npg

ORIGINAL ARTICLE

Seasonal dynamics of SAR11 populations in the euphotic and mesopelagic zones of the northwestern Sargasso Sea

Craig A Carlson[1,5], Robert Morris[1,2,5], Rachel Parsons[3,5], Alexander H Treusch[4,5], Stephen J Giovannoni[4] and Kevin Vergin[4]

[1]Department of Ecology, Evolution, and Marine Biology, University of California, Santa Barbara, CA, USA; [2]Department of Oceanography, University of Washington, Seattle, WA, USA; [3]Bermuda Institute of Ocean Sciences, Ferry Reach, St George's GE01, Bermuda and [4]Department of Microbiology, Oregon State University, Corvallis, OR, USA

Bacterioplankton belonging to the SAR11 clade of α-proteobacteria were counted by fluorescence *in situ* hybridization (FISH) over eight depths in the surface 300 m at the Bermuda Atlantic Time-series Study (BATS) site from 2003 to 2005. SAR11 are dominant heterotrophs in oligotrophic systems; thus, resolving their temporal dynamics can provide important insights to the cycling of organic and inorganic nutrients. This quantitative time-series data revealed distinct annual distribution patterns of SAR11 abundance in the euphotic (0–120) and upper mesopelagic (160–300 m) zones that were reproducibly correlated with seasonal mixing and stratification of the water column. Terminal restriction fragment length polymorphism (T-RFLP) data generated from a decade of samples collected at BATS were combined with the FISH data to model the annual dynamics of SAR11 subclade populations. 16S rRNA gene clone libraries were constructed to verify the correlation of the T-RFLP data with SAR11 clade structure. Clear vertical and temporal transitions were observed in the dominance of three SAR11 ecotypes. The mechanisms that lead to shifts between the different SAR11 populations are not well understood, but are probably a consequence of finely tuned physiological adaptations that partition the populations along physical and chemical gradients in the ecosystem. The correlation between evolutionary descent and temporal/spatial patterns we describe, confirmed that a minimum of three SAR11 ecotypes occupy the Sargasso Sea surface layer, and revealed new details of their population dynamics.

Introduction

Over two decades ago, small subunit rRNA gene clones recovered from the surface waters of the Sargasso Sea revealed the presence of SAR11, a diverse clade of heterotrophic α-proteobacteria (Giovannoni *et al.*, 1990). Since then SAR11 genes have been found ubiquitously distributed throughout the World's oceans. They represent approximately one quarter of all rRNA genes identified in clone libraries from marine environments (Rappé and Giovannoni, 2003). The prevalence of SAR11

cells in oceanic systems has been confirmed quantitatively by fluorescence *in situ* hybridization (FISH) cell counts and account for approximately one-third of the prokaryotic cells in the surface waters of the northwestern Sargasso Sea (Morris *et al.*, 2002).

Field *et al.* (1997) were the first to identify ecotypes of the SAR11 clade using 16S rRNA gene sequencing and RNA hybridization probes. The ecotypes had highly depth-specific distributions. One SAR11 ecotype, later identified as subclade II, was more abundant in the upper mesopelagic, whereas the original SAR11 ecotype (subclade Ia) was more prevalent in the euphotic zone. They concluded that the differing depth profiles were indicative of niche partitioning. Later work with additional sequences described the structure of the SAR11 clade in greater detail (Suzuki *et al.*, 2001; Rappé and Giovannoni, 2003). Morris *et al.* (2005) then used terminal restriction fragment length

polymorphisms (T-RFLP) to confirm earlier findings and to show that SAR11 subclade Ib was also found in surface waters.

The analogous discovery of ecotypes of *Prochlorococcus* (Moore *et al.*, 1998) was supported by genome sequences and physiological characteristics that helped to explain their differing vertical distributions. In particular, adaptations to high-light and low-nutrients were observed in the shallow ecotypes, whereas deeper ecotypes were suited to low-light, high-nutrient environments. Studies of ecotype distributions help provide oceanographers with an understanding of how phylogenetic microdiversity impacts estimates of productivity and nutrient cycling, and are being used to develop new models to explore relationships between climate change, community structure and biogeochemical cycles (Follows *et al.*, 2007; Hood *et al.*, 2006).

Resolving the dynamics of SAR11 ecotypes in the ocean surface is particularly important because the flux of organic carbon through bacterioplankton is comparable to daily primary production in this environment (Ducklow, 2000), where SAR11 are the dominant heterotrophs. Insights into the interactions between SAR11 populations and nutrient fields have emerged from interpretations of genomic data, proteomic data and the study of cultured isolates (Giovannoni *et al.*, 2005a). These studies have shown that the SAR11 cells have light-dependent proton pumps and unusual nutritional requirements for reduced sulfur compounds and glycine (Giovannoni *et al.*, 2005b; Tripp *et al.*, 2008a, b). They appear to have sacrificed nutritional versatility for a streamlined metabolism (Giovannoni *et al.*, 2005a). These findings were largely developed through study of cultured SAR11 isolates. However, little is known about how natural populations of SAR11 and its ecotypes vary through time and over depth in response to physical and chemical variability. Temporal variability of hydrography and biogeochemical processes is highly complex. Consequently, progress in understanding the correspondence between biogeochemical processes and microbial diversity is constrained by the practical problem of implementing experimental designs that sample with sufficiently fine scale to resolve processes (for example, Karl *et al.*, 2003).

The goal of this study was to quantitatively resolve temporal patterns in SAR11 cell densities throughout the euphotic and mesopelagic zone. Here, we present a 3-year time-series of SAR11 cell distributions, as determined by FISH, collected from the upper 300 m at the Bermuda Atlantic Time-series (BATS) site. We frame these dynamics in the context of physical and biogeochemical data from other long-term studies at the site. To resolve SAR11 ecotype population dynamics in greater detail, we used 16S rRNA gene sequence and T-RFLP data to measure the relative abundance of SAR11 subclades Ia, Ib and II, and used these data to model the ecotype population's dynamics. The data reveal many new features of SAR11 population dynamics in the ocean surface layer and show that they occur reproducibly each year.

Materials and methods

Study site and sample collection

Samples were collected aboard the R. V. Weatherbird II at the BATS site ($31° 40′$ N, $64° 10′$ W). All cruises were conducted as part of the larger BATS program and sampled at least monthly with biweekly sampling each year between January and March. This sampling strategy has been successful in revealing the major temporal microbial and biogeochemical patterns (Carlson *et al.*, 1996; Steinberg *et al.*, 2001; Morris *et al.*, 2005). A broader assessment of the BATS biogeochemical data is presented elsewhere (see Deep Sea Research II volumes 43 (2–3) and 48 (8–9) for relevant studies).

Samples for FISH were collected from eight depths over the surface 300 m from January 2003 to December 2005. Seawater was fixed with 0.2 μm filtered formalin (10% concentration) and stored in liquid nitrogen until processing (within 2 weeks of collection). No appreciable decay of probe signal was detected when stored in liquid nitrogen for periods up to 6 months.

High molecular weight DNA was collected regularly during BATS cruises (97 separate casts) from 1 and 200 m from August 1991 to January 1994, and September 1997 to January 2004. An additional 36 profiles over 6–8 depths were sampled periodically throughout this time-series (Table 1). All 341 DNA samples were collected through CTD rosette or *in situ* pumps (McLane Research Laboratories Inc., East Falmouth, MA, USA). Sixteen to seventy liters of seawater were filtered onto 0.2 μm Supor filters (Pall Life Sciences, East Hills, NY, USA) and extracted as described earlier (Giovannoni *et al.*, 1996).

Ancillary data used to provide the context for the annual patterns of mixed layer depth (MLD) and primary production are available from the public data set (http://www.bios.edu). The MLD during experimental periods was determined as the depth where potential density (sigma-*t*) of the water was equal to sea surface sigma-*t* plus an increment in sigma-*t* equivalent to a 0.2 °C temperature decrease (Sprintall and Tomczak, 1992). Contour plots were created in Ocean Data View (Schlitzer, R., http://www.awi-bremerhaven.de/GEO/ODV, 2004). Tukey–Kramer HSD statistical analyses were performed using the statistical package JMP (SAS Institute Inc., Cary, NC, USA).

FISH

Samples (10 ml) were filtered onto 0.2 μm 25 mm polycarbonate filters and stored desiccated at -20 °C until hybridization. Filters were cut into quarters

Table 1 Annual composite of all 341 DNA samples collected between 1991 and 2004 from which T-RFLP data were generated

Depth (m)	Month relative to timing of deep mixing												
	−1	Deep mixing	+1	+2	+3	+4	+5	+6	+7	+8	+9	+10
1	9	9	6	6	7	7	6	8	7	8	8	9
40	3	6	1	2	1	3		3	2	2	1	2
80	3	7	2	2	1	4		3	2	2	1	3
120	3	6	2	2	1	3	1	3	3	2	2	3
160	2	6	2	2	1	2	1	3	3	2	2	3
200	10	6	7	6	7	6	6	7	9	8	8	8
250	2	4	2	2	1	2	1	3	2	2	2	2
300	1	4	1	2	2	1	2	2	2	2	1	2

The values represent the number of samples collected for each month relative to the time of deep mixing at each depth throughout the DNA time-series. 1 m and 200 m samples were sampled at the highest frequency and other depths were sampled periodically. Blanks mean no samples exist for that month or depth.

and hybridization reactions were performed on one-quarter membrane sections as described in Morris et al. (2002). Briefly, four Cy3-labelled oligonucleotide probes targeting the SAR11 clade including SAR11 152R-Cy3 (5′ ATTAGCACAAGTTTCCY CGTGT), SAR11 441R-Cy3, (5′-TACAGTCATTTTCT TCCCCGAC); SAR11-542R (5′-TCCGAACTACGCTA GGTC) and SAR11-732R (GTCAGTAATGATCCAGA AAGYTG). A nonsense probe 338F- (TGAGGATGC CCTCCGTCG) served as a negative control. Forty microliters of hybridization solution containing $2 \, ng \, \mu l^{-1}$ of probe in a buffer of $900 \, mmol \, l^{-1}$ NaCl, 15% formamide, $20 \, mmol \, l^{-1}$ Tris-HCl (pH 7.4) and 0.01% sodium dodecyl sulfate (SDS) was spotted onto each quarter filter and incubated at $37 \, °C$ in a Techne hybridizer (HB-1D, NJ) for 16 h. Optimal hybridization stringency was achieved by washing the membranes in hybridization wash ($20 \, mmol \, l^{-1}$ Tris-HCl (pH 7.4), $150 \, mmol \, l^{-1}$ NaCl, $5 \, mmol \, l^{-1}$ EDTA and 0.01% SDS) for two 10-min intervals at experimentally determined temperatures of dissociation (SAR11, $55 \, °C$; 338F, $50 \, °C$) in Coplin jars. The hybridized filters were then air-dried and mounted on dry slides using a mounting media containing $1.67 \, \mu g \, ml^{-1}$ 0.6-diamidino-2-phenylindole dihydrochloride (DAPI) in Citiflour solution (Ted Pella Inc., Redding, CA, USA).

Microscopy and image analysis
Total prokaryotic cell abundance was determined through DAPI direct counting through epifluorescence microscopy (Porter and Feig, 1980). Ten to twelve fields of view encompassing $\geqslant 500$ cells per sample were counted from one of the filter quarters. Detection of CY3-positive cells and their ratio to DAPI-positive cells was aided by image analysis using an Olympus AX70 microscope ($\times 1000$) (Olympus, Japan) equipped with a Toshiba 3CCD video camera (IK-TU40A Toshiba, Japan), a computer assisted frame grabber and appropriate dichroic filters (Morris et al., 2002). Briefly, exposure times of 1 and 5 seconds were used for DAPI and CY3 images, respectively. CY3 images were segmented with Image Pro Plus software (Media Cybernetics, Bethesda, MD, USA) and overlaid onto corresponding segmented DAPI images. Objects with overlapping signals in both Cy3 and DAPI images were counted as probe positive. The negative control was determined similarly and subtracted from the positive probe counts to correct for autofluorescence and non-specific binding. The percent probe positive cells were determined for 300 to >1000 DAPI-positive cells. The ratio of probe-positive to DAPI-positive cells from the hybridized filter membrane was then multiplied by total DAPI count to derive SAR11 abundances (Morris et al., 2002). The coefficient of variability for positive probe enumeration was on the order of 15% in the euphotic zone where SAR11 comprised >30% of total DAPI counts and increased as SAR11 densities decreased over depth.

T-RFLP analysis of 16S rDNA
T-RFLP analysis was used for comparison of bacterial community composition between samples (Liu et al., 1997). Ribosomal RNA genes from mixed communities were amplified by PCR with Taq polymerase (Fermentas Inc., Glen Burnie, MD, USA) and the commonly used bacterial primers, 8F-FAM (5′-AGRGTTYGATYMTGGCTCAG-3′), and 519R (5′-GWATTACCGCGGCKGCTG- 3′) (Morris et al., 2005). The 8F-FAM primer was 5′ end-labeled with the phosphoramidite fluorochrome 5-carboxyfluorescein (6-FAM). Amplifications were performed with 5 ng of template DNA per $50 \, \mu l$ reaction using the following conditions: 30 cycles, denaturation at $94 \, °C$ for 15 s, annealing at $55 \, °C$ for 30 s and elongation at $72 \, °C$ for 1 min. Acetamide was added to a final concentration of 5% to reduce $G + C$ content-induced bias.

PCR products were restricted for 6 hours at $37 \, °C$ in reactions containing $10 \, mmol \, l^{-1}$ $MgCl_2$, and 10 U *Bsu*RI restriction enzyme (Fermentas, MD, USA). PCR and restriction digest were cleaned up according to QIAquick PCR purification kit instructions (Qiagen, Valencia, CA, USA). Terminal restriction

fragments (T-RFs) were resolved on an ABI 3100 Genetic Analyzer (Applied Biosystems, Forest City, CA, USA), and ABI Genescan software was used to size fragments based on MapMarker 1000 size standards (Bioventures, Murfeesboro, TN, USA). According to Kitts (2001) large and small T-RFLP fragments are more comparable and accurate when integrated peak areas are used compared with peak height. To be consistent with our earlier study (Morris et al., 2005) we used the peak area for the determination of percent contribution of the different fragments in all downstream analyses. T-RFLP data was standardized according to Osborn et al. (2006).

T-RFLP fragments were putatively identified using the sequences from our 16S rDNA clone libraries and publicly available sequences from marine samples. The T-RFs were predicted from the available sequence data using ARB (Ludwig et al., 2004) and TRF-CUT (Ricke et al., 2005). T-RF matches were verified by applying T-RFLP to identified clones (Morris et al., 2005). We focused specifically on T-RFLP fragments that putatively identified members of the SAR11 clade. One or more T-RFs within each electropherogram represented the predominant SAR11 subclades, Ia, Ib and II and were putatively designated as follows: SAR11 subclade Ia, 113 bp, SAR11 subclade Ib, 227, 228 and 293 bp and SAR11 subclade II 292 bp. The sum of all T-RF areas for each subclade was normalized by total SAR11 peak area to determine the relative contribution of each subclade. The phylogenetic affiliations assigned to the T-RFLP fragments in this study were the same as described in Morris et al. (2005) except that fragments 292 and 293, which coincided with subclade II and Ib, respectively, were 1 bp longer than reported by Morris et al. (2005). This slight discrepancy is likely due to subtle differences in base pair calling and standardization between the two studies.

Cloning and sequencing
Clone libraries were constructed from 40 m samples collected in the vicinity of the BATS site in March 2004 and July 2004. Ribosomal RNA genes were amplified from environmental DNA for cloning by PCR as described above with the exception that triplicate PCRs were pooled following amplification. A single band of the predicted length was excised and purified from a 1% agarose gel using a QIAquick PCR purification kit (Qiagen). Clone libraries were constructed using the pGEM-T-Easy (Promega Corporation, Madison, WI, USA) vector following the manufacturer's instructions. Transformations were submitted for plating and single-pass sequencing at the High-Throughput Genomics Unit (University of Washington, Seattle, WA, USA).

Partial 16S rRNA gene clone sequences were obtained and added to an aligned database containing greater than 350 SAR11 sequences maintained

with the ARB software package (Ludwig et al., 2004). Evolutionary distance analysis methods identified phylogenetic relationships within the SAR11 dataset, using the program PAUP* 4.0 beta 10 (Swofford, 2002). The sequence from 'Candidatus Pelagibacter ubique' (Giovannoni et al., 2005a) was used as a type of strain for alignment and fragment designation.

Nucleotide sequence accession numbers
Gene sequences were deposited in Genbank and given accession numbers EU984307 through EU984496.

Results

SAR11 abundance measurements revealed regular annual dynamics of cell density increase and redistribution in the upper 300 m that coincided with water column mixing and stratification (Figure 1a). SAR11 densities were significantly greater within the euphotic zone than in the upper mesopelagic zone. The general temporal trend for all years demonstrated a significant decrease in surface SAR11 densities upon onset of deep mixing in winter (Figure 1a). As the MLD began to shoal in
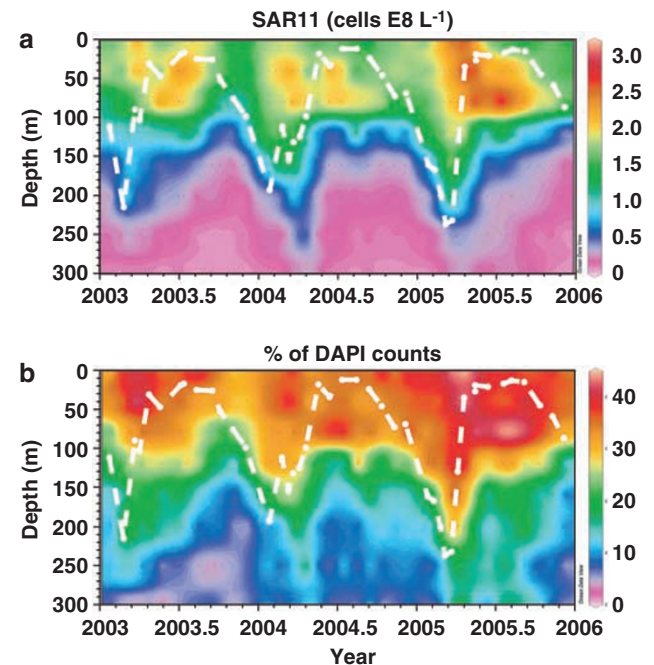
Figure 1 Contours of SAR11 cell densities (E8 cells L^{-1}) determined through FISH (**a**) and percentage of DAPI-positive cells identified as SAR11 (**b**) in the surface 300 m at BATS from January 2003 to December 2005. Mixed layer depth (MLD) estimate (dashed white line) was based on a variable sigma-*t* (potential density) criterion (Sprintall and Tomczak, 1992). MLD was determined as the depth where sigma-*t* is equal to sea surface sigma-*t* plus an increment in sigma-*t* equivalent to a 0.2 °C temperature decrease.

spring of each year, the SAR11 densities increased significantly (Student's t-test, P value < 0.0001) from winter lows to mean cell densities of 2.2×10^8 cells l^{-1}. During periods of summer stratification (MLD < 50 m) the SAR11 maximum was most pronounced just below the MLD between 40 and 80 m and reached cell densities as high as $2.5–3.25 \times 10^8$ cells l^{-1} each year. Within the upper mesopelagic (160–300 m) SAR11 cell densities increased significantly during and following the maximum MLD, and persisted at elevated densities for months after restratification (Figure 1a).

The contribution of SAR11 to total DAPI cell counts also varied through time and over depth. The percentage of DAPI-positive cells identified as SAR11 ranged from a mean of $33 \pm 8\%$ in the surface 120 m ($n = 176$) to $\sim 13 \pm 8\%$ within the upper mesopelagic zone ($n = 176$) (Figure 1b). In the mesopelagic zone the percent contribution of SAR11 cells to total DAPI counts increased to as high as 30% after deep mixing and then decreased to $< 10\%$ during summer-stratified periods.

*Annual variability of depth-integrated SAR11 stocks*
An annual linear harmonic model was fit to the temporal trend of the depth-integrated SAR11 data—it explained 44% of the variability with significance $> 99\%$ confidence interval (CI) throughout this time-series. Annual winter deep mixing extended the maximal MLD below the euphotic zone (0.1% light level ~ 140 m) to depths between 190 m to 240 m (Figure 1a). Deep mixing entrains nutrients into the surface waters stimulating spring phytoplankton blooms that reached a productivity maximum a month after maximal MLD (Figure 2a). The annual deep mixing and subsequent phytoplankton blooms initiate important annual organic and inorganic nutrient dynamics at this site (see Steinberg et al., 2001; Hansell and Carlson, 2001). The timing of deep mixing varied between early January and late February during this study. Therefore, to compare SAR11 dynamics in the context of the mixing event, we collapsed the 3-year time-series into a single 12-month composite. We aligned each year's data to the timing of the deep mixing event and determined the integrated SAR11 means for the corresponding months prior to and following deep mixing (Figure 2). Integrated SAR11 composite (0-300 m) revealed a significant two-fold increase in SAR11 stocks following deep mixing (Figure 2b).

Significant regular annual dynamics of SAR11 were observed within the euphotic zone (Figures 1 and 2c). The integrated SAR11 stocks (0–120 m) revealed significant accumulation rates of 0.78 E6 cells l^{-1} d^{-1} ($r^2 = 0.72$; $n = 5$), 0.99 E6 cells l^{-1} d^{-1} ($r^2 = 0.81$; $n = 6$) and 1.45 E6 cells l^{-1} d^{-1} ($r^2 = 0.68$; $n = 6$), in the springs of 2003, 2004 and 2005, respectively. The euphotic zone SAR11 maximum lagged the deep mixing event by 2 months
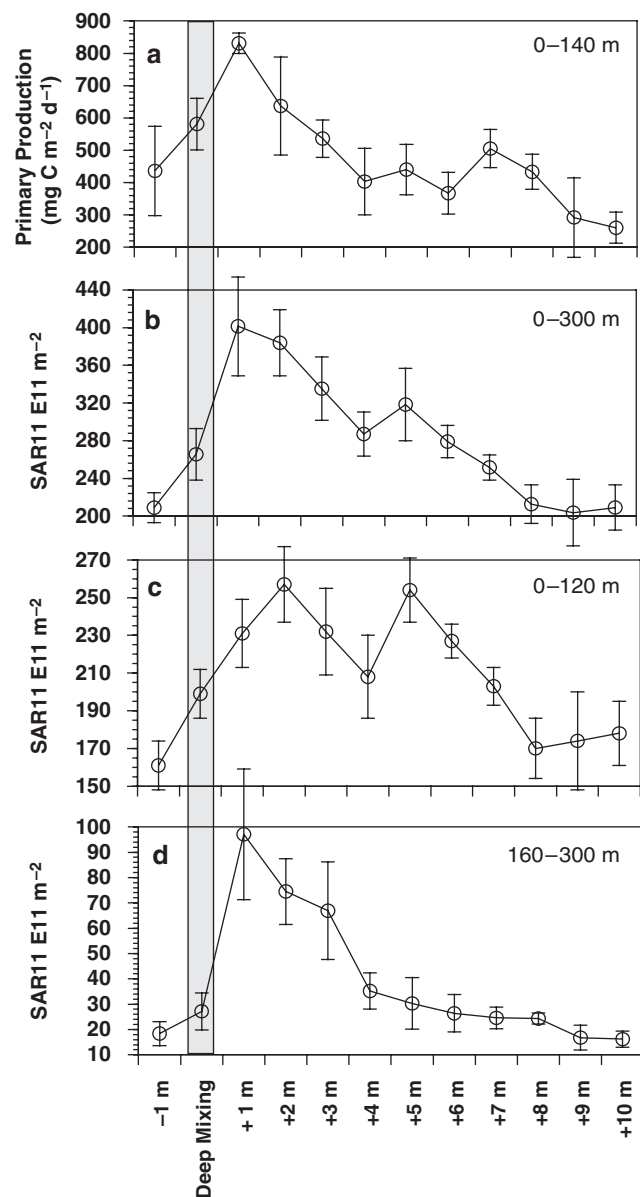
Figure 2 Annual composite of the depth-integrated means of (**a**) primary production (0–140 m) and SAR11 stocks for each month plotted relative to the deep convective mixing event. The various integration depth horizons for SAR11 include (**b**) the surface 300 m (**c**) the euphotic zone (0–120 m) (**d**) and the upper mesopelagic zone (160–300 m). The negative and positive values on the x axis refer to the month (s) prior to or following the deep convective mixing event. The gray bar is the reference and it represents the period of convective mixing for each year. The error bars represent s.e. of the 3-year mean for each corresponding month.

(Figure 2c). After the spring maximum, the euphotic zone SAR11 stocks decreased in the third and fourth months following deep mixing (Figure 2c). Each year after stratification, SAR11 stocks increased again, suggestive of a secondary summer maximum. However, the interannual variability of the summer maxima rendered this composite trend statistically insignificant (Figure 2c).

SAR11 ecotypes in the Sargasso Sea
CA Carlson et al
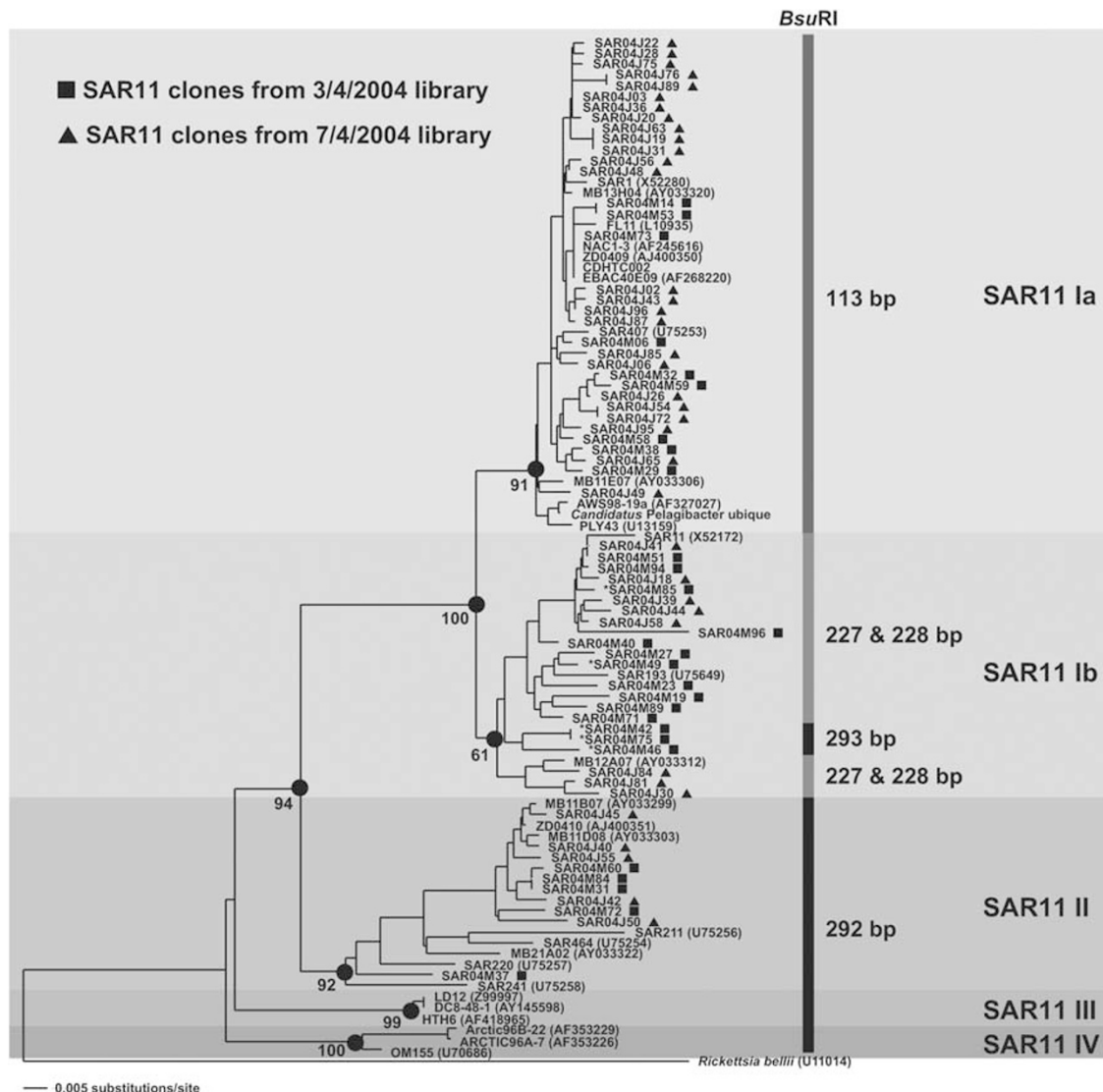
Figure 3 Neighbor-joining 16S rRNA phylogenetic tree showing terminal restriction fragment sizes and relationships among representative members of the SAR11 clade. Clones obtained from this study are shown as squares and triangles for 2004, 40 m spring and summer clone libraries, respectively. Representative sequences derived from public databases are labeled with clone names and accession numbers. Bootstrap proportions over 60% from the neighbor-joining analysis are shown. SAR11 Ia, SAR11 Ib and SAR11 II subclade designations were determined by terminal restriction sites and by tree topology. *Rickettsia bellii* (U11014) was used as the outgroup. Terminal restriction fragment sizes of the 16S rRNA gene from clones are shown to the right of each subclade. Scale bar, 0.05 substitutions per nucleotide position. SAR11 subclade Ia, 113 bp, SAR11 subclade Ib, 227, 228 and 293 bp and SAR11 subclade II 292 bp.

Regular annual SAR11 dynamics were also observed within the upper mesopelagic zone. Mesopelagic SAR11 stocks remained $\leqslant 30 \times 10^{11}$ cells m^{-2} from late summer (>6 months post mixing) through the mixing event of each year contributing $<10\%$ of the SAR11 stocks in the surface 300 m. However, following deep mixing, mesopelagic SAR11 stocks increased >4 fold from the annual minimum accounting for $\sim 25\%$ of the total 0–300 m integrated SAR11 stock (Figure 2d).

16S rDNA terminal restriction fragment designation
Our 2004 spring and summer 40 m clone libraries added to existing 16S rDNA sequence data for 1 m and 200 m (Field *et al.*, 1997; Suzuki *et al.*, 2001; Morris *et al.*, 2005). From each library, 94 clones were screened, of which 28 and 38 were identified as members of the SAR11 clade from March 2004 and July 2004, respectively. Subclade designations for fragments predicted from sequence data and for those observed by T-RFLP analyses were determined as described by Morris *et al.* (2005). T-RFLP assignments are shown in Figure 3 (see Materials and methods). We designated fragment 292 bp as a proxy for SAR11 subclade II; however, a few of the clones belonged to the phylum Cyanobacteria. The contribution of SAR11 subclade II was most pronounced below the euphotic zone ($\geqslant 160$ m) (Table 2), where the cyanobacteria are expected to

Table 2 Mean relative contribution (%) matrix determined from SAR11 T-RFLP fragments representative of each subclade

Depth	Observed T-RFLP	Putative SAR11	Month relative to the timing of deep mixing											
M	Fragment length (bp)	Subclade	−1	Mixing event	+1	+2	+3	+4	+5	+6	+7	+8	+9	+10
0	113	Ia	37 (6)	39 (5)	45 (9)	51 (9)	58 (9)	54 (11)	61 (10)	63 (14)	53 (12)	42 (9)	41 (8)	40 (7)
40			39 (4)	33 (7)	*37 (10)*	42 (21)	*47 (10)*	49 (4)	**51 (11)**	53 (10)	48 (5)	43 (7)	*43 (5)*	43 (5)
80			47 (9)	38 (12)	33 (4)	45 (14)	*43 (10)*	43 (11)	**40 (12)**	37 (4)	38 (3)	35 (4)	*37 (4)*	38 (4)
120			40 (6)	29 (9)	29 (13)	44 (2)	*38 (11)*	41 (10)	*39 (9)*	39 (10)	37 (9)	33 (14)	22 (2)	30 (14)
160			21 (9)	28 (11)	24 (9)	24 (10)	*24 (10)*	25 (17)	*22 (10)*	19 (9)	24 (6)	26 (14)	7 (10)	12 (1)
200			10 (6)	17 (8)	13 (9)	15 (8)	13 (10)	17 (10)	14 (6)	16 (3)	14 (5)	15 (11)	12 (6)	11 (4)
250			9 (2)	15 (14)	11 (5)	13 (3)	*17 (6)*	17 (6)	*17 (6)*	18 (8)	12 (4)	12 (1)	10 (1)	4 (6)
300			*8 (1)*	11 (7)	*11 (6)*	12 (7)	20 (5)	*18 (6)*	13 (5)	16 (5)	14 (11)	14 (6)	*11 (4)*	8 (1)
0	227, 228 and 293	Ib	58 (6)	54 (6)	48 (7)	43 (5)	34 (6)	38 (8)	36 (10)	28 (10)	40 (12)	47 (11)	51 (8)	53 (7)
40			60 (1)	56 (11)	*53 (5)*	47 (16)	*44 (8)*	43 (4)	**41 (14)**	39 (13)	47 (12)	44 (8)	*46 (11)*	46 (11)
80			52 (9)	51 (5)	51 (2)	49 (14)	*49 (6)*	50 (10)	**51 (10)**	52 (2)	53 (6)	62 (8)	*55 (10)*	53 (11)
120			56 (3)	49 (7)	45 (2)	48 (4)	*47 (8)*	47 (12)	*50 (8)*	52 (3)	51 (8)	55 (8)	47 (7)	47 (8)
160			43 (7)	47 (4)	50 (5)	36 (2)	*34 (7)*	26 (2)	*39 (14)*	46 (15)	40 (11)	35 (18)	29 (21)	30 (2)
200			25 (9)	37 (13)	28 (14)	35 (11)	44 (15)	33 (10)	31 (11)	32 (12)	25 (3)	27 (12)	19 (7)	21 (5)
250			21 (12)	28 (8)	24 (6)	29 (14)	*29 (7)*	31 (1)	*26 (8)*	27 (8)	20 (2)	25 (6)	20 (1)	19 (2)
300			*14 (5)*	25 (10)	*28 (9)*	28 (10)	32 (14)	*30 (11)*	29 (18)	25 (16)	18 (3)	30 (10)	*22 (9)*	16 (5)
0	292	II	5 (8)	7 (6)	7 (5)	6 (6)	8 (5)	8 (7)	2 (5)	9 (6)	8 (8)	11 (8)	8 (7)	7 (7)
40			2 (3)	11 (6)	*11 (5)*	10 (4)	*9 (3)*	8 (2)	**8 (8)**	8 (8)	6 (8)	16 (7)	*11 (10)*	9 (10)
80			1 (2)	10 (9)	16 (3)	7 (1)	*8 (2)*	6 (8)	**8 (8)**	11 (2)	10 (1)	3 (5)	*8 (7)*	9 (5)
120			4 (7)	22 (8)	26 (15)	8 (1)	*15 (10)*	12 (2)	*11 (6)*	9 (8)	12 (10)	12 (6)	32 (10)	23 (22)
160			37 (16)	24 (7)	27 (4)	40 (11)	*42 (11)*	49 (15)	*39 (18)*	35 (23)	36 (17)	38 (31)	64 (30)	58 (2)
200			65 (12)	46 (19)	59 (22)	50 (10)	43 (14)	51 (14)	55 (7)	52 (12)	61 (5)	59 (22)	70 (12)	68 (7)
250			70 (10)	57 (22)	65 (11)	59 (17)	55 (10)	52 (5)	*58 (14)*	55 (15)	68 (2)	63 (5)	70 (1)	77 (4)
300			*78 (6)*	64 (16)	*61 (14)*	60 (17)	48 (19)	52 (16)	58 (23)	60 (11)	68 (8)	56 (16)	*67 (13)*	76 (6)

The annual mean composite represents means percent contribution for each subclade over each depth for each month relative to the time of annual mixing. The mean composite was generated from 341 T-RFLP samples collected over a 10-year period at BATS. Each year's data was aligned to the timing of the deep mixing event and are presented as the month prior to mixing through the tenth month following mixing. Values in parenthesis represent one s.d. of the mean. Values in bold did not have any representative samples and values were interpolated from previous and following month following. Values in italic were samples that only contained one sample and were averaged with the months prior and following to smooth data.
The values for each ecotype in the same depth—month field add up to 100%.

290

be inconsequential (DuRand *et al.*, 2001). However, we cannot rule out the possibility that SAR11 subclade II was overestimated in the euphotic zone due to a small contribution of cyanobacteria T-RF 292 bp. The clone libraries demonstrated a spring-to-summer transition in the relative contribution of the subclades: the spring library contained 32% subclade Ia, 50% subclade Ib and 17% subclade II; the summer library was comprised of 66% subclade Ia, 21% subclade Ib and 13% subclade II (Figure 3).

Vertical and temporal variation in SAR11 ecotype distributions
The SAR11 oligonucleotide FISH probes hybridize broadly to the SAR11 clade. To model the contributions of the SAR11 subclades to the quantitative SAR11 cell counts, we incorporated the T-RFLP data from a decade of BATS samples. The percent contribution of each SAR11 subclade was determined from T-RFLP patterns from 1 and 200 m samples (90 and 88 samples, respectively). The decade of SAR11 T-RFLP data was collapsed to a single annual composite, again with the 12 months of each year aligned to the timing of the deep mixing event. The data yielded clear trends over depth and through time (Figure 4). The percent contribution of SAR11 subclades Ia and Ib to SAR11 surface populations averaged 48 ± 12% and 45 ± 12%, respectively, whereas subclade II averaged only 7 ± 7% of SAR11 surface populations (Figure 4a). Maxima for subclades Ia and Ib at 1 m alternated between periods of mixing (late winter early spring) and stratification (summer), resulting in a significant inverse relationship between the two subclades (Figure 4a). The seasonal transition from SAR11 subclade Ib to subclade Ia in the surface waters corresponded to water column stratification and MLD shoaling to < 50 m (Figures 1a and 4a). In contrast, subclades Ia and Ib contributed less to SAR11 populations at 200 m, where they averaged 14 ± 7% and 29 ± 12%, respectively. SAR11 subclade II, dominated at 200 m, averaging 57 ± 15% of total SAR11 populations. The periodicity in relative contribution of the various subclades was less pronounced at 200 m compared with surface samples (Figure 4b).

Quantitative modeling of temporal variation in SAR11 ecotype distributions
T-RFLP patterns at 1 and 200 m indicated that combining the qualitative T-RFLP data with the quantitative FISH data could resolve the dynamics of the three major SAR11 subclades. However, the FISH data were collected from 2003 to 2005 and T-RFLP data was collected between 1991 to 2004 with an additional 36 depth profiles sampled at depths other than 1 and 200 m at a lower temporal resolution. We generated a T-RFLP percent contribution matrix from the 341 samples by determining
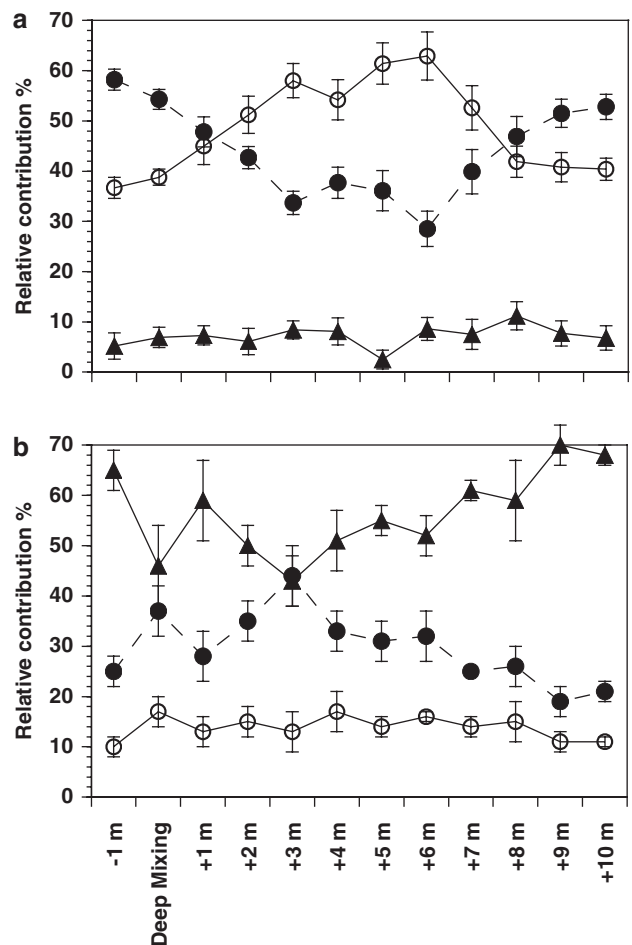

Figure 4 Annual composite of the percent contribution (%) of SAR11 Subclade Ia (open circles) Ib (closed circles) and II (closed triangles) to total SAR11 T-RFLP peak area (see text for details) for the 1 m (**a**) and 200 m (**b**) samples at BATS. The annual composite was generated by aligning each of the 10 years of data to the timing of the deep mixing event and taking the mean of the month prior and 10 months following the deep mixing event. The error bars represent s.e.m.

the mean subclade contribution for each depth and month, aligned to the timing of each year's deep mixing event (Table 2). For the fifth month (post mixing) there was no data for the 40 m and 80 m depths; thus, we interpolated data between the fourth and sixth months (bold in Table 2).

The T-RFLP percent composition matrix (Table 2) was used to factor the quantitative FISH data, producing models of the population dynamics of the subclades in units of cell density (Figure 5). Figure 6 shows the three SAR11 subclade densities integrated over various depth horizons presented as an annual composite relative to the timing of deep mixing. The asterisks in Figure 6 indicate months in which all three subclades were significantly different from each other (Tukey–Kramer HSD at 95% CI).

The integration of all SAR11 subclades over the surface 300 m predicted that subclade II was
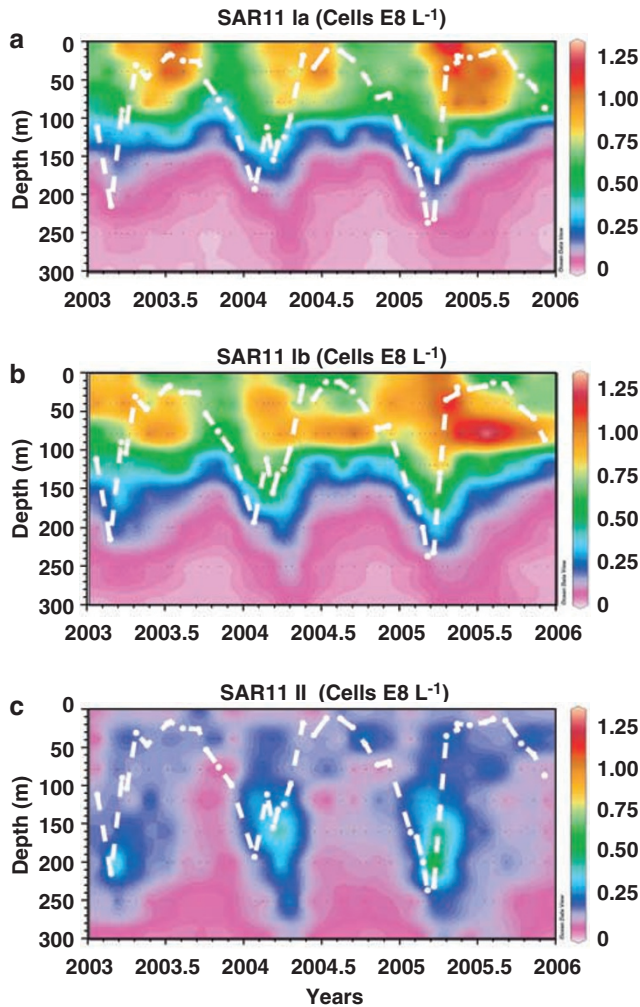
Figure 5 Contours of modeled subclade cell densities in the surface 300 m from 2003 through 2005 for subclade Ia (**a**), Ib (**b**) and II (**c**). The data reported in this figure were determined by multiplying the quantitative SAR11 densities in Figure 1a by the percent contribution matrix presented in Table 2. White dashed line represents MLD as described in Figure 1 and is used to examine distribution patterns in the context of mixing and stratification.

significantly less abundant than the other two subclades (Tukey–Kramer HSD at 95% CI), except for the first month post mixing. Subclade Ib stocks were greater than subclade Ia in the month prior through to the month following deep mixing and then became similar in magnitude within the upper 300 m (Figure 6a).

To further examine how the subclade dynamics varied over depth we partitioned the modeled data into three depth horizons—the euphotic zone (that is, 0–120 m and 0–40 m) and the mesopelagic zone (160–300 m). During periods of mixing the SAR11 densities and stocks were dominated by subclade Ib (Figures 5b and 6b), but, as the MLD shoaled to < 50 m, the euphotic zone stocks of SAR11 subclades Ia and Ib became comparable (0–120 m) (Figure 6b). However, focus-
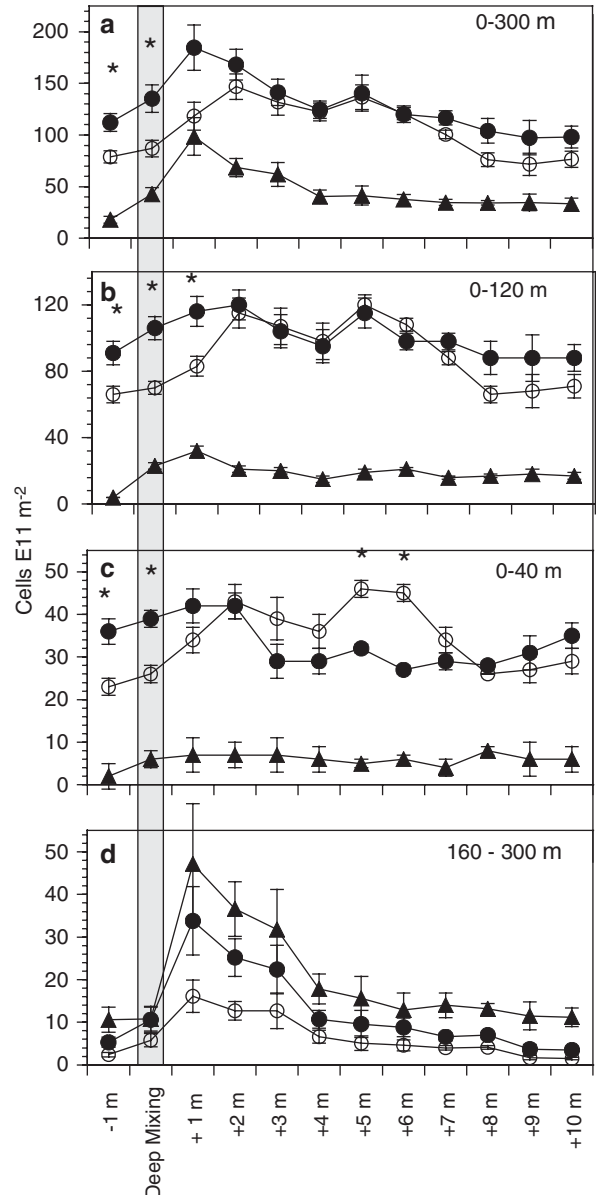
Figure 6 Annual composite of 3 years of modeled SAR11 data aligned relative to the timing of the deep convective mixing event. The data were integrated over various depth horizons (**a**) 0–300 m, (**b**) the euphotic zone 0–120 m, (**c**) the surface 40 m of the euphotic zone and (**d**) the upper mesopelagic zone 160–300 m. Open circles are SAR11 subclade Ia, filled circles are SAR11 subclade Ib and triangles are SAR11 subclade II. Gray bar is a reference that represents the timing of the deep mixing in the composite plot. Error bars represent s.e. Asterisk represents months when all subclade stocks were significantly different from each other by Tukey–Kramer HSD (95% CI).

ing on the surface 40 m revealed a striking transition from subclade Ib dominance during the mixing period to subclade Ia dominance during the summer-stratified period (5–6 months post mixing) (Figure 6c).

Subclade II dynamics were relatively minor within the euphotic zone but became a significant contributor in the upper mesopelagic zone (160–300 m). Members of subclade II were significantly

greater than subclade Ia, but not subclade Ib, in the mesopelagic during the 2 months following deep mixing (Tukey–Kramer HSD at 90% CI) (Figures 5c and 6d). Cell densities of subclade II were greatest and its relative contribution most significant between 200 and 300 m (Figure 5c). Integrated stocks of subclade Ib, and to a lesser extent Ia, were dynamic in the mesopelagic zone (Figure 6d); however, the majority of their contribution was localized between 160 and 200 m (Figure 5).

Discussion

Abundance

The high relative abundance of SAR11 cells we report for the BATS site is consistent with previous studies in the Sargasso Sea (Morris *et al.*, 2002; Malmstrom *et al.*, 2004), the eastern North Atlantic subtropical gyre (Mary *et al.*, 2006) and the coastal oligotrophic Mediterranean Sea (Alonso-Saez *et al.*, 2007). The ubiquity and numerical dominance of SAR11, supported by genomic and culture evidence revealing their metabolism (Rappé *et al.*, 2002; Giovannoni *et al.*, 2005a) and cell-specific radioactive substrate uptake experiments (Malmstrom *et al.*, 2004; Mou *et al.*, 2007) is compelling evidence that these cells play a major role in the oxidation of organic carbon in the oceans.

Evidence for ecotypes: vertical stratification and seasonality

Ecotypes are similar phylogenetic subgroups of bacteria that differ in physiological details that determine niche specificity (Cohan, 2006). Patterns of variation in time and space are often used to identify ecotypes. Subclades Ia and Ib are 3–4% different in 16S rRNA sequence, making them different species by commonly employed standards. SAR11 subclade II differs in 16S rRNA gene sequence by 10–11% from the SAR11 Ia and Ib, which is more than enough sequence divergence to justify placing them in a different genus or family. It is assumed that divergence of ecotypes is caused by selection for specific traits, although in most cases the traits are unknown (Koeppel *et al.*, 2008; Hunt *et al.*, 2008). The best example is *Prochlorococcus*, where the distributions of ecotypes and some traits that contribute to niche specificity are both known (Moore *et al.*, 1998).

The pronounced vertical and seasonal patterns in the SAR11 subclade distributions we observed provide evidence of at least three SAR11 ecotypes. All three of these ecotypes were reported previously (Suzuki *et al.*, 2001; Morris *et al.*, 2005), but this analysis revealed many previously unseen details of their ecology. Our new findings show that SAR11 subclade Ia is not just a surface ecotype, but an ecotype that blooms in the extreme environment of the summer mixed layer, where sunlight (UV) is intense (Zafiriou *et al.*, 2008) and macronutrients are scarce (Steinberg *et al.*, 2001). Recent genome sequence data from a Sargasso Sea summer isolate of SAR11 subclade Ia (strain HTCC7211; unpublished data) indicated that they have nutrient acquisition systems that are not present in coastal strains. Likewise, subclade II is not only a mesopelagic ecotype; it specifically blooms in the upper mesopelagic zone following deep mixing. SAR11 Ib is prominent throughout the surface 300 m year round, but becomes dominant in the euphotic zone during periods of water column instability and blooms in the mesopelagic after mixing. We observed a rhythmic succession each year of subclade Ia replacing Ib in the upper 40 m with the progression of stratification, with the maximum in Ib populations moving downward to 80 m at the height of summer stratification (Figures 5 and 6). There could be finer patterns of evolutionary divergence within the SAR11 clade at BATS, for example the three subclades we delineate could each be a composite of related ecotypes that further subdivide the habitat, but the data could not resolve this.

The partitioning of microbial populations with depth is a dominant theme in microbial plankton ecology. For example, partitioning between the euphotic and aphotic regions has been amply documented for bacterioplankton processes (Carlson *et al.*, 2004; Reinthaler *et al.*, 2006) and for many major microbial plankton clades (Gordon and Giovannoni, 1996; Giovannoni *et al.*, 1996; Wright *et al.*, 1997; DeLong *et al.*, 2006). Evolutionary adaptive radiation into ecotypes that are specialized to specific depth horizons has also been previously described as niche partitioning for SAR11 and *Prochlorococcus* (Field *et al.*, 1997; Moore *et al.*, 1998). Many factors that impact microbial metabolism, including light, temperature and organic and inorganic nutrient availability, vary in consistent patterns with depth. Below, we discuss the potential significance of one of the least obvious of these factors—the export and vertical stratification of dissolved organic matter (DOM).

We note however, that all of our evaluations were made in two dimensions—depth and time—because we had inadequate spatial data to assess the impact of horizontal mixing and advective processes to the results. We, therefore, cannot rule out the possibility that horizontal advection interfered with some of the temporal variability we observed. Nonetheless, we demonstrated that the dynamics of SAR11 stocks fit an annual harmonic model with a high level of significance, indicative of regular seasonal variability.

Geochemical patterns at BATS and SAR11 metabolism

The factors that determine the distributions of SAR11 populations are unknown, but information is accumulating that may eventually lead to a mechanistic understanding of how physical, chemical and biological processes combine to cause the reproducible population dynamics we observed. So

far the genomic and physiological properties that distinguish the ecotypes of SAR11 have not been reported; most of what is known was determined from the study of coastal isolates from subclade Ia (Giovannoni *et al.*, 2005a; Tripp *et al.*, 2008a, b) and metagenomic and metaproteomic data (Venter *et al.*, 2004; Wilhelm *et al.*, 2007; Sowell *et al.*, 2008). These organisms have the smallest genomes known for free-living heterotrophic cells, an apparent consequence of genome streamlining and reduction driven by the selection for efficient growth in oligotrophic ocean habitats (Giovannoni *et al.*, 2005a). Mary *et al.* (2006) found that SAR11 dominated the low nucleic acid fraction of bacterioplankton in the North Atlantic subtropical gyre, consistent with the prediction that small genomes are common to this clade.

The unusual requirements of SAR11 for reduced sulfur compounds (Tripp *et al.*, 2008a) and glycine or serine (Tripp *et al.*, 2008b) may play a role in spatial and temporal control of SAR11 populations, although too little chemical data exists at present to test this hypothesis. The proteome composition of SAR11 appears to be another facet of a survival strategy related to small cell dimensions (Sowell *et al.*, 2008). Recent environmental proteomic data from a summer surface sample at BATS demonstrate that SAR11 ecotypes highly express genes for high-affinity transport systems, involved with phosphate, amino acid and carbohydrate uptake, relative to other genes (Wilhelm *et al.*, 2007; Sowell *et al.*, 2008). The unusual regulation of the SAR11 glyoxylate cycle reported by Tripp *et al.* (2008b) supports the interpretation that osmolytes, amino acids and other precursors are important sources of DOM fueling SAR11 metabolism.

Interactions between bacterioplankton cells and the natural DOM pool are important, but difficult to directly measure, although DOM quality has been shown to affect the community structure of heterotrophic bacterioplankton (Cottrell and Kirchman, 2000). Although it has not yet been possible to directly measure heterotrophic interactions between microbial populations and DOM at a comprehensive molecular level, described patterns of DOM concentration, composition and reactivity in the surface layer suggest they are factors that influence distributions of SAR11 populations.

Photoautotrophy and subsequent food web interactions provide a source of new DOC to the surface waters. A significant portion of daily DOM production is labile, cycled rapidly (that is, minutes to hours) (Nagata, 2000; Carlson, 2002) and does not accumulate appreciably in stratified surface waters (that is, labile monomers <5 nM) (Keil and Kirchman, 1999). The tight coupling between production and contemporaneous uptake of labile compounds largely fuels heterotrophic bacterial production within the euphotic zone (Ducklow, 2000) leaving behind a recalcitrant or 'semi-labile' pool of DOM resistant to rapid microbial degradation by surface

populations (Hansell and Carlson, 2001; Carlson *et al.*, 2004). Members of SAR11 ecotypes Ia and Ib appear to be competitive at utilizing labile organic compounds under low inorganic nutrient conditions. Previous work has demonstrated that surface SAR11 out-compete other bacterioplankton for labile compounds such as amino acids, sugars, aromatic monomers and DMSP (Malmstrom *et al.*, 2004; Mary *et al.*, 2006; Mou *et al.*, 2007).

Our findings also suggest that SAR11 ecotypes II and Ib are implicated in the continued diagenetic alteration and remineralization of euphotic zone produced semi-labile DOM after it is exported from the surface to the upper mesopelagic (Figure 5c). Although resistant to rapid microbial remineralization within the euphotic zone, the seasonally produced semi-labile DOM is more 'bioreactive' compared to the resident mesopelagic DOM (Goldberg *et al.*, 2008). Once exported to the upper mesopelagic, a fraction of the surface semi-labile DOM is oxidized and diagenetically altered (Carlson *et al.*, 1994; Hansell and Carlson, 2001; Carlson *et al.*, 2004; Goldberg *et al.*, 2008). A bloom of SAR11, marine *Actinobacteria*, SAR202 and OCS116 occurs in the mesopelagic (Morris *et al.*, 2005) coincident with this spring episode of DOM oxidation (Hansell and Carlson, 2001), suggesting that the exported DOM provides substrates that drive the spring increases in SAR11 subclade we report here, as well as blooms of other bacterioplankton (Carlson *et al.*, 1996; Steinberg *et al.*, 2001; Carlson *et al.*, 2004).

It seems very likely that the supply, quality and quantity of DOM is different for each of the SAR11 subpopulations we observed, but further experimentation will be needed to measure patterns in the use of specific organic compounds by SAR11 populations.

Further considerations
The time-series data presented here demonstrate the power of combining quantitative FISH abundance data with relative abundance T-RFLP data to assess transitions of microbial ecotypes in time and space. Several recent time-series studies have demonstrated recurring interannual patterns in bacterioplankton community structure linked to regular ecosystem variability (Morris *et al.*, 2005; Fuhrman *et al.*, 2006), but the relative numbers provided by ARISA and T-RFLP methods do not provide accurate estimates of cell abundance or biomass. FISH provides a quantitative measure of cell density (Morris *et al.*, 2002), but it is a targeted approach. The combination of the two forms of data, as we do here, provides estimates of cell numbers that are often critical for geochemical models, but also yield insights into population dynamics.

Significant insight can be gained from such a combination of these approaches. However, there are limitations that must be considered. First,

despite 3 years of relatively high frequency sampling, the temporal coverage is still considered short to predict long-term trends by traditional time-series analyses. Second, primer specificity, high 'species' richness and biases associated with PCR amplification and the restriction digest can lead to over or underrepresentation of specific ribotypes or the generation of 'pseudo-T-RFs' (Suzuki *et al.*, 1998; Egert and Friedrich, 2003; Castle and Kirchman, 2004). Third, phylogenetic resolution by T-RFLP may not be sufficient to resolve genetically distinct populations, thus the distributions of any one of the three SAR11 ecotypes we present may be the sum of additional distinct subclades of SAR11.

Comparative studies of the physiological properties and genomes of cultured isolates and metagenomics have been effective tools for identifying specific mechanisms that link microbial ecotypes with environmental variability. These tools, combined with controlled experiments using natural assemblages and isolates, will test if factors such as macronutrient availability, organic matter quality or light spectrum are responsible for controlling the response of SAR11 ecotypes. Further insights provided by the continued combination of field and experimental approaches will lead to a better understanding of the relationship between gene and genome sequence diversity and the biogeochemical properties of microbial plankton cells.

Acknowledgements

References

Alonso-Saez L, Balague V, Sa EL, Sanchez O, Gonzalez JM, Pinhassi J *et al.* (2007). Seasonality in bacterial diversity in north-west Mediterranean coastal waters: assessment through clone libraries, fingerprinting and FISH. *Fems Microbiol Ecol* **60**: 98–112.

Castle D, Kirchman DL. (2004). Composition of estuarine bacterial communities assessed by denaturing gradient gel electrophoresis and fluorescence in situ hybridization. *Limnol Oceanogr Methods* **2**: 303–314.

Carlson CA. (2002). Production and Removal Processes. In: Hansell DA and Carlson CA (eds). *Biogeochemistry of Marine Dissolved Organic Matter*. Academic Press: San Diego, pp 91–151.

Carlson CA, Ducklow HW, Sleeter TD. (1996). Stocks and dynamics of bacterioplankton in the Northwestern Sargasso Sea. *Deep Sea Res II* **43**: 491–515.

Carlson CA, Ducklow HW, Michaels AF. (1994). Annual flux of dissolved organic carbon from the euphotic zone in the northwestern Sargasso Sea. *Nature* **371**: 405–408.

Carlson CA, Giovannoni SJ, Hansell DA, Goldberg SJ, Parsons RV, Vergin K. (2004). Interactions between DOC, microbial processes, and community structure in the mesopelagic zone of the northwestern Sargasso Sea. *Limnol and Oceanog* **49**: 1073–1083.

Cohan FM. (2006). Toward a conceptual and operational union of bacterial systematics, ecology, and evolution. *Phil Trans R Soc B* **361**: 1985–1996.

Cottrell MT, Kirchman DL. (2000). Natural assemblages of marine proteobacteria and members of the Cytophaga-Flovobacter cluster consuming low and high molecular weight dissolved organic matter. *Appl Environ Microbiol* **66**: 1692–1697.

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.

Ducklow H. (2000). Bacterial production and biomass in the ocean. In: Kirchman DL (ed). *Microbial Ecology of the Oceans*. Wiley-Liss Inc.: New York, pp 85–120.

DuRand MD, Olson RJ, Chisholm SW. (2001). Phytoplankton population dynamics at the Bermuda Atlantic Time-series station in the Sargasso Sea. *Deep Sea Res II* **48**: 1983–2004.

Egert M, Friedrich MW. (2003). Post-amplification Klenow fragment treatment alleviates PCR bias caused by partially single-stranded amplicons. *Aquat Environ Microbiol* **69**: 2555–2562.

Field KG, Gordon D, Wright T, Rappé M, Urbach E, Vergin K *et al.* (1997). Diversity and depth-specific distribution of SAR 11 cluster rRNA genes from marine planktonic bacteria. *App Environ Microbiol* **63**: 63–70.

Follows MJ, Dutiewicz S, Grant S, Chisholm SW. (2007). Emergent biogeography of microbial communities in a model ocean. *Science* **30**: 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.

Giovannoni SJ, Britschgi TB, Moyer CL, Field KG. (1990). Genetic diversity in Sargasso Sea bacterioplankton. *Nature* **345**: 60–63.

Giovannoni SJ, Rappé MS, Vergin K, Adair N. (1996). 16S rRNA genes reveal stratified open ocean bacterioplankton populations related to the green non-sulfur bacteria phylum. *Proc Natl Acad Sci USA* **93**: 7979–7984.

Giovannoni SJ, Tripp HJ, Givan S, Podar M, Vergin KL, Baptista D *et al.* (2005a). Genome streamlining in a cosmopolitan oceanic bacterium. *Science* **309**: 1242–1245.

Giovannoni SJ, Bibbs L, Cho JC, Stapels MD, Desiderio R, Vergin KL *et al.* (2005b). Proteorhodopsin in the ubiquitous marine bacterium SAR11. *Nature* **438**: 82–85.

Goldberg SJ, Carlson CA, Hansell DA, Nelson NB, Siegel DA. (2008). Temporal dynamics of dissolved combined neutral sugars and the quality of dissolved organic matter in the northwestern Sargasso Sea. *Deep Sea Res I* (in press).

Gordon D, Giovannoni SJ. (1996). Detection of stratified microbial populations related to Chlorobium and Fibrobacter species in the Atlantic and Pacific oceans. *Appl Environ Microbio* **62**: 1171–1177.

Hansell DA, Carlson CA. (2001). Biogeochemistry of total organic carbon and nitrogen in the Sargasso Sea: Control by Convective Overturn. *Deep Sea Res II* **48**: 1649–1667.

Hood RR, Laws EA, Armstrong R, Bates N, Brown CW, Carlson CA *et al.* (2006). Functional group modeling: Progress, challenges and prospects. *Deep Sea Research II* **53**: 459–512.

Hunt DE, David LA, Gevers D, Preheim SP, Alm EJ, Polzl MF. (2008). Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science* **320**: 1081–1085.

Karl DM, Laws EA, Morris P, Williams PJl, Emerson S. (2003). Metabolic balance of the open sea. *Nature* **426**: 32.

Koeppel A, Perry EB, Sikorski J, Krizanc D, Warner A, Ward DM *et al.* (2008). Identifying the fundamental units of bacterial diversity: A paradigm shift to incorporate ecology into bacterial systematics. *Proc Natl Acad Sci USA* **105**: 2504–2509.

Kitts CL. (2001). Terminal Restriction Fragment Patterns—a tool for comparing microbial communities and assessing community dynamics. *Curr Iss Int Microbiol* **2**: 17–25.

Keil RG, Kirchman DL. (1999). Utilization of dissolved protein and amino acids in the northern Sargasso Sea. *Aquat Microbial Ecol* **18**: 293–300.

Liu WT, Marsh TL, Cheng H, Forney LJ. (1997). Characterization of microbial diversity by determining terminal restriction fragment length polymorphism of genes encoding 16s rRNA. *Appl Environ Microbiol* **63**: 4516–4522.

Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar A *et al.* (2004). ARB: a software environment for sequence data. *Nucleic Acids Res* **32**: 1363–1371.

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.

Mary I, Heywood JL, Fuchs BM, Amann R, Tarran GA, Burkill PH *et al.* (2006). SAR11 dominance among metabolically active low nucleic acid bacterioplankton in surface waters along an Atlantic meridional transect. *Aquat Microbial Ecol* **45**: 107–113.

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

Morris RM, Connon SA, Rappé M, Vergin KL, Siebold WA, Carlson CA *et al.* (2002). High cellular ab unbundance of the SAR11 bacterioplankton clade in seawater. *Nature* **420**: 806–809.

Morris RM, Vergin KL, Cho JC, Rappe MS, Carlson CA, Giovannoni SJ. (2005). Temporal and spatial response of bacterioplankton lineages to annual convective overturn at the Bermuda Atlantic Time-series Study site. *Limnol Oceanogr* **50**: 1687–1696.

Mou XZ, Hodson RE, Moran MA. (2007). Bacterioplankton assemblages transforming dissolved organic compounds in coastal seawater. *Environ Microbiol* **9**: 2025–2037.

Nagata T. (2000). Production mechanisms of dissolved organic matter. In: Kirchman DL (ed). *Microbial Ecology of the Oceans*. Wiley-Liss: New York, pp 121–152.

Osborn AM, Moore ERB, Timmis KN. (2006). An evaluation of terminal-restriction fragment length polymorphism (T-RFLP) analysis for the study of microbial community structure and dynamics. *Environ Microbiol* **2**: 39–50.

Porter KG, Feig YS. (1980). The use of DAPI for identifying and counting aquatic microflora. *Limnol Oceanogr* **25**: 943–948.

Rappé MS, Connon SA, Vergin KL, Giovannoni SJ. (2002). Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. *Nature* **418**: 630–633.

Rappé MS, Giovannoni SJ. (2003). The uncultured microbial majority. *Annu Rev Microbiol* **57**: 369–394.

Reinthaler T, van Aken H, Veth C, Aristegui J, Robinson C, Williams P *et al.* (2006). Prokaryotic respiration and production in the meso- and bathypelagic realm of the eastern and western North Atlantic basin. *Limnol Oceanogr* **51**: 1262–1273.

Ricke P, Kolb S, Braker G. (2005). Application of a newly developed ARB software-integrated tool for in silico terminal restriction fragment length polymorphism analysis reveals the dominance of a novel pmoA cluster in a forest soil. *Appl Environ Microbiol* **71**: 1671–1673.

Sowell SM, Wilhelm LJ, Norbeck AD, Lipton MS, Nicora CD, Barofsky DF *et al.* (2008). Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *ISME J* **3**: 93–105.

Sprintall J, Tomczak M. (1992). Evidence of the barrier layer in the surface layer of the tropics. *J Geophys Res* **97**: 7305–7316.

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 II* **48**: 1405–1447.

Suzuki MT, Preston CM, Chavez FP, DeLong EF. (2001). Quantitative mapping of bacterioplankton populations in seawater: field tests across an upwelling plume in Monterey Bay. *Aquat Microbial Ecol* **24**: 117–127.

Suzuki MT, Rappé MS, Giovannoni SJ. (1998). Kinetic bias in estimate of costal picoplankton community structure obtained by measurements of small-subunit rRNA gene PCR amplicon length heterogeneity. *Appl Environ Microbiol* **63**: 983–989.

Swofford DL. (2002). *PAUP* Phylogenetic Analysis Using Parsimony (*and other methods)*. Sinauer Associates: Sunderland, MA.

Tripp HJ, Kitner JB, Schwalbach MS, Dacey JWH, Wilhelm LJ, Giovanonni SJ. (2008a). SAR11 marine bacteria require exogenous reduced sulfur for growth. *Nature* **452**: 741–744.

Tripp HJ, Schwalbach MS, Meyer MM, Kitner JB, Breaker RR, Giovannoni SJ. (2008b). 'Candidatus Pelagibacter 'ubique' is a functional glycine-serine auxotroph with a glycine riboswitch preceding malate synthase. *Environ Microbiol* (in press).

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**: 66–72.

Wilhelm L, Tripp HJ, Givan S, Smith D, Giovannoni SJ. (2007). Natural variation in SAR11 marine bacterioplankton genomes inferred from metagenomic data. *Biology Direct* **2**: 27.

Wright TD, Vergin K, Boyd P, Giovannoni SJ. (1997). A novel delta-proteobacterial lineage from the lower ocean surface layer. *Applied and Environmental Microbiology* **63**: 1441–1448.

Zafiriou OC, Xie H, Nelson NB, Najjar RG, Wang W. (2008). Diel carbon monoxide cycling in the upper Sargasso Sea near Bermuda at the onset of spring and in midsummer. *Limnol Oceanog* **53**: 835–850.