Temporal dynamics of Prochlorococcus cells with the potential for nitrate assimilation in the subtropical Atlantic and Pacific oceans

Temporal dynamics of *Prochlorococcus* cells with the potential for nitrate assimilation in the subtropical Atlantic and Pacific oceans

Paul M. Berube[1,*], Allison Coe[1], Sara E. Roggensack[1], and Sallie W. Chisholm[1,2,*]

[1] Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, 02139, USA

[2] Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts, 02139, USA

[*] Correspondence: SW Chisholm, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Bldg 48-419, Cambridge, Massachusetts, 02139, USA. E-mail: chisholm@mit.edu; PM Berube, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Bldg 48-424, Cambridge, Massachusetts, 02139, USA. E-mail: pmberube@gmail.com

Running Title: Temporal dynamics of *Prochlorococcus narB*

ABSTRACT

Utilization of nitrate as a nitrogen source is broadly conserved among marine phytoplankton, yet many strains of *Prochlorococcus* lack this trait. Among cultured strains, nitrate assimilation has only been observed within two clades of *Prochlorococcus*: the high-light adapted HLII clade and the low-light adapted LLI clade. To better understand the frequency and dynamics of nitrate assimilation potential among wild *Prochlorococcus,* we measured seasonal changes in the abundance of cells containing the nitrate reductase gene (*narB)* in the subtropical North Atlantic and North Pacific oceans. At the Atlantic station, the proportion of HLII cells containing *narB* varied with season, with the highest frequency observed in stratified waters during the late summer, when inorganic nitrogen concentrations were lowest. The Pacific station, with more persistent stratification and lower N:P ratios, supported a perennially stable subpopulation of HLII cells containing *narB*. Approximately 20-50% of HLII cells possessed *narB* under stratified conditions at both sites. Since HLII cells dominate the total *Prochlorococcus* population in both ecosystems, nitrate potentially supports a significant fraction of the *Prochlorococcus* biomass in these waters. The abundance of LLI cells containing *narB* was positively correlated with nitrite concentrations at the Atlantic station. These data suggest that *Prochlorococcus* may contribute to the formation of primary nitrite maxima through incomplete nitrate reduction and highlight the potential for interactions between *Prochlorococcus* and sympatric nitrifying microorganisms. Further examination of these relationships will help clarify the selection pressures shaping nitrate utilization potential in low-light and high-light adapted *Prochlorococcus*.

INTRODUCTION

The cyanobacterium *Prochlorococcus* is often the numerically dominant phototroph in the tropical and subtropical oceans where it contributes substantially to net primary production (Flombaum et al. 2013). This genus encompasses several phylogenetic clades consisting of cells that are high-light adapted (clades HLI – HLVI) or low-light adapted (clades LLI – LLVI) (Huang et al. 2012; Biller et al. 2014). As such, several *Prochlorococcus* clades are associated with sets of physiological properties (Moore and Chisholm 1999; Ahlgren et al. 2006), reflecting adaptations that apparently influence the distributions of these clades along environmental gradients (Bouman et al. 2006; Johnson et al. 2006; West and Scanlan 1999). For example, cells belonging to the high-light adapted HLII clade (eMIT9312 ecotype) have a relatively high optimal temperature for growth and dominate *Prochlorococcus* populations in warmer latitudes (Johnson et al. 2006). Cells belonging to the low-light adapted LLI clade (eNATL2A ecotype) are relatively tolerant of light shock, and unlike cells belonging to other low-light adapted clades, can persist during vertical mixing events that expose cells to higher photon fluxes at the surface (Malmstrom et al. 2010).

The closest relatives of *Prochlorococcus* are the marine *Synechococcus* (Rocap et al. 2002; Scanlan et al. 2009). Although similar in many respects, *Prochlorococcus* differs from *Synechococcus* in several ways, including having smaller and more streamlined genomes (Partensky and Garczarek 2010) and using divinyl chlorophylls *a* and *b* for harvesting light energy (Goericke and Repeta 1992). Most cultured strains of *Synechococcus* are capable of nitrate assimilation (Ahlgren and Rocap 2006), whereas initial *Prochlorococcus* isolates lacked this ability (Moore et al. 2002; García-Fernández et al. 2004; Kettler et al. 2007). As a result, modeling efforts at that time used the absence of nitrate assimilation by *Prochlorococcus* as a defining feature of this phytoplankton group (Follows et al. 2007). That said, this feature of

Prochlorococcus remained an enigma given that nitrogen is often the proximal limiting nutrient controlling phytoplankton growth in marine ecosystems (Tyrrell 1999), and most cyanobacteria have pathways for utilizing nitrate as a source of nitrogen (Ohashi et al. 2011).

More recently, field studies have revealed nitrate uptake by wild populations of *Prochlorococcus* in the North Atlantic Subtropical Gyre (Casey et al. 2007) as well as the presence of nitrate assimilation genes in uncultivated *Prochlorococcus* genomes (Martiny et al. 2009*b*). We then identified and isolated strains capable of nitrate assimilation and have shown that axenic cultures of *Prochlorococcus* are capable of growth using nitrate as the sole nitrogen source (Berube et al. 2015). Comparative genomics of these isolates indicated the possibility of multiple gains and losses of nitrate utilization genes during the divergence of *Prochlorococcus* from *Synechococcus* (Berube et al. 2015); even so, the capacity for nitrate assimilation appears to be more closely tied to ribotype phylogeny than functions associated with phosphorus assimilation (Martiny et al. 2009*a*; Berube et al. 2015). It is now clear that within this single genus, there are cells with distinct differences in their ability to access nitrate in the oceans.

Although nitrate is one of the more abundant forms of nitrogen available to marine phytoplankton (Gruber 2008), nitrate utilization is tightly regulated by cyanobacteria (Ohashi et al. 2011). Cyanobacteria discern the nitrogen status of the cell by sensing changes in the cellular level of 2-oxoglutarate which serves as an acceptor molecule for newly assimilated nitrogen. 2-oxoglutarate accumulates in cells under nitrogen limitation and activates the NtcA transcriptional regulator which is responsible for inducing the expression of multiple nitrogen assimilation genes (Lopatovskaya et al. 2011). Like other cyanobacteria, *Prochlorococcus* responds to nitrogen starvation by activating a putative NtcA regulon (Tolonen et al. 2006). Cyanobacteria generally prefer ammonium as a nitrogen source and quickly inhibit the expression of nitrate assimilation genes in the presence of sufficient concentrations of ammonium (Ohashi et al. 2011). This

preference is often explained by the difference in oxidation state of these compounds (Ohashi et al. 2011). Yet, recent studies of marine *Synecococcus* suggest that a limiting step in the nitrate assimilation pathway, rather than the availability of reducing power, might be partly responsible for this preference (Collier et al. 2012).

What selection pressures control the distribution of nitrate assimilation genes in wild *Prochlorococcus* populations? Metagenomic data from surface waters have indicated that *Prochlorococcus* genes conferring the ability to use nitrate are relatively more abundant in the Caribbean Sea and Indian Ocean (Martiny et al. 2009*b*) – areas with lower than average nitrate concentrations. Yet, how nitrate utilization potential is distributed within wild *Prochlorococcus* populations, and how this distribution is related to temporal and vertical variations in the availability of nitrogen, remain open questions. To address this issue, we looked at seasonal changes in the distribution and abundance of cells containing the gene for nitrate reductase (*narB*) in wild *Prochlorococcus* populations in two open ocean ecosystems: the North Atlantic Subtropical Gyre (Sargasso Sea, Bermuda Atlantic Time-series Study, BATS) and the North Pacific Subtropical Gyre (Station ALOHA, Hawai'i Ocean Time-series, HOT) (Karl and Lukas 1996; Steinberg et al. 2001). *Prochlorococcus* are abundant at both sites (Campbell et al. 1994; DuRand et al. 2001) and numerically dominated by the high-light adapted HLII clade (eMIT9312 ecotype); when averaged over the entire year during our study period, the HLII clade constitutes approximately 60% and 75% of the total *Prochlorococcus* population at BATS and HOT, respectively, as derived from the work of Malmstrom et al. (Malmstrom et al. 2010).

While each of these long standing time-series stations is located in oligotrophic, open ocean waters, they differ in the finer details of their physics and nutrient dynamics. The Atlantic site displays substantial seasonal variations in temperature and nutrient concentrations because it is subject to stronger convective mixing during the winter and early spring; these events disrupt

stratification in the euphotic zone and transport nutrient rich water to the surface (Steinberg et al. 2001). In contrast, the Pacific site is characterized by more consistent stratification of the water column and higher concentrations of inorganic phosphorus (Wu et al. 2000; Cavender-Bares et al. 2001; Steinberg et al. 2001). Indeed, inorganic N:P ratios are often well below the Redfield ratio of 16:1 at the Pacific site, suggesting that cells here are typically limited by nitrogen availability (Wu et al. 2000).

Studies comparing these two sites have been instrumental in understanding how environmental factors influence the population dynamics of *Prochlorococcus* clades and the genes they carry (Malmstrom et al. 2010; Coleman and Chisholm 2010). Here we expand on this comparative approach by examining the presence of the *Prochlorococcus* nitrate reductase gene (*narB*) as a proxy for nitrate assimilation potential by *Prochlorococcus* at these sites.

