environmental microbiology

Environmental Microbiology (2017) **00**(00), 000–000 doi:10.1111/1462-2920.14007

A denitrifying community associated with a major, marine nitrogen fixer

Christopher J. Coates[1,2] and Michael Wyman[1]*

[1]*Biological and Environmental Sciences, Faculty of Natural Sciences, University of Stirling, Stirling FK9 4LA, Scotland, UK.*

[2]*Department of Biosciences, College of Science, Swansea University, Swansea SA2 8PP, Wales, UK.*

Summary

The diazotrophic cyanobacterium, *Trichodesmium*, is an integral component of the marine nitrogen cycle and contributes significant amounts of new nitrogen to oligotrophic, tropical/subtropical ocean surface waters. *Trichodesmium* forms macroscopic, fusiform (tufts), spherical (puffs) and raft-like colonies that provide a pseudobenthic habitat for a host of other organisms including marine invertebrates, microeukaryotes and numerous other microbes. The diversity and activity of denitrifying bacteria found in association with the colonies was interrogated using a series of molecular-based methodologies targeting the gene encoding the terminal step in the denitrification pathway, nitrous oxide reductase (*nosZ*). *Trichodesmium* spp. sampled from geographically isolated ocean provinces (the Atlantic Ocean, the Red Sea and the Indian Ocean) were shown to harbor highly similar, taxonomically related communities of denitrifiers whose members are affiliated with the *Roseobacter* clade within the Rhodobacteraceae (Alphaproteobacteria). These organisms were actively expressing *nosZ* in samples taken from the mid-Atlantic Ocean and Red Sea implying that *Trichodesmium* colonies are potential sites of nitrous oxide consumption and perhaps earlier steps in the denitrification pathway also. It is proposed that coupled nitrification of newly fixed N is the most likely source of nitrogen oxides supporting nitrous oxide cycling within *Trichodesmium* colonies.

Received 4 July, 2017; revised 15 November, 2017; accepted 21 November, 2017. *For correspondence. E-mail michael.wyman@stir.ac.uk; Tel. (+44) 1786 467784; Fax (+44) 1786 467843.

Introduction

It is over twenty years since Louis Codispoti posed the question 'Are the oceans losing nitrate?' (Codispoti, 1995). A simple enough query to have merited an unequivocal answer by now, perhaps, but our conceptual grasp on the biological processes involved in adding to or removing fixed N from the oceans has undergone several major revolutions since. We know now that denitrification (the reduction of nitrate/nitrite to dinitrogen gas) is just one biochemical route by which fixed N is lost to the atmosphere. Indeed anammox (the anaerobic oxidation of ammonium fuelled by nitrite) may dominate N losses in ODZs (oxygen deficient zones) in some regions (Voss *et al.*, 2013). Likewise, new inputs of fixed N via biological fixation are more diverse in their sources and wider in their overall oceanic distribution than was once believed (Farnelid *et al.*, 2011). Most recently, nitrifiers capable of conserving energy by oxidizing ammonium completely to nitrate (rather than only as far as nitrite) have been isolated from wastewater and aquaculture recirculating systems (Daims *et al.*, 2015; van Kessel *et al.*, 2015). Signature genes characteristic of these organisms are present in soils and freshwaters but have not been detected to date in marine metagenomic libraries (Daims *et al.*, 2015), although their presence has been predicted from energetic considerations (Kuypers, 2015).

With this new found complexity comes an even greater impetus to constrain the marine N budget. The concern that lies at the heart of Codispoti's original question is itself straightforward but, in the context of global change, could not be more important. If the oceans are indeed losing nitrate then (i) their future productivity will diminish, (ii) their capacity to drawdown atmospheric CO_2 will decline and (iii) their vulnerability to large-scale damage through acidification will intensify. Current estimates of annual marine denitrification rates (that include a large sediment component as well as anammox) range from 230–240 to in excess of 400 Tg N whereas pelagic N fixation adds just 103–177 Tg N per annum (Codispoti, 2007; Canfield *et al.*, 2010; De Vries *et al.*, 2012; Großkopf *et al.*, 2012). A somewhat higher estimate of N fixation rates (285 Tg N per year) adopted in one modeling study (De Vries *et al.*, 2012) produces a more or less balanced N budget at the

lower end of the range of marine denitrification rates. More conservative estimates of biological N inputs, including those revised upward to correct for past methodological inaccuracies (Großkopf *et al*., 2012), imply a deficit of at least 53 Tg N if one assumes pelagic N losses occur solely in ODZs. In arriving at a figure in excess of 400 Tg N per year (and a deficit of $>$ 230 Tg N), however, Codispoti (2007) envisaged low rates of denitrification in oxygen-deficient regions within the $>$ 99% of the ocean volume that lies outside of the three major ODZs.

While this large expanse of ocean is generally well ventilated, potential suboxic microsites that might support active denitrification are known and include marine snow particles, zooplankton guts and fecal pellets, for example. Heterotrophic activity associated with these organically rich microenvironments can reduce internal oxygen tensions sufficiently to enable oxygen-sensitive processes to proceed, as has been demonstrated recently for both denitrification and N fixation in association with copepods (Glud *et al*., 2015; Scavotto *et al*., 2015). Earlier work from this laboratory showed that the colonial cyanobacterium, *Trichodesmium*, also harbors denitrifying bacteria (Wyman *et al*., 2013): a finding that hints at a more intimate spatial coupling of oceanic nitrogen sources and sinks than that envisaged by current models of ocean geochemistry (Deutsch *et al*., 2007). *Trichodesmium* accounts for up to 50% of biologically fixed N inputs to the ocean: that is, $\sim$ 70–90 Tg N *per annum* (Mahaffey *et al*., 2005). If at least some of this fixed N is denitrified *in situ*, then the potential losses of N are considerable given the wide distribution of *Trichodesmium* in the global ocean (Capone *et al*., 1997; Bergman *et al*., 2013). Critically, high rates of endogenous respiration within *Trichodesmium* colonies can promote the development of very low oxygen tensions at their center in subdued light and at night (Paerl and Bebout, 1988). The colony microenvironment may not always become suboxic under these conditions, however, and oxygen saturation state may vary with colony size, form (spherical 'puffs' or fusiform 'tufts' or rafts), and the metabolic activity of *Trichodesmium*, and its associated organisms, in particular (Eichner *et al*., 2017).

Perhaps ironically, up to 50–80% of the N fixed by *Trichodesmium* is liberated in the form of ammonium and dissolved organic N (including significant quantities of free amino acids) (Mulholland *et al*., 2006). Ammonium concentrations are elevated in the vicinity of *Trichodesmium* blooms (Karl *et al*., 1997), and the fate of this released fixed N is assumed to enhance the productivity of nearby organisms (Nausch, 1996) and the members of the colony consortium, in particular. Isotopic evidence shows that at least some of this liberated ammonium and remineralized N supports nitrification and the localized production of nitrogen oxides (Gandhi *et al*., 2010; Sutka *et al*., 2004). If just a tenth of the N released by *Trichodesmium* should be subsequently lost to the atmosphere through denitrification this would increase pelagic N losses (currently estimated to be $\sim$ 66 Tg N per year; De Vries *et al*., 2012) by 11–14%. Since any *Trichodesmium*-associated denitrification is most likely to occur within the colony under suboxic rather anoxic conditions, there is the potential for the end product to be nitrous oxide (N_2O) rather than dinitrogen. Nitrous oxide is a powerful greenhouse gas ($\sim$ 300-fold the warming potential of CO_2 over a 100-year period) and is presently the single most important source of emissions contributing to ozone depletion in the stratosphere (Ravishankara *et al*., 2009).

In this communication, we show that that the association between *Trichodesmium* and its epibiotic denitrifiers is global, specific and active. A cluster of related, denitrifying alphaproteobacteria are involved, they are invariably present, and they express at least one of the O_2-regulated genes (*nosZ*, encoding nitrous oxide reductase) of the denitrification pathway *in situ*.

Results and discussion

Amplification of nosZ by the PCR: Primer design and optimization

Preliminary experiments with *Trichodesmium* samples originating from the Arabian Sea showed that while *nosZ* could be amplified routinely from gDNA extracted from the colonies tested, the overall diversity of the amplicons retrieved was low when their DNA sequences were compared. This raised the concern that the *nosZ* PCR primer pair (*nosZ*F1 and *nosZ*R) developed by us previously (Wyman *et al*., 2013) may not capture the full complexity of the denitrifier community associated with *Trichodesmium*. To investigate this possibility, the target specificity and universality of the primers used was reassessed, and as a result, a new set of primers was designed (Supporting Information Table S1) that target a much broader range of organisms (see Supporting Information – NosZ sequence analysis and PCR primer design).

One set of the new primers targets organisms producing the typical Z-type NosZ protein (i.e., Clade I; Graf *et al*., 2014) found primarily among the Proteobacteria, while the second is designed to amplify genes encoding the atypical form of the enzyme (Clade II; Graf *et al*., 2014) that occurs in other Bacteria and Archaeal groups. The Z-type protein has been named the 'True Denitrifier' NosZ (Sanford *et al*., 2012). It is found in the majority ($\sim$ 90%) of genomes that also contain recognizable copies of either *nirK* or *nirS* (encoding dissimilatory nitrite reductase) and *nor* (encoding nitric oxide reductase) (Sanford *et al*., 2012; Graf *et al*., 2014). By contrast, the atypical form of NosZ occurs in a significant number of species that do not harbor other denitrification genes (Sanford *et al*., 2012; Graf *et al*., 2014). Accordingly, we limit our definition of denitrifiers to those

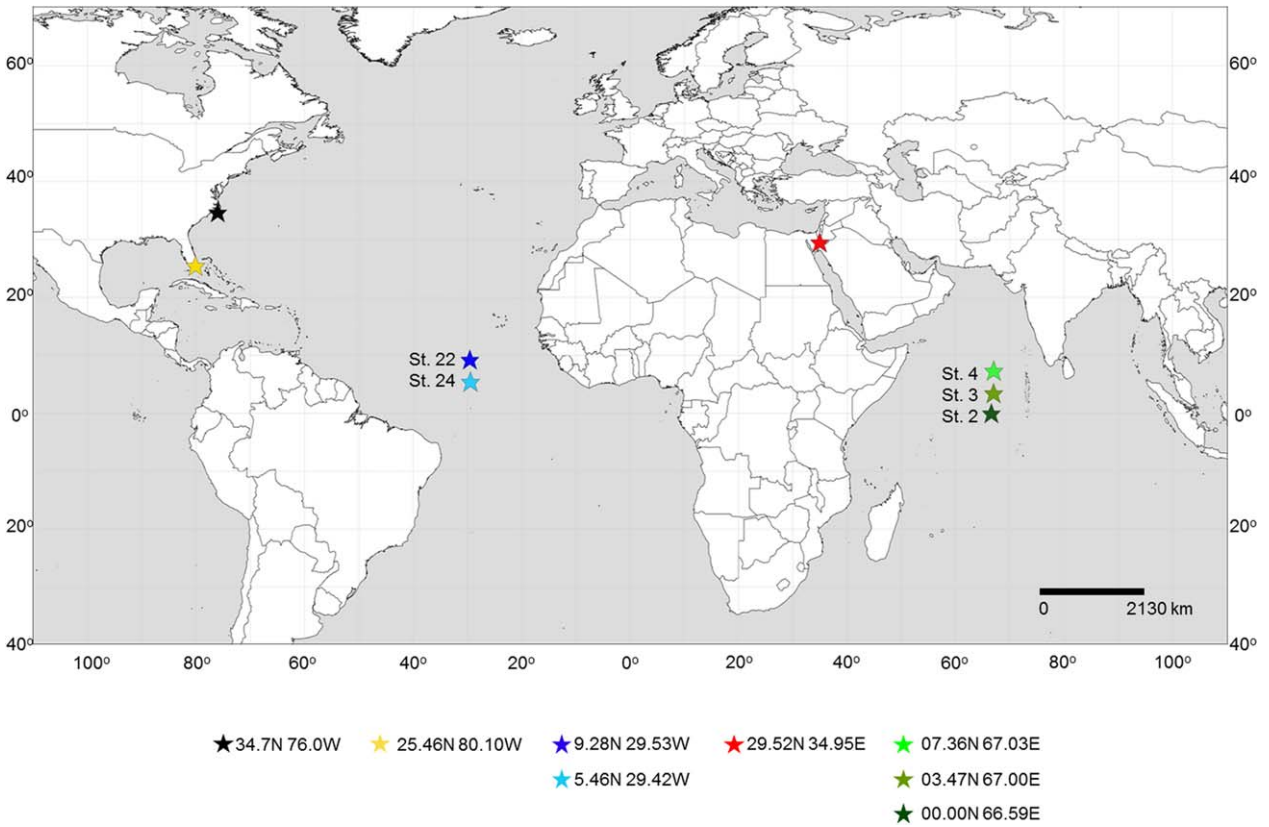

Fig. 1. Geographical locations and positions of the stations where DNA/RNA samples from *Trichodesmium* colonies were obtained for use in this study. The site sampled for the isolation of the *Trichodesmium* sp. strain IMS101 culture (Prufert-Bebout *et al.*, 1993) is also shown (black star).

organisms that produce the Z-type NosZ while recognizing that this excludes others (primarily those harboring *nirK*) capable of partial denitrification but which lack NosZ and, in some cases, Nor also (Graf *et al.*, 2014).

Diversity of the denitrifier community associated with Trichodesmium *spp*

Total genomic DNA was extracted from *Trichodesmium* colonies sampled at a number of tropical and subtropical sites from the world ocean and also from cultures of *Trichodesmium* sp. strain IMS101, originally isolated from the eastern seaboard of the USA (Fig. 1). The natural samples were all obtained from warm (26–30°C), stratified, oligotrophic, surface waters, highly depleted in combined nitrogen and characterized by low to very low chlorophyll concentrations ($<$ 0.1–0.2 mg m^{-3}). Further detail of the environmental conditions and microbial communities present at the Indian Ocean stations is given by Mazard and colleagues (2004), Bird and colleagues (2005), Zubkov and colleagues (2006) and Wyman and colleagues (2013). For an overview of the conditions and the microbial communities characteristic of the mid-Atlantic Ocean in late summer/autumn, see Schattenhofer and colleagues (2009) and Rees and colleagues (2015), and for the Gulf of Aqaba and wider Red Sea, see Rahav and colleagues (2015) and Pearman and colleagues (2017).

Fragments of *nosZ* were amplified from all samples interrogated and the peptide sequences (188 residues) derived from one hundred and fifty three *Trichodesmium*-associated *nosZ* gDNA clones were included in phylogenetic analysis (Fig. 2). The derived sequences shared considerable similarity ($>$ 95% mean peptide similarity; $>$ 80% DNA identity) and, without exception, they were all affiliated with the Z-type NosZ protein (i.e., the form of the enzyme found most frequently in complete denitrifiers; Sanford *et al.*, 2012) from members of the class, Alphaproteobacteria. In particular, the majority of the translated clones were similar to the NosZ from denitrifiers belonging to the order Rhodobacteraceae that includes the ubiquitous *Roseobacter* clade, and most closely among cultured isolates, to two representatives of the genus, *Labrenzia*: *Labrenzia aggregata* strain IAM 12614 and *Labrenzia*. sp. strain DG 1229 (Table 1).

Labrenzia (formerly *Stappia*) spp. have been isolated from a range of saline, benthic and planktonic habitats, and they frequently form close associations with other marine organisms (Weber and King, 2007). These

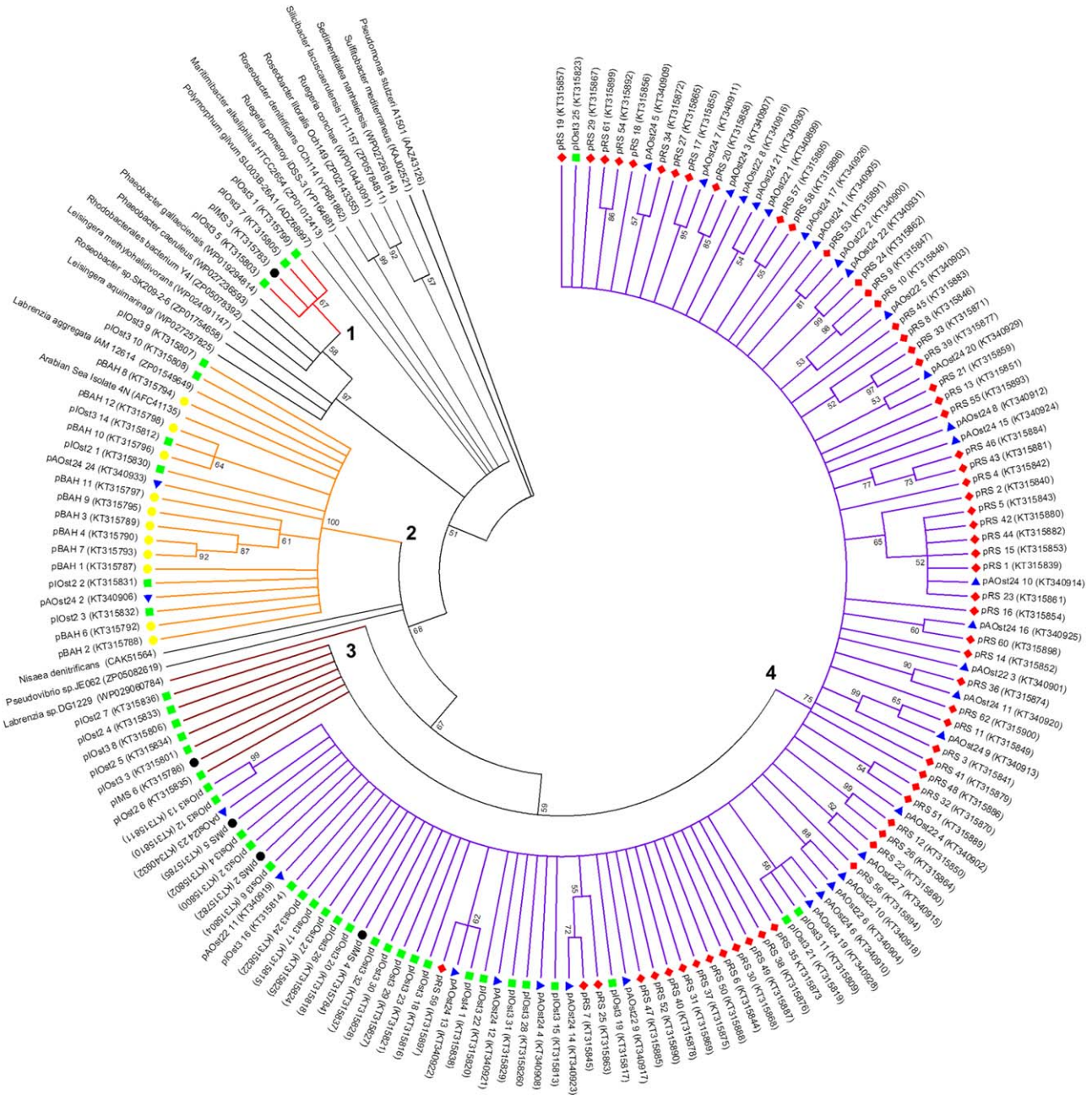

Fig. 2. Consensus cladogram (1000 bootstrap replicates) of translated NosZ peptide sequences derived from *Trichodesmium* colony total genomic DNA obtained from samples collected at two stations in the mid-Atlantic Ocean (clone prefix pAO followed by the station number; blue triangles), three stations in the Indian Ocean (clone prefix pIO followed by the station number; green squares), the Red Sea (clone prefix pRS; red diamonds), Bahama Islands (clone prefix pBAH; yellow circles) and from archived DNA from the isolate, *Trichodesmium* IMS101 (clone prefix pIMS; black circles). Groups 1 (red lines), 2 (orange lines), 3 (brown lines) and 4 (purple lines) are highlighted. The numbers at the nodes indicate the bootstrap support (%) received for each partition. The tree is rooted with the NosZ peptide sequence from *Pseudomonas stutzeri.*

associations include invertebrates such as molluscs, corals and sponges, and a variety of photosynthetic partners including seaweeds, diatoms, dinoflagellates, green and red algae. They are motile, mesophilic, halophytic, nonfermentative, chemoheterotrophs and, in the case of *L. alexandrii*, at least, are capable of aerobic, anoxygenic phototrophy (Biebl *et al.*, 2007). The capacity to oxidize carbon monoxide at ambient seawater concentrations is widespread and many isolates also denitrify (King, 2003; Weber and King, 2007). *L. aggregata* strain IAM 12614

Table 1. DNA and translated peptide sequence identity/similarity between nitrous oxide reductase genes from *Labrenzia aggregata* strain IAM 12614, *Labrenzia* sp. strain DG 1229 and the gDNA/cDNA clones obtained from *Trichodesmium* spp. colonies collected at stations in the Red Sea, Atlantic and Indian Oceans.

Labrenzia sp. strain	IAM 12614		DG 1229	
Percentage value	Mean	Max/min	Mean	Max/min
gDNA translated peptide identity	88.4	100/70.2	91.7	96.8/71.8
gDNA translated peptide similarity	94.1	100/84.0	95.2	98.4/85.6
gDNA nucleotide sequence identity	83.9	94.9/67.7	84.0	89.2/69.7
cDNA translated peptide identity	87.5	100/75.0	92.4	94.1/79.3
cDNA translated peptide similarity	93.8	100/84.0	95.7	97.3/85.6
cDNA nucleotide sequence identity	83.1	94.9/77.7	84.2	85.1/78.4

The maximum and minimum (max/min) range for each comparison is also shown.

was isolated originally from sediments in the Western Baltic Sea (Ahrens, 1968), while *Labrenzia* sp. strain DG 1229 was obtained by enrichment culture of the bacterial community associated with the dinoflagellate, *Scrippsiella trochoidea* (Hatton *et al*., 2012).

Labrenzia sp. Strain DG 1229 is capable of metabolizing dimethylsulfoniopropionate (DMSP) and dimethylsulfide (DMS) to produce dimethylsulfoxide and can assimilate DMSP in support of growth but not DMS (Hatton *et al*., 2012). It harbors genes encoding enzymes of both the cleavage (DddP) and the demethylation (DmdA) pathways for the utilization of DMSP, whereas *L. aggregata* strain IAM 12614 has the machinery of the cleavage pathway only, using a different form of the lyase (DddL) to generate DMS (Dickschat *et al*., 2015). In common with several other *Labrenzia* species (and, more widely, marine Alphaproteobacteria), both strains harbor *dsyB* (encoding a methyltransferase involved in DMSP synthesis) and are DMSP producers as well as consumers (Curson *et al*., 2017).

Five of the translated *nosZ* gDNA clones from the Indian Ocean (pBAH8, pIOst3_9, pIOst3_10, pIOst2_1 and pIOst2_2) and one from the mid-Atlantic (pAOst24_24) were 100% matches to the peptide sequence from *L. aggregata* strain IAM 12614 and were 94–95% identical at the nucleotide level also. These clones shared even closer DNA identity ($>$ 98.9%) with the gene from the Arabian Sea isolate, alphaproteobacterium strain 4N. This full denitrifying isolate is related to *Labrenzia* spp. and has been detected previously in close association with colonies of *Trichodesmium* spp. (Wyman *et al*., 2013). Not surprisingly, all six *Trichodesmium*-associated phylotypes occurred alongside NosZ from *L. aggregata* IAM 12614 and strain 4N in the phylogram where they are located in a cluster of sequences designated, Group 2 (Fig. 2). This group also includes translated gDNA sequences from *Trichodesmium* colonies from most of the other locations sampled (and cDNA clones from the Red Sea and mid-Atlantic, see below) and the clones obtained from *Trichodesmium thiebautii* colonies collected offshore of the Bahama Islands, exclusively. All 10 of the sequences from the Bahamas shared $>$ 87% identity ($\geq$ 98.9% similarity) with the peptide sequence from *L. aggregata* IAM 12614 and $>$ 95% with that from strain 4N.

While none of the translated *Trichodesmium*-associated gDNA clones were complete matches to NosZ from *Labrenzia* sp. strain DG 1229, more than 50% were $>$ 95% similar. Compared to *L. aggregata* strain IAM 12614, the overall mean similarity of the peptide sequence from the strain DG 1229 to the *Trichodesmium* phylotypes was higher (95.2% vs 94.1%) as was the median similarity (95.2% vs 93.6%) and the mean peptide sequence identity (91.7% vs 88.4%, Table 1). Although closer in peptide sequence identity to *L. aggregata* IAM 12614 (91.7%), *Labrenzia* sp. strain DG 1229 NosZ also shares $>$ 88% identity with the protein from *Nisaea denitrificans* and *Pseudovibrio* sp. strain JE062 with which it clusters in the phylogram between Groups 2 and 3 (Fig. 2). The six environmental gDNA sequences found in Group 3 are from the Indian Ocean and are $>$ 97.9% similar to NosZ from *Labrenzia* sp. strain DG 1229. This small group also contains a single phylotype from *Trichodesmium* sp. IMS101, a cultured isolate from North Carolina coastal waters (Prufert-Bebout *et al*., 1993) that is most closely related to the species, *Trichodesmium erythraeum* (Orcutt *et al*., 2002; Lundgren *et al*., 2005; Hynes *et al*., 2012).

The largest cluster, designated Group 4, harbors the majority of the translated gDNA clones recovered from *Trichodesmium* colonies obtained from the Red Sea, the mid-Atlantic and the Indian Ocean, in addition to three further *Trichodesmium* sp. IMS101-associated NosZ sequences (Fig. 2). The latter share high peptide sequence identity ($>$ 97.6% for the 83 residues available for comparison) with the clone, pIMS1 (GenBank Accession number KC205088), obtained previously from this same *Trichodesmium* isolate using our original *nosZ* primer pair (Wyman *et al*., 2013), and at the nucleotide level, they are virtually identical ($>$ 98.8% over 252 nucleotides vs $\sim$ 81% for strain 4N). Most members of Group 4 share the greatest mean amino acid sequence similarity with NosZ from *Labrenzia* sp. strain DG 1229 (95.9%) compared to *L. aggregata* strain IAM 12614 (93%), the strain most closely affiliated overall (95.4% mean similarity vs 93.8% for strain 4N) with the Indian Ocean clones. Group 4, however, also includes a few sequences that are related more closely to the species, *Labrenzia alba*, or to other alphaproteobacterial genera. For example, clones pAOst24_18 and pAOst22_4 are more similar ($>$ 93%) to *L. alba* strain CECT 7551 than they are to the two other *Labrenzia* species.

The last cluster, Group 1, harbors the most divergent NosZ phylotypes recovered during the study including a further clone from *Trichodesmium* IMS101 and three

related sequences from station 3 in the Indian Ocean (Fig. 2). All four phylotypes share ~ 93% similarity with *L. aggregata* strain IAM 12614 but they are nearly identical (> 99%) to NosZ from *Leisingera* (*Phaeobacter*) *aquaemixtae*. This recently described member of the *Roseobacter* clade was isolated from coastal waters off Jeju Island in the Korea Strait in the mixing zone of an inflowing freshwater spring and the ocean (Park *et al*., 2014). In agreement with their high sequence identity, Group 1 *Trichodesmium* phylotypes cluster with *Leisingera* and *Phaeobacter* species in the phylogram rather than with either of the *Labrenzia* species most closely affiliated with the other three groups.

Apart from the few exceptional phylotypes found in Groups 1 and 4, what emerges most strikingly from the phylogenetic analysis of the gDNA clones is the remarkably low overall diversity of the denitrifying community associated with *Trichodesmium*. While there is some evidence for restricted biogeographic range (e.g., the Bahamas phylotypes of Group 2), there is strong, underlying similarity in the community composition of denitrifiers between samples taken from a geographically widespread range of ocean provinces. To verify that this was not an artefact arising from the generally conserved nature of the NosZ peptide region analyzed, all *Trichodesmium*-associated *nosZ* gDNA nucleotide sequences were also aligned using MatGAT (Campanella *et al*., 2003) and individual OTUs were assigned at a cutoff of 5% (i.e., each OTU shared ≥ 95% DNA identity). Classical indices of diversity for the five OTUs identified by this procedure confirmed the low overall diversity of the community from all samples (Shannon–Weiner $H = 0.704$, Simpson's $D = 0.654$; $1/D = 1.528$).

For comparison, the same diversity indices based on *nirS* (encoding another component of the denitrification pathway, dissimilatory nitrite reductase) using a similar cutoff (95% identity) to define OTUs were $H = 4.24$ and $D = 0.04$ ($1/D = 25$) for five benthic denitrifier assemblages sampled along the Chesapeake Bay estuary (Francis *et al*., 2013). While denitrifier communities in open waters tend to be less species rich than either coastal or benthic populations, the overall *nirS* diversity found at three open ocean sites in the Arabian Sea was also notably higher ($H = 1.56$–1.95; Jayakumar *et al*., 2009) than that associated with the *Trichodesmium* consortia analyzed and reported here. This is the case, even though these authors used a less stringent cutoff (90% DNA identity) to define OTUs than that applied in the present study and by Francis and colleagues (2013).

By broadening the specificity of the *nosZ* primers used at the start of this study, it was anticipated that a much wider variety of *Trichodesmium*-associated sequences might be encountered. In practice, the majority of *nosZ* DNA sequences (141/153) were assigned to just two OTUs (corresponding to Groups 2 and 4, Fig. 2) resulting in a dominance value (1 – J, where J is Pielou's evenness index) of 0.56. This is comparable to the least diverse coastal *nirS* community ($1 - J = 0.48$, $H = 0.35$) encountered by Jayakumar and colleagues (2009) for fully denitrified, neritic waters off Vengurla, just north of Goa, India. These results, while surprising, are presented with the important caveat that the new primers designed and used in this study may still underrepresent the true diversity of the *nosZ*-containing community present within the consortium despite their wider target range.

The 19 gDNA nucleotide sequences from Group 2 were > 93% identical to *nosZ* from *L. aggregata* strain IAM 12614 and were very closely related also to the gene from the Arabian Sea strain 4N (> 95% identical). The consensus nucleotide sequence derived for the largest NosZ cluster, Group 4 (122/153 sequences), however, shares only 84.6% identity with that of *Labrenzia* sp. strain DG 1229 and 82.8% with *nosZ* from strain 4N. Two environmental DNA sequences from a coastal marine sediment bordering the East China Sea (GenBank references FJ227186.1 and FJ227160.1) are somewhat closer matches overall (~ 92% identity) than the two *Labrenzia* reference sequences used here but probably not significantly so as their translated products are only 89.4% similar (vs 95.9% for *Labrenzia* sp. strain DG 1229) and ~ 80% identical. The members of the largest Group 4, therefore, most likely belong to a novel lineage of denitrifying alphaproteobacteria that, in common with the members of Group 2, associate in consortium with *Trichodesmium* spp. quite specifically.

Enrichment of the denitrifier community associated with Trichodesmium

A quantitative polymerase chain reaction (qPCR) protocol was developed to determine the concentrations of Group 2 and 4 denitrifiers found within *Trichodesmium* colonies and to compare these with the combined population numbers present in surrounding waters (Supporting Information Fig. S2 and Supporting Information Table S1). At Indian Ocean stations 2 and 3, the target sequences for both groups were highly enriched within *Trichodesmium* colonies but were at or below the level of detection in the bulk water column (Table 2). To verify that the latter result was not due to the presence of PCR inhibitors, the same samples (equivalent to the DNA extracted from bacterioplankton recovered from 25 ml of seawater; ~ 10 ng DNA per reaction) were used successfully to amplify 16S rRNA genes without notable inhibition (Supporting Information Fig. S3A). This positive result confirms that free-living representatives of Groups 2 and 4 *Trichodesmium*-associated denitrifiers were not present in appreciable numbers within the upper mixed layer at these highly oligotrophic stations. Indeed, previous estimates of the abundance of Group 2 denitrifiers in near surface waters to the north and south of stations 2

Table 2. Enrichment of Group 2 and Group 4 *nosZ* phylotypes associated with *Trichodesmium* colonies from the Indian Ocean.

Station and depth	Group 2 *nosZ*	Group 4 *nosZ*	Sum of *nosZ* copies ml^{-1}
Stat 2, colonies 10 m	1.6×10^5 (78.6%)	4.37×10^4 (21.4%)	2.04×10^7–1.02×10^8
Stat 2, water column 20 m	< 100	< 100	< 8
Stat 3, colonies 5 m	8.98×10^4 (65.1%)	4.82×10^4 (34.9%)	1.38×10^7–6.9×10^7
Stat 3, water column 20 m	< 100	< 100	< 8
Stat 3, water column 25 m	< 50	< 100	< 6

Groups 2 and 4 cell concentrations for *Trichodesmium* are expressed per colony while those for the water column are for bacterioplankton ($>$ 0.2 μm) filtered from a volume of 25 ml seawater and are at or below the limit of detection for the assays. The numbers in parentheses are the percentage of the total represented by each Group found in association with the *Trichodesmium* colonies. The total number of *nosZ* copies per milliliter in the last column is based on a colony volume of 0.002–0.01 ml (Sheridan *et al*., 2002).

and 3 and sampled during the same cruise were of the order of $\sim$ 1–5 $\times$ 10^4 *nosZ* copies l^{-1} (Wyman *et al*., 2013).

Assuming an average colony volume of 0.002–0.01 ml (as estimated by Sheridan *et al*., 2002), the combined concentration of the target *nosZ* sequences from the two groups found in association with *Trichodesmium* colonies was highly enriched at 1.38×10^7–1.02×10^8 copies ml^{-1} (Table 2). To put these numbers in context, the concentration of all bacteria within the upper mixed layer at stations 2 and 3 was markedly lower at $\sim$ 0.8–1.2 $\times$ 10^6 cells ml^{-1} (Zubkov *et al*., 2006). The average bacterial density associated with *Trichodesmium* colonies collected from the Sargasso Sea, by contrast, has been estimated at 8.2×10^8 (range 8.1×10^7–3.5×10^9) cells ml^{-1} (Sheridan *et al*., 2002) and may be as high as 2.6×10^{11} cells ml^{-1} elsewhere in the Atlantic Ocean (Paerl, 1982). By extrapolation, this suggests that Groups 2 and 4 denitrifiers account for $\sim$ 0.3% to as high as 10% of the bacterial cells present within the *Trichodesmium* consortium if *nosZ* is present at one copy per cell. While many proteobacteria including *Escherichia coli* are indeed monoploid or (mero-)oligoploid, genome copy number can be much higher (e.g., 40–80 genome equivalents per cell) in some exceptional species like *Azotobacter vinelandii* (Pecoraro *et al*., 2011). Therefore, as their ploidy is uncertain, the actual contribution of Groups 2 and 4 denitrifiers to the consortium may well be lower than these estimates. Nevertheless, it is apparent that the *Trichodesmium*-associated denitrifiers are present within the colonies at an enrichment factor of at least several orders of magnitude when compared to the much lower population concentrations found in surrounding waters (Wyman *et al*., 2013).

Does the Trichodesmium*-associated denitrifying community express* nosZ *within the colony environment?*

To explore whether the *Trichodesmium* consortium denitrifiers might be capable of expressing *nosZ in situ*, cDNA synthesis reactions were performed with the *nosZ*HMRevA/G primer combination (Supporting Information Table S1) using total RNA purified from *Trichodesmium* colonies obtained from the Red Sea and mid-Atlantic Ocean (Supporting Information Fig. S3B and C). The samples from the Red Sea were collected a few hours after dusk while those from mid-Atlantic stations 22 and 24 were obtained from pre-dawn hydrocasts; that is, at sampling times when no *Trichodesmium* photosynthesis should be taking place and when the oxygen concentrations within the colonies were most likely to be well below ambient (Paerl and Bebout, 1988; 1992). Nitrous oxide reductase, along with other components of the denitrification apparatus, is tightly regulated by oxygen in most denitrifiers (Zumft, 1997) and, hence, avoiding samples taken during the day when *Trichodesmium* colonies are producing photosynthetically generated oxygen was thought to maximize the chances of detecting *nosZ* transcripts. In an earlier study, it was shown that *nosZ* mRNAs are present at low concentrations in surrounding surface waters where *Trichodesmium* is present (Wyman *et al*., 2013) and so, as a precaution, the colonies were washed three times in sterile filtered seawater to minimize the carryover of unattached bacteria prior to RNA extraction.

Transcripts corresponding to *nosZ* were detected in all three samples interrogated but not in the control reactions in which the reverse transcriptase enzyme was omitted (Supporting Information Fig. S3B and C). A total of 265 cDNA clones (Red Sea 68; Atlantic Ocean Station 22–102 clones and Station 24–95 clones) were subsequently obtained from these samples and their derived peptide sequences were incorporated alongside all 153 translated *Trichodesmium*-associated *nosZ* gDNA clones in an enlarged phylogenetic analysis (Supporting Information Fig. S4). Without exception, the translated cDNAs clustered with either Groups 2 or 4; that is, within the clusters representing the most frequently encountered NosZ gDNA sequences retrieved from these locations. Not only are representatives from these two groups the most abundant, therefore, but they also dominate the transcriptionally active denitrifier populations within the colonies. If significant numbers of other denitrifiers are also present, they either remain undetected using the new primer sets developed in this study or are incapable of denitrifying nitrogen

oxides as far as N_2 due to the absence of NosZ. It is known, for example, that members of the *Alteromonas* and *Pseudoaltreromonas* genera (which are found in close association with *Trichodesmium* spp.; Hewson *et al*., 2009; Lee *et al*., 2017) include some denitrifying representatives but not all can reduce nitrate as far as nitrous oxide or dinitrogen (Enger *et al*., 1987).

The NosZ peptide sequences from the majority of the cDNAs sequenced were placed in the largest cluster, Group 4 (Supporting Information Fig. S4) and, overall, were most similar to *Labrenzia* sp. strain DG 1229 although clearly not identical at either the nucleotide or peptide sequence level (Table 1). Six of the sequences (pRSc_34, pAOcSt24_76, pAOcSt22_95, pAOcSt22_98, pAOcSt22_102 and pAOcSt24_90), however, were marginally more closely related to NosZ from *L. alba* strain CECT 5094. Like *L. alba* strain CECT 7551 (see above), this strain was originally isolated from oysters growing off the Spanish Mediterranean coast and its NosZ sequence shares 99% identity with that of *Labrenzia* sp. strain DG 1229. Given that the latter strain was obtained from a dinoflagellate enrichment culture (Hatton *et al*., 2012), it is conceivable that the origin of the two Spanish *L. alba* isolates may have been from food particles ingested by the oyster hosts.

The detection of *nosZ* mRNAs originating from the denitrifier community associated with *Trichodesmium* raises questions concerning the biogeochemical significance of these organisms. Is the detection of transcripts indicative of colony associated N_2O reduction (and even, perhaps, earlier steps in the denitrification pathway) *in situ* or could the genes still be transcribed without always giving rise to an active enzyme? Traditionally, nitrous oxide reductase is regarded as the most oxygen-sensitive and extensively regulated enzyme of the denitrification pathway, only being transcribed and expressed under denitrifying conditions (Bonin *et al*., 1989; Zumft, 1997). With the recognition of the wider potential for aerobic denitrification in a variety of habitats, however, this clearly does not apply universally. Nevertheless, even under these more relaxed atmospheric conditions *nosZ* (and other components of the denitrification apparatus) are only transcribed and translated when oxygen and nitrogen oxides are being co-utilized as terminal electron acceptors (Härtig and Zumft, 1999; Gao *et al*., 2010; Miyahara *et al*., 2010; Chen and Strous, 2013).

With Group 2 alphaproteobacterium strain 4N available to hand (Wyman *et al*., 2013), it was possible to establish experimentally whether *nosZ* is expressed constitutively in this representative strain (see Fig. 2) or whether it is only induced under suboxia when the organism is actively denitrifying. Both *nosZ* and *nirS* mRNAs and their cognate gene products were upregulated very substantially (*nosZ* mRNA $>$ 2000-fold, NosZ 117-fold; *nirS* mRNA 718-fold and NirS 695-fold) under anoxia (in the presence of nitrate as electron acceptor) confirming that both genes are very tightly regulated in this strain under aerobic, nondenitrifying conditions (Supporting Information Fig. S5). If gene regulation in the *Trichodesmium* associated denitrifiers is at all similar in this regard then the detection of *nosZ* mRNAs within natural populations raises the possibility of active N_2O reduction occurring within colonies *in situ*. While bioenergetically this would be less favorable than aerobic respiration, the additional contribution from N_2O reduction (and, conceivably, auxiliary nitrate/nitrite reduction also) should increase overall ATP yields (Chen and Strous, 2013) under the oxygen-depleted conditions found within the colonies in dimly-lit waters or at night (Paerl and Bebout, 1988). It is less likely that this process would be operating during full daylight hours, however, because of the production of photosynthetically generated oxygen by *Trichodesmium* and its accumulation to supersaturated concentrations within the colonies (Paerl and Bebout, 1988; Eichner *et al*., 2017).

While the detection of *nosZ* transcripts is at least consistent with active, colony-associated N_2O-reduction/denitrification, measurements of N_2O cycling within the *Trichodesmium* consortium are needed to confirm this. There is also uncertainty concerning the source(s) of nitrogen oxides needed to drive these processes. It is well established, however, that nitrous oxide can be produced via nitrification (and especially so under reduced oxygen conditions) and abiotic sources of N_2O have also been identified (Codispoti and Christensen, 1985; Samarkin *et al*., 2010; Ji *et al*., 2015; Trimmer *et al*., 2016). The low concentrations of combined nitrogen in the illuminated surface waters where *Trichodesmium* is found, however, would appear to militate against significant rates of colony associated nitrate/nitrite reduction, although, of course, deeper-dwelling populations should encounter higher concentrations closer to the nitracline.

Perhaps it is worth noting in this context that *Trichodesmium* colonies liberate substantial quantities of ammonium and other forms of reduced N (including amino acids) into surrounding waters (Karl *et al*., 1997, Mulholland *et al*., 2006). Some of this fixed N is assimilated by the epibiotic community directly (Eichner *et al*., 2017), but it has also been shown that ammonium derived directly or indirectly (i.e., as remineralized DON) from *Trichodesmium* blooms supports upper water column nitrification leading to the transient appearance of isotopically light nitrate (Gandhi *et al*., 2010; Sutka *et al*., 2004). Elevated nitrite concentrations have been observed also in the tropical surface waters of the SE Indian Ocean where *Trichodesmium* blooms frequently occur alongside high numbers of *Rhodobacteraceae* bacteria harboring *nosZ* (Raes *et al*., 2016).

To our knowledge, no study has looked for nitrifiers growing in association with *Trichodesmium* specifically but the genomes of the clade 2 isolate, Alphaproteobacterium strain 4N (Wyman *et al*., 2013), *Labrenzia* spp., and several other related marine alphaproteobacteria contain the nitrification

Table 3. The distribution of N metabolism genes in Group 2 isolate 4N and related alphaproteobacteria.

Organism	*napA*	*nirS/K*	*nosZ*	*amoA*	*hao*	*cycA*	NO_2^-
Alphaproteobacterium 4N	+	+	+	+	–	+	+
Labrenzia aggregate IAM 12614	+	+	+	+	–	+	?
Polymorphum gilvum	+	+	+	+	–	–	?
Rugeria pomroyi	–	+	+	–	–	–	–
Lesingera caerulea	+	+	+	–	–	–	–
Nisaea denitrificans	+	+	+	+	–	+	?
Silicibacter lacuscaerulensis	–	+	+	+	–	–	?
Silicibacter TrichCH4B	–	+	–	+	–	–	?
Rosebacter denitrificans OCh 114	+	+	+	+	–	–	?

The genes listed encode: periplasmic nitrate reductase (*napA*); dissimilatory nitrite reductase (*nirS* or *nirK*); nitrous oxide reductase (*nosZ*); ammonium monoxygenase (*amoA*); hydroxylamine oxidoreductase (*hao*); and cytochrome c554 (*cycA*). The ability to nitrify (i.e., to produce NO_2^- via ammonium oxidation) is indicated in the last column.

genes ammonia monoxygenase (*amoA*) and that (*cycA*) for the cytochrome c554 which accepts electrons from hydroxylamine oxidoreductase (Table 3). While recognizable copies of genes encoding the latter enzyme (used by autotrophic, ammonia-oxidizing bacteria to oxidize hydroxylamine to nitrite) seem not to be present, nitrite accumulates transiently in late exponential/stationary phase cultures of the Group 2 denitrifier, strain 4N when grown with ammonium as sole N source (Table 3). The genome of at least one other denitrifying *Trichodesmium*-associated heterotroph, *Silicibacter* TrichCH4B (see Rao *et al*., 2015) also contains *amoA* in addition to *nirK* (nitrite reductase) and a nitric oxide reductase (*norB*), although it lacks *nosZ* (Table 3).

Firm evidence in support of N_2O reduction and denitrification outside of ODZs is scant apart from that associated with suboxic microenvironments such as living/dead copepods and the guts of other animals (Glud *et al*., 2015; Scavotto *et al*., 2015). However, another recent study has also reported the presence of an active *nosZ* assemblage in oxygenated surface waters (Sun *et al*., 2017). The distribution of these nitrous oxide-consuming bacteria was correlated with elevated nitrite and nitrous oxide concentrations within the water column and also with chlorophyll fluorescence, suggesting a link with the autotrophic community. The present study demonstrates the potential for N_2O reduction to N_2 in oxic waters by bacteria living in intimate association with *Trichodesmium* populations. If the ultimate source of N_2O is nitrified and/or denitrified fixed N directly transferred from *Trichodesmium* to the consortium members (see Eichner *et al*., 2017), there is the potential for its rapid return to the overlying atmosphere through the completion of the marine N cycle at the microscale.

Experimental procedures

Sample collection, nucleic acid extraction and purification

Trichodesmium colonies from two stations in the central Atlantic Ocean (Station 22: 9°28′N, 29°53′W; Station 24: 5°46′N, 29°42′W) were collected in October 2013 during the NERC-funded AMT 23 cruise aboard *RRS James Clark Ross* (JR300). The samples were obtained during predawn, vertical bongo net hauls integrated through the upper 200 m of the water column using two side-by-side WP-2 nylon nets (200 μm mesh) with 0.57 m openings and 200 μm cod-end mesh windows. The individual colonies were picked from the material concentrated in the cod-ends using a sterile plastic inoculating loop. After rinsing in three changes of 0.2 μm filtered surface seawater, the samples were preserved in RNAlater (Thermo Fisher Scientific, Loughborough, UK) and stored at −80°C. On return of the research vessel to the UK, the frozen samples were shipped on dry ice to Stirling and stored at −20°C prior to the extraction of DNA and RNA.

*S*amples of *Trichodesmium* from near-shore waters of the Red Sea were collected passively in November/December 2014 with a 200 μm mesh net boomed out from the seaward end of the 50 m long pier at the Interuniversity Institute of Marine Science at Eilat, Israel (29°59N 34°95′E). The individual colonies were sorted from the net haul with sterile plastic dropper pipettes, washed three times in sterile surface seawater and transferred to RNAlater as described above and then shipped on wet ice to the UK by air before storage at −20°C prior to DNA and RNA extraction. The samples used for RNA extraction were obtained from net samples deployed in late afternoon/early evening that were sorted and preserved a few hours after darkness.

To prepare samples from both the Red Sea and central Atlantic Ocean for nucleic acid extraction, 10–20 *Trichodesmium* colonies were captured on sterile, 25 mm diameter, 0.2 μm pore size polycarbonate membranes (Nuclepore; GE Healthcare Life Sciences, Little Chalfont, UK) under gentle vacuum. The colonies were then rinsed twice with 5 ml of sterile artificial seawater (ASW) medium (Wyman *et al*., 1985) to remove any traces of RNAlater. After the last drops of ASW were removed, the membranes were transferred aseptically to 2 ml screw-cap tubes and DNA was extracted using a ZR Fungal/Bacteria DNA Mini-prep Kit (Zymo Research, Irvine California, USA) according to the manufacturer's guidelines. Total RNA was purified using the Direct-zol™ RNA mini-prep kit (Zymo Research) and then treated postextraction with Ambion Turbo DNase (Applied Biosystems, Warrington, UK) followed by Genomic Wipeout reagent (Qiagen, Crawley, UK) using the sample

preparation and incubation conditions recommended by the suppliers.

Archived DNA from samples obtained during earlier studies and stored at −20°C in TE buffer (10 mM Tris-HCl, pH 8.0; 1 mM EDTA) was also used during this investigation. DNA from *Trichodesmium* colonies originally collected in September 2001 from near-surface waters in the northern Indian Ocean (Stations 2, 3 and 4; 00°00′N, 66°59′E, 03°47′N 67°00′E and 07°36′N 67°03′E, respectively) were purified as described by Wyman and colleagues (2013). DNA extracted from unfractionated bacterioplankton samples at depths between 20 and 45 m at the same stations was collected on 0.2 μm pore size polycarbonate filters and was purified as outlined by Bird and colleagues (2005). DNA from *T. thiebautii* colonies collected from near-surface waters off the Bahama Islands in September 1991 (24°08′N, 75°53′W) was extracted as described by Kramer and colleagues (1996). DNA obtained from the non-axenic isolate, *T. erythraeum* strain IMS101, which was brought into culture from waters off the coast of North Carolina, USA (∼ 34°36′N 76°42′W) by Prufert-Bebout and colleagues (1993), was donated to the Wyman laboratory in late 1997 by Dr Jon G. Kramer (University of Maryland).

Reverse transcription of RNA

Purified RNA (5 ng) from *Trichodesmium* colonies was reverse transcribed using the High Capacity cDNA reverse transcriptase (RT) kit (Applied Biosystems) and an equimolar (100 pmol) mix of the *nosZ* reverse primers nosHMRevA and nosHMRevG (Supporting Information Table S1) described below. In some instances, the cDNAs synthesized were then amplified prior to the PCR using an Illustra™ Genomiphi™ V2 DNA amplification kit (GE Healthcare Life Sciences) according to the supplier's guidelines. A template-free reagent control was set-up with all of the RT reactions alongside a second control reaction including template RNA but omitting the RT enzyme. Both controls were used to verify that the RNA templates, and RT reagents were free of amplifiable DNA during the preamplification step where this was performed and also during subsequent PCR detection of cDNAs.

Primer design, PCR amplification and DNA and cDNA sequencing of nosZ from Trichodesmium *consortia*

Two novel sets of oligonucleotide primers targeting an internal fragment of *nosZ* were designed at the start of this investigation using an alignment of the 178 complete NosZ sequences available at that time from the GenBank database. The first primer set (NosHMFor and NosHMRev) target *nosZ* from organisms that have a recognizable twin-arginine translocation (TAT) leader signal (consensus – SRRXFLK; Palmer and Berks, 2012) close to the N-terminus of the peptide sequence (Supporting Information Fig. S1A and B). The second primer set (NosHQFor2 and NosHQRev2) targets organisms producing a presumptive Sec-dependent NosZ; i.e., those lacking a recognizable TAT motif in their leader sequences.

Trichodesmium consortium DNA or cDNA (plus the RT and template-free controls from the cDNA synthesis reactions) was interrogated for the presence of *nosZ* by the PCR using MyTaq™ Red DNA polymerase master mix (Bioline Reagents, London, UK) in reaction volumes of 25 μl. Each reaction contained in addition, 1 μl genomic DNA (from 100 μl extract) or 1 μl of a 1 in 50 dilution of the cDNA reactions, 2–3.5 mM $MgCl_2$, plus 50 or 200 pmol oligonucleotide primer mix (see Supporting Information Table S1). In all instances (cDNA or gDNA), the final amount of nucleic acids added ranged from 9.2 to 10.8 ng per reaction. After an initial denaturation step at 95°C for 2 min, cycling conditions consisted of three steps: 94°C for 30 s; 46 or 56°C for 30 s and 72°C for 1 min for 30 cycles followed by a final extension at 72°C for 10 min. The resulting PCR products (∼ 750 bp) were resolved by electrophoresis through 2% (wt/vol) agarose gels in TAE buffer, excised using sterile razor blades and gel-purified using the Wizard® SV Gel and PCR Cleanup kit (Promega, Southampton, UK).

The gel-purified products were TA cloned in pCR2.1®TOPO® vector (Invitrogen-Life Technologies, Paisley, UK) and transformed into α-Select Gold Efficiency competent *E. coli* (Bioline Reagents). Transformants were selected on Luria Bertani (LB) agar plates containing 100 μg ml^{-1} ampicillin and recombinants identified by alpha complementation (blue white screening). To verify the presence of the targeted *nosZ* fragment, selected recombinant clones were grown overnight at 37°C in LB broth containing 100 μg ml^{-1} ampicillin and/or 50 μg ml^{-1} kanamycin). Plasmid DNA was purified using the NucleoSpin® plasmid purification kit (Macherey-Nagal; CamLab, Cambridge, UK) and screened for the presence of an insert by digestion of the vector with the restriction enzyme, *EcoR1*. Plasmids that harbored inserts of the expected size were DNA sequenced on both strands with M13 primers by a commercial provider (BioScience, Bo'Ness, UK) to confirm their identity.

Following this preliminary screen, recombinant clones were grown overnight in 2× YT medium (containing 50 μg ml^{-1} kanamycin) in 96-well plates and plasmid DNA was purified using the SeqPrep™ 96 plasmid preparation kit (EdgeBiosystems; VH Bio, Gateshead, UK). High-throughput DNA sequencing was performed with BigDye reagents on an ABI 3730 capillary sequencer by the NERC Biomolecular Analysis Facility, University of Edinburgh, Edinburgh, UK.

Quantification of nosZ copy number in Trichodesmium *colonies and bacterioplankton samples*

A qPCR assay was developed to estimate the abundance of *nosZ*-containing organisms associated with *Trichodesmium* colonies obtained from the Indian Ocean. The samples were assayed using *nosZ* group-specific primers (Supporting Information Table S1 and Supporting Information Fig. S2A) and double-dye-labeled oligonucleotide probes. The primers and probes target the two most frequently encountered *nosZ* phylotypes (named Groups 2 and 4) that were found associated with *Trichodesmium* colonies in this study. Group 2 primers (NosGrp2F/2R) amplify a product of 112 bp, whereas Group 4 primers (NosGrp4F/2R) amplify a product of 72 bp. The specificity of the primers was confirmed by Sanger DNA sequencing of the PCR products and verified also by testing each primer set against representative clones from all four nosZ groups (Fig. 2) identified in this study (Supporting Information Fig. S2B and C).

The qPCR assays were performed with a Stratagene Mx3000p thermocycler (Agilent Technologies, Stockport, UK) using SensiFAST™ Probe Lo-ROX reagent mix (Bioline Reagents) in 20 μl reactions containing 1 μl DNA (~ 10 ng), 50 pmol primer mix and 150 nM of labeled probe as appropriate. Both oligonucleotide probes were labeled with fluorescein at the 5′ end and Elipse dark quencher at their 3′ prime end (Eurofins Genomics, Ebersberg, Germany). Following activation at 95°C for 10 min, reaction mixtures were cycled at 95°C for 15 s and 60°C for 15 s (combined elongation and extension) for a total of 35 cycles. The quantification cycle (C_q) was determined automatically with the MxPro v6.22 software supplied with the instrument. Initial template quantity (copies of *nosZ* per assay) was extrapolated from standard curves of ten-fold dilution series (1×10^1–1×10^7 copies per reaction) of the target *nosZ* amplicons.

Standard DNA was amplified by the PCR from clones of representative of Groups 2 and 4 *nosZ* phylotypes using the group-specific primers for each template DNA. The amplicons were then gel purified using the Wizard® SV Gel and PCR Cleanup kit as described previously and their purity (A260:A280) and concentration was estimated using a Picodrop microliter UV/Vis spectrophotometer (Cambridge Bioscience, Cambridge, UK). Copy number was calculated using the known length (bp) and measured concentration of the PCR products assuming that the molecular mass of a single base pair is 650 Da. The detection limit for both *nosZ* standard DNAs was consistently between 10 and 100 copies per reaction. PCR efficiency was between 98% and 105% for both Group 2 ($r^2 = 0.99$) and Group 4 ($r^2 = 0.98$).

DNA from bacterioplankton samples collected at stations 2 (45 m depth) and 3 (20 and 25 m depths) in the Indian Ocean were also assayed for the presence of Group 2 and 4 *nosZ* as described above. The reactions included 1 μl of template DNA (~ 10 ng); equivalent to the yield of DNA extracted from a volume of 25 ml of each seawater sample analyzed (Bird *et al*., 2005). To demonstrate that amplifiable DNA was present, the bacterioplankton DNA samples were also screened for the presence of 16S rRNA genes using the primer pair 341F and 518R (Muyzer *et al*., 1993). The PCRs were performed with 1× MyTaq™ Red DNA polymerase master mix, 1 μl DNA, 50 pmol primer mix and 1.5 mM $MgCl_2$. Following an initial denaturation step at 95°C for 2 min, the reactions were cycled 30 times at 94°C for 15 s, 58°C for 15 s and 72°C for 30 s with a final elongation step of 72°C for 5 min. A prominent band of the expected product size of ~ 180 bp was detected in each bacterioplankton sample (Supporting Information Fig. S2a).

Phylogenetic analyses

Nucleotide sequences of *nosZ* recovered from *Trichodesmium* spp. were compiled manually, trimmed of the primer regions and then translated *in silico*. Reference sequences of NosZ from known alphaproteobacterial denitrifiers from the marine environment that were closest in peptide sequence to the *Trichodesmium*-associated clones were downloaded from GenBank using the Basic Local Alignment Search Tool at NCBI (Altschul *et al*., 1990). The NosZ sequence from the gammaproteobacterium, *Pseudomonas stutzeri*, was used to root the trees. The trimmed amplified region of *nosZ* corresponds to a translated peptide sequence of ~ 236 residues. To enable the direct inclusion of several *Trichodesmium*-associated sequences amplified previously with a different primer set (*nosZ*F1 and *nosZ*R, Supporting Information Table S1; Wyman *et al*., 2013), the trimmed sequences were edited further at the N-terminal end to leave a region of 188 informative amino acids for analysis.

A multiple alignment of the edited peptide sequences was carried out using ClustalX 2.1 (Larkin *et al*., 2007). Evolutionary analyses and phylogenetic reconstructions were performed in MEGA 6 (Tamura *et al*., 2013) using a Maximum Likelihood routine based on the Dayhoff matrix model. Consensus trees with the highest log likelihood values were produced based on 1000 resamplings in the bootstrap. Global alignments of nucleotide and peptide sequences were performed using MatGAT (Matrix Global Alignment Tool) software (Campanella *et al*., 2003).

Investigation of the effect of suboxia on nosZ and nirS (nitrite reductase) expression by qRT-PCR in the Arabian Sea isolate 4N

Experimental cultures ($n = 3$) of Group 2 isolate 4N from the Arabian Sea (Wyman *et al*., 2013) were grown in 25 ml of three quarter strength ASW medium (Wyman *et al*., 1985) that was amended with 2 g l^{-1} sodium acetate, 1 g l^{-1} yeast extract, 2 g l^{-1} sodium nitrate and 1 g l^{-1} ammonium chloride. The cultures were grown aerobically in loose-bunged 250 ml conical flasks in a rotary shaking incubator at 200 rpm at a growth temperature of 33°C.

For the establishment of anaerobic cultures ($n = 3$), fresh sterile medium was boiled in a microwave oven to drive off dissolved gases. The original volume was restored by the addition of boiled deionized water and then used to fill sterile 25 ml glass universal bottles to the rim. The culture medium was cooled and supplemented with 20 mg l^{-1} sodium dithionite prior to inoculation to remove residual oxygen. The bottles were inoculated and then sealed with gas tight caps and incubated at 33°C in a static incubator until the cultures had reached the mid-logarithmic phase of batch growth.

Aerobic and anaerobic cultures were harvested by centrifugation (8000*g* for 5 min) at 4°C and RNA extracted from the cell pellets using the Direct-zol™ RNA mini-prep kit (Zymo Research). Following removal of any contaminating DNA from the RNA preparations as described, cDNA synthesis was performed using 100 ng of RNA in reactions containing the random primers supplied with the High Capacity cDNA reverse transcriptase (RT) kit according to the suppliers (Applied Biosystems) recommended protocol. Control reactions omitting either the template or the RT enzyme (see above) were set up alongside the cDNA synthesis reactions for each replicate for both treatments.

The concentration of *nosZ* cDNAs was estimated using the qPCR conditions and QnosF/R primers plus probe combination described previously (Wyman *et al*., 2013) with SensiFAST™ Probe Lo-ROX reagents containing a 1 in 50 dilution of cDNA. The same diluted cDNA reactions were employed to quantify *nirS* mRNAs using the primer pair and cycling conditions given in Supporting Information Table S1. These reagents were used also to estimate 16S rRNA concentrations (in order to normalize *nosZ* and *nirS* cDNA

concentrations) using the PCR cycling conditions previously described for the 341F/518R primer pair (Muyzer *et al*., 1993). Owing to the high concentration of target molecules, 1 in 1000 dilutions (rather than 1 in 50) of the cDNA reactions were added to the 16S rRNA qPCRs. A fresh dilution series (1×10^1–1×10^7 copies per reaction) of the previously described *nosZ* standard was used for quantification. Likewise, a similar range of standards was prepared for 16S rRNA and *nirS* using the cloned genes from the Arabian Sea isolate 4N (Wyman *et al*., 2013) as template DNA for PCR and the reaction conditions just described.

Nucleotide sequence accession numbers

Genomic DNA sequences of *nosZ* amplified from *Trichodesmium* colonies from the Atlantic Ocean (Stations 22 and 24; KT340899–KT340933), the Red Sea (KT315839–KT315900), the Indian Ocean (Stations 2, 3 and 4; KT315799–KT315838), the Bahama Islands (KT315787–KT315798) and the cultured *Trichodesmium* isolate IMS101 (KT315782–KT315786) have been deposited in GenBank under the respective accession numbers. Complementary DNA sequences of *nosZ* mRNA from *Trichodesmium* colonies collected from the Red Sea (KT340934–KT341001) and the Atlantic Ocean (KT340702–KT340898) have also been deposited in GenBank.

Acknowledgements

This research was supported by a grant from the Natural Environment Research Council (NERC) of the UK (NE/K015095/1) awarded to MW. The DNA sequencing carried out during this study was performed by the NERC Biomolecular Analysis Facility at the University of Edinburgh (NERC grant number NBAF843). We are grateful to Dr Andy Rees (Plymouth Marine Laboratory) and Prof Mike Zubkov (National Oceanography Centre, Southampton) for facilitating the collection of *Trichodesmium* during the AMT 23 cruise and to Priscilla Lange (University of Oxford) for sorting the samples from the net hauls. We are most grateful also to Drs Subhajit Basu and Yeala Shaked (The Hebrew University of Jerusalem) for collecting *Trichodesmium* samples from the Red Sea. We would like to thank Sylvia Hodgson for technical assistance.

References

Ahrens, R. (1968) Taxonomische Untersuchungen an sternbildenden *Agrobacterium-Arten* aus der westlichen Ostsee. *Kiel Meeresforsch* **24:** 147–173.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment tool. *J Mol Biol* **215:** 403–410.

Bergman, B., Sandh, G., Lin, S., Larsson, J., and Carpenter, E.J. (2013) *Trichodesmium* – a widespread marine cyanobacterium with unusual nitrogen fixation properties. *FEMS Microbiol Rev* **37:** 286–302.

Biebl, H., Pukall, R., Lünsdorf, H., Schulz, S., Allgaier, M., Tindall, B., and Wagner-Döbler, I. (2007) Description of *Labrenzia alexandrii* gen. nov., sp. nov., a novel alphaproteobacterium containing bacteriochlorophyll a, and a proposal for reclassification of *Stappia aggregata* as *Labrenzia aggregata* comb. nov., of *Stappia marina* as *Labrenzia marina* comb. nov. and of *Stappia alba* as *Labrenzia alba* comb. nov., and emended descriptions of the genera *Pannonibacter*, *Stappia* and *Roseibium*, and of the species *Roseibium denhamense* and *Roseibium hamelinense*. *Int J Syst Evol Microbiol* **57:** 1095–1107.

Bird, C., Martinez Martinez, J., O'donnell, A.G., and Wyman, M. (2005) Spatial distribution and transcriptional activity of an uncultured clade of planktonic diazotrophic gammaproteobacteria in the Arabian sea. *Appl Environ Microbiol* **71:** 2079–2085.

Bonin, P., Gilewicz, M., and Bertrand, J.C. (1989) Effects of oxygen on each step of denitrification in *Pseudomonas nautica*. *Can J Microbiol* **35:** 1061–1064.

Canfield, D.E., Glazer, A.N., and Falkowski, P.G. (2010) The evolution and future of the Earth's nitrogen cycle. *Science* **330:** 192–196.

