Title
Abundance of Broad Bacterial Taxa in the Sargasso Sea Explained by Environmental Conditions but Not Water Mass

Journal
Applied and Environmental Microbiology, 80(9)

ISSN
0099-2240

Authors
Sjöstedt, Johanna
Martiny, Jennifer BH
Munk, Peter
et al.

Publication Date
2014-05-01

DOI
10.1128/aem.00099-14

Peer reviewed

AEM
Journals.ASM.org

Abundance of Broad Bacterial Taxa in the Sargasso Sea Explained by Environmental Conditions but Not Water Mass

Johanna Sjöstedt,[a] Jennifer B. H. Martiny,[b] Peter Munk,[c] Lasse Riemann[d]

Department of Natural Sciences, Linnaeus University, Kalmar, Sweden[a]; Department of Ecology and Evolutionary Biology, University of California, Irvine, California, USA[b]; National Institute of Aquatic Resources, Technical University of Denmark, Charlottenlund, Denmark[c]; Marine Biological Section, University of Copenhagen, Helsingør, Denmark[d]

To explore the potential linkage between distribution of marine bacterioplankton groups, environmental conditions, and water mass, we investigated the factors determining the abundance of bacterial taxa across the hydrographically complex Subtropical Convergence Zone in the Sargasso Sea. Based on information from 16S rRNA gene clone libraries from various locations and two depths, abundances of the predominant taxa (eubacteria, *Archaea*, *Alphaproteobacteria*, *Gammaproteobacteria*, *Bacteroidetes*, and the *Roseobacter*, SAR11, and SAR86 clades) were quantified by real-time PCR. In addition, the abundances of *Synechococcus*, *Prochlorococcus*, and picoalgae were determined by flow cytometry. Linear multiple-regression models determining the relative effects of eight environmental variables and of water mass explained 35 to 86% of the variation in abundance of the quantified taxa, even though only one to three variables were significantly related to any particular taxon's abundance. Most of the variation in abundance was explained by depth and chlorophyll *a*. The predominant phototrophs, *Prochlorococcus* and picoalgae, were negatively correlated with phosphate, whereas eubacteria, heterotrophic bacteria, and SAR86 were negatively correlated with nitrite. Water mass showed limited importance for explaining the abundance of the taxonomical groups (significant only for *Roseobacter*, explaining 14% of the variation). The results suggest the potential for predicting the abundance of broad bacterioplankton groups throughout the Sargasso Sea using only a few environmental parameters.

Different oceanic regions harbor distinct bacterioplankton communities (1–4). Such differences between communities may be explained by environmental variation and possibly dispersal limitation (5–7). Indeed, a number of recent studies found that the taxonomic composition of marine bacterioplankton correlates closely to the water mass from which the community is sampled (8–11). The characteristics of a given water mass encapsulate a variety of abiotic parameters—including temperature and salinity—which may influence bacterioplankton composition (10). In addition, density gradients across water masses may be hydrographic barriers to microbial dispersal, which could further influence microbial distributions (8).

Most of the above-mentioned studies investigated variation in bacterial composition by fingerprinting techniques (9, 11) or by clustering nucleotide sequences based on sequence similarity (e.g., see references 4, 8, and 10). The authors then tested whether water mass or other abiotic conditions explained variation in the relative abundances of taxa among samples (9, 10). In contrast, fewer studies examined how these environmental factors influence the absolute abundance of particular taxa or broader clades (but see, e.g., references 2 and 12). However, such quantitative relationships provide a model of a taxon's realized niche (13), which can then be incorporated into biogeochemical models.

For the marine environment, studies that quantify microbial abundance often focus on broad taxa that can be counted by microscopy or flow cytometry, including *Prochlorococcus*, *Synechococcus*, and heterotrophic bacterioplankton (e.g., 14, 15). Such studies have provided a good understanding of the regional and global distributions of these particular groups (e.g., phytoplankton and picoheterotrophs [16–18]), and *Synechococcus* and *Prochlorococcus* are indeed already included in ocean models (19, 20). In contrast, quantifying the absolute abundance of particular subgroups is more difficult. While rRNA gene amplification methods are powerful for comparing relative abundances among samples, primer selectivity across groups can result in biased estimates of absolute abundances (see, e.g., references 21 and 22). As a result, quantitative PCR (qPCR) or fluorescent *in situ* hybridization (FISH) methods are preferred for absolute quantification (see, e.g., reference 23). For instance, qPCR has been used to relate various environmental factors to the abundance of cyanobacterial subgroups (ecotypes) (24–26) and specific heterotrophic taxa, such as SAR11 and the *Roseobacter* clade affiliation (RCA) (27–30). Similarly, researchers used catalyzed reporter deposition (CARD)-FISH to demonstrate that flavobacterial clade abundances are highly homogenous within water masses in the North Atlantic (12).

Despite these recent advances, however, our quantitative understanding of what factors control the abundance of marine (particularly heterotrophic) bacterial taxa remains limited. Thus, the aim of this study was 3-fold: (i) to quantify the abundance of the dominant, broadly defined (i.e., subphylum and above) bacterial taxa across water masses in the Sargasso Sea region, (ii) to identify the environmental factors, including water mass identity, that best explain the variation in the abundance of the particular taxa, and

Received 9 January 2014 **Accepted** 17 February 2014
Published ahead of print 21 February 2014
Editor: K. E. Wommack
Address correspondence to Lasse Riemann, LRiemann@bio.ku.dk.
This article is contribution no. P104 from Galathea 3.
Supplemental material for this article may be found at http://dx.doi.org/10.1128/AEM.00099-14.

doi:10.1128/AEM.00099-14

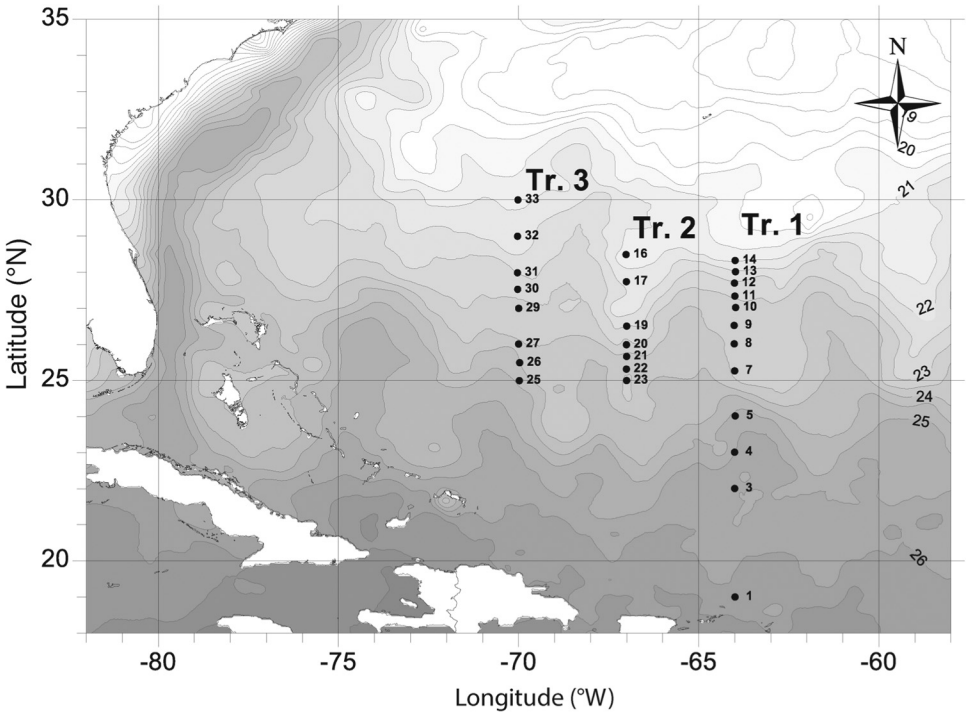

FIG 1 Map showing the three sampled transects (Tr.) in the Sargasso Sea. Data on sea surface temperatures are from satellite observations on 4 April 2007 as described by Riemann et al. (34), contoured at 0.5°C intervals. Clone libraries were constructed from the following stations (St.) and depths: St. 11, 10 m; St. 16, 10 m; St. 23, 10 m; St. 30, 10 m; and St. 32, 80 m.

(iii) to ascertain whether the relative influences of the investigated factors differ among taxa. To do this, we used multiple linear regression followed by variance partitioning of the regressor variables. This statistical approach has several advantages over commonly used pairwise correlations. First, multiple regression aims to disentangle the influence of correlated parameters. For instance, chlorophyll (Chl) *a* usually decreases with depth, and two separate correlation tests would not tease apart whether one, the other, or both explain a group's abundance. The analysis can also incorporate categorical variables, such as water mass. In this way, we can test whether water mass identity helps to predict bacterial abundance, beyond the information already contained in the other variables included, such as temperature and salinity. In addition, the approach allows a direct quantitative comparison of the importance of each environmental factor to the abundance of different bacterial groups. Finally, multiple regression eliminates the need to correct for multiple comparisons, as all factors are tested at the same time.

The Sargasso Sea is bounded by the ocean currents that form the North Atlantic Gyre: the Gulf Stream to the north and west, the North Equatorial Current to the south, and a region of weak recirculation flow to the east. Cold waters from the north and warm tropical water from the south meet at the Subtropical Convergence Zone (STCZ) and generate a complex hydrography with frontal zones and sharp gradients in environmental parameters (31). At 27 stations that crossed these water masses, we quantified the abundance of eubacteria, *Archaea*, and several heterotrophic taxa by qPCR as well as phytoplankton groups by flow cytometry. Our analysis revealed that just two or three parameters usually explain more than half (and up to 86%) of the variation in the abundance of these broad microbial taxa in this region.

MATERIALS AND METHODS

Sampling and DNA extraction. Sampling was carried out from RV Vædderen during a part of the Danish Galathea 3 expedition from 29 March to 10 April 2007 along three transects traversing the STCZ in the Sargasso Sea (Fig. 1). Samples of 4 to 5 liters were filtered through 0.22-μm Sterivex capsule filters (Millipore) via peristaltic pump (<100 ml min^{-1}). Lysis buffer was added (1 ml of 20% sucrose, 50 mM EDTA, 50 mM Tris hydrochloride [pH 8.0]), and the filters were sealed with Parafilm and frozen at −80°C. DNA was extracted using an enzyme-phenol-chloroform protocol (32) but with a 30-min lysozyme digestion (5-mg ml^{-1} final concentration) at 37°C and an overnight proteinase K digestion (100-μg ml^{-1} final concentration) at 55°C (33). DNA was quantified using PicoGreen (Molecular Probes).

Environmental parameters and water masses. Data for temperature, salinity, nutrients, Chl *a*, and abundance of *Synechococcus*, *Prochlorococcus*, and picoeukaryotes from these samples were first reported in reference 34. The methods by which they were collected are therefore only briefly described here. Vertical profiles of salinity and temperature were measured using a Seabird 9/11 CTD equipped with a 12 Niskin (30-liter) bottle rosette sampler. Nutrient concentrations (phosphorus, nitrate, nitrite, ammonia, and silicate) were determined using an automatic nutrient analyzer (Dansk Havteknik). Chl *a* concentrations were determined using ethanol extraction and a Turner Designs model 700 fluorometer. Bacterial abundance was determined by flow cytometry (FACSCalibur; Becton, Dickinson) after staining with SYBR green I (Molecular Probes). Picoalgal and cyanobacterial abundances (*Prochlorococcus* and *Synechococcus*) were determined as for bacteria but without staining. Gating was defined using cultures of *Prochlorococcus*, *Synechococcus*, and picoplankton as references.

Water masses were classified as Northern Sargasso water, subtropical mode water, or the mixing zone based on temperature and salinity characteristics (33) (Fig. 2). Areas with a temperature below 22.5°C were defined as Northern Sargasso water, and areas with temperatures above

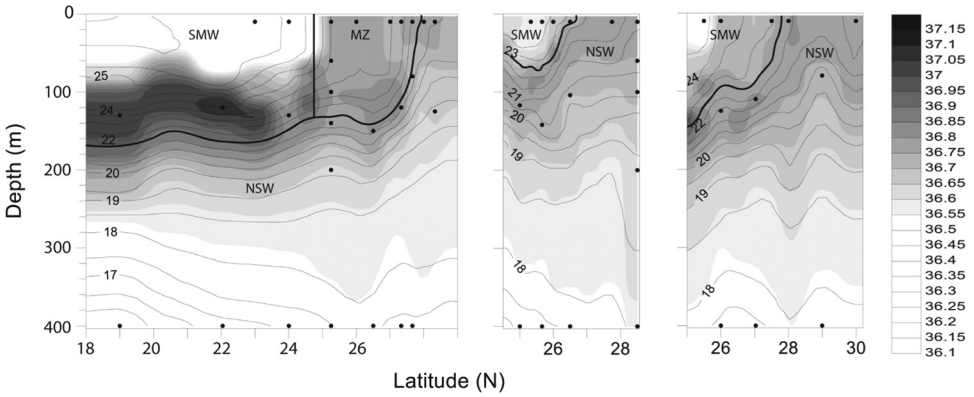

FIG 2 Vertical sections of temperature (contour lines, °C) and salinity (shading) along transects 1, 2, and 3 (left to right) and distribution of the water masses: subtropical mode water (SMW), Northern Sargasso water (NSW), and the mixing zone (MZ). Dots indicate where DNA samples for community composition analyses were taken.

22.5°C were defined as subtropical mode water. Along transect 1 a mixing zone was formed between the Northern Sargasso water and the subtropical mode water with a temperature between 22.5°C and 24.5°C (approximately 24.5°N to 27.5°N, from the surface to a maximum of 130 m [Fig. 2]).

Clone libraries. To examine bacterial community composition, clone libraries were constructed from 5 samples representing the three transects and two different depths (Fig. 1). Bacterial 16S rRNA genes were amplified using primers 27F and 1492R (35), and products were purified using a gel extraction kit (E.Z.N.A.) followed by the Cycle-Pure kit (E.Z.N.A.). Products were cloned (TOPO TA cloning kit; Invitrogen), plasmid DNA was extracted (R.E.A.L. Prep 96 plasmid kit; Qiagen), and inserts were sequenced (Macrogen, South Korea). Sequences shorter than 700 bp or containing undetermined nucleotides or detected as chimeras (using ChimeraSlayer utility with default settings [36]) were removed, resulting in 348 sequences (58 to 83 per library). Sequences were clustered into operational taxonomic units (OTUs) at 97% similarity using SeqMan (DNA Star). Consensus sequences from each OTU was imported into an ARB database and aligned with their closest relatives using the integrated alignment module within the ARB package (37).

qPCR. Seven nonexclusive bacterial groups (eubacteria, *Alphaproteobacteria*, the SAR11 clade, the *Roseobacter* clade, *Gammaproteobacteria*, the SAR86 clade, and *Bacteroidetes*) and *Archaea* were quantified using a Step One Plus real-time PCR system (Applied Biosystems), StepOne software v. 2.0, and MicroAmp fast optical 96-well reaction plates. Primers and conditions used are described in Table 1. Linearized recombinant plasmids from the clone libraries were used as standards for qPCR by making dilution series covering 5×10^1 to 5×10^6 copies. Extracted plasmid DNA (Plasmid Miniprep kit I; E.Z.N.A.) was linearized using the restriction enzyme NotI (Roche) according to the manufacturer's instructions, purified (Cycle-Pure kit; E.Z.N.A.), and quantified (NanoDrop; Thermo Scientific). The number of molecules per microliter was calcu-

TABLE 1 Primers, conditions, PCR efficiencies, and R^2 values for standard curves (± standard deviations) in the quantitative PCR assays

Target group	Primer name (concn, in nM) and sequence	Annealing temp, in °C (time)	Extension temp, in °C (time)	Efficiency	R^2	Reference
Eubacteria	Eub338F (250), ACTCCTACGGGAGGCAGCAG Eub518R (250), ATTACCGCGGCTGCTGG	60 (15 s)	72 (20 s)	87.5 ± 4.1	0.988 ± 0.004	80 81
Alphaproteobacteria	Eub338F (250) Alpha685R (250), TCTACGRATTTCACCYCTAC	54 (15 s)	72 (22 s)	83.3 ± 3.7	0.980 ± 0.01	80 80
Bacteroidetes	Cfb319F (250), GTACTGAGACACGGACCA Eub518R (250)	63 (15 s)	72 (30 s)	98.5 ± 1.7	0.982 ± 0.01	82 81
Gammaproteobacteria	Gamma395F (250), CMATGCCGCGTGTGTGAA Eub518R (250)	64 (15 s)	72 (30 s)	96.3 ± 5.0	0.988 ± 0.01	83 81
SAR11 clade	S11-433F (500), CTCTTTCGTCGGGGAAGAAA S11-588R (1,500), CCACCTACGWGCTCTTTAAGC	59 (1 min)	59 (1 min)	100.2 ± 3.1	0.992 ± 0.005	15 15
SAR86 clade	S86-492F (250), CAGAATAAGSACCGGCTAATTC Gamma680mod (250),[a] ATTCCACCGCTACACTRTG	59 (1 min)	59 (1 min)	105 ± 3.6	0.982 ± 0.01	15 84
Roseobacter clade	ROS292F (1,500), GGTTTKAGAGGATGATCAGCMAC ROS567R (500), CCAGTAATTCCGAACAACGCTAA	67 (1 min)	67 (1 min)	88.3 ± 11.3	0.95 ± 0.06	15 15
Archaea	ARCH349F (1,000), GYGCASCAGKCGMGAAW ARCH806R (1,000), GGACTACVSGGGTATCTAAT	50 (30 s)	72 (30 s)	80.7 ± 4.6	0.95 ± 0.02	85 85

[a] Modified from Gamma680 (84) to target all sequences in the clone libraries classified as SAR86.

lated using the formula $[a/(b \times 660)] \times 6.022 \times 10^{23}$, where a is the plasmid DNA concentration (in g μl^{-1}), b is the plasmid length (in bp; vector and insert), 660 is the average molecular weight of 1 bp, and 6.022×10^{23} is the molar constant (38).

Annealing temperatures were initially determined by gradient endpoint PCR analysis (50 to 70°C) of each primer set with standards containing the target sequence and further refined in qPCR assays after primer concentration optimization. Optimal primer concentrations were determined using a matrix of concentrations of forward and reverse primers (250 nM, 500 nM, 1,000 nM, and 1,500 nM) and 5×10^4 copies of the corresponding target. The specificity of each primer pair was evaluated using 5×10^4 copies of the standard for the six other quantified taxa. Nonspecific amplification did not exceed 300 copies for any of the taxa.

Inhibition tests were performed for all samples using the *Bacteroidetes* primers. Each sample was spiked with 5×10^4 copies of the *Bacteroidetes* standard, and inhibition was quantified using the formula $1 - [(C_{t\text{sample}} - C_{t\text{standard}})/C_{t\text{standard}}] \times 100$ (39), where C_t is the cycle threshold. These tests showed that inhibition could be avoided by using 0.5 or 0.05 ng of DNA.

Triplicate reactions were conducted for each DNA sample and for each dilution of the standard. Reaction mixtures contained 1× SYBR green PCR master mix (Applied Biosystems), forward and reverse primers (250 nM to 1,500 nM [Table 1]), and 0.5 or 0.05 ng of environmental DNA. Thermal cycling conditions were 95°C for 10 min and 40 cycles of 95°C for 15 s and annealing and extension temperatures as described in Table 1, followed by a melting curve. Both PCR efficiencies (80.7 to 105%) and R^2 values (0.95 to 0.992) for the standard curves were within an acceptable range (40).

To account for variations in bacterial abundance and DNA extraction efficiencies between samples when examining distributions of taxa, the abundance of each taxon was calculated based on the relative abundance (estimated from the gene copy number) and flow cytometry counts. Based on qPCR data, the fractions of eubacteria and *Archaea* were calculated as eubacteria/(eubacteria + *Archaea*) and *Archaea*/(eubacteria + *Archaea*). Absolute abundances of eubacteria and *Archaea* were then estimated by multiplying these fractions by total flow cytometry cell counts. The fractions of all other groups targeted by qPCR were calculated in a similar manner; e.g., SAR11 = SAR11/eubacteria. Absolute abundances were then calculated by multiplying with the abundance of eubacteria. The abundance of heterotrophic bacteria was calculated by subtracting counts for *Prochlorococcus and Synechococcus* from that for eubacteria.

Statistics. To investigate the relative importance of different environmental variables on the abundance of the bacterial taxa, we used multiple linear regressions followed by variance partitioning of the regressor variables. For our purposes here, there are three important components of the regression results: (i) which variables are significant (which explain any of the variation in a taxon's abundance), (ii) the total variation explained by all significant variables, that is, the predictive "power" of the model given by the R^2 value, and (iii) the sensitivity of the group's abundance to the significant variables, represented by the sign (positive or negative) and value of the partial regression coefficients.

Before this analysis, we first assessed the redundancy of the abiotic variables by examining their pairwise correlations in PRIMER6 (41). Nitrate and phosphate (P) were highly correlated, and we decided to remove nitrate because two other nitrogen sources were available and because phosphate concentrations are low in the Sargasso Sea and have previously been shown to be an important factor for the distribution of *Prochlorococcus*. Water mass, depth, nitrite, ammonium, phosphate, salinity, silicate, temperature, and Chl *a* were retained in the analysis. The distribution of temperature and salinity is illustrated in Fig. 2, whereas the distribution of the other environmental parameters is illustrated in Fig. S1 in the supplemental material. All variables were continuous except water mass, which was coded into the three categories described above. We purposely excluded the abundance of bacterial taxa as independent variables in the models, as it would be impossible to disentangle the importance of interactions with these groups versus indirect correlations with other environmental variables. The one exception is that we included Chl *a* in models of the heterotrophic taxa, because it may be a proxy for carbon availability.

To improve normality, data on nitrite, phosphate, ammonium, and Chl *a* were $\log(x)$ transformed. Similarly, the abundance of eubacteria, *Archaea*, the different bacterial taxa, and picoalgae were $\log(x + 1)$ transformed. Multiple linear regressions were performed on the transformed data using the lm function in R (42). Nonsignificant variables were removed after the first analysis, and the partial regression coefficients reported are those from a secondary regression. Partial residual plots were used to visualize the relationships of some of the significant variables; these plots were created in R using the cr.plots function in the "car" package (43). The relative importance of the significant variables was evaluated with the calc.relimp function in the "relaimpo" R package using the LMG method (44). To compare the results across taxa, we report the nonnormalized variance explained by each variable.

Nucleotide sequence accession numbers. The partial 16S rRNA gene sequences obtained from the clone libraries have been deposited in GenBank and assigned accession numbers KC197360 to KC197707.

RESULTS AND DISCUSSION

Abundance of main bacterioplankton taxa in the Sargasso Sea. The clone libraries of 16S rRNA genes amplified from five stations and two different depths revealed a bacterioplankton community that was comparable to those described in previous studies from the area (see, e.g., references 45 to 47). *Alphaproteobacteria* was the dominant class in the libraries, with the SAR11 clade as the main subgroup (Fig. 3A). In addition, *Rhodobacterales* (family *Rhodobacteraceae*), which contains the *Roseobacter* clade, was detected in all surface samples (Fig. 3B). In three of the libraries from surface water, *Gammaproteobacteria* was the second most common class, with the SAR86 clade accounting for 53 to 100% of the clones (data not shown). *Actinobacteria* accounted for 5 to 7% of the clones at three stations, and cyanobacterial sequences were also widely represented.

Based on the results from the clone libraries, we selected published primers for quantifying key bacterial groups and *Archaea* by qPCR (Table 1). Conditions were optimized for all primer pairs; however, a robust actinobacterial qPCR assay could not be developed. In addition to the qPCR assays, we used flow cytometry to quantify *Synechococcus*, *Prochlorococcus*, picoalgae (less than ~4 μm in size), and heterotrophic bacteria (and *Archaea*). Gene copy numbers from the qPCR were used to calculate the relative abundances of the different bacterial taxa, and the abundances were obtained by multiplying with flow cytometric counts. It is important to note that heterotrophic bacteria and picoalgae are not monophyletic, in contrast to the other taxa quantified.

Alphaproteobacteria were, on average, the most abundant class enumerated (1.2×10^5 cells ml^{-1}), and most of these cells were from the SAR11 clade (8.5×10^4 cells ml^{-1}), with fewer from the *Roseobacter* clade (3.1×10^4 cells ml^{-1} [Table 2]). *Synechococcus*, *Prochlorococcus*, *Bacteroidetes*, and *Gammaproteobacteria* were almost equally abundant (on average, 1×10^4 to 7×10^4 cells ml^{-1} [Table 2]), and *Gammaproteobacteria* encompassed mostly SAR86 cells. In a few samples, the estimated abundances of SAR11 and SAR86 exceeded those of *Alphaproteobacteria* and *Gammaproteobacteria*, respectively. A likely explanation is that these clades contain sequences not efficiently targeted by the class level primers. Indeed, we observed mismatches between the alphaproteobacterial primer and the majority of the SAR11 sequences obtained from our clone libraries (data not shown). The average

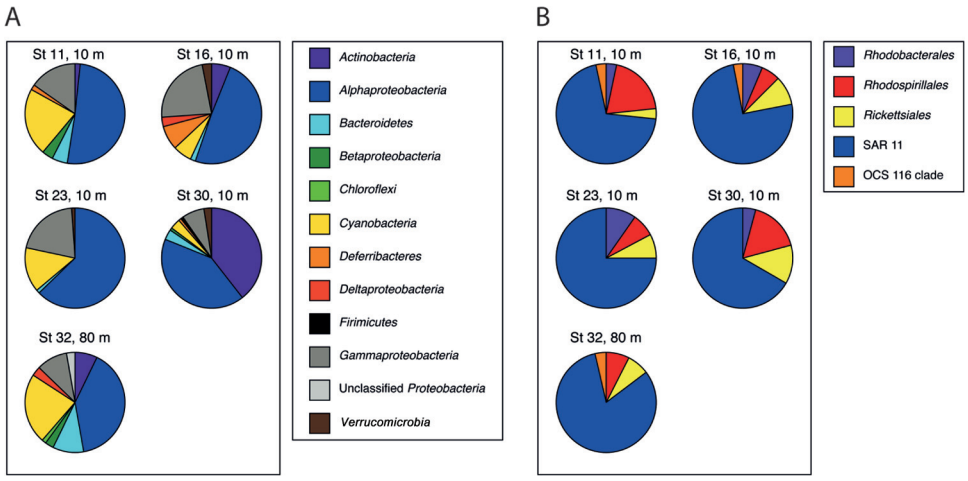

FIG 3 Bacterial community composition at five stations as analyzed by 16S rRNA gene clone libraries. Shown are the distribution of phyla/classes (A) and orders within *Alphaproteobacteria* (B). Each clone library contained 58 to 83 clones. For station numbers, see Fig. 1.

abundances of picoalgae and *Archaea* were 1.4×10^3 cells ml^{-1} and 1×10^5 cells ml^{-1}, respectively. In agreement with a previous study (48), *Archaea* generally contributed only a few percent in the upper waters, but the abundance was usually similar to or exceeded bacterial abundance below 100 m.

Importance of water mass and environmental factors for taxon abundance. Our results show that only a few abiotic variables had significant influence on the abundance of a variety of broad groups of bacterioplankton across the Sargasso Sea. Although we included nine variables in the multiple-regression models (water mass identity plus 8 continuous variables), only one to three variables at a time were significantly related to any taxon's abundance (Fig. 4). Yet these simple models performed quite well. The significant variables explained between 35 and 86% of the variation in the taxon's abundance (R^2 between 0.35 and 0.86 [Fig. 4]). Some of the unknown variation is possibly due to biotic interactions (e.g., virus, predation, and competition) as well as unmeasured environmental variables (49) such as dissolved oxygen, iron concentration, or turbulence.

Chl *a* concentration, a proxy for phytoplankton biomass, was significantly related to the abundance of all the heterotrophic taxa (positive partial regression coefficients in Fig. 4). Thus, we can compare the sensitivities of the groups to Chl *a*. The abundance of the *Roseobacter* clade appears to be particularly sensitive to Chl *a* concentrations, as the partial regression coefficient is relatively high. In contrast, eubacteria are much less sensitive to changes in Chl *a* (Fig. 5). In other words, for the same change in Chl *a* concentration, *Roseobacter* abundance is predicted to change more than that of most other heterotrophic groups. This result is perhaps unsurprising given that the highest abundances of *Roseobacter* previously have been reported above the upper mixed layer, where phytoplankton are most abundant (50–52). In addition, roseobacters thrive in association with marine phytoplankton blooms (32, 51, 53). On the other hand, it was unexpected that SAR11, given its typical oligotrophic lifestyle (54) and potentially reduced capacity to respond to changing conditions (55), was equally sensitive to changes in Chl *a* as other heterotrophic groups (e.g., eubacteria and *Gammaproteobacteria*) (Fig. 4).

To examine the relative importance (i.e., the predictive power) of the variables for the abundance of each bacterial taxon, we used an additional analysis to partition the variance explained by the regression models. These results show not only that both depth and Chl *a* are significantly correlated with heterotrophic taxa but also that these two parameters alone can explain a large proportion of the variation in their abundances (Fig. 4). Depth is well known as an important determinant of bacterial community composition. For phototrophs, depth in our models is presumably related to light availability, as we did not include a direct measure of light availability in the models but did include nutrients and temperature (which also vary with depth). For heterotrophic taxa, depth may be a proxy for the amount of substrate available because the concentration and chemical composition of dissolved organic matter (DOM) change with depth (56, 57) and the bioavailability of DOM is higher by the surface than in deep water (58). This pattern may be linked to the DOM-generating processes (59) but also to sunlight-driven transformation of recalcitrant

TABLE 2 Minimum, maximum, and average abundances with standard deviations for the different taxonomic groups[a]

Taxon	Abundance (cells ml^{-1})			
	Minimum	Maximum	Avg	SD
Eubacteria	1.7×10^4	7.4×10^5	2.9×10^5	1.9×10^5
Heterotrophic bacteria	1.5×10^4	5.2×10^5	2.1×10^5	1.5×10^5
Alphaproteobacteria	5.4×10^3	4.5×10^5	1.2×10^5	1.1×10^4
SAR11 clade	3.3×10^3	2.6×10^5	8.5×10^4	6.8×10^4
Roseobacter group	4	1.7×10^5	3.1×10^4	4.5×10^4
Gammaproteobacteria	5.5×10^2	6.5×10^4	2.2×10^4	1.7×10^4
SAR86 clade	BD	2.3×10^5	2.9×10^4	4.9×10^4
Bacteroidetes	7.5×10^2	1.3×10^5	3.6×10^4	3.2×10^4
Synechococcus	BD	3.6×10^5	7.1×10^4	9.9×10^4
Prochlorococcus	BD	4.6×10^4	1.1×10^4	1.2×10^4
Picoalgae	BD	8×10^3	1.4×10^3	1.6×10^3
Archaea	BD	5.1×10^5	1.0×10^5	1.2×10^5

[a] *Synechococcus*, *Prochlorococcus*, and picoalgal abundances were determined using flow cytometry, while abundances of other groups were calculated from a combination of qPCR and flow cytometry data. See Materials and Methods for details. $n = 59$ for all groups except *Roseobacter*, for which $n = 58$. BD, below detection; the detection limit for qPCR ranged from ~2 to 137 gene copies ml^{-1} of seawater.

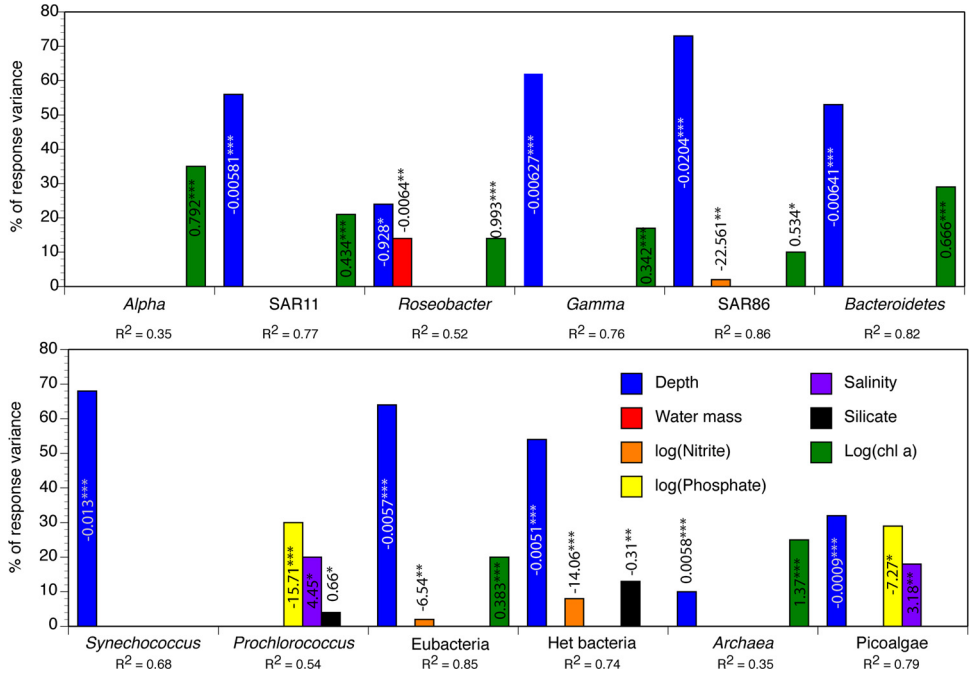

FIG 4 Relative importance of the significant independent variables explaining the variance in abundance of the different taxonomic groups. Partial regression coefficients are written on the bars and shown only for statistically significant variables. *Alpha*, *Alphaproteobacteria*; *Gamma*, *Gammaproteobacteria*; Het bacteria, heterotrophic bacteria. $n = 59$ for all groups except *Roseobacter*, for which $n = 58$. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

DOM to more labile forms available for heterotrophic bacteria (60).

Remarkably, depth explained 68 and 73% of the variation in *Synechococcus* and SAR86 abundance, respectively (Fig. 4). While it is well known that *Synechococcus* is restricted to the upper euphotic zone (61, 62), less is known about the distribution and ecology of the SAR86 clade (63). Our data show that the SAR86 clade is also confined to the euphotic zone (Fig. 6A), which is consistent with findings of SAR86 clones predominantly in upper waters (46, 52) and presence of genes encoding proteorhodopsin in SAR86 genomes (64) that may allow for energy acquisition from light (65). This points to a high autecological importance of phototrophy for members of this clade. Moreover, given the predominance of this group in surface waters (accounting for, on average, 18% of the eubacteria above the Chl *a* maximum), our data highlight a potentially high importance of phototrophic energy acquisition for noncyanobacterial bacterioplankton in the Sargasso Sea.

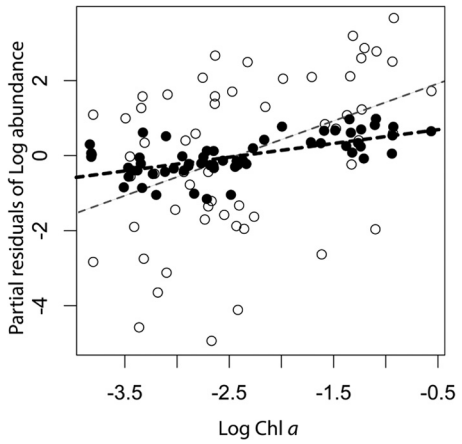

FIG 5 Partial residual plots of the relationship between Chl *a* and the abundance of the *Roseobacter* clade (thin dashed line and open circles) and eubacteria (thick dashed line and solid circles).

In contrast to heterotrophic taxa, *Prochlorococcus* and picoalgae were negatively correlated with phosphate (P) (Fig. 4), suggesting that their overall abundance is nutrient limited. Indeed, *Prochlorococcus* is often abundant in ocean regions where phosphate regulates primary productivity (66). For instance, in the Sargasso Sea (<0.1 μmol of P liter^{-1}), *Prochlorococcus* contains multiple genes involved in orthophosphate uptake, organic P utilization, and regulation, whereas these genes are almost absent in *Prochlorococcus* from regions with higher P concentrations (66). *Prochlorococcus* and picoalgae were also positively correlated with salinity, whereas no other bacterial taxa were affected. The distribution of both of these groups has been linked to salinity in other oceanic regions (67–69); however, given that salinity differed less than one unit across our data set, the linkage may reflect a correlation with other unmeasured variables, rather than a selective pressure exhibited by salinity *per se*.

Although temperature ranged widely across the sampling area (from 15.6 to 25.7°C), this parameter did not explain the abundance of any of the bacterial groups examined. Nevertheless, it is likely that temperature affects the distribution of bacterial taxa at finer phylogenetic resolution. For instance, specific lineages within SAR11 (70, 71) and *Prochlorococcus* (24) are associated with different temperatures.

The percentage of variation explained by the model was lowest for *Alphaproteobacteria*, followed by the *Roseobacter* clade and *Prochlorococcus* (Fig. 4). Correlations between environmental pa-

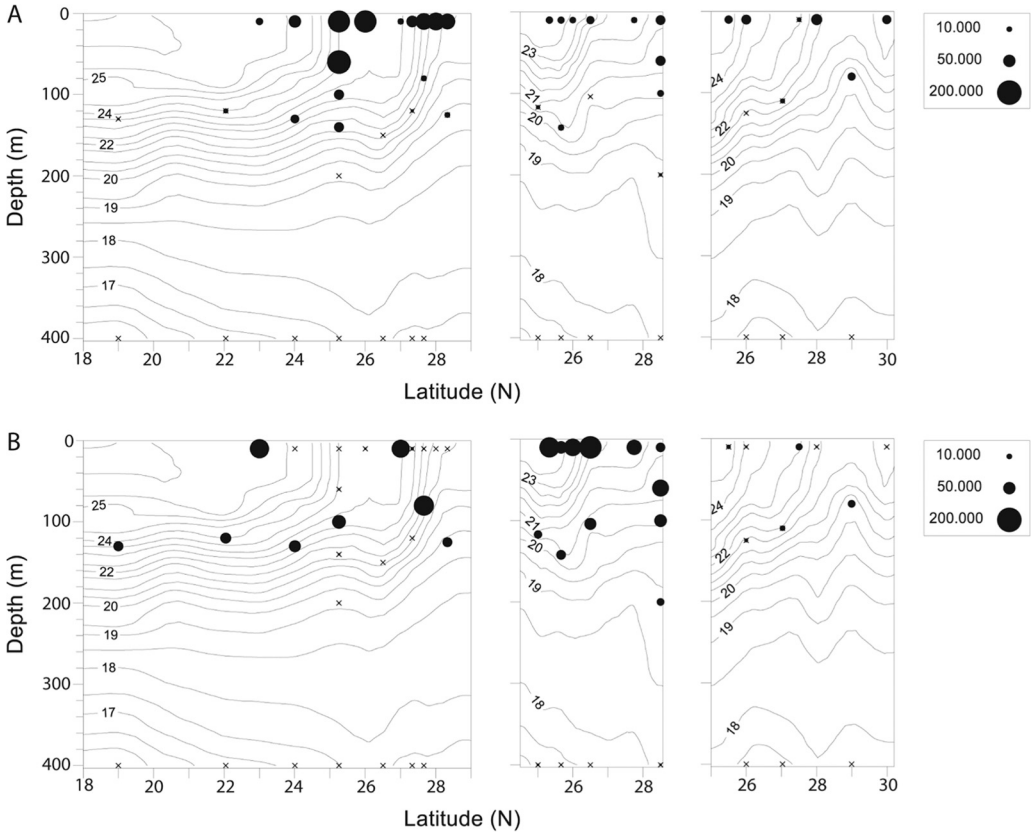

FIG 6 Distribution (cells ml^{-1}) of SAR86 (A) and *Roseobacter* (B) along transects 1, 2, and 3 (left to right) illustrated by the relative area of circles superimposed on temperature isotherms. Samples with very low abundances, i.e., nonvisible circles, are marked with "x."

rameters and subgroups within *Roseobacter* have previously been shown, for example, for the RCA cluster (28, 29). However, since predominance of *Roseobacter* clusters other than RCA have previously been found in the Sargasso Sea (72, 73), we applied a more general primer pair targeting *Roseobacter*, but with the risk of coamplification of the *Rhodobacter* group (15). Our limited ability to explain variations in *Roseobacter* abundance may therefore be linked to the amplification of a broad taxonomic group containing multiple phenotypes occupying different ecological niches (74, 75). However, it might also be that in contrast to the other taxa quantified, we may have missed the most important variables for explaining the *Roseobacter* abundance. Indeed, an earlier study explained 98.1% of the RCA cluster abundance by including Chl *a*, bacterial abundance, phaeopigments, particulate organic matter, suspended particulate matter, and salinity change of the trajectories of the water mass (29), factors which except for Chl *a* were not included in our analysis.

Finally, water mass, defined by temperature and salinity characteristics, had limited importance for explaining the abundance of the taxonomical groups when the other environmental variables measured were included in the models. The *Roseobacter* clade, which was the only taxon significantly correlated with water mass, showed highest abundance in the subtropical mode water (Fig. 6B). Only 14% of the variation in *Roseobacter* abundances was explained by water mass, possibly representing one or several specific variables not included in the regression model. Previous studies showing correlations by water mass do not account for environmental variables at the same time. Thus, it may be the case that these water mass correlations are due to the covarying environmental parameters, rather than water mass history. Alternatively, linkages between water mass and bacterial community composition are based on presence or absence (9) or relative abundances (8, 10, 11) of finer-scale OTUs. In contrast, our study was based on the total abundance of broad bacterial taxa, which are probably not affected by dispersal limitation between water masses because they are ubiquitous. Moreover, most of our samples are from the upper 200 m, where community composition is presumably driven by current environmental variables, whereas the circulation history of water masses may be more important for deep communities (4). Indeed, oceanic studies reporting a coupling between water mass and bacterial community composition often include more samples from bathypelagic layers (8–10).

In the present study, more than 50% of the variation in abundance for a variety of the broad bacterial groups could be predicted from environmental variables, like Chl *a* and depth. This indicates relevant ecological coherence at these high taxonomical levels (76). Hence, although many functional traits of microbes are phylogenetically dispersed at a much finer level than targeted in this study (77), our results indicate that even broad taxonomic categories of marine bacterioplankton may represent ecologically meaningful units sharing general life strategies or traits that distinguish them from other taxa (76, 77). This suggests hope for modeling marine bacterial communities, for which the tremendous fine-scale phylogenetic diversity poses a seemingly insur-

mountable problem (78). Further development of our model for the Sargasso Sea requires incorporation of temporal variation to examine whether dynamics of broad taxonomic groups are seasonally predictable, as has been shown for finer-scale taxa (79). Nevertheless, linkages of broad bacterioplankton taxa with a few easily quantified parameters indicate that modeling the abundance of these groups in the oligotrophic ocean may be a promising endeavor.

ACKNOWLEDGMENTS

We thank A. C. Martiny and two anonymous reviewers for helpful comments on an earlier version of the manuscript.

The cruise was supported by grants from Knud Højgaards Fond, the Danish Natural Sciences Research Council, Nordea Fond, Villum Kann Rasmussen Fond, and Elisabeth and Knud Petersens Fond as part of the Galathea 3 expedition under the auspices of Dansk Expeditions Fond. This work was supported by grant 621-20085385 from the Swedish Research Council to Åke Hagström, by grants 09-066396 and 11-105450 from The Danish Council for Independent Research, Natural Sciences, to L.R., and by visiting scientist grants from the Danish Agency for Science Technology and Innovation and the VELUX Foundation to L.R. and J.B.H.M.

REFERENCES

1. **Pommier T, Canbäck B, Riemann L, Boström KH, Simu K, Lundberg P, Tunlid A, Hagström Å.** 2007. Global patterns of diversity and community structure in marine bacterioplankton. Mol. Ecol. **16:**867–880. http://dx.doi.org/10.1111/j.1365-294X.2006.03189.x.
2. **Wietz M, Gram L, Jorgensen B, Schramm A.** 2010. Latitudinal patterns in the abundance of major marine bacterioplankton groups. Aquat. Microb. Ecol. **61:**179–189. http://dx.doi.org/10.3354/ame01443.
3. **Zinger L, Amaral-Zetter LA, Fuhrman JA, Horner-Devine MC, Huse SM, Welch DBM, Martiny JBH, Sogin ML, Boetius A, Ramette A.** 2011. Global patterns of bacterial beta-diversity in seafloor and seawater ecosystems. PLoS One **6:**e24570. http://dx.doi.org/10.1371/journal.pone.0024570.
4. **Ghiglione J, Galand PE, Pommier T, Pedrós-Alió C, Maas EW, Bakker K, Bertilsson S, Kirchman DL, Lovejoy C, Yager PL, Murray AE.** 2012. Pole-to-pole biogeography of surface and deep marine bacterial communities. Proc. Natl. Acad. Sci. U. S. A. **109:**17633–17638. http://dx.doi.org/10.1073/pnas.1208160109.
5. **Martiny JB, Bohannan BJ, Brown JH, Colwell RK, Fuhrman JA, Green JL, Horner-Devine MC, Kane M, Krumins JA, Kuske CR, Morin PJ, Naeem S, Ovreas L, Reysenbach AL, Smith VH, Staley JT.** 2006. Microbial biogeography: putting microorganisms on the map. Nat. Rev. Microbiol. **4:**102–112. http://dx.doi.org/10.1038/nrmicro1341.
6. **Barberán A, Casamayor E.** 2010. Global phylogenetic community structure and β-diversity patterns in surface bacterioplankton metacommunities. Aquat. Microb. Ecol. **59:**1–10. http://dx.doi.org/10.3354/ame01389.
7. **Pommier T, Douzery EJP, Mouillot D.** 2012. Environment drives high phylogenetic turnover among oceanic bacterial communities. Biol. Lett. **8:**562–566. http://dx.doi.org/10.1098/rsbl.2011.0990.
8. **Galand PE, Potvin M, Casamayor EO, Lovejoy C.** 2010. Hydrography shapes bacterial biogeography of the deep Arctic Ocean. ISME J. **4:**564–576. http://dx.doi.org/10.1038/ismej.2009.134.
9. **Yokokawa T, De Corte D, Sintes E, Herndl G.** 2010. Spatial patterns of bacterial abundance, activity and community composition in relation to water masses in the eastern Mediterranean Sea. Aquat. Microb. Ecol. **59:**185–195. http://dx.doi.org/10.3354/ame01393.
10. **Agogué H, Lamy D, Neal PR, Sogin ML, Herndl GJ.** 2011. Water mass-specificity of bacterial communities in the North Atlantic revealed by massively parallel sequencing. Mol. Ecol. **20:**258–274. http://dx.doi.org/10.1111/j.1365-294X.2010.04932.x.
11. **Hahnke RL, Probian C, Fuchs BM, Harder J.** 2013. Variations in pelagic bacterial communities in the North Atlantic Ocean coincide with water bodies. Aquat. Microb. Ecol. **71:**131–140. http://dx.doi.org/10.3354/ame01668.
12. **Gómez-Pereira PR, Fuchs BM, Alonso C, Oliver MJ, van Beusekom JE, Amann R.** 2010. Distinct flavobacterial communities in contrasting water masses of the North Atlantic Ocean. ISME J. **4:**472–487. http://dx.doi.org/10.1038/ismej.2009.142.
13. **Pearson RG, Dawson TP.** 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? Glob. Ecol. Biogeogr. **12:**361–371. http://dx.doi.org/10.1046/j.1466-822X.2003.00042.x.
14. **Partensky F, Hess WR, Vaulot D.** 1999. *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. Microbiol. Mol. Biol. Rev. **63:**106–127.
15. **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. Microb. Ecol. **24:**117–127. http://dx.doi.org/10.3354/ame024117.
16. **Buitenhuis ET, Li WKW, Vaulot D, Lomas MW, Landry MR, Partensky F, Karl DM, Ulloa O, Campbell L, Jacquet S, Lantoine F, Chavez F, Macias D, Gosselin M, McManus GB.** 2012. Picophytoplankton biomass distribution in the global ocean. Earth Syst. Sci. Data **4:**37–46. http://dx.doi.org/10.5194/essd-4-37-2012.
17. **Flombaum P, Gallegos JL, Gordillo RA, Rincón J, Zabala LL, Jiao N, Karl DM, Li WKW, Lomas MW, Veneziano D, Vera CS, Vrugt JA, Martiny AC.** 2013. Present and future global distribution of the marine cyanobacteria *Prochlorococcus* and *Synechococcus*. Proc. Natl. Acad. Sci. U. S. A. **110:**9824–9829. http://dx.doi.org/10.1073/pnas.1307701110.
18. **Buitenhuis ET, Li WKW, Lomas MW, Karl DM, Landry MR, Jacquet S.** 2012. Picoheterotroph (*Bacteria* and *Archaea*) biomass distribution in the global ocean. Earth Syst. Sci. Data **4:**101–106. http://dx.doi.org/10.5194/essd-4-101-2012.
19. **Salihoglu B, Hofmann EE.** 2007. Simulations of phytoplankton species and carbon production in the equatorial Pacific Ocean. 2. Effects of physical and biogeochemical processes. J. Mar. Res. **65:**275–300.
20. **Follows MJ, Dutkiewicz S, Grant S, Chisholm SW.** 2007. Emergent biogeography of microbial communities in a model ocean. Science **315:**1843–1846. http://dx.doi.org/10.1126/science.1138544.
21. **Amend AS, Seifert KA, Bruns TD.** 2010. Quantifying microbial communities with 454 pyrosequencing: does read abundance count? Mol. Ecol. **19:**5555–5565. http://dx.doi.org/10.1111/j.1365-294X.2010.04898.x.
22. **Pinto AJ, Raskin L.** 2012. PCR biases distort bacterial and archaeal community structure in pyrosequencing datasets. PLoS One **7:**e43093. http://dx.doi.org/10.1371/journal.pone.0043093.
23. **Schattenhofer M, Fuchs BM, Amann R, Zubkov MV, Tarran GA, Pernthaler J.** 2009. Latitudinal distribution of prokaryotic picoplankton populations in the Atlantic Ocean. Environ. Microbiol. **11:**2078–2093. http://dx.doi.org/10.1111/j.1462-2920.2009.01929.x.
24. **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. http://dx.doi.org/10.1126/science.1118052.
25. **Zinser ER, Johanson 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. http://dx.doi.org/10.4319/lo.2007.52.5.2205.
26. **Tai V, Palenik B.** 2009. Temporal variation of *Synechococcus* clades at a coastal Pacific Ocean monitoring site. ISME J. **3:**903–915. http://dx.doi.org/10.1038/ismej.2009.35.
27. **Giebel HA, Brinkhoff T, Zwisler W, Selje N, Simon M.** 2009. Distribution of *Roseobacter* RCA and SAR11 lineages and distinct bacterial communities from the subtropics to the Southern Ocean. Environ. Microbiol. **11:**2164–2178. http://dx.doi.org/10.1111/j.1462-2920.2009.01942.x.
28. **Giebel H, Kalhoefer D, Lemke A, Thole S, Gahl-Janssen R, Simon M, Brinkhoff T.** 2011. Distribution of *Roseobacter* RCA and SAR11 lineages in the North Sea and characteristics of an abundant RCA isolate. ISME J. **5:**8–19. http://dx.doi.org/10.1038/ismej.2010.87.
29. **Sperling M, Giebel H, Rink B, Grayek S, Staneva J, Stanev E, Simon M.** 2012. Differential effects of hydrographic and biochemical properties on the SAR11 clade and the *Roseobacter* RCA cluster in the North Sea. Aquat. Microb. Ecol. **67:**25–34. http://dx.doi.org/10.3354/ame01580.
30. **Campbell BJ, Yu L, Straza TRA, Kirchman DL.** 2009. Temporal changes in bacterial rRNA and rRNA genes in Delaware (USA) coastal waters. Aquat. Microb. Ecol. **57:**123–135. http://dx.doi.org/10.3354/ame01335.
31. **Voorhis AD, Hersey JB.** 1964. Oceanic thermal fronts in the Sargasso Sea. J. Geophys. Res. **69:**3809–3814. http://dx.doi.org/10.1029/JZ069i018p03809.
32. **Riemann L, Steward GF, Azam F.** 2000. Dynamics of bacterial community composition and activity during a mesocosm diatom bloom. Appl.

Environ. Microbiol. **66:**578–587. http://dx.doi.org/10.1128/AEM.66.2.578-587.2000.
33. **Boström HK, Simu K, Hagström Å, Riemann L.** 2004. Optimization of DNA extraction for quantitative marine bacterioplankton community analysis. Limnol. Oceanogr. Methods **2:**365–373. http://dx.doi.org/10.4319/lom.2004.2.365.
34. **Riemann L, Nielsen TG, Kragh T, Richardson K, Parner H, Jakobsen HH, Munk P.** 2011. Distribution and production of plankton communities in the subtropical convergence zone of the Sargasso Sea. I. Phytoplankton and bacterioplankton. Mar. Ecol. Prog. Ser. **426:**57–70. http://dx.doi.org/10.3354/meps09001.
35. **Giovannoni SJ.** 1991. The polymerase chain reaction, p 177–201. *In* Stackebrandt E, Goodfellow M (ed), Nucleic acid techniques in bacterial systematics. John Wiley & Sons, New York, NY.
36. **Haas B, Gevers D, Earl A, Feldgarden M, Ward D, Giannokous G, Ciulla D, Tabbaa D, Highlander S, Sodergren E, Methe B, DeSantis T, Petrosino J, Knight R, Birren B.** 2011. Chimeric 16S rRNA sequences formation and detection in Sanger and 454-pyrosequencing PCR amplicons. Genome Res. **21:**494–504. http://dx.doi.org/10.1101/gr.112730.110.
37. **Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar Buchner A, Lai T, Steppi S, Jobb G, Förster W, Brettske I, Gerber S, Ginhart AW, Gross O, Grumann S, Hermann S, Joat R, König A, Liss T, Lüssman R.** 2004. ARB: a software environment for sequence data. Nucleic Acids Res. **32:**1363–1371. http://dx.doi.org/10.1093/nar/gkh293.
38. **Zhu F, Massana R, Not F, Marie D, Vaulot D.** 2005. Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. FEMS Microbiol. Ecol. **52:**79–92. http://dx.doi.org/10.1016/j.femsec.2004.10.006.
39. **Goebel NL, Turk KA, Achilles KM, Paerl R, Hewson I, Morrison AE, Montoya JP, Edwards CA, Zehr JP.** 2010. Abundance and distribution of major groups of diazotrophic cyanobacteria and their potential contribution to N fixation in the tropical Atlantic Ocean. Environ. Microbiol. **12:**3272–3289. http://dx.doi.org/10.1111/j.1462-2920.2010.02303.x.
40. **Zhang T, Fang HH.** 2006. Applications of real-time polymerase chain reaction for quantification of microorganisms in environmental samples. Appl. Microbiol. Biotechnol. **70:**281–289. http://dx.doi.org/10.1007/s00253-006-0333-6.
41. **Clarke KR, Gorley RN.** 2006. PRIMER v6: user manual/tutorial. PRIMER-E, Plymouth, United Kingdom.
42. **Development Core Team R.** 2008. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
43. **Fox J.** 1997. Applied regression analysis, linear models, and related methods. Sage Publications, Thousand Oaks, CA.
44. **Grömping U.** 2006. Relative importance for linear regression in R: the package relaimpo. J. Stat. Softw. **17:**1–27. http://dx.doi.org/10.1360/jos170001.
45. **Morris RM, Vergin KL, Cho J, Rappé 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. http://dx.doi.org/10.4319/lo.2005.50.5.1687.
46. **Treusch AH, Vergin KL, Finlay LA, Donatz MG, Burton RM, Carlson CA, Giovannoni SJ.** 2009. Seasonality and vertical structure of microbial communities in an ocean gyre. ISME J. **3:**1148–1163. http://dx.doi.org/10.1038/ismej.2009.60.
47. **Giovannoni SJ, Vergin KL.** 2012. Seasonality in the ocean microbial communities. Science **335:**671–676. http://dx.doi.org/10.1126/science.1198078.
48. **Herndl GJ, Reinthaler T, Teira E, van Aken H, Veth C, Pernthaler A, Pernthaler J.** 2005. Contribution of *Archaea* to total prokaryotic production in the deep Atlantic Ocean. Appl. Environ. Microbiol. **71:**2303–2309. http://dx.doi.org/10.1128/AEM.71.5.2303-2309.2005.
49. **Larsen PE, Field D, Gilbert JA.** 2012. Predicting bacterial community assemblages using an artificial neural network approach. Nat. Methods **9:**621–625. http://dx.doi.org/10.1038/nmeth.1975.
50. **Acinas SG, Antón J, Rodríguez-Valera F.** 1999. Diversity of free-living and attached bacteria in offshore western Mediterranean waters as depicted by analysis of genes encoding 16S rRNA. Appl. Environ. Microbiol. **65:**514–522.
51. **González JM, Simó R, Massana R, Covert JS, Casamayor EO, Pedrós-Alió C, Moran MA.** 2000. Bacterial community structure associated with a dimethylsulfoniopropionate-producing North Atlantic algal bloom. Appl. Environ. Microbiol. **66:**4237–4246. http://dx.doi.org/10.1128/AEM.66.10.4237-4246.2000.
52. **Suzuki MT, Preston CM, Béjà O, de la Torre JR, Steward GF, DeLong EF.** 2004. Phylogenetic screening of ribosomal RNA gene-containing clones in bacterial artificial chromosome (BAC) libraries from different depths in Monterey Bay. Microb. Ecol. **48:**473–488. http://dx.doi.org/10.1007/s00248-004-0213-5.
53. **Buchan A, González JM, Moran MA.** 2005. Overview of the marine *Roseobacter* lineage. Appl. Environ. Microbiol. **71:**5665–5677. http://dx.doi.org/10.1128/AEM.71.10.5665-5677.2005.
54. **Lauro FM, McDougald D, Thomas T, Williams TJ, Egan S, Rice S, DeMaere MZ, Ting L, Ertan H, Johnson J, Ferriera S, Lapidus A, Anderson I, Kyrpides N, Munk AC, Detter C, Han CS, Brown MV, Robb FT, Kjelleberg S, Cavicchioli R.** 2009. The genomic basis of trophic strategy in marine bacteria. Proc. Natl. Acad. Sci. U. S. A. **106:**15527–15533. http://dx.doi.org/10.1073/pnas.0903507106.
55. **Yooseph S, Nealson KH, Risch DB, McCrow JP, Ferriera S, Beeson K, Williams SJ, Tovchigrechko A, Allen AE, Zeigler LA, Sutton G, Eisenstadt E, Rogers YH, Friedman R, Frazier M, Venter JC.** 2010. Genomic and functional adaptation in surface ocean planktonic prokaryotes. Nature **468:**60–67. http://dx.doi.org/10.1038/nature09530.
56. **Pakulski JD, Benner R.** 1994. Abundance and distribution of carbohydrates in the ocean. Limnol. Oceanogr. **39:**930–940. http://dx.doi.org/10.4319/lo.1994.39.4.0930.
57. **McCarthy M, Hedges J, Benner R.** 1995. Major biochemical composition of dissolved high molecular weight organic matter in seawater. Mar. Chem. **55:**281–297.
58. **Benner R, Biddanda B.** 1998. Photochemical transformations of surface and deep marine dissolved organic matter: effects on bacterial growth. Limnol. Oceanogr. **43:**1373–1378. http://dx.doi.org/10.4319/lo.1998.43.6.1373.
59. **Nagata T.** 2008. Organic matter-bacteria interactions in seawater, p 207–242. *In* Kirchman DL (ed), Microbial ecology of the oceans, 2nd ed. John Wiley & Sons, Inc, Hoboken, NJ.
60. **Moran MA, Zepp RG.** 1997. Role of photoreactions in the formation of biologically labile compounds from dissolved organic matter. Limnol. Oceanogr. **42:**1307–1316. http://dx.doi.org/10.4319/lo.1997.42.6.1307.
61. **Campbell L, Liu H, 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. I **44:**167–192. http://dx.doi.org/10.1016/S0967-0637(96)00102-1.
62. **Partensky F, Blanchot J, Vaulot D.** 1999. Differential distribution and ecology of *Prochlorococcus* and *Synechococcus* in oceanic waters: a review. Bull. Inst. Oceanogr. **19:**457–475.
63. **Dupont CL, Rusch DB, Yooseph S, Lombardo M, Richter RA, Valas R, Novotny M, Yee-Greenbaum J, Selengut JD, Haft DH, Halpern AL, Lasken RS, Nealson K, Friedman R, Venter JG.** 2012. Genomic insights to SAR86, an abundant uncultivated marine bacterial lineage. ISME J. **6:**1186–1199. http://dx.doi.org/10.1038/ismej.2011.189.
64. **Béjà O, Aravind L, Koonin EV, Suzuki MT, Hadd A, Nguyen LP, Jovanovich SB, Gates CM, Feldman RA, Spudich JL, Spudich EN, DeLong EF.** 2000. Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. Science **289:**1902–1906. http://dx.doi.org/10.1126/science.289.5486.1902.
65. **DeLong EF, Béjà O.** 2010. The light-driven proton pump proteorhodopsin enhances bacterial survival during tough times. PLoS Biol. **8:**e1000359. http://dx.doi.org/10.1371/journal.pbio.1000359.
66. **Martiny AC, Huang Y, Li W.** 2009. Occurrence of phosphate acquisition genes in *Prochlorococcus* cells from different ocean regions. Environ. Microbiol. **11:**1340–1347. http://dx.doi.org/10.1111/j.1462-2920.2009.01860.x.
67. **Calvo-Díaz A, Morán XAG, Nogueira E, Bode A, Varela M.** 2004. Picoplankton community structure along the northern Iberian continental margin in late winter-early spring. J. Plankton Res. **26:**1069–1081. http://dx.doi.org/10.1093/plankt/fbh098.
68. **Jiao N, Yang Y, Hong N, Ma Y, Harada S, Koshikawa H, Watanabe M.** 2005. Dynamics of autotrophic picoplankton and heterotrophic bacteria in the East China Sea. Cont. Shelf Res. **25:**1265–1279. http://dx.doi.org/10.1016/j.csr.2005.01.002.
69. **Fengfeng L, Xiuren N, Chenggang L, Xiaobo N, Qiang H.** 2012. Picoplankton abundance and biomass in the East China Sea during autumn and winter. Oceanology **52:**48–59. http://dx.doi.org/10.1134/S0001437011060026.
70. **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. http://dx.doi.org/10.1038/ismej.2008.117.

71. **Brown MV, Lauro FM, DeMaere MZ, Muir L, Wilkins D, Thomas T, Riddle MJ, Fuhrman JA, Andrews-Pfannkoch C, Hoffman JM, McQuaid JB, Allen A, Rintoul SR, Cavicchioli R.** 2012. Global biogeography of SAR11 marine bacteria. Mol. Syst. Biol. **8:**595. http://dx.doi.org/10.1038/msb.2012.28.
72. **Moran MA, Buchan A, González JM, Heidelberg JF, Whitman WB, Kiene RP, Henriksen JR, King GM, Belas R, Fuqua C, Brinkac L, Lewis M, Johri S, Weaver B, Pai G, Eisen JA, Rahe E, Sheldon WM, Ye W, Miller TR, Carlton J, Rasko DA, Paulsen IT, Ren Q, Daugherty SC, Deboy RT, Dodson RJ, Durkin AS, Madupu R, Nelson WC, Sullivan SA, Rosovitz MJ, Haft DH, Selengut J, Ward N.** 2004. Genome sequence of *Silicibacter pomeroyi* reveals adaptations to the marine environment. Nature **432:**910–913. http://dx.doi.org/10.1038/nature03170.
73. **Venter JG, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA, Wu D, Paulsen I, Nelson KE, Nelson W, Fouts DE, Levy S, Knap AH, Lomas MW, Nealson K, White O, Peterson J, Hoffman J, Parsons R, Baden-Tillson H, Pfannkoch C, Rogers YH, Smith HO.** 2004. Environmental genome shotgun sequencing of the Sargasso Sea. Science **304:**66–74. http://dx.doi.org/10.1126/science.1093857.
74. **Moran MA, Belas R, Schell MA, González JM, Sun F, Sun S, Binder BJ, Edmonds J, Ye W, Orcutt B, Howard EC, Meile C, Palefsky W, Goesmann A, Ren Q, Paulsen I, Ulrich LE, Thompson LS, Saunders E, Buchan A.** 2007. Ecological genomics of marine roseobacters. Appl. Environ. Microbiol. **73:**4559–4569. http://dx.doi.org/10.1128/AEM.02580-06.
75. **Brinkhoff T, Giebel H, Simon M.** 2008. Diversity, ecology, and genomics of the *Roseobacter* clade: a short overview. Arch. Microbiol. **189:**531–539. http://dx.doi.org/10.1007/s00203-008-0353-y.
76. **Philippot L, Bru D, Saby NP, Cuhel J, Arrouays D, Simek M, Hallin S.** 2009. Spatial patterns of bacterial taxa in nature reflect ecological traits of deep branches of the 16S rRNA bacterial tree. Environ. Microbiol. **11:**3096–3104. http://dx.doi.org/10.1111/j.1462-2920.2009.02014.x.
77. **Martiny AC, Treseder K, Pusch G.** 2013. Phylogenetic conservatism of functional traits in microorganisms. ISME J. **7:**830–838. http://dx.doi.org/10.1038/ismej.2012.160.
78. **Hood RR, Laws EA, Follows MJ, Siegel DA.** 2007. Modeling and prediction of marine microbial populations in the genomic era. Oceanography **20:**155–165. http://dx.doi.org/10.5670/oceanog.2007.61.
79. **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. U. S. A. **103:**13104–13109. http://dx.doi.org/10.1073/pnas.0602399103.
80. **Lane D.** 1991. 16s/23s rRNA sequencing, p 115–175. *In* Stackebrandt E, Goodfellow M (ed), Nucleic acid techniques in bacterial systematics. Wiley & Sons, New York, NY.
81. **Muyzer G, de Waal EC, Uitterlinden AG.** 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl. Environ. Microbiol. **59:**695–700.
82. **Manz W, Amann R, Ludwig W, Vancanneyt M, Schleifer K-H.** 1996. Application of a suite of 16S rRNA-specific oligonucleotide probes designed to investigate bacteria of the phylum cytophaga-flavobacter-bacteroides in the natural environment. Microbiology **142:**1097–1106. http://dx.doi.org/10.1099/13500872-142-5-1097.
83. **Mühling M, Woolven-Allen J, Murrell JC, Joint I.** 2008. Improved group-specific PCR primers for denaturing gradient gel electrophoresis analysis of the genetic diversity of complex microbial communities. ISME J. **2:**379–392. http://dx.doi.org/10.1038/ismej.2007.97.
84. **Morales SE, Holben WE.** 2009. Empirical testing of 16S rRNA gene PCR primer pairs reveals variance in target specificity and efficacy not suggested by in silico analysis. Appl. Environ. Microbiol. **75:**2677–2683. http://dx.doi.org/10.1128/AEM.02166-08.
85. **Takai K, Horikoshi K.** 2000. Rapid detection and quantification of members of the archaeal community by quantitative PCR using fluorogenic probes. Appl. Environ. Microbiol. **66:**5066–5072. http://dx.doi.org/10.1128/AEM.66.11.5066-5072.2000.