MATERIALS AND METHODS

Gene target for nitrate assimilation potential. Nitrate reductase (NarB) is the first enzyme in the nitrate assimilation pathway and its gene (*narB*) is a useful marker for inferring the potential for this trait in wild phytoplankton populations (Paerl et al. 2011; Paerl et al. 2012). Analysis of metagenomic data from the Global Ocean Sampling (GOS) expedition revealed two variants of *Prochlorococcus narB* (Martiny et al. 2009*b*) – one with a GC content of ~30% and adjacent sequences most closely related to HLII *Prochlorococcus* genomes, and the other with a GC content of ~40% and adjacent sequences most closely related to LLI *Prochlorococcus* genomes. These two variants will hereafter be referred to as 'HLII *narB'* and 'LLI *narB'*. To date, nitrate assimilation genes have only been observed in cultured isolates of *Prochlorococcus* belonging to the HLII and LLI clades (Berube et al. 2015). Based on genomic and metagenomic data, these genes are typically colocalized in a single region of the chromosome in these clades (Martiny et al. 2009*b*; Berube et al. 2015), suggesting that *Prochlorococcus* with the potential for nitrate assimilation usually contain a single copy of *narB*. Thus, gene copy number is assumed to generally represent *narB*-containing cell abundance. More recently, a metagenomics study of anoxic marine zones uncovered evidence of nitrate assimilation genes within the genomes of low-light adapted *Prochlorococcus* possibly belonging to the LLV or LLVI clades (Astorga-Eló et al. 2015). While we continue to search for these genes in other clades, our analyses here are restricted to the HLII and LLI clades. Fortunately, these two clades represent the most abundant high- and low-light adapted *Prochlorococcus* groups at our study sites (Malmstrom et al. 2010), and thus, among *Prochlorococcus*, they likely have the strongest biogeochemical imprint on these systems.

Primer design. The design of quantitative polymerase chain reaction (qPCR) primers for the detection of cells containing HLII *narB* and LLI *narB* (Table 1) was based on an alignment of

Prochlorococcus narB sequences derived from cultures (Berube et al. 2015) and several ocean metagenome databases in which *narB* gene sequences could be linked to *Prochlorococcus* DNA (Venter et al. 2004; Rusch et al. 2007; DeLong et al. 2006; Martiny et al. 2009*b*). Sequences were also identified in HOT and BATS metagenomes (Coleman and Chisholm 2010) based on similarity to the *narB* sequences observed in *Prochlorococcus* genomes. Since the number of LLI clade sequences are eclipsed by the relatively high number of HLII clade sequences in most surface water metagenomes, additional *narB* sequences were obtained from *narB* clone libraries prepared from DNA obtained at depths of 75m and 125m on the HOT179 cruise at Station ALOHA. These clone libraries were constructed using degenerate primers narB34F and narB2099R (Table 1) which target both variants of the nearly full length *Prochlorococcus narB*. The *narB* amplicons were cloned into pCR4 (Life Technologies, Grand Island, NY, USA) and sequenced at the Dana-Farber/Harvard Cancer Center DNA Resource Core. Partial putative *narB* sequences have been deposited in GenBank (accession numbers KM411385-KM411429).

Prochlorococcus and *Synechococcus narB* nucleotide sequences were aligned by codon using transAlign (Bininda-Emonds 2005) and manually curated. The DISTMAT tool in the EMBOSS software suite (Rice et al. 2000) was used to evaluate sequence similarity. Sequences within the HLII *narB* group were 96.3 ± 4.5 percent identical and sequences within the LLI *narB* group were 95.3 ± 2.2 percent identical. Sequences between the HLII and LLI *narB* groups were 62.3 ± 3.2 percent identical. Using this alignment, qPCR primers were designed against the *Prochlorococcus* HLII and LLI *narB* consensus sequences using NCBI Primer-BLAST (Ye et al. 2012). Degenerate nucleotides at variable positions were introduced as long as they did not significantly impact specificity or amplification efficiency. Based on the aligned sequence data, these primers are expected to capture at least 80% of the diversity of each *narB* variant (the proportion of *Prochlorococcus narB* sequences with exact matches to our primer sequences).

Primer specificity was further examined with the in-silico polymerase chain reaction (isPCR) algorithm (Kuhn et al. 2013) using the NCBI RefSeq Release 71 database in order to assess the possibility that these primers could amplify sequences other than *Prochlorococcus narB*. No other potential targets were observed in this database when using the default isPCR parameters (4000 bp maximum amplicon size, a minimum of 15 perfect matches at the 3' end of each primer, and a minimum length of 15 bp where there must be 2 matches for each mismatch). Relaxing the stringency to a minimum of 5 perfect matches at the 3' end of each primer only identified weak hits to terrestrial animals, soil bacteria, and *Halomonas* sp. GFAJ-1 (an extremophile isolated from a hypersaline, alkaline lake). None of these are predicted to result in PCR amplicons given the high proportions of mismatches across the length of the primers and none of these organisms are found at significant concentrations in the oligotrophic open ocean. No *Synechococcus* targets were observed, which is likely a reflection of the high degree of divergence in GC content between *Synechococcus* and the HLII and LLI clades of *Prochlorococcus* (Scanlan et al. 2009). Thus, our primers are expected to be specific for each variant of *Prochlorococcus narB*.

qPCR standards. Cells of *Prochlorococcus* SB (Shimada et al. 1995; Berube et al. 2015) and *Prochlorococcus* MIT0917, both containing a single copy of *narB*, were used as standards in qPCR assays to measure the abundance of cells containing the HLII and LLI variants of *narB* respectively. These standards were processed in the same manner as field samples as previously described (Ahlgren et al. 2006; Zinser et al. 2007; Malmstrom et al. 2010). While a single *Prochlorococcus* strain was observed to contain duplicate copies of *narB*, genomic and metagenomic evidence suggests that this is a rare occurrence and that most *narB*-containing *Prochlorococcus* contain a single copy of this gene (Berube et al. 2015). The MIT0917 strain is a derivative of the P0903-H212 enrichment culture (Berube et al. 2015) and has been deemed

199 unialgal based on observations of a single *Prochlorococcus* population using flow cytometry and
200 by the presence of a single 16S-23S rRNA internal transcribed spacer (ITS) sequence as
201 determined by direct sequencing of its ITS PCR amplicon. The ITS sequence for MIT0917 has
202 been deposited in GenBank under accession number KM281884.
203 Each strain was maintained at 24°C in Pro99 medium (Moore et al. 2007) containing 800
204 μmol L^{-1} of nitrate as the sole nitrogen source. SB was grown at a constant illumination of ~30
205 μmol quanta m^{-2} s^{-1} and MIT0917 was grown at a constant illumination of ~10 μmol quanta m^{-2}
206 s^{-1}. Cells were enumerated using an Influx Cell Sorter (BD Biosciences, San Jose CA, USA) as
207 previously described (Olson et al. 1985; Cavender-Bares et al. 1999).
208 ***qPCR assay conditions***. Reaction parameters for qPCR were optimized on a set of
209 plasmid clones containing full-length *Prochlorococcus narB* sequences, genomic DNA from
210 *Prochlorococcus* strains MIT0604, SB, and MIT0917, and environmental DNA obtained at
211 depths of 75 and 125m on the HOT179 cruise. Plasmids containing the *narB* sequence from
212 *Synechococcus* WH8102 and *Synechococcus* WH7803 were used as negative controls; these
213 strains belong to the 5.1A and 5.1B groups of *Synechococcus* subcluster 5.1, the dominant
214 subcluster of *Synechococcus* inhabiting oligotrophic open ocean waters. All assays were
215 performed on a LightCycler 480 System (Roche Applied Science, Indianapolis, IN, USA).
216 Reaction conditions were optimized across gradients of temperature and primer concentrations.
217 Standards and samples were processed and analyzed as described previously (Ahlgren et al.
218 2006; Zinser et al. 2007; Malmstrom et al. 2010) in order to facilitate comparison with ITS qPCR
219 measurements of *Prochlorococcus* ecotype abundance (Malmstrom et al. 2010). Assays were
220 performed in 15 µl reaction volumes with 6 µl template and the following final concentrations of
221 reaction components: 1x QuantiTect SYBR Green PCR Mix (Qiagen, Germantown, MD, USA)
222 and 0.5 μmol L^{-1} of each forward and reverse primer. Reactions were pre-incubated at 95°C for

15 min to activate the polymerase and then cycled at 95°C for 15 s, 60°C for 15 s, and 72°C for 30 s. The HLII *narB* assay (using primers ProHLIInarB-qPCR-F and ProHLIInarB-qPCR-R) used 45 amplification cycles and the LLI *narB* assay (using primers ProLLInarB-qPCR-F and ProLLInarB-qPCR-R) used 50 amplification cycles. The HLII *narB* and LLI *narB* primer sets did not amplify the *Synechococcus narB* negative controls. PCR amplification efficiencies and limits of quantification (defined as the minimum number of cells per reaction that remained within the linear portion of standard curves) were determined using serial dilutions of *Prochlorococcus* cells that were processed as qPCR standards as previously described (Ahlgren et al. 2006; Zinser et al. 2007; Malmstrom et al. 2010). The amplification efficiency for the HLII *narB* assay was 102% and 95% respectively for strains SB and MIT0604, with a limit of quantification of 7 cells per reaction (equivalent to 8 cells ml^{-1} seawater). The amplification efficiency for the LLI *narB* assay was 89% and 92% respectively for MIT0917 and MIT0915, with a limit of quantification of 29 cells per reaction (equivalent to 32 cells ml^{-1} seawater). The MIT0915 strain is a derivative of the P0902-H212 enrichment culture (Berube et al. 2015). Melting curve analysis was used to assess specificity in each reaction as previously described (Malmstrom et al. 2010).

