



# Diazotroph Diversity in the Sea Ice, Melt Ponds, and Surface Waters of the Eurasian Basin of the Central Arctic Ocean

Mar Fernández-Méndez<sup>1,2\*</sup>, Kendra A. Turk-Kubo<sup>3</sup>, Pier L. Buttigieg<sup>1</sup>, Josephine Z. Rapp<sup>1,2</sup>, Thomas Krumpfen<sup>4</sup>, Jonathan P. Zehr<sup>3</sup> and Antje Boetius<sup>1,2</sup>

<sup>1</sup> HGF-MPG Group for Deep Sea Ecology and Technology, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany, <sup>2</sup> HGF-MPG Group for Deep Sea Ecology and Technology, Max Planck Institute for Marine Microbiology, Bremen, Germany, <sup>3</sup> Department of Ocean Sciences, University of California at Santa Cruz, Santa Cruz, CA, USA, <sup>4</sup> Sea Ice Physics Section, Climate Sciences Department, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany

## OPEN ACCESS

### Edited by:

Angela Landolfi,  
GEOMAR Helmholtz Centre for Ocean  
Research Kiel, Germany

### Reviewed by:

James T. Hollibaugh,  
University of Georgia, USA  
Mikkel Bentzon-Tilia,  
Technical University of Denmark,  
Denmark  
Isabelle C. Biegala,  
Institut de Recherche pour le  
Développement (IRD), France

### \*Correspondence:

Mar Fernández-Méndez  
[mar.fdez.mendez@gmail.com](mailto:mar.fdez.mendez@gmail.com)

### Specialty section:

This article was submitted to  
Aquatic Microbiology,  
a section of the journal  
Frontiers in Microbiology

**Received:** 29 June 2016

**Accepted:** 09 November 2016

**Published:** 23 November 2016

### Citation:

Fernández-Méndez M,  
Turk-Kubo KA, Buttigieg PL,  
Rapp JZ, Krumpfen T, Zehr JP and  
Boetius A (2016) Diazotroph Diversity  
in the Sea Ice, Melt Ponds,  
and Surface Waters of the Eurasian  
Basin of the Central Arctic Ocean.  
*Front. Microbiol.* 7:1884.  
doi: 10.3389/fmicb.2016.01884

The Eurasian basin of the Central Arctic Ocean is nitrogen limited, but little is known about the presence and role of nitrogen-fixing bacteria. Recent studies have indicated the occurrence of diazotrophs in Arctic coastal waters potentially of riverine origin. Here, we investigated the presence of diazotrophs in ice and surface waters of the Central Arctic Ocean in the summer of 2012. We identified diverse communities of putative diazotrophs through targeted analysis of the *nifH* gene, which encodes the iron protein of the nitrogenase enzyme. We amplified 529 *nifH* sequences from 26 samples of Arctic melt ponds, sea ice and surface waters. These sequences resolved into 43 clusters at 92% amino acid sequence identity, most of which were non-cyanobacterial phylotypes from sea ice and water samples. One cyanobacterial phylotype related to *Nodularia* sp. was retrieved from sea ice, suggesting that this important functional group is rare in the Central Arctic Ocean. The diazotrophic community in sea-ice environments appear distinct from other cold-adapted diazotrophic communities, such as those present in the coastal Canadian Arctic, the Arctic tundra and glacial Antarctic lakes. Molecular fingerprinting of *nifH* and the intergenic spacer region of the rRNA operon revealed differences between the communities from river-influenced Laptev Sea waters and those from ice-related environments pointing toward a marine origin for sea-ice diazotrophs. Our results provide the first record of diazotrophs in the Central Arctic and suggest that microbial nitrogen fixation may occur north of 77°N. To assess the significance of nitrogen fixation for the nitrogen budget of the Arctic Ocean and to identify the active nitrogen fixers, further biogeochemical and molecular biological studies are needed.

**Keywords:** nitrogen fixation, *nifH*, Arctic, non-cyanobacterial diazotrophs, sea ice, bacterial diversity

## INTRODUCTION

Arctic marine ecosystems are rapidly changing due to climate change (Wassmann et al., 2011). In summer 2012, Arctic sea ice extent declined to its minimum ever recorded (Parkinson and Comiso, 2013), following the general trend in Arctic sea-ice cover decline (Stroeve et al., 2012). This led to vast areas of water being exposed to higher light intensities, which enhances phytoplankton

primary production (Arrigo and van Dijken, 2015). Primary production in this highly stratified ocean is generally limited by nitrogen, most of which is delivered in the form of nitrate via the inflow of Pacific and Atlantic waters (Codispoti et al., 2013). Riverine input of nitrogen, especially from the Lena river that discharges into the Laptev Sea, may be another important source of nitrate in the Eurasian Basin (Holmes et al., 2011), alongside atmospheric deposition by snowfall (Beine et al., 2003), recycling of organic matter, and nitrogen fixation in ice masses. These different nitrogen sources should be balanced by nitrogen sinks, such as primary production and export, as well as microbial denitrification along the broad continental shelves that surround the Arctic Ocean (Devol et al., 1997; Rysgaard et al., 2004). However, current nutrient budgets indicate an imbalance of Arctic nitrogen supply, demand and export, suggesting that nitrogen fixation may have a key role in closing the budget (Torres-Valdés et al., 2013).

The biological fixation of gaseous dinitrogen ( $N_2$ ) by diazotrophs is an important source of bioavailable nitrogen in nutrient-limited pelagic ecosystems (LaRoche and Breitbarth, 2005), thereby influencing primary productivity and carbon export to the seafloor (Codispoti et al., 2001; Arrigo, 2005). Nitrogen fixation rates in surface waters of tropical, subtropical and some temperate oceans are primarily driven by diazotrophic cyanobacteria (Langlois et al., 2008; Moisander et al., 2010; Turk-Kubo et al., 2012). Non-cyanobacterial diazotrophs have also been detected in numerous marine, pelagic environments (Falcón et al., 2004; Riemann et al., 2010; Farnelid et al., 2011; Moisander et al., 2014), although their contribution to oceanic nitrogen fixation remains poorly understood (Turk-Kubo et al., 2014). In contrast to temperate and tropical seas, the specific contribution of nitrogen fixation and its key microbial agents in ice-covered seas is not well known (Tremblay et al., 2009; Blais et al., 2012; Luo et al., 2012; Torres-Valdés et al., 2013). This study aims at providing a first survey of the nitrogen-fixing potential of the eastern Central Arctic Ocean.

Several factors are thought to constrain oceanic nitrogen fixation: low temperatures (Brauer et al., 2013), high dissolved oxygen, high N:P ratios, and iron or phosphate limitation (Paerl and Zehr, 2000; Mills and Arrigo, 2010; Riemann et al., 2010; Monteiro et al., 2011). The generally low temperatures and the high dissolved oxygen in Arctic waters (Bates et al., 2014) support the long-held assumption that nitrogen fixation is unlikely in the Arctic Ocean. However, diazotrophs have developed strategies to withstand freezing and high salinity by producing antifreeze proteins (Schmidt et al., 1991), and to avoid oxygen by developing endo-symbiosis, like *Rhizobiales*, which include mainly *Alpha-proteobacteria* and *Beta-proteobacteria* and have been identified also in polar soils and frost flowers (Bordeleau and Prévost, 1994; Bowman et al., 2013). Symbiotic associations between diatoms and diazotrophs have been described in other oceans (Villareal, 1992; Foster et al., 2011). We hypothesize that sea-ice related environment of the Eurasian basin of the Central Arctic Ocean might hold diazotrophs.

Arctic waters often have low N:P ratios (Tremblay et al., 2008), which have been hypothesized to favor diazotrophs (Tyrrell, 1999). They are relatively rich in iron (1–3 nM) from river

input (Klunder et al., 2012), which could well support the iron demand of nitrogen fixing enzymes. Nitrogen fixation rates and putative diazotrophs from riverine origin (microorganisms containing at least the *nifH* gene) have been identified close to the Mackenzie River and in the water column of the Canadian Arctic shelves (Farnelid et al., 2011; Blais et al., 2012). Relatively high cyanobacterial *nifH* gene diversity (e.g., *Trichodesmium* and *Cyanothece*) has also been described in the sea ice and waters of the Fram Strait in the Eurasian Basin (Díez et al., 2012). In contrast, to date there is no record of marine diazotroph occurrence in the Central Arctic Ocean north of 76°N. Non-diazotrophic cyanobacteria, such as *Prochlorococcus*, which are very abundant in other oceanic regions (Partensky and Hess, 1999), are rare in polar marine waters (Vincent, 2000; Lovejoy and Potvin, 2010) although they appear in both polar regions (Nadeau et al., 2001). Diazotrophic cyanobacteria, which represent a small percentage of marine cyanobacteria, are also the most common autotrophic diazotroph in other oceans (Moisander et al., 2010; Luo et al., 2012) and on glaciers (Yallop et al., 2012), but their diversity and abundance in the marine environments of the polar regions is still understudied.

In the Arctic Ocean, only a few *nifH* gene surveys have been conducted in the past decade, and they focused on coastal regions. In this study, we investigated *nifH* gene diversity in different Arctic environments, including Arctic sea ice and surface water from under-ice and open water locations. To assess diazotroph diversity and discover specific phylotypes that might be unique for the Arctic marine environment, we chose a targeted analysis of the *nifH* gene, coding for the iron protein of the nitrogenase enzyme. In addition, we used molecular fingerprinting analyses of the intergenic spacer region of the rRNA operon to characterize the complete bacterial community, and to reveal the dissimilarity patterns between different sea-ice related environments. Furthermore, we measured the physical-chemical properties of each environment type and we included data about the origin of the sea ice, measured by its drift trajectory.

These datasets allowed us to assess the distribution and diversity of diazotrophs in melt pond, sea-ice and surface waters of the Central Arctic and to evaluate the genetic nitrogen-fixing potential of the microbial communities. Specifically, we addressed the hypotheses that (1) putative diazotrophs of the Central Arctic occupy a niche in the nitrogen-limited sea ice and waters of the Eurasian Basin, (2) they originate from coastal areas such as the Laptev Sea where the ice is formed and transported with the Transpolar Drift, and (3) the Central Arctic marine sea-ice related diazotrophs are distinct from other cold-adapted diazotrophic communities and those of adjacent oceanic regions.

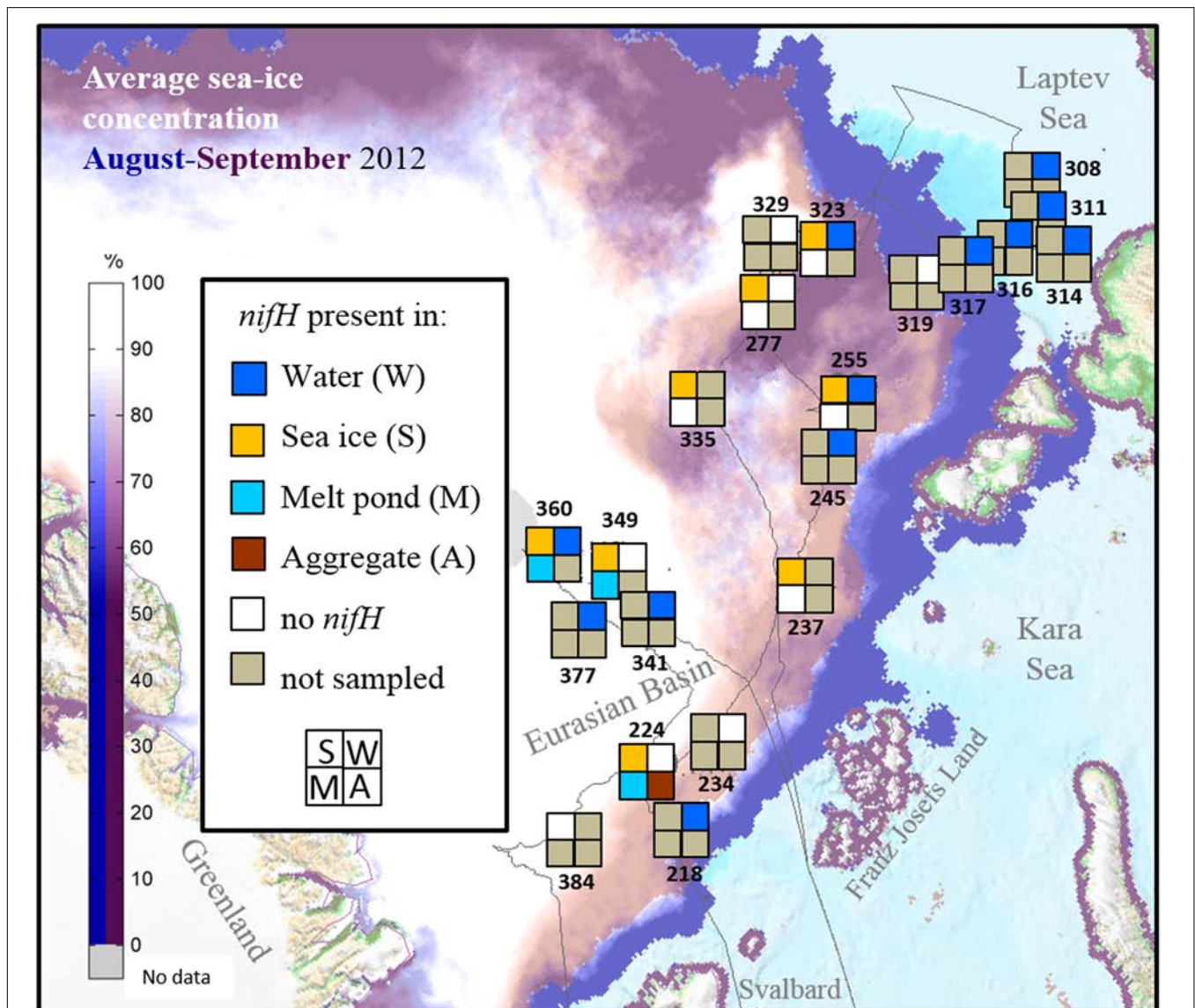
## MATERIALS AND METHODS

### Sampling

A total of 53 samples were taken for this investigation during the ARKXXVII/3 cruise to the Central Arctic (77–88°N and 30–133°E) from the 7th of August to the 30th of September, 2012, onboard the German icebreaker R/V POLARSTERN

(Supplementary Table S1). Sea ice, melt ponds and seawater were sampled at 9 ice stations and 13 water stations located in different parts of the central Eurasian Basin. Surface water samples from ice-covered waters ( $n = 13$ ) were differentiated from those from open water from the Laptev Sea region ( $n = 6$ ) for the analysis. Our samples come from a wide range of ice conditions and nutrient regimes (Figure 1). Samples of rapidly melting sea ice ( $n = 21$ ) were taken using an ice corer (Kovacs Enterprise, Roseburg, OR, USA) and cut into two sections (top and bottom) prior to being transferred to plastic containers that had been rinsed with ultrapure water and ethanol and melted in the dark

at 4°C. Samples of melt pond water ( $n = 8$ ) and algal aggregates ( $n = 5$ ) found in melt ponds were collected using a hand pump (Model 6132-0010, Nalgene, Penfield, NY, USA) and stored in clean Nalgene bottles. Samples of water under the ice ( $n = 5$ ) were collected using a peristaltic pump (Masterflex® E/S™ portable sampler, 115 VAC, Oldham, UK). Surface water (2–5 m depth) samples at all other stations ( $n = 13$ ) were collected with a rosette sampler equipped with Niskin bottles and a Conductivity Temperature Depth (CTD) profiler (Sea-Bird Electronics Inc., Bellevue, WA, USA). In addition, ultrapure water was sampled to check for possible contaminations from the onboard water



**FIGURE 1 | Stations sampled for *nifH* analysis in the Eurasian Basin of the Central Arctic during summer 2012.** Station numbers correspond to those in Supplementary Table S1. Each square correspond to one environment. If a particular environment was not samples at one station it appears in gray. If no *nifH* gene could be amplified it is colored in white, and if *nifH* genes could be amplified in the color corresponding to each environment (Dark blue = water, Yellow = sea ice, Light Blue = melt pond and Brown = algal aggregates). The thin gray line corresponds to the cruise track. Average sea-ice concentration data source: [www.meereisportal.de](http://www.meereisportal.de).

purification system (Milli-Q Gradient A10, Millipore) when using this water to rinse the sampling equipment. Between 0.3 and 2 L volume of each sample were filtered through Sterivex filters of 0.2  $\mu\text{m}$  pore size (Durapore, Millipore, Darmstadt, Germany) using a multichannel peristaltic pump (Model PD 51; Heidolph, Schwabach, Germany). Filters were stored at  $-80^{\circ}\text{C}$  until further processing.

## Characterization of Central Arctic Environments

Temperature and salinity were measured in sea ice and melt pond water with a hand-held conductivity meter (315i with TetraCon electrode cell, WTW GmbH, Weilheim in Oberbayern, Germany), and in the water column with a CTD profiler. The data is available in the databases of the Data Publisher for Earth and Environmental Science, PANGAEA, doi: 10.1594/PANGAEA.819452 (Rabe et al., 2012). Concentrations of nitrate, phosphate and silicate were measured with a standard photometric method using a Technicon TRAACS 800 continuous flow auto analyzer (Technicon Corporation) as described in Fernández-Méndez et al. (2015). Raw nutrient concentrations and integrated stocks are available in PANGAEA doi: 10.1594/PANGAEA.834081 (Fernández-Méndez et al., 2014a; Bakker, 2014, unpublished).

Data concerning ice thickness and melt pond coverage are stored in PANGAEA, doi: 10.1594/PANGAEA.803221 (Hendricks et al., 2012). The origin of sea ice was determined using ice drift information obtained from satellites. In this study, two different sets of ice drift products were used: The first data set, Polar Pathfinder Sea Ice Motion Vectors (Version 2) obtained from the National Snow and Ice Data Center (NSIDC) was chosen because of its year round availability. We used it to calculate ice drift trajectories during summer months (June–August). The second data set, sea ice motion provided by the Center for Satellite Exploitation and Research (CERSAT) at the Institut Francais de Recherche pour l'Exploitation de la Mer (IFREMER), shows a good performance on the Siberian shelf (Krumpfen et al., 2013) and was therefore used to complement the calculation of ice drift trajectories between September and May. To determine drift trajectories and source areas of sampled sea ice a specific ice area is tracked backward until: (a) the ice reaches a position next to a coastline, (b) the ice concentration at a specific location reaches a threshold value of ( $>15\%$ ) when ice parcels are considered lost, or (c) the tracking time exceeds 4 years. A more detailed method description is provided in Krumpfen et al. (2016).

## *nifH* Gene Molecular Analysis

Total community DNA was extracted using the DNeasy Plant Mini Kit (QIAGEN, Valencia, CA, USA) and the QIAcube extraction instrument following the manufacturer's instructions (Bombar et al., 2013). To amplify the *nifH* gene, a nested polymerase chain reaction (PCR) employing a high number of amplification cycles ( $n = 50$ ) with two sets of degenerate primers was used since it is a high-sensitivity method (Zehr and Turner, 2001). The four degenerate primers used, *nifH1-nifH4*, cover  $>94\%$  of available *nifH* sequences, allowing our approach to

access a high percentage of the known diazotroph diversity (Gaby and Buckley, 2012) at the limited availability of sample volume in this study. PCR amplifications were performed in a MyCycler Thermal Cycler (BioRad, Berkeley, CA, USA). The first PCR amplification of *nifH* in each sample was performed with 2  $\mu\text{l}$  of DNA template in 24  $\mu\text{l}$  of PCR reaction containing 4 mM  $\text{MgCl}_2$ , 0.4 mM dNTPs, 10X Buffer, 1.25 U (0.2  $\mu\text{l}$ ) Platinum TaqDNA Polymerase (Invitrogen) and 0.5  $\mu\text{M}$  of the forward *nifH3* primer (5'-ATR TTR TTN GCN GCR TA-3') and reverse *nifH4* primer (5'-TTY TAY GGN AAR GGN GG-3') (Zehr and Turner, 2001). Amplifications proceeded with 25 cycles of 3 min at  $95^{\circ}\text{C}$ , 30 s of denaturation at  $95^{\circ}\text{C}$ , 30 s of annealing at  $55^{\circ}\text{C}$  and 45 s of elongation at  $72^{\circ}\text{C}$ . The second amplification procedure was performed as described above, save the use of 1  $\mu\text{l}$  of the PCR product from the first amplification process as the template and the forward *nifH1* primer (5'-TGY GAY CCN AAR GCN GA-3') and reverse *nifH* primer (5'-ADN GCC ATC ATY TCN CC-3') (Zehr and Turner, 2001). In addition, the annealing temperature was increased to  $57^{\circ}\text{C}$ . In both steps, negative controls were performed using milliQ water in place of the DNA template.

A 15  $\mu\text{l}$  aliquot of the PCR products from the second amplification was used for electrophoresis separation (1,4% low-melt agarose gel). The DNA band of the appropriate size (360 bp) was extracted with a gel extractor (X-tracta gel extractor, USA Scientific, Ocala, FL, USA). The gel was purified with a QIAquick Gel Extraction Kit (Qiagen) and the DNA cloned with a TOPO<sup>®</sup> TA Cloning<sup>®</sup> Kit for Sequencing with One Shot<sup>®</sup> TOP10 Chemically Competent *E. coli* (Invitrogen) following the manufacturer's guidelines. Dependent on outcomes of the cloning process, plasmids from between 12 and 48 clones per sample were purified using the Millipore Montage Plasmid Miniprep<sub>96</sub>KitsMiniprep kit (Millipore, Darmstadt, Germany) and the inserts were sequenced using the Sanger method (Sanger et al., 1977) at the University of California, Berkeley.

Nucleic acid sequence data were trimmed and quality checked using the Sequencher<sup>®</sup> sequence analysis software (Gene Codes Corporation, Ann Arbor, MI, USA). Quality-controlled sequences were imported into the software program ARB (Ludwig et al., 2004), translated to amino acid sequences and imported into a publically available, curated GenBank database containing all *nifH* sequences submitted to the nr database that have been validated to be *nifH* based on analysis (Heller et al., 2014; April 2014 release). Amino acid sequences were aligned using the program HMMER which contains a Hidden Markov Model from the protein family database PFAM (Finn et al., 2010). Subsequently, the nucleotide sequences were re-aligned according to the aligned amino acid sequences using ARB. To confirm that the sequences were *nifH* gene, the alignments were visually checked for conserved regions. Our 572 confirmed *nifH* sequences included one sequence from the PCR blank and 28 sequences from the ultrapure water used to rinse the plastic containers where the samples were stored. These sequences clustered at  $>94\%$  amino acid sequence identity with 15 sequences from our environmental samples, so these were considered putative contaminants and removed from further analyses. We also checked if any of the retrieved sequences were closely related to other sequences reported as

contaminants on other studies. The 45 contaminant amino acid sequences reported in the April 2014 *nifH* database (Farnelid et al., 2009, 2013) were all less than 90% similar to our sequences retrieved from the environment. Despite these checks, one cannot exclude the possibility that some of the sequences reported might be contaminants. From the 529 confirmed environmental *nifH* sequences, a total of 43 clusters with  $\geq 92\%$  amino acid sequence identity were identified using the CD-HIT program suite (Huang et al., 2010). Each cluster was represented by one representative sequence. A maximum likelihood tree of partial *nifH* sequences was built in RAxML version 8.1 (Stamatakis, 2014). This tree included the 43 representative sequences (92% amino acid similarity using the matrix BLOSUM62) of our *nifH* Arctic sequences and their closest cultivated relatives. We used the RAxML function *auto prot* to determine the most appropriate protein substitution model. Following the function's recommendation, we used the LG matrix substitution model to determine relatedness (represented by branch length) and a custom mask for the *nifH* amplicon region (Mask Makyyy in the ARB publicly available database). The stability of the RAxML phylogenetic reconstruction was assessed by bootstrapping 1000 times with RAxMLs *-b* function. The tree and its associated metadata were visualized using ITOL (Letunik and Bork, 2007). The phylogenetic affiliation of each *nifH* cluster, defined by CD-HIT, was determined based on the phylogenetic affiliation of the closest cultivated relative, which have been assigned according to the convention suggested by Zehr et al. (2003) and provided in the curated *nifH* database (Heller et al., 2014).

For the comparison of different regions, sequences from nine independent studies, including the present one, were selected from studies submitted to the GenBank database prior to April 2014. All available Arctic sequences from different environments were selected, as well as all Antarctic sequences, which were exclusively from lake microbial mats. In addition, sequences from one study in the North Atlantic (constrained to surface water samples as in our study) and one in the subtropical Atlantic (water samples collected at 8 m depth) were selected for further comparison outside of the polar regions. In total, 1523 sequences were chosen to characterize the diazotroph community structure from: the Central Arctic (this study) ( $n = 572$ ), the coastal Canadian Arctic (Blais et al., 2012) ( $n = 21$ ), the coastal Eurasian Basin (Díez et al., 2012) ( $n = 69$ ), the Arctic tundra soil (Izquierdo and Nüsslein, 2006) ( $n = 24$ ), the Antarctic ice shelf microbial mats in lakes (Olson et al., 1998; Jungblut and Neilan, 2010) ( $n = 15+43$ ), the subtropical Atlantic Ocean (Langlois et al., 2005) ( $n = 175$ ), and the North Atlantic Ocean (Turk et al., 2011) ( $n = 604$ ). All studies except Díez et al., 2012 used the same primers as used in this study. Nucleotide sequences from these studies were clustered using CD-HIT EST (Li and Godzik, 2006) at 97% nucleotide similarity (Turk-Kubo et al., 2014). From the 243 representative sequences (of 1523 initial sequences), 86 originated from *nifH* genes present in the Central Arctic, 19 from the Canadian Arctic, 20 from the Eurasian Basin, 16 from the Arctic tundra, 25 from the Antarctic microbial mats, 51 from the North Atlantic, and 26 from the subtropical Atlantic. The representative sequences at 97% nt identity retrieved from the Central Arctic sea-ice related environments as well as the

ultrapure water blanks were submitted to GenBank and assigned accession numbers KT354077-KT354180. Only five of the CD-HIT clusters represented sequences from both the North Atlantic and the subtropical Atlantic and one CD HIT cluster represented sequences from the Eurasian Basin and the Subtropical Atlantic. All other CD-HIT clusters represented sequences exclusively from one oceanic region. A maximum likelihood tree of these 243 representative sequences was built using RAxML v8.1. The best performing model was the WAG matrix substitution model. The stability of the RAxML tree was assessed by bootstrapping 1000 times with RAxMLs *-b* function. The tree and its associated metadata were visualized using ITOL (Letunik and Bork, 2007).

## Intergenic Spacer 16SrRNA Molecular Analysis

The bacterial community structure was investigated targeting the Intergenic Spacer (ITS) region of the 16SrRNA by Automated Ribosomal Intergenic Spacer Analysis (ARISA) (Fisher and Triplett, 1999) in the 26 samples where *nifH* could be amplified. ARISA was chosen as a rapid and effective method to analyze bacterial community structure across this large dataset. Results of this method were previously shown to be coherent with diversity patterns retrieved using next generation sequencing approaches (e.g., Gobet et al., 2014). Each ARISA PCR contained 2.5  $\mu\text{l}$  of 10x reaction buffer S (PEQLAB Biotechnologie GmbH, Erlangen, Germany), 1  $\mu\text{l}$  of 25 mM  $\text{MgCl}_2$  (PEQLAB Biotechnologie GmbH, Erlangen, Germany), 0.625  $\mu\text{l}$  of a 10 mM dNTP mix (PEQLAB Biotechnologie GmbH, Erlangen, Germany), 0.75  $\mu\text{l}$  of 3 mg  $\text{ml}^{-1}$  bovine serum albumin (Sigma-Aldrich Chemie GmbH; Munich, Germany), 0.25  $\mu\text{l}$  of 40  $\mu\text{M}$  universal forward primer ITSF (5'-GTCGTAACAAGGTAGCCGTA-3') (Biomers.net, Ulm, Germany), labeled with 6-carboxyfluorescein (FAM), 0.25  $\mu\text{l}$  of 40  $\mu\text{M}$  ITSReub reverse primer (5'-GCCAAGGCATCCACC-3') (Biomers.net, Ulm, Germany), 0.25  $\mu\text{l}$  of 5 units  $\mu\text{l}^{-1}$  Taq polymerase (PEQLAB Biotechnologie GmbH, Erlangen, Germany) and approximately 10 ng environmental DNA as determined by spectrophotometry (Infinite® M200 NanoQuant, Tecan Group Ltd., Switzerland). PCR water was added to each reaction mix to a final volume of 25  $\mu\text{l}$ . All reactions were conducted in triplicate. PCR conditions were set to 3 min at 94°C, 30 cycles of first 94°C for 45 s, 55°C for 45 s, and 72°C for 90 s followed. Final extension time was 5 min at 72°C.

