npg

ORIGINAL ARTICLE

Co-occurring *Synechococcus* ecotypes occupy four major oceanic regimes defined by temperature, macronutrients and iron

Jill A Sohm[1,4], Nathan A Ahlgren[1,2,4], Zachary J Thomson[2], Cheryl Williams[2], James W Moffett[1], Mak A Saito[3], Eric A Webb[1] and Gabrielle Rocap[2]

[1]*Department of Biological Sciences, University of Southern California, Los Angeles, CA, USA;* [2]*School of Oceanography, University of Washington, Seattle, WA, USA and* [3]*Marine Chemistry and Geochemistry Department, Woods Hole Oceanographic Institution, Woods Hole, MA, USA*

Marine picocyanobacteria, comprised of the genera *Synechococcus* and *Prochlorococcus*, are the most abundant and widespread primary producers in the ocean. More than 20 genetically distinct clades of marine *Synechococcus* have been identified, but their physiology and biogeography are not as thoroughly characterized as those of *Prochlorococcus*. Using clade-specific qPCR primers, we measured the abundance of 10 *Synechococcus* clades at 92 locations in surface waters of the Atlantic and Pacific Oceans. We found that *Synechococcus* partition the ocean into four distinct regimes distinguished by temperature, macronutrients and iron availability. Clades I and IV were prevalent in colder, mesotrophic waters; clades II, III and X dominated in the warm, oligotrophic open ocean; clades CRD1 and CRD2 were restricted to sites with low iron availability; and clades XV and XVI were only found in transitional waters at the edges of the other biomes. Overall, clade II was the most ubiquitous clade investigated and was the dominant clade in the largest biome, the oligotrophic open ocean. Co-occurring clades that occupy the same regime belong to distinct evolutionary lineages within *Synechococcus*, indicating that multiple ecotypes have evolved independently to occupy similar niches and represent examples of parallel evolution. We speculate that parallel evolution of ecotypes may be a common feature of diverse marine microbial communities that contributes to functional redundancy and the potential for resiliency.
The ISME Journal (2016) **10**, 333–345; doi:10.1038/ismej.2015.115; published online 24 July 2015

Introduction

Microbes have vital roles in marine biogeochemical cycles, and describing their biogeography is critical to our understanding of where and how particular taxa contribute to these cycles (for example, the carbon cycle). As marine microbes carry out nearly half of primary production on earth, understanding the diversity and activity of phytoplankton is paramount to informing global models (Field *et al.,* 1998). The most abundant of these phytoplankton are the picocyanobacteria, which are broadly classified into the genera *Prochlorococcus* and *Synechococcus* based on their size and photosynthetic pigments: *Synechococcus* (0.6–2 µm) are slightly

larger than *Prochlorococcus* (0.6–0.8 µm) and *Synechococcus* possess light-harvesting phycobilisomes while *Prochlorococcus* does not and instead has divinyl chlororophyll *a* and *b* (Waterbury et al., 1986; Coleman and Chisholm, 2007; Scanlan *et al.,* 2009). Where the two coexist in the open ocean, *Prochlorococcus* are frequently more abundant than *Synechococcus*, often by 10-fold or more. However, *Synechococcus* can be temporally or regionally important contributors to carbon fixation, as they can be seasonally dominant (Flombaum *et al.,* 2013) and their larger cell size means they can fix an order of magnitude more carbon than *Prochlorococcus* cells (Jardillier *et al.,* 2010). *Synechococcus* is also present in a wider variety of environments—open-ocean and coastal waters from tropical to polar environments (Waterbury *et al.,* 1986; Zwirglmaier *et al.,* 2007; Huang *et al.,* 2011)—suggesting that the genus *Synechococcus* encompasses a wider range of physiological diversity.

Considerable genetic and genomic diversity within these two genera (Ferris and Palenik, 1998; Moore *et al.,* 1998; Fuller et al., 2003; Palenik *et al.,* 2006; Coleman and Chisholm, 2007; Kettler *et al.,* 2007;

Correspondence: EA Webb, Department of Biological Sciences, University of Southern California, 3616 Trousdale Pkwy, Los Angeles, CA 90089, USA.
or G Rocap, School of Oceanography, University of Washington, Box 357940, Seattle, WA 98195, USA.
E-mail: eawebb@usc.edu or rocap@ocean.washington.edu
[4]Co-authors.

Dufresne *et al.*, 2008; Scanlan *et al.*, 2009; Ahlgren and Rocap, 2012) suggests that each genus contains subpopulations with physiological differences that could allow them to occupy different environments. In fact, at least six phylogenetic clades in *Prochlorococcus* have been shown to represent ecotype subpopulations that differ in physiology and occupy distinct niches in the oceans (Moore *et al.*, 1998; Johnson *et al.*, 2006; Coleman and Chisholm, 2007; Martiny *et al.*, 2009). This niche differentiation apparently occurred in a successive or stepwise manner as basal *Prochlorococcus* clades are better adapted to low-light conditions, a more recently emerged lineage (LLI) can be found throughout the water column, and the most derived clades are high-light adapted. There is evidence of further divergence within the high-light lineage, resulting in ecotypes that occupy different temperature and nutrient regimes (Johnson *et al.*, 2006; Coleman and Chisholm, 2007; Martiny *et al.*, 2009; Malmstrom *et al.*, 2012). Such integration of information about the biogeographic and phylogenetic relatedness of ecotypes provides insight into evolutionary processes underlying niche differentiation in marine microbes.

In contrast, there is no clear model describing how the phylogeny of the marine *Synechococcus* group relates to the physiology or ecology of clades, in part because there are >20 clades (Ahlgren and Rocap, 2006; Huang *et al.*, 2011; Mazard *et al.*, 2012a) and only a few recognized adaptations described for cultured isolates. Examples include differences in light-harvesting phycobilisome pigments and structures, and the capacity to alter phycobilisomes to capture different wavelengths of light (Palenik, 2001; Ahlgren and Rocap, 2006; Six *et al.*, 2007), differing abilities to grow on nitrate (NO_3^-) (Moore *et al.*, 2002; Ahlgren and Rocap, 2006; Fuller *et al.*, 2006), or on different organic P sources (Moore *et al.*, 2005; Mazard *et al.*, 2012b). While identification of such traits is valuable, in many cases physiological information is available for only a subset of clades currently in culture and several clades do not yet have cultured representatives.

Variation in genome content among 10 cultured isolates is also suggestive of ecological differences between *Synechococcus* clades (Palenik et al., 2006; Dufresne *et al.*, 2008; Scanlan *et al.*, 2009). For example, specific gene differences explain *in vivo* variations in pigment composition and N utilization (Palenik *et al.*, 2006; Six *et al.*, 2007; Dufresne *et al.*, 2008; Scanlan *et al.*, 2009). In addition, *Synechococcus* strains have different complements of predicted Fe sensing and acquisition genes (Palenik *et al.*, 2006; Rivers *et al.*, 2009; Scanlan *et al.*, 2009). However, as is the case with physiological data from cultures, it is difficult to predict the realized niches and biogeography of *Synechococcus* clades solely from their genomic potential.

Molecular detection of *Synechococcus* clades *in situ* has provided valuable insight into their biogeography and factors controlling their distributions (Fuller *et al.*, 2003, 2006; Zwirglmaier *et al.*, 2007, 2008; Tai and Palenik, 2009). Clone libraries of cyanobacterial genes (for example, the 16 S–23 S internally transcribed spacer (ITS), *rpoC1*, *narB* and *petB*) have been used to identify the breadth of the diversity of natural communities (Toledo and Palenik, 1997; Rocap *et al.*, 2002; Paerl *et al.*, 2008; Choi and Noh, 2009; Huang *et al.*, 2011; Mazard *et al.*, 2012a), while assays designed to measure the abundance of specific clades have provided quantitative information about their distributions (Fuller *et al.*, 2006; Zwirglmaier *et al.*, 2007, 2008; Ahlgren and Rocap, 2012; Ahlgren *et al.*, 2014; Gutierrez-Rodriguez *et al.*, 2014). The most comprehensive of these latter studies, in terms of global biogeography, revealed an unprecedented view of clades that are preferentially more abundant in cold (clades I and IV) vs warmer (clade II) coastal and shelf habitats, concurrent with latitude. Additionally, this study identified clade III as an oligotroph restricted to open ocean waters (Zwirglmaier *et al.*, 2008).

Subsequent studies have expanded the known genetic and genomic diversity of *Synechococcus* (Huang *et al.*, 2011; Mazard *et al.*, 2012a), so there is an opportunity to update our view of clade biogeography, particularly for newly described clades not detected by previous assays. Additionally, while it is clear that temperature and macronutrient availability are strong drivers of *Synechococcus* and *Prochlorococcus* clade niche adaptation, these environmental parameters often do not fully explain the variance in clade abundances (for example, Johnson *et al.*, 2006; Zwirglmaier *et al.*, 2007, 2008), suggesting that other biotic or abiotic factors and interactions are also important in the niche partitioning of marine cyanobacteria. Specifically, trace metals like iron (Fe) and cobalt (Co) are hypothesized to influence niche adaptation of cyanobacteria ecotypes (Palenik *et al.*, 2006; Zwirglmaier *et al.*, 2008). For example, two new *Prochlorococcus* clades have been identified that are apparently adapted to high nutrient, low chlorophyll (HNLC) regions known to be Fe limited (Rusch *et al.*, 2010; West *et al.*, 2011), and Fe and cobalt are implicated in shaping *Synechococcus* clade composition in the tropical Pacific (Ahlgren *et al.*, 2014). However, because trace metals have not been measured alongside *Prochlorococcus* or *Synechococcus* clade abundances on a global scale, their importance in shaping niche adaptation of ecotypes has not been clarified.

To elucidate the ecology of *Synechococcus* populations, we have mapped the distribution of 10 *Synechococcus* clades in the global surface ocean using clade-specific qPCR assays targeting the 16S–23S rDNA ITS region (Ahlgren and Rocap, 2012). These data were quantitatively related to environmental parameters including temperature, macronutrients, and the trace metal Fe to discern how these variables affect the distribution of *Synechococcus* clades and to better define their

specific niches. Our study expands upon previous *Synechococcus* biogeography surveys (Fuller *et al.,* 2006; Zwirglmaier *et al.,* 2007, 2008; Tai and Palenik, 2009; Mella-Flores *et al.,* 2011; Mazard *et al.,* 2012a) by measuring four additional clades (XV, XVI, CRD1 and CRD2) that were not specifically enumerated previously (Ahlgren and Rocap, 2012) and concurrently measuring dissolved Fe for samples from four out of nine cruises. Finally, the distinct niches of clades inferred from biogeographic surveys were examined in light of their phylogenetic relatedness to investigate how adaptation to shared niches relates to the broader evolutionary history of marine *Synechococcus*.

Materials and methods

Samples were collected on nine research cruises. Details of cruise locations, dates, parameters measured and source of data (if previously published) are listed in Supplementary Table S1. Temperature and salinity data were extracted from conductivity, temperature, depth sensor data taken at the time of sampling. Macronutrients were measured with standard colorimetric methods (Strickland and Parsons, 1968). For four of the nine cruises dissolved Fe (Fe that passes through a 0.4-µm filter, hereafter simply referred to as Fe) was measured using trace metal clean protocols (Chappell *et al.,* 2012; Noble *et al.,* 2012) from the corresponding water columns (same cruises, stations and depths) for which DNA samples were taken (Supplementary Table S1, Supplementary Figure S1). In total, Fe was measured at 53% of the surface stations used in our analyses.

Abundances of 10 *Synechococcus* clades were measured using clade-specific qPCR assays applied to DNA extracted from 100 ml of filtered seawater as described in Ahlgren and Rocap (2012) except that for some samples DNA was extracted from 1 to 2 l of filtered seawater (volumes were recorded and accounted for in the calculation of clade cells ml^{-1}). All standards and samples were run in triplicate and melt curves showed no evidence of non-specific amplification. When triplicate reactions had poor replication ($\geqslant 3\%$ standard error) outliers were identified and removed using the Q test (Dean and Dixon, 1951). We note that > 20 *Synechococcus* clades have been identified using a variety of different loci (Ahlgren and Rocap, 2012); however, many of those not investigated with our qPCR assays are rarely observed in clone libraries, are restricted to estuarine or polar/subpolar waters, or are not represented by enough sequences to confidently design discriminative primers (Zwirglmaier *et al.,* 2008; Huang *et al.,* 2011).

Linear correlations (Pearson's *r*) were computed between log (x+1) clade abundances to investigate patterns of clade co-occurrence. Local similarity analysis (LSA) (Ruan *et al.,* 2006) was used to investigate linear correlations between environmental

factors and clade abundances using two data sets: (1) samples that had associated temperature, NO_3^- and PO_4^{3-} data or (2) samples that also had Fe data. Spearman rank correlations were computed to investigate non-linear relationships between clade abundances and environmental factors. For each clade, box and whisker plots were generated for temperature, NO_3^-, PO_4^{3-}, Fe and clade richness (the number of *Synechococcus* clades detected) for samples where that clade was detected. Significance levels for multiple pairwise linear and Spearman tests and pairwise *t*-tests of environmental parameters means were corrected with the Benjamini and Hochberg procedure ($P < 0.05$) (Benjamini and Hochberg, 1995).

Non-metric multidimensional scaling analyses were conducted with the package vegan in R (Oksanen *et al.,* 2012) on the data sets with and without Fe. Bray-Curtis dissimilarities were computed using Wisconsin double standardization of log (x+1) transformed clade abundances. Samples in the without Fe analysis clustered into four groups with dissimilarities of $\leqslant 0.5$, and this grouping of samples was significantly supported by analysis of similarities ($P < 0.001$, 999 permutations). Significant correlation of environmental parameters to sample ordination was tested with the envfit function in vegan. Significant relationships were also supported by Permutational Multivariate Analysis of Variance Using Distance Matrices as computed with the function adonis using 1000 permutations ($P < 0.001$ for each correlation to temperature, NO_3^- and PO_4^{-3} in the analysis without Fe data and $P < 0.01$ for the correlation to Fe in the analysis with Fe data).

Results and Discussion

Relationship between qPCR and cell counts
We analyzed 301 samples from 92 stations sampled on 9 research cruises in the Atlantic and Pacific Oceans from 2002 to 2010. To determine how well the abundance of these 10 clades (as measured by qPCR) represents the total abundance of *Synechococcus* present at each station, the sum of the qPCR clade abundances was compared with *Synechococcus* abundance from either microscope counts (Waterbury *et al.,* 1986) or flow cytometry (Marie *et al.,* 2005), when available ($n = 96$). The data cluster around a line with a slope of 0.84 ($R^2 = 0.54$), and 80% of samples are less than four-fold different between counts and summed qPCR indicating that, for the majority of samples for which we have cell count data, most of the *Synechococcus* present are accounted for with the 10 clades investigated (Figure 1a). Small overestimation observed in our data and other qPCR assays of natural samples may be due in part to natural variation in genome copies per cell (Zinser *et al.,* 2007; Tai and Palenik, 2009; Ahlgren and Rocap, 2012). Notably, in five samples total cell counts were more than 10-fold higher than the sum of the qPCR clade abundances, suggesting
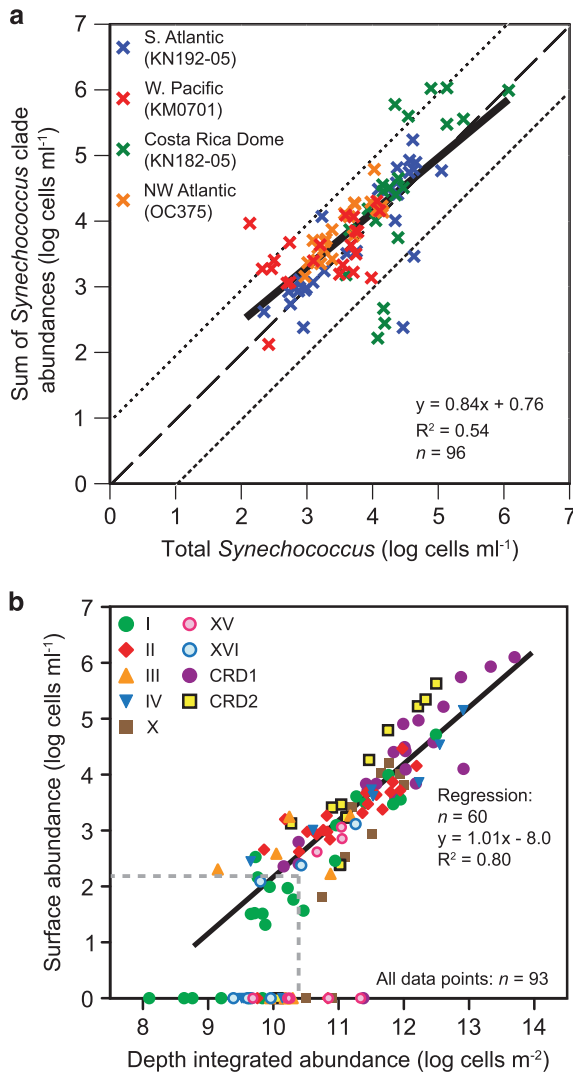
Figure 1 Validity of qPCR methodology and use of surface values in analysis. (**a**) Comparison of total *Synechococcus* abundance (determined with microscope counts or flow cytometry (FCM)) to the sum of all 10 *Synechococcus* clade abundances measured by qPCR. Points are color coded by region and research cruise (Supplementary Table S1). Because the limits of detection for measuring total *Synechococcus* by FCM or microscopy or for measuring individual clades qPCR assays are all about 100 cells ml^{-1}, samples where either FCM/microscopy counts ($n=6$) or summed qPCR results ($n=4$) were below 100 cells ml^{-1} were excluded. The dashed line depicts the 1:1 line, dotted lines depict 10:1 and 1:10 lines, and the solid line depicts the linear regression of the data. Most samples (80%) were within 4-fold of the 1:1 line. (**b**) Water column integrated abundance vs surface abundance of *Synechococcus* clades for all cruises where profiles were available (TN-224, KM0405, KM0701, KN182-05, OC375 and KN192-05, see Supplementary Table S1). Trapezoidal integration was used to generate water column values down to 150 m for 34 depth profiles. The horizontal dashed line indicates the median detection limit for the qPCR assays (150 cells ml^{-1}), and the vertical dashed line depicts a corresponding lower limit for depth-integration values assuming cells are at the 150 cells ml^{-1} detection limit in a water column of 150 m (2.3×10^{10} cells m^{-2}). The regression only includes data points where surface abundances were above the detection limit.

the presence of additional *Synechococcus* clades not detected by our qPCR primers. These locations may harbor novel clades and be productive targets for future isolation efforts.

Horizontal vs vertical partitioning of Synechococcus *clades*
Previous surveys indicate that *Synechococcus* partition the oceans more strongly along the horizontal scale than vertically with depth (Fuller *et al.,* 2003, 2006; Zwirglmaier *et al.,* 2007, 2008; Choi and Noh, 2009; Huang *et al.,* 2011; Mazard *et al.,* 2012a), and we found similar trends in our study. Globally, surface clade abundance was strongly correlated to depth integrated clade abundances (Figure 1b); hence, the *Synechococcus* community structure at the surface appears to generally be representative of the community at lower depths. Notable exceptions include six instances where a clade was not detected in the surface but had depth integrated abundances significantly above the limit of detection (Figure 1b). These cases were primarily attributed to clade XV, which has been found at higher abundances below surface waters (Ahlgren *et al.,* 2014). Most other anomalies from the otherwise strong relationship between surface and depth integrated abundance represent a clade detected in a subsurface sample just above the limit of detection but not detected in surface waters (for example, see Supplementary Figure S2: clade I in station 3 and clade III in station 13). These occurrences result in depth integrated abundances below the theoretical limit of detection and thus represent artifacts of detecting populations near the limit of detection rather than an indication that surface populations were not representative of those at depth.

Specifically, in the North Pacific, clade composition was fairly uniform with depth but changed dramatically with latitude (Figure 2), from domination by clades I and IV above 37 °N, to clade II in subtropical waters, and then to clade CRD1 south of ~20 °N. A similar shift from clade II to CRD1 throughout the water column was seen from the western to eastern basins in the South Atlantic (Supplementary Figure S2, station 3 vs station 13). In contrast, vertical partitioning of clades was not observed in these transects, as clade abundance almost always decreased with depth along with total *Synechococcus* abundance (Figure 2; Supplementary Figure S2). In summary, our data indicate surface clade abundances are generally representative of deeper populations. We therefore focused on studying the global biogeography of surface populations for which we had a broader distribution of samples.

The biogeography, co-occurrence and niche adaptation of Synechococcus *clades*
Synechococcus was found widely distributed in surface waters of the global ocean, detected with at least one set of qPCR primers at 93% of stations (Figure 3). At the majority of stations (55%) a single
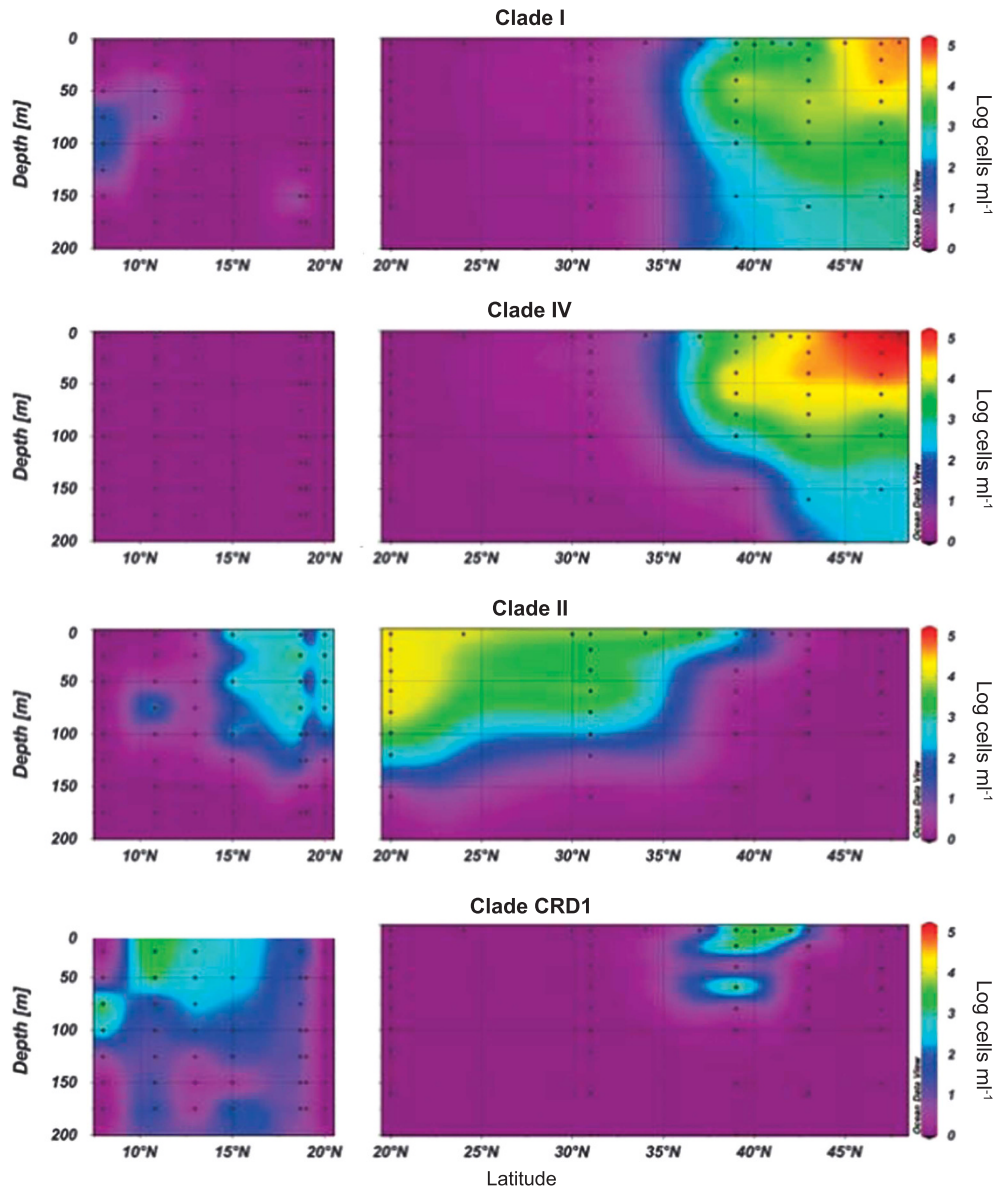
Figure 2 Sections from two cruises in the North Pacific (left, R/V Kilo Moana, February 2004, 202 °E; right, R/V Thompson, September 2008, 208 °E) showing abundance of dominant clades in the region with depth (clades III and X, not shown, were often present with clade II, but in lower abundance). The transition from clades I and IV in the subarctic to clade II in subtropical waters to clade CRD1 in the equatorial upwelling region is clearly seen both at the surface and at lower depths. Clade CRD1 is also present at the surface where the population transitions from clades I and IV to clade II at ∼ 40 ºN.

clade made up > 80% of the total *Synechococcus* population (Figure 3, Supplementary Figure S3). Nevertheless, multiple *Synechococcus* clades often co-occurred in surface waters, with at least three clades detected at 43% of stations. Four distinct groups of frequently co-occurring clades emerged from global distribution patterns and were supported by significant linear correlations between clade abundances (Figure 4a; see Supplementary Table S2 for numerical values): clades I and IV; II, III and X; CRD1 and CRD2; and XV and XVI. Members of Clade VIII are generally found in euryhaline or estuarine waters (Dufresne *et al.,* 2008; Huang *et al.,* 2011) and unsurprisingly were not detected in our oceanic samples.

Ordination of samples based on clade composition resulted in four significantly distinct clusters, each dominated by a different group of clades (Figure 5). These sites were significantly correlated to gradients in temperature, nutrients and Fe (Figure 5). Thus, each group of clades occupied a distinct marine biome or habitat: (1) clades I and IV dominated in cold, nutrient-rich waters ('cold, high nutrient'); (2) clades II, III and X were dominant in warm, oligotrophic, open-ocean habitats in the tropical and subtropical oceans ('warm, low nutrient'); (3) clades CRD1 and CRD2 ('low Fe') were most successful in low Fe waters that overlap with previously defined HNLC provinces (Martin *et al.,* 1991; Fung *et al.,* 2000) including equatorial upwelling regions and
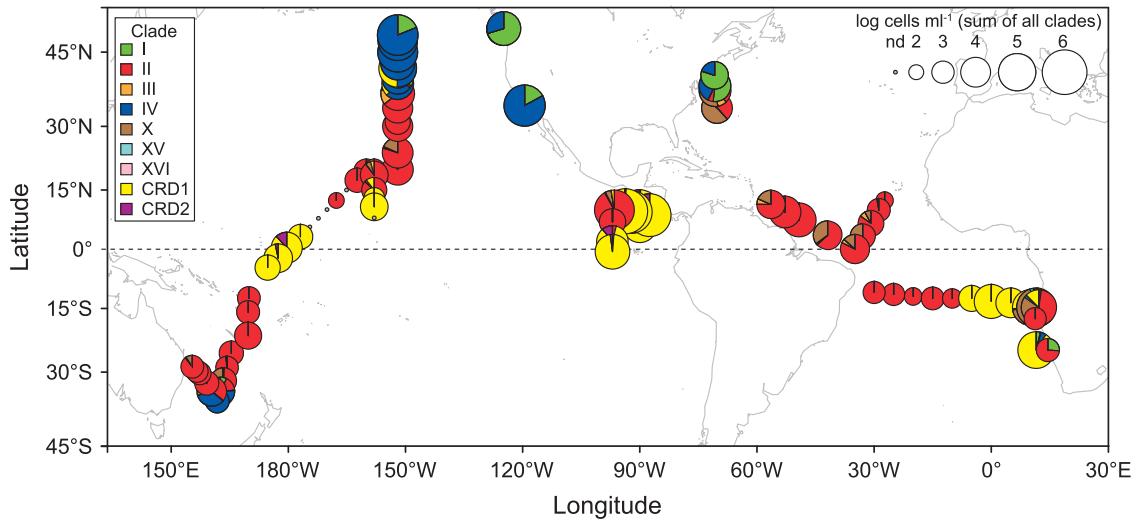
Figure 3 Global map of *Synechococcus* clade distributions in surface waters. Pie charts depict the relative abundance of each clade, and the radius of each pie chart is scaled to the sum of clade abundances at that location (scale at top right, log cells ml^{-1}). Small gray circles indicate stations where none of the 10 clades were above the limits of detection of the assays, generally ~ 100 cells ml^{-1} ('nd'). Maps of individual clade abundances are shown in Supplementary Figure S3.
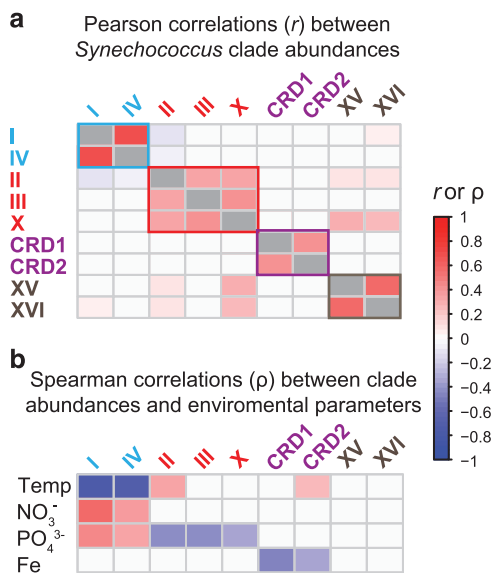

Figure 4 (**a**) Linear Pearson's correlation coefficients (*r*) between log transformed *Synechococcus* clade abundances and (**b**) Spearman correlation coefficients (*ρ*) between *Synechococcus* clade abundances and environmental parameters. Only significant correlations (*P* < 0.05) are shown. Clade names are colored according to the four major groups of co-occurring clades that are evident from the linear correlations (these groups are also highlighted by corresponding colored boxes in the matrix).

North Pacific sites; and (4) clades XV and XVI occurred at low abundances and were only detected at ecotone sites with intermediate conditions that often occurred near junctions of biomes ('transitional').

Clades I and IV were highly correlated to each other (Figure 4a) and were mainly restricted to high latitude open ocean sites and temperate coastal zones, as has been previously well described

(Zwirglmaier *et al.,* 2008; Tai and Palenik, 2009). These sites are typified by significantly higher nutrient levels and colder temperatures (Figures 4 and 6). Linear LSA (Supplementary Figure S4) and non-linear Spearman rank correlations (Figure 4b) confirm positive and negative relationships of clades I and IV to macronutrients and temperature, respectively. The latter agrees with recent characterization of clade I strains that exhibit greater tolerance to cold temperatures than those of clade II (Pittera *et al.,* 2014). Clades I and IV had higher mean abundances than clades II, III and X (although only significantly higher for clade III, *P* < 0.05) and higher maximum abundances than clades II and III (Supplementary Table S3) as they were typically found in waters that support higher total *Synechococcus* abundance.

The relative abundances of clades I and IV have been shown to vary between ocean regions, but controls on this ratio are unclear (Zwirglmaier *et al.,* 2008). Temporal studies at a coastal site suggest that clade I can take advantage of rapid increases in nutrients, while clade IV is more ubiquitous year-round (Tai and Palenik, 2009). Genomic, physiological and transcriptomic studies suggest that perhaps these clades are better adapted to fluctuating metal conditions in coastal habitats (Palenik *et al.,* 2006; Stuart *et al.,* 2009). Our analysis appears to show a lower Fe range for clade I (Figure 6c); however, it is important to note that Fe was not measured in our study at sites frequently dominated by clades I and IV (Supplementary Figure S1) so we cannot be conclusive about their Fe niches. Although both clades I and IV exhibit a low Fe range in our data set (Figure 6c) and are dominant in the subarctic, HNLC North Pacific (Martin *et al.,* 1991; Fung *et al.,* 2000), we suspect that the actual Fe range for both clades I and IV is broader than depicted here given that they
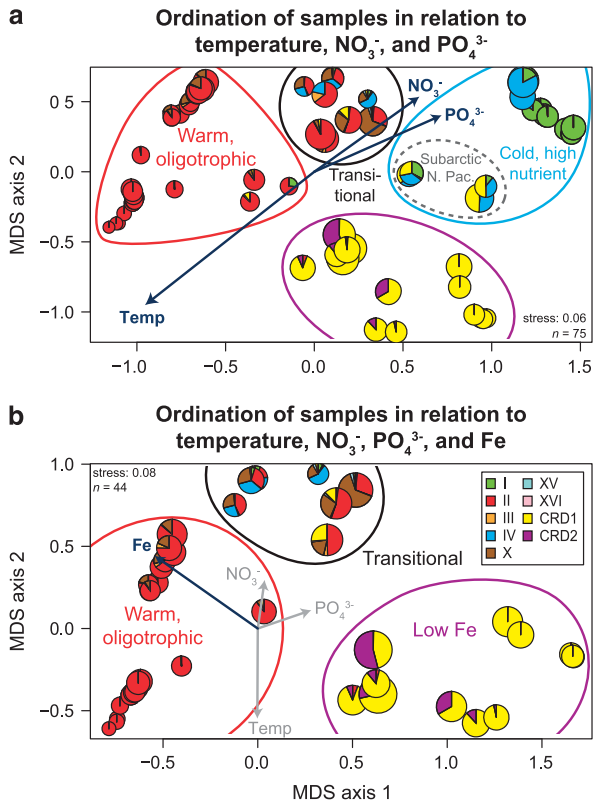
Figure 5 Non-metric multidimensional scaling (nMDS) ordination of samples according to Bray-Curtis dissimilarity calculated from clade abundances. Pie charts depict relative clade composition and total summed abundance as in Figure 3 for each sample. (**a**) Analysis with temperature and nutrients ($n = 75$ samples). (**b**) Analysis of samples with temperature, nutrient and Fe data ($n = 44$). Note that samples dominated by clades I and IV are absent from (**b**) because Fe was not measured at these sites. Solid lines indicate clusters of samples where samples within each cluster have Bray-Curtis dissimilarities of $\leqslant 0.5$, and this clustering of samples was supported by analysis of similarities ($P < 0.01$). Clusters are classified into four regime types defined by their biogeography and environmental parameter ranges. Parameters significantly correlated to sample ordination are depicted in blue ($P < 0.05$). A dashed gray line outlines low Fe, HNLC biome samples from the N. Pacific.

also frequently dominate in coastal sites where trace metal concentrations can vary greatly (Bruland *et al.,* 2001; Johnson *et al.,* 2001).

Clades II, III and X were most commonly found at low latitude, open ocean sites (Figure 3) typified by significantly warmer temperatures and lower macronutrient levels than where clades I and IV dominate (Figures 4 and 6). Clade II in particular was the most ubiquitous clade, but not the most abundant (Supplementary Table S3), likely because it most often occurs in the open ocean where total *Synechococcus* is less abundant. This may explain why LSA networks showed positive relationships between clade II and PO_4^{3-} and NO_3^- (Supplementary Figure S4) while clade II was negatively correlated to PO_4^{3-} by Spearman correlation, a non-linear statistic (Figure 4). Spearman correlations suggested that clade II most often occurred in oligotrophic waters while LSA networks highlight that when clade

II is detected its abundance increases with higher nutrient availability (confirmed by linear regression of only samples where clade II was found, $P < 0.05$).

