

ORIGINAL ARTICLE

Temporal dynamics of *Prochlorococcus* ecotypes in the Atlantic and Pacific oceansRex R Malmstrom¹, Allison Coe¹, Gregory C Kettler², Adam C Martiny^{1,3},
Jorge Frias-Lopez^{1,4}, Erik R Zinser^{1,5} and Sallie W Chisholm¹¹Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA;²Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA; ³Departments of Earth System Science and Ecology and Evolutionary Biology, University of California, Irvine, CA, USA; ⁴The Forsyth Institute, Boston, MA, USA and ⁵Department of Microbiology, University of Tennessee, Knoxville, TN, USA

To better understand the temporal and spatial dynamics of *Prochlorococcus* populations, and how these populations co-vary with the physical environment, we followed monthly changes in the abundance of five ecotypes—two high-light adapted and three low-light adapted—over a 5-year period in coordination with the Bermuda Atlantic Time Series (BATS) and Hawaii Ocean Time-series (HOT) programs. Ecotype abundance displayed weak seasonal fluctuations at HOT and strong seasonal fluctuations at BATS. Furthermore, stable ‘layered’ depth distributions, where different *Prochlorococcus* ecotypes reached maximum abundance at different depths, were maintained consistently for 5 years at HOT. Layered distributions were also observed at BATS, although winter deep mixing events disrupted these patterns each year and produced large variations in ecotype abundance. Interestingly, the layered ecotype distributions were regularly reestablished each year after deep mixing subsided at BATS. In addition, *Prochlorococcus* ecotypes each responded differently to the strong seasonal changes in light, temperature and mixing at BATS, resulting in a reproducible annual succession of ecotype blooms. Patterns of ecotype abundance, in combination with physiological assays of cultured isolates, confirmed that the low-light adapted eNATL could be distinguished from other low-light adapted ecotypes based on its ability to withstand temporary exposure to high-intensity light, a characteristic stress of the surface mixed layer. Finally, total *Prochlorococcus* and *Synechococcus* dynamics were compared with similar time series data collected a decade earlier at each location. The two data sets were remarkably similar—testimony to the resilience of these complex dynamic systems on decadal time scales.

The ISME Journal (2010) 4, 1252–1264; doi:10.1038/ismej.2010.60; published online 13 May 2010

Subject Category: Microbial population and community ecology

Keywords: *Prochlorococcus*; *Synechococcus*; ecotype; time-series; HOT; BATS

Introduction

A key challenge facing marine microbiology is to understand how microbial diversity and biogeochemical cycles are linked, and to eventually incorporate this understanding into conceptual and predictive ocean models. Physiological and genetic analyses of cultured isolates, as well as metagenomic studies of whole communities (Venter *et al.*, 2004; DeLong *et al.*, 2006), are uncovering more and more about the metabolic potential of microbes comprising these assemblages. In addition, time series studies are revealing how the composition of microbial communities varies over both time

and space (Fuhrman *et al.*, 2006; Carlson *et al.*, 2009; Treusch *et al.*, 2009). Coupling the spatial and temporal dynamics of specific microbial groups with insights into their metabolic potential is essential for developing a quantitative understanding of the roles these microbes have in marine ecosystems.

Over the past 20 years, investigations of the unicellular cyanobacterium *Prochlorococcus* have provided insight into both the biogeography and metabolic potential of this group. *Prochlorococcus* is typically the most abundant photoautotroph in tropical and subtropical waters (Campbell *et al.*, 1994; Partensky *et al.*, 1999), and its abundance varies seasonally at some locations (Campbell *et al.*, 1997; DuRand *et al.*, 2001). Studies of cultured isolates have revealed the optimal light and temperature levels differ among the strains (Moore *et al.*, 1998, 2002; Moore and Chisholm, 1999; Zinser *et al.*, 2007), as do the nutrient pools available to them (Moore *et al.*, 2002, 2005).

Correspondence: S Chisholm, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 15 Vassar St, Cambridge, MA, 02139, USA.
E-mail: chisholm@mit.edu

Received 4 January 2010; revised 22 March 2010; accepted 24 March 2010; published online 13 May 2010

Genomic analyses of isolates (Rocap *et al.*, 2003; Coleman *et al.*, 2006; Kettler *et al.*, 2007), and metagenomic analyses of natural communities (DeLong *et al.*, 2006; Martiny *et al.*, 2006, 2009a), have provided additional insights into the genetic diversity, metabolic potential and evolutionary history of the group. The ability to examine the abundance and distribution of *Prochlorococcus* in the wild, and measure its genomic and metabolic properties in the lab and field, make *Prochlorococcus* a model system for advancing our understanding the ecology of marine microbes.

Prochlorococcus is composed of several clades that are physiologically and phylogenetically distinct with respect to their optimal light and temperature environments (Moore *et al.*, 1998; Moore and Chisholm, 1999; Rocap *et al.*, 2002; Johnson *et al.*, 2006; Kettler *et al.*, 2007; Zinser *et al.*, 2007). These clades have been referred to as 'ecotypes' (Moore *et al.*, 1998; Moore and Chisholm, 1999; Rocap *et al.*, 2002) following the broader historical designation for genetically distinct subgroups within a species that are adapted to specific environments (Turesson, 1922; Clausen *et al.*, 1940). They do not necessarily conform to the more recent and narrowly defined 'ecotype' concept developed by Cohan and others (Cohan, 2001; Cohan and Perry, 2007), which has become a notable model for exploring the theoretical basis for divergence among bacteria (Ward *et al.*, 2006; Frasier *et al.*, 2009). A more detailed discussion of the different uses of the term 'ecotype' is provided by Coleman and Chisholm, (2007).

The abundance of *Prochlorococcus* ecotypes in various oceanic regions has been studied extensively (West and Scanlan, 1999; Ahlgren *et al.*, 2006; Bouman *et al.*, 2006; Johnson *et al.*, 2006; Zinser *et al.*, 2006, 2007). Members of the two high-light adapted (HL) ecotypes, eMIT9312 and eMED4, are most abundant in the upper regions of the euphotic zone, whereas low-light adapted (LL) ecotypes such as eNATL and eMIT9313 are most abundant in the lower euphotic zone (West *et al.*, 2001; Johnson *et al.*, 2006; Zinser *et al.*, 2007). These distribution patterns agree well with differences in the optimal light levels for representative ecotype strains (Moore and Chisholm, 1999; Zinser *et al.*, 2007). Furthermore, eMED4 tends to dominate in cooler, higher latitude waters, whereas eMIT9312 dominates warmer, lower latitude waters (Johnson *et al.*, 2006; Zwirgmaier *et al.*, 2007); also in good agreement with temperature optima of representative strains. Water column stability and nutrient concentrations have also been correlated with ecotype abundance (Bouman *et al.*, 2006; Johnson *et al.*, 2006), and community structure (Martiny *et al.*, 2009b), although the underlying causalities of these relationships are not well understood.

The environmental factors influencing *Prochlorococcus* ecotype abundance, such as light, temperature and water column mixing, vary over time and

can display seasonal patterns, thus we might expect ecotype dynamics to do the same. Although we have analyzed ecotype variability over a span of a few days (Zinser *et al.*, 2007), extensive time-series studies have not been conducted, and thus little is known about the dynamics of *Prochlorococcus* ecotypes on the scale of months to years. Analyses on these time scales should help refine our understanding of ecotype/environment interactions, and provide data sets for testing models designed to explore the dynamics of phytoplankton community structure (Follows *et al.*, 2007).

To this end, we followed the spatial and temporal dynamics of *Prochlorococcus* ecotypes at monthly intervals over 5 years in coordination with the Bermuda Atlantic Time Series (BATS) and the Hawaii Ocean Time-series (HOT) programs (Karl and Lukas, 1996; Steinberg *et al.*, 2001). Both programs are focused on oligotrophic, open ocean sites where *Prochlorococcus* is found in abundance (Campbell *et al.*, 1994; DuRand *et al.*, 2001). However, the physics and chemistry of these locations differ, most notably by the stronger seasonal mixing events at BATS and the higher inorganic phosphate concentrations at HOT (Wu *et al.*, 2000; Cavender-Bares *et al.*, 2001; Steinberg *et al.*, 2001). Here, we explore how changes in environmental factors such as mixing, light and temperature are related to ecotype abundance and distribution, as well as how they co-vary temporally over 5 years. We also examine some of the emergent patterns in the field data through studies of light-shock tolerance in cultured isolates of different *Prochlorococcus* ecotypes.

Materials and methods

Sample collection

Beginning in November 2002, flow cytometry and qPCR samples were collected over a 5-year period during monthly cruises for the Bermuda Atlantic Time Series (BATS) and Hawaii Ocean Time-series (HOT) programs. Additional samples were collected bi-weekly between February and April at BATS. Samples were collected from 12 depths (1, 10, 20, 40, 60, 80, 100, 120, 140, 160, 180 and 200 m) at the BATS site (~5 nautical mile radius around 31° 40'N, 64° 10' W), and from 12 depths (5, 25, 45, 60, 75, 85, 100, 115, 125, 150, 175 and 200 m) at Sta. ALOHA (~5 nautical mile radius around 22° 45'N, 158° 00'W). These locations are referred to BATS and HOT throughout the article for simplicity.

Flow cytometry

Whole seawater samples were immediately fixed with glutaraldehyde (final conc. 0.125% v/v) for 10 min, frozen in liquid nitrogen, and stored at -80°C until samples could be processed in the laboratory using an influx flow cytometer

(Becton Dickinson, Franklin Lakes, NJ, USA). *Prochlorococcus* and *Synechococcus* populations were identified and quantified based on their unique autofluorescence and scatter signals (Olson *et al.*, 1990a, 1990b). *Prochlorococcus* could not always be clearly distinguished in the upper 40 m at BATS from June–September, thus these profiles were excluded from depth-integrated counts in Figure 2. Flow cytometry profiles from early 2003 were not processed.

Primer re-design for ecotype eSS120

New primers for the eSS120 ecotype were designed in ARB (Ludwig *et al.*, 2004), using a large database of environmental ITS sequences (Martiny *et al.*, 2009b). These primers, 5'-AACAACTTTCTCCTGGCT-3' and 5'-AGTTGATCAGTGGAGGTAAG-3', matched 87% of known eSS120 ITS sequences, but did not target MIT9211. The specificity of the primers was initially tested against cultured isolates from other ecotypes such as MIT9313, MIT9312, MED4 and NATL2a. These tests confirmed specificity within the dynamic range of the assay (~ 5 to 5×10^5 cells ml⁻¹). Specificity was also confirmed by cloning and sequencing ITS regions amplified from field samples from BATS and HOT using the new primers. All amplified ITS sequences clustered with strains SS120 and MIT9211, both members of the eSS120 ecotype (Kettler *et al.*, 2007), in a bootstrapped ($n = 100$) neighbor-joining tree constructed with the Bosque software package (Ramirez-Flandes and Ulloa, 2008) (Supplementary Figure 1). The new primers substantially increased counts of the SS120 ecotype (Supplementary Materials).

Quantitative PCR

Samples for qPCR were collected and processed as described previously (Ahlgren *et al.*, 2006; Zinser *et al.*, 2006, 2007) with two small modifications. First, reaction volumes were reduced from 25 μ l to 15 μ l and performed in 384-well plates on the Light Cycler 480 for samples collected after 2003. Second, concentrations of eMIT9313-specific primers were increased to 5 μ M to improve sensitivity. Estimated abundances that fell below the lowest value of the standard curve were set to the theoretical detection limit of 0.65 cells ml⁻¹. Samples were excluded if their melt curves contained multiple peaks or peaks different from those in the DNA standards to ensure only the targeted ecotypes were quantified. Missing data were determined by linear interpolation when abundance estimates were available for the depths immediately above and below the missing value.

Environmental data extraction

Temperature, salinity and potential density were downloaded from the HOT and BATS websites. Missing data were determined by linear interpolation

when values were available for the depths immediately above and below the missing data point. Mixed layer depths, light attenuation coefficients and *Prochlorococcus* and *Synechococcus* abundance (1991–1995) were also downloaded directly from the HOT website. Mixed layer depths and attenuation coefficients at BATS were calculated from bottle-derived profiles and SeaWiFS Profiling Multi-channel Radiometer profiles of photosynthetically active radiation (PAR) collected by Bermuda Bio-Optics Program. The mixed layer depth was determined when potential density differed from surface values by > 0.125 kg m⁻³. Attenuation coefficients at BATS were calculated by linear regression of log-transformed PAR values; only profiles with an $R^2 > 0.993$ were used.

Solar irradiance

Surface irradiance was determined from SeaWiFS-derived estimates of daily-integrated PAR. Eight-day means of daily integrated PAR values from March 2000 to July 2006 were calculated from a 27 km by 27 km region around BATS and HOT (White *et al.*, 2007). A two component Fourier model with a period of 1 year was fitted to PAR data, producing an R^2 of 0.90 and 0.93 at BATS and HOT, respectively. Modeled PAR data were used to estimate surface irradiance on the day of sample collection. This model was necessary to account for seasonal variability in solar flux due to changes in day length and solar azimuth. These changes result in a roughly two-fold difference in daily-integrated solar flux between summer and winter (for example, $\sim 20\,000$ – $56\,000$ mE m⁻² d⁻¹ at BATS, and $\sim 31\,000$ – $58\,000$ mE m⁻² d⁻¹ at HOT, using this model).

The relationships between ecotype abundance and PAR plotted in Figure 1 and Supplementary Figure 2 were determined using robust locally weighted linear regression (LOWESS) in MATLAB. Robust LOWESS is more resistant to outliers (Cleveland, 1979), which are defined in MATLAB's robust LOWESS function as data outside six mean absolute deviations.

Time series and other statistical analyses

Integrated ecotype abundance, surface PAR, and mixed layer depth were log-transformed, detrended, and resampled at a regular monthly interval to meet the mathematical requirements for time series analyses (Legendre and Legendre, 1998). Detrended data were calculated as the residuals of linear regression of log-transformed data against time. Detrended data were resampled every 30.44 days, which is equivalent to 12 measurements per year, using linear interpolation. Coefficients of autocorrelation and cross-correlation were determined in MATLAB, and s.d. calculated as $n^{-1/2}$, where n is the number of resampled data points. Spectral

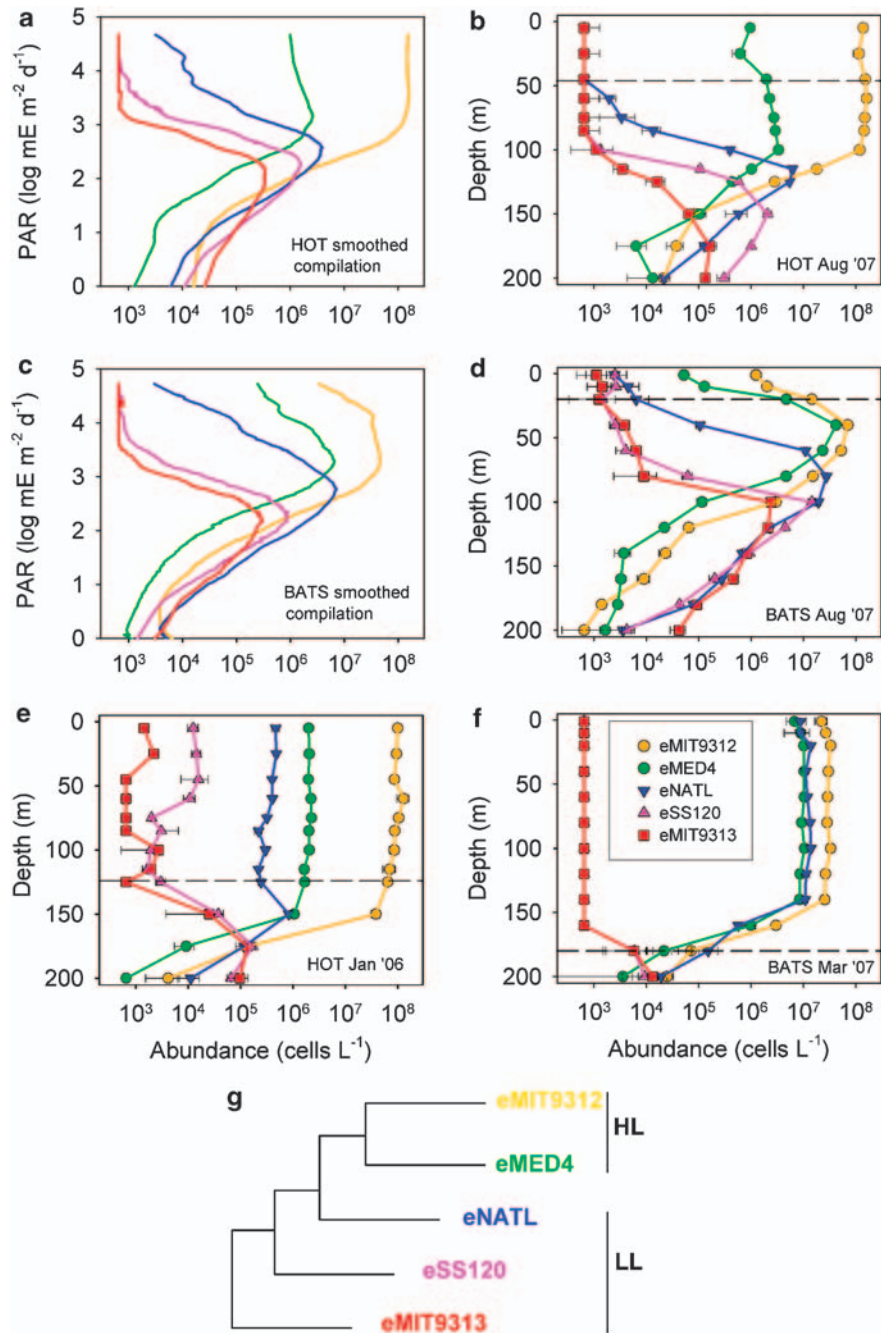


Figure 1 Ecotype distribution patterns in relation to depth and photosynthetically active radiation (PAR). A smoothed illustration of the relationship between ecotype abundance and irradiance, calculated from a compilation of data from all 5 years, was plotted for HOT (a) and BATS (c). Data were smoothed by locally weighted regression of ecotype abundance against irradiance. Smoothed data exclude profiles with a mixed layer depth > 100 m. Representative depth profiles of ecotype abundance at HOT (b, e) and BATS (d, f) during periods of stratification and deep mixing. Dashed lines mark the mixed layer depth, and error bars represent one s.d. An illustration of the color scheme and general phylogenetic relationships among ecotypes after (Rocap *et al.*, 2002; Kettler *et al.*, 2007) is displayed in (g). This tree serves only as reference to the other panels; branch lengths are not to scale.

analysis by discrete Fourier transformation calculated using the Fast Fourier Transformation (FFT) algorithm, was also performed in MATLAB. The power spectral density was estimated as the absolute value of FFT².

Differences among ecotypes in the average depth of maximum abundance were tested using repeated-

measures ANOVA, followed by a Tukey *post-test* ($\alpha = 0.05$), using MATLAB.

Non-parametric partial correlation coefficients (Spearman R) were calculated in MATLAB to determine the relationship between abundance and temperature while controlling for the influence of light.

Light shock experiments

Serial batch cultures of axenic *Prochlorococcus* strains MED4, NATL2a and SS120 were grown in Sargasso seawater-based Pro99 media (Moore *et al.*, 2007) and illuminated by cool white fluorescent lamps. Cultures were transferred at least four times to acclimatize them to $35 \mu\text{E m}^{-2} \text{s}^{-1}$ of continuous light. Duplicate acclimated cultures that were in log-phase growth were then exposed to $400 \mu\text{E m}^{-2} \text{s}^{-1}$ for 4 h before being returned to their initial light levels. *In vivo* chlorophyll fluorescence was measured before and after light shock with a 10-AU fluorometer (340–500 nm excitation and 680 nm emission filters), and cell counts determined by flow cytometry as described above.

Results and Discussion

Ecotype distribution with light and temperature

Striking similarities between HOT and BATS emerged when data from all depths and all 5 years were combined for a synoptic analysis of *Prochlorococcus* ecotype abundance along irradiance/depth gradients (Figures 1a and b; Supplementary Figure 2). As expected, the two high-light adapted ecotypes, eMIT9312 and eMED4, were most abundant at higher irradiances, with abundance dropping off sharply below $1500 \text{mE m}^{-2} \text{d}^{-1}$ of photosynthetically active radiation (Figures 1a and b; Supplementary Figures 2a and e). In contrast, two low-light adapted ecotypes, eSS120 and eMIT9313, usually reached maximum abundances between 100 and $250 \text{mE m}^{-2} \text{d}^{-1}$, and were typically at or near detection limits at irradiances $>1500 \text{mE m}^{-2} \text{d}^{-1}$. The eNATL group, also low-light adapted had an intermediate distribution, reaching maximum abundance at $300\text{--}600 \text{mE m}^{-2} \text{d}^{-1}$. Unlike the other low light ecotypes, eNATL abundance was occasionally high at irradiance levels $>1500 \text{mE m}^{-2} \text{d}^{-1}$ (Figures 1a and b; Supplementary Figures 2c and h), which is consistent with hypothesis that eNATL can tolerate exposure to higher light levels than eSS120 and eMIT9313 (Coleman and Chisholm, 2007; Zinser *et al.*, 2007).

This consistent relationship between irradiance and ecotype abundance results in a 'layered' depth distribution at both locations (Figures 1a–d). At HOT, for example, HL ecotype eMIT9312 tended to reach maximum abundance at shallower depths than did its fellow HL ecotype eMED4 (Table 1). LL ecotypes also partitioned the water column at HOT, with eNATL abundance peaking at significantly shallower depths than eSS120 and eMIT9313 (Table 1). Similar patterns in ecotype distribution were also observed at BATS (Table 1; Figures 1b and d), except for during deep mixing events, defined here as when the mixed layer depth was $>100 \text{m}$. During periods of deep mixing physical homogenization appears to overwhelm biological partitioning of the water column, resulting in

Table 1 Average depth of maximum ecotype abundance at HOT and BATS (mean \pm STD)

Ecotype	HOT (m)	BATS (m)
eMIT9312	42 \pm 27 ^A	54 \pm 28 ^A
eMED4	71 \pm 29 ^B	60 \pm 26 ^A
eNATL	105 \pm 18 ^C	87 \pm 22 ^B
eSS120	118 \pm 26 ^D	101 \pm 28 ^C
eMIT9313	128 \pm 30 ^D	107 \pm 25 ^C

Samples collected when the mixed layer depth was $>100 \text{m}$ were excluded. Values with different superscripts are significantly different (Tukey post test of repeated measures ANOVA; $\alpha = 0.05$).

uniform depth distributions (Figures 1e and f). Therefore, data from periods of deep mixing were removed from statistical analysis of depth distributions.

It is remarkable that, with the exception of periods of deep mixing, the general patterns of ecotype abundance appear relatively consistent in the Atlantic and Pacific despite substantial differences in the chemical and physical environment. Furthermore, these patterns are consistent with those found previously in a variety of ocean regions (West and Scanlan, 1999; West *et al.*, 2001; Bouman *et al.*, 2006; Johnson *et al.*, 2006; Zinser *et al.*, 2006, 2007). That is, the ecotypes tend to partition the water column by depth, with eNATL reaching maximum abundance in between the peaks of HL ecotypes MED4 and MIT9312, and other LL ecotypes. The similarities in the distributions along depth/light gradients suggest that *Prochlorococcus* ecotypes are responding in a consistent fashion to irradiance regardless of geography.

Temporal dynamics: depth-integrated *Prochlorococcus* and *Synechococcus* populations

Using flow cytometry, we measured the abundance of *Prochlorococcus*, and its close relative *Synechococcus* (Rocap *et al.*, 2002), to provide an overall framework for exploring *Prochlorococcus* ecotype dynamics. At BATS, the depth-integrated *Prochlorococcus* population displayed a strong seasonal pattern, reaching the highest levels in the late summer and fall, and the lowest in the late winter during the annual deep mixing events (Figure 2a). *Synechococcus* displayed the inverse pattern, and even occasionally exceeded the abundance of *Prochlorococcus* during deep mixing events (Figure 2b). This is the same pattern reported by DuRand *et al.* (2001) for *Prochlorococcus* and *Synechococcus* at BATS from 1990 to 1994 (Figures 2a and b). The concordance between these two data sets is remarkable, given that they are separated by more than a decade.

Variations in *Prochlorococcus* and *Synechococcus* abundance were much less dramatic at HOT. Integrated abundance of *Prochlorococcus* varied just over 2-fold throughout the time series, and did not always reach maximal abundance in the summer or minimal abundance in the winter (Figure 2c).