PCR performance was examined using gel electrophoresis and the length of successfully amplified and purified ITS fragments was analyzed via capillary electrophoresis on an ABI Prism 3130 XL – Genetic Analyzer (Applied Biosystems, Carlsbad, CA, USA). Signals were evaluated using the GeneMapper Software v3.7 (Applied Biosystems, Carlsbad, CA, USA) and subsequent generation of operational taxonomic units (OTUs) was done with custom R scripts<sup>1</sup> as previously described (Ramette, 2009).

## Statistical Analysis and Ordination

All statistical tests were performed in R version 3.1.1 (R Core Team, 2015). Differences in diazotroph community structure –

<sup>1</sup><http://www.ecology-research.com>

based on the presence or absence of *nifH*-derived operational taxonomic units – were expressed using the Jaccard dissimilarity measure, which was computed as  $2B/(1+B)$ , where  $B$  is the Bray–Curtis dissimilarity measure:

$$B = \frac{\sum_{i=1}^S (x_{ij} - x_{ik})}{\sum_{i=1}^S (x_{ij} + x_{ik})}$$

and visualized by non-metric multidimensional scaling (NMDS) using the metaMDS routine of the “vegan 2.3” package (Oksanen et al., 2013). Differences in the total bacterial community structure – represented through variations in the internal transcribed spacer region (ITS) fingerprint – were similarly visualized, but calculated using the Bray–Curtis dissimilarity measure (Legendre and Legendre, 1998). Guided by the Shepherd stress of our NMDS results, we chose to ordinate ranked dissimilarities from both analyses in two-dimensional space. In each plot, dissimilarity between samples is approximated by the distances between the points representing them. Stress values reflect the degree of correspondence between the distances between points in the NMDS plot and values in the original dissimilarity matrix. An analysis of similarity (ANOSIM) was used to evaluate how strongly our representations of bacterial community structure differed between *a priori* defined groups. The groups defined were: Laptev Sea ( $n = 4$ ), Surface Waters ( $n = 6$ ), Melt Ponds ( $n = 3$ ), Ice Bottom ( $n = 3$ ), Ice Top ( $n = 8$ ), Brown ice ( $n = 1$ ), and Aggregate ( $n = 1$ ). Groups with one sample only were excluded from the analysis. First, we performed an omnibus test followed by a pairwise ANOSIM as *post hoc* testing. Individual  $p$ -values were Bonferroni-corrected. A Mantel test was performed using the mantel routine of the “vegan 2.3” package to compare the two dissimilarity matrices (*nifH* and ARISA).

To complement our dissimilarity-based methods, which may confound location and spread (Warton et al., 2012), redundancy analysis (RDA) using type I scaling was used to estimate the degree of linear association between variation in the diazotroph community and temperature, salinity, nitrate, phosphate and silicate. These environmental variables were standardized by  $z$ -scoring prior to use in RDA. The response data was Hellinger transformed to standardize it (Legendre and Gallagher, 2001), using the *decostand* function of the “vegan 2.3” package. The significance of the RDA analysis was calculated using ANOVA and the residuals were plotted in a histogram (Supplementary Figure S3).

To compare the diazotroph community of the Central Arctic Ocean described in this study to the diazotroph community from Arctic adjacent oceanic regions (Canadian and Eurasian Arctic shelves), other polar environments (Antarctic lake microbial mats and Arctic tundra soil), and other oceanic regions (North and Subtropical Atlantic), we performed phylogenetic distance analysis (Unifrac) (Lozupone et al., 2006) on the maximum likelihood phylogenetic tree containing only representative sequences from those environments (CD-HIT clustering at a 97% nucleotide similarity threshold). To calculate dissimilarities based on phylogenetic distances between the diazotrophic communities we used the package “phyloseq” in R (McMurdie

and Holmes, 2013). Principal coordinate analysis (PCoA) was used to ordinate Unifrac distances using the functions *cmdscale()* and *ordiplot()* in the R package “vegan 2.3.” An automatic correction for negative eigenvalues was used. Rarefaction curves were computed across Hill numbers (Chao et al., 2014) using the iNEXT package (Hsieh et al., 2016), both for 92% amino acid similarity representative sequences in the different sea-ice related environments of the Central Arctic, as well as for the 97% nucleotide identity representative sequences in the different oceanic regions. Hill number 0 refers to species richness, 1 to the number of ‘typical’ species in the community (exponential of Shannon entropy), and 2 to the number of very abundant species in a community (inverse Simpson).

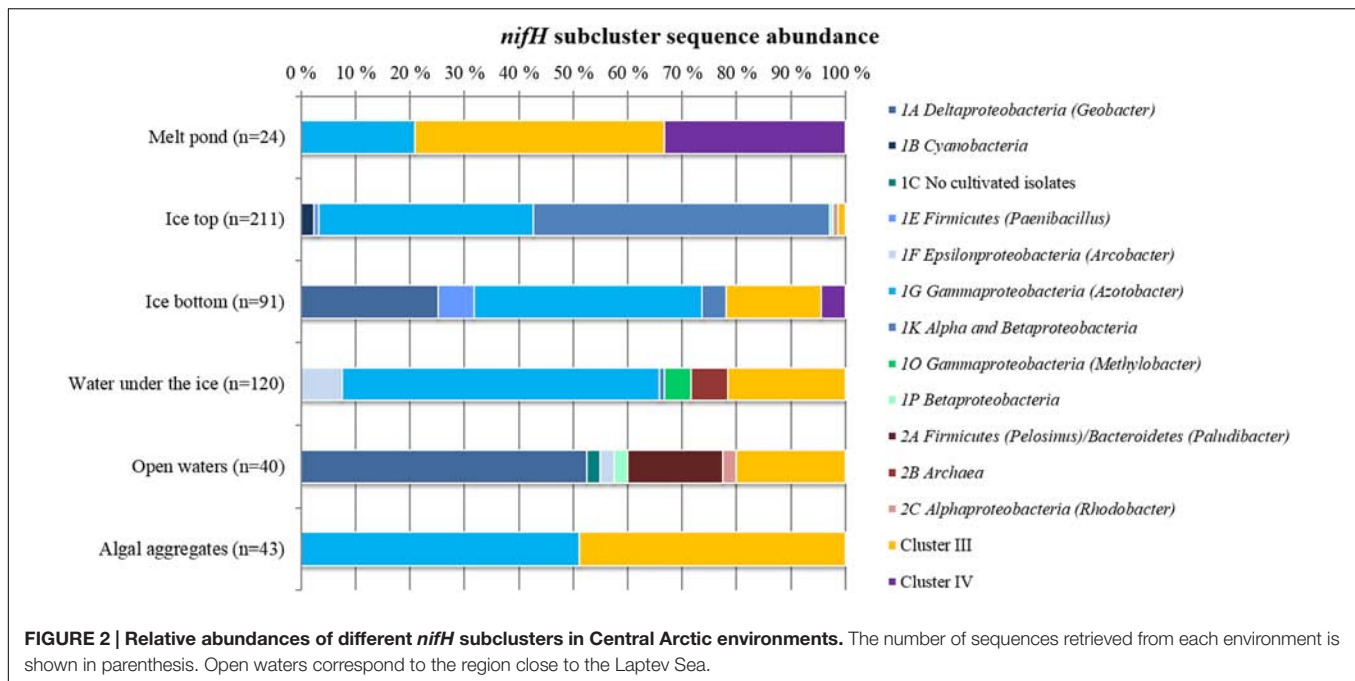
## RESULTS

### Characterization of Arctic Diazotroph Communities Based on *nifH* sequences

During our study in summer 2012, *nifH* gene fragments were amplified from 26 of the 53 samples collected in different Arctic environments: sea ice, melt ponds and surface water column (Figure 1). A total of 529 sequences were retrieved, 40% of them from the upper part of the sea ice and only 5% from melt ponds (Table 1; Supplementary Table S1). These sequences were clustered into 43 clusters at 92% amino acid similarity and were distributed across all four main *nifH* clusters I–IV as defined by Zehr et al. (2003), including both cyanobacterial and non-cyanobacterial phylotypes (Figures 2 and 3). Central Arctic sequences affiliated with Cluster I contained mainly *Proteobacteria* (1G and 1K), *Firmicutes* (1E), *Cyanobacteria* (1B) and several uncultivated microorganisms. Cluster II contained *Proteobacteria* (2C), *Firmicutes* (2A), and members of the Archaea (2B). Cluster III contained putative anaerobes including sulfate reducing genera of the *Deltaproteobacteria*, and genera such as *Clostridium*. Cluster IV contained *nifH* paralogs that are thought to function in metabolic processes other than nitrogen fixation (e.g., Young, 2005; Staples et al., 2007).

**TABLE 1 | Number of samples screened for *nifH* gene amplification and confirmed *nifH* gene sequences retrieved.**

Environment	Total Samples	Samples with <i>nifH</i>	Sequences retrieved
Melt ponds	8	3	24
Ice top	8	8	211
Ice bottom	13	4	91
Water under the ice	13	6	120
Surface open waters Laptev Sea	6	4	40
Algal aggregates	5	1	43
Total	53	26	529



The vast majority of sequences retrieved across the different environments of the Central Arctic Ocean belonged to non-cyanobacterial diazotrophs of Cluster I (Figure 2). In this cluster, 53% of the sequences belonged to the subcluster 1G that contains sequences from genera such as *Azotobacter*, *Brenneria*, *Teredinibacter*, and *Pseudomonas*. Subcluster 1K comprised 29% of the sequences, containing both *Alpha*- and *Beta*-proteobacteria such as *Bradyrhizobium* and *Azospirillum* sp. From subcluster 1B, which contains exclusively cyanobacterial diazotrophs, only five sequences were amplified from the upper layer of the sea ice (top 50 cm of a snow-free core) at ice station 224 (Subcluster 1B, Representative sequence N\_224IT\_002 in Figure 3). These sequences were all closely related ( $\geq 92\%$  amino acid sequence similarity) to *Nodularia*. Cluster III sequences accounted for up to 20% of the sequences retrieved by clone libraries from all samples together (Figure 2). Only 12 sequences corresponding to *nifH* paralogs (Cluster IV) were retrieved from sea-ice and melt pond samples, but they were not closely related to any cultivated organism (Figure 3). Since they do not provide any insight into nitrogen fixation capabilities of the Arctic microbial community, we will not develop them further.

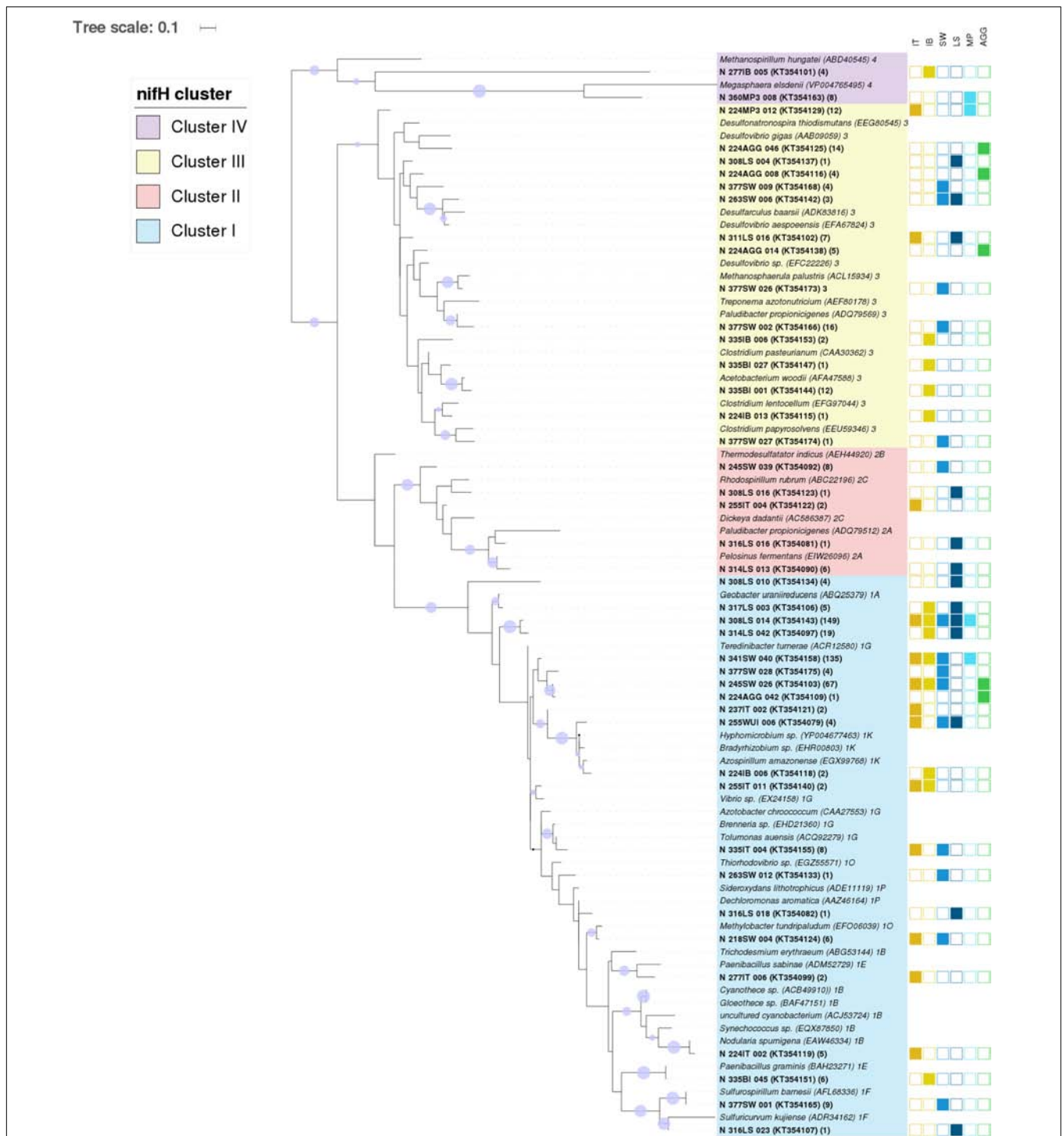
## Presence of *nifH* Genes in the Central Arctic and Link to Temperature and Nutrients

Two of the main environmental factors that might affect the presence of diazotrophs, temperature and nutrient concentrations, were quite variable in the environments and stations sampled. Seawater temperatures ranged from  $-1.7^{\circ}\text{C}$  below the ice and up to  $3^{\circ}\text{C}$  in open waters of the Laptev Sea shelf (Supplementary Figure S1). Sea-ice temperatures were between  $-0.2$  and  $-1.8^{\circ}\text{C}$  and melt pond temperatures between  $-0.7$

and  $0.3^{\circ}\text{C}$ . Nitrogen fixation genes from putative diazotrophs were amplified from samples along the entire temperature range sampled (Figure 1).

