40°W and the distributions of Gamma A and *Trichodesmium* were weakly correlated (R^2 = 0.39, Figure 4a,c). Further south, *Trichodesmium* can also be abundant in the western Atlantic at the boundary between oligotrophic waters and Amazon River outflow (Subramanian et al. 2008). Although not found in our study, Langlois et al. (2008) detected a Cluster III *nifH* phylotypes in the North Atlantic at higher latitudes than those sampled in our study.

4.2. Analysis of the community structure using the multivariate statistics

4.2.1 East-West segregation

Our statistical analyses confirmed the observation that the diazotrophic phylotypes detected in our study inhabit primarily the SML (Figure 2) and hence we carried out the statistical analysis on the SML samples only. The geographically segregated east and west diazotrophic communities (confirmed by ANOSIM; Table 1), were dominated by different phylotypes. The SIMPER analysis indicated that the *nifH* copys mL^{-1} of significant phylotypes contributing to the different community structure were unicellular cyanobacteria (UCYN A and UCYN B) and slightly higher abundances of *Trichodesmium* and Gamma A phylotypes east of 40°W, while the *Hemiaulus-Richelia*

symbiont was the discriminant phylotype west of 40°W (Table 2). The diatom *Hemiaulus-Richelia* symbiont dominance, near the North American as estimated from the nifH copies mL^{-1} coast was associated with much lower water temperature but higher nutrient concentrations, indicating a preference for these environmental conditions. However, the *Hemiaulus-Richelia* association has often been detected in the warmer waters north of Brazil and in the Caribbean (Carpenter et al. 1999, Foster et al. 2007, Subramanian et al. 2008, Goebel et al. 2010). The geographical distribution of the *Hemiaulus-Richelia* along the American Eastern seaboard could also result from the transport of the diatom and its symbiont from the tropical Atlantic by eddy formation and spin-off from the Gulf Stream (Lee, Yoder and Atkinson 1991).

An exception to the eastern weighted distribution pattern of *Trichodesmium* was observed at BATS, corresponding with the recent passage of tropical storm Sean on Nov 11th, 2011 (Figure 1) at that station (Nov 19th 2011) which caused high winds and rainfall (gusts of 91 kph and 0.05 – 0.12" rain at BATS) followed by days with wind speeds of 10 – 30 kph; the resultant water column mixing and higher availability of nutrients may have contributed to higher *Trichodesmium nifH* gene copy numbers at that location.

4.2.2 Influence of aerosol concentrations on diazotroph distribution

High and low aerosol Fe concentration and aerosol origin superimposed onto the SML PCA plot (Figure 5) provided information on the preference of high aerosol supply for specific *nifH* phylotypes present in different clusters. We assume here that measurements of high aerosol load equate to proportionally higher aerosol flux to

surface waters. Eastern samples received high aerosol loads of North African/Saharan origin whereas western samples received low aerosol loading of either North American or Marine origin. As confirmed by ANOSIM (Table 1) and SIMPER (Table 2), and with the exception of *Hemiaulus-Richelia,* aerosol loading significantly correlated with diazotroph distribution in the SML with all *nifH* phylotypes being significantly associated with high aerosol loading. In addition UCYN A, UCYN B, UCYN C, *Rhizosolenia-Richelia* and *Trichodesmium* were significantly associated with a North African/Saharan dust origin (Figure 5 and Supplemental Figure 1). Episodic dust storms deposit 10 – 50 g of dust m^{-2} to the eastern North Atlantic annually (Lawrence and Neff 2009) and hence supply this area with a variety of nutrients. Monthly averaged remote sensing data from MODIS over the time period of the research cruise (supplemental figure 3a and b) show high optical depth at 550 nm at the same stations that were recorded to have high aerosol load by Shelley et al. (2014). Optical depth decreased to < 0.1 at 32°W, coinciding with the observed drop in diazotrophic diversity and abundance (figure 3b). Precipitation data from TRMM (Tropical Rainfall Measuring Mission; http://trmm.gsfc.nasa.gov/) showed that the ocean west of 40°W received high precipitation, again creating different environmental conditions in the North Atlantic east and west of 40°W during our study. The dust deposited on the ocean surface is a composite of many trace elements (Jickells 1999, Goudie and Middleton 2001, Viana et al. 2002, Baker et al. 2006, Buck et al., 2010; Shelley at al.2014), macronutrients (Donaghay et al. 1991, Guerzoni et al. 1999, Duarte et al. 2006, Duce et al. 2008) and organic material (Mahowald et al. 2008; Wozniak et al., 2013). A combination of high dust deposition, high temperatures and nitrate-limited surface waters near CVOO

(Figure 3d), provided conditions favorable for diazotroph growth in the SML. These results support previous findings (Mills et al. 2004, Moore et al. 2009, Rubin et al. 2011, Langlois et al. 2012). However, our multivariate analyses including the suite of measurements from the GEOTRACES data base, provides further statistical evidence that the high abundance and diversity of diazotrophs in the eastern subtropical North Atlantic is linked to areas where surface waters receive high mineral dust deposition.

4.2 Presence of diazotroph *nifH* phylotypes below the Surface Mixed Layer

High *nifH* phylotypes abundance below the SML occurred primarily at the water mass boundaries between the sub-tropical gyre and oxygen depleted, high nutrient waters from the northwest African upwelling (Supplemental Figure 2), which occurs mid-spring till mid-autumn at 15°N (Marcello at al. 2009) This oxygen minimum zone reaches as far west as the Cape Verde Islands (minimum 40 µM at 400 m depth; Stramma et al. 2008), and supplies intermediate waters with macro- and micro-nutrients (Rijkenberg et al. 2012, Fitzsimmons et al. 2013). The high dissolved Fe concentrations (> 1.5 nM) and slight excess of PO_4 (shown as N*; Figure 3d and Supplemental Figure 2i and j; Zimmer and Cutter 2012; Wurl et al. 2013;) that were detected below the SML on the eastern side of the transect provide optimal growth conditions for diazotrophs found at the boundary of the gradient between the water masses. This indicates again that diazotrophs thrive in waters that are enriched in dissolved phosphate relative to fixed nitrogen. In our study, nutrients favoring diazotrophic growth were supplied to the SML either from above, through atmospheric input or from below the SML through upwelling of nutrient-rich waters with a N* ratio indicating an excess of PO_4 over fixed N.