Campanella, J.J., Bitincks, L., and Smalley, J. (2003). MatGAT: an application that generates similarity/identity matrices using protein or DNA sequences. *BMC Bioinformatics* **4:** 29. doi:10.1186/1471-2105-4-29.

Capone, D.G., Zehr, J.P., Paerl, H.W., Bergman, B., and Carpenter, E.J. (1997) *Trichodesmium*, a globally significant marine cyanobacterium. *Science* **276:** 1221–1229.

Chen, J., and Strous, M. (2013) Denitrification and aerobic respiration, hybrid electron transport chains and co-evolution. *BBA Bioenergetics* **1827:** 136–144.

Codispoti, L.A. (1995) Is the ocean losing nitrate? *Nature* **376:** 724.

Codispoti, L.A. (2007) An oceanic fixed nitrogen sink exceeding 400 Tg N a-1 vs. the concept of homeostasis in the fixed nitrogen inventory. *Biogeosciences* **4:** 233–253.

Codispoti, L.A., and Christensen, J.P. (1985) Nitrification, denitrification and nitrous oxide cycling in the eastern tropical South Pacific Ocean. *Mar Chem* **16:** 277–300.

Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004) WebLogo: a sequence logo generator. *Genome Res* **14:** 1188–1190.

Curson, A.R.J., Liu, J., Bermejo Martínez, A., Green, R.T., Chan, Y., and Carrión, O. (2017) Dimethylsulfoniopropionate biosynthesis in marine bacteria and identification of the key gene in this process. *Nat Microbiol* **2:** 17009. doi: 10.1038/nmicrobiol.2017.9.

Daims, H., Lebedeva, E.V., Pjevac, P., Han, P., Herbold, C., and Albertsen, M. (2015) Complete nitrification by *Nitrospira* bacteria. *Nature* **528:** 504–509.

Deutsch, C., Sarmiento, J.L., Sigman, D.M., Gruber, N., and Dunne, J.P. (2007) Spatial coupling of nitrogen inputs and losses in the ocean. *Nature* **445:** 163–167.

De Vries, T., Deutsch, C., Primeau, F., Chang, B., and Devol, A. (2012) Global rates of water column denitrification derived from nitrogen gas measurements. *Nat Geosci* **5:** 547–550.

Dickschat, J.S., Rabe, P., and Citron, C.A. (2015) The chemical biology of dimethylsulfoniopropionate. *Org Biomol Chem* **13:** 1954–1968.

Eichner, M.J., Klawonn, I., Wilson, S.T., Littmann, S., Whitehouse, M.J., Church, M.J., *et al*. (2017) Chemical microenvironments and single-cell carbon and nitrogen uptake in field-collected colonies of *Trichodesmium* under different pCO2. *ISME J* **11:** 1305–1317.

Enger, Ø., Nygaard, H., Solberg, M., Schei, G., Nielsen, J., and Dundas, I. (1987) Characterization of *Alteromonas denitrificans* sp. nov. *Int J Syst Bacteriol* **37:** 416–421.

Farnelid, H., Andersson, A.F., Bertilsson, S., Abu Al-Soud, W., 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.

Francis, C.A., O'mullan, G.D., Cornwell, J.C., and Ward, B.B. (2013) Transitions in *nirS*-type denitrifier diversity, community composition, and biogeochemical activity along the Chesapeake Bay estuary. *Front Microbiol* **4:** 237. doi: 10.3389/fmicb.2013.00237.

Gandhi, N., Ramesh, R., Srivastava, R., Sheshshayee, M.S., Dwivedi, R.M., and Raman, M. (2010) Nitrogen uptake rates during Spring in the NE Arabian Sea. *Int J Oceanogr* **2010:** 1. doi:10.1155/2010/127493.

Gao, H., Schreiber, F., Collins, G., Jensen, M.M., Kostka, J.E., Lavik, G., *et al*. (2010) Aerobic denitrification in permeable Wadden Sea sediments. *ISME J* **4:** 417–442.

Glud, R.N., Grossart, H.P., Larsen, M., Tang, K.W., Arendt, K.E., Rysgaard, S., *et al*. (2015) Copepod carcasses as microbial hot spots for pelagic denitrification. *Limnol Oceanogr* **60:** 2026–2036.

Graf, D.R.H., Jones, C.M., Hallin, S., and de Crécy-Lagard, V. (2014) Intergenomic comparisons highlight modularity of the denitrification pathway and underpin the importance of community structure for N_2O emissions. *PLoS One* **9:** e114118. doi.org/10.1371/journal.pone.0114118.

Großkopf, T., Mohr, W., Baustian, T., Schunck, H., Gill, D., Kuypers, M.M.M., *et al*. (2012) Doubling of marine dinitrogen-fixation rates based on direct measurements. *Nature* **488:** 361–364.

Härtig, E., and Zumft, W.G. (1999) Kinetics of *nirS* expression (Cytochrome cd 1 Nitrite Reductase) in *Pseudomonas stutzeri* during the transition from aerobic respiration to denitrification: evidence for a denitrification-specific nitrate- and nitrite-responsive regulatory system. *Appl Environ Microbiol* **181:** 161–166.

Hatton, A.D., Shenoy, D.M., Hart, M.C., Mogg, A., and Green, D.H. (2012) Metabolism of DMSP, DMS and DMSO by the cultivable bacterial community associated with the DMSP-producing dinoflagellate *Scrippsiella trochoidea*. *Biogeochemistry* **110:** 131–146.

Hewson, I., Poretsky, R.S., Dyhrman, S.T., Zielinski, B., White, A.E., Tripp, H.J., *et al*. (2009) Microbial community gene expression within colonies of the diazotroph, *Trichodesmium*, from the Southwest Pacific Ocean. *ISME J* **3:** 1286–1300.

Hynes, A.M., Webb, E.A., Doney, S.C., and Waterbury, J.B. (2012) Comparison of cultured Trichodesmium (Cyanophyceae) with species characterized from the field. *J Phycol* **48:** 196–210.

Jayakumar, A., Naqvy, S.W.A., and Ward, B.B. (2009) Distribution and relative quantification of key genes involved in fixed nitrogen loss from the Arabian Sea oxygen minimum zone. In *Indian Ocean Biogeochemical Processes and Ecological Variability. Geophysical Monograph Series 185*. Wiggert, J.D., Hood, R.R., Naqvi, S.W.A., Brink, K.H., and Smith, S.L. (eds.), Washington, DC, USA: American Geophysical Union, pp.187–203.

Ji, Q., Babbin, A.R., Jayakumar, A., Oleynik, S., and Ward, B.B. (2015) Nitrous oxide production by nitrification and denitrification in the Eastern Tropical South Pacific oxygen minimum zone. *Geophys Res Lett* **42:** 10,755–10,764.

Karl, D., Letelier, R., Tupas, L., Dore, J., Christian, J., and Hebel, D. (1997) The role of nitrogen fixation in biogeochemical cycling in the subtropical North Pacific Ocean. *Nature* **388:** 533–538.

King, G.M. (2003) Molecular and culture-based analyses of aerobic carbon monoxide oxidizer diversity. *Appl Environ Microbiol* **69:** 7257–7265.

Kramer, J.G., Wyman, M., Zehr, J.P., and Capone, D.G. (1996) Diel variability in transcription of the structural gene for glutamine synthetase (*glnA*) in natural populations of the marine diazotrophic cyanobacterium *Trichodesmium thiebautii*. *FEMS Microbiol Ecol* **21:** 187–196.

Kuypers, M.M.M. (2015) Microbiology: a division of labour combined. *Nature* **528:** 487–488.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., *et al*. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23:** 2947–2948. doi:10.1093/bioinformatics/btm404.

Lee, M.D., Walworth, N.G., McParland, E.L., Fu, F.-X., Mincer, T.J., Levine, N.M., *et al*. (2017) The *Trichodesmium* consortium: conserved heterotrophic co-occurrence and genomic signatures of potential interactions. *ISME J* **11:** 1813–1824. doi:10.1038/ismej.2017.49.

Lundgren, P., Janson, S., Jonasson, S., Singer, A., and Bergman, B. (2005) Unveiling of novel radiations within *Trichodesmium* cluster by *hetR* gene sequence analysis. *Appl Environ Microbiol* **71:** 190–196.

Mahaffey, C., Michaels, A.F., and Capone, D.G. (2005) The conundrum of marine N2 fixation. *Am J Sci* **305:** 546–595.

Mazard, S.L., Fuller, N.J., Orcutt, K.M., Bridle, O., and Scanlan, D.J. (2004) PCR analysis of the distribution of unicellular cyanobacterial diazotrophs in the Arabian Sea. *Appl Environ Microbiol* **70:** 7355–7364.

Miyahara, M., Kim, S.-W., Fushinobu, S., Takaki, K., Yamada, T., Watanabe, A., *et al*. (2010) Potential of aerobic denitrification by *Pseudomonas stutzeri* TR2 to reduce nitrous oxide emissions from wastewater treatment plants. *Appl Environ Microbiol* **76:** 4619–4625.

Mulholland, M.R., Bernhardt, P.W., Heil, C.A., Bronk, D.A., and O'neil, J.M. (2006) Nitrogen fixation and release of fixed nitrogen by *Trichodesmium* spp. in the Gulf of Mexico. *Limnol Oceanogr* **51:** 1762–1776.

Muyzer, G., de Waal, E.C., and Uitterlinden, A.G. (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* **59:** 695–700.

Nausch, M. (1996) Microbial activities on *Trichodesmium* colonies. *Mar Ecol Prog Ser* **141:** 173–181.

Orcutt, K.M., Rasmussen, U., Webb, E.A., Waterbury, J.B., Gundersen, K., and Bergman, B. (2002) Characterization of *Trichodesmium* spp. by genetic techniques. *Appl Environ Microbiol* **68:** 2236–2245.

Paerl, H.W. (1982). Interactions with bacteria. In *The Biology of Cyanobacteria*. Carr, N.G., and Whitton, B.A. (eds). Oxford, UK: Blackwell, pp. 441−461.

Paerl, H.W., and Bebout, B.M. (1988) Direct measurement of O_2-depleted micro-zones in marine *Oscillatoria*: relation to N_2 fixation. *Science* **241:** 442–445.

Paerl, H.W., and Bebout, B.M. (1992) Oxygen dynamics in *Trichodesmium* spp. aggregates. In *Marine Pelagic Cyanobacteria: Trichodesmium and Other Diazotrophs*. Carpenter, E. J., Capone, D. G. and Rueter, J. G. (eds). Dordrecht, Netherlands: Kluyver Academic Publishers, pp. 43–59.

Palmer, T., and Berks, B.C. (2012) The twin-arginine translocation (TAT) protein export pathway. *Nat Rev Microbiol* **10:** 483–496.

Park, S., Park, D.S., Bae, K.S., and Yoon, J.H. (2014) *Phaeobacter aquaemixtae* sp. nov., isolated from the junction between the ocean and a freshwater spring. *Int J Syst Evol Microbiol* **64:** 1378–1383.