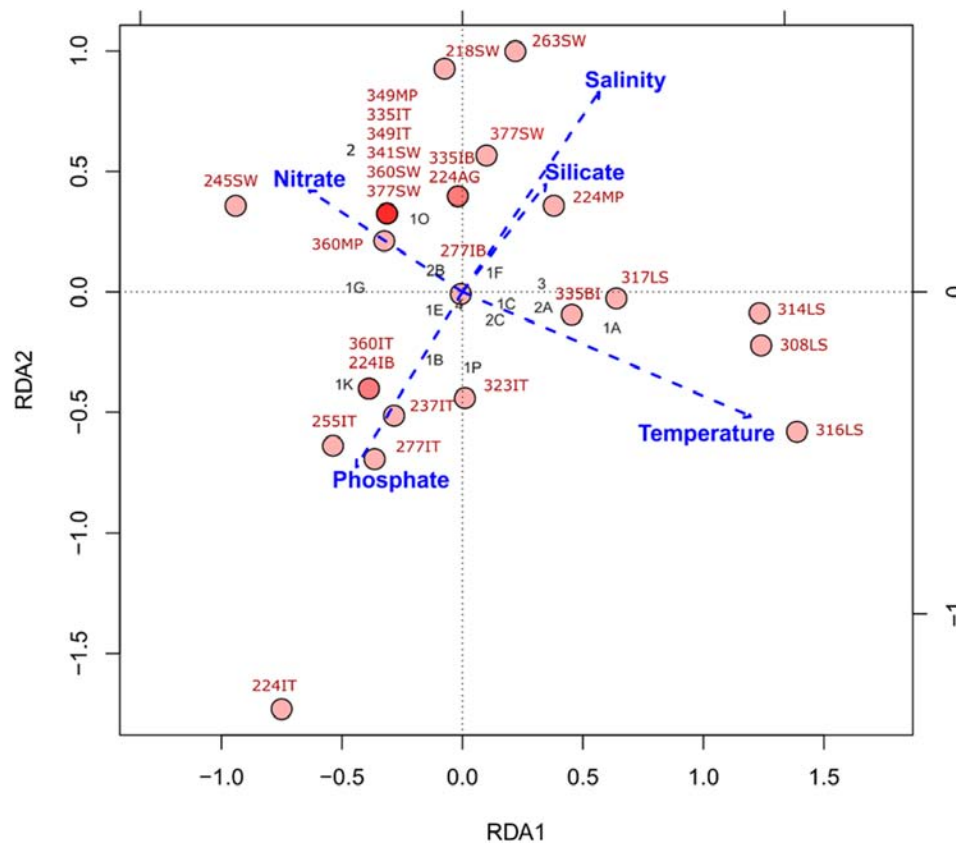
Regarding nutrients, nitrate concentrations in summer 2012 were very similar in sea ice and melt ponds ranging between 0.2 and  $1.5\ \mu\text{M}$ , while being more variable in surface waters ranging between 0.02 and  $6.4\ \mu\text{M}$  (with highest concentrations at stations 218 and 245). Phosphate, the most relevant nutrient for diazotrophs, was more variable in sea ice ( $0.02$ – $1.95\ \mu\text{M}$ ) reaching its highest concentrations at ice station 224. Melt ponds had in general very low phosphate concentrations ( $0.04$ – $0.2\ \mu\text{M}$ ) while in surface waters the concentrations were higher ( $0.15$ – $0.49\ \mu\text{M}$ ). As indicated by the N:P molar ratio, during summer in the Central Arctic, two nutrient regimes were identified in the euphotic zone of the water column (Supplementary Figure S2A). All N:P molar ratios in surface waters were below Redfield ( $<16:1$ ) indicating general nitrogen limitation. The Atlantic-influenced ice margin in the Nansen Basin and the Laptev Sea had N:P ratios around 10, while the more Central Arctic waters in the Amundsen Basin had N:P ratios below 5. In sea ice, nutrient concentrations were in general lower than in the water column and the integrated N:P ratios were more variable. N:P ratios at stations 335 and 349 were close to Redfield, while the rest ranged between 5 and 11 (Supplementary Figure S2B). In melt ponds the N:P ratio ranged between 1 and 16. The genes responsible for nitrogen fixation could be amplified in all environments regardless of their N:P ratio at the time of sampling.

Our RDA suggests that our current set of explanatory variables are only able to account for just over half ( $\sim 57\%$ ) of the variation in our *nifH* subcluster data (Figure 4). Salinity and silicate increasing concentrations showed strong positive covariation, while phosphate concentrations showed a negative covariation with the other variables (Figure 4). Nitrate concentrations and



**FIGURE 3 | Central Arctic *nifH* gene diversity in summer 2012.** Maximum likelihood phylogenetic tree of *nifH* partial amino acid sequences (104 positions) containing representative sequences from the Central Arctic (bold) and their closest cultivated relatives (italics). The Central Arctic *nifH* amino acid sequences were clustered using CD-HIT at 92% amino acid similarity to determine representative sequences. The origin of our Central Arctic representative sequences is indicated in the name. The number of sequences represented by each CD-HIT subcluster are indicated in brackets. The station number is followed by the environment from which it originates: Ice Top (IT), Ice Bottom (IB), Brown Ice (BI), Melt Pond (MP), Algal aggregate (AGG), Water under the ice (WUI), Surface water (SW), and Laptev Sea (LS). The *nifH* subcluster assignment for each cultivated diazotroph is indicated after the name. Branch lengths were determined using the LG protein substitution model as recommended by RAxML Version 8. The percentage of calculated trees in which sequences clustered together in the bootstrap test (1000 replicates) is shown next to the branch as a purple circle when greater than 50% (smallest circle corresponds to 50% and greater to 100%). GenBank accession numbers for the 43 representative sequences are in parenthesis after the sequence name.





**FIGURE 4 | Redundancy analysis (RDA) of *nifH* subcluster presence absence in Central Arctic samples and environmental variables.** Environmental variables: temperature, salinity, nitrate, phosphate and silicate were standardized by z-scoring prior to RDA. The red circles represent the different samples and the intensity of the color indicates if there is one (light red) or many (dark red) samples at that coordinate in the plot. Sample names in red correspond to those in Supplementary Table S1. The black labels correspond to the different subclusters of *nifH* (Cluster number 1, 2, 3, or 4; and the subcluster letter A-K). The explanatory variables constrained ~57% of the variance in the *nifH* subclusters across samples. The significance of this type I scaling was 0.001 and the residuals were mostly distributed around zero (Supplementary Figure S3).

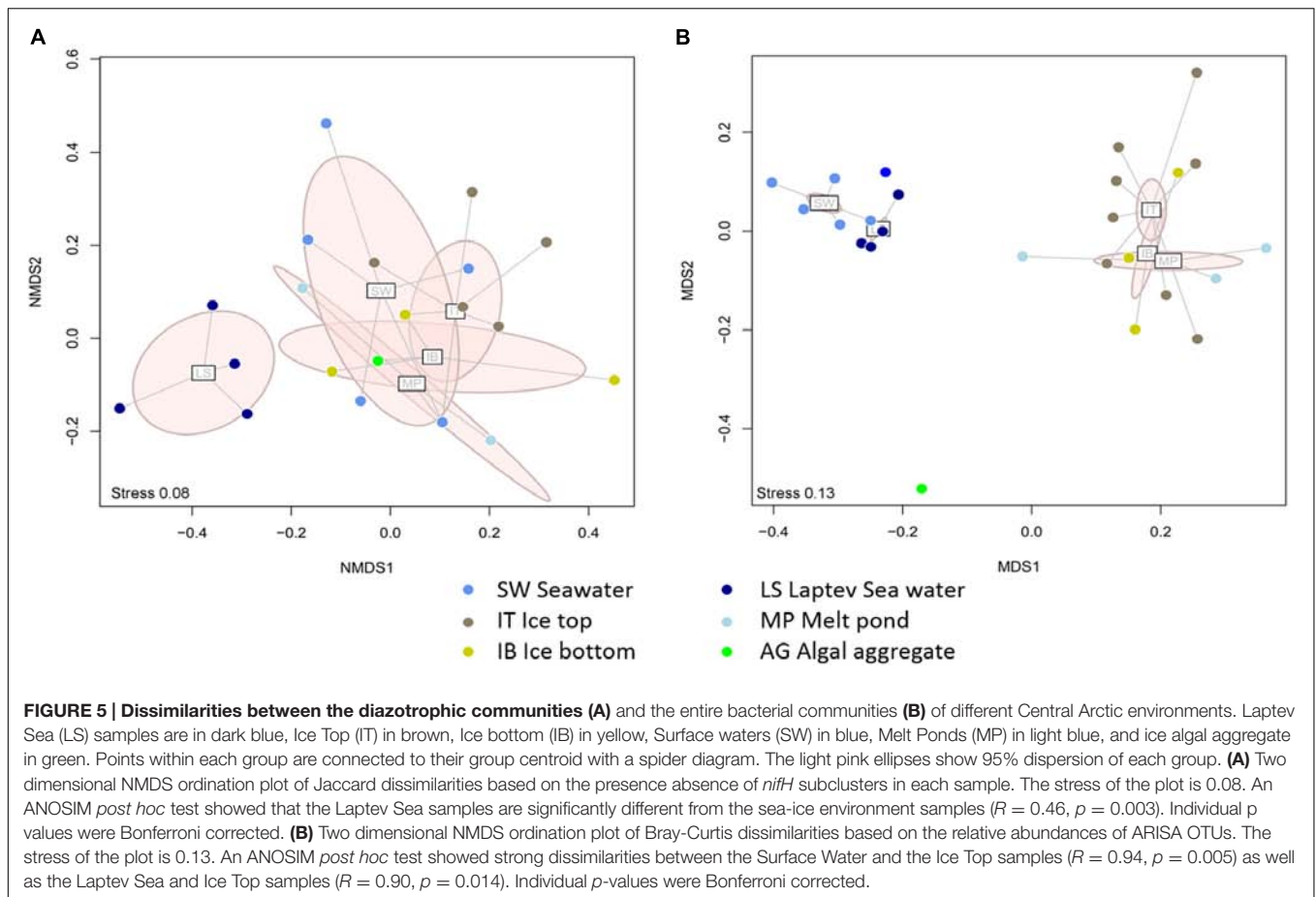
temperature showed strong negative covariation. In general, *nifH* genes are found across a wide range of temperature and nutrient values, but with changing community structure. For example, Laptev Sea samples, containing sequences from subclusters 1A (*Deltaproteobacteria*) and 2A (*Firmicutes and Bacteroidetes*), were more associated with increased temperatures and decreased nitrate concentrations relative to the other samples analyzed (Figure 4). The presence of cyanobacterial *nifH* genes at the upper layer of the ice at station 224 (labeled as 224IT in Figure 4), is associated with the high phosphate concentrations.

## Diazotroph and Total Microbial Community Diversity Patterns in Different Environments in the Central Arctic

Comparing the distribution of the *nifH* subclusters across the Central Arctic environments sampled, we observed differences between the Laptev Sea open waters and the sea-ice related environments. According to the dissimilarities between the *nifH* subclusters present in each sample, we detected a cluster of Laptev

Sea samples that was distinctly separate from samples from sea ice environments (Figure 5A) (ANOSIM  $R = 0.46$ ; *Post hoc* test  $p = 0.003$ ). The rarefaction curve for each environment (Supplementary Figure S4A) shows that the Laptev Sea is under sampled, but the abundant types that we captured (represented in the panels for Hill numbers 1 and 2) were very different from the others. In addition, the 16S rRNA analysis also shows that the total bacterial community of the Laptev Sea is different from the sea-ice related environments (Figure 5B). However, in this case, the Laptev Sea clusters together with the other surface water samples, which is not the case for the diazotrophic community. The ice top environment and the Laptev Sea open water communities (ANOSIM  $R = 0.90$ ;  $p = 0.014$ ) showed moderate differences, as well as the ice top and the ice-covered surface waters (ANOSIM  $R = 0.94$ ;  $p = 0.005$ ).

Sea ice and Laptev Sea open waters had the largest number of unique representative sequences. These two environments had four *nifH* subclusters in common (1A, 1K, 1P, and 2C) (Figure 2). Most of the recovered *nifH* sequences in the top half of the ice belonged to subclusters 1G and 1K, which contain *Gamma*- and *Alphaproteobacteria*, respectively. We detected the greatest



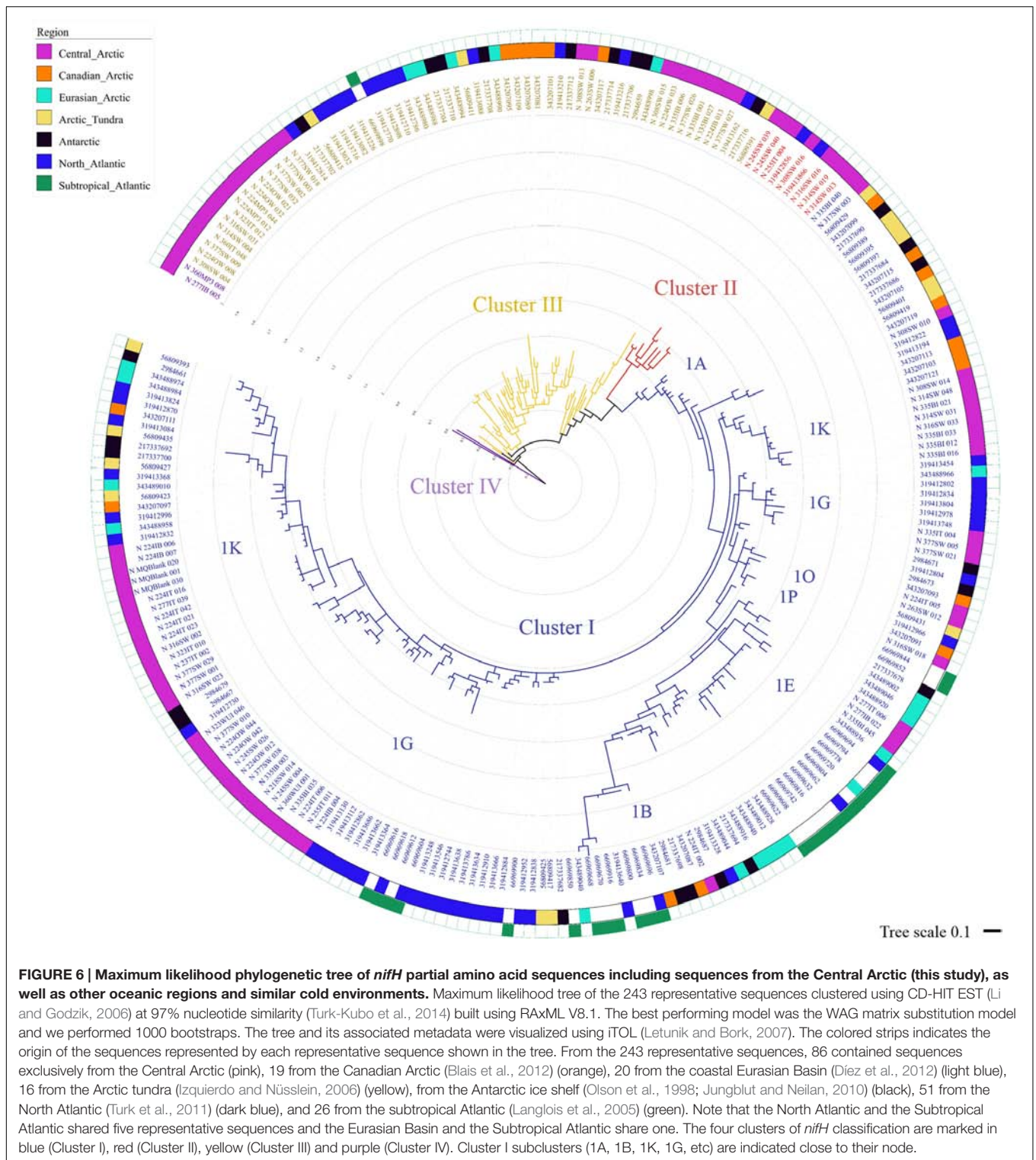
number of unique representative sequences in samples acquired from the lower portion of the ice, which contained sequences affiliated with subcluster 1A (*Deltaproteobacteria*) and Cluster III (Figure 2). Melt ponds, formed on top of the ice, only shared 1G subcluster sequences with the upper part of the ice, and additionally contained sequences from Cluster III (Figure 2). We detected sequences from subcluster 1G and from Cluster III – also found in our melt pond samples – in our samples of algal aggregates. These aggregates were primarily composed of algal species associated with sea ice. However, at a higher phylogenetic resolution, sequences in Cluster III differed between both environments (Figure 3).

More than half of the *nifH* sequences retrieved from water collected below the ice clustered together with sequences from the bottom part of the ice, an anticipated result because these two environments are in constant connection. However, the rest of the sequences detected in our water samples belonged to subclusters not detected in ice or melt pond samples: 1F (*Epsilonproteobacteria*), 1O (*Gammaproteobacteria*) and 2B (*Archaea*). The 1F subcluster also appeared in the open water samples from the Laptev Sea region (latitudes 77–79°N). The major subclusters present in our Laptev Sea open water samples were 1A and 2A. Subcluster 1A contains non-sulfate reducing *Deltaproteobacteria* such as *Geobacter* sp. and subcluster 2A contains fermenting bacteria from the genus *Pelosinus* and

*Paludibacter*. *Deltaproteobacteria* from Cluster III composed around 20% of the sequences from open waters (Figure 2).

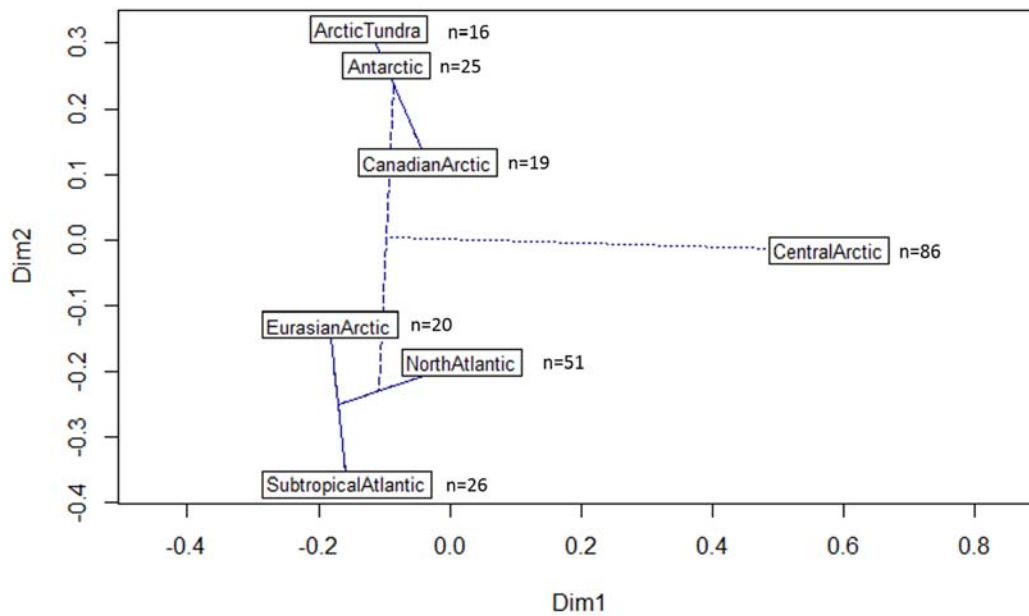
## Comparison of Central Arctic Diazotrophic Communities across Adjacent Polar Environments and Oceanic Regions

The diazotrophic community of the Central Arctic is highly dissimilar to all other regions and environments we compared it to as unweighted unifracs distances were the highest (0.85–0.78), when the other regions were between (0.75–0.58) (Supplementary Table S2). At 97% nucleotide identity almost no sequences from different oceanic regions clustered together (Figure 6). Only five phylotypes were shared between the North Atlantic and the Subtropical Atlantic (Figure 6). The *nifH* diversity in sea ice, melt ponds and surface waters of the Central Arctic is different from the coastal Canadian Arctic marine diazotrophic community (Mackenzie river, Baffin bay) (Figure 7). Furthermore, it also differs from the sea ice, snow and water column communities from the coastal Eurasian Arctic (close to Svalbard) (Figure 7). The rarefaction curves show that the diversity in regions from which we acquired a greater number of sequences (i.e., the Central Arctic, the North Atlantic and the Subtropical Atlantic) appears to have been



reasonably well sampled (Supplementary Figure S4B). However, the other oceanic regions are still in the rapid accumulation phase of their respective curves and deeper sequencing is likely to strongly affect their diversity profiles (Supplementary Figure S4B). We therefore place more interpretive weight on

dissimilarities between our well-sampled regions (marked in the dissimilarities Supplementary Table S2), while noting that dissimilarities bearing on the Arctic tundra, Canadian Arctic, Eurasian Basin and Antarctic might be subject to change when the sampling effort is increased.



**FIGURE 7 | Dissimilarities between different oceanic regions based on their *nifH* phylogenetic signature.** Plot of the first two principal coordinate axes of the Principal coordinates analysis (PCoA) derived from the Unifrac distances between different oceanic regions included in the maximum likelihood tree (Figure 6). The further away the samples are, the more dissimilar they are in terms of *nifH* phylogenetic signature. An automatic correction for negative eigenvalues was used. A cluster dendrogram is superimposed. The lines indicate the strength of the dissimilarity. Oceanic regions connected with a solid line have a dissimilarity less than 0.69 (Supplementary Table S2). The dashed line indicates a dissimilarity of 0.76 and the dotted line between the Central Arctic and the subtropical Atlantic indicates the maximum dissimilarity of 0.85 (Supplementary Table S2). The number of representative sequences (97% nucleotide identity) from each oceanic region included in the analysis is indicated next to the region's name.

## DISCUSSION

### Distribution of Diazotrophs in the Central Arctic Ocean

Confirming our first hypothesis, we found putative diazotrophs in the nitrogen-limited sea ice and waters of the Central Arctic Ocean (77–89°N). Most of the *nifH* sequences retrieved were non-cyanobacterial diazotrophs, mainly distributed between Cluster I and III as the bacterial phylotypes described in the marine Arctic studies by Farnelid et al. (2011), Blais et al. (2012), and Díez et al. (2012). All prominent nitrogen-fixing taxa at the subclade level present in other oceans were also present in the Central Arctic Ocean. However, most of the *nifH* genes amplified in the Central Arctic belonged to non-cyanobacterial diazotrophs, while cyanobacteria were under-represented. Previous Arctic diazotroph diversity studies also retrieved only a few cyanobacterial phylotypes. *Cyanothece* sp. and *Trichodesmium* sp., were found in the coastal Eurasian Basin in early May (Díez et al., 2012) and *Nostocales* sp., was found in the Beaufort Sea in July–August (Blais et al., 2012). The fact that we recovered only a single cyanobacterial phylotype in late summer, *Nodularia* sp., a cyanobacterial genus that contributes to nitrogen fixation in the Baltic Sea (Bostrom et al., 2007), is consistent with the results of Farnelid et al. (2011), who observed that their Arctic station sampled in July (in Baffin Bay) only yielded non-cyanobacterial diazotrophs from Cluster III. This implies that cyanobacterial diazotrophs may not thrive in

Central Arctic waters late in the productive season, far away from the coast line where there is riverine influence. However, *nifH* clone-library based approaches need to be interpreted with caution due to possible primer bias toward certain non-cyanobacterial diazotrophs (Turk et al., 2011; Turk-Kubo et al., 2015). Furthermore, the use of cyanobacteria-selective primers by Díez et al. (2012) might explain the high cyanobacterial diazotroph diversity reported in their study, compared to others.

The non-cyanobacterial diazotrophic community inhabiting the water column in the Central Arctic Ocean was dominated by Cluster I sequences: mainly *Alphaproteobacteria* (subcluster 1K) and *Deltaproteobacteria* (subcluster 1G) (Figure 2). While *Alpha*- and *Gammaproteobacteria* dominated non-cyanobacterial diazotrophic assemblages in the Atlantic and Pacific Oceans (Falcón et al., 2004; Langlois et al., 2005) and were present in lower abundances in the water column above the Canadian Arctic shelves (Blais et al., 2012). Sequences related to anaerobic bacteria (Cluster III) made up to 20% of under ice and open water communities in the Central Arctic (Figure 2). Cluster III sequences were once thought to be rare in brackish or marine surface waters (Moisander et al., 2007), but they are increasingly being recovered from ocean waters (Langlois et al., 2008; Turk-Kubo et al., 2014) including the Canadian Arctic shelf region (Farnelid et al., 2011).

Among the sea-ice related environments studied, floating aggregates are conglomerations of sea ice diatoms sticking together in a mucous matrix that can have an anoxic interior

and, therefore, have the potential to provide an environment for anaerobic processes such as denitrification and nitrogen fixation (Fernández-Méndez et al., 2014b; Lehto et al., 2014). Therefore, it is important to study *nifH* diversity in this environment. It is difficult to infer phylogeny from Cluster III sequences; however, some of our sequences were closely related to cultivated *Deltaproteobacteria* (Figure 3). For example, nine sequences retrieved from the floating algal aggregate (AGG) were related (88% aa similarity) to the genus *Desulfovibrio*, a common genus in marine coastal sediments and anoxic zones (Herbert, 1999). These aggregates form in late summer below melting sea-ice (Assmy et al., 2013; Fernández-Méndez et al., 2014b) and are not connected with coastal areas at the time of formation. This could indicate that their anaerobic bacterial community does not originate from coastal sediments.

### Impact of Temperature and Nutrient Concentrations on Arctic Diazotrophs

The low temperatures and wide range of salinities present in sea ice environments shape their bacterial communities, leading to class-level differences between sea ice and water column communities (Boetius et al., 2015). To further test our first hypothesis regarding the presence of diazotrophs in the Central Arctic, we explored if the distribution and diversity of cyanobacterial and non-cyanobacterial diazotrophs was restricted by low temperatures or a specific nutrient ratio. During our cruise in August–September 2012, which covered a temperature gradient ranging from  $-1.7$  to  $3^{\circ}\text{C}$ , nitrogen fixation genes from putative non-cyanobacterial diazotrophs were consistently amplified. This indicates that temperature limitation did not exclude non-cyanobacterial diazotrophs, an observation in agreement with previous studies (Farnelid et al., 2011; Blais et al., 2012; Díez et al., 2012). Non-cyanobacterial diazotrophs are known to thrive at low water temperatures (Riemann et al., 2010) and have been reported to be relevant for nitrogen fixation in other warmer oligotrophic oceans (Bombar et al., 2013; Moisaner et al., 2014) and in the Baltic Sea (Farnelid et al., 2013). Increasing temperature seems to be linked to the structure of diazotroph community composition in the Laptev Sea samples (Figure 4) and could imply that a shift in surface water temperatures in the Arctic Ocean may have corresponding consequences on the regional diazotroph community. For example, non-cyanobacterial *nifH* subclusters 1A (*Geobacter*) and 2A (*Pelosinus* and *Paludibacter*) seem to be linked to increased temperatures (Figure 4). However, we are aware of the instability of our RDA solution due to the low number of sequences in our dataset and it is possible that deeper and more even sequencing will reveal somewhat different relationships.