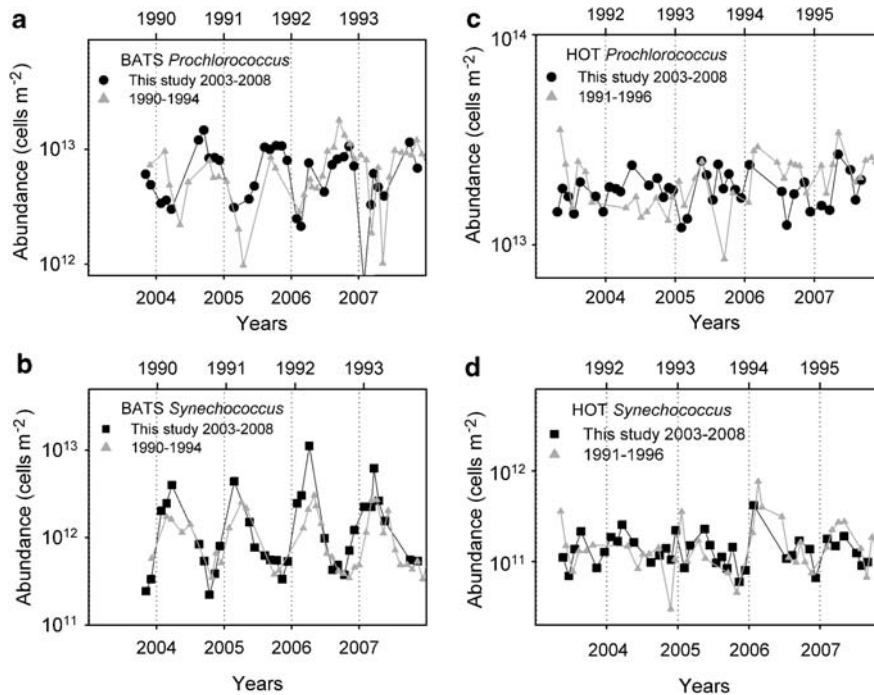


Figure 2 Abundance of *Prochlorococcus* and *Synechococcus* at BATS and HOT determined by flow cytometry (bottom x axis). Abundance levels from this study are compared with those collected at BATS from 1990 to 1994 (DuRand *et al.*, 2001) (a, b), and at HOT from 1991 to 1995 (c, d) (top x axis), which were also determined by flow cytometry.

Synechococcus still displayed an annual pattern, tending to reach peak abundance during winter, but they were never as abundant as *Prochlorococcus*, in contrast to what was observed at BATS. These abundance and variability levels are also consistent with those observed over a decade ago at HOT (Figures 2c and d)

Temporal dynamics: depth-integrated ecotype abundance patterns

As was seen in the total *Prochlorococcus* population, the depth-integrated (0–200 m) abundance of all five ecotypes followed clear annual patterns at BATS (Figure 3a). Spectral analysis of each ecotype revealed dominant peaks in the power spectrum at a period of 1 year (Supplementary Figure 3), and autocorrelations displayed a sinusoidal pattern, with peaks in autocorrelation every 12 months (Supplementary Figure 4). Furthermore, fitting a single component Fourier series with a period of 1 year to each ecotype produced R^2 values ranging from 0.48 (eMIT9313) to 0.67 (eSS120), indicating that most of the variability in integrated abundance could be accounted for by an annual oscillation.

Annual variations in temperature and mixed layer depth were smaller at HOT than at BATS, as were the variations in ecotype abundance (Figure 3b). Although spectral analysis did reveal that all ecotypes had a peak in the power spectrum at a period of 1 year, this was not the only strong signal,

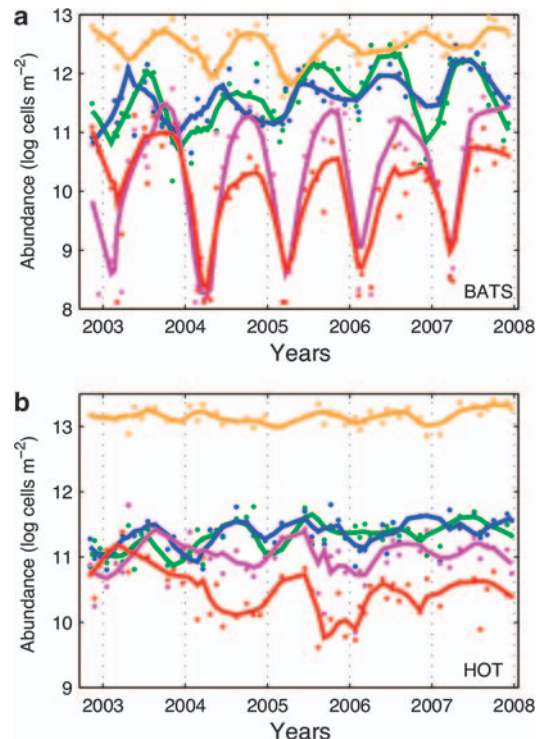


Figure 3 Integrated ecotype abundance (0–200 m) at BATS (a) and HOT (b) from 2003 to 2008. Solid lines represent abundances smoothed by locally weighted regression for eMIT9312 (yellow), eMED4 (green), eNATL (blue), eSS120 (purple) and eMIT9313 (red).

particularly for the LL ecotypes (Supplementary Figure 5). Fitting a single component Fourier series with a period of 1 year to each ecotype produced R^2 values ranging from only 0.17 (eMIT9312) to 0.33 (eMIT9313). In addition, autocorrelations did not display the same strong sinusoidal pattern as seen at BATS (compare Supplementary Figure 3 with Supplementary Figure 6). Thus, while there was a component of annual variability to ecotype abundance at HOT, there was also a variability at periods greater and less than 1 year. This suggests that when annual environmental variations are more moderate, the impact of intra- and inter-annual events, such as passage of mesoscale eddies or El Niño/La Niña oscillations—both known to influence primary production and phytoplankton community composition at HOT (Karl *et al.*, 1995; Letelier *et al.*, 2000; Corno *et al.*, 2007; Bibby *et al.*, 2008)—may become more pronounced.

Although ecotype abundance followed a clear annual cycle at BATS, the cycles were not synchronized among ecotypes, that is, different ecotypes reached peak abundance at different times. In each of the 5 years, the eNATL ecotype reached its maximum integrated abundance about 4 months after winter deep mixing events (Table 2), typically in June (Supplementary Figure 7). Abundance of eMED4 peaked roughly 1 month later, whereas the LL-clades eSS120 and eMIT9313 reached maximal abundance about 8 months after the deep mixing event (Table 2). HL ecotype eMIT9312 also reached maximal abundances at around the same time as eSS120 and eMIT9313, typically in October or November (Supplementary Figure 7). The regularity of this pattern indicates that each ecotype responded to changes in environmental conditions in different, yet consistent ways, resulting in an annually repeating succession of ecotypes.

The succession of *Prochlorococcus* ecotypes is in some ways reminiscent of classical phytoplankton succession models, although it is occurring at a much smaller phylogenetic scale; all *Prochlorococcus* differ in 16S rRNA sequence by <3% (Moore *et al.*, 1998), which would collectively constitute a single

bacterial species by conventional standards (Stackebrandt and Goebel, 1994). Interestingly, two recent time series studies at BATS have also uncovered annual cycles and succession patterns in other microbial groups, most notably the SAR11 clade (Carlson *et al.*, 2009; Treusch *et al.*, 2009). For example, one SAR11 subgroup reaches peak abundance in surface waters during the summer, whereas another subgroup peaks in the winter (Carlson *et al.*, 2009). SAR11 bacteria are the most abundant heterotrophs at BATS and are major consumers of dissolved organic compounds (Morris *et al.*, 2002; Malmstrom *et al.*, 2005), whereas *Prochlorococcus* is the most abundant photoautotroph and a substantial source of dissolved organics (DuRand *et al.*, 2001; Bertilsson *et al.*, 2005). In addition, SAR11 bacteria and *Prochlorococcus* are both major consumers of small compounds like amino acids and dimethylsulfoniopropionate (DMSP) (Zubkov *et al.*, 2003; Malmstrom *et al.*, 2004; Vila-Costa *et al.*, 2006; Michelou *et al.*, 2007), which are significant sources of C, N and S to marine microbial communities. Therefore, it seems plausible that succession in these two abundant groups could be linked through the production of, and competition for, dissolved organic compounds. Uncovering potential links in the dynamics of the dominant microbial groups presents a future challenge.

Ecotype abundance in different regions of the euphotic zone

Variations in light and temperature levels, which impact the growth of *Prochlorococcus* (Moore and Chisholm, 1999; Johnson *et al.*, 2006; Zinser *et al.*, 2007), are greater in surface waters than at depth. Thus integrating abundance across the entire water column likely obscures important features in ecotype dynamics. To get a more detailed understanding of the temporal and spatial dynamics of the ecotypes, we analyzed integrated abundance in three sections of the euphotic zone (0–60 m, 60–120 m and 120–200 m).

The HL-adapted ecotype eMIT9312 displayed similar abundance patterns in the upper (0–60 m) and middle (60–120 m) euphotic zone at BATS, but below 120 m the seasonal cycle was out of phase with surface cycles by several months (Figure 4a; Figure 5a). Abundance in the lower euphotic was positively correlated with mixed layer depth (Spearman $R = 0.5$; $P < 0.05$), and abundance peaks occurred simultaneously with annual deep mixing events (Figure 4a). This suggests that it is the transport of surface populations below 120 m, and not *in situ* growth that may be responsible for most of the annual variations in abundance of eMIT9312 in the lower euphotic zone at BATS.

As with eMIT9312, patterns of eMED4 abundance in the upper and middle euphotic zone also differed from those in the lower zone at BATS (Figure 5b), with strong spikes in abundance below 120 m

Table 2 Correlation and cross-correlation between integrated ecotype abundance (0–200 m) and mixed layer depth at BATS

Ecotype	Correlation coeff. (lag in months)	Max. cross-correlation coeff. (lag in months)
eMIT9312	−0.06 (0)	0.61* (8)
eMED4	−0.62* (0)	0.58* (5)
eNATL	−0.54* (0)	0.65* (4)
eSS120	−0.60* (0)	0.66* (8)
eMIT9313	−0.46* (0)	0.61* (8)

The coefficient of cross-correlation represents the relationship between abundance at 1 month and the mixed layer depth from previous months. The maximum positive cross-correlation is reported along with the delay in months (lag) between abundance and mixed layer depth when correlation is greatest. *Indicates $P < 0.05$.

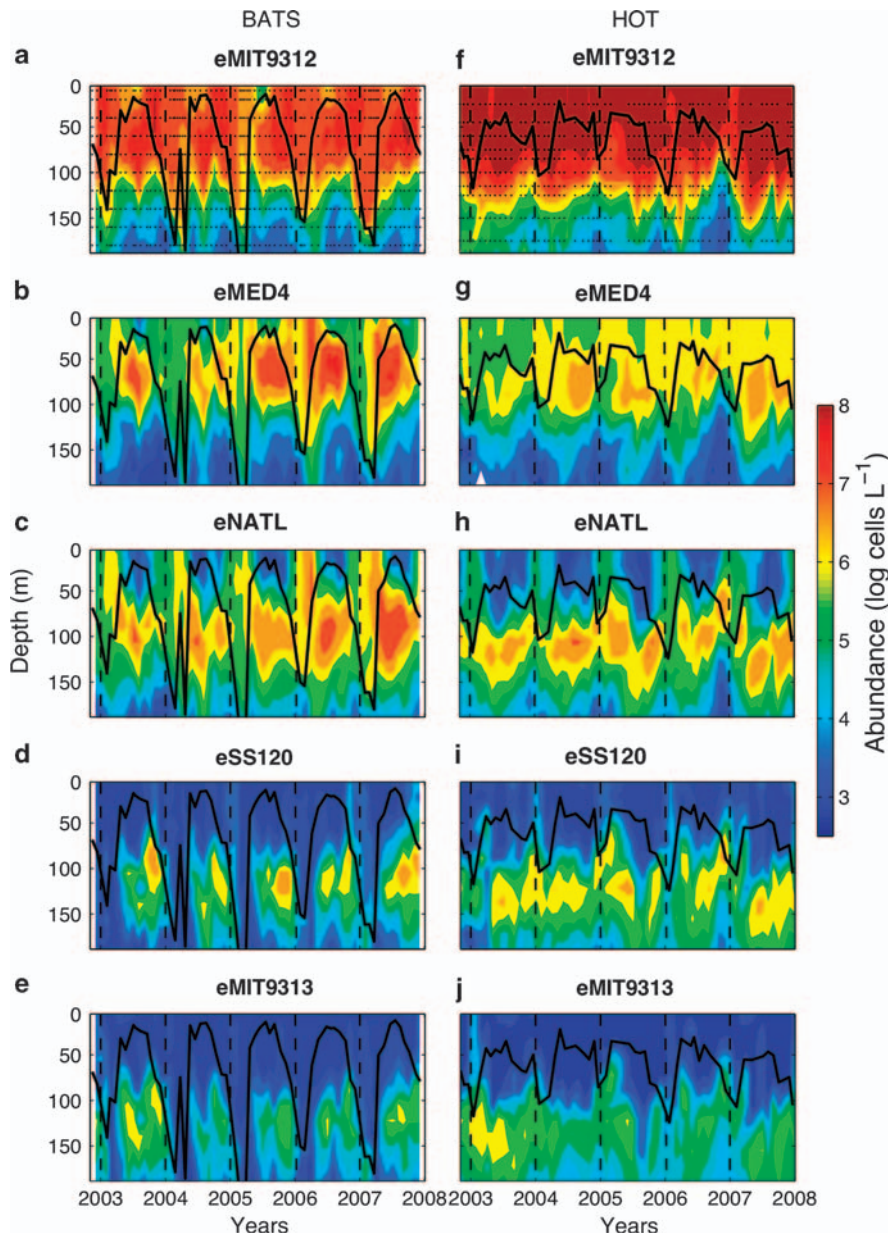


Figure 4 Ecotype abundance at BATS (a–e) and HOT (f–j) from 2003 to 2008. Sampling depths at HOT and BATS are indicated by the points overlaying (a and f). The solid lines indicate the mixed layer depth.

accompanying deep mixing (Figure 4b). But while the eMED4 and eMIT9312 dynamics were synchronized below 120 m, these HL-adapted ecotypes were out of synchronization by several months in the upper 120 m (Figures 5a and b). That is, each year at BATS, eMED4 typically reached peak abundance in July–August, whereas eMIT9312 reached maximum abundance in October–November (Supplementary Figure 7). At HOT, in contrast, these two HL ecotypes did not display these offset repeating patterns.

We hypothesize that different temperature sensitivities of eMED4 and eMIT9312 explain, at least in part, the differences in their temporal dynamics at BATS. That is, strains belonging to the eMED4 clade

have lower temperature optima than those belonging to the eMIT9312 clade (Johnson *et al.*, 2006; Zinser *et al.*, 2007). If this differential trait is universal among cells belonging to the eMED4 and eMIT9313 clades, then this would allow eMED4 cells to accumulate during the first half of the year when temperatures were low, whereas the higher temperature optimum of eMIT9312 would limit their accumulation until later in the season when temperatures were high—as was observed at BATS. In fact, the abundance of eMED4 in the upper 60 m was negatively correlated with temperature when light levels were taken into account (partial correlation coeff. -0.43 ; $P < 0.05$), whereas eMIT9312 abundance was positively correlated with temperature

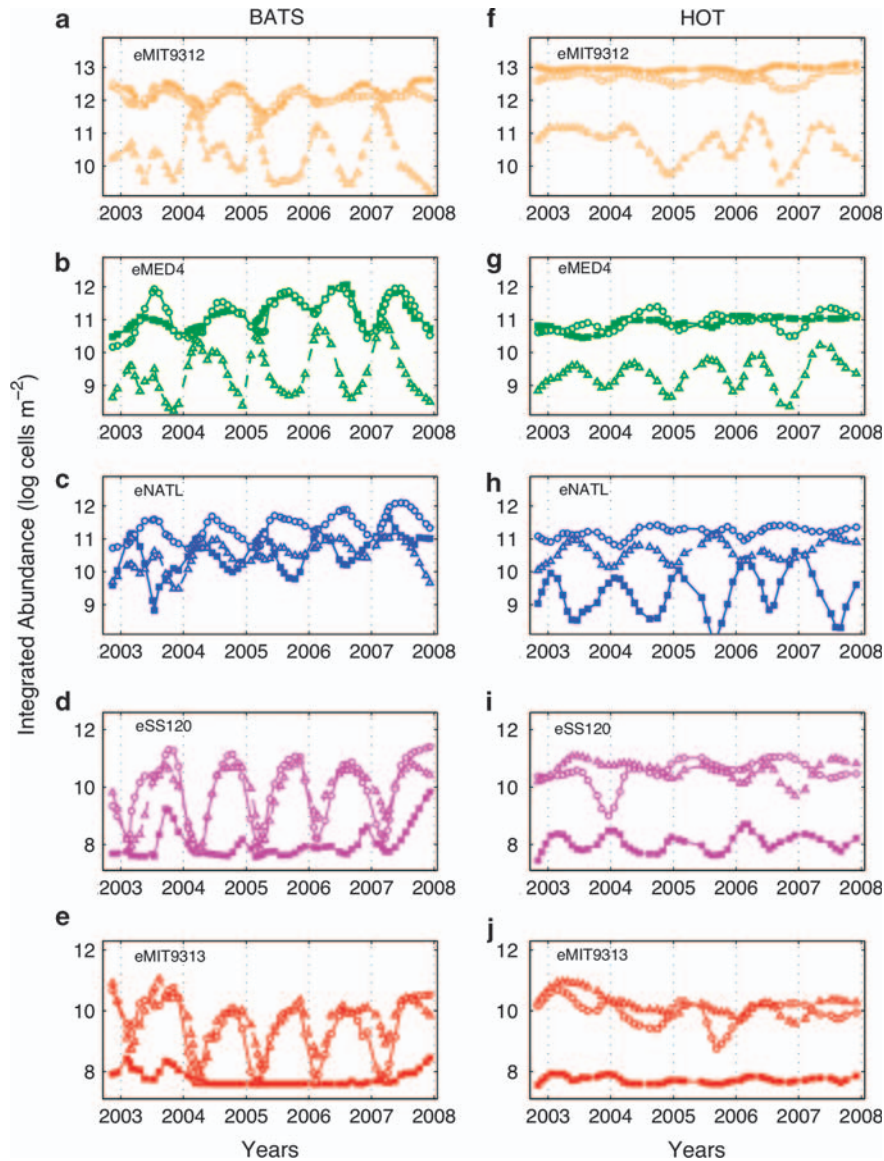


Figure 5 Integrated ecotype abundance in different regions of the euphotic zone at BATS (a–e) and HOT (f–j). The regions from 0–60 m, 60–120 m and 120 m are identified by filled squares, open circles and open triangles, respectively. Abundance was smoothed using locally weighted regression to eliminate low frequency variation.

(partial correlation coeff. 0.30; $P < 0.05$). The potential influence of temperature on the temporal distribution of HL ecotypes at BATS is analogous to its inferred influence on their geographic distribution: eMED4 dominates in cooler, higher latitude waters, and eMIT9312 dominates warmer, lower latitude waters (Johnson *et al.*, 2006; Zwirgmaier *et al.*, 2007).

Abundance patterns of eNATL also differed among regions of the euphotic zone, and the similarity of these patterns between BATS and HOT was striking. At both locations, abundance in the upper 60 m typically peaked 3–4 months earlier than at 60–120 m and 120–200 m (Figures 4c and h; Figures 5c and h). Integrated abundance in the upper 60 m reached maximum levels during deep

mixing events, and decreased as the mixed layer depth shoaled (Figures 4c and h). This suggests that annual peaks in abundance in the upper 60 m were due, at least in part, to vertical transport of deeper cells to surface waters via mixing. While it remains unclear if members of the LL-adapted eNATL clade were able to grow at high light levels found in the upper euphotic zone, the net accumulation eNATL cells throughout the water column during periods of deep mixing confirmed that the eNATL clade was able to at least tolerate temporary exposure to high irradiance.

In contrast to eNATL, the abundance of the other two LL-adapted ecotypes, eSS120 and eMIT9313, did not increase dramatically in the upper 60 m during periods of deep mixing. In fact, their

abundance was typically near or below detection levels throughout the entire water column at BATS when the well-mixed layer spanned the euphotic zone (Figures 4d and e). It appears that, unlike eNATL, members of the eSS120 and eMIT9313 ecotypes could not tolerate temporary high light exposure when transported to surface waters via mixing. At HOT, however, the mixed layer depth rarely exceeded 100 m, and therefore a substantial fraction of eSS120 and eMIT9313 cells were not transported to surface waters during winter mixing events. Thus, it appears that eSS120 and eMIT9313 maintained relatively stable abundance levels throughout the time series, in comparison with BATS, as they were not subjected to inhibitory, or lethal, high light exposure during mixing events.

Response to light shock in cultured isolates

To directly test the hypothesis that members of the eNATL clade can withstand temporary exposure to high intensity light better than members of the other LL-ecotypes, we performed light shock experiments with cultured isolates from the eMED4, eNATL and eSS120 clades. Strains acclimated to an irradiance of $35 \mu\text{E m}^{-2} \text{s}^{-1}$ were exposed to light levels of $400 \mu\text{E m}^{-2} \text{s}^{-1}$ for 4 h and then returned back to $35 \mu\text{E m}^{-2} \text{s}^{-1}$. For perspective, these light levels are equivalent to those that would be found at 90 m and 36 m at mid-day, assuming a surface irradiance of $2000 \mu\text{E m}^{-2} \text{s}^{-1}$ and an attenuation coefficient of -0.045 m^{-1} , which are typical of HOT and BATS. Furthermore, $400 \mu\text{E m}^{-2} \text{s}^{-1}$ is a lethal intensity for NATL2a and SS120, but not MED4, when applied continuously (Moore and Chisholm, 1999).

In vivo chlorophyll fluorescence of all three cultures dropped significantly relative to the control after the light shock. Fluorescence in MED4 and NATL2a cultures rebounded within 24 h and continued to increase over the next 48 h (Figure 6), whereas fluorescence of the SS120 culture contin-

ued to decline for the following 72 h. Cell concentrations of MED4 and NATL2a also rose following the light shock, whereas the number of SS120 cells declined throughout the recovery period (Supplementary Figure 8). These results indicate that NATL2a can withstand temporary light shock better than SS120, which is consistent with the temporal and spatial distribution patterns of their respective ecotypes at HOT and BATS described above. These results are also consistent with a similar light shock experiment comparing isolates SS120 and PCC 9511, a member of the HL-adapted eMED4 ecotype. In this study, the capacity to repair damage to photosystem II caused by light shock was much greater in PCC 9511 than in SS120 (Six *et al.*, 2009). Although it was not tested specifically, it is plausible that NATL2a is also better able to reverse photoinactivation of photosystem II than SS120. If so, then this would present a possible mechanism to explain, at least in part, the differences in temporary light shock tolerance and the subsequent environmental distributions of LL ecotypes.

Analysis of the genomes of cultured isolates provides some insight into the genes that may be responsible for differences in light physiology among LL-adapted ecotypes. For example, the two sequenced eNATL isolates each have 41 genes encoding high light-inducible proteins (HLIP), whereas the other LL-adapted ecotypes only have 9–13 HLIP genes (Coleman and Chisholm, 2007). This protein family, also called small cab-like proteins, aids in high-light survival and photoacclimation in *Synechocystis* (He *et al.*, 2001; Havaux *et al.*, 2003), and may also have a role in light-shock tolerance in eNATL. The mechanism by which they do this remains unclear, but high light-inducible proteins are thought to physically associate with either photosystem and allow it to shed excess energy as heat and thereby reduce photoinactivation (Promnares *et al.*, 2006; Yao *et al.*, 2007). As a photosynthetic organism, reducing and reversing the effects of photosystem inactivation is crucial for the survival of *Prochlorococcus* exposed to high light levels.

Differences in their ability to repair UV-damaged DNA may also partially explain differences in the environmental distributions of eNATL and other LL ecotypes. The six sequenced isolates from the two HL-adapted ecotypes and the two sequenced isolates from the eNATL ecotype contain genes that encode photolyase (Coleman and Chisholm, 2007; Kettler *et al.*, 2007), an enzyme helps repair UV-damaged DNA (Sancar, 2000). Photolyase genes are absent in the genomes of the other four LL-adapted isolates that have been sequenced. Although these other LL-adapted isolates do encode for an alternate putative UV-repair enzyme, pyrimidine dimer glycosylase (Goosen and Moolenaar, 2008; Partensky and Garczarek, 2010), its function in *Prochlorococcus* has not been confirmed. This enzyme also has a modification of the Arg-26

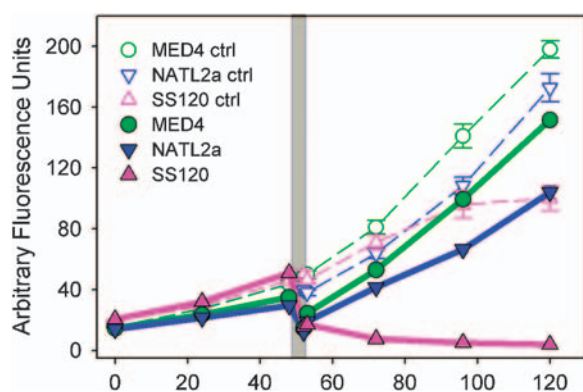


Figure 6 Response of *Prochlorococcus* strains MED4, NATL2a and SS120 to light-shock. *In vivo* chlorophyll fluorescence of duplicate cultures acclimated to $35 \mu\text{E m}^{-2} \text{s}^{-1}$, exposed to $400 \mu\text{E m}^{-2} \text{s}^{-1}$ for 4 hours (gray area), and returned to $35 \mu\text{E m}^{-2} \text{s}^{-1}$. Controls did not experience light shock. Error bars represent one s.d.

residue that is known to dramatically reduce activity in homologs (Doi *et al.*, 1992; Goosen and Moolenaar, 2008), suggesting a diminished ability to repair UV-damaged DNA in eSS120 and eMIT9313 ecotypes. Indeed, the high-light adapted strain MED4, which encodes photolyase, has a greater tolerance to UV exposure than low-light adapted strain MIT9313 (Osburne *et al.*, 2010), which lacks photolyase but encodes pyrimidine dimmer glycosylase. Thus, the protection from UV exposure provided by photolyase may explain, at least in part, why the eNATL clade can better survive transport to UV-rich surface waters.