Validation of the qPCR assay. The qPCR assay was validated on a set of qPCR samples obtained on the HOT186 and BATS216 cruises for which there also exists metagenome sequence data from depths of 25, 75, and 110m for HOT186 and depths of 50 and 100m for BATS216 (Coleman and Chisholm 2010). These metagenome sequence libraries enabled us to use an independent method to examine the frequency of *narB* in the total *Prochlorococcus* population at each site and evaluate the specificity of our qPCR assay (Table 2). For the metagenome data sets, sequence reads were mapped to consensus HLII and LLI *narB* sequences using BLASTN with a minimum bit score of 40 and a minimum read coverage of 20%. Hits were then screened against the NCBI nt database using BLASTN to ensure they did not have better matches to taxa other

than *Prochlorococcus* or to genes other than *narB*. The counts of single copy core *Prochlorococcus* genes (Coleman and Chisholm 2010) and *Prochlorococcus narB* genes in each metagenome library were normalized to gene length to account for the probability of detecting the respective gene. The frequency of each *narB* variant in the *Prochlorococcus* population was determined by dividing the normalized occurrence of *Prochlorococcus narB* genes by the average normalized occurrence of single copy core *Prochlorococcus* genes (Table 2). For the qPCR assay, the frequency of each *narB* variant in the total *Prochlorococcus* population was determined by dividing the abundance of cells containing each *narB* variant by the total abundance of *Prochlorococcus* measured by flow cytometry (Table 2). In the event that qPCR samples were not obtained at the same depth as the metagenome library, the closest depth was also evaluated. Each method resulted in similar measurements of *narB* gene frequency, thus indicating that our qPCR assay is specific for each variant of *Prochlorococcus narB* and that it detects most of the HLII *narB* and LLI *narB* sequences at our study sites (Table 2).

Analysis and interpretation of Prochlorococcus narB abundance data. The abundance of *Prochlorococcus* cells containing HLII *narB* and LLI *narB* genes was determined at the Atlantic and Pacific sites over 2 full seasonal cycles (October 2005 – December 2007), using the identical field samples examined by Malmstrom et al. (Malmstrom et al. 2010). In addition, the abundance of cells belonging to the eMIT9312 and eNATL2A ecotypes of *Prochlorococcus* for the HOT186 cruise were determined for the first time using the methods described by Malmstrom et al. (2010). *Prochlorococcus* cell abundance data based on qPCR have been deposited with the Biological and Chemical Oceanography Data Management Office (BCO-DMO).

Correspondence between the abundance of HLII cells (eMIT9312 ecotype) measured by qPCR and total *Prochlorococcus* cells measured by flow cytometry indicates that ITS primers detect the majority of HLII cells at both sites (Zinser et al. 2006; Malmstrom et al. 2010). Thus, it

was possible to estimate the proportion of HLII cells with the potential for nitrate assimilation by normalizing the abundance of cells containing the HLII *narB* gene to the abundance of *Prochlorococcus* cells belonging to the HLII clade (Malmstrom et al. 2010). Proportions were only determined if the absolute concentrations of both HLII cells and cells containing HLII *narB* (assuming single copies of the ITS sequence and *narB* gene per cell, respectively) were greater than 50 cells ml^{-1} in order to avoid potentially aberrant ratios.

In contrast, we could not normalize the abundance of cells containing the LLI *narB* gene to the abundance of *Prochlorococcus* cells belonging to the LLI clade (eNATL2A ecotype) because our ITS primers are not able to capture all of the low-light adapted *Prochlorococcus*. The sum of ITS qPCR measurements for 5 *Prochlorococcus* ecotypes tends to underestimate total *Prochlorococcus* abundance in deep waters where low-light adapted cells comprise a significant fraction of the total population (Zinser et al. 2006; Ahlgren et al. 2006; Malmstrom et al. 2010). To confirm the suspicion that ITS primers may be missing LLI cells, we tested the ability of LLI clade ITS qPCR primers (NATL3f and NATL2r) (Ahlgren et al. 2006; Malmstrom et al. 2010) to detect MIT0917, which has an ITS sequence that falls within the LLI clade. Our ITS qPCR primers could not detect MIT0917. While we have tried to redesign the LLI clade ITS qPCR primer set using an extensive data set of sequences from ITS sequence clones (Martiny et al. 2009*a*), we have not been successful. Thus, in the context of this study, we could only determine the absolute abundance of LLI cells containing the *narB* gene (assuming a single copy of *narB* per cell) and not their frequency of occurrence within the total LLI population.

Environmental data and statistical analyses. *Prochlorococcus* ecotype abundance, mixed layer depths, and photosynthetically active radiation (PAR) estimates were obtained from Malmstrom et al. (Malmstrom et al. 2010). Chemiluminescent based measurements of low-level nitrogen (nitrate + nitrite) concentrations from the HOT program and colorimetric based

measurements of nitrite and nitrate + nitrite concentrations from the BATS program were obtained from the Ocean Data View website (http://odv.awi.de/en/data/ocean/). Nitrite concentration data were not collected by the HOT program during the period of our study. Nitrate + nitrite concentrations were interpolated using weighted-average gridding and then mapped onto our sampling depths using 2D estimation in Ocean Data View 4.6.2. All data were log (x + 1) transformed before statistical analysis. Spearman rank correlation coefficients were calculated in R using the stats package to assess the relationship between the abundance of cells containing HLII *narB* or LLI *narB* and the abundance of *Prochlorococcus* cells belonging to the HLII and LLI clades. Partial Spearman correlation coefficients were calculated in R using the ppcor package to determine the relationship between the abundance of *narB*-containing *Prochlorococcus* and nitrite or total inorganic nitrogen (nitrate + nitrate) concentrations while controlling for the influence of light.

RESULTS AND DISCUSSION

Seasonal changes in narB occurrence in high-light adapted Prochlorococcus cells. The concentrations of both total *Prochlorococcus* (measured by flow cytometry) and those belonging to the HLII clade (measured by qPCR) are typically higher at the Pacific site compared to the Atlantic site (Malmstrom et al. 2010) (Fig. 1a), as were the abundances of cells in this clade that contained *narB* (Fig. 1b). Total HLII clade cells and cells that contained the HLII *narB* gene were co-localized as indicated by a strong positive correlation between their abundances (Spearman correlation; Pacific, $R = 0.89$, $p < 0.01$; Atlantic, $R = 0.95$, $p < 0.01$). Twenty to 50% of HLII clade cells were estimated to possess *narB* throughout much of the year at the Pacific site, but the same was true only during the late summer and autumn at the Atlantic site (Fig. 1c). During the winter months at the Atlantic site, the proportion of HLII cells containing *narB* drops to less than 10% of the total HLII population (Fig. 1c).

We examined these trends in the context of nutrient supply using nitrate + nitrite concentrations (Fig. 1d) as a proxy for the availability of inorganic nitrogen (ammonium concentrations were not measured by the HOT and BATS programs during the period of our study). While we expect nitrogen limiting conditions to be more prevalent when inorganic nitrogen concentrations are at their lowest, assessing the limiting nutrient for phytoplankton growth is complicated by a number of factors: the availability of measurements for other potentially limiting nutrients (e.g. phosphorus or iron) which are often below limits of detection, knowledge of the bioavailability of different chemical forms of nutrients, and consideration of the flexible elemental stoichiometry of marine microorganisms (Mulholland and Lomas 2008). Further, nutrient fluxes rather than concentrations are likely more important determinants of nutrient limitation. That said, evidence suggests that *Prochlorococcus* is limited by nitrogen at the Pacific site (Van Mooy and Devol 2008), where phosphate concentrations are often more than

an order of magnitude greater than at the Atlantic site (Wu et al. 2000; Cavender-Bares et al. 2001). In contrast, both nitrogen and phosphorus are known to limit primary production at the Atlantic site (Moore et al. 2013), where seasonal dynamics contribute to alternating states of phosphorus limitation in the winter and nitrogen limitation in the summer (Wu et al. 2000).

Cyanobacteria experiencing nitrogen limitation typically induce the expression of alternative nitrogen assimilation pathways; thus we expect that cells with the genomic potential for the assimilation of nitrate will be at a selective advantage under nitrogen limiting conditions. Our data are generally consistent with this hypothesis. At the Atlantic site, the abundance of HLII cells containing *narB* increased with decreasing concentrations of inorganic nitrogen (nitrate + nitrite) when the effect of light availability was taken into account (Spearman partial correlation coefficient, $R = -0.55$, $p < 0.01$), but was independent of nitrite alone (Spearman partial correlation coefficient, $R = -0.02$, $p = 0.71$). Thus, as intense nutrient cycling is expected to drive the system towards nitrogen limitation when the water column is stratified (Wu et al. 2000), we observe an increase in the proportion of HLII cells with the potential for nitrate assimilation (Fig. 1c). A weaker, but also significant negative correlation between inorganic nitrogen concentrations and the abundance of HLII cells containing *narB* was observed at the Pacific site (Spearman partial correlation coefficient, $R = -0.26$, $p < 0.01$), which has low inorganic nitrogen concentrations throughout the year. Overall, these data are consistent with more general patterns observed in metagenomic data sets which have shown greater occurrence of the *Prochlorococcus narB* gene among sequences obtained from low nitrogen waters (Martiny et al. 2009*b*).

While most *narB*-containing HLII cells were found in the nitrogen depleted surface layers at both sites, we also observed several instances of HLII populations with relatively high proportions of *narB*-containing cells in the lower reaches of the euphotic zone (Fig. 1c). Short-lived (<10 days) vertical transport events supplying nitrate to the base of the euphotic zone at the

355 Pacific site have been previously observed (Johnson et al. 2010). Given that *Prochlorococcus*
356 cells are primarily nitrogen limited at the Pacific site (Van Mooy and Devol 2008), these
357 injections of nitrate into the euphotic zone may result in higher relative growth rates of *narB*-
358 containing cells at the base of the euphotic zone. While the HOT program does not measure
359 ammonium concentrations because they are below the 50 nM detection limit of their colorimetric
360 assay (Chiswell et al. 1990), it is expected that extremely low ammonium concentrations would
361 result in the expression of nitrate assimilation genes in cyanobacteria (Ohashi et al. 2011). Thus,
362 it is possible that low ammonium availability coupled with intermittent nitrate injections might
363 contribute to the episodic enrichment of *narB* genotypes in HLII populations under these higher
364 nitrate conditions. Greater temporal and spatial sampling resolution, facilitated by automated
365 instrumentation (Ottesen et al. 2011), as well as a better understanding of ammonium
366 concentrations and fluxes, could help resolve the causality of these features.
367 Selection for genes conferring traits that facilitate access to additional pools of a particular
368 limiting nutrient has been observed previously in *Prochlorococcus.* For example, phosphorus
369 assimilation genes such as alkaline phosphatase and phosphonate transporters are more abundant
370 in phosphorus limited environments relative to phosphorus replete environments (Rusch et al.
371 2007; Coleman and Chisholm 2010; Feingersch et al. 2012). Similarly, *Prochlorococcus* cyanate
372 transporter genes are highly abundant in stratified and nitrogen depleted waters, but not in mixed
373 water columns, of the Red Sea (Kamennaya et al. 2008; Kamennaya and Post 2013). Given that
374 the capacity for nitrate assimilation appears more closely tied to ribotype than many phosphorus
375 acquisition functions (Martiny et al. 2009*a*; Berube et al. 2015), our observations could be
376 explained by covariation of nitrate assimilation potential and other genes under selection. We
377 consider it more likely, however, that nitrogen limitation drives the selection of high-light

adapted *Prochlorococcus* cells capable of nitrate assimilation and potentially other nitrogen assimilation pathways.

Distribution patterns of low-light adapted Prochlorococcus with the potential for nitrate assimilation. Cells belonging to all low-light adapted clades of *Prochlorococcus* reach their maximum abundances deeper in the water column (Malmstrom et al. 2010) where fewer photons penetrate and nutrients are typically at higher concentrations (Letelier et al. 2004; West and Scanlan 1999). Accordingly, low-light adapted *Prochlorococcus* are usually at negligible concentrations in the surface mixed layer. The LLI clade (eNATL2A ecotype) however, which is the focus of our analyses here, stands out among low-light adapted cells in that they tolerate intermittent periods of excess light intensity during deep mixing events that transport cells to the well lit surface (Zinser et al. 2007; Malmstrom et al. 2010). At the same time, the LLI cells are distinct from high-light adapted *Prochlorococcus* with regard to their gene content, photophysiology, and phylogenetic similarity to other low-light adapted clades (Kettler et al. 2007; Malmstrom et al. 2010).

Although we could not normalize the abundance of cells containing LLI *narB* to total LLI cell abundance, as we were able to do for the HLII cells (see methods), we can report absolute counts of cells containing LLI *narB* (Fig. 2,3). This provides information on where this subgroup of LLI cells are located in the water column, and how their distribution and abundance changes with season. The spatial distribution of cells containing LLI *narB* overlaps with that of total LLI cells (Spearman correlation; Pacific, $R = 0.97$, $p < 0.01$; Atlantic, $R = 0.94$, $p < 0.01$) (Fig. 2a,b). Cells containing LLI *narB* are almost fully restricted to intermediate depths of the euphotic zone at both sites, except during periods of deep mixing when their range expands into the surface waters, albeit in low numbers. This is consistent with our understanding of the photophysiology

401 of members of this clade – e.g. their tolerance of light stress, relative to other low-light adapted
402 *Prochlorococcus*, when mixed into surface waters.

403 We observed that cells containing the LLI *narB* gene are found in greatest abundance just
404 above or within the top layer of the nitracline (Fig. 2b,c), and in close proximity to the primary
405 nitrite maximum at the Atlantic site (Fig. 3). Because these are low-light adapted cells, however,
406 we know that they are relatively more abundant in deeper waters because of their adaptation to
407 low light intensities. Controlling for the inverse relationship between light and inorganic nitrogen
408 in the water column, there is no significant relationship between the abundance of cells
409 containing LLI *narB* and nitrate + nitrite concentration at either site (Spearman partial correlation
410 coefficient; Pacific, $R = 0.03$, $p = 0.70$; Atlantic, $R = 0.00$, $p = 0.95$). At the Atlantic site
411 however, there is a positive correlation between the abundance of cells containing LLI *narB* and
412 nitrite concentration when light is taken into account (Spearman partial correlation coefficient;
413 Atlantic, $R = 0.54$, $p < 0.01$). Although independent nitrite measurements are not available from
414 the Pacific site, historical data indicate that maximal nitrite concentrations occur at a mean depth
415 of 126 m at the Pacific site (Dore and Karl 1996) – close to depths where cells containing LLI
416 *narB* are most abundant.

417 What can the proximity of the LLI *narB* populations to the nitracline and the primary
418 nitrite maximum suggest about the selection pressures determining the distribution and
419 abundance of *narB* in low-light adapted cells? The primary nitrite maximum is a common feature
420 of both the Pacific and the Atlantic sites (Dore and Karl 1996; Lomas and Lipschultz 2006). This
421 peak in nitrite concentration may result from either the production of nitrite by ammonia
422 oxidizing microorganisms (Dore and Karl 1996; Newell et al. 2013) or by the incomplete
423 reduction of nitrate by phytoplankton (Dore and Karl 1996; Lomas and Lipschultz 2006). At
424 present, it is unclear which process is the primary driver of the formation of the nitrite maximum,

and it may be that these processes are subject to diel and seasonal variability (Mackey et al. 2011). But, given that LLI clade cells and ammonia oxidizing microorganisms potentially coexist within the primary nitrite maximum, competition for ammonium may be intense enough to favor the selection of LLI cells capable of assimilating nitrate. Further, *Prochlorococcus* could contribute to the formation of the primary nitrite maximum through incomplete nitrate reduction and excretion of nitrite.

Implications for nitrate assimilation by Prochlorococcus in subtropical gyres. We have shown that high-light adapted *Prochlorococcus* cells belonging to the HLII clade (eMIT9312 ecotype) – the most abundant group of *Prochlorococcus* in the oceans by an order of magnitude – are more likely to possess the *narB* gene in the surface of highly stratified waters characterized by nitrogen depletion (i.e. during the summer in the Sargasso Sea and throughout much of the year in the North Pacific Subtropical Gyre). This suggests that *narB*-containing HLII *Prochlorococcus* have a greater selective advantage under low nitrogen conditions and contrasts with a hypothesis we put forth some years ago, derived from a global simulation model, which posited that *Prochlorococcus* cells would be more likely to lose *narB* under similar conditions (Bragg et al. 2010). In light of our current study, this model could possibly be improved by parameterizing different concentrations of ammonium under which nitrate assimilation genes are expressed for different phytoplankton groups. Evaluating the costs and benefits of maintaining the capacity for nitrate assimilation as well as considering the potential role of frequency dependent selection (Bragg et al. 2010; Cordero and Polz 2014) could also aid in the refinement of this model.

Potential biogeochemical and ecological interactions between *Prochlorococcus* and nitrifying microorganisms should also be examined in light of our evidence that *narB*-carrying *Prochlorococcus* are abundant under potentially nitrogen limiting conditions. Nitrification occurs

449 throughout the euphotic zone and is a significant source of nitrate for marine phytoplankton
450 (Yool et al. 2007; Clark et al. 2008). Given the high substrate affinities exhibited by oligotrophic
451 ammonia oxidizing microorganisms (Martens-Habbena et al. 2009; Newell et al. 2013),
452 *Prochlorococcus* could face intense competition for ammonium. Under these conditions,
453 *Prochlorococcus* cells that are capable of nitrate assimilation may be at a selective advantage
454 because they can utilize the end products of both ammonia oxidation (nitrite) and nitrite oxidation
455 (nitrate). A greater understanding of *Prochlorococcus'* growth kinetics on these nitrogen sources,
456 as well as knowledge of the co-occurrence patterns of *narB*-containing *Prochlorococcus* and
457 nitrifiers, will be needed to elucidate these potential interactions.

458 Further, our understanding of the occurrence of *narB* in other high-light adapted
459 *Prochlorococcus* clades is limited given that cells belonging to the HLII clade are over-
460 represented in culture collections (Biller et al. 2014; Biller et al. 2015) and also dominate regions
461 from which most metagenomic data sets have been collected (Rusch et al. 2007). There is
462 evidence, however, that cells belonging to the HLIII and HLIV clades do not contain the genes
463 necessary for nitrate assimilation due to the high iron requirement of this pathway and the
464 observed dominance of these clades in iron limited and hence high nitrogen environments (Rusch
465 et al. 2010; Malmstrom et al. 2013). Our understanding of the distribution of nitrate assimilation
466 potential among low-light adapted *Prochlorococcus* is also influenced by the limitations of our
467 culture collection (Biller et al. 2015), but we are optimistic that targeted isolations or single cell
468 genome analysis will uncover more *narB*-containing low-light adapted cells. Single cell
469 sequencing designed to explore the evolutionary history of the *narB* gene among diverse
470 *Prochlorococcus* ribotypes will be particularly powerful in resolving how genes conferring the
471 potential for nitrate assimilation may co-vary with genetic markers for other ecologically relevant
472 attributes as it is clear that the distribution of specific traits among cells within the

473 *Prochlorococcus* 'federation' (Biller et al. 2015) is the result of a complex interplay between
474 many different selective pressures in the marine environment.

475 Finally, given *Prochlorococcus*' numerical dominance and its potential for efficient
476 nutrient uptake facilitated by its small size (Chisholm 1992), it likely contributes to setting the
477 lower bounds of inorganic nitrogen concentrations in the surface waters of these systems. We
478 note that *Prochlorococcus* cells containing *narB* were particularly abundant in the water column
479 of the Atlantic site when Casey et al. first observed nitrate uptake by wild *Prochlorococcus* in the
480 autumn of 2005 (Casey et al. 2007). Thus, it is possible that the *Prochlorococcus* genotypes
481 detected in our study could be assimilating nitrate in the wild. The fact that at least 50% of
482 *Prochlorococcus* lack the ability to utilize nitrate in these environments, however, suggests that
483 carrying nitrate assimilation genes imposes a fitness cost to some cells, perhaps associated with
484 maintaining these genes in the genome or regulating their expression.

REFERENCES

Ahlgren, N. A., and G. Rocap. 2006. Culture isolation and culture-independent clone libraries reveal new marine *Synechococcus* ecotypes with distinctive light and N physiologies. Appl. Environ. Microbiol. **72**: 7193-7204.

Ahlgren, N. A., G. Rocap, and S. W. Chisholm. 2006. Measurement of *Prochlorococcus* ecotypes using real-time polymerase chain reaction reveals different abundances of genotypes with similar light physiologies. Environ. Microbiol. **8**: 441-454.

Astorga-Eló, M., S. Ramírez-Flandes, E. F. DeLong, and O. Ulloa. 2015. Genomic potential for nitrogen assimilation in uncultivated members of *Prochlorococcus* from an anoxic marine zone. ISME J. **9**: 1264-1267.

Berube, P. M., S. J. Biller, A. G. Kent, J. W. Berta-Thompson, S. E. Roggensack, K. H. Roache-Johnson, M. Ackerman, L. R. Moore, J. D. Meisel, D. Sher, L. R. Thompson, L. Campbell, A. C. Martiny, and S. W. Chisholm. 2015. Physiology and evolution of nitrate acquisition in *Prochlorococcus*. ISME J. **9**: 1195-1207.

Biller, S. J., P. M. Berube, J. W. Berta-Thompson, L. Kelly, S. E. Roggensack, L. Awad, K. H. Roache-Johnson, H. Ding, S. J. Giovannoni, G. Rocap, L. R. Moore, and S. W. Chisholm. 2014. Genomes of diverse isolates of the marine cyanobacterium *Prochlorococcus*. Scientific Data **1**: 140034, doi:10.1038/sdata.2014.34

Biller, S. J., P. M. Berube, D. Lindell, and S. W. Chisholm. 2015. *Prochlorococcus*: the structure and function of collective diversity. Nat. Rev. Microbiol. **13**: 13-27.

Bininda-Emonds, O. R. 2005. transAlign: using amino acids to facilitate the multiple alignment of protein-coding DNA sequences. BMC Bioinformatics **6**: 156.

Bouman, H. A., O. Ulloa, D. J. Scanlan, K. Zwirglmaier, W. K. Li, T. Platt, V. Stuart, R. Barlow, O. Leth, L. Clementson, V. Lutz, M. Fukasawa, S. Watanabe, and S. Sathyendranath. 2006.

509 Oceanographic basis of the global surface distribution of *Prochlorococcus* ecotypes. Science
510 **312**: 918-921.
511 Bragg, J. G., S. Dutkiewicz, O. Jahn, M. J. Follows, and S. W. Chisholm. 2010. Modeling
512 selective pressures on phytoplankton in the global ocean. PLoS ONE **5**: e9569.
513 Campbell, L., H. A. Nolla, and D. Vaulot. 1994. The importance of *Prochlorococcus* to
514 community structure in the central North Pacific Ocean. Limnol. Oceanogr. **39**: 954-961.
515 Casey, J. R., M. W. Lomas, J. Mandecki, and D. E. Walker. 2007. *Prochlorococcus* contributes
516 to new production in the Sargasso Sea deep chlorophyll maximum. Geophys. Res. Lett. **34**:
517 L10604.
518 Cavender-Bares, K. K., D. M. Karl, and S. W. Chisholm. 2001. Nutrient gradients in the western
519 North Atlantic Ocean: relationship to microbial community structure and comparison to
520 patterns in the Pacific Ocean. Deep-Sea Res. Part I-Oceanogr. Res. Pap. **48**: 2373-2395.
521 Cavender-Bares, K. K., E. L. Mann, S. W. Chisholm, M. E. Ondrusek, and R. R. Bidigare. 1999.
522 Differential response of equatorial Pacific phytoplankton to iron fertilization. Limnol.
523 Oceanogr. **44**: 237-246.
524 Chisholm, S. W. 1992. Phytoplankton size, p. 213-237. *In* P. G. Falkowski and A. D. Woodhead
525 [eds.], Primary productivity and biogeochemical cycles in the sea. Plenum Press.
526 Chiswell, S., E. Firing, D. Karl, R. Lukas, and C. Winn. 1990. Hawaii Ocean Time-series
527 Program Data Report 1, 1988-1989. SOEST Tech. Rept. #1, School of Ocean and Earth
528 Science and Technology, Univ. of Hawaii, Honolulu, HI.
529 Clark, D. R., A. P. Rees, and I. Joint. 2008. Ammonium regeneration and nitrification rates in the
530 oligotrophic Atlantic Ocean: implications for new production estimates. Limnol. Oceanogr.
531 **53**: 52-62.

532 Coleman, M. L., and S. W. Chisholm. 2010. Ecosystem-specific selection pressures revealed
533 through comparative population genomics. Proc. Natl. Acad. Sci. USA **107**: 18634-18639.
534 Collier, J. L., R. Lovindeer, Y. Xi, J. C. Radway, and R. A. Armstrong. 2012. Differences in
535 growth and physiology of marine *Synechococcus* (Cyanobacteria) on nitrate versus ammonium
536 are not determined solely by nitrogen source redox state. J. Phycol. **48**: 106-116.
537 Cordero, O. X., and M. F. Polz. 2014. Explaining microbial genomic diversity in light of
538 evolutionary ecology. Nat. Rev. Microbiol. **12**: 263-273.
539 DeLong, E. F., C. M. Preston, T. Mincer, V. Rich, S. J. Hallam, N. U. Frigaard, A. Martinez, M.
540 B. Sullivan, R. Edwards, B. R. Brito, S. W. Chisholm, and D. M. Karl. 2006. Community
541 genomics among stratified microbial assemblages in the ocean's interior. Science **311**: 496-
542 503.
543 Dore, J. E., and D. M. Karl. 1996. Nitrite distributions and dynamics at Station ALOHA. Deep-
544 Sea Res. Part II-Top. Stud. Oceanogr. **43**: 385-402.
545 DuRand, M. D., R. J. Olson, and S. W. Chisholm. 2001. Phytoplankton population dynamics at
546 the Bermuda Atlantic Time-series station in the Sargasso Sea. Deep-Sea Res. Part II **48**: 1983-
547 2003.
548 Feingersch, R., A. Philosof, T. Mejuch, F. Glaser, O. Alalouf, Y. Shoham, and O. Béjà. 2012.
549 Potential for phosphite and phosphonate utilization by *Prochlorococcus*. ISME J. **6**: 827-834.
550 Flombaum, P., J. L. Gallegos, R. A. Gordillo, J. Rincón, L. L. Zabala, N. Jiao, D. M. Karl, W. K.
551 Li, M. W. Lomas, D. Veneziano, C. S. Vera, J. A. Vrugt, and A. C. Martiny. 2013. Present and
552 future global distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus*.
553 Proc. Natl. Acad. Sci. USA **110**: 9824-9829.
554 Follows, M. J., S. Dutkiewicz, S. Grant, and S. W. Chisholm. 2007. Emergent biogeography of
555 microbial communities in a model ocean. Science **315**: 1843-1846.

556 García-Fernández, J. M., N. T. de Marsac, and J. Diez. 2004. Streamlined regulation and gene
557 loss as adaptive mechanisms in *Prochlorococcus* for optimized nitrogen utilization in
558 oligotrophic environments. Microbiol. Mol. Biol. Rev. **68**: 630-638.
559 Goericke, R., and D. J. Repeta. 1992. The pigments of *Prochlorococcus marinus*: The presence
560 of divinyl chlorophyll *a* and *b* in a marine prokaryote. Limnol. Oceanogr. **37**: 425-433.
561 Gruber, N. 2008. The marine nitrogen cycle: overview and challenges, p. 1-50. *In* D. G. Capone,
562 D. A. Bronk, M. R. Mulholland, and E. J. Carpenter [eds.], Nitrogen in the Marine
563 Environment. Academic Press.
564 Huang, S., S. W. Wilhelm, H. R. Harvey, K. Taylor, N. Jiao, and F. Chen. 2012. Novel lineages
565 of *Prochlorococcus* and *Synechococcus* in the global oceans. ISME J. **6**: 285-297.
566 Johnson, K. S., S. C. Riser, and D. M. Karl. 2010. Nitrate supply from deep to near-surface
567 waters of the North Pacific subtropical gyre. Nature **465**: 1062-1065.
568 Johnson, Z. I., E. R. Zinser, A. Coe, N. P. McNulty, E. M. S. Woodward, and S. W. Chisholm.
569 2006. Niche partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental
570 gradients. Science **311**: 1737-1740.
571 Kamennaya, N. A., and A. F. Post. 2013. Distribution and expression of the cyanate acquisition
572 potential among cyanobacterial populations in oligotrophic marine waters. Limnol. Oceanogr.
573 **58**: 1959-1971.
574 Kamennaya, N. A., M. Chernihovsky, and A. F. Post. 2008. The cyanate utilization capacity of
575 marine unicellular Cyanobacteria. Limnol. Oceanogr. **53**: 2485-2494.
576 Karl, D. M., and R. Lukas. 1996. The Hawaii Ocean Time-series (HOT) program: background,
577 rationale and field implementation. Deep-Sea Res. Part II-Top. Stud. Oceanogr. **43**: 129-156.
578 Kettler, G. C., A. C. Martiny, K. Huang, J. Zucker, M. L. Coleman, S. Rodrigue, F. Chen, A.
579 Lapidus, S. Ferriera, J. Johnson, C. Steglich, G. M. Church, P. Richardson, and S. W.