4.3 Trace metals in the water column

Within the SML, correlation of *nifH* abundances with dissolved trace metals (e.g. Al, Mn, Ga, Ba, Pb) appeared to be related more to ocean circulation than to biological requirements. Some trace metal concentrations (e.g. Al, Ga, As) were higher in the Sargasso Sea than in the eastern basin close to the Saharan dust source, likely due to longer residence times in the oligotrophic gyre. Low nutrient concentrations in the Sargasso Sea result in lower productivity and hence leads to reduced scavenging and uptake rates (Dammshäuser et al. 2011). Dissolved Fe concentrations were patchy and generally high (in the SML 28 out of 37 stations had concentrations above 0.2 nM; Supplemental Figure 2j), which may have contributed to the finding that dissolved Fe did not significantly influence diazotrophic distribution. In contrast aerosol Al and Fe showed significant positive correlations with high diazotrophic abundances in the SML. Dissolved Al has historically been used as a tracer for dust deposition because it is a major component of mineral dust and was thought to be biologically inactive (Grand et al, 2014). However, recent studies suggests that variable dissolution of Al from wet and dry dust deposition as well as increased scavenging of Al in more productive ocean regions (Dammshäuser et al. 2011) will affect the usefulness of Al as a tracer for atmospheric Fe sources. Thus, dissolved Al concentrations in surface waters may only be an accurate representation of dust deposition under specific conditions. . Thus, dissolved Al concentrations in surface waters may only be an accurate representation of dust deposition under specific conditions. In our study, where both dissolved and atmospheric data were available, Fe and Al concentrations in aerosols were a better

predictor of *nifH* phylotype abundances than dissolved Al concentrations in surface waters.

5.0 Conclusions

Basin wide *nifH* phylotype measurements from samples collected during US GEOTRACES cruises were used as a proxy to assess the large scale distribution patterns of several abundant marine diazotrophs found in the Atlantic Ocean. The west east transect spanned the North Atlantic from 10°W to 70°W and 20°N to 40°N, from the surface down to 800 m. The distribution patterns of the *nifH* phylotypes showed that the communities on the eastern and western side of the Atlantic were significantly different. The western Atlantic diazotrophic community was characterized by the presence of *Hemiaulus-Richelia* association. In contrast, the eastern Atlantic diazotrophic community was dominated by the unicellular cyanobacteria groups (UCYN A, B and C), *Trichodesmium* and Gamma A. The eastern Atlantic community was associated withtemperatures >22 °C in regions of high North African dust deposition, confirming other studies that have previously suggested the importance of aeolian dust deposition to the tropical eastern Atlantic ecosystem. Diazotroph abundance below the SML were associated with water masses with higher concentrations of remineralized nutrients, slightly enriched in PO_4 from either the OMZ near the African Coast or the Gulf Stream on the western side of the Atlantic. Associations with other biologically relevant trace metals could not be conclusively demonstrated and dissolved Al concentrations could not be shown to predict the occurrence of *nifH* phylotypes.

Acknowledgements

We thank Robert Anderson, head of the US GEOTRACES program, for providing access to samples from the cruises, Chris Measures and Mariko Hatta for providing dissolved Al and Fe concentrations and the calculation of the SML depths and Douglas Wallace for discussion. We also thank Melanie Hammer for helping with the remote sensing data and Jennifer Tolman for helping with the data analysis and commenting on the manuscript. Funding for the nutrient and aerosol measurements was provided by the US National Science Foundation (grant OCE-0926092 to G. Cutter; grant OCE-0752351, 0929919 and 1132766 to W. M. Landing), and also *nifH* analysis by an NSERC Discovery Grant to Julie LaRoche. Jenni-Marie Ratten received financial support from the DAAD (Deutscher Akademischer Austausch Dienst), the SKS (Stiftung für Kanadastudien) and received a PhD stipend from the Transatlantic Ocean System Science and Technology (TOSST). Rebecca Langlois was supported by the CERC Grant to D.W.R Wallace.

554 **References**

555 Altabet, M. (2006). Constraints on oceanic N balance/imbalance from sedimentary ^{15}N
556 records. *Biogeosciences Discussions, 3*(4), 1121-1155.

557 Baker, A., Jickells, T., Witt, M., Linge, K. (2006). Trends in the solubility of iron,
558 aluminium, manganese and phosphorus in aerosol collected over the Atlantic
559 Ocean. *Marine Chemistry, 98*(1), 43-58.

560 Blais, M., Tremblay, J., Jungblut, A. D., Gagnon, J., Martin, J., Thaler, M., Lovejoy, C.
561 (2012). Nitrogen fixation and identification of potential diazotrophs in the Canadian
562 Arctic. *Global Biogeochemical Cycles, 26*(3), GB3022.

563 Bonnet, S., Guieu, C., Chiaverini, J., Ras, J., Stock, A. (2005). Effect of atmospheric
564 nutrients on the autotrophic communities in a low nutrient, low chlorophyll system.
565 *Limnology and Oceanography, 50*(6), 1810-1819.

566 Breitbarth, E., Oschlies, A., LaRoche, J. (2007). Physiological constraints on the global
567 distribution of *Trichodesmium*–effect of temperature on diazotrophy.
568 *Biogeosciences, 4*(1), 53-61.

569 Buck, C. S., Landing, W. M., Resing, J. A., Measures, C. I. (2010). The solubility and
570 deposition of aerosol Fe and other trace elements in the North Atlantic Ocean:
571 Observations from the A16N CLIVAR/CO_2 repeat hydrography section. *Marine*
572 *Chemistry, 120*(1), 57-70.

573 Capone, D. G., Zehr, J. P., Paerl, H. W., Bergman, B., Carpenter, E. J. (1997).
574 *Trichodesmium*, a globally significant marine cyanobacterium. *Science*, 276 (5316),
575 1221-1229.

576 Capone, D. G., Burns, J. A., Montoya, J. P., Subramaniam, A., Mahaffey, C.,
577 Gunderson, T., Carpenter, E. J. (2005). Nitrogen fixation by *Trichodesmium spp*.:
578 An important source of new nitrogen to the tropical and subtropical North Atlantic
579 Ocean. *Global Biogeochemical Cycles, 19*(2), GB2024.

580 Carpenter, E. J., Montoya, J. P., Burns, J., Mulholland, M. R., Subramaniam, A.,
581 Capone, D. G. (1999). Extensive bloom of a N_2-fixing diatom/cyanobacterial
582 association in the tropical Atlantic Ocean. *Marine Ecology Progress Series, 185*,
583 273-283.

584 Church, M. J., Bjorkman, K. M., Karl, D. M., Saito, M. A., Zehr, J. P. (2008). Regional
585 distributions of nitrogen-fixing bacteria in the Pacific Ocean. *Limnology and*
586 *Oceanography, 53*(1), 63.

587 Church, M. J., Jenkins, B. D., Karl, D. M., Zehr, J. P. (2005). Vertical distributions of
588 nitrogen-fixing phylotypes at stn ALOHA in the oligotrophic North Pacific Ocean.
589 *Aquatic Microbial Ecology, 38*(1), 3-14.

590 Clarke K. R., Gorley R. N. (2006). PRIMER v6: User manual/tutorial. *PRIMER-E,*
591 *Plymouth.*

592 Codispoti, L. (2006). An oceanic fixed nitrogen sink exceeding 400 tg N a^{-1} vs the
593 concept of homeostasis in the fixed-nitrogen inventory. *Biogeosciences*
594 *Discussions, 3*(4), 1203-1246.

595 Dammshäuser, A., Wagener, T., Croot, P. L. (2011). Surface water dissolved aluminum
596 and titanium: Tracers for specific time scales of dust deposition to the Atlantic.
597 *Geophysical Research Letters, 38*(24) L24601.

598 Donaghay, P. L., Liss, P. S., Duce, R. A., Kester, D. R., Hanson, A. K., Villareal, T.,
599 Gifford, D. J. (1991). The role of episodic atmospheric nutrient inputs in the
600 chemical and biological dynamics of oceanic ecosystems. *Oceanography, 4*(2), 62-
601 70.

602 Duarte, C. M., Dachs, J., Llabrés, M., Alonso-Laita, P., Gasol, J. M., Tovar-Sánchez, A.,
603 Agustí, S. (2006). Aerosol inputs enhance new production in the subtropical
604 northeast atlantic. *Journal of Geophysical Research, 111*(G4), G04006.

605 Duce, R. A., LaRoche, J., Altieri, K., Arrigo, K. R., Baker, A. R., Capone, D. G., Zamora,
606 L. (2008). Impacts of atmospheric anthropogenic nitrogen on the open ocean.
607 *Science, 320*(5878), 893-897.

608 Falkowski, P. G. (1997). Evolution of the nitrogen cycle and its influence on the
609 biological sequestration of CO_2 in the ocean. *Nature, 387*(6630), 272-275.

610 Farnelid, H., Andersson, A. F., Bertilsson, S., Al-Soud, W. A., Hansen, L. H., Sørensen,
611 S., Riemann, L. (2011). Nitrogenase gene amplicons from global marine surface
612 waters are dominated by genes of non-cyanobacteria. *PloS One, 6*(4), e19223.

613 Fernández, A., Graña, R., Mouriño-Carballido, B., Bode, A., Varela, M., Domínguez-
614 Yanes, J. F., Marañón, E. (2013). Community N_2 fixation and *Trichodesmium spp*.
615 abundance along longitudinal gradients in the eastern subtropical North Atlantic.
616 *ICES Journal of Marine Science: Journal Du Conseil, 70*(1), 223-231.

617 Fitzsimmons, J. N., Zhang, R., Boyle, E. A. (2013). Dissolved iron in the tropical north
618 atlantic ocean. *Marine Chemistry, 154*, 87-99.

619 Foster, R., Subramaniam, A., Mahaffey, C., Carpenter, E., Capone, D., Zehr, J. (2007).
620 Influence of the Amazon River plume on distributions of free-living and symbiotic
621 cyanobacteria in the western tropical North Atlantic Ocean. *Limnology and*
622 *Oceanography, 52*(2), 517-532.

623 Franchy, G., Ojeda, A., López-Cancio, J., Hernández-León, S. (2013). Plankton
624 community response to Saharan dust fertilization in subtropical waters off the
625 Canary Islands. *Biogeosciences Discussions, 10*(11), 17275-17307.

Gao, Y., Kaufman, Y., Tanre, D., Kolber, D., Falkowski, P. (2001). Seasonal distributions of aeolian iron fluxes to the global ocean. *Geophysical Research Letters, 28*(1), 29-32.

Goebel, N. L., Edwards, C. A., Church, M. J., Zehr, J. P. (2007). Modeled contributions of three types of diazotrophs to nitrogen fixation at station ALOHA. *The ISME Journal, 1*(7), 606-619.

Goebel, N. L., Turk, K. A., Achilles, K. M., Paerl, R., Hewson, I., Morrison, A. E., Montoya, J. P., Edwards, C. A., Zehr, J. P. (2010). Abundance and distribution of major groups of diazotrophic cyanobacteria and their potential contribution to N2 fixation in the tropical Atlantic Ocean. *Environmental Microbiology, 12*(12), 3272-3272.

Goudie, A., Middleton, N. (2001). Saharan dust storms: Nature and consequences. *Earth-Science Reviews, 56*(1), 179-204.

Grand, M.M., Buck, C. S., Landing W.M., Measures, C.I., Hatta, M., Hiscock, W.T., Brown, M., Resing. J.A. (2014). Quantifying the impact of atmospheric deposition on the biogeochemistry of Fe and Al in the upper ocean: A decade of collaboration with the US CLIVAR-CO2 Repeat Hydrography Program. *Oceanography 27*(1):62–65.

Gruber, N., Sarmiento, J. L. (1997). Global patterns of marine nitrogen fixation and denitrification. *Global Biogeochemical Cycles, 11*(2), 235-266.

Guerzoni, S., Chester, R., Dulac, F., Herut, B., Loÿe-Pilot, M., Migon, C., Saydam, C. (1999). The role of atmospheric deposition in the biogeochemistry of the Mediterranean Sea. *Progress in Oceanography, 44*(1), 147-190.

Hans Wedepohl, K. (1995). The composition of the continental crust. *Geochimica Et Cosmochimica Acta, 59*(7), 1217-1232.

Herut, B., Zohary, T., Krom, M., Mantoura, R. F. C., Pitta, P., Psarra, S., Frede Thingstad, T. (2005). Response of east Mediterranean surface water to Saharan dust: On-board microcosm experiment and field observations. *Deep Sea Research Part II: Topical Studies in Oceanography, 52*(22), 3024-3040.

Jickells, T. (1999). The inputs of dust derived elements to the Sargasso Sea; a synthesis. *Marine Chemistry, 68*(1), 5-14.

Jickells, T., An, Z., Andersen, K., Baker, A., Bergametti, G., Brooks, N., Torres, R. (2005). Global iron connections between desert dust, ocean biogeochemistry, and climate. *Science, 308*(5718), 67-71.

Karl, D., Michaels, A., Bergman, B., Capone, D., Carpenter, E., Letelier, R., Stal, L. (2002). Dinitrogen fixation in the world's oceans. *Springer*.

662 Krupke, A., Musat, N., LaRoche, J., Mohr, W., Fuchs, B. M., Amann, R. I., Foster, R. A.
663 (2013). In situ identification and N_2 and C fixation rates of uncultivated
664 cyanobacteria populations. *Systematic and Applied Microbiology,* 36(4), 259–271.

665 Langlois, R. J., Hümmer, D., LaRoche, J. (2008). Abundances and distributions of the
666 dominant *nifH* phylotypes in the northern Atlantic Ocean. *Applied and*
667 *Environmental Microbiology, 74*(6), 1922-1931.

668 Langlois, R. J., LaRoche, J., Raab, P. A. (2005). Diazotrophic diversity and distribution
669 in the tropical and subtropical Atlantic Ocean. *Applied and Environmental*
670 *Microbiology, 71*(12), 7910-7919.

671 Langlois, R. J., Mills, M. M., Ridame, C., Croot, P., LaRoche, J. (2012). Diazotrophic
672 bacteria respond to Saharan dust additions. *Marine Ecology Progress Series, 470*,
673 1-14.

674 Lawrence, C. R., Neff, J. C. (2009). The contemporary physical and chemical flux of
675 aeolian dust: A synthesis of direct measurements of dust deposition. *Chemical*
676 *Geology, 267*(1), 46-63.

677 Lee, T. N., Yoder, J. A., Atkinson, L. P. (1991). Gulf stream frontal eddy influence on
678 productivity of the southeast US continental shelf. *Journal of Geophysical*
679 *Research: Oceans (1978–2012), 96*(C12), 22191-22205.

680 Luo, Y., Doney, S., Anderson, L., Benavides, M., Bode, A., Bonnet, S., Carpenter, E.
681 (2012). Database of diazotrophs in global ocean: Abundances, biomass and
682 nitrogen fixation rates. *Earth System Science Data Discussions, 5*(1), 47-106.

683 Mahowald, N., Jickells, T. D., Baker, A. R., Artaxo, P., Benitez-Nelson, C. R.,
684 Bergametti, G., Herut, B. (2008). Global distribution of atmospheric phosphorus
685 sources, concentrations and deposition rates, and anthropogenic impacts. *Global*
686 *Biogeochemical Cycles, 22*(4), GB4026.

687 Maranón, E., Fernández, A., Mourino-Carballido, B., Martínez-García, S., Teira, E.,
688 Cermeno, P., Calvo-Díaz, A. (2010). Degree of oligotrophy controls the response of
689 microbial plankton to Saharan dust. *Limnology and Oceanography, 55*(6), 2339-
690 2352.

691 Marcello, J., Hernandez-Guerra, A., Eugenio, F., Fonte, A. (2011). Seasonal and
692 temporal study of the northwest African upwelling system. *International Journal of*
693 *Remote Sensing, 32*(7), 1843-1859.

694 Mills, M. M., Ridame, C., Davey, M., La Roche, J., Geider, R. J. (2004). Iron and
695 phosphorus co-limit nitrogen fixation in the eastern tropical North Atlantic. *Nature,*
696 *429*(6989), 292-294.

Moisander, P. H., Beinart, R. A., Hewson, I., White, A. E., Johnson, K. S., Carlson, C. A., Zehr, J. P. (2010). Unicellular cyanobacterial distributions broaden the oceanic N_2 fixation domain. *Science, 327*(5972), 1512-1514.

Monteiro, F., Follows, M., Dutkiewicz, S. (2010). Distribution of diverse nitrogen fixers in the global ocean. *Global Biogeochemical Cycles, 24*(3), GB4026.

Montoya, J., Voss, M., Capone, D. (2007). Spatial variation in N_2-fixation rate and diazotroph activity in the tropical Atlantic. *Biogeosciences, 4*(3), 369-376.

Moore, C. M., Mills, M. M., Achterberg, E. P., Geider, R. J., LaRoche, J., Lucas, M. I., Rijkenberg, M. J. (2009). Large-scale distribution of Atlantic nitrogen fixation controlled by iron availability. *Nature Geoscience, 2*(12), 867-871.

Needoba, J. A., Foster, R. A., Sakamoto, C., Zehr, J. P., Johnson, K. S. (2007). Nitrogen fixation by unicellular diazotrophic cyanobacteria in the temperate oligotrophic North Pacific Ocean. *Limnology and Oceanography, 52*(4), 1317.

Ramakers, C., Ruijter, J. M., Deprez, R. H. L., Moorman, A. F. (2003). Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neuroscience Letters, 339*(1), 62-66.

Rijkenberg, M. J., Steigenberger, S., Powell, C. F., Haren, H., Patey, M. D., Baker, A. R., Achterberg, E. P. (2012). Fluxes and distribution of dissolved iron in the eastern (sub-) tropical North Atlantic Ocean. *Global Biogeochemical Cycles, 26*(3), GB3004.

Rubin, M., Berman-Frank, I., Shaked, Y. (2011). Dust- and mineral-iron utilization by the marine dinitrogen-fixer *Trichodesmium*. *Nature Geoscience*, *4*, 529–534.

Shelley, R. U., Morton, P. L., Landing, W. M. (2014). Elemental ratios and enrichment factors in aerosols from the US-GEOTRACES North Atlantic transects. *Deep Sea Res. II,* (this issue).

Shiller, A. M. (1997). Manganese in surface waters of the Atlantic Ocean. *Geophysical Research Letters*, *24*(1), 1495-1498.

Sohm, J. A., Subramaniam, A., Gunderson, T. E., Carpenter, E. J., Capone, D. G. (2011). Nitrogen fixation by *Trichodesmium spp*. and unicellular diazotrophs in the North Pacific subtropical gyre. *Journal of Geophysical Research: Biogeosciences (2005–2012), 116*, G03002.

Staal, M., Hekkert, S. T. L., Brummer, G. J., Veldhuis, M., Sikkens, C., Persijn, S., Stal, L. J. (2007). Nitrogen fixation along a north-south transect in the eastern Atlantic ocean. *Limnology and Oceanography, 52*(4), 1305.

Stal, L. J. (2009). Is the distribution of nitrogen-fixing cyanobacteria in the oceans related to temperature? *Environmental Microbiology, 11*(7), 1632-1645.

Stramma, L., Johnson, G. C., Sprintall, J., Mohrholz, V. (2008). Expanding oxygen-minimum zones in the tropical oceans. *Science, 320*(5876), 655-658.

Subramaniam, A., Yager, P. L., Carpenter, E. J., Mahaffey, C., Bjorkman, K., Cooley, S., Capone, D. G. (2008). Amazon River enhances diazotrophy and carbon sequestration in the tropical North Atlantic Ocean. *Proceedings of the National Academy of Sciences of the United States of America, 105*(30), 10460-10465.

Thompson, A. W., Zehr, J. P. (2013). Cellular interactions: Lessons from the nitrogen-fixing cyanobacteria. *Journal of Phycology, 49*(6), 1024-1035.

Turk, K. A., Rees, A. P., Zehr, J. P., Pereira, N., Swift, P., Shelley, R., Gilbert, J. (2011). Nitrogen fixation and nitrogenase (*nifH*) expression in tropical waters of the eastern North Atlantic. *The ISME Journal, 5*(7), 1201-1212.

Viana, M., Querol, X., Alastuey, A., Cuevas, E., Rodrıguez, S. (2002). Influence of African dust on the levels of atmospheric particulates in the Canary Islands air quality network. *Atmospheric Environment, 36*(38), 5861-5875.

Vitousek, P. M., Howarth, R. W. (1991). Nitrogen limitation on land and in the sea: How can it occur? *Biogeochemistry, 13*(2), 87-115.

Wozniak, A. S., Shelley, R. U., Sleighter, R. L., Abdulla, H. A. N., Morton, P. L., Landing, W. M. and Hatcher, P. G. (2013). Relationships among aerosol water soluble organic matter, iron and aluminum in European, North African, and Marine air masses from the 2010 US GEOTRACES cruise. *Marine Chemistry, 154*, 24-33.

Wurl, O., Zimmer, L., Cutter, G. A. 2013. Arsenic and phosphorus biogeochemistry in the ocean: Arsenic species as proxies for P-limitation. *Limnology and. Oceanogr*aphy, *58*(2), 729-740.

Zehr, J. P., Mellon, M. T., Zani, S. (1998). New nitrogen-fixing microorganisms detected in oligotrophic oceans by amplification of nitrogenase (*nifH*) genes. *Applied and Environmental Microbiology, 64*(9), 3444-3450.

Zimmer, L.A. and G. A. Cutter. 2012. High resolution determination of nanomolar concentrations of dissolved reactive phosphate in ocean surface waters using long path liquid waveguide capillary cells (LWCC) and spectrometric detection. *Limnology and Oceanography: Methods*, *10*, 568-580.

Tables

Table 1: Statistical comparison (ANOSIM) of SML *nifH* phylotype abundances with environmental variables.

	Threshold	R statistic[1)]	Significance level %
Physical parameters			
Temperature[2)]	22 °C	0.256	**0.5**
East – West	40°W	0.669	**0.1**
rain	presence	0.023	38.4
Nutrients (all low)			
N:16P ratio	1	0.723	5.9
Trace metals			
Dissolved Al[3)]	20 nM	0.053	27
Dissolved Fe[4)]	0.2 nM	-0.034	62.9
Dissolved Co	50 pM	-0.009	49.1
Dissolved Mn[5)]	2 nM	0.008	43.4
Dust			
Al aerosol[6)] (air sample)	50 ng m^{-3}	0.582	**0.1**
Fe aerosol[7)] (air sample)	50 ng m^{-3}	0.548	**0.1**
Aerosol origin		0.639	**0.1**

1) An R value with a significance level lower than 5% indicates that groupings are significantly different from each other. R increases with significance. Significant variables are highlighted in bold.
2) Breitbarth et al. (2007)
3) Dammshäuser et al. (2011)
4) Moore et al. (2009)
5) Shiller (1997)
6) Buck et al. (2010)
7) Buck et al. (2010)

778 Table 2: Average log abundances of discriminatory SML *nifH* phylotypes that contributed to the
779 overall dissimilarity between sample groupings (Dissimilarity/Standard Deviation >1) determined
780 using SIMPER.

Sample groupings	**Het 1**[1)]	**Het 2**[2)]	***Trichodesmium***	**UCYN A**	**UCYN B**	**UCYN C**	**Gamma A**
East of 40°W[3)]	***2.27***[4)]	1.33	4.22	***3.29***	***2.86***	1.76	3.73
West of 40°W	***0.90***	2.51	3.57	***0.32***	***0.39***	0.08	2.18
Marine aerosol	***0.96***	***1.76***	***3.64***	[5)]			
N. American aerosol	***1.05***	***3.31***	***3.30***				
Marine aerosol	0.96		***3.64***	***0.34***	***0.75***		2.11
N. African aerosol	1.18		***2.89***	***2.45***	***1.76***		3.02
N. American aerosol	1.05	***3.31***	3.30	***0.74***	***0.25***	0.00	
N. African aerosol	1.18	***1.39***	2.89	***2.54***	***1.75***	1.21	
high Fe aerosol[6)]	2.47	1.48	4.50	***3.62***	***2.99***	***2.03***	3.75
low Fe aerosol[7)]	1.04	2.35	3.56	***0.49***	***0.56***	***0.00***	2.35
High Temperature[8)]	1.52	***1.79***	***3.91***	***1.59***	1.61		2.87
Low Temperature[9)]	1.03	***3.60***	***3.35***	***1.09***	0.00		2.36

781 1) *Rhizosolenia-Richelia* symbiont
782 2) *Hemiaulus-Richelia* symbiont
783 3) Compared abundances are separated by dashed line.
784 4) The three major contributors to differences are printed in bold+italic
785 5) Not contributing phylotypes and phylotypes with a Dissimilarity/Standard Deviation below
786 1 are not displayed
787 6) Above 50 ng m^{-3}
788 7) Below 50 ng m^{-3}
789 8) Above 22 °C
790 9) Below 22 °C

791

Figure Captions

Figure 1: Cruise tracks and stations of USGT10 (triangles) and USGT11 (circles) in 2010 and 2011. Labelled are the time series stations BATS and CVOO as well as the mid-Atlantic ridge (MAR). Stations with very high *Trichodesmium nifH* abundance are indicated with open circles. The track of hurricane Sean (at BATS on 11 Nov. 2011) is overlaid in open squares.

Figure 2: Abundances of total *nifH* copies mL^{-1} measured in all samples in relation to the difference between SML depth and sample depth.

Figure 3: a) Sum of *nifH* copy numbers, b) the Shannon diversity index, c) temperature, d) N* and e) phosphate in the SML during cruises USGT10 (gray shaded) and USGT11 (no shading). Stations BATS, MAR and CVOO are indicated by open diamonds. If more than one sample was taken from the SML, averages are shown.

Figure 4: Average *nifH* copy (mL^{-1}) numbers of a) *Trichodesmium*, b) UCYN A, c) Gamma A (black diamonds) UCYN B (grey triangles) *Hemiaulus-Richelia* (open squares), and d) UCYN C (open triangles) *Rhizosolenia-Richelia* (black circles) in the SML and e) aluminium (black triangles) and iron (black crosses) concentrations from aerosol samples (ng m^{-3}) from USGT10 (gray shading) and USGT11 (no shading). Stations BATS, MAR and CVOO are indicated by white symbols.

Figure 5: Principal Components Analysis (PCA) of SML samples from USGT10 and USGT11 showing variables that contributed to significant clustering of samples. Significant clusters are traced with a line representing a Euclidean-distance of 3 obtained from a hierarchical cluster analysis of the samples. a) Samples west of 40°W are indicated with open squares and eastern samples with black circles, b) high aerosol iron concentrations (above 50 ng m^{-3}) plotted as open squares, low iron aerosol concentrations (below 50 ng m^{-3}) plotted as black circles (Shelley et al. 2014), c) aerosol origin as back trajectories over the past 5 days (Shelley et al. 2014): Marine (black circles), North African (black diamonds), North American (open squares)

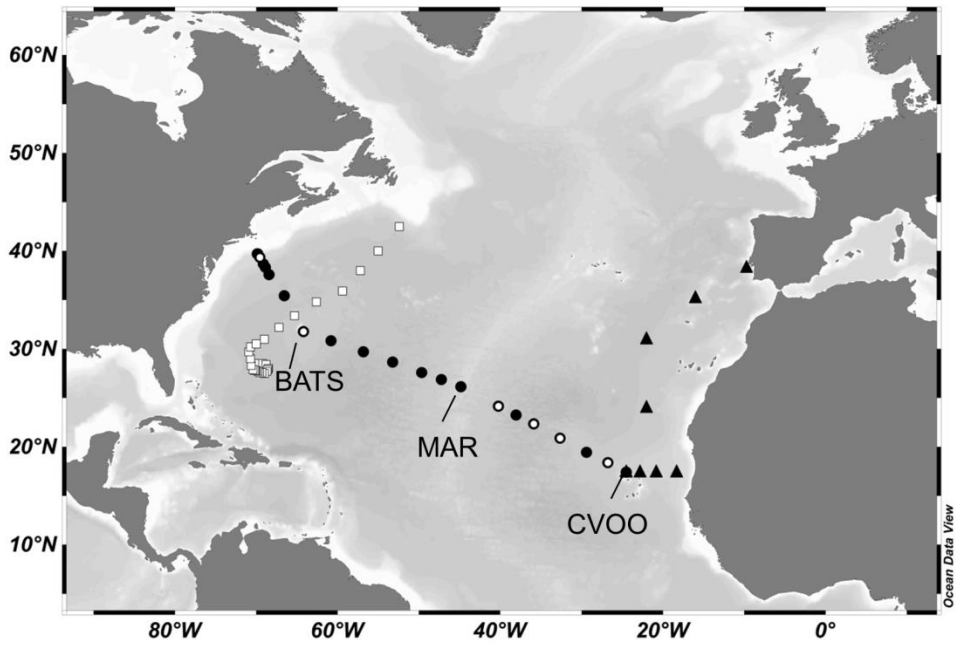
60°N
50°N
40°N
30°N
20°N
10°N
80°W
60°W
40°W
20°W
0°
BATS
MAR
CVOO
Ocean Data View

Figure 1

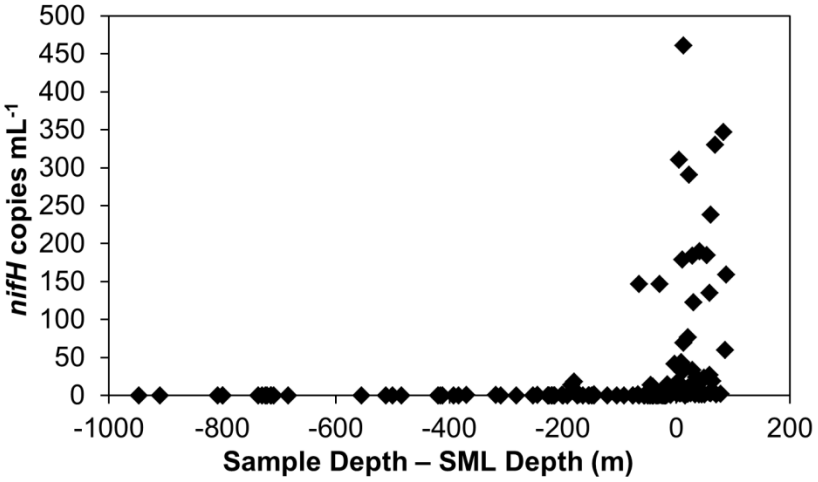
500
450
400
350
300
250
200
150
100
50
0
nifH copies mL-1
-1000
-800
-600
-400
-200
0
200
Sample Depth – SML Depth (m)

Figure 2

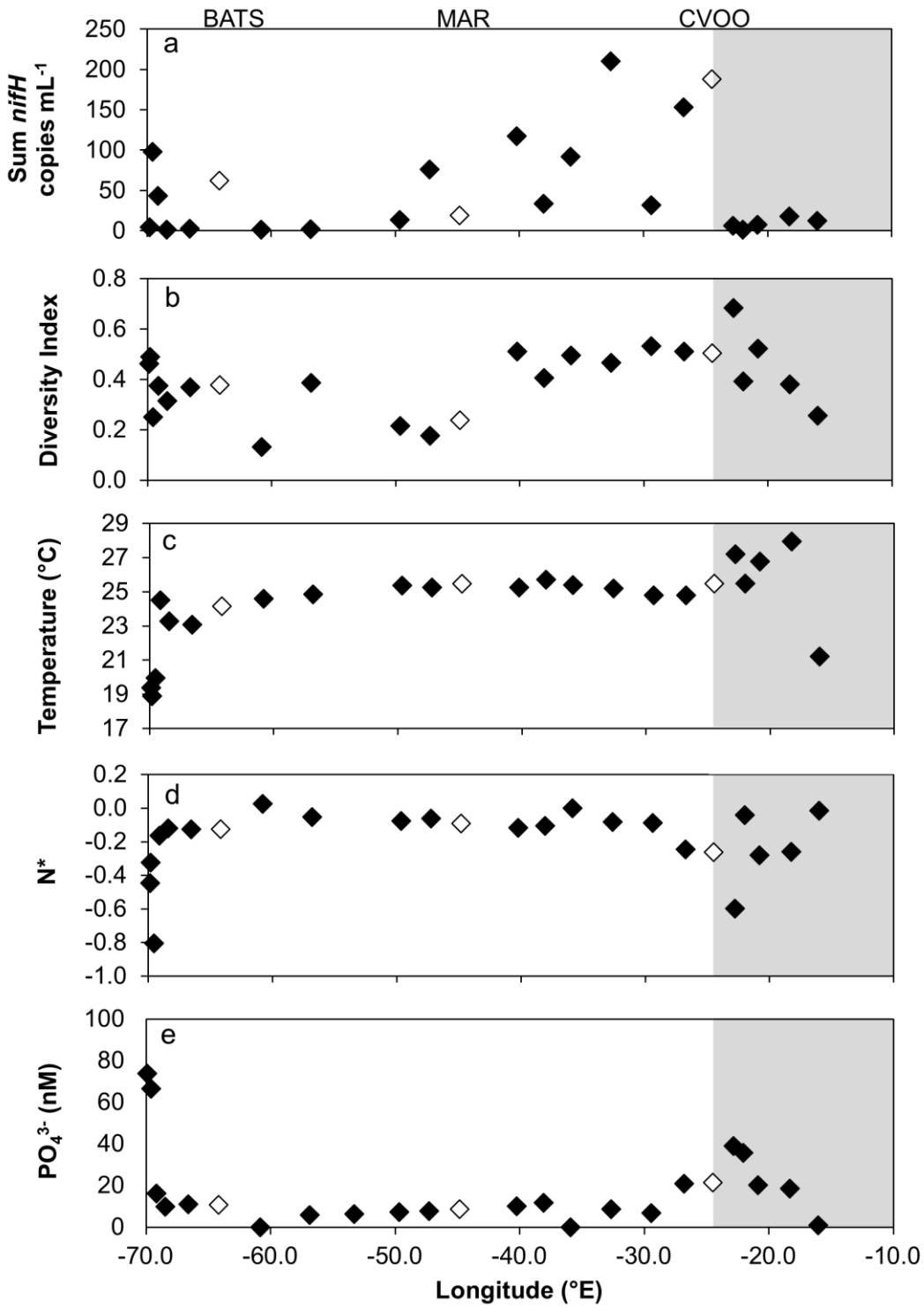
BATS
MAR
CVOO
a
b
c
d
e
Sum nifH copies mL-1
Diversity Index
Temperature (°C)
N*
PO43- (nM)
Longitude (°E)
250
200
150
100
50
0
0.8
0.6
0.4
0.2
0.0
29
27
25
23
21
19
17
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
100
80
60
40
20
0
-70.0
-60.0
-50.0
-40.0
-30.0
-20.0
-10.0

