Variations in pelagic bacterial communities in the North Atlantic Ocean coincide with water bodies

Richard L. Hahnke[1], Christina Probian[1], Bernhard M. Fuchs[2], Jens Harder[1,*]

[1]Department of Microbiology and [2]Department of Molecular Ecology, Max Planck Institute for Marine Microbiology, Celsiusstr. 1, 28359 Bremen, Germany

ABSTRACT: Physical and chemical characteristics define oceanographic regions. The potential for a distinct biogeography of bacterial communities in these oceanic provinces was studied in epipelagic and upper mesopelagic water bodies of the North Atlantic Ocean by terminal restriction fragment length polymorphism (T-RFLP) analysis of bacterial 16S rRNA genes and flow cytometry. Water samples from 67° N to 34° N along the 30° W meridian contained epipelagic populations of *Synechococcus* in the north and *Prochlorococcus* in the south. Bacterial communities were generally more diverse in phototrophic layers above the pycnocline. Communities significantly differed in the epipelagic zone along the latitudinal transect through the different oceanic provinces and between the epipelagic and the upper mesopelagic zone. Differences in the T-RFLP patterns coincided well with differences in the physico-chemical conditions of the sampling sites. Changes in bacterial communities were traced to characteristic terminal restriction fragments (TRFs). *In silico* assignments of phylogenetic groups to TRFs, e.g. populations of high-light and low-light ecotypes of *Prochlorococcus*, supported our T-RFLP analysis of bacterial communities. Distinct bacterial communities in water bodies of the North Atlantic Ocean hosted different bacterial populations, which may serve as biological markers for oceanic provinces.

KEY WORDS: *Synechococcus* · *Prochlorococcus* · North Atlantic Ocean · Gulf Stream · Canonical correspondence analysis · T-RFLP · 16S rRNA gene fingerprint

INTRODUCTION

The open ocean harbours a diversity of microorganisms, which have often a regional distribution. For example, unicellular cyanobacteria affiliating with *Synechococcu*s and *Prochlorococcus* prevail in different regions of the oceans (Li 1994, Liu et al. 1997, Veldhuis et al. 1997). The variety of habitats results from annual seasonal changes, intense atmospheric events, thermohaline circulation and currents throughout the ocean (Platt & Sathyendranath 1999, Teeling et al. 2012). Longhurst et al. (1995) partitioned the ocean into 56 ecological provinces based on physical forcing and provided static definitions of the province boundaries. These ecological provinces are regions defined by physico-chemical (e.g. tem-

perature, salinity and bathymetry) and biological (e.g. chlorophyll *a* concentration and vertical distribution of bacterial communities) characteristics and by a common history (Emery & Meincke 1986, Devred et al. 2007). Because ocean surface color significantly correlates with water column-integrated chlorophyll concentrations, photic depth and nutrients, ecological provinces can be discriminated by a global time series of satellite ocean color and sea surface temperature (Esaias et al. 2000, Oliver & Irwin 2008).

The North Atlantic Current and its prolongation, the North Atlantic Drift Current, divide the North Atlantic Ocean into a northern and a southern part. Two branches extend at 38° N, 44° W northeastward along the continental slope to the pole and southeastward along the continental slope feeding the current

around the North Atlantic Gyre (Mann 1967). Along the 30° W meridian from 67° N to 34° N, the North Atlantic Ocean contains 9 water bodies in 4 Longhurstian provinces (see Fig. S1 in Supplement 1 at www.int-res.com/articles/suppl/a071p131_supp/): 1 Boreal Polar (BPLR), 4 Atlantic Arctic (ARCT), 2 North Atlantic Drift (NADR) and 2 North Atlantic Subtropical Gyre (NAST) (Longhurst et al. 1995, Gómez-Pereira et al. 2010). Physical (temperature and salinity), chemical (nutrients) and biological data (chlorophyll *a*, picoplankton, nanophytoplankton and enyzme activities) clearly indicated gradients along the transect, from cold and nutrient-rich water bodies in the north to warm, oligotrophic water bodies in the south (Gómez-Pereira et al. 2010, Schattenhofer et al. 2011, Arnosti et al. 2012). Coincidence of ecological provinces in surface waters of the North Atlantic Ocean and local bacterial populations was recently shown for *Flavobacteria* clades (Gómez-Pereira et al. 2010) and picoplanktonic populations (Schattenhofer et al. 2011).

To date, a fine-scale characterization of the bacterial diversity within water bodies and with depth has not been reported for the North Atlantic Ocean. A first massive sequencing study had 3 samples from the North Atlantic Ocean and many from the Central Atlantic Ocean (Agogué et al. 2011). We applied terminal restriction fragment analyses (T-RFLP) that has less taxonomic resolution but is well suited for diversity studies on many samples (Forney et al. 2004). We hypothesized a strong correlation of bacterial communities with water bodies and conducted a cruise from 67° N to 34° N along a latitudinal gradient along 30° W, thus north of cruises of the Atlantic meridional transect program (Aiken et al. 2000). Epipelagic and upper mesopelagic bacterial communities were investigated by T-RFLP of 16S rRNA gene amplicons and flow cytometry counting of *Prochlorococcus* and *Synechococcus* populations.