Conclusions

Clear patterns in the temporal and spatial distribution of *Prochlorococcus* ecotypes emerge from this study. For example, ecotype abundance follows a strong annual pattern at BATS, whereas ecotype abundance has only a weak annual pattern at HOT. In addition, ecotypes at BATS follow an annual succession pattern, but a similar pattern is not observed at HOT. These patterns are consistent with what we have learned from physiological assays on cultured isolates with regards to the light optima, temperature optima, and light-shock tolerance of different ecotypes. That is, the distinct distribution patterns at both HOT and BATS can be explained, at least in general terms, by the consistent and predictable responses of ecotypes to changes in the light, temperature and mixing at each location.

Analyses of inorganic nutrient concentrations and ecotype abundance were not possible as nutrient levels were below detection in the upper 100 m at BATS throughout most of this study. However, while nutrients undoubtedly influence growth rates and standing stocks of *Prochlorococcus*, their influence on specific ecotypes may not be apparent, as evidenced by the fact that general patterns can be explained without them. Indeed, recent work suggests that nitrate concentrations might impact the composition of *Prochlorococcus* populations at finer phylogenetic levels than the ecotypes examined in our study (Martiny *et al.*, 2009b). Exploring how, and at which temporal, spatial and phylogenetic scales, the chemical and biological environment influence *Prochlorococcus* abundance and diversity presents a future challenge.

Results from this study help set the stage for coupling patterns in temporal and spatial dynamics of ecotypes with insights into their metabolic potential. The next step is to understand how *Prochlorococcus* ecotypes, or even sub-groups within these ecotypes, might differ in terms of nutrient usage, dissolved organic matter production and consumption, and other metabolic processes. This understanding will come from additional studies involving strain isolation, metagenomic comparisons, and large-scale single-cell genomics.

Acknowledgements

We thank Michael Lomas and the BATS team for sample collection at Bermuda, and David Karl, Matthew Church, and the HOT team for sample collection at Hawaii. We also thank Angel White and Ricardo Letelier for assistance with deriving solar flux data and attenuation coefficients. Norm Nelson and David Court generously provided light data from the Bermuda Bio Optics Program. We thank Daniele Veneziano for advice on statistical analysis. This work was funded by grants from the National Science Foundation, NSF STC Center for Microbial Oceanography: Research and Education, and the Gordon and Betty Moore Foundation.

References

- Ahlgren NA, Rocap G, Chisholm SW. (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.
- Bertilsson S, Berglund O, Pullin MJ, Chisholm SW. (2005). Release of dissolved organic matter by *Prochlorococcus*. *Vie Et Milieu-Life and Environment* **55**: 225–231.
- Bibby TS, Gorbunov MY, Wyman KW, Falkowski PG. (2008). Photosynthetic community responses to upwelling in mesoscale eddies in the subtropical north atlantic and pacific oceans. *Deep-Sea Res Part II-Top Stud Oceanogr* **55**: 1310–1320.
- Bouman HA, Ulloa O, Scanlan DJ, Zwirgmaier K, Li WK, Platt T *et al.* (2006). Oceanographic basis of the global surface distribution of *Prochlorococcus* ecotypes. *Science* **312**: 918–921.
- Campbell L, Liu HB, Nolla HA, Vaulot D. (1997). Annual variability of phytoplankton and bacteria in the subtropical north pacific ocean at station ALOHA during the 1991–1994 ENSO event. *Deep-Sea Res Part I-Oceanogr Res Pap* **44**: 167.
- Campbell L, Nolla HA, Vaulot D. (1994). The importance of *prochlorococcus* to community structure in the central north pacific-ocean. *Limnol Oceanogr* **39**: 954–961.
- Carlson CA, Morris R, Parsons R, Treusch AH, Giovannoni SJ, Vergin K. (2009). Seasonal dynamics of SAR11 populations in the euphotic and mesopelagic zones of the northwestern Sargasso Sea. *ISME J* **3**: 283–295.
- Cavender-Bares KK, Karl DM, Chisholm SW. (2001). Nutrient gradients in the western north atlantic ocean: relationship to microbial community structure and comparison to patterns in the pacific ocean. *Deep-Sea Res Part I-Oceanogr Res Pap* **48**: 2373–2395.
- Clausen J, Keck DD, Hiesey WM. (1940). Experimental studies on the nature of species. I. Effects of varied environments on western North American plants. *Carnegie Institute of Washington* **520**.
- Cleveland WS. (1979). Robust locally weighted regression and smoothing scatterplots. *J Am Stat Assoc* **74**: 829–836.
- Coleman ML, Chisholm SW. (2007). Code and context: *Prochlorococcus* as a model for cross-scale biology. *Trends Microbiol* **15**: 398–407.
- Coleman ML, Sullivan MB, Martiny AC, Steglich C, Barry K, Delong EF *et al.* (2006). Genomic islands

- and the ecology and evolution of *Prochlorococcus*. *Science* **311**: 1768–1770.
- Cohan FM. (2001). Bacterial species and speciation. *Systematic Biol* **50**: 513–524.
- Cohan FM, Perry EB. (2007). A systematics for discovering the fundamental units of bacterial diversity. *Curr Biol* **17**: R373–R386.
- Corno G, Karl DM, Church MJ, Letelier RM, Lukas R, Bidigare RR *et al.* (2007). Impact of climate forcing on ecosystem processes in the North Pacific Subtropical Gyre. *J Geophys Res* **112**: C04021, doi:10.1029/2006JC003730.
- Doi T, Recktenwald A, Karaki Y, Kikuchi M, Morikawa K, Ikehara M *et al.* (1992). The role of the basic amino acid cluster and Glu-23 in pyrimidine dimer glycosylase activity of T4-endonuclease-V. *Proc Natl Acad Sci USA* **89**: 9420–9424.
- DeLong EF, Preston CM, Mincer T, Rich V, Hallam SJ, Frigaard NU *et al.* (2006). Community genomics among stratified microbial assemblages in the ocean's interior. *Science* **311**: 496–503.
- DuRand MD, Olson RJ, Chisholm SW. (2001). Phytoplankton population dynamics at the bermuda atlantic time-series station in the sargasso sea. *Deep-Sea Res Part II-Top Stud Oceanogr* **48**: 1983–2003.
- Frasier C, Alm EJ, Polz MF, Spratt BG, Hanage WP. (2009). The bacterial species challenge: making sense of genetic and ecological diversity. *Science* **323**: 741–746.
- Follows MJ, Dutkiewicz S, Grant S, Chisholm SW. (2007). Emergent biogeography of microbial communities in a model ocean. *Science* **315**: 1843–1846.
- Fuhrman JA, Hewson I, Schwalbach MS, Steele JA, Brown MV, Naeem S. (2006). Annually reoccurring bacterial communities are predictable from ocean conditions. *Proc Natl Acad Sci USA* **103**: 13104–13109.
- Goosen N, Moolenaar GF. (2008). Repair of UV damage in bacteria. *DNA Repair* **7**: 353–379.
- Havaux M, Guedeney G, He QF, Grossman AR. (2003). Elimination of high-light-inducible polypeptides related to eukaryotic chlorophyll a/b-binding proteins results in aberrant photoacclimation in *Synechocystis* PCC6803. *Biochimica Et Biophysica Acta-Bioenergetics* **1557**: 21–33.
- He QF, Dolganov N, Bjorkman O, Grossman AR. (2001). The high light-inducible polypeptides in *Synechocystis* PCC6803 - Expression and function in high light. *J Biol Chem* **276**: 306–314.
- 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.
- Karl DM, Letelier R, Hebel D, Tupas L, Dore J, Christian J *et al.* (1995). Ecosystem changes in the north pacific subtropical gyre attributed to the 1991–92 El-Nino. *Nature* **373**: 230–234.
- Karl DM, Lukas R. (1996). The hawaii ocean time-series (HOT) program: background, rationale and field implementation. *Deep-Sea Res Part II-Top Stud Oceanogr* **43**: 129–156.
- 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 Genetics* **3**: 2515–2528.
- Legendre P, Legendre L (eds). (1998). Ecological data series. In: *Numerical Ecology: Second English Addition*. Elsevier Science BV: Amsterdam, pp 637–691.
- Letelier RM, Karl DM, Abbott MR, Flament P, Freilich M, Lukas R *et al.* (2000). Role of late winter mesoscale events in the biogeochemical variability of the upper water column of the north pacific subtropical gyre. *J Geophysical Res-Oceans* **105**: 28723–28739.
- Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar Buchner A *et al.* (2004). ARB: a software environment for sequence data. *Nucleic Acids Res* **32**: 1363–1371.
- Malmstrom RR, Cottrell MT, Elifantz H, Kirchman DL. (2005). Biomass production and dissolved organic matter assimilation by SAR11 bacteria in the northwest atlantic ocean. *Appl Environ Microbiol* **71**: 2979–2986.
- Malmstrom RR, Kiene RP, Cottrell MT, Kirchman DL. (2004). Contribution of SAR11 bacteria to dissolved dimethylsulfoniopropionate and amino acid uptake in the North Atlantic Ocean. *Appl Environ Microbiol* **70**: 4129–4135.
- Martiny AC, Coleman ML, Chisholm SW. (2006). Phosphate acquisition genes in *Prochlorococcus* ecotypes: Evidence for genome-wide adaptation. *Proc Natl Acad Sci USA* **103**: 12552–12557.
- Martiny AC, Kathuria S, Berube PM. (2009a). Widespread metabolic potential for nitrite and nitrate assimilation among *Prochlorococcus* ecotypes. *Proc Natl Acad Sci USA* **106**: 10787–10792.
- Martiny AC, Tai APK, Veneziano D, Primeau F, Chisholm SW. (2009b). Taxonomic resolution, ecotypes and the biogeography of *Prochlorococcus*. *Environ Microbiol* **11**: 823–832.
- Michelou VK, Cottrell MT, Kirchman DL. (2007). Light-stimulated bacterial production and amino acid assimilation by cyanobacteria and other microbes in the north atlantic ocean. *Appl Environ Microbiol* **73**: 5539–5546.
- Moore LR, Chisholm SW. (1999). Photophysiology of the marine cyanobacterium *Prochlorococcus*: ecotypic differences among cultured isolates. *Limnol Oceanogr* **44**: 628–638.
- Moore LR, Coe A, Zinser ER, Saito MA, Sullivan MB, Lindell D *et al.* (2007). Culturing the marine cyanobacterium *Prochlorococcus*. *Limnol Oceanogr* **5**: 353–362.
- 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.
- 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, Rocap G, Chisholm SW. (1998). Physiology and molecular phylogeny of coexisting *Prochlorococcus* ecotypes. *Nature* **393**: 464–467.
- Morris RM, Rappé MS, Connon SA, Vergin KL, Siebold WA, Carlson CA *et al.* (2002). SAR11 clade dominates ocean surface bacterioplankton communities. *Nature* **420**: 806–810.
- Olson RJ, Chisholm SW, Zettler ER, Altabet MA, Dusenberry JA. (1990a). Spatial and temporal distributions of prochlorophyte picoplankton in the north-atlantic ocean. *Deep-Sea Res Part a-Oceanographic Res Papers* **37**: 1033–1051.
- Olson RJ, Chisholm SW, Zettler ER, Armbrust EV. (1990b). Pigments, size, and distribution of *synechococcus*