Figure 3

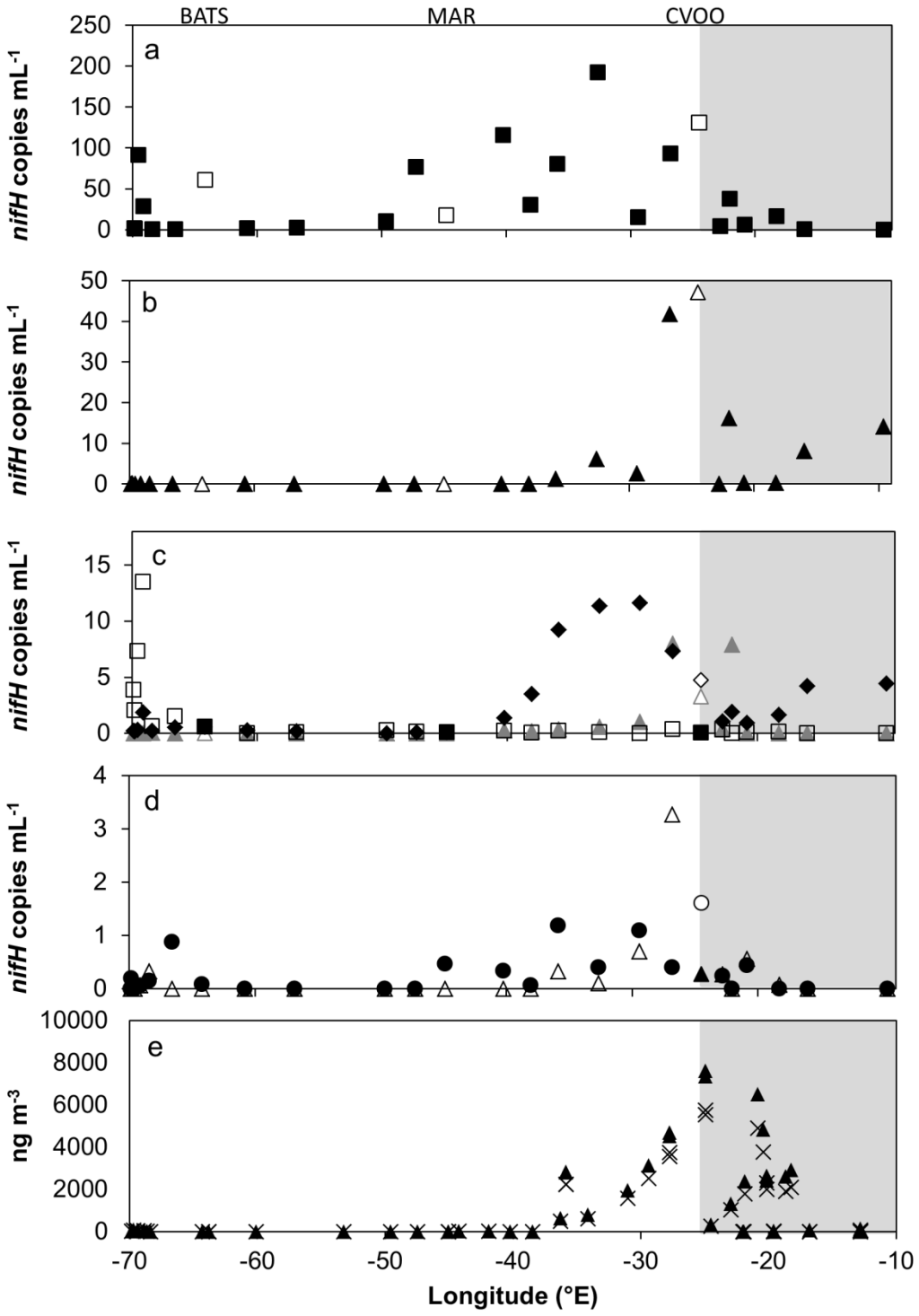

Figure 4

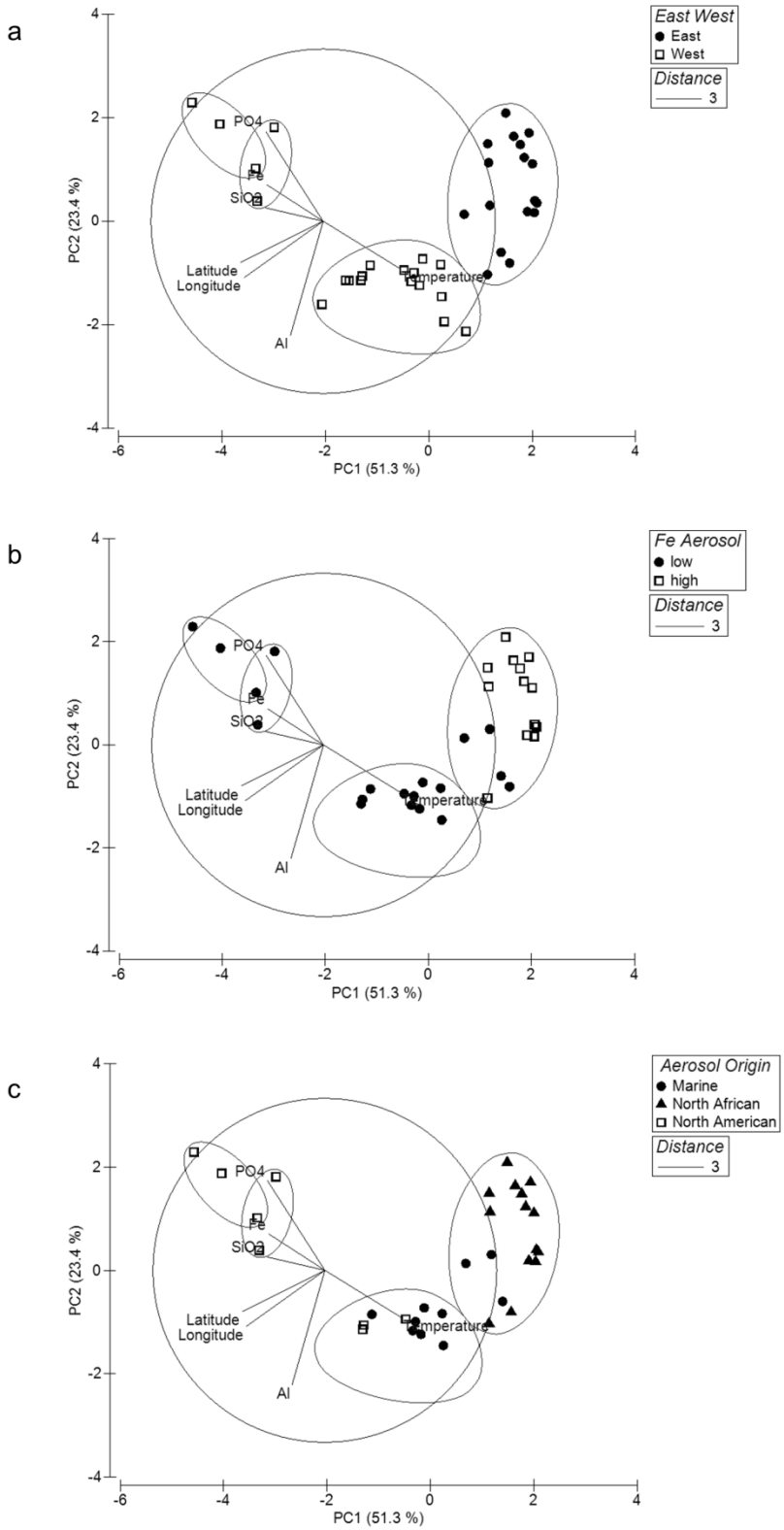

Figure 5

827
828