580 Chisholm. 2007. Patterns and implications of gene gain and loss in the evolution of
581 *Prochlorococcus.* PLoS Genet. **3**: e231.
582 Kuhn, R. M., D. Haussler, and W. J. Kent. 2013. The UCSC genome browser and associated
583 tools. Brief. Bioinform. **14**: 144-161.
584 Le Borgne, R., R. T. Barber, T. Delcroix, H. Y. Inoue, D. J. Mackey, and M. Rodier. 2002.
585 Pacific warm pool and divergence: temporal and zonal variations on the equator and their
586 effects on the biological pump. Deep-Sea Res. Part II-Top. Stud. Oceanogr. **49**: 2471-2512.
587 Letelier, R. M., D. M. Karl, M. R. Abbott, and R. R. Bidigare. 2004. Light driven seasonal
588 patterns of chlorophyll and nitrate in the lower euphotic zone of the North Pacific Subtropical
589 Gyre. Limnol. Oceanogr. **49**: 508-519.
590 Lomas, M. W., and F. Lipschultz. 2006. Forming the primary nitrite maximum: Nitrifiers or
591 phytoplankton? Limnol. Oceanogr. **51**: 2453-2467.
592 Lopatovskaya, K. V., A. V. Seliverstov, and V. A. Lyubetsky. 2011. NtcA and NtcB regulons in
593 cyanobacteria and rhodophyta chloroplasts. Mol. Biol. **45**: 522-526.
594 Mackey, K. R., L. Bristow, D. R. Parks, M. A. Altabet, A. F. Post, and A. Paytan. 2011. The
595 influence of light on nitrogen cycling and the primary nitrite maximum in a seasonally
596 stratified sea. Prog. Oceanogr. **91**: 545-560.
597 Malmstrom, R. R., A. Coe, G. C. Kettler, A. C. Martiny, J. Frias-Lopez, E. R. Zinser, and S. W.
598 Chisholm. 2010. Temporal dynamics of *Prochlorococcus* ecotypes in the Atlantic and Pacific
599 oceans. ISME J. **4**: 1252–1264.
600 Malmstrom, R. R., S. Rodrigue, K. H. Huang, L. Kelly, S. E. Kern, A. Thompson, S.
601 Roggensack, P. M. Berube, M. R. Henn, and S. W. Chisholm. 2013. Ecology of uncultured
602 *Prochlorococcus* clades revealed through single-cell genomics and biogeographic analysis.
603 ISME J. **7**: 184-198.

Martens-Habbena, W., P. M. Berube, H. Urakawa, J. R. de la Torre, and D. A. Stahl. 2009. Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and Bacteria. Nature **461**: 976-979.

Martiny, A. C., A. P. Tai, D. Veneziano, F. Primeau, and S. W. Chisholm. 2009a. Taxonomic resolution, ecotypes and the biogeography of *Prochlorococcus*. Environ. Microbiol. **11**: 823-832.

Martiny, A. C., S. Kathuria, and P. M. Berube. 2009b. Widespread metabolic potential for nitrite and nitrate assimilation among *Prochlorococcus* ecotypes. Proc. Natl. Acad. Sci. USA **106**: 10787-10792.

Moore, C. M., M. M. Mills, K. R. Arrigo, I. Berman-Frank, L. Bopp, P. W. Boyd, E. D. Galbraith, R. J. Geider, C. Guieu, and S. L. Jaccard. 2013. Processes and patterns of oceanic nutrient limitation. Nat. Geosci. **6**: 701-710.

Moore, L. R., A. Coe, E. R. Zinser, M. A. Saito, M. B. Sullivan, D. Lindell, K. Frois-Moniz, J. Waterbury, and S. W. Chisholm. 2007. Culturing the marine cyanobacterium *Prochlorococcus*. Limnol. Oceanogr. Meth. **5**: 353-362.

Moore, L. R., A. F. Post, G. Rocap, and S. W. Chisholm. 2002. Utilization of different nitrogen sources by the marine cyanobacteria *Prochlorococcus* and *Synechococcus*. Limnol. Oceanogr. **47**: 989-996.

Moore, L. R., and S. W. Chisholm. 1999. Photophysiology of the marine cyanobacterium *Prochlorococcus*: ecotypic differences among cultured isolates. Limnol. Oceanogr. **44**: 628-638.

Mulholland, M. R., and M. W. Lomas. 2008. Nitrogen uptake and assimilation, p. 303-384. *In* D. G. Capone, D. A. Bronk, M. R. Mulholland, and E. J. Carpenter [eds.], Nitrogen in the Marine Environment. Academic Press.

Newell, S. E., S. E. Fawcett, and B. B. Ward. 2013. Depth distribution of ammonia oxidation rates and ammonia-oxidizer community composition in the Sargasso Sea. Limnol. Oceanogr. **58**: 1491-1500.

Ohashi, Y., W. Shi, N. Takatani, M. Aichi, S. I. Maeda, S. Watanabe, H. Yoshikawa, and T. Omata. 2011. Regulation of nitrate assimilation in cyanobacteria. J. Exp. Bot. **62**: 1411-24

Olson, R. J., D. Vaulot, and S. W. Chisholm. 1985. Marine phytoplankton distributions measured using shipboard flow cytometry. Deep-Sea Res. **32**: 1273-1280.

Ottesen, E. A., R. Marin, C. M. Preston, C. R. Young, J. P. Ryan, C. A. Scholin, and E. F. DeLong. 2011. Metatranscriptomic analysis of autonomously collected and preserved marine bacterioplankton. ISME J. **5**: 1881-1895.

Paerl, R. W., K. A. Turk, R. A. Beinart, F. P. Chavez, and J. P. Zehr. 2012. Seasonal change in the abundance of *Synechococcus* and multiple distinct phylotypes in Monterey Bay determined by *rbcL* and *narB* quantitative PCR. Environ. Microbiol. **14**: 580-593.

Paerl, R. W., K. S. Johnson, R. M. Welsh, A. Z. Worden, F. P. Chavez, and J. P. Zehr. 2011. Differential distributions of *Synechococcus* subgroups across the California current system. Front. Microbiol. **2**: 59, doi: 10.3389/fmicb.2011.00059

Partensky, F., and L. Garczarek. 2010. *Prochlorococcus*: advantages and limits of minimalism. Annu. Rev. Mar. Sci. **2**: 305-331.

Rice, P., I. Longden, and A. Bleasby. 2000. EMBOSS: the European molecular biology open software suite. Trends Genet. **16**: 276-277.

Rocap, G., D. L. Distel, J. B. Waterbury, and S. W. Chisholm. 2002. Resolution of *Prochlorococcus* and *Synechococcus* ecotypes by using 16S-23S ribosomal DNA internal transcribed spacer sequences. Appl. Environ. Microbiol. **68**: 1180-1191.

651 Rodrigue, S., R. R. Malmstrom, A. M. Berlin, B. W. Birren, M. R. Henn, and S. W. Chisholm.
652 2009. Whole genome amplification and de novo assembly of single bacterial cells. PLoS ONE
653 **4**: e6864.
654 Rusch, D. B., A. C. Martiny, C. L. Dupont, A. L. Halpern, and J. C. Venter. 2010.
655 Characterization of *Prochlorococcus* clades from iron-depleted oceanic regions. Proc. Natl.
656 Acad. Sci. USA **107**: 16184-16189.
657 Rusch, D. B., A. L. Halpern, G. Sutton, K. B. Heidelberg, S. Williamson, S. Yooseph, D. Wu, J.
658 A. Eisen, J. M. Hoffman, K. Remington, K. Beeson, B. Tran, H. Smith, H. Baden-Tillson, C.
659 Stewart, J. Thorpe, J. Freeman, C. Andrews-Pfannkoch, J. E. Venter, K. Li, S. Kravitz, J. F.
660 Heidelberg, T. Utterback, Y. H. Rogers, L. I. Falcón, V. Souza, G. Bonilla-Rosso, L. E.
661 Eguiarte, D. M. Karl, S. Sathyendranath, T. Platt, E. Bermingham, V. Gallardo, G. Tamayo-
662 Castillo, M. R. Ferrari, R. L. Strausberg, K. Nealson, R. Friedman, M. Frazier, and J. C.
663 Venter. 2007. The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through
664 eastern tropical Pacific. PLoS Biol. **5**: e77.
665 Scanlan, D. J., M. Ostrowski, S. Mazard, A. Dufresne, L. Garczarek, W. R. Hess, A. F. Post, M.
666 Hagemann, I. Paulsen, and F. Partensky. 2009. Ecological genomics of marine
667 picocyanobacteria. Microbiol. Mol. Biol. Rev. **73**: 249-299.
668 Shimada, A., M. Nishijima, and T. Maruyama. 1995. Seasonal appearance of *Prochlorococcus* in
669 Suruga Bay, Japan. J. Oceanogr **51**: 289-300.
670 Steinberg, D. K., C. A. Carlson, N. R. Bates, R. J. Johnson, A. F. Michaels, and A. H. Knap.
671 2001. Overview of the US JGOFS Bermuda Atlantic Time-series Study (BATS): a decade-
672 scale look at ocean biology and biogeochemistry. Deep-Sea Res. Part II-Top. Stud. Oceanogr.
673 **48**: 1405-1447.

674 Tolonen, A. C., J. Aach, D. Lindell, Z. I. Johnson, T. Rector, R. Steen, G. M. Church, and S. W.
675 Chisholm. 2006. Global gene expression of *Prochlorococcus* ecotypes in response to changes
676 in nitrogen availability. Mol. Syst. Biol. **2**: 53.
677 Tyrrell, T. 1999. The relative influences of nitrogen and phosphorus on oceanic primary
678 production. Nature **400**: 525-531.
679 Van Mooy, B. A., and A. H. Devol. 2008. Assessing nutrient limitation of *Prochlorococcus* in
680 the North Pacific subtropical gyre by using an RNA capture method. Limnol. Oceanogr. **53**:
681 78-88.
682 Venter, J. C., K. Remington, J. F. Heidelberg, A. L. Halpern, D. Rusch, J. A. Eisen, D. Wu, I.
683 Paulsen, K. E. Nelson, W. Nelson, D. E. Fouts, S. Levy, A. H. Knap, M. W. Lomas, K.
684 Nealson, O. White, J. Peterson, J. Hoffman, R. Parsons, H. Baden-Tillson, C. Pfannkoch, Y.
685 H. Rogers, and H. O. Smith. 2004. Environmental genome shotgun sequencing of the Sargasso
686 Sea. Science **304**: 66-74.
687 West, N. J., and D. J. Scanlan. 1999. Niche-partitioning of *Prochlorococcus* populations in a
688 stratified water column in the eastern North Atlantic Ocean. Appl. Environ. Microbiol. **65**:
689 2585-2591.
690 Wu, J., W. Sunda, E. A. Boyle, and D. M. Karl. 2000. Phosphate depletion in the western North
691 Atlantic Ocean. Science **289**: 759.
692 Ye, J., G. Coulouris, I. Zaretskaya, I. Cutcutache, S. Rozen, and T. L. Madden. 2012. Primer-
693 BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC
694 Bioinformatics **13**: 134.
695 Yool, A., A. P. Martin, C. Fernández, and D. R. Clark. 2007. The significance of nitrification for
696 oceanic new production. Nature **447**: 999-1002.

Zinser, E. R., A. Coe, Z. I. Johnson, A. C. Martiny, N. J. Fuller, D. J. Scanlan, and S. W. Chisholm. 2006. *Prochlorococcus* ecotype abundances in the North Atlantic Ocean as revealed by an improved quantitative PCR method. Appl. Environ. Microbiol. **72**: 723-732.

Zinser, E. R., Z. I. Johnson, A. Coe, E. Karaca, D. Veneziano, and S. W. Chisholm. 2007. Influence of light and temperature on *Prochlorococcus* ecotype distributions in the Atlantic Ocean. Limnol. Oceanogr. **52**: 2205-2220.

ACKNOWLEDGEMENTS

We thank Michael Lomas and the Bermuda Atlantic Time-series Study team for sample collection at Bermuda; David Karl, Matthew Church, and the Hawai'i Ocean Time-series team for sample collection at Hawai'i; and the captain and crew of the *R/V Kilo Moana* (KM0915). We are also grateful to the DeLong Lab (MIT) for providing DNA samples from the HOT179 cruise, Maureen Coleman (University of Chicago) for providing *Prochlorococcus* core gene occurrence data for the HOT186 and BATS216 metagenome libraries, Julie Miller (MIT) for assistance with DNA sequencing, and Jason Bragg (Australian National University) for thoughtful discussion. We also thank two anonymous reviewers for their insightful and constructive criticism. This work was funded in part by the Gordon and Betty Moore Foundation through Grant GBMF495 to SWC and by the National Science Foundation (OCE-1153588 and DBI-0424599) to SWC. This article is a contribution from the NSF Center for Microbial Oceanography: Research and Education (C-MORE).

716 **Table 1**. Oligonucleotide primers used in this study.

Name	Sequence	Reference
ITS-F	5'-CCGAAGTCGTTACTYYAACCC-3'	(Rodrigue et al. 2009)
ITS-R	5'-TCATCGCCTCTGTGTGCC-3'	(Rodrigue et al. 2009)
NATL3f	5'-ACCTAGCTTCTTGTCATCTTTTTAT-3'	(Ahlgren et al. 2006)
NATL2r	5'-CATGAGATGCTTTATTCTTTCTAATC-3'	(Ahlgren et al. 2006)
narB34F	5'-TGCCCWTATTGYGGTGTWGGHTG-3'	This study
narB2099R	5'-ATBGGRCATGWYTKYTCRTGC-3'	This study
ProHLIInarB-qPCR-F	5'-AGGCAAGAGGGTACAGCAGCAG-3'	This study
ProHLIInarB-qPCR-R	5'-GCRTTTGGTTGGCCAGTYAAGG-3'	This study
ProLLInarB-qPCR-F	5'-TGMGACACCTAATGGTCGAGCCC-3'	This study
ProLLInarB-qPCR-R	5'-TGCCATTGTCCDAGGTAACGYC-3'	This study

717

718 **Table 2**. Validation of the *narB* qPCR assay by comparison with metagenomic sequence data.
719 Using qPCR, we determined the proportion of *Prochlorococcus* cells containing HLII *narB* or
720 LLI *narB* relative to the total concentration of *Prochlorococcus* measured by flow cytometry.
721 Using metagenomic data (Coleman and Chisholm 2010), the proportion of total *Prochlorococcus*
722 containing HLII *narB* or LLI *narB* was determined by comparing the occurrence of
723 *Prochlorococcus narB* genes with the average occurrence of single copy core *Prochlorococcus*
724 genes [sensu (Coleman and Chisholm 2010)]. When qPCR samples were not available for the
725 same depth as the metagenome library, qPCR measurements of *narB* were obtained for the
726 nearest depth. Note that qPCR and metagenome samples were not necessarily obtained on the
727 same hydrocast. Agreement between these two methods indicates that the qPCR assay is
728 detecting the majority of cells containing HLII *narB* or LLI *narB*.

	% *Prochlorococcus* containing *narB* estimated by qPCR or bioinformatics approaches			
	HL *narB*		**LL *narB***	
Depth (m)	**qPCR** *narB* qPCR count per total cell count	**Metagenomic** *narB* hits per average core gene hit	**qPCR** *narB* qPCR count per total cell count	**Metagenomic** *narB* hits per average core gene hit
Pacific Site (Hawai'i Ocean Time-series; HOT cruise 186; October 18-24, 2006)				
25	28.4%	28.9%	0.0%	0.5%
75	10.7%	15.8%	0.0%	0.0%
100	3.4%	n.a.	4.1%	n.a.
110	n.a.	3.5%	n.a.	5.7%
Atlantic Site (Bermuda Atlantic Ocean Time-series Study; BATS cruise 216; October 10-14, 2006)				
40	21.4%	n.a.	0.1%	n.a.
50	n.a.	9.5%	n.a.	0
60	0.0%	n.a.	2.1%	n.a.
100	0.7%	2.5%	7.6%	7.4%

729 n.a. samples not available for this depth.

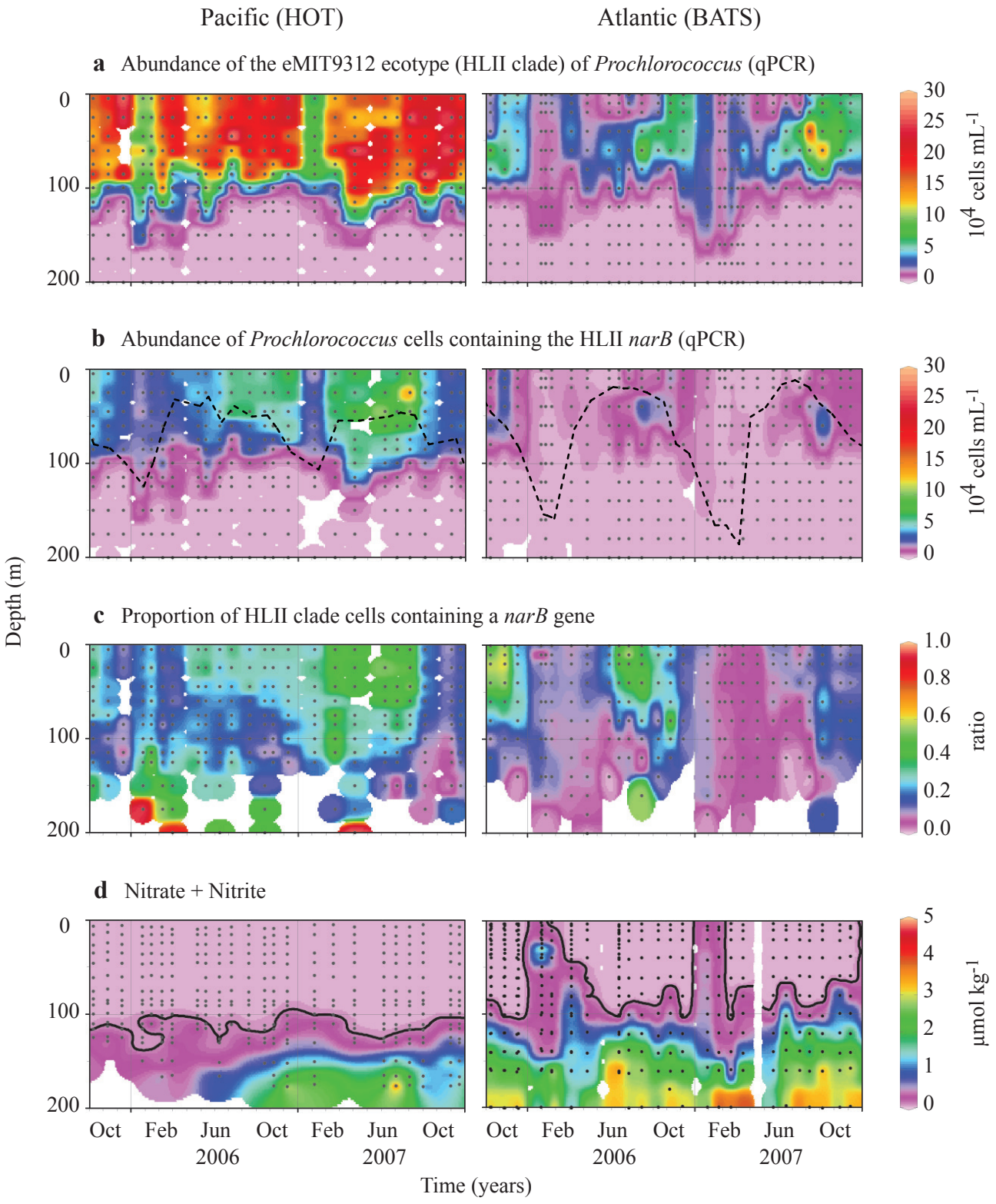

Fig. 1. Distribution and abundance of total HLII *Prochlorococcus* cells and HLII cells containing *narB* in the Pacific and Atlantic Oceans. (**a**) the abundance of HLII clade cells (eMIT9312 ecotype) determined by qPCR (Malmstrom et al. 2010), (**b**) the abundance of HLII clade cells containing *narB* determined by qPCR, (**c**) the fraction of HLII clade cells containing *narB* determined by normalizing the abundance of *narB*-containing HLII clade cells to total HLII clade cells, (**d**) nitrate + nitrite concentrations obtained from the HOT and BATS programs (Karl and Lukas 1996; Steinberg et al. 2001). Dashed lines represent the depth of the mixed layer (Malmstrom et al. 2010) and solid lines represent the nitracline depth defined as the depth at which nitrate + nitrite equals 0.1 µmol kg^{-1} (Le Borgne et al. 2002). Plots were created using Ocean Data View 4.6.2 and data were interpolated using weighted-average gridding.

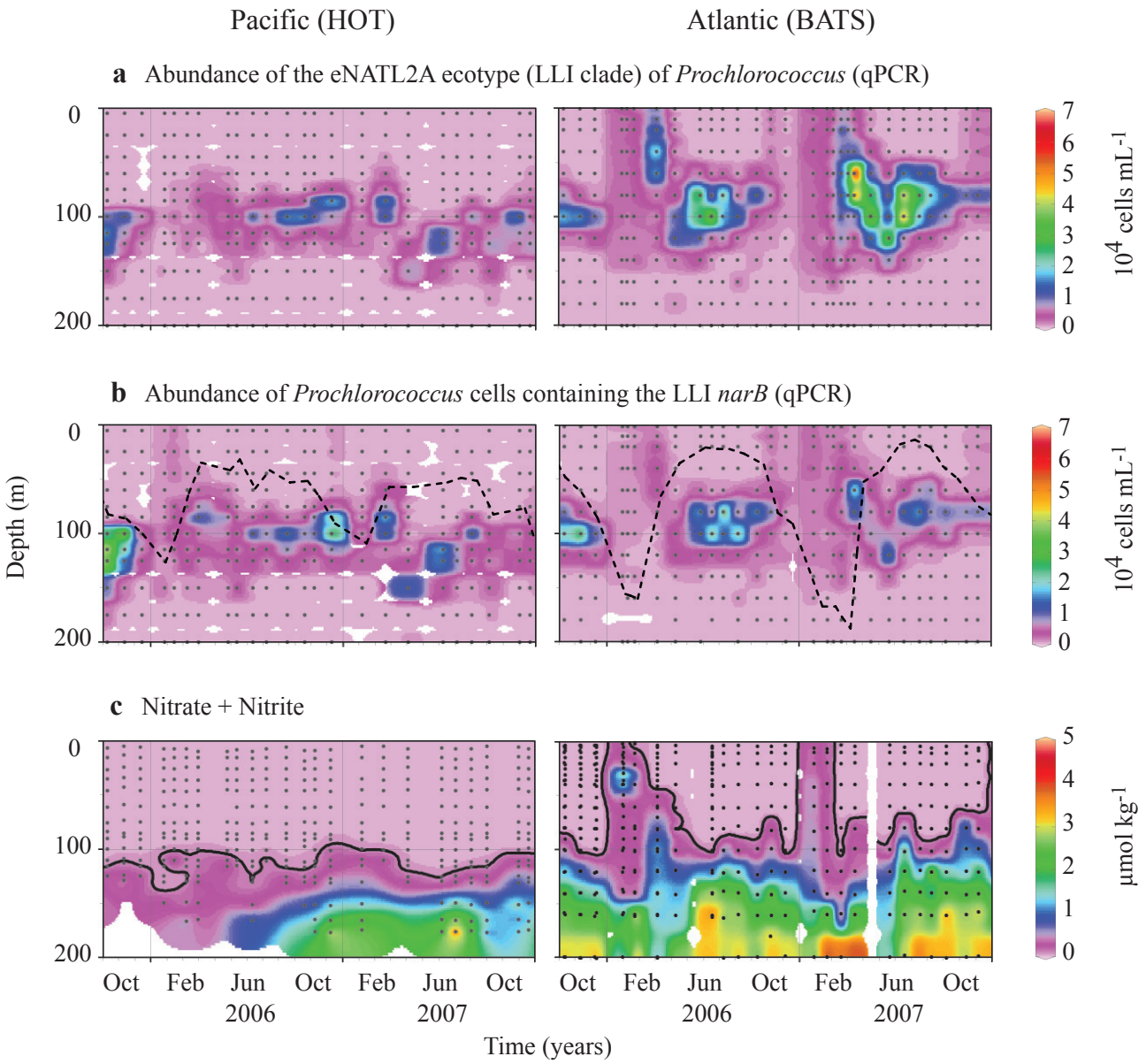

Fig. 2. Distribution and abundance of total LLI *Prochlorococcus* cells and LLI cells containing *narB* in the Pacific and Atlantic Oceans. (**a**) the abundance of LLI clade cells (eNATL2A ecotype) determined by qPCR (Malmstrom et al. 2010), (**b**) the abundance of LLI clade cells containing *narB* determined by qPCR, (**c**) nitrate + nitrite concentrations obtained from HOT and BATS. Dashed lines represent the depth of the mixed layer (Malmstrom et al. 2010) and solid lines represent the nitracline depth defined as the depth at which nitrate + nitrite equals 0.1 µmol kg^{-1} (Le Borgne et al. 2002). Plots were created using Ocean Data View 4.6.2 with data interpolated using weighted-average gridding.

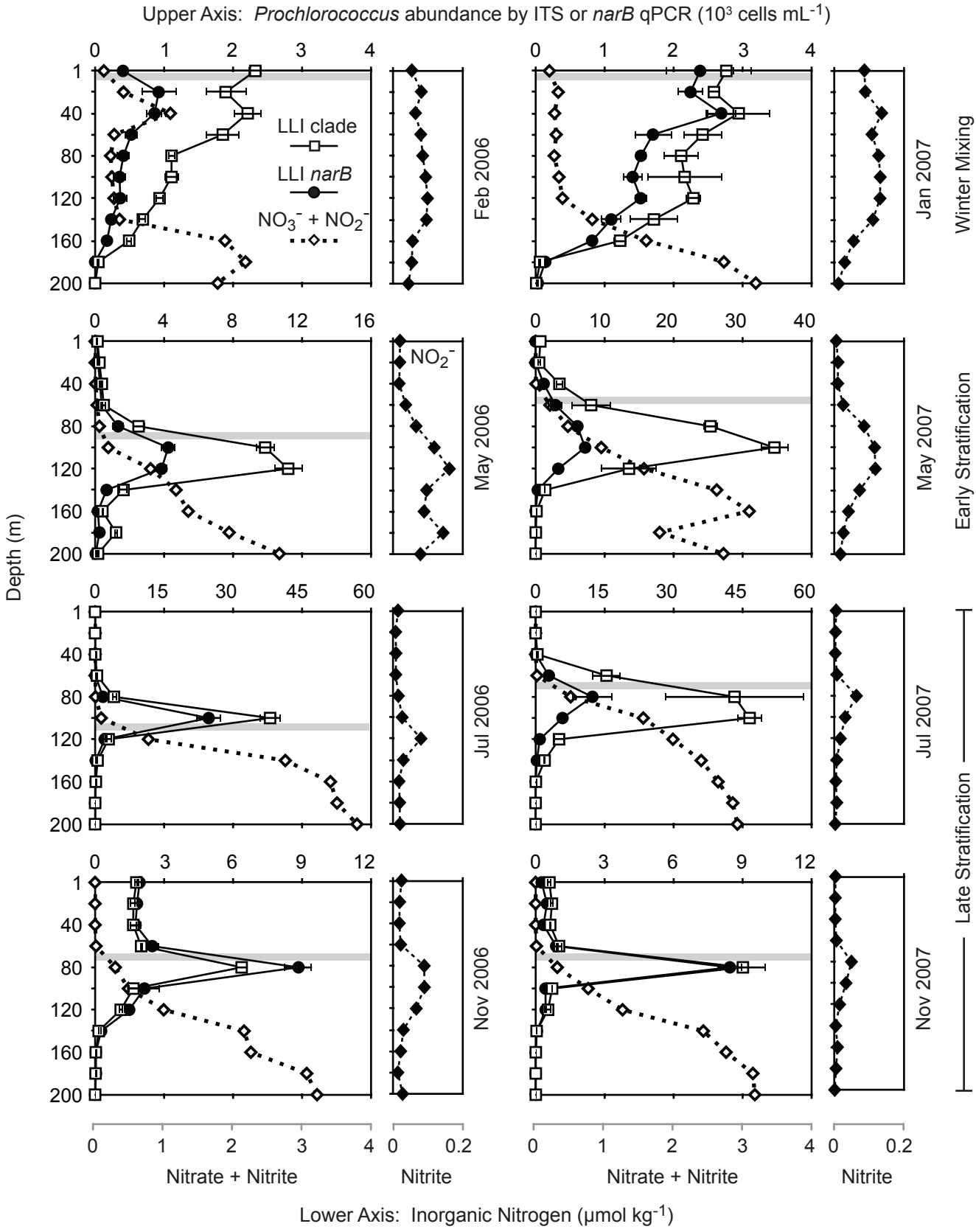

Fig. 3. Distribution of total LLI *Prochlorococcus* cells and LLI cells containing *narB* relative to the nitracline and primary nitrite maximum at the Atlantic site (BATS Station). The abundances of LLI clade cells (ITS qPCR) and *narB*-containing LLI clade cells (*narB* qPCR) are plotted for 8 representative depth profiles. Inorganic nitrogen (nitrate + nitrite) and nitrite concentrations were interpolated in Ocean Data View 4.6.2 in order to match depths of the qPCR samples. Gray bars represent the nitracline depth defined as the depth at which nitrate + nitrite equals 0.1 µmol kg^{-1} (Le Borgne et al. 2002). Within each water column, the greatest abundance of cells containing LLI *narB* are found immediately above or within the top layer of the nitracline and in close proximity to the primary nitrite maximum under stratified conditions.