MATERIALS AND METHODS

Sampling

Water samples were obtained on the RV 'Maria S. Merian' during the VISION cruise MSM03/1 (September 2006) with a CTD rosette equipped with 24 Niskin bottles (Fig. S1 in Supplement 1, Table S1 in Supplement 2 at www.int-res.com/articles/suppl/a071 p131_supp/). At each depth, seawater aliquots were sampled in triplicate: the biomass of a 200 ml aliquot was concentrated on a 0.2 µm Isopore filter with a di-

ameter of 45 mm (Millipore), and the filter was frozen immediately and stored at −20°C. Salinity, temperature and the concentrations of phosphate, ammonium, nitrite and nitrate were taken from Gómez-Pereira et. al. (2010). The density σ_θ was calculated using the equation of state for sea water (Stewart 2008). The cell numbers of *Synechococcus*, *Prochlorococcus* and total bacterial communities were determined by flow cytometric analyses as described by Tarran et al. (2006).

DNA extraction and T-RFLP

From each station depth, 3 biological replicates were analyzed. Genomic DNA was isolated from half a filter, representing 100 ml water sample, based on a protocol of Boström et al. (2004). The filter half was placed in a 2.2 ml sample vial and extracted with 525 µl lysis buffer and 11 µl lysozyme (50 mg ml^{-1}) for 30 min at 37°C in an overhead shaker. After addition of 60 µl 10% SDS and 3 µl proteinase K (20 mg ml^{-1}), the extraction was continued for 12 h at 55°C in an overhead shaker. The supernatant was transferred and incubated together with 100 µl isopropanol for 1 h at room temperature. The DNA was precipitated with 15000 × *g* for 30 min at 4°C. The pellet was washed with 100 µl cold ethanol, precipitated a second time and air dried. The DNA was dissolved in 50 µl water, and quantification yielded 10 to 50 ng genomic DNA per sample. Amplification of the partial 16S rRNA gene was performed with the fluorescently labelled primers 27F (FAM, 5′-AGA GTT TGA TYM TGG CTC AG-3′) and 907R (HEX, 5′-CCG TCA ATT CCT TTR AGT TT-3′), targeting all bacteria (Muyzer et al. 1995). The PCR reaction contained 12.5 µl PCR Master Mix (Promega), 4 µM of forward and of reverse primer, and 1 to 5 ng DNA template in 25 µl. The cycle program was 95°C for 1 min, 33 cycles of 95°C for 1 min, 60°C for 1 min and 72°C for 3 min, followed by 60°C for 60 min. PCR amplicons were purified on Sephadex columns (Sephadex™ G-50 Superfine, Amersham Biosciences). Approximately 25 ng of PCR amplicon were digested in a total volume of 10 µl using 5 U of the restriction enzyme *Alu* I (Fermentas) at 37°C for 3 h, followed by heat inactivation at 65°C for 30 min. After purification on Sephadex columns, terminal restriction fragments (TRFs) were detected on an ABI Prism 3130 XL Genetic Analyzer (Applied Biosystems) equipped with an 80 cm capillary, a POP-7 polymer and the filter set D (Filter DS-30). The ROX-labelled MapMarker® 1000 (Eurogentec) served as a size standard between 50 and 1000 bp.

TRF pattern analyses

T-RFLP patterns were inspected manually with the software GenMapper 3.7 (Applied Biosystems). The fluorescence intensity threshold was set to 20 units, and the fragments with a size between 50 and 1000 nucleotides were identified and sized (Local Southern, normalization within each run, sum of signals) with the internal size standard. For comparative analyses, the individual patterns were processed by applying the interactive binner (Ramette 2009). The binning size was 1 nucleotide, and the binning shift was 0.5 nucleotides. Because each TRF was named by its start of the binning window, we added 0.5 bases to the TRF length when naming TRFs. The resulting pattern with normalized peak areas (relative fluorescence intensity [RFI] of 100% corresponds to the sum of the peak areas in each T-RFLP profile) were visualized in rank versus cumulated abundance curves with the *k*-dominance plot in PRIMER-E (v.6, PRIMER-E) (Clarke 1993). Inspection by GenMapper and the *k*-dominance plots served to remove outliers within the triplicates and identify the final T-RFLP data set (Fig. S2 in Supplement 1).

Constrained (canonical) correspondence analysis (CCA) (TerBraak 1986) was used to relate the compositional variation in the bacterial community of sampling sites as χ^2 (chi-squared) distances to the observed environmental variation by canonical correlations and was performed as a weighted linear mapping without including depth and longitude. For comparability, a sampling site-similarity matrix was generated using the Bray-Curtis coefficient by comparing the RFI of each TRF with regard to every pairwise combination of all stations and depth, with 999 permutations. Non-parametric multivariate statistical analysis was performed using PRIMER-E and the R package VEGAN (v.1.8-3) (Dixon 2003). Visual comparisons between bacterial communities of predefined oceanic provinces (BPLR, ARCT, NADR and NAST) were explored by ordination using non-metric multidimensional scaling (nMDS) with 100 random restarts and 999 iterations. As a third method, hierarchical clustering was used to group the sampling sites. Visualization was performed by adding the information of the hierarchical tree into the nMDS plot. A consistent biplot was obtained in a fitting of the environmental conditions into the nMDS plot by applying the function envfit of the R package VEGAN with 1000 permutations and p-values < 0.001 but without including depth and longitude (Dixon 2003). Analysis of similarity (ANOSIM) in PRIMER-E was used to verify the significance of water body-specific clustering of bacterial communities by testing the null hypothesis that bacterial communities from the same water body were more similar to each other than to bacterial communities in different water bodies. Differences in bacterial communities between water bodies in the epipelagic zone were tested by 1-way ANOSIM, and differences in epipelagic and upper mesopelagic bacterial communities were tