The ISME Journal (2013) 7, 1413–1423

www.nature.com/ismej

ORIGINAL ARTICLE

Active nitrogen-fixing heterotrophic bacteria at and below the chemocline of the central Baltic Sea

Hanna Farnelid[1], Mikkel Bentzon-Tilia[2], Anders F Andersson[3], Stefan Bertilsson[4], Günter Jost[5], Matthias Labrenz[5], Klaus Jürgens[5] and Lasse Riemann[1,2]

[1]*Department of Natural Sciences, Linnaeus University, Kalmar, Sweden;* [2]*Marine Biological Section, Department of Biology, University of Copenhagen, Helsingør, Denmark;* [3]*Division of Gene Technology, KTH Royal Institute of Technology, Science for Life Laboratory, School of Biotechnology, Solna, Sweden;* [4]*Department of Ecology and Genetics, Limnology, Uppsala University, Uppsala, Sweden and* [5]*Department of Biological Oceanography, Leibniz Institute for Baltic Sea Research Warnemünde (IOW), Rostock, Germany*

The Baltic Sea receives large nitrogen inputs by diazotrophic (N_2-fixing) heterocystous cyanobacteria but the significance of heterotrophic N_2 fixation has not been studied. Here, the diversity, abundance and transcription of the *nifH* fragment of the nitrogenase enzyme in two basins of the Baltic Sea proper was examined. N_2 fixation was measured at the surface (5 m) and in anoxic water (200 m). Vertical sampling profiles of >10 and $<10\,\mu m$ size fractions were collected in 2007, 2008 and 2011 at the Gotland Deep and in 2011 in the Bornholm Basin. Both of these stations are characterized by permanently anoxic bottom water. The 454-pyrosequencing *nifH* analysis revealed a diverse assemblage of *nifH* genes related to *alpha*-, *beta*- and *gammaproteobacteria* (*nifH* cluster I) and anaerobic bacteria (*nifH* cluster III) at and below the chemocline. Abundances of genes and transcripts of seven diazotrophic phylotypes were investigated using quantitative polymerase chain reaction revealing abundances of heterotrophic *nifH* phylotypes of up to 2.1×10^7 *nifH* copies l^{-1}. Abundant *nifH* transcripts (up to 3.2×10^4 transcripts l^{-1}) within *nifH* cluster III and co-occurring N_2 fixation (0.44 ± 0.26 nmol l^{-1} day^{-1}) in deep water suggests that heterotrophic diazotrophs are fixing N_2 in anoxic ammonium-rich waters. Our results reveal that N_2 fixation in the Baltic Sea is not limited to illuminated N-deplete surface waters and suggest that N_2 fixation could also be of importance in other suboxic regions of the world's oceans.

The ISME Journal (2013) **7,** 1413–1423; doi:10.1038/ismej.2013.26; published online 28 February 2013
Subject Category: microbial ecology and functional diversity of natural habitats
Keywords: Baltic Sea; chemocline; *nifH*; nitrogen fixation; oxic-anoxic interface; pyrosequencing

Introduction

Biological N_2 fixation is the main source of fixed nitrogen (N) in the oceans and a key process in regulating primary productivity and thereby carbon fluxes (Gruber, 2005). N_2 fixation is carried out by a limited, but diverse, group of bacteria and Methanotrophic *Archaea,* known as diazotrophs (Zehr *et al.*, 2003b). The Baltic Sea is among the world's largest brackish seas and regularly experiences extensive blooms of diazotrophic heterocystous cyanobacteria (Stal *et al.*, 2003) providing N-input almost as large as the riverine load (Larsson *et al.*, 2001). Cyanobacterial N_2 fixation is also significant in open oceans (Goebel *et al.*, 2007), but molecular techniques targeting the *nifH* gene of the nitrogenase enzyme have revealed that heterotrophic diazotrophs are also present and *nifH* transcribed in diverse marine and estuarine environments (reviewed in Riemann *et al.*, 2010). However, their significance to N_2 fixation and the factors controlling their activity and distribution patterns are not yet understood. A better understanding of their ecology is therefore essential for resolving the role of this functional group in oceanic and coastal nutrient cycling.

Nitrogenase activity is thought to be tightly regulated by ammonium (NH_4^+) availability (Kanemoto and Ludden, 1984; Klugkist and Haaker, 1984) and is inactivated and destroyed by oxygen (O_2; Dixon and Kahn, 2004). Consequently, heterotrophic diazotrophs likely require suboxic conditions for active N_2 fixation. In the oxygenated water column, such loci may be associated with live or dead particulate matter (reviewed in Riemann *et al.*, 2010). Alternatively, these organisms may exploit low O_2 biomes, such as pelagic transition

Correspondence: L Riemann, Marine Biological Section, University of Copenhagen, Strandpromenaden 5, Helsingør, DK-3000, Denmark.
E-mail: lriemann@bio.ku.dk
Received 30 October 2012; revised 15 January 2013; accepted 18 January 2013; published online 28 February 2013

zones between oxic and anoxic layers and oceanic oxygen minimum zones (OMZs). Such areas, representing zones with large N-losses through anammox and heterotrophic denitrification (Lam and Kuypers, 2011), can be possible sites for heterotrophic N_2 fixation (Zehr *et al.*, 2006; Riemann *et al.*, 2010). Concordantly, N_2 fixation rates and diverse putative heterotrophic diazotrophs were recently reported from hypoxic waters in the eastern tropical South Pacific (Fernandez *et al.*, 2011) and the Southern Californian Bight (Hamersley *et al.*, 2011). Similarly, *nifH* genes and transcripts related to heterotrophs were reported from the Arabian Sea OMZ (Jayakumar *et al.*, 2012). Interestingly, a recent study documented the co-occurrence of N_2 fixation and denitrification in a meromictic lake (Halm *et al.*, 2009). Combined, these findings suggest that two central N transformation processes, thus far thought to be segregated, can co-occur in OMZs, challenging the conventional view of N cycling in the oceans and unveiling the intriguing possibility of NH_4^+-rich suboxic and anoxic waters as environments for heterotrophic N_2 fixation. However, knowledge on N_2 fixation in suboxic waters is limited (Ulloa *et al.*, 2012) and molecular data and N_2 fixation rates in truly anoxic marine systems is lacking.

The Baltic Sea proper is characterized by a permanent halocline, which prevents vertical mixing. Organic matter mineralization causes O_2 deficiency ($<2\,ml\,l^{-1}$) in deep waters, covering an area of 12 000–70 000 km^2 (Conley *et al.*, 2002). At and below the chemocline, hydrogen sulfide (H_2S) and NH_4^+ accumulates and the potential for denitrification and anammox has been demonstrated (Hietanen *et al.*, 2012). The aim of this study was to explore the chemocline and anoxic waters as potential loci for heterotrophic N_2 fixation. The 454-pyrosequencing provided high resolution data on the distribution of *nifH* phylotypes and through quantitative polymerase chain reaction (QPCR) and reverse transcription QPCR (RT-QPCR), we determined the abundance and expression of seven prominent *nifH* gene clusters. The analysis of five sampling profiles, four from the Gotland Deep and one from the Bornholm Basin, from three sampling years, revealed a large diversity of heterotrophic diazotrophs at and below the chemocline. We also report transcription of *nifH* and N_2 fixation rates in anoxic NH_4^+-rich water and thereby extend the currently known regime of marine N_2 fixation.

Materials and methods

Environmental sampling

Samples were collected at the Gotland Deep (240 m total depth, 57°20.09'N, 20°03.09'E) in the summers of 2007 (9 July, morning), 2008 (15 August, noon) and 2011 (14 and 18 July, morning). Samples were collected between 5 and 233 m using 5 l Free-Flow bottles (Hydrobios) attached to a PVC-coated stainless steel rosette equipped with sensors for O_2 (Sea-Bird SBE 43), conductivity, temperature and depth (Seabird Inc., Bellevue, WA, USA), with high depth resolution around the chemocline (129–132 m in 2007, 121–127 m in 2008 and 110–121 m in 2011, Figure 1). In 2011 (13 July), samples were collected at 10 m intervals from the Bornholm Basin (89 m total depth, 55°15.05'N, 15°59.06'E). For each sample, 1–3.8 l were filtered onto a 10 μm polycarbonate filter (47 mm, GE Water & Process Technologies, Trevose, PA, USA) and 0.6–2 l of the filtrate was filtered onto a 0.2 μm (47 mm) Supor filter (PALL Corporation, Port Washington, NY, USA). After filtration, RNA filters

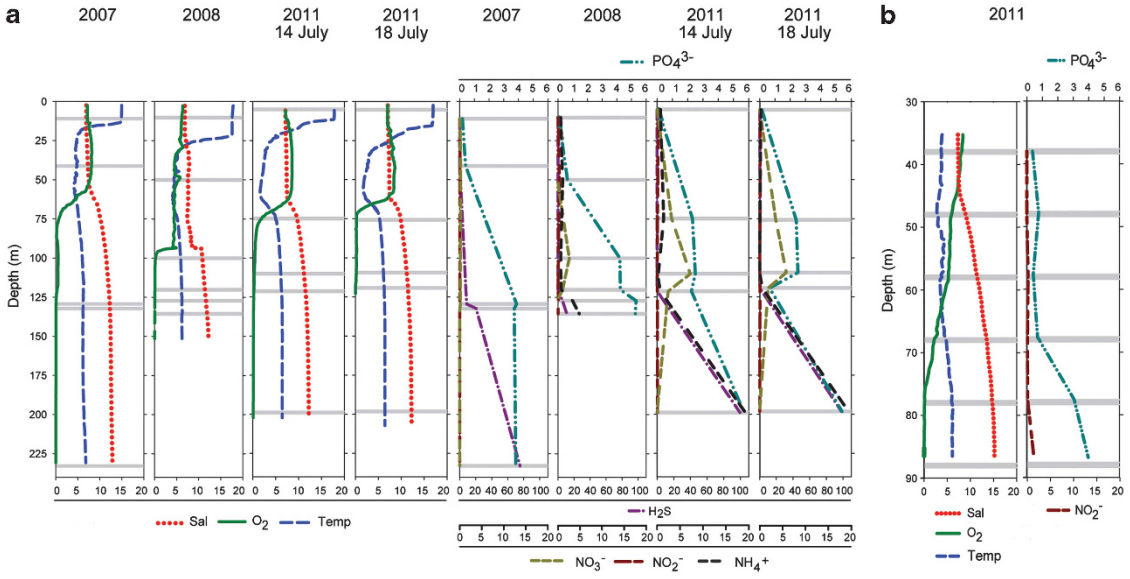

Figure 1 Environmental data from the sampled depth profiles at the Gotland Deep (**a**) and the Bornholm Basin (**b**) in the Baltic Sea proper. Depths of sampling for nutrients and nucleic acids are indicated by gray horizontal lines. Concentrations of inorganic nutrients are in $\mu mol\,l^{-1}$, temperature in °C and O_2 in $ml\,l^{-1}$.

Table 1 Summary of samples obtained and analyzed by 454-pyrosequencing and/or QPCR

Gotland Deep								*Bornholm Basin*	
Depth (m)	*2007 9 July*	*Depth (m)*	*2008 15 August*	*Depth (m)*	*2011 14 July*	*Depth (m)*	*2011 18 July*	*Depth (m)*	*2011 13 July*
5	I	5	I	5	I, II	5	I, II	38	I, II
11	A, B, I	15	A, I	75	I, II	75	I, II	48	I, II
41	A, B, C, I	50	I	110	I, II	110	I, II	58	I, II
101	I	100	A, I	121	I, II	119	I, II	68	I, II
129	A, B, C, I	121	A, I	200	I, II	200	I, II	78	I, II
132	A, B, C, I	127	A, I					88	I, II
181	I	135	A, I						
233	A, B, C, I								

Abbreviations: cDNA, complementary DNA; QPCR, quantitative polymerase chain reaction.
454-Pyrosequencing DNA $<10\,\mu m$ (A), 454-pyrosequencing DNA $>10\,\mu m$ (B), 454-pyrosequencing cDNA $<10\,\mu m$ (C), QPCR DNA $<10\,\mu m$ and DNA $>10\,\mu m$ (I), QPCR cDNA $<10\,\mu m$ and cDNA $>10\,\mu m$ (II).

were soaked in RNAlater (Ambion, Carlsbad, CA, USA). Filters were placed in 0.5 ml RNAlater or 1 ml Tris-EDTA (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) for RNA and DNA filters, respectively, and the cryotubes were frozen in liquid N or at $-20\,^{\circ}C$ (2011). A summary of nucleic acid samples and analytical methods is provided in Table 1. Samples for bacterial enumeration were fixed with glutaraldehyde (1% final concentration) and counted using flow cytometry (Troussellier *et al*., 1999). Chemical profiles of H_2S, NH_4^+ (not analyzed in 2007), phosphate (PO_4^{3-}), nitrite (NO_3^-) and nitrate (NO_2^-) were determined for the Gotland Deep and PO_4^{3-} and NO_2^- for the Bornholm Basin (Grasshoff *et al*., 1983).

N_2 fixation rates

On the 14 July 2011, rates of N_2 fixation were calculated using $^{15}N_2$-incorporation (Montoya *et al*., 1996) in samples from the oxygenated surface and anoxic waters at the Gotland Deep. Water was transferred directly from the Free-Flow bottle into triplicate 1.2 l acid-washed polycarbonate bottles pre-flushed with argon. The bottles were filled from the bottom up using silicone tubing. 1.0 ml of $^{15}N_2$ tracer gas (98%, Campro Scientific, Berlin, Germany) was added through a gas-tight septum. The bottles were inverted 50 times and incubated for 24 h in an on-deck container with running surface seawater covered with fabric reducing the light to 13% of the surface light intensity (5 m samples) or at 8 °C in the dark (200 m samples). Post-incubation size-fractionation was performed to quantify free-living N_2 fixation: 600 ml of water was filtered through a 10 μm polycarbonate filter (GE Water & Process Technologies) and thereafter through a pre-combusted (450 °C for 8 h) GF/F filter (0.7 μm, Whatman, GE Healthcare, Little Chalfont, UK). The remaining 600 ml were filtered through a GF/F filter. The background $\delta^{15}N$ levels, two time-zero controls, and un-amended controls were included for each depth. Filters were stored at $-20\,^{\circ}C$, dried at 60 °C for 24 h, and analyzed on an isotope ratio mass spectrometer (PDZ Europa, Northwich, UK) at the Laboratory of Applied Physical Chemistry, Gent, Belgium. In these analyses, particulate N concentrations were $>33\,\mu g\,N\,l^{-1}$ and did not limit detection. No significant differences were observed between background $\delta^{15}N$ levels and the time zero and un-amended controls.

Nucleic acids extractions and complementary DNA (cDNA) synthesis

Community DNA was extracted using an enzyme/phenol-chloroform protocol (Riemann *et al*., 2008). Community RNA was extracted using the RNeasy mini kit (Qiagen, Hilden, Germany; Boström *et al*., 2007b). For each sampling year and filter type, blank filters were included as negative controls. DNA and RNA concentrations were measured using NanoDrop 2000 (Thermo Scientific, Waltham, MA, USA). First-strand cDNA synthesis was done using the nifH3 primer (Zehr and McReynolds, 1989) and the TaqMan reverse transcription kit (Applied Biosystems, Foster City, CA, USA) including a negative control with only RT-PCR grade water (Ambion) and a second set of reactions without reverse transcriptase. The DNA and cDNA was stored at $-20\,^{\circ}C$.

PCR amplification and 454-pyrosequencing

A total of 16 DNA and 4 cDNA samples were amplified with nifH3 and nifH4 primers (Zehr and Turner, 2001) followed by 30 cycles with custom bar-coded nifH1 and nifH2 primers (Supplementary Table S1) using Pure Taq Ready-To-Go PCR Beads (GE Healthcare). As the cDNA samples from 41, 132 and 233 m yielded faint PCR products they were first amplified with nifH1 and nifH2 primers (Zehr and McReynolds, 1989) followed by 10 PCR cycles with barcoded primers (Farnelid *et al*., 2011). Non-transcribed RNA samples were included as control for incomplete DNA digestion. The PCR products were gel purified (Gel extraction kit, Qiagen) and quantified (NanoDrop 2000). For PCR samples from

negative extraction and RT-PCR grade water controls, although no bands were visible, a slice of the expected product size (402 bp) was excised, purified, and cloned (TOPO TA, Invitrogen, Carlsbad, CA, USA). No clones were obtained. The PCR products were mixed in equal amounts and sequenced using a GS FLX pyrosequencing system (Roche, Basel, Switzerland). Sequences are deposited in the National Center for Biotechnology Information Sequence Read Archive under accession number SRA058502.

Sequence quality controls and analysis

The data set was filtered to only include sequences with correct tag and primer sequences. These were then trimmed, sequences were clipped at 219 bp, translated into amino acids (73 aa), and sequences having in-frame stop codon(s) were removed. Unique reads were aligned as previously described (Farnelid *et al.*, 2011). Reads were screened to check for homology to all reported *nifH* contaminant sequences, but no identical reads were found. To study the phylogeny of the dominant clusters, the reads were clustered at 96% similarity using CD-hit (Li and Godzik, 2006) and clusters containing <100 sequences were removed. The remaining 222 clusters and related sequences in the database were aligned in MEGA (Tamura *et al.*, 2007) and a phylogenetic tree was created (p-distance). To allow for cross-sample comparisons, a random re-sampling to identical sequencing depth (5000 sequences) was done using an in-house developed Perl script. The S_{Chao1} richness estimator (Chao, 1984) and the Shannon diversity index (Shannon, 1948) were calculated and rarefaction curves were generated using Analytical Rarefaction 1.3 (http://www.uga.edu/strata/software/index.html). Bray–Curtis distances between samples were calculated from the abundances of the *nifH* clusters using the R package Vegan (http://www.r-project.org/). Principal coordinates analysis and hierarchical clustering based on these pairwise distances were performed with the R packages Made4 and Cluster, respectively.

Quantitative PCR and RT-QPCR

Eight of the identified *nifH* clusters were targeted by QPCR. For seven of the clusters, primers and probes were designed using the software Primer Express 3.0 (Applied Biosystems). The primer/probe set for the ALHOU cluster was designed to also amplify the B24DA cluster (Supplementary Table S2) and is referred to as ALHOU hereafter. For the *Nodularia*-like *nifH* cluster (CM8AH), the forward primer and probe developed by Boström *et al.* (2007b) and a reverse primer within the 219 bp 454-read was used (Supplementary Table S2). Primers and probes had mismatches with all non-target sequences (Supplementary Table S2), and the specificity was verified through BLAST searches against the NCBI database. A QPCR standard for the EVHVF target was obtained from an available clone (EU916338, Farnelid *et al.*, 2009). For the other targets, the 219 bp long sequences were synthesized and cloned into pUC19 vectors (Epoch Life Science, Missouri City, TX, USA). Plasmids were extracted (EZNA Plasmid Mini Kit, Omega bio-tek, Norcross, GA, USA), linearized using *Hind*III (10 U μl^{-1}, Roche), purified (EZNA gel extraction kit, Omega bio-tek) and quantified (NanoDrop 2000). The specificities of the primer and probe sets were evaluated for each target with 10^5 copies of the non-target standards. Unspecific amplification was very low (<34 copies of 10^5 detected). To examine potential PCR inhibition in the samples, 5, 2 or 0.5 ng of mixed DNA from each profile and size fraction was run in QPCR with 10^4 copies of standard. There was no inhibition in any of the samples at 2 ng per reaction.

The QPCR and RT-QPCR reactions were run on a StepOnePlus Real-Time PCR System (Applied Biosystems). Each reaction (25 µl) consisted of TaqMan Universal PCR Master Mix (Applied Biosystems), 900 nM of each primer (HPLC purified, Sigma-Aldrich, St Louis, MO, USA), 250 mM probe (Applied Biosystems), RT-PCR grade water and 1–2 µl standard or sample (DNA 2 ng, cDNA 1.2 ng). The thermal cycling conditions were 2 min at 50 °C, 10 min at 95 °C followed by 40 cycles of 15 s at 95 °C and 1 min at 60 °C. All DNA samples, standard dilutions (10^6–10^1 copies), and no-template controls (RT-PCR grade water) were run in triplicate and cDNA samples were run in duplicate. No-template controls were undetectable. The efficiency of the primer/probe sets was 80–103% (Supplementary Table S3). In some samples, at least one of the replicates did not produce an amplification signal above the lowest point on the standard curve (10 *nifH* copies); in these, targets were considered detectable but not quantifiable. To calculate the percentage of heterotrophic *nifH* phylotypes, the abundances of the quantified *nifH* clusters were summed and related to the bacterial abundance in the sample, assuming one *nifH* copy per genome and one genome per bacterium (Langlois *et al.*, 2008).

Results and Discussion

Environmental data

The Gotland Deep station is characterized by a stable halocline (60–80 m) dividing the water column into an upper oxygenated layer and an underlying suboxic and anoxic layer. In 2007, the chemocline was at 127–130 m (Figure 1a). In 2008, the halocline was deeper (90–100 m), possibly due to a recent mixing event, and the chemocline was at 121–127 m. In 2011, the halocline and the chemocline were located around 65–75 m and 120–125 m, respectively (Figure 1a). At 200 m, NH_4^+

concentrations were high (19.7 and 19.4 μM; Figure 1a). Coinciding with the disappearance of O_2, peaks in NO_3^- concentration were observed in 2008 (2.57 μM, 100 m) and in 2011 (7.25 μM, 110 m and 5.82 μM, 109 m). At the Bornholm Basin, O_2 was below detection at 75 m (Figure 1b). For all sampling profiles, the bacterial abundance was highest at the surface ($2.6 \times 10^6 - 4.2 \times 10^6$ ml^{-1}) and lowest at the halocline ($1.3 \times 10^5 - 3.8 \times 10^5$ ml^{-1}; Supplementary Figure S1). No peaks were observed directly at the chemocline, as previously reported (Höfle and Brettar, 1995). Instead the abundance increased with depth or remained stable below the chemocline.

Diversity and composition of nifH DNA and cDNA libraries

Pyrosequencing of *nifH* genes PCR amplified from the Gotland Deep in 2007 and 2008 (Table 1) produced a total of 311 857 sequences encoding 28 812 unique proteins (Supplementary Figure S2). The number of high-quality sequences varied from 5814 to 40 670 per sample. Rarefaction analysis at 96% similarity clustering indicated that the 5000 resampled sequences were a good representation of the total diversity (Lundin *et al.*, 2012; Supplementary Figure S3). The cDNA samples featured the least number of clusters (Supplementary Figure S4D) suggesting that only a limited number of *nifH* gene clusters were actively transcribed. Similar differences between DNA and RNA libraries were previously reported (for example, Man-Aharonovich *et al.*, 2007). We speculate that the discrepancy is caused by temporal changes in environmental conditions with resulting transient niches for activity of specific N_2-fixers. However, it should also be noted that the current sampling and extraction protocol may lead to biases in gene expression (Feike *et al.*, 2012) and thereby underestimation of the occurrence of transcripts.

The *nifH* gene and transcript libraries showed that the diazotrophic community consisted mainly of *alpha-*, *beta-* and *gammaproteobacteria* belonging to the canonical *nifH* cluster I and diverse anaerobes within *nifH* cluster III (Chien and Zinder, 1996), including, for example, *Desulfovibrio* spp. and archaeal methanogens (Figure 2). Previous studies have reported diverse non-cyanobacterial *nifH* phylotypes within these clusters in both surface (Man-Aharonovich *et al.*, 2007; Farnelid *et al.*, 2011; Mulholland *et al.*, 2012) and deep waters (Hamersley *et al.*, 2011; Jayakumar *et al.*, 2012). Among the *nifH* sequences within cluster III in this study (Figure 2), none were identical to previously reported clones from the surface of the Baltic Sea proper (Farnelid *et al.*, 2009). Interestingly, cluster III phylotypes were not enriched in the >10 μm size fraction relative to the <10 μm size fraction (Supplementary Figure S2B). Hence, in contrast with previous suggestions (for example, Braun *et al.*, 1999; Man-Aharonovich *et al.*, 2007) particles did not appear to provide an advantageous environment for these bacteria.

The occurrence of alternative nitrogenases (cluster II) in the *nifH* libraries was surprisingly low (208 sequences) considering their prevalence among clones (Farnelid *et al.*, 2009) or isolates from the Baltic Sea (Boström *et al.*, 2007a). Although libraries consisted mostly of non-cyanobacterial sequences, *nifH* reads affiliated with *Nodularia*, known to dominate cyanobacterial blooms (Stal *et al.*, 2003; Boström *et al.*, 2007b), were common in surface samples in 2007. *Pseudanabaena*-like *nifH* gene sequences were highly represented in oxic samples both sampling years (Figure 2 and Supplementary Figure S2). For the non-cyanobacterial clusters, the highest relative abundance generally occurred at or below the chemocline and interestingly, among the ones detected in cDNA, all feature the highest relative abundance at the chemocline (Figure 2). Thus, the chemocline provides environmental conditions that promote growth of diverse assemblages of putative heterotrophic diazotrophs that actively express nitrogenase.

The principal coordinates analysis showed that the composition of putative diazotrophs in the oxic samples was distinct from those at the chemocline and anoxic waters. In contrast, the latter two groups were intermingled (Figure 3 and Supplementary Figure S5). The oxic samples displayed large variation in composition, partly related to the size fraction analyzed (with the exception of 2007, 41 m, <10 μm; Figure 3). Generally, the *nifH* diversity increased with depth and the gradual decrease in O_2 (Supplementary Figures S3 and S4). The Chao richness estimators and Shannon diversity indexes, respectively, were similar in 2007 and 2008, decreasing abruptly around the chemocline in <10 μm samples (Supplementary Figures S4A, B). Our observations of increased *nifH* diversity with depth suggest that suboxic water contains numerous ecological niches for putative heterotrophic diazotrophs. In addition, most of the identified clusters were found during both years (2007 and 2008; Figure 2), suggesting that they are a stable component of the Baltic Sea bacterial community.

Distribution and abundances of nifH phylotypes and transcripts

One drawback of the nested PCR approach is the high susceptibility to contamination through handling or presence of *nifH* genes in PCR reagents (Zehr *et al.*, 2003a; Goto *et al.*, 2005). To circumvent this potential risk, *nifH* phylotypes may be quantified using QPCR (Hewson *et al.*, 2007). We specifically designed primers and probes for the EVHVF cluster (100% aa similarity to *Burkholderia vietnamiensis*; AM110708; Figure 2) known to affiliate with reported contaminants (Farnelid *et al.*, 2009). The distribution of EVHVF was patchy among samples and sampling years with generally low abundance (Figure 4 and Supplementary Table S3) suggesting that its presence was not due to contamination of extraction or PCR reagents. Nevertheless,

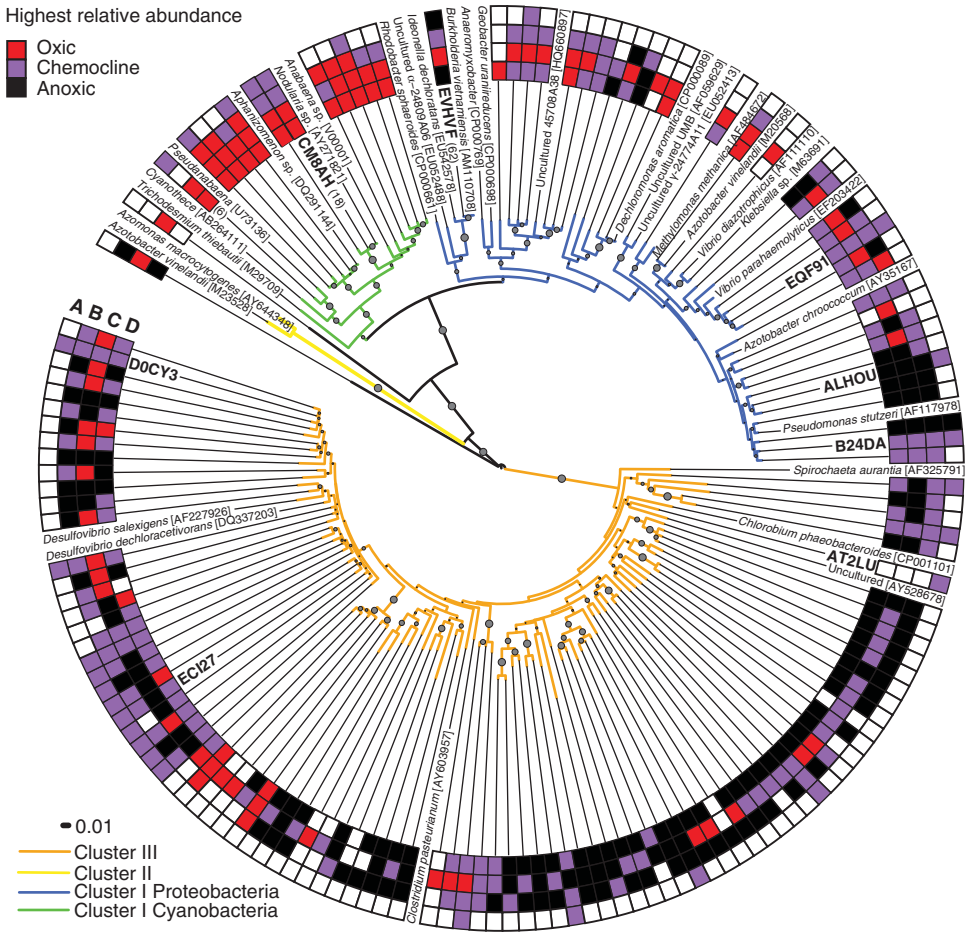

Figure 2 Neighbor-joining phylogenetic tree of 96% aa similarity clustered *nifH* sequences with affiliating canonical *nifH* clusters (Chien and Zinder, 1996) indicated by colored branches. The highest relative abundance was calculated by normalizing the number of sequences in each cluster and is shown as being in the oxic (red), at the chemocline (purple) or in anoxic waters (black), derived from the O_2 concentrations in Figure 1, for each type of sample; 2007 cDNA $<10\,\mu m$ (**A**), 2007 DNA $>10\,\mu m$ (**B**), 2008 DNA $<10\,\mu m$ (**C**) and 2007 DNA $<10\,\mu m$ (**D**). White indicates that no sequences within the cluster were detected. Accession numbers for reference sequences are indicated in squared brackets and clusters targeted by QPCR in this study are marked in bold. Bootstrap values (1000 replicates) >50% are indicated with size proportional gray circles. For clarity, only clusters with >100 sequences are included in the tree. The EVHVF, CM8AH and *Pseudanabaena*-like clusters were collapsed, respectively, and the numbers of subclusters are indicated in parentheses. The full branching of the subclusters is shown in Supplementary Figure S2A.

the high relative abundance of EVHVF in three of the cDNA libraries (Supplementary Figure S2A) does not mirror the abundance suggested by QPCR. Similarly, the AT2LU cluster (99% aa similarity to AY223933; Zehr *et al.*, 2003b) could not be quantified in any of the samples although it was present in the 129 m cDNA library (2731 sequences; Figure 2 and Supplementary Figure S2B). Hence, in concordance with recent findings (Turk *et al.*, 2011), some *nifH* phylotypes may be preferentially amplified. In addition, the coverage of degenerate *nifH* primers differs between phylogenetic groups (Gaby and Buckley, 2012). Therefore, the relative proportions of clusters in sequence libraries should be interpreted with caution.

Of the clusters quantified, only *nifH* from the cyanobacterium *Nodularia sp.* (CM8AH) was present in surface samples. Abundances were up to 1.9×10^6 copies l^{-1} and transcripts up to 2.1×10^5 copies l^{-1} (Figure 4 and Supplementary Table S3), which is comparable to a previous report from the Baltic Sea proper (Boström *et al.*, 2007b). N_2 fixation by *Nodularia* sp. is likely limited to illuminated waters and, therefore, *nifH* gene copies from these cyanobacteria at and below the chemocline (Figure 4 and Supplementary Table S3) is probably derived from settling senescing cells.

Two *gammaproteobacteria*-like *nifH* clusters, EQF91 and ALHOU, were targeted by QPCR (Figure 2 and Supplementary Table S2). EQF91 was most closely related to clones from hypoxic water in the Southern Californian Bight (97% aa similarity; HQ660860; Hamersley *et al.*, 2011) suggesting that these phylotypes may be characteristic for suboxic regions. Accordingly, EQF91 was never detected above 100 m at the Gotland Deep or above 68 m at the Bornholm Basin, but was

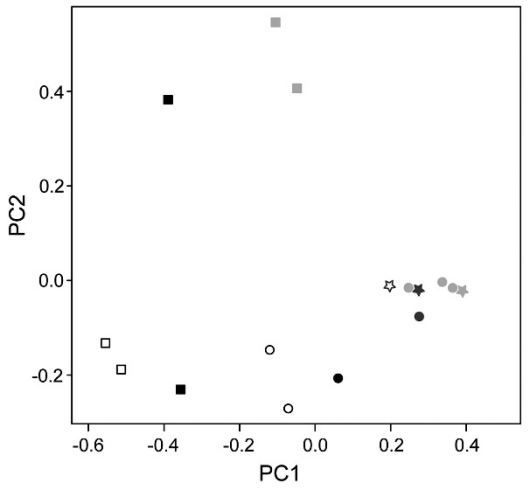

Figure 3 Principal coordinates analysis (PCoA) plot based on the Bray–Curtis distances calculated between samples. The samples are indicated as being in the oxic (squares), at the chemocline (circles) or in anoxic waters (stars), derived from the O_2 concentrations in Figure 1. Symbols for samples from 2007 and 2008 are black and gray, respectively. Open symbols show samples from the $>10\,\mu m$ size fraction.

consistently present at and below the chemocline, with abundances up to 5.2×10^5 copies l^{-1} (Figures 4 and 5 and Supplementary Table S3). On the 18 July 2011 at 119 m, EQF91 transcripts were detected but below the limit of quantification (Figure 4). The ALHOU cluster was related to *Pseudomonas stutzeri* (96% aa similarity) and a prevalent cluster from the South Pacific OMZ (93% aa similarity; HM801245; Fernandez *et al.*, 2011). The abundance of ALHOU reached up to 3.0×10^6 copies l^{-1}. The cluster was consistently present in suboxic/anoxic waters in all sampling profiles but transcription was only detected at 88 m in the Bornholm Basin (Figures 4 and 5 and Supplementary Table S3). However, as transcripts within the ALHOU and B24DA clusters were present in both the 129 and 132 m cDNA libraries (Supplementary Figure S2), constraints on the amount of template cDNA or disturbances during sampling (see Feike *et al.*, 2012) may explain why transcripts were not more frequently detected or quantified.

Three clusters within *nifH* cluster III were targeted by QPCR (Figure 2). The D0CY3 cluster was 99% similar (aa) to sequences from North American coastal waters (Mulholland *et al.*, 2012) while the ECI27 cluster was 100% similar (aa) to *nifH* clones (JN638704) from an unpublished study of the redoxcline of the Black Sea. Very high abundances of the D0CY3 (up to 3.3×10^6 copies l^{-1}) and ECI27 (up to 2.2×10^7 copies l^{-1}) clusters were observed at and below the chemocline (Figure 4 and Supplementary Table S3). These abundances are among the highest reported (but see Halm *et al.*, 2012; Mulholland *et al.*, 2012) and magnitudes higher than previous reports on cluster III phylotypes in the Pacific and Atlantic oceans (Church *et al.*, 2005; Langlois *et al.*, 2008). The activity of anaerobic N_2-fixers is largely unknown but interestingly, transcripts of a cluster III phylotype were recently quantified to 1.9×10^3 copies l^{-1} at 200 m in the South Pacific Gyre (Halm *et al.*, 2012). In our study, transcripts of D0CY3 (3.2×10^4 and 6.5×10^3 copies l^{-1}) at the two consecutive samplings for 200 m Gotland Deep samples (Figure 4 and Supplementary Table S3) and the co-occurrence of genes and transcripts also at the Bornholm Basin (88 m, Figure 5) further highlights the potential significance of cluster III phylotypes as N_2-fixers in suboxic/anoxic waters. In summary, the QPCR results show that abundances of several heterotrophic phylotypes increased dramatically at and below the chemocline. *nifH* transcripts of both *Proteobacteria* and cluster III were also detected indicating that bacteria within these groups are active in N_2 fixation (Figures 4 and 5 and Supplementary Table S3).

N fixation

Rates of N_2 fixation were measured in surface (5 m) and deep waters (200 m) of the Gotland Deep on 14 July 2011 using the traditional $^{15}N_2$-incorporation method; an approach yielding conservative estimates of N_2 fixation (Mohr *et al.*, 2010). The mean rate of N_2 fixation in surface water was 7.6 ± 1.76 nmol l^{-1} day^{-1} of which 2.0 ± 0.74 nmol l^{-1} day^{-1} was in the $<10\,\mu m$ size fraction. As some *Nodularia* cells may occur solitary or in short series (Lehtimäki *et al.*, 2000) and *Pseudanabaena* may pass a 10 μm filter (Stal *et al.*, 2003; Farnelid *et al.*, 2009), N_2 fixation in the $<10\,\mu m$ size fraction is not evidence for heterotrophic N_2 fixation. In the anoxic sample, the mean rate was 0.44 ± 0.26 nmol l^{-1} day^{-1} of which 0.24 ± 0.26 nmol l^{-1} day^{-1} was in the $<10\,\mu m$ size fraction. These rates are comparable to those of hypoxic waters in the Southern Californian Bight (average 0.07 nmol l^{-1} day^{-1}; Hamersley *et al.*, 2011) and the Peruvian OMZ (average 1.27 nmol l^{-1} day^{-1}; Fernandez *et al.*, 2011). In the same water sample, transcription of the D0CY3 cluster was 3.2×10^4 copies l^{-1} (Figure 4a) strongly suggesting that these anaerobic bacteria were largely responsible for the N_2 fixation at 200 m.

In anoxic water, N_2 fixation took place at a NH_4^+ concentration of 19.7 μM (Figure 1). Similarly, N_2 fixation has been measured off Peru and in a meromictic lake at measurable NH_4^+ levels (Halm *et al.*, 2009; Fernandez *et al.*, 2011). In the laboratory, NH_4^+ downregulates or shuts down nitrogenase expression (for example, Kanemoto and Ludden, 1984; Klugkist and Haaker, 1984). However, the high NH_4^+ concentration at 200 m did not reduce the abundance or inhibit transcription of ECI27 and D0CY3 (Figure 4a). Similarly, N_2 fixation and/or transcripts (notably within *nifH* cluster III) were reported from NH_4^+-rich microbial mats (Steppe and Paerl, 2002; Omoregie *et al.*, 2004), representing a natural

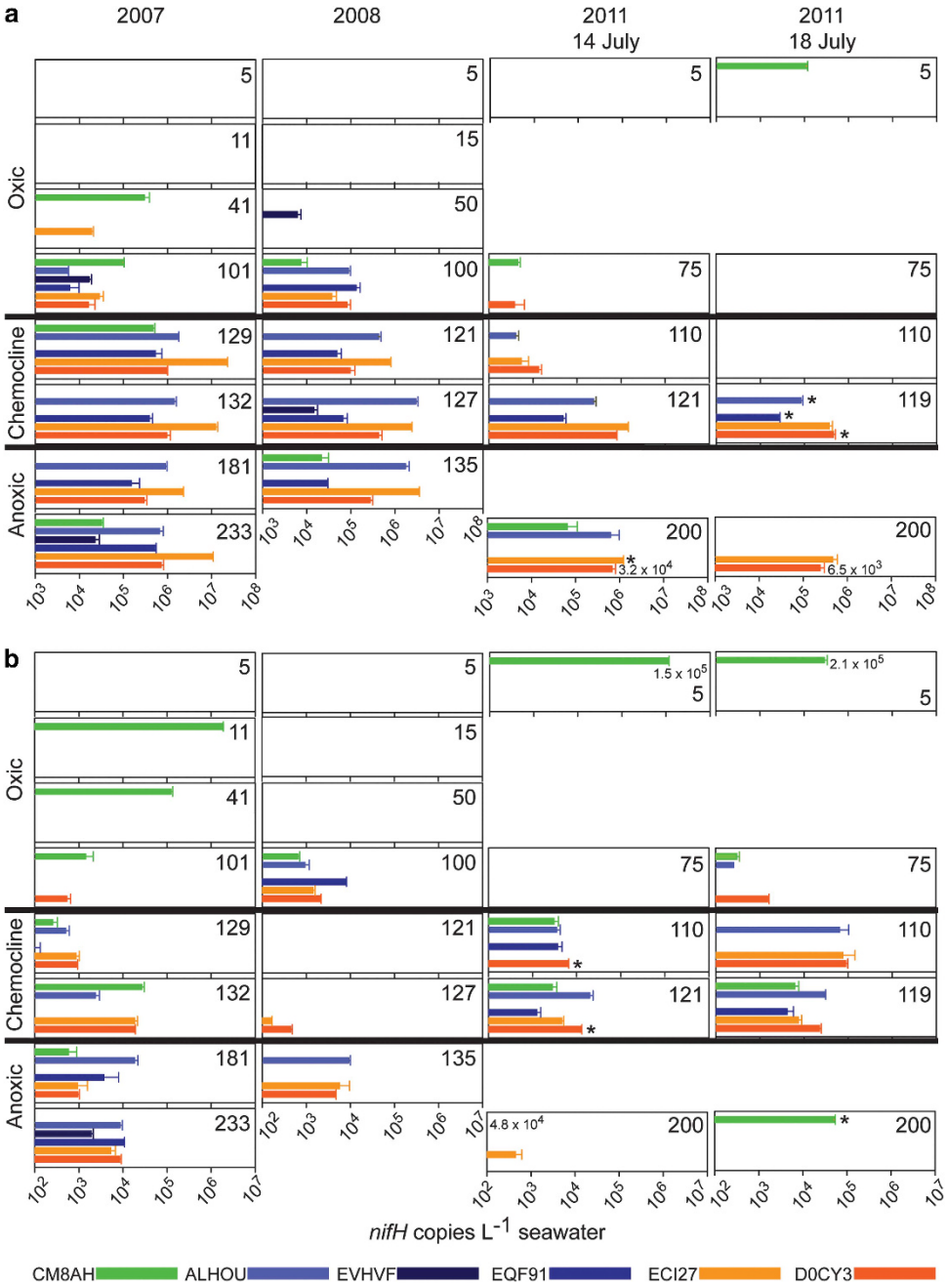

Figure 4 Abundances of specific *nifH* gene clusters in the free-living (<10 μm) fractions (**a**) and the particle-associated (>10 μm) fractions (**b**) determined by QPCR at the Gotland Deep. Cluster AT2LU is not shown as it was never quantifiable. Detection of transcripts is indicated with a star and when quantified, the number of transcripts l^{-1} is indicated next to the horizontal bars. Sample depth (m) is indicated in each panel. Based on data shown in Figure 1, the profiles are divided into the oxic, the chemocline or the anoxic depth strata. Note the different scales on the X axis for **a** and **b**.

microscale analog of hypoxic waters columns (Zehr *et al.*, 2006). Sustained N_2 fixation following NH_4^+ amendment has also been observed in a salt marsh (Hanson, 1977) and a gammaproteobacterial culture (Boström *et al.*, 2007a). Thus, despite the high energy demand of N_2 fixation compared with NH_4^+ uptake, availability of NH_4^+ does not always preclude N_2 fixation (reviewed in Knapp, 2012). An intriguing explanation could be that some nitrogenases, because of their early origin before the oxygenation of the atmosphere when high ambient levels of NH_4^+ prevailed (Raymond *et al.*, 2004), may not be tightly regulated by NH_4^+ (Zehr *et al.*, 2006). Nonetheless, the maintenance of N_2 fixation rates and transcripts in NH_4^+-replete waters highlights the incomplete understanding of the biology of these organisms, and underlines the need for surveys of diazotrophic activities in marine environments hitherto believed to be void of diazotrophy.

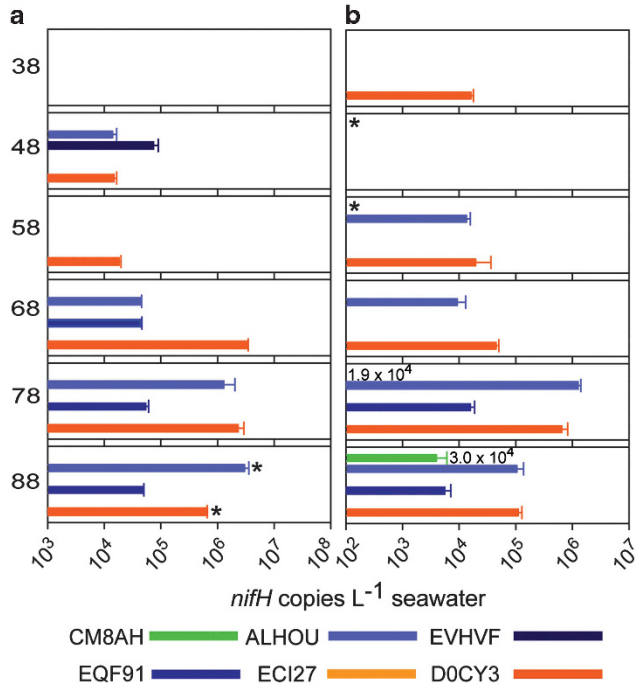

Figure 5 Abundances of specific *nifH* gene clusters in the free-living (<10 μm) fractions (**a**) and the particle-associated (>10 μm) fractions (**b**) determined by QPCR at the Bornholm Basin. Cluster AT2LU is not shown as it was never quantifiable. Detection of transcripts is indicated with a star and when quantified, the number of transcripts l^{-1} is indicated next to the horizontal bars. Sample depth (m) is indicated to the left. Note the different scales on the X axis for **a** and **b**.

Implications of heterotrophic N_2 fixation in suboxic waters

In this study, we demonstrate that suboxic and anoxic waters of the Baltic Sea harbor diverse and active heterotrophic N_2-fixing communities. The chemocline of the Gotland Deep features high microbial activity (Höfle and Brettar, 1995), dark CO_2 fixation (Jost *et al.*, 2008) and loss of inorganic N (Hietanen *et al.*, 2012). The growth of bacteria involved in N-loss processes is often carbon limited (Lam and Kuypers, 2011). It is thus perplexing to find microorganisms investing in the expensive process of N_2 fixation under conditions that are N-replete and most probably carbon limited.

N_2 fixation in the anoxic waters was low compared with surface waters (Larsson *et al.*, 2001). However, unlike cyanobacteria, heterotrophic diazotrophs are not limited by light or water temperature. Hence, they could potentially fix N_2 in a large pelagic realm. Several of the *nifH* phylotypes that we quantified were abundant in suboxic/anoxic waters during all sampling years and in both basins, ~340 km apart (Figures 4 and 5). Accordingly, if consistently present, they may occupy a volume of up to 3000 km^3 in the Baltic Sea (<2 ml O_2 l^{-1}; Hansson *et al.*, 2011), and account for a fixation of up to 11 Gg N $year^{-1}$, equal to up to 6% of the N fixed yearly in surface waters of the Baltic Sea proper (Larsson *et al.*, 2001), if assuming that our rate measured at 200 m is representative. Although this rough and speculative calculation merely illustrates the large potential of heterotrophic diazotrophs, determining their exact contribution to N input is difficult because of their generally low concentration, the very low transcript copy numbers, and accordingly, rather low rates of N_2 fixation. Still, heterotrophic N_2-fixers accounted for up to 2.1% of the total bacteria at the chemocline (Supplementary Figure S1). In comparison, biogeochemically significant anammox bacteria ranged from 2.2% to 3.1% of the bacterial community in the Peruvian OMZ (Hamersley *et al.*, 2007). Hence, future N_2 fixation measurements with extensive spatio-temporal resolution including also the deeper suboxic and anoxic zones of the Baltic Sea, could potentially reveal integrated rates of high N budgetary significance.

Heterotrophic diazotrophs are almost ubiquitous in marine surface waters (Riemann *et al.*, 2010; Farnelid *et al.*, 2011) and occur in OMZs (Fernandez *et al.*, 2011; Hamersley *et al.*, 2011; Jayakumar *et al.*, 2012) around the globe. Combined with the present extension of their niches to NH_4^+-rich sulfidic-anoxic waters of one of the largest brackish waters on earth, the prevailing perception of vertical and geographical boundaries for marine N_2 fixation has been dramatically challenged; thus, we foresee a future broadening of the marine N_2 fixation regime with potentially large consequences for N biogeochemistry estimates.

Acknowledgements

We thank the captains and crews of the cruises 2007: RV Professor A Penck, 2008: RV Poseidon and 2011: RV Elisabeth Mann-Borgese. This work was supported by the Swedish Research Council FORMAS 217-2006-342 to LR and the K&A Wallenberg foundation via the former Center for Metagenomic Sequence Analysis at KTH to LR and SB. AFA was supported by grants from the Swedish Research Councils VR and FORMAS.

References

Boström KH, Riemann L, Kühl M, Hagström Å. (2007a). Isolation and gene quantification of heterotrophic N_2-fixing bacterioplankton in the Baltic Sea. *Environ Microbiol* **9**: 152–164.

Boström KH, Riemann L, Zweifel UL, Hagström Å. (2007b). *Nodularia* sp. *nifH* gene transcripts in the Baltic Sea proper. *J Plankton Res* **29**: 391–399.

Braun ST, Proctor LM, Zani S, Mellon MT, Zehr JP. (1999). Molecular evidence for zooplankton-associated nitrogen-fixing anaerobes based on amplification of the *nifH* gene. *FEMS Microbiol Ecol* **28**: 273–279.

Chao A. (1984). Nonparametric estimation of the number of classes in a population. *Scand J Stat* **11**: 265–270.

Chien Y-T, Zinder SH. (1996). Cloning, functional organization, transcript studies, and phylogenetic analysis of

the complete nitrogenase stuctural genes (*nifHDK2*) and associated genes in the Archaeon *Methanosarcina barkeri* 227. *J Bacteriol* **178**: 143–148.

Church MJ, Jenkins BD, Karl DM, Zehr JP. (2005). Vertical distributions of nitrogen-fixing phylotypes at Stn ALOHA in the oligotrophic North Pacific Ocean. *Aquat Microb Ecol* **38**: 3–14.

Conley DJ, Humborg C, Rahm L, Savchuk OP, Wulff F. (2002). Hypoxia in the Baltic Sea and basin-scale changes in phosphorus biogeochemistry. *Environ Sci Technol* **36**: 5315–5320.

Dixon R, Kahn D. (2004). Genetic regulation of biological nitrogen fixation. *Nature Rev Microbiol* **2**: 621–631.

Farnelid H, Andersson AF, Bertilsson S, Al-Soud WA, Hansen LH, Sørensen S *et al.* (2011). Nitrogenase gene amplicons from global marine surface waters are dominated by genes of non-cyanobacteria. *PLoS One* **6**: e19223.

Farnelid H, Öberg T, Riemann L. (2009). Identity and dynamics of putative N_2-fixing picoplankton in the Baltic Sea proper suggest complex patterns of regulation. *Env Microbiol Rep* **1**: 145–154.

Feike J, Jürgens K, Hollibaugh JT, Krüger S, Jost G, Labrenz M. (2012). Measuring unbiased metatranscriptomics in suboxic waters of the central Baltic Sea using a new *in situ* fixation system. *ISME J* **6**: 461–470.

Fernandez C, Farías L, Ulloa O. (2011). Nitrogen fixation in denitrified marine waters. *PLoS ONE* **6**: e20539.

Gaby JC, Buckley DH. (2012). A comprehensive evaluation of PCR primers to amplify the *nifH* gene of nitrogenase. *PLoS One* **7**: e42149.

Goebel NL, Edwards CA, Church MJ, Zehr JP. (2007). Modeled contributions of three types of diazotrophs to nitrogen fixation at Station ALOHA. *ISME J* **1**: 606–619.

Goto M, Ando S, Hachisuka Y, Yoneyama T. (2005). Contamination of diverse *nifH* and *nifH*-like DNA into commercial PCR primers. *FEMS Microbiol Lett* **246**: 33–38.

Grasshoff K, Erhardt M, Kremling K. (1983). *Methods of Seawater Analysis*, 2nd edn. Verlag Chemie: Weinheim, pp 125–157.

Gruber N. (2005). A bigger nitrogen fix. *Nature* **436**: 786–787.

Halm H, Lam P, Ferdelman TG, Lavik G, Dittmar T, LaRoche J *et al.* (2012). Heterotrophic organisms dominate nitrogen fixation in the South Pacific Gyre. *ISME J* **6**: 1238–1249.

Halm H, Musat N, Lam P, Langlois R, Musat F, Peduzzi S *et al.* (2009). Co-occurence of denitrification and nitrogen fixation in a meromictic lake, Lake Cadagno (Switzerland). *Env Microbiol* **11**: 1945–1958.

Hamersley R, Lavik G, Woebken D, Rattray JE, Lam P, Hopmans EC *et al.* (2007). Anaerobic ammonium oxidation in the Peruvian oxygen minimum zone. *Limnol Oceanogr* **52**: 923–933.

Hamersley R, Turk KA, Leinweber A, Gruber N, Zehr JP *et al.* (2011). Nitrogen fixation within the water column associated with two hypoxic basins within the Southern California Bight. *Aquat Microb Ecol* **63**: 193–205.

Hanson RB. (1977). Nitrogen fixation (acetylene reduction) in a salt marsh amended with sewage sludge and organic carbon and nitrogen compounds. *Appl Environ Microbiol* **33**: 846–852.

Hansson M, Andersson L, Axe P. (2011). Areal extent and volume of anoxia and hypoxia in the Baltic Sea, 1960–2011. SMHI Report Series: Report Oceanography No. 42.

Hewson I, Moisander P, Achilles K, Carlson C, Jenkins BD, Mondragon E *et al.* (2007). Characteristics of diazotrophs in surface to abyssopelagic waters of the Sargasso Sea. *Aquat Microb Ecol* **46**: 15–30.

Hietanen S, Jäntti H, Buizert C, Jürgens K, Labrenz M, Voss M *et al.* (2012). Hypoxia and nitrogen processing in the Baltic Sea water column. *Limnol Oceanogr* **57**: 325–337.

Höfle MG, Brettar I. (1995). Taxonomic diversity and metabolic activity of microbial communities in the water column of the central Baltic Sea. *Limnol Oceanogr* **40**: 868–874.

Jayakumar A, Al-Rshaidat MMD, Ward BB, Mulholland MR. (2012). Diversity, distribution, and expression of diazotroph *nifH* genes in oxygen-deficient waters of the Arabian Sea. *FEMS Microbiol Ecol* **82**: 597–606.

Jost G, Zubkov MV, Yakushev E, Labrenz M, Jürgens K. (2008). High abundance and dark CO_2 fixation of chemolithoautotrophic prokaryotes in anoxic waters of the Baltic Sea. *Limnol Oceanogr* **53**: 14–22.

Kanemoto RH, Ludden PW. (1984). Effect of ammonia, darkness, and phenazine methosulfate on whole-cell nitrogenase activity and Fe protein modification in *Rhodospirillum rubrum*. *J Bacteriol* **158**: 713–720.

Klugkist J, Haaker H. (1984). Inhibition of nitrogenase activity by ammonium chloride in *Azotobacter vinelandii*. *J Bacteriol* **157**: 148–151.

Knapp A. (2012). The sensitivity of marine N_2 fixation to dissolved inorganic nitrogen. *Front Microbiol* **3**: 374.

Lam P, Kuypers MMM. (2011). Microbial nitrogen cycling processes in oxygen minimum zones. *Annu Rev Mar Sci* **3**: 317–345.

Langlois RJ, Hümmer D, La Roche J. (2008). Abundances and distributions of the dominant *nifH* phylotypes in the Northern Atlantic Ocean. *Appl Environ Microbiol* **74**: 1922–1931.

Larsson U, Hajdu S, Walve J, Elmgren R. (2001). Baltic Sea nitrogen fixation estimated from the summer increase in upper mixed layer total nitrogen. *Limnol Oceanogr* **46**: 811–820.

Lehtimäki J, Lyra C, Suomalainen S, Sundman P, Rouhiainen L, Paulin L *et al.* (2000). Characterization of *Nodularia* strains, cyanobacteria from brackish waters, by genotypic and phenotypic methods. *Int J Syst Evol Micr* **50**: 1043–1053.

Li W, Godzik A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**: 1658–1659.

Lundin D, Severin I, Logue JB, Östman Ö, Andersson AF, Lindström ES. (2012). Which sequencing depth is sufficient to describe patterns in bacterial α- and β-diversity? *Environ Microbiol Rep* **4**: 367–372.

Man-Aharonovich D, Kress N, Zeev EB, Berman-Frank I, Beja O. (2007). Molecular ecology of *nifH* genes and transcripts in the eastern Mediterranean Sea. *Environ Microbiol* **9**: 2354–2363.

Mohr W, Großkopf T, Wallace DWR, LaRoche J. (2010). Methodological underestimation of oceanic nitrogen fixation rates. *PLoS ONE* **5**: e12583.

Montoya JP, Voss M, Kähler P, Capone DG. (1996). A simple, high-precision, high-sensitivity tracer assay for N_2 fixation. *Appl Environ Microbiol* **62**: 986–993.

Mulholland MR, Bernhardt PW, Blanco-Garcia JL, Mannino A, Hyde K, Mondragon E *et al.* (2012). Rates of dinitrogen fixation and the abundance of diazotrophs in North American coastal waters

between Cape Hatteras and Georges Bank. *Limnol Oceanogr* **57**: 1067–1083.

Omoregie EO, Crumbliss LL, Bebout BM, Zehr JP. (2004). Determination of nitrogen-fixing phylotypes in *Lyngbya* sp. and *Microcoleus chthonoplastes* cyanobacterial mats from Guerrero Negro, Baja California, Mexico. *Appl Environ Microbiol* **70**: 2119–2128.

Raymond J, Siefert JL, Staples CR, Blankenship RE. (2004). The natural history of nitrogen fixation. *Mol Biol Evol* **21**: 541–554.

Riemann L, Farnelid H, Steward GF. (2010). Nitrogenase genes in non-cyanobacterial plankton: prevalence, diversity and regulation in marine waters. *Aquat Microb Ecol* **61**: 235–247.

Riemann L, Leitet C, Pommier T, Simu K, Holmfeldt K, Larsson U *et al.* (2008). The native bacterioplankton community in the central Baltic Sea is influenced by freshwater bacterial species. *Appl Environ Microbiol* **74**: 503–515.

Shannon CE. (1948). A mathematical theory of communication. *Bell System Tech J* **27**: 379–423.

Stal LJ, Albertano P, Bergman B, von Bröckel K, Gallon JR, Hayes PK *et al.* (2003). BASIC: Baltic Sea cyanobacteria. An investigation of the structure and dynamics of water blooms of cyanobacteria in the Baltic Sea - responses to a changing environment. *Cont Shelf Res* **23**: 1695–1714.

Steppe TF, Paerl HW. (2002). Potential N_2 fixation by sulfate-reducing bacteria in a marine intertidal microbial mat. *Aquat Microb Ecol* **28**: 1–12.

Tamura K, Dudley J, Nei M, Kumar S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**: 1596–1599.

Troussellier M, Courties C, Lebaron P, Servais P. (1999). Flow cytometric discrimination of bacterial populations in seawater based on SYTO 13 staining of nucleic acids. *FEMS Microbiol Ecol* **29**: 319–330.

Turk KA, Rees AP, Zehr JP, 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.

Ulloa O, Canfield DE, Delong EF, Letelier RM, Stewart FJ. (2012). Microbial oceanography of anoxic oxygen minimum zones. *Proc Natl Acad Sci USA* **109**: 15996–16003.

Zehr JP, Church MJ, Moisander P. (2006). Diversity, distribution and biogeochemical significance of nitrogen-fixing microorganisms in anoxic and suboxic environments. In: Neretin LN (ed) *Past and Present Water Column Anoxia*. Springer: AZ Dordrecht, The Netherlands, pp 337–369.

Zehr JP, Crumbliss LL, Church MJ, Omoregie EO, Jenkins BD. (2003a). Nitrogenase genes in PCR and RT-PCR reagents: implications for studies of diversity of functional genes. *BioTechniques* **35**: 996–1005.

Zehr JP, Jenkins BD, Short SM, Steward GF. (2003b). Nitrogenase gene diversity and microbial community structure: a cross-system comparison. *Environ Microbiol* **5**: 539–554.

Zehr JP, McReynolds LA. (1989). Use of degenerate oligonucleotides for amplification of the *nifH* gene from the marine cyanobacterium *Trichodesmium thiebautii*. *Appl Environ Microbiol* **55**: 2522–2526.

Zehr JP, Turner PJ. (2001). Nitrogen fixation: nitrogenase genes and gene expression. In: Paul JH (ed) *Methods in Microbiology*. Academic Press: New York, USA, pp 271–285.

Supplementary Information accompanies the paper on The ISME Journal website (http://www.nature.com/ismej)