Our data show that clade II is the dominant open-ocean ecotype (Figure 3). This agrees with smaller-scale quantitative surveys and clone libraries and analysis of global metagenomic data sets (Ferris and Palenik, 1998; Fuller *et al.,* 2003, 2006; Zwirglmaier *et al.,* 2007; Huang *et al.,* 2011; Mazard *et al.,* 2012a) but differs from a previous global survey where clade II was not detected in many open-ocean stations and rather was suggested to be a tropical/subtropical coastal/shelf ecotype (Zwirglmaier *et al.,* 2007, 2008). This latter survey employed 16S rDNA dot-blot probing methods for detection and it is possible that the discrepancy is due to that particular probe not capturing the full diversity of natural populations of clade II. In addition, clade X frequently co-occurs with clades II and III in warm, oligotrophic waters (Figures 3 and 4a); it was detected at 78% of sites where both clades II and III occur or 44% of sites where either clade II or III is found. Clades III and X were generally less abundant and prevalent than clade II, and clade III was the least abundant of the three clades in this group (Figure 3, Supplementary Table S3). While the representative clade III strain, WH8102, has historically been considered as the type strain for the oligotrophic ocean (Palenik *et al.,* 2006; Zwirglmaier *et al.,* 2008), we show that clade II is in fact the dominant clade in waters of the largest marine biome and suggest that future studies focus on type strain(s) from clade II.

Clades CRD1 and CRD2, which are newly enumerated with our assays, were mostly restricted to tropical/subtropical upwelling waters (that is, the Costa Rica Dome, equatorial upwelling and the Benguela upwelling; Figure 3, Supplementary Figure S3). At these upwelling sites, they reached high abundances, with clade CRD1 showing the highest mean and maximum (1.1×10^6 cells ml^{-1}) abundances of all clades (Supplementary Table S3). Accordingly, clades CRD1 and CRD2 had somewhat higher nutrient means than those of the other warm water clades II, III and X (Figure 6; Supplementary Figure S5), and clade CRD1 had a positive relationship to PO_4^{3-} in LSA networks (Supplementary Figure S4). It is noteworthy that clade CRD2 but not CRD1 was positively correlated to temperature (Figure 4b). This probably reflects the fact that clade CRD2 was found at low latitude sites only, while clade CRD1 was found at low latitude stations as well as few cold water sites in the subarctic North Pacific.

Among the predominantly warm water clades (II, III, X, CRD1 and CRD2) that were all detected in samples where Fe was measured concurrently (Supplementary Figure S1), clades CRD1 and CRD2 were found at sites with noticeably lower Fe levels (Figure 6c) and fittingly were negatively correlated with Fe (Figure 4b). In particular, clade CRD2 exhibited a statistically lower Fe range than
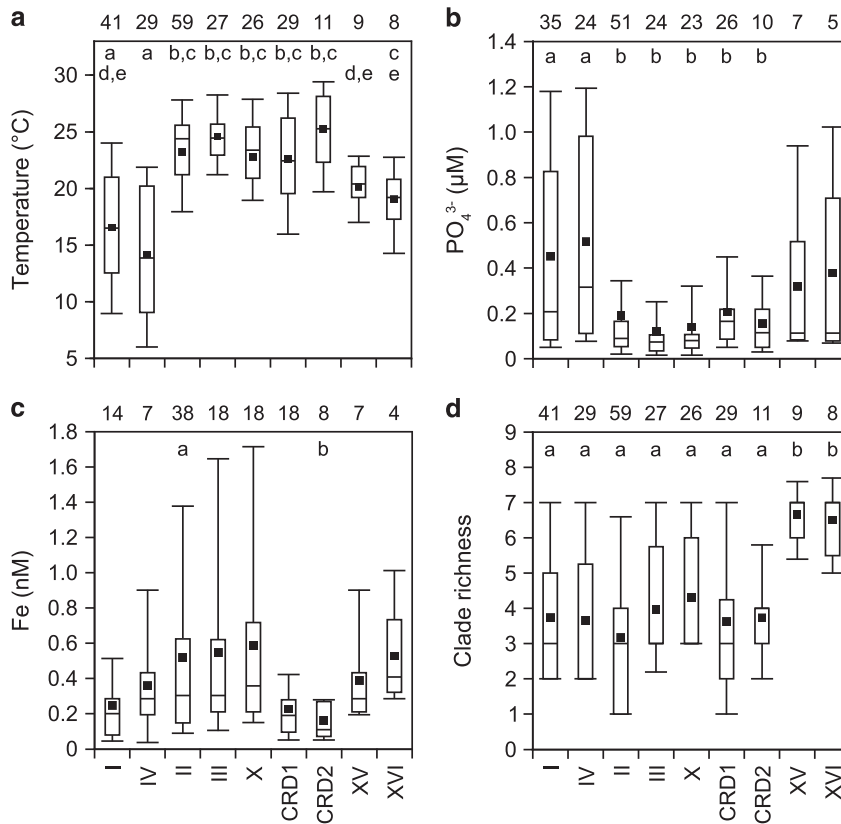
Figure 6 Box and whisker diagrams of the environmental conditions and clade richness (number of clades detected) of surface samples where each clade was detected. (**a**) Temperature, (**b**) PO_4^{3-} concentrations, (**c**) Fe concentrations and (**d**) clade richness. The box shows the 25–75% range, the whiskers show the 10–90% range, the bar shows the median and the square symbol shows the mean. Numbers at the top indicate the number of sites for which that clade was detected and the environmental parameter was measured. Letters above bars indicate the level of significance in pairwise comparisons of means. Clades that do not share letters have means that are significantly different from each other; clades with the same letter have means that are not significantly different from each other, and clades without letters have means that are not significantly different from any other clade (pairwise *t*-tests, $P<0.05$). Box plots for NO_3^- were similar to those of PO_4^{3-} (Supplementary Figure S5).

clade II, another predominantly warm-water clade (Figure 6c). Low Fe was also significantly correlated to the distinct clustering of samples dominated by clades CRD1 and CRD2 (Figure 5b). Thus, these clades appear to be specifically adapted to low Fe habitats in comparison with clades II, III and X, which otherwise have comparable temperature and nutrient ranges (Figure 6).

Consistent with a low Fe niche for clades CRD1 and CRD2 is the fact that these clades dominate at sites in the Equatorial Pacific previously identified as HNLC oceanic regions where low Fe availability limits phytoplankton growth (Martin *et al.,* 1991; Coale *et al.,* 1996; Fung *et al.,* 2000). Specifically clades CRD1 and CRD2 dominate at equatorial sites previously described as HNLC (Martin *et al.,* 1991; Fung *et al.,* 2000) in the Western Pacific (near 180º W) while clade II dominates in adjacent gyre waters (Figure 3) and there is a clear transition in dominance from clade II to CRD1 on the transect going directly south of the Hawaiian Islands (Figures 2 and 3) in parallel with a decrease in Fe away from the islands (measured on a different cruise along this same transect; Supplementary

Figure S6). Although Fe was not measured for North Pacific samples, clade CRD1 was also abundant (up to 6×10^4 cells ml^{-1}) at ~ 40 °N (Figures 2 and 3, Supplementary Figure S3), in the subarctic North Pacific HNLC region (Fung *et al.,* 2000). Interestingly, many of the sites where clade CRD2 was found are also regions overlying oxygen minimum zones with elevated Co (Noble *et al.,* 2012; Ahlgren *et al.,* 2014), suggesting that high Co and low Fe availability may influence this clade's habitat range in tandem (Ahlgren *et al.,* 2014). However, more data are needed to determine this.

Fe availability is known to limit total phytoplankton abundance in HNLC regions (Boyd *et al.,* 2007) and some larger phytoplankton species are adapted to low Fe conditions (Marchetti *et al.,* 2006). Two *Prochlorococcus* clades, HNLC1 and HNLC2, have been observed in low Fe HNLC habitats (Rusch *et al.,* 2010; West *et al.,* 2011); these low Fe *Synechococcus* clades have therefore emerged in parallel to similarly adapted HNLC *Prochlorococcus* clades. More broadly, our results complement genomic and physiological work that indicate trace metals are important in shaping the evolution

and ecology of phytoplankton ecotypes (Mann *et al.*, 2002; Palenik *et al.*, 2006; Rivers *et al.*, 2009; Scanlan *et al.*, 2009; Stuart *et al.*, 2009; Ahlgren *et al.*, 2014). Different clades are known to contain different complements of Fe stress-related genes, suggesting that the presence or absence of these genes could give cells an advantage under Fe stress (Palenik *et al.*, 2006; Rivers *et al.*, 2009; Scanlan *et al.*, 2009). It will be interesting to see what Fe-related gene clades CRD1 and CRD2 possess and if they exhibit similar genomic signatures of low Fe adaptation to those in *Prochlorococcus* HNLC clades, namely fewer genes encoding Fe-binding proteins. Unfortunately, there are no published genomes of clade CRD1 strains, and there are no cultured isolates of clade CRD2.

The last group of clades, XV and XVI, were neither ubiquitous nor particularly abundant (Figure 3, Supplementary Figure S3, Supplementary Table S3). They were only detected at samples belonging to the 'transitional' non-metric multidimensional scaling cluster (Figure 5) that exhibit intermediate temperatures, nutrients and Fe levels (Figure 6). These sites also had significantly higher clade richness (Figure 6d) and often occurred near the junctions of the other three biomes (for example, in waters between Australia and New Zealand, Supplementary Figure S3). This evidence suggests that clades XV and XVI are adapted to life in marine ecotones—regions at the edges or transitions of other biomes that are typified by intermediate environmental conditions and higher diversity. Clades XV and XVI have also been shown to occupy subsurface waters elsewhere (Ahlgren *et al.*, 2014; Gutierrez-Rodriguez *et al.*, 2014), so perhaps higher mixing in ecotone sites brings them to the surface and explains why they were rarely detected in surface waters. In fact, in the Sargasso Sea, clades XV and XVI were present through the water column at the onset of water column stratification following winter deep mixing (Ahlgren and Rocap, 2006).

Overall, our findings demonstrate coherent biogeographic patterns in the distributions of the major *Synechococcus* clades. Notably, several clades share similar environmental niches that are broadly defined by temperature, macronutrient and dissolved Fe concentrations. We confirm the dominance of clades I and IV in cold, mesotrophic waters (Zwirglmaier *et al.*, 2008; Tai and Palenik, 2009; Ahlgren and Rocap, 2012); but more importantly we reveal the niches of remaining clades that were previously undetermined or poorly defined (II, X, XV, XVI, CRD1 and CRD2). In particular, this study makes significant advances to *Synechococcus* clade ecology by establishing that clade II is the dominant ecotype in the world's oceans; clade X often co-occurs with clade II in warm, oligotrophic habitats; clades CRD1 and CRD2 dominate in low Fe habitats; and less abundant clades XV and XVI are often found at the transitions of major biomes.
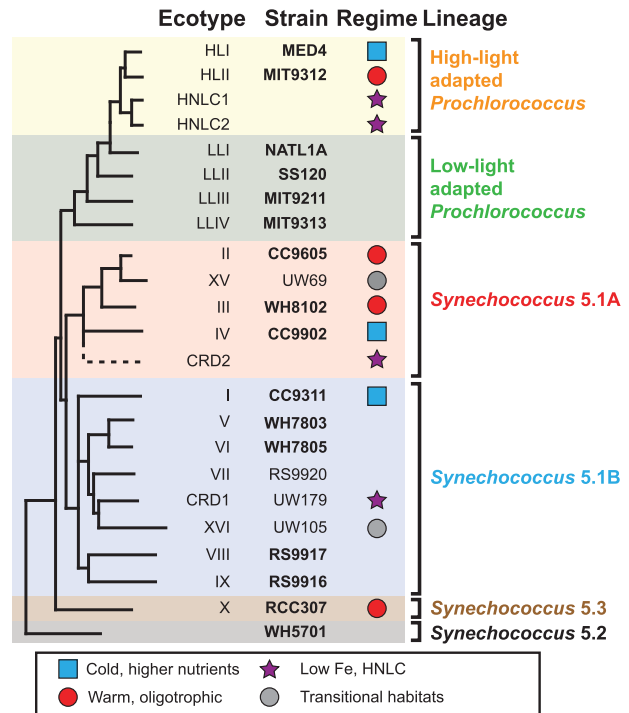

Figure 7 The phylogeny of *Prochlorococcus* and *Synechococcus* ecotypes, representative strains, and the oceanic regimes they occupy. The regimes that ecotypes occupy are paraphyletic, consistent with a pattern of parallel evolution whereby clades have independently adapted to similar habitats. The schematic tree depicted is supported by congruent concatenated core gene phylogenies (Ahlgren and Rocap, 2012; Mazard *et al.*, 2012a) and phylogenomic analysis of >1100 core genes (Dufresne *et al.*, 2008). Clade CRD2 groups closely with 5.1 A strains based on ITS phylogeny and is provisionally assigned to this subcluster (dashed line). Representative strains that have had their genomes sequenced are shown in bold.

Evolutionary relationships of co-occurring clades

Having elucidated the niches of several major *Synechococcus* clades, we examined the evolutionary relatedness of co-occurring clades and found that many of the clades that share similar niches are not closely related within the marine *Synechococcus* group (Figure 7). Although the form genus *Synechococcus* is dispersed polyphyletically across the cyanobacterial tree, most marine *Synechococcus* clades belong to the larger phylogenetic cluster 5.1 that comprises two distinct subgroups of clades, labeled 5.1 A (II, III, IV and XV) and 5.1B (I, V, VI, VII, VIII, IX, XVI and CRD1). Clade X belongs to a third related lineage, cluster 5.3 (Dufresne *et al.*, 2008; Scanlan *et al.*, 2009; Huang *et al.*, 2011; Mazard *et al.*, 2012a). Thus, clades I and IV, which co-occur in colder, nutrient-rich waters, belong to different lineages within cluster 5.1. Similarly, clades XV and XVI, which co-occur in transitional waters, also belong to different subclusters. Likewise, although both clades II and III are found in subcluster 5.1 A, they share a similar environmental niche with clade X, which belongs to a separate cluster (5.3). Clades CRD1 and CRD2 that are prevalent in low Fe waters are likely a fourth

example. Clade CRD1 is a member of subcluster 5.1B (Ahlgren and Rocap, 2012; Mazard *et al.,* 2012a), while ITS phylogeny suggests clade CRD2 is closely related to 5.1 A clades (Huang *et al.,* 2011; Ahlgren and Rocap, 2012).

Therefore clades within each of the phylogenetically defined subclusters 5.1 A and 5.1B do not share a single, cohesive habitat or niche. It had been previously suggested that clades within the same subcluster have similar ecological strategies because of their shared evolutionary history, with 5.1 A clades described as 'oligotrophs/specialists' and 5.1B clades as 'generalists/opportunists' partly because 5.1 A strains have significantly fewer enzymes that sense and respond to the environment than 5.1B strains (Palenik *et al.,* 2006; Dufresne *et al.,* 2008). Our data suggest that *Synechococcus* clades occupying the same regimes have independently arisen multiple times and represent examples of convergent or parallel evolution. We herein use the latter term since it refers to convergent evolution that has occurred among closely related organisms (Freeman and Herron, 1998).

Parallel evolution also describes the overall pattern of ecotype radiation of both *Prochlorococcus* and *Synechococcus* together. Ecotypes that occupy the three major biomes have emerged from within at least four major lineages within the marine picocyanobacteria (Figure 7). Specifically, ecotypes that dominate in low-Fe, HNLC waters have evolved in both *Synechococcus* and *Prochlorococcus*. Similarly, cold- and warm-adapted ecotypes that partition the high or low latitude portions of the open-ocean gyres have emerged independently in both genera. Interestingly, the shift from warm-water *Synechococcus* clades (II, III and X) to cold-water clades (I and IV) occurs at ~ 21 °C in the open ocean, similar to the thresholds for the shifts in dominance of warm-adapted HLII to cold-adapted HLI *Prochlorococcus* ecotypes (Supplementary Figure S7; Johnson *et al.,* 2006).

Parallel evolution of ecotypes may be facilitated by the horizontal transfer of adaptive genes, from both other marine bacteria and among picocyanobacteria (Kettler *et al.,* 2007; Dufresne *et al.,* 2008; Scanlan *et al.,* 2009; Zhaxybayeva *et al.,* 2009). Such gene transfer is at least partially mediated by cyanophage, which can cross-infect multiple *Prochlorococcus* ecotypes and *Synechococcus* clades (Sullivan *et al.,* 2006). Genome comparisons within the picocyanobacteria reveal paraphyletic distributions of potentially adaptive genes as a result of horizontal gene transfer (Dufresne *et al.,* 2008; Scanlan *et al.,* 2009; Zhaxybayeva *et al.,* 2009). For example, phycobilisome genes appear to have been horizontally transferred between *Synechococcus* clades, resulting in a diversity of pigment types that are shared across lineages (Six *et al.,* 2007; Everroad and Wood, 2012). Similarly, the distribution of phosphate utilization genes among *Prochlorococcus* strains is better explained by P availability in the oceanic basin from

which they were isolated than their core-gene phylogeny (Martiny *et al.,* 2006). The recent discovery of nitrate utilization in *Prochlorococcus* is another example where gene acquisition occurred independently across ecotypes (Berube *et al.,* 2015). In addition to sharing genes across ecotypes that are important for adaptation to particular niches, co-occurring ecotypes could also possess distinct genomic changes (as a result of horizontal gene transfer, *de novo* mutation or gene loss) that allow for adaptation to the same environment, but by different mechanisms. Since currently only the genomes of one or two strains have been sequenced within each *Synechococcus* clade, it is not possible to assess intra- vs inter-clade genomic differences (Zhaxybayeva *et al.,* 2009). As more genomes are sequenced, it will be feasible to investigate how distribution of specific accessory genes may confer adaptation to their niches to understand the mechanisms behind this observed parallel evolution. In particular, it will be intriguing to determine whether low Fe *Synechococcus* CRD1 and CRD2 identified herein (for which there are currently no genomes publically available) share adaptations observed in low Fe, HNLC *Prochlorococcus* ecotypes (Rusch *et al.,* 2010).

Recently described ecotypes in SAR11, an abundant lineage of pelagic *Alphaproteobacteria*, also exhibit parallel evolution (Vergin *et al.,* 2013). SAR11 ecotypes occupy distinct niches according to depth and season but deep and surface types are paraphyletic rather than monophyletic. Thus, rather than a model of evolution whereby invasion of niches occurs in a sequential manner, ecotypes that occupy deep or shallow portions of the water column appear to have independently arisen multiple times within different major lineages of SAR11. Parallel evolution of ecotypes may prove to be a common theme for marine microbial populations.

Ecological significance of co-occurring ecotypes
Co-occurrence of closely related ecotypes initially seems unexpected as outlined in the classic 'paradox of the plankton' argument, which suggests competition for the same resources in a relatively well-mixed environment should produce a small number of winning taxa (Hutchinson, 1961). Other bottom-up factors likely differentiate the niches of co-occurring clades such as the ability to use organic N and P or differences in the ratio at which multiple nutrients are utilized (Roy and Chattopadhyay, 2007). Top-down pressures such as grazing (Apple *et al.,* 2011) and virus infection (Suttle, 2007) or 'lateral' mechanisms such as commensalism (Morris *et al.,* 2012) and allelopathy (Paz-Yepes *et al.,* 2013) could also contribute to the maintenance of multiple ecotypes. Regardless of the mechanism for co-existence, a potential consequence of the parallel evolution of co-occurring populations is a level of functional redundancy in the community that may influence

overall stability (Tilman and Downing, 1994). Multiple ecotypes within a biome may provide a larger pool of genetic variants from which adaptive strains could be selected and the potential loss of one ecotype within a biome could be buffered by the persistence of the other member(s). Thus co-occurring ecotypes perhaps create the potential for resiliency within marine microbial communities in the face of impending climate change.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

References

Ahlgren NA, Rocap G. (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 NA, Rocap G. (2012). Diversity and distribution of marine *Synechococcus*: multiple gene phylogenies for consensus classification and development of qPCR assays for sensitive measurement of clades in the ocean. *Front Microbiol* **3**: 213.

Ahlgren NA, Noble AE, Patton AP, Roache-Johnson K, Jackson L, Robinson D *et al.* (2014). The unique trace metal and mixed layer conditions of the Costa Rica upwelling dome support a distinct and dense community of *Synechococcus*. *Limnol Oceanogr* **59**: 2166–2184.

Apple JK, Strom SL, Palenik B, Brahamsha B. (2011). Variability in protist grazing and growth rates on different marin *Synechococcus* isolates. *Appl Environ Microbiol* **77**: 3074–3084.

Benjamini Y, Hochberg Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol* **57**: 289–300.

Berube PM, Biller SJ, Kent AG, Berta-Thompson JW, Roggensack SE, Roache-Johnson KH *et al.* (2015).

Physiology and evolution of nitrate acquisition in *Prochlorococcus*. *ISME J* **9**: 1195–1207.

Boyd PW, Jickells T, Law CS, Blain S, Boyle EA, Buesseler KO *et al.* (2007). Mesoscale iron enrichment experiments 1993-2005: synthesis and future directions. *Science* **315**: 612–617.

Bruland KW, Rue EL, Smith GJ. (2001). Iron and macronutrients in California coastal upwelling regimes: Implications for diatoms blooms. *Limnol Oceanogr* **46**: 1661–1674.

Chappell PD, Moffett JW, Hynes AM, Webb EA. (2012). Molecular evidence of iron limitation and availability in the global diazotroph *Trichodesmium*. *ISME J* **6**: 1728–1739.

Choi DH, Noh JH. (2009). Phylogenetic diversity of *Synechococcus* strains isolated from the East China Sea and the East Sea. *FEMS Microbiology Ecology* **69**: 439–448.

Coale KH, Johnson KS, Fitzwater SE, Gordon RM, Tanner S, Chavez FP *et al.* (1996). A massive phytoplankton bloom induced by an ecosystem-scale iron fertilization experiment in the equatorial Pacific Ocean. *Nature* **383**: 495–501.

Coleman M, Chisholm SW. (2007). Code and context: *Prochlorococcus* as a model for cross scale biology. *Trends Microbiol* **15**: 398–407.

Dean RB, Dixon WJ. (1951). Simplified statistics for small numbers of obesrvations. *Anal Chem* **23**: 636–638.

Dufresne A, Ostrowski M, Scanlan DJ, Garczarek L, Mazard S, Palenik BP *et al.* (2008). Unraveling the genomic mosaic of a ubiquitous genus of marine cyanobacteria. *Genome Biol* **9**: R90.

Everroad RC, Wood AM. (2012). Phycoerythrin evolution and diversification of spectral phenotype in marine *Synechococcus* and related picocyanobacteria. *Mol Phylogenet Evol* **64**: 381–392.

Ferris MJ, Palenik B. (1998). Niche adaptation in ocean cyanobacteria. *Nature* **396**: 226–228.

Field CB, Behrenfeld MJ, Randerson JT, Falkowski P. (1998). Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science* **281**: 237–240.

Flombaum P, Gallegos JL, Gordillo RA, Rincon J, Zabala LL, Jiao N *et al.* (2013). Present and future global distributions of the marine Cyanobacteria Prochlorococcus and Synechococcus. *Proc Natl Acad Sci USA* **10**: 9824–9829.

Freeman S, Herron JC. (1998). *Evolutionary Analysis*. Prentice HallUpper Saddle River, NJ.

Fuller NJ, Marie D, Partensky F, Vaulot D, Post AF, Scanlan DJ. (2003). Clade-specific 16 S ribosomal DNA oligonucleotides reveal the predominance of a single marine *Synechococcus* clade throughout a stratified water column in the Red Sea. *Appl Environ Microbiol* **69**: 2430–2443.

Fuller NJ, Tarran GA, Yallop M, Orcutt KM, Scanlan DJ. (2006). Molecular analysis of picocyanobacterial community structure along an Arabian Sea transect reveals distinct spatial separation of lineages. *Limnol Oceanogr* **51**: 2515–2526.

Fung IY, Meyn SK, Tegen I, Doney SC, John JG, Bishop JKB. (2000). Iron supply and demand in the upper ocean. *Glob Biogeochem Cycles* **14**: 281–295.

Gutierrez-Rodriguez A, Slack G, Daniels EF, Selph KE, Palenik B, Landry MR. (2014). Fine spatial structure of genetically distinct picocyanobacterial populations

344

across environmental gradients in the Costa Rica Dome. *Limnol Oceanogr* **59**: 705–723.

Huang S, Wilhelm SW, Harvey HR, Taylor K, Jiao N, Chen F. (2011). Novel lineages of *Prochlorococcus* and *Synechococcus* in the global oceans. *ISME J* **6**: 285–297.

Hutchinson GE. (1961). The paradox of the plankton. *Am Nat* **95**: 137–145.

Jardillier L, Zubkov MV, Pearman J, Scanlan DJ. (2010). Significant CO_2 fixation by small prymnesiophytes in the subtropical and tropical northeast Atlantic Ocean. *ISME J* **4**: 1180–1192.

Johnson KS, Chavez FP, Elrod VA, Fitzwater SE, Pennington JT, Buck KR et al. (2001). The annual cycle of iron and the biological response in central California coastal waters. *Geophys Res Lett* **28**: 1247–1250.

Johnson ZI, Zinser ER, Coe A, McNulty NP, Woodward EMS, Chisholm SW. (2006). Niche partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental gradients. *Science* **311**: 1737–1740.

Kettler GC, Martiny AC, Huang K, Zucker J, Coleman ML, Rodrigue S et al. (2007). Patterns and implications of gene gain and loss in the evolution of *Prochlorococcus*. *PLoS Genet* **3**: 2515–2528.

Malmstrom RR, Rodrigue S, Huang KH, Kelly L, Kern SE, Thompson A et al. (2012). Ecology of uncultured *Prochlorococcus* clades revealed through single-cell genomics and biogeographic analysis. *ISME J* **7**: 184–198.

Mann EL, Ahlgren NA, Moffett JW, Chisholm SW. (2002). Copper toxicity and cyanobacteria ecology in the Sargasso Sea. *Limnol Oceanogr* **47**: 976–988.