- in the north-atlantic and pacific oceans. *Limnol Oceanogr* **35**: 45–58.
- Osburne MS, Holmbeck BM, Frias-Lopez J, Steen R, Huang K, Kelly L *et al.* (2010). UV hyper-resistance in *Prochlorococcus* MED4 results from a single base pair deletion just upstream of an operon encoding nudix hydrolase and photolyase. *Environ Microbiol* Published Online: 23 March 2010; DOI: 10.1111/j.1462-2920.2010.02203.x.
- Partensky F, Garczarek L. (2010). Prochlorococcus: advantages and limits of minimalism. *Annu Rev Marin Sci* Vol. 2 305–331.
- Partensky F, Hess WR, Vaulot D. (1999). Prochlorococcus: a marine photosynthetic prokaryote of global significance. *Microbiol Mol Biol Rev* **63**: 106.
- Promnares K, Komenda J, Bumba L, Nebesarova J, Vacha F, Tichy M. (2006). Cyanobacterial small chlorophyll-binding protein ScpD (HliB) is located on the periphery of photosystem II in the vicinity of PsbH and CP47 subunits. *J Biol Chem* **281**: 32705–32713.
- Ramirez-Flandes S, Ulloa O. (2008). Bosque: integrated phylogenetic analysis software. *Bioinformatics* **24**: 2539–2541.
- Rocap G, Distel DL, Waterbury JB, Chisholm SW. (2002). Resolution of *Prochlorococcus* and *Synechococcus* ecotypes by using 16S-23S ribosomal DNA internal transcribed spacer sequences. *Appl Environ Microbiol* **68**: 1180–1191.
- Rocap G, Larimer FW, Lamerdin J, Malfatti S, Chain P, Ahlgren NA *et al.* (2003). Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. *Nature* **424**: 1042–1047.
- Sancar GB. (2000). Enzymatic photoreactivation: 50 years and counting. *Mutation Res-Fundamental Mol Mechanisms Mutagenesis* **451**: 25–37.
- Six C, Finkel ZV, Irwin AJ, Campbell DA. (2009). Light variability illuminates niche-partitioning among marine picocyanobacteria. *PLoS ONE* **2**: e1341.
- Stackebrandt E, Goebel BM. (1994). Taxonomic note: a place for DNA:DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int J Syst Bacteriol* **44**: 846–849.
- Steinberg DK, Carlson CA, Bates NR, Johnson RJ, Michaels AF, Knap AH. (2001). Overview of the US JGOFS Bermuda atlantic time-series study (BATS): a decade-scale look at ocean biology and biogeochemistry. *Deep-Sea Res Part II-Top Stud Oceanogr* **48**: 1405–1447.
- Treusch AH, Vergin KL, Finlay LA, Donatz MG, Burton RM, Carlson CA *et al.* (2009). Seasonality and vertical structure of microbial communities in an ocean gyre. *ISME J* **3**: 1148–1163.
- Tureson G. (1922). Species and the variety as ecological units. *Hereditas* **3**: 100–113.
- Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA *et al.* (2004). Environmental genome shotgun sequencing of the sargasso sea. *Science* **304**: 66–74.
- Vila-Costa M, Simo R, Harada H, Gasol JM, Slezak D, Kiene RP. (2006). Dimethylsulfoniopropionate uptake by marine phytoplankton. *Science* **314**: 652–654.
- Ward DM, Bateson MM, Ferris MJ, Kühl M, Wieland A, Koeppel A *et al.* (2006). Cyanobacterial ecotypes in the microbial mat community of mushroom spring (yellowstone national park, wyoming) as species-like units linking microbial community composition, structure and function. *Philos Trans R Soc London (Biol)* **361**: 1997–2008.
- West NJ, Scanlan DJ. (1999). Niche-partitioning of *Prochlorococcus* populations in a stratified water column in the eastern North Atlantic Ocean. *Appl Environ Microbiol* **65**: 2585–2591.
- West NJ, Schönhuber WA, Fuller NJ, Amann RI, Rippka R, Post AF *et al.* (2001). Closely related *Prochlorococcus* genotypes show remarkably different depth distributions in two oceanic regions as revealed by *in situ* hybridization using 16S rRNA-targeted oligonucleotides. *Microbiology* **147**: 1731–1744.
- White AE, Spitz YH, Letelier RM. (2007). What factors are driving summer phytoplankton blooms in the North Pacific Subtropical Gyre? *J Geophys Res* **112**: C12006, doi:10.1029/2007JC004129.
- Wu JF, Sunda W, Boyle EA, Karl DM. (2000). Phosphate depletion in the western north atlantic ocean. *Science* **289**: 759–762.
- Yao D, Kieselbach T, Komenda J, Promnares K, Prieto MA, Tichy M *et al.* (2007). Localization of the small CAB-like proteins in photosystem II. *J Biol Chem* **282**: 267–276.
- Zinser ER, Coe A, Johnson ZI, Martiny AC, Fuller NJ, Scanlan DJ *et al.* (2006). Prochlorococcus ecotype abundances in the North Atlantic Ocean as revealed by an improved quantitative PCR method. *Appl Environ Microbiol* **72**: 723–732.
- 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.
- Zubkov MV, Fuchs BM, Tarran GA, Burkill PH, Amann R. (2003). High rate of uptake of organic nitrogen compounds by *Prochlorococcus* cyanobacteria as a key to their dominance in oligotrophic oceanic waters. *Appl Environ Microbiol* **69**: 1299–1304.
- Zwirgmaier K, Heywood JL, Chamberlain K, Woodward EMS, Zubkov MV, Scanlan DJ. (2007). Basin-scale distribution patterns of picocyanobacterial lineages in the Atlantic Ocean. *Environ Microbiol* **9**: 1278–1290.

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