Pearman, J.K., Ellis, J., Irigoien, X., Sarma, Y.V.B., Jones, B.H., and Carvalho, S. (2017) Microbial planktonic communities in the Red Sea: high levels of spatial and temporal variability shaped by nutrient availability and turbulence. *Sci Rep* **7:** 6611. doi:10.1038/s41598-017-06928-z.

Pecoraro, V., Zerulla, K., Lange, C., Soppa, J., and Aziz, S. (2011) Quantification of ploidy in Proteobacteria revealed the existence of monoploid, (mero-)oligoploid and polyploid species. *PLoS One* **6:** e16392. doi:10.1371/journal.pone.0016392.

Prufert-Bebout, L., Paerl, H.W., and Lassen, C. (1993) Growth, nitrogen fixation and spectral attenuation in cultivated *Trichodesmium* species. *Appl Environ Microbiol* **59:** 1367–1375.

Raes, E.J., Bodrossy, L., Van de Kamp, J., Holmes, B., Hardman-Mountford, N., and Thompson, P.A. (2016) Reduction of the powerful greenhouse gas N_2O in the South-Eastern Indian Ocean. *PLoS One.* doi:10.1371/journal.pone.0145996.

Rahav, E., Herut, B., Mulholland, M.R., Belkin, N., Elifantz, H., and Berman-Frank, I. (2015) Heterotrophic and autotrophic contribution to dinitrogen fixation in the Gulf of Aqaba. *Mar Ecol Prog Ser* **522:** 67–77.

Rao, M., Smith, B.C., and Marletta, M.A. (2015) Nitric oxide mediates biofilm formation and symbiosis in *Silicibacter* sp. strain TrichCH4B. *MBio* **6:** e00206–e00215.

Ravishankara, A.R., Daniel, J.S., and Portmann, R.W. (2009) Nitrous oxide (N_2O): the dominant ozone depleting substance emitted in the 21st century. *Science* **326:** 123–125.

Rees, A., Robinson, C., Smyth, T., Aiken, J., Nightingale, P., and Zubkov, M. (2015) 20 years of the Atlantic Meridional Transect – AMT. *Limnol Oceanogr Bull* **24:** 101–124. doi:10.1002/lob.10073.

Samarkin, V.A., Madigan, M.T., Bowles, M.W., Casciotti, K.L., Priscu, J.C., McKay, C.P., and Joye, S.B. (2010) Abiotic nitrous oxide emission from the hypersaline Don Juan Pond in Antarctica. *Nat Geosci* **3:** 341–344.

Sanford, R.A., Wagner, D.D., Wu, Q., Chee-Sanford, J.C., Thomas, S.H., Cruz-Garcia, C., *et al*. (2012) Unexpected nondenitrifier nitrous oxide reductase gene diversity and abundance in soils. *Proc Natl Acad Sci USA* **109:** 19709–19714.

Scavotto, R.E., Dziallas, C., Bentzon-Tilia, M., Riemann, L., and Moisander, P.H. (2015) Nitrogen-fixing bacteria associated with copepods in coastal waters of the North Atlantic Ocean. *Environ Microbiol* **17:** 3754–3765.

Schattenhofer, M., Fuchs, B.M., Amann, R., Zubkov, M.V., Tarran, G.A., and Pernthaler, J. (2009) Latitudinal distribution of prokaryotic picoplankton populations in the Atlantic Ocean. *Environ Microbiol* **11:** 2078–2093.

Sheridan, C., Steinberg, D., and Kling, G. (2002) The microbial and metazoan community associated with colonies of *Trichodesmium* spp.: a quantitative survey. *J Plankton Res* **24:** 913–922.

Sutka, R.L., Ostrom, N.E., Ostrom, P.H., and Phanikumar, M.S. (2004) Stable isotope dynamics of dissolved nitrate in a transect from the North Pacific Subtropical Gyre to the Eastern Tropical North Pacific. *Geochim Cosmochim Acta* **68:** 517–527.

Sun, X., Jayakumar, A., and Ward, B.B. (2017) Community composition of nitrous oxide consuming bacteria in the oxygen minimum zone of the Eastern Tropical South Pacific. *Front Microbiol* **8:** 1183. doi:10.3389/fmicb.2017.01183.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* **30:** 2725–2729.

Trimmer, M., Chronopoulou, P.M., Maanoja, S.T., Upstill-Goddard, R.C., Kitidis, V., and Purdy, K.J. (2016) Nitrous oxide as a function of oxygen and archaeal gene abundance in the North Pacific. *Nat Commun* **7:** 13451. doi:10.1038/ncomms13451.

van Kessel, M.A., Speth, D.R., Albertsen, M., Nielsen, P.H., den Camp, H.J.O., and Kartal, B. (2015) Complete nitrification by a single microorganism. *Nature* **528:** 555–559.

Voss, M., Bange, H.W., Dippner, J.W., Middelburg, J.J., Montoya, J.P., and Ward, B. (2013) The marine nitrogen cycle: recent discoveries, uncertainties and the potential relevance of climate change. *Philos Trans Roy Soc B* **368:** 20130121.

Weber, C.F., and King, G.M. (2007) Physiological, ecological, and phylogenetic characterization of Stappia, a marine CO-oxidizing bacterial genus. *Appl Environ Microbiol* **73:** 1266–1276.

Wyman, M., Gregory, R.P.F., and Carr, N.G. (1985) Novel role for phycoerythrin in a marine cyanobacterium, *Synechococcus* strain DC2. *Science* **230:** 818–820.

Wyman, M., Hodgson, S., and Bird, C. (2013) Denitrifying alphaproteobacteria from the Arabian Sea that express *nosZ*, the gene encoding nitrous oxide reductase, in oxic and suboxic waters. *Appl Environ Microbiol* **79:** 2670–2681.

Zubkov, M.V., Tarran, G.A., and Burkill, P.H. (2006) Bacterioplankton of low and high DNA content in the suboxic waters of the Arabian Sea and the Gulf of Oman: abundance and amino acid uptake. *Aquat Microb Ecol* **43:** 23–32.

Zumft, W.G. (1997) Cell biology and molecular basis of denitrification. *Microbiol Mol Biol Rev* **61:** 533–616.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.

Table S1. Oligonucleotide primers used in this study, PCR cycling conditions and expected product size in base pairs (bp).

Fig. S1. A. Consensus phylogram (1000 bootstrap replicates) of NosZ constructed using a Neighbor-Joining routine in MEGA 6. The tree is based on 464 amino acid residues following the removal of indels and non-conserved regions at the C and N termini from an alignment of 178

published peptide sequences. Two major and three minor clades were identified. The largest (Clade 5, 81 sequences) contains proteobacteria almost exclusively (and three Archaea) that all produce a TAT-dependent NosZ and have the closest match to the target sequence of the *nosZ*F1 primer (DVH/QYQPGH as indicated). The second largest (Clade 1, 70 sequences) incudes members of the Bacteroidetes primarily and shows several variable substitutions in the primer target (consensus PA/TF/YYSP/VGH) as is evident for the three remaining minor clades as indicated after the clade designations. Like the members of clade 1, the N termini of NosZ from these minor clades lack the TAT signal sequence (consensus SRRXFLK; Palmer and Berks, 2012), and they are Sec dependent in localization. The bootstrap support ($>$ 50%) received for each partition is shown at the nodes. B. Graphical representations of the conserved NosZ peptide regions targeted by the forward (left) and reverse (right) primer sets. The graphics were drawn using WebLogo (Crooks *et al*., 2004) and display the consensus for all 174 NosZ sequences analyzed (top) and that for either the TAT-dependent (HM-type, middle) or Sec-dependent (HQ-type, bottom) forms of the enzyme. The nucleotide sequences shown below the middle and bottom rows are reverse translated (amino acid to DNA) versions of the peptides. In the case of the reverse primer target (right) the sequences are reverse complementary to the peptide sequences shown and together with the forward sequences (left) guided the design of the primer sets shown in Supporting Information Table S1.

Fig. S2. A. Alignment of *nosZ* from members of Groups 1, 2, 3 and 4 to show the targets of Group 2 (green) and Group 4 (blue) specific primers and probes (italics letters). Non-matching nucleotides within the target regions of the primers and probes are colored in red. B and C. Ethidium bromide stained agarose gel showing PCR products amplified with (B) Group 2 and (C) Group 4 primers. The DNA template used for each PCR is indicated by the group numbers shown in each lane (e.g., Group 1 in lanes a and b and Group 2 in lanes c and d). Note that the starting quantities of template DNA were not equivalent but were adequate to generate PCR products using the cycling protocols listed in Supporting Information Table S1.

Fig. S3. A. PCR products obtained from Indian Ocean stations 2 (lane a), 3 (lanes b and c) and 4 (lane d) amplified from $>$ 0.2 μm bacterioplankton DNA samples using 16S rRNA gene-specific universal primers (see Supporting Information Table S1). A prominent product of the expected size ($\sim$ 190 bp) is present in each lane indicating the lack of PCR inhibitors in the samples. B. Amplified *nosZ* cDNA products obtained from *Trichodesmium* colony RNA samples from the mid-Atlantic Ocean station 22 (lane b; the minus RT control is shown in lane a) and station 24 (lane d; lane c minus RT control). PCR products from gDNA samples from each of the stations are also shown (lanes e and f respectively). C. *nosZ* cDNAs amplified from two *Trichodesmium* samples obtained from the Red Sea (lanes b and d; the corresponding minus RT reactions are shown in lanes a and c respectively). PCR products from gDNA samples obtained from the same location are also shown (lanes a and b respectively).

Fig. S4. Consensus phylogram (1000 bootstrap replicates) of NosZ peptide sequences derived from total genomic DNA or RNA (cDNA) from *Trichodesmium* collected at two stations in the mid-Atlantic Ocean (clone prefix pAO followed by the station number; blue triangles), three stations in the Indian Ocean (clone prefix pIO followed by the station number; green squares), the Red Sea (clone prefix pRS; red diamonds), Bahama Islands (clone prefix pBAH; yellow circles) and from archived DNA from the isolate, *Trichodesmium* IMS101 (clone prefix pIMS; black circles). The cDNA clones are from the mid-Atlantic and Red Sea samples only and are indicated by closed symbols (clone prefix pAOc; blue triangles and pRSc; red diamonds, respectively) whereas the gDNA clones are highlighted with open symbols. Groups 1 (red lines), 2 (orange lines), 3 (brown lines) and 4 (blue lines) are highlighted. The numbers at the nodes indicate the bootstrap support (%) received for each partition. The tree is rooted with the corresponding NosZ peptide sequence from *Pseudomonas stutzeri*.

Fig. S5. Mean abundance ($\pm$ SE) of nirS and *nosZ* mRNAs in cultures (n = 3) of Group 2 alphaproteobacterium strain 4N under aerobic (O2) and anaerobic (ANO2) growth conditions. The numbers in parentheses above the bars show the corresponding degree of upregulation (fold-differences in abundance vs air-grown cells) of the encoded proteins nitrite reductase (NirS) and nitrous oxide reductase (NosZ) in the absence of oxygen (Matallana-Surget and Wyman, in preparation)