Marchetti A, Maldonado MT, Lane ES, Harrison PJ. (2006). Iron requirements of the pennate diatom *Pseudonitzschia*: Comparison of oceanic (high-nitrate, low-chlorophyll waters) and coastal species. *Limnol Oceanogr* **51**: 2092–2101.

Marie D, Simon N, Vaulot D. (2005). Phytoplankton cell counting by flow cytometry. In: Anderson RA (ed) *Algal Culturing Techniques*. Academic Press: Burlington, MA, USA, pp 253–268.

Martin JH, Gordon RM, Fitzwater SE. (1991). The case for iron. *Limnol Oceanogr* **36**: 1793–1802.

Martiny AC, Coleman ML, Chisholm SW. (2006). Phopshate acquisition genes in Prochlrococcus ecotypes: Evidence for genome-wide adaptation. *Proc Natl Acad Sci USA* **103**: 12552–12557.

Martiny AC, Tai APK, Veneziano D, Primeau F, Chisholm SW. (2009). Taxonomic resolution, ecotypes and the biogeography of *Prochlorococcus*. *Environ Microbiol* **11**: 823–832.

Mazard S, Ostrowski M, Partensky F, Scanlan DJ. (2012a). Multi-locus sequence analysis, taxonomic resolution and biogeography of marine Synechococcus. *Environ Microbiol* **14**: 372–386.

Mazard S, Wilson WH, Scanlan DJ. (2012b). Dissection the physiological response to phosphorus stress in marine Synechococcus isolates (CYANOPHYCEAE). *J Phycol* **48**: 94–105.

Mella-Flores D, Mazard S, Humily F, Partensky F, Mahe F, Bariat L et al. (2011). Is the distribution of Prochlorococcus and Synechococcus ecotypes in the Mediterranean Sea affected by global warming? *Biogeosci Discuss* **8**: 4281–4330.

Moore LR, Rocap G, Chisholm SW. (1998). Physiology and molecular phylogeny of coexisting *Prochlorococcus* ecotypes. *Nature* **393**: 464–467.

Moore LR, Post AF, Rocap G, Chisholm SW. (2002). Utilization of different nitrogen sources by the marine cyanobacteria *Prochlorococcus* and *Synechococcus*. *Limnol Oceanogr* **47**: 989–996.

Moore LR, Ostrowski M, Scanlan DJ, Feren K, Sweetsir T. (2005). Ecotypic variation in phosphorus acquisition mechanisms within marine picocyanobacteria. *Aquat Microb Ecol* **39**: 257–269.

Morris JJ, Lenski RE, Zinser ER. (2012). The black queen hypothesis: evolution of dependencies through adaptive gene loss. *mBio* **3**: e00036-12.

Noble AE, Lamborg CH, Ohnemus DC, Lam PJ, Goepfert TJ, Measures C et al. (2012). Basin-scale inputs of cobalt, iron and manganese from the Benguela-Angola front to the South Atlantic Ocean. *Limnol Oceanogr* **57**: 989–1010.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB et al. (2012). vegan: Community Ecology Package 2.0-3.

Paerl RW, Foster RA, Jenkins BD, Montoya JP, Zehr JP. (2008). Phylogenetic diversity of cyanobacterial narB genes from various marine habitats. *Environ Microbiol* **10**: 3377–3387.

Palenik B. (2001). Chromatic adaptation in marine Synechococcus strains. *Appl Environ Microbiol* **67**: 991–994.

Palenik B, Ren QH, Dupont CL, Myers GS, Heidelberg JF, Badger JH et al. (2006). Genome sequence of Synechococcus CC9311: Insights into adaptation to a coastal environment. *Proc Natl Acad Sci USA* **103**: 13555–13559.

Paz-Yepes J, Brahamsha B, Palenik B. (2013). Role of a Microcin-C-like biosynthetic gene cluster in allelopathic interactions in marine *Synechococcus*. *Proc Natl Acad Sci USA* **110**: 12030–12035.

Pittera J, Humily F, Thorel M, Grulois D, Garczarek L, Six C. (2014). Connecting thermal physiology and latitudinal niche partitioning in marine *Synechococcus*. *ISME J* **8**: 1221–1236.

Rivers AR, Jakuba RW, Webb EA. (2009). Iron stress genes in marine *Synechococcus* and the development of a flow cytometric iron stress assay. *Environ Microbiol* **11**: 382–396.

Rocap G, Distel DL, Waterbury JB, Chisholm SW. (2002). Resolution of Prochlorococcus and Synechococcus ecotypes by using 16 S-23 S ribosomal DNA internal transcribed spacer sequences. *Appl Environ Microbiol* **68**: 1180–1191.

Roy S, Chattopadhyay J. (2007). Towards a resolution of ˜the paradox of the plankton': A brief overview of the proposed mechanisms. *Ecol Complex* **4**: 26–33.

Ruan Q, Dutta D, Schwalbach MS, Steele JA, Fuhrman JA, Sun F. (2006). Local similarity analysis reveals unique associations among marine bacterioplankton species and environmental factors. *Bioinformatics* **22**: 2532–2538.

Rusch DB, Martiny AC, Dupont CL, Halpern AL, Venter JC. (2010). Characterization of *Prochlorococcus* clades from iron-depleted oceanic regions. *Proc Natl Acad Sci USA* **107**: 16184–16189.

Scanlan DJ, Ostrowski M, Mazard S, Dufresne A, Garczarek L, Hess WR et al. (2009). Ecological genomics of marine Picocyanobacteria. *Microbiol Mol Biol Rev* **73**: 249–299.

Six C, Thomas J, Garczarek L, Ostrowski M, Dufresne A, Blot N et al. (2007). Diversity and evolution of phycobilisomes in marine *Synechococcus* spp.: a comparative genomics study. *Genome Biol* **8**: R259.

Strickland JDH, Parsons TR. (1968). *A Practical Handbook of Seawater Analysis*. Fisheries Resource Board: Ottawa, Canada.

Stuart RK, Dupont CL, Johnson DA, Paulsen IT, Palenik B. (2009). Coastal strains of marine *Synechococcus* species exhibit increased tolerance to copper shock and a distinctive transcriptional response relative to those of open-ocean strains. *Appl Environ Microbiol* **75**: 5047–5057.

Sullivan MB, Lindell D, Lee JA, Thompson LR, Bielawski JP, Chisholm SW. (2006). Prevalence and evolution of core photosystem II genes in marine cyanobacterial viruses and their hosts. *Plos One* **4**: e234.

Suttle CA. (2007). Marine viruses - major players in the global ecosystem. *Nat Rev Microbiol* **5**: 801–812.

Tai V, Palenik B. (2009). Temporal variation of *Synechococcus* clades at a coastal Pacific Ocean monitoring site. *ISME J* **3**: 903–915.

Tilman D, Downing JA. (1994). Biodiversity and stability in grasslands. *Nature* **367**: 363–365.

Toledo G, Palenik B. (1997). Synechococcus diversity in the California current as seen by RNA polymerase (rpoC1) gene sequences of isolated strains. *Appl Environ Microbiol* **63**: 4298–4303.

Vergin KL, Beszteri B, Monier A, Thrash JC, Kilpert F, Worden AZ *et al.* (2013). High-resolution SAR11 ecotype dynamics at the Bermuda Atlantic Time-series Study site by phylogenetic placement of pyrosequences. *ISME J* **7**: 1322–1332.

Waterbury JB, Watson SW, Valois FW, Franks DG. (1986). Biological and ecological characterization of the marine unicellular cyanobacterium Synechococcus. *Can Bull Fish Aquat Sci* **214**: 71–120.

West NJ, Lebaron P, Strutton PG, Suzuki MT. (2011). A novel clade of *Prochlorococcus* found in high nutrient low chlorophyll waters in the South and Equatorial Pacific Ocean. *ISME J* **5**: 933–944.

Zhaxybayeva O, Doolittle WF, Thane Papke F, Peter Gogarten JG. (2009). Intertwined evolutionary histories of marine Synechococcus and Prochlorococcus marinus. *Genome Biol Evol* **1**: 325–339.

Zinser ER, Johnson ZI, Coe A, Karaca E, Veneziano D, Chisholm SW. (2007). Influence of light and temperature on Prochlorococcus ecotype distributions in the Atlantic Ocean. *Limnol Oceanogr* **52**: 2205–2220.

Zwirglmaier K, Heywood JL, Chamberlain K, Woodward EMS, Zubkov MV, Scanlan DJ. (2007). Basin-scale distribution patterns lineages in the Atlantic Ocean. *Environ Microbiol* **9**: 1278–1290.

Zwirglmaier K, Jardillier L, Ostrowski M, Mazard S, Garczarek L, Vaulot D *et al.* (2008). Global phylogeography of marine Synechococcus and Prochlorococcus reveals a distinct partitioning of lineages among oceanic biomes. *Environ Microbiol* **10**: 147–161.

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