The temperature ranges and optima of nitrogenase activities are only known for a few nitrogen-fixing cyanobacteria, such as *Trichodesmium* (Breitbarth et al., 2006) and *Cyanothece* (Brauer et al., 2013), which grow better at warm temperatures. Putative nitrogen-fixing cyanobacteria were only amplified in one sample in the Central Arctic Ocean (upper part of the ice at  $-0.2^{\circ}\text{C}$ ), supporting previous hypotheses that this group of nitrogen-fixers has not realized a niche in ice-covered polar open

oceans (Murphy and Haugen, 1985; Koh et al., 2012b). However, other nitrogen-fixing cyanobacteria have been detected in snow (Harding et al., 2011; Boetius et al., 2015), glacial environments (Yallop et al., 2012; Vonnahme et al., 2015), hydrothermal vents (Mehta et al., 2003), and in other cold environments, such as Antarctic lakes (Olson et al., 1998) and sea ice (Koh et al., 2012a). Therefore, it remains unclear why they have not populated nitrogen-limited marine Arctic waters. The absence of cyanobacterial sequences in most of our samples might be due to undersampling (Supplementary Figure S4A), or to the primer bias of the primers we used toward non-cyanobacterial phylotypes (Turk-Kubo et al., 2015). However, the fact that cyanobacteria are rare in waters of the Arctic Ocean close to the ice (Lovejoy et al., 2002; Sherr et al., 2003) supports our findings.

Besides temperature, nutrient availability also plays a role in the occurrence of diazotrophs. The N:P ratios in surface waters of the Eurasian Basin range between 11 and 16 (Sakshaug et al., 2004). However, in summer the ratio can drop to 2–10, particularly in the Central Arctic north of  $80^{\circ}\text{N}$  (Supplementary Figure S2A), where nitrate availability limits algal growth (Fernández-Méndez et al., 2015). Silicate, which can also be limiting for diatom growth at the end of the productive season, showed no direct impact on *nifH* diversity. The *nifH* profile of the sample where we found cyanobacterial diazotrophs was associated with high phosphate concentrations supporting the low N:P hypothesis for autotrophic diazotrophs. In Central Arctic sea ice the occurrence of nitrogen-fixing cyanobacteria, such as *Nodularia* sp., seems to be driven by phosphate concentrations. However, as discussed previously this is a preliminary conclusion based on the very scarce data available at the moment.

We could retrieve *nifH* genes mostly of non-cyanobacterial putative diazotrophs from environments with a wide range of physical and chemical conditions, including summer sea ice with high N:P molar ratio ( $>16$ ) and low salinities (0.5–3.6), and surface seawaters with low temperatures ( $-1.7^{\circ}\text{C}$ ) and higher salinity (33). Further research is required to determine whether any of the putative diazotrophs detected in this study are actively transcribing the *nif* operon and fixing nitrogen in the water, sea ice, and melt ponds of the Eurasian Basin.

### The Origin of Marine Diazotrophs in the Central Arctic

Our second hypothesis was that putative diazotrophs in Central Arctic sea ice would come from the coastal areas influenced by rivers, such as the Lena River. The Lena River is a source of organic matter and iron that could provide favorable conditions for nitrogen fixation in the Laptev Sea (Lara et al., 1998). Our results, however, show a distinct entire bacterial community and diazotrophic subcommunity in the Laptev Sea region, close to the Lena Delta, when compared to that present in sea ice communities (Figure 5).

The ice floes sampled at the end of the productive season were formed during freeze-up or in polynyas located along the fast ice edge in the Laptev and Kara seas (Supplementary Figure S5). This, together with winds that might transport dust and microorganisms offshore (Harding et al., 2011), might explain the high relative abundance of *nifH* sequences related

to *Bradyrhizobium* sp. (*Alphaproteobacteria* 1K subcluster). *Bradyrhizobium* sp. is a symbiotic soil bacterium present in the nodules of leguminous plants (Hennecke, 1990), that can also fix nitrogen as free-living bacteria (Dreyfus et al., 1988), and is likely to be transported by rivers or wind. Indeed it has also been found in surface waters of the Mediterranean Sea (Le Moal et al., 2011). Some sea ice sequences clustered with the endospore-forming *Paenibacillus* (subcluster 1E) that has also been found in high Arctic soils (Jordan et al., 1978). In both environments, polar soils and sea ice, high salinities can be reached in pore fluids (Tedrow, 1966; Ewert and Deming, 2014), favoring halotolerant diazotrophs. Therefore, halotolerant diazotrophs can be expected. However, the riverine adapted freshwater diazotrophic bacteria are unlikely to thrive in the marine environment partially due to its high salinity (Fernandes et al., 1993). Subcluster 1F, which was present in both waters under the ice and those of the Laptev Sea, includes diazotrophs from the genus *Arcobacter* which is known to occur in roots of salt marsh plants, but also in sulfidic environments such as cold seeps and hydrothermal vents (Mehta et al., 2003). The dispersal of this genus from the soil of Arctic coastal areas to the Central Arctic is possible although other mechanisms might occur.

Overall, our assays of microbial communities based on the 16S rRNA gene and the *nifH* gene showed little support for the notion of the riverine origins of Central Arctic diazotrophs: the most abundant types in the bacterial and diazotrophic communities of the Laptev Sea water were notably dissimilar to those of Central Arctic sea ice and those present in the water under the ice. Sequences retrieved from brown ice (due to a high a concentration of phaeopigments from diatoms and possibly some sediments) are an exception: these clustered with sequences from the Laptev Sea waters. The subcluster represented by sequence 317LS\_003 from the Laptev Sea in **Figure 3** includes sequences from sample 335BI (brown ice from station 335). Hence, a coastal (land or riverine) origin of the putative diazotrophs and other microbes in the Central Arctic is possible, but our data indicates that it is not likely the main process for diazotroph dispersal in the Eurasian Basin.

## The Distinct Diazotroph Community of the Central Arctic Ocean

To assess potential causes of dissimilarity on the basis of the limited data set available, we grouped all available sequences from the Central Arctic, coming from different environments, and compared them to sequences from other oceanic regions and polar environments such as microbial mats in Antarctic lakes or Arctic tundra soils (**Figure 6**). This comparison assumes that the variability between the different environmental types in one region has a smaller effect than the variability between regions in the phylogenetic signal of *nifH* caused by the environment type.

The Eurasian Arctic communities were more similar to the North Atlantic and Subtropical Atlantic communities than to the Canadian Arctic (**Figure 7**). This is to be expected since these oceanic regions are tightly connected through the Fram Strait. The Arctic tundra soil communities were similar to communities isolated from Antarctic lake microbial mats at

the *nifH* subcluster level. These Antarctic microbial mats were described as submerged in sediments of a meltwater pond in the McMurdo ice shelf (Jungblut and Neilan, 2010), or embedded in the permanent ice cover of Lake Bonney in the Antarctic McMurdo Dry Valley in summer (Olson et al., 1998). This indicates that soil-related communities have higher similarity to one another than to marine environments with similar temperature conditions. However, the dissimilarities between the Arctic tundra, the Antarctic and the Canadian Arctic must be interpreted with caution since the *nifH* diversity in these regions is under sampled (Supplementary Figure S4B).

Despite the limitations of this preliminary comparative study due to limited data, the Central Arctic diazotrophic community appears to have a different phylogenetic composition from adjacent regions and similar cold environments (**Figure 7**). Furthermore, our *nifH* gene analysis and the ITS region analysis (**Figures 5A,B**) indicate that Central Arctic sea ice environments harbor a microbial community distinct from that of the water column. We also noted that the marine diazotrophic communities in the Arctic were distinct from the terrestrial diazotrophic communities of the Arctic tundra (Izquierdo and Nüsslein, 2006) and the Antarctic ice shelf (Olson et al., 1998; Jungblut and Neilan, 2010). This bacterial Central Arctic endemism seems to be also reflected in the eukaryotic composition after the polar night when no phototrophs are dominating (Bachy et al., 2011). The diazotrophic communities of the Eurasian Arctic north of Svalbard (Díez et al., 2012), the North Atlantic (Turk et al., 2011) and the Subtropical Atlantic (Langlois et al., 2005) appeared similar to one another. This similarity is probably due to the presence of shared phylotypes (5 CD-HIT clusters 97% nucleotide identity contained sequences from the North Atlantic and the Subtropical Atlantic) and of cyanobacterial diazotrophs in these regions (**Figure 6**). One should note that Díez et al. (2012) used slightly different primers from all other studies and this might be partially responsible for the dissimilarities between the Eurasian Basin and the other regions. As more *nifH* sequences from different Arctic environments and regions will be added to the *nifH* database, more robust conclusions can be drawn from comparisons of diazotrophic community composition among different environments. For example, although cyanobacteria have been found thriving on glaciers and snow (Harding et al., 2011; Yallop et al., 2012) no analyses of *nifH* diversity were performed during these studies. Nevertheless, our initial comparisons, which included all currently available diazotroph sequences in the Arctic, strongly suggest that the Central Arctic Ocean harbors a community of putative diazotrophic microorganisms, which is distinct to those present in other oceanic regions and similarly cold environments.

## CONCLUSION AND OUTLOOK

This study reveals the potential for nitrogen fixation far away from the coastal shelves in the Central Arctic, where diazotrophs were previously not detected. Nitrogen availability in the euphotic zone constrains the potential increase of

primary production especially in the central oligotrophic basins (Tremblay et al., 2015). Hence it is relevant to better understand the role of nitrogen fixers. Most of the *nifH* sequences retrieved belonged to putative non-cyanobacterial diazotrophs from Clusters I and III while diazotrophic cyanobacteria were rare. The origin of the sea-ice diazotrophic community in the Central Arctic seems to be marine and not riverine, since we found significant differences between the river-influenced Laptev Sea communities, and the sea-ice related diazotrophic communities. Assuming that the upper limit of nitrogen fixation estimated by Blais et al. (2012) in the Canadian Arctic (up to  $0.14 \text{ nmol N L}^{-1} \text{ d}^{-1}$  transformed to carbon using a C:N ratio of 7.3:  $0.6 \text{ mg C m}^{-2} \text{ d}^{-1}$ ) would also take place in the Central Arctic, the nitrogen fixed by the diazotrophs could sustain  $\sim 7\%$  of the new primary production in the region ( $9.4 \pm 3.6 \text{ g C m}^{-2} \text{ year}^{-1}$ , Fernández-Méndez et al., 2015). However, the presence and diversity of diazotrophs does not imply high nitrogen fixation rates (Moisander et al., 2007). With current trends of warming in Arctic waters (Polyakov et al., 2010; IPCC, 2013), higher marine nitrogen fixation rates can be expected in summer ice-free waters (Blais et al., 2012). However, the real contribution of these diazotrophs to nitrogen fixation still needs to be assessed before any further conclusions can be drawn regarding their role in the ecosystem and relevance for Arctic primary production.

## AUTHOR CONTRIBUTIONS

MF-M and AB collected the samples. MF-M, KT-K, and JR performed the laboratory molecular analysis and the phylogenetic bioinformatic analysis. MF-M, JR, and PB performed the statistical analysis. TK provided the sea ice physics background information and the ice drift analysis. AB and JZ designed the study, organized the analysis and contributed to the

discussion of the results. MF-M prepared the manuscript with contributions from all-coauthors.

## FUNDING

This study was supported by the Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung and the Max Planck Society, as well as the ERC Advanced Grant Abyss (no.294757) to AB. PB's work on this project is supported through the Micro B3 project, funded by the European Union's Seventh Framework Programme (Joint Call OCEAN.2011-2: marine microbial diversity – new insights into marine ecosystems functioning and its biotechnological potential) under the grant agreement no 287589.

## ACKNOWLEDGMENTS

We thank the captain and crew of the RV Polarstern for their support during the ARK XXVII/3 expedition. We are particularly thankful to Anique Stecher, Christiane Uhlig, Ben Lange, Heidi L. Sørensen, Ilka Peeken, and Hauke Flores for their help during the sampling. The technical help of Mary Hogan and Brandon Carter at UCSC is greatly appreciated. We also thank Antonio Fernández-Guerra and the bioinformatics group at the Max Planck Institute for their support and advice.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2016.01884/full#supplementary-material>

## REFERENCES

- Arrigo, K. R. (2005). Marine microorganisms and global nutrient cycles. *Nature* 437, 349–356. doi: 10.1038/nature04158
- Arrigo, K. R., and van Dijken, G. L. (2015). Continued increases in Arctic Ocean primary production. *Prog. Oceanogr.* 136, 60–70. doi: 10.1016/j.pocean.2015.05.002
- Assmy, P., Ehn, J. K., Fernández-Méndez, M., Hop, H., Katlein, C., Sundfjord, A., et al. (2013). Floating ice-algal aggregates below melting arctic sea ice. *PLoS ONE* 8:e76599. doi: 10.1371/journal.pone.0076599
- Bachy, C., López-García, P., Vereshchaka, A., and Moreira, D. (2011). Diversity and vertical distribution of microbial eukaryotes in the snow, sea ice and seawater near the north pole at the end of the polar night. *Front. Microbiol.* 2:106. doi: 10.3389/fmicb.2011.00106
- Bates, N. R., Garley, R., Frey, K. E., Shake, K. L., and Mathis, J. T. (2014). Sea-ice melt CO<sub>2</sub>-carbonate chemistry in the western Arctic Ocean: meltwater contributions to air-sea CO<sub>2</sub> gas exchange, mixed layer properties and rates of net community production under sea ice. *Biogeosci. Dis.* 11, 1097–1145. doi: 10.5194/bgd-11-1097-2014
- Beine, H. J., Domine, F., Ianniello, A., Nardino, M., Allegrini, I., Teinilä, K., et al. (2003). Fluxes of nitrates between snow surfaces and the atmosphere in the European high Arctic. *Atmos. Chem. Phys.* 3, 335–346. doi: 10.5194/acp-3-335-2003
- Blais, M., Tremblay, J.-É., Jungblut, A. D., Gagnon, J., Martin, J., Thaler, M., et al. (2012). Nitrogen fixation and identification of potential diazotrophs in the Canadian Arctic. *Glob. Biogeochem. Cycles* 26:GB3022. doi: 10.1029/2011GB004096
- Boetius, A., Anesio, A. M., Deming, J. W., Mikucki, J., and Rapp, J. Z. (2015). Microbial ecology of the cryosphere: sea ice and glacial habitats. *Nat. Rev. Microbiol.* 13, 677–690. doi: 10.1038/nrmicro3522
- Bombar, D., Turk-Kubo, K. A., Robidart, J., Carter, B. J., and Zehr, J. P. (2013). Non-cyanobacterial *nifH* phylotypes in the North Pacific Subtropical Gyre detected by flow-cytometry cell sorting. *Environ. Microbiol. Rep.* 5, 705–715. doi: 10.1111/1758-2229.12070
- Bordeleau, L. M., and Prévost, D. (1994). "Nodulation and nitrogen fixation in extreme environments," in *Symbiotic Nitrogen Fixation*, eds P. H. Graham, M. J. Sadowsky, and C. P. Vance (Dordrecht: Springer), 115–125. doi: 10.1007/978-94-011-1088-4\_13
- Bostrom, K. H., Riemann, L., Zweifel, U. L., and Hagstrom, A. (2007). *Nodularia* sp. *nifH* gene transcripts in the Baltic Sea proper. *J. Plankton Res.* 29, 391–399. doi: 10.1093/plankt/fbm019
- Bowman, J. S., Larose, C., Vogel, T. M., and Deming, J. W. (2013). Selective occurrence of Rhizobiales in frost flowers on the surface of young sea ice near Barrow, Alaska and distribution in the polar marine rare biosphere. *Environ. Microbiol. Rep.* 5, 575–582. doi: 10.1111/1758-2229.12047
- Brauer, V. S., Stomp, M., Rosso, C., Van Beusekom, S. A. M., Emmerich, B., Stal, L. J., et al. (2013). Low temperature delays timing and enhances the cost of nitrogen fixation in the unicellular cyanobacterium *Cyanothece*. *ISME J.* 7, 2105–2115. doi: 10.1038/ismej.2013.103

- Breitbarth, E., Oschlies, A., and LaRoche, J. (2006). Physiological constraints on the global distribution of *Trichodesmium*-effect of temperature on diazotrophy. *Biogeosci. Dis.* 3, 779–801. doi: 10.5194/bg-3-779-2006
- Chao, A., Gotelli, N. J., Hsieh, T. C., Sander, E. L., Ma, K. H., Colwell, R. K., et al. (2014). Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecol. Monogr.* 84, 45–67. doi: 10.1890/13-0133.1
- Codispoti, L. A., Brandes, J. A., Christensen, J. P., Devol, A. H., Naqvi, S. W. A., Paerl, H. W., et al. (2001). The oceanic fixed nitrogen and nitrous oxide budgets: moving targets as we enter the anthropocene? *Sci. Mar.* 65, 85–105. doi: 10.3989/scimar.2001.65s285
- Codispoti, L. A. A., Kelly, V., Thessen, A., Matrai, P., Suttles, S., Hill, V., et al. (2013). Synthesis of primary production in the Arctic Ocean: III. Nitrate and phosphate based estimates of net community production. *Prog. Oceanogr.* 110, 126–150. doi: 10.1016/j.pocean.2012.11.006
- Devol, A., Codispoti, L. A., and Christensen, J. (1997). Summer and winter denitrification rates in western Arctic Shelf sediments. *Cont. Shelf Res.* 17, 1029–1050. doi: 10.1016/S0278-4343(97)00003-4
- Diez, B., Bergman, B., Pedrós-Alió, C., Antó, M., and Snoeijis, P. (2012). High cyanobacterial *nifH* gene diversity in Arctic seawater and sea ice brine. *Environ. Microbiol. Rep.* 4, 360–366. doi: 10.1111/j.1758-2229.2012.00343.x
- Dreyfus, B., Garcia, J. L., and Gillis, M. (1988). Characterization of *Azorhizobium caulinodans* gen. nov., sp. nov., a stem-nodulating nitrogen-fixing bacterium isolated from *Sesbania rostrata*. *Int. J. Syst. Bacteriol.* 38, 89–98. doi: 10.1099/00207713-38-1-89
- Ewert, M., and Deming, J. W. (2014). Bacterial responses to fluctuations and extremes in temperature and brine salinity at the surface of Arctic winter sea ice. *FEMS Microbiol. Ecol.* 89, 476–489. doi: 10.1111/1574-6941.12363
- Falcón, L. L., Carpenter, E. J., Cipriano, F., Bergman, B., and Capone, D. G. (2004). N<sub>2</sub> fixation by unicellular bacterioplankton from the Atlantic and Pacific Oceans: Phylogeny and in situ rates. *Appl. Environ. Microbiol.* 70, 765–770. doi: 10.1128/AEM.70.2.765
- Farnelid, H., Andersson, A. F., Bertilsson, S., Al-Soud, W. A., Hansen, L. H., Sørensen, S., et al. (2011). Nitrogenase gene amplicons from global marine surface waters are dominated by genes of non-cyanobacteria. *PLoS ONE* 6:e19223. doi: 10.1371/journal.pone.0019223
- Farnelid, H., Bentzon-Tilia, M., Andersson, A. F., Bertilsson, S., Jost, G., Labrenz, M., et al. (2013). Active nitrogen-fixing heterotrophic bacteria at and below the chemocline of the central Baltic Sea. *ISME J.* 7, 1413–1423. doi: 10.1038/ismej.2013.26
- Farnelid, H., Oberg, T., and Riemann, L. (2009). Identity and dynamics of putative N<sub>2</sub>-fixing picoplankton in the Baltic Sea proper suggest complex patterns of regulation. *Environ. Microbiol. Rep.* 1, 145–154. doi: 10.1111/j.1758-2229.2009.00021.x
- Fernandes, T. A., Iyer, V., and Apte, S. K. (1993). Differential responses of nitrogen-fixing cyanobacteria to salinity and osmotic stresses. *Appl. Environ. Microbiol.* 59, 899–904.
- Fernández-Méndez, M., Katlein, C., Rabe, B., Nicolaus, M., Peeken, I., Bakker, K., et al. (2015). Photosynthetic production in the central Arctic Ocean during the record sea-ice minimum in 2012. *Biogeosciences* 12, 3525–3549. doi: 10.5194/bg-12-3525-2015
- Fernández-Méndez, M., Peeken, I., and Boetius, A. (2014a). *Integrated Nutrients Concentrations Calculated for CTD Water Samples During POLARSTERN Cruise ARK-XXVII/3 (IceArc) in 2012*. Colaba Mumbai: PANGAEA. doi: 10.1594/PANGAEA.834296
- Fernández-Méndez, M., Wenzhöfer, F., Peeken, I., Sørensen, H. L., Glud, R. N., and Boetius, A. (2014b). Composition, buoyancy regulation and fate of ice algal aggregates in the Central Arctic Ocean. *PLoS ONE* 9:e107452. doi: 10.1371/journal.pone.0107452
- Finn, R., Mistry, J., Tate, J., Coggill, P., Heger, A., and Je, P. (2010). The Pfam protein families database. *Nucleic Acids Res.* 38, D211–D222. doi: 10.1093/nar/gkp985
- Fisher, M. M., and Triplett, E. W. (1999). Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to freshwater bacterial communities. *Appl. Environ. Microbiol.* 65, 4630–4636.
- Foster, R. A., Kuypers, M. M. M., Vagner, T., Paerl, R. W., Musat, N., and Zehr, J. P. (2011). Nitrogen fixation and transfer in open ocean diatom-cyanobacterial symbioses. *ISME J.* 5, 1484–1493. doi: 10.1038/ismej.2011.26
- Gaby, J. C., and Buckley, D. H. (2012). A comprehensive evaluation of PCR primers to amplify the *nifH* gene of nitrogenase. *PLoS ONE* 7:e42149. doi: 10.1371/journal.pone.0042149
- Gobet, A., Boetius, A., and Ramette, A. (2014). Ecological coherence of diversity patterns derived from classical fingerprinting and Next Generation Sequencing techniques. *Environ. Microbiol.* 16, 2672–2681. doi: 10.1111/1462-2920.12308
- Harding, T., Jungblut, A. D., Lovejoy, C., and Vincent, W. F. (2011). Microbes in high arctic snow and implications for the cold biosphere. *Appl. Environ. Microbiol.* 77, 3234–3243. doi: 10.1128/AEM.02611-10
- Heller, P., Tripp, H. J., Turk-Kubo, K., and Zehr, J. P. (2014). ARBITRATOR: a software pipeline for on-demand retrieval of auto-curated *nifH* sequences from GenBank. *Bioinformatics* 30, 2883–2890. doi: 10.1093/bioinformatics/btu417
- Hendricks, S., Nicolaus, M., and Schwegmann, S. (2012). *Sea Ice Conditions During POLARSTERN Cruise ARK-XXVII/3 (IceArc) in 2012*. Bremerhaven: Alfred Wegener Institute for Polar and Marine Research.
- Hennecke, H. (1990). Nitrogen fixation genes involved in the Bradyrhizobium japonicum- soybean symbiosis. *FEBS* 268, 422–426. doi: 10.1016/0014-5793(90)81297-2
- Herbert, R. A. (1999). Nitrogen cycling in coastal marine ecosystems. *FEMS Microbiol. Rev.* 23, 563–590. doi: 10.1111/j.1574-6976.1999.tb00414.x
- Holmes, R. M., McClelland, J. W., Peterson, B. J., Tank, S. E., Bulygina, E., Eglinton, T. I., et al. (2011). Seasonal and annual fluxes of nutrients and organic matter from large rivers to the Arctic Ocean and surrounding seas. *Estuaries Coasts* 35, 369–382. doi: 10.1007/s12237-011-9386-6
- Hsieh, T. C., Ma, K. H., and Chao, A. (2016). *iNEXT: iNterpolation and EXTrapolation for Species Diversity. R package*. Available at: <http://chao.stat.nthu.edu.tw/blog/software-download>
- Huang, Y., Niu, B., Gao, Y., Fu, L., and Li, W. (2010). CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* 26, 680–682. doi: 10.1093/bioinformatics/btq003
- IPCC (2013). *Climate Change 2013: The Physical Science Basis. Contribution of working group I to the Fifth Assessment Report of the IPCC*. Geneva: IPCC
- Izquierdo, J. A., and Nüsslein, K. (2006). Distribution of extensive *nifH* gene diversity across physical soil microenvironments. *Microb. Ecol.* 51, 441–452. doi: 10.1007/s00248-006-9044-x
- Jordan, D. C., McNicol, P. J., and Marshall, M. R. (1978). Biological nitrogen fixation in the terrestrial environment of a high Arctic ecosystem (Truelove Lowland, Devon Island, N.W.T.). *Can. J. Microbiol.* 24, 643–649. doi: 10.1139/m78-108
- Jungblut, A. D., and Neilan, B. A. (2010). *NifH* gene diversity and expression in a microbial mat community on the McMurdo Ice Shelf, Antarctica. *Antarct. Sci.* 22, 117–122. doi: 10.1017/S0954102009990514
- Klunder, M. B., Bauch, D., Laan, P., de Baar, H. J. W., van Heuven, S., Ober, S., et al. (2012). Dissolved iron in the Arctic shelf seas and surface waters of the central Arctic Ocean: Impact of Arctic river water and ice-melt. *J. Geophys. Res.* 117:C01027. doi: 10.1029/2011JC007133
- Koh, E. Y., Cowie, R. O. M., Simpson, A. M., O'Toole, R., and Ryan, K. G. (2012a). The origin of cyanobacteria in Antarctic sea ice: marine or freshwater? *Environ. Microbiol. Rep.* 4, 479–483. doi: 10.1111/j.1758-2229.2012.00346.x
- Koh, E. Y., Martin, A. R., McMinn, A., and Ryan, K. G. (2012b). Recent advances and future perspectives in microbial phototrophy in antarctic sea ice. *Biology (Basel)* 1, 542–556. doi: 10.3390/biology1030542
- Kruppen, T., Gerdes, R., Haas, C., Hendricks, S., Herber, A., Selyuzhenok, V., et al. (2016). Recent summer sea ice thickness surveys in the Fram Strait and associated volume fluxes. *Cryosph. Discuss.* 9, 5171–5202. doi: 10.5194/tcd-9-5171-2015
- Kruppen, T., Janout, M., Hodges, K. I., Gerdes, R., Girard-Arduin, F., Hölemann, J. A., et al. (2013). Variability and trends in laptev sea ice outflow between 1992–2011. *Cryosphere* 7, 349–363. doi: 10.5194/tc-7-349-2013
- Langlois, R. J., Hümmel, D., and LaRoche, J. (2008). Abundances and distributions of the dominant *nifH* phylotypes in the Northern Atlantic Ocean. *Appl. Environ. Microbiol.* 74, 1922–1931. doi: 10.1128/AEM.01720-07
- Langlois, R. J., LaRoche, J., and Raab, P. A. (2005). Diazotrophic diversity and distribution in the tropical and subtropical Atlantic Ocean. *Appl. Environ. Microbiol.* 71, 7910–7919. doi: 10.1128/AEM.71.12.7910
- Lara, R. J. R. J., Rachold, V., Kattner, G., Hubberten, H. W., Guggenberger, G., Skoog, A., et al. (1998). Dissolved organic matter and nutrients in the Lena



- river, siberian arctic: characteristics and distribution. *Mar. Chem.* 59, 301–309. doi: 10.1016/S0304-4203(97)00076-5
- LaRoche, J., and Breitbart, E. (2005). Importance of the diazotrophs as a source of new nitrogen in the ocean. *J. Sea Res.* 53, 67–91. doi: 10.1016/j.seares.2004.05.005
- Le Moal, M., Collin, H., and Biegala, I. C. (2011). Intriguing diversity among diazotrophic picoplankton along a Mediterranean transect: a dominance of rhizobia. *Biogeosciences* 8, 827–840. doi: 10.5194/bg-8-827-2011
- Legendre, P., and Gallagher, E. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia* 129, 271–280. doi: 10.1007/s004420100716
- Legendre, P., and Legendre, L. (1998). *Numerical Ecology*. 2nd English. Amsterdam: Elsevier Science BV.
- Lehto, N., Glud, R., Nordi, G., Zang, H., and Davison, W. (2014). Anoxic microniches in marine sediments induced by aggregate settlement: Biogeochemical dynamics and implications. *Biogeochemistry* 119, 307–327.
- Letunik, I., and Bork, P. (2007). Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* 23, 127–128. doi: 10.1093/bioinformatics/btl529
- Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22, 1658–1659. doi: 10.1093/bioinformatics/btl158
- Lovejoy, C., Legendre, L., Martineau, M.-J., Bacle, J., and von Quillfeldt, C. H. (2002). Distribution of phytoplankton and other protists in the North Water. *Deep Sea Res. Part II* 49, 5027–5047. doi: 10.1016/S0967-0645(02)00176-5
- Lovejoy, C., and Potvin, M. (2010). Microbial eukaryotic distribution in a dynamic Beaufort Sea and the Arctic Ocean. *J. Plankton Res.* 33, 431–444. doi: 10.1093/plankt/fbq124
- Lozupone, C., Hamady, M., and Knight, R. (2006). UniFrac—an online tool for comparing microbial community diversity in a phylogenetic context. *BMC Bioinformatics* 7:371. doi: 10.1186/1471-2105-7-371
- Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, et al. (2004). ARB: a software environment for sequence data. *Nucleic Acids Res.* 32, 1363–1371. doi: 10.1093/nar/gkh293
- Luo, Y.-W. Y.-W., Doney, S. C., Anderson, L. A., Benavides, M., Berman-Frank, I., Bode, A., et al. (2012). Database of diazotrophs in global ocean: abundance, biomass and nitrogen fixation rates. *Earth Syst. Sci. Data* 4, 47–73. doi: 10.5194/essd-4-47-2012
- McMurdie, P. J., and Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* 8:e61217. doi: 10.1371/journal.pone.0061217
- Mehta, M. P., Butterfield, D. A., and Baross, J. A. (2003). Phylogenetic diversity of Nitrogenase (*nifH*) genes in deep-sea and hydrothermal vent environments of the Juan de Fuca Ridge. *Appl. Environ. Microbiol.* 69:960. doi: 10.1128/AEM.69.2.960
- Mills, M. M., and Arrigo, K. R. (2010). Magnitude of oceanic nitrogen fixation influenced by the nutrient uptake ratio of phytoplankton. *Nat. Geosci.* 3, 412–416. doi: 10.1038/ngeo856
- Moisander, P. H., Beinart, R. A., Hewson, I., White, A. E., Johnson, K. S., Carlson, C. A., et al. (2010). Unicellular cyanobacterial distributions broaden the oceanic N<sub>2</sub> fixation domain. *Science* 327, 1512–1514. doi: 10.1126/science.1185468
- Moisander, P. H., Morrison, A. E., Ward, B. B., Jenkins, B. D., and Zehr, J. P. (2007). Spatial-temporal variability in diazotroph assemblages in Chesapeake Bay using an oligonucleotide *nifH* microarray. *Environ. Microbiol.* 9, 1823–1835. doi: 10.1111/j.1462-2920.2007.01304.x
- Moisander, P. H., Serros, T., Paerl, R. W., Beinart, R. A., and Zehr, J. P. (2014). Gammaproteobacterial diazotrophs and *nifH* gene expression in surface waters of the South Pacific Ocean. *ISME J.* 8, 1962–1973. doi: 10.1038/ismej.2014.49
- Monteiro, F. M., Dutkiewicz, S., and Follows, M. J. (2011). Biogeographical controls on the marine nitrogen fixers. *Glob. Biogeochem. Cycles* 25:GB2003. doi: 10.1029/2010GB003902
- Murphy, L., and Haugen, E. (1985). The distribution and abundance of phototrophic ultraplankton in the North Atlantic. *Limnol. Oceanogr.* 30, 47–58. doi: 10.4319/lo.1985.30.1.0047
- Nadeau, T. L., Milbrandt, E. C., and Castenholz, R. W. (2001). Evolutionary relationships of cultivated antarctic oscillatorians (cyanobacteria). *J. Phycol.* 37, 650–654. doi: 10.1046/j.1529-8817.2001.037004650.x
- Oksanen, J., Blanchet, F., Guillaume Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., et al. (2013). *Vegan: Community Ecology Package. R package version 2.0-10*. Available at: <http://cran.r-project.org/package=vegan>
- Olson, J. B., Stegge, T. F. T., Litaker, R. W. R., and Paerl, H. W. (1998). N<sub>2</sub>-fixing microbial consortia associated with the ice cover of Lake Bonney, Antarctica. *Microb. Ecol.* 36, 231–238. doi: 10.1007/s002489900110
- Paerl, H. W., and Zehr, J. P. (2000). “Marine nitrogen fixation,” in *Microbial Ecology of the Oceans*, ed. D. L. Kirchman (New York, NY: Wiley-Liss), 387–426.
- Parkinson, C. L., and Comiso, J. C. (2013). On the 2012 record low Arctic sea ice cover: Combined impact of preconditioning and an August storm. *Geophys. Res. Lett.* 40, 1356–1361. doi: 10.1002/grl.50349
- Partensky, F., and Hess, W. R. (1999). Prochlorococcus, a marine photosynthetic prokaryote of global significance. *Microbiol. Mol. Biol. Rev.* 63, 106–127.
- Polyakov, I., Timokhov, L., Alexeev, V., Bacon, S., Dmitrenko, I., Fortier, L., et al. (2010). Arctic Ocean warming contributes to reduced polar ice cap. *J. Phys. Oceanogr.* 40, 2743–2756. doi: 10.1175/2010JPO4339.1
- R Core Team (2015). *R: A language and Environment for Statistical Computing*. Available at: <http://www.r-project.org/>
- Rabe, B., Wisotzki, A., Rettig, S., Somavilla Cabrillo, R., and Sander, H. (2012). *Physical Oceanography Measured on Water Bottle Samples During POLARSTERN Cruise ARK-XXVII/3*. Bremerhaven: Alfred Wegener Institute for Polar and Marine Research.
- Ramette, A. (2009). Quantitative community fingerprinting methods for estimating the abundance of operational taxonomic units in natural microbial communities. *Appl. Environ. Microbiol.* 75, 2495–2505. doi: 10.1128/AEM.02409-08
- Riemann, L., Farnelid, H., and Steward, G. (2010). Nitrogenase genes in non-cyanobacterial plankton: prevalence, diversity and regulation in marine waters. *Aquat. Microb. Ecol.* 61, 235–247. doi: 10.3354/ame01431
- Rysgaard, S., Glud, R. N., Risgaard-Petersen, N., and Dalsgaard, T. (2004). Denitrification and anammox activity in Arctic marine sediments. *Limnol. Oceanogr.* 49, 1493–1502. doi: 10.4319/lo.2004.49.5.1493
- Sakshaug, E., Stein, R., and Macdonald, R. W. (2004). “Primary and secondary production in the Arctic Seas,” in *The Organic Carbon Cycle in the Arctic Ocean*, eds R. Stein and R. W. Macdonald (Berlin: Springer), 57–82.
- Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* 74, 5463–5467.
- Schmidt, S., Moskall, W., De Mora, S. J., Howard-Williams, C., and Vincent, W. F. (1991). Limnological properties of Antarctic ponds during winter freezing. *Antarct. Sci.* 3, 379–388. doi: 10.1017/S0954102091000482
- Sherr, E. B., Sherr, B. F., Wheeler, P. A., and Thompson, K. (2003). Temporal and spatial variation in stocks of autotrophic and heterotrophic microbes in the upper water column of the central Arctic Ocean. *Deep Sea Res. Part I Oceanogr. Res. Pap.* 50, 557–571. doi: 10.1016/S0967-0637(03)00031-1
- Stamatakis, A. (2014). RAxML Version 8: a tool for Phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033
- Staples, C. R., Lahiri, S., Raymond, J., Von Herbulis, L., Mukhopadhyay, B., and Blankenship, R. E. (2007). Expression and association of group IV nitrogenase *NifD* and *NifH* homologs in the non-nitrogen-fixing archaeon *Methanocaldococcus jannaschii*. *J. Bacteriol.* 189, 7392–7398. doi: 10.1128/JB.00876-07
- Stroeve, J. C., Kattsov, V., Barrett, A., Serreze, M., Pavlova, T., Holland, M., et al. (2012). Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations. *Geophys. Res. Lett.* 39, L16502. doi: 10.1029/2012GL052676
- Tedrow, J. C. F. (1966). Polar desert soils. *Soil Sci. Soc. Am. J.* 30, 381. doi: 10.2136/sssaj1966.03615995003000030024x
- Torres-Valdés, S., Tsubouchi, T., Bacon, S., Naveira-Garabato, A. C., Sanders, R., McLaughlin, F. A., et al. (2013). Export of nutrients from the Arctic Ocean. *J. Geophys. Res. Ocean* 118, 1625–1644. doi: 10.1002/jgrc.20063
- Tremblay, G., Belzile, C., Gosselin, M., Poulin, M., and Roy, S. (2009). Late summer phytoplankton distribution along a 3500 km transect in Canadian Arctic waters: strong numerical dominance by picoeukaryotes. *Aquat. Microb. Ecol.* 54, 55–70. doi: 10.3354/ame01257
- Tremblay, J.-É., Anderson, L. G., Matrai, P., Coupel, P., Bélanger, S., Michel, C., et al. (2015). Global and regional drivers of nutrient supply, primary production and CO<sub>2</sub> drawdown in the changing Arctic Ocean. *Prog. Oceanogr.* 139, 171–196. doi: 10.1016/j.pocean.2015.08.009

- Tremblay, J. -É., Simpson, K., Martin, J., Miller, L., Gratton, Y., Barber, D., et al. (2008). Vertical stability and the annual dynamics of nutrients and chlorophyll fluorescence in the coastal, southeast Beaufort Sea. *J. Geophys. Res.* 113:C07S90. doi: 10.1029/2007JC004547
- Turk, K. A., Rees, A. P., Zehr, J. P., Pereira, N., Swift, P., Shelley, R., et al. (2011). Nitrogen fixation and nitrogenase (*nifH*) expression in tropical waters of the eastern North Atlantic. *ISME J.* 5, 1201–1212. doi: 10.1038/ismej.2010.205
- Turk-Kubo, K. A., Achilles, K. M., Serros, T. R. C., Ochiai, M., Montoya, J. P., and Zehr, J. P. (2012). Nitrogenase (*nifH*) gene expression in diazotrophic cyanobacteria in the Tropical North Atlantic in response to nutrient amendments. *Front. Microbiol.* 3:386. doi: 10.3389/fmicb.2012.00386
- Turk-Kubo, K. A., Frank, I. E., Hogan, M. E., Desnues, A., Bonnet, S., and Zehr, J. P. (2015). Diazotroph community succession during the VAHINE mesocosm experiment (New Caledonia lagoon). *Biogeosciences* 12, 7435–7452. doi: 10.5194/bg-12-7435-2015
- Turk-Kubo, K. A., Karamchandani, M., Capone, D. G., and Zehr, J. P. (2014). The paradox of marine heterotrophic nitrogen fixation: abundances of heterotrophic diazotrophs do not account for nitrogen fixation rates in the Eastern Tropical South Pacific. *Environ. Microbiol.* 16, 3095–3114. doi: 10.1111/1462-2920.12346
- Tyrrell, T. (1999). The relative influences of nitrogen and phosphorus on oceanic primary production. *Nature* 400, 525–531. doi: 10.1038/22941
- Villareal, T. A. (1992). “Marine nitrogen-fixing diatom-cyanobacteria symbiosis,” in *Marine Pelagic Cyanobacteria: Trichodesmium and other Diazotrophs*, eds E. J. Carpenter, D. G. Capone, and J. G. Rueter (Dordrecht: Springer), 163–175. doi: 10.1007/978-94-015-7977-3\_10
- Vincent, W. F. (2000). “Cyanobacterial dominance in the polar regions,” in *The Ecology of Cyanobacteria*, eds A. Whitton and M. Potts (Amsterdam: Kluwer Academic Publishers), 321–340.
- Vonnamme, T. R., Devetter, M., Žárský, J. D., Šabacká, M., and Elster, J. (2015). Controls on microalgal community structures in cryoconite holes upon high Arctic glaciers, Svalbard. *Biogeosci. Discuss.* 12, 11751–11795. doi: 10.5194/bg-12-11751-2015
- Warton, D. I., Wright, S. T., and Wang, Y. (2012). Distance-based multivariate analyses confound location and dispersion effects. *Methods Ecol. Evol.* 3, 89–101. doi: 10.1111/j.2041-210X.2011.00127.x
- Wassmann, P., Duarte, C. M., Agustí, S., and Sejr, M. K. (2011). Footprints of climate change in the Arctic marine ecosystem. *Glob. Chang. Biol.* 17, 1235–1249. doi: 10.1111/j.1365-2486.2010.02311.x
- Yallop, M. L., Anesio, A. M., Perkins, R. G., Cook, J., Telling, J., Fagan, D., et al. (2012). Photophysiology and albedo-changing potential of the ice algal community on the surface of the Greenland ice sheet. *ISME J.* 6, 2302–2313. doi: 10.1038/ismej.2012.107
- Young, J. P. W. (2005). “The phylogeny and evolution of nitrogenases,” in *Genomes and Genomics of Nitrogen Fixing Organisms*, eds R. Palacios and W. E. Newton (Dordrecht: Springer), 221–241. doi: 10.1007/1-4020-3054-1\_14
- Zehr, J. P., Jenkins, B. D., Short, S. M., and Steward, G. F. (2003). Nitrogenase gene diversity and microbial community structure: a cross-system comparison. *Environ. Microbiol.* 5, 539–554. doi: 10.1046/j.1462-2920.2003.00451.x
- Zehr, J. P., and Turner, P. J. (2001). “Nitrogen fixation: Nitrogenase genes and gene expression,” in *Methods in Microbiology*, ed. J. H. Paul (San Francisco, CA: Elsevier), 271–286.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Fernández-Méndez, Turk-Kubo, Buttigieg, Rapp, Krumpfen, Zehr and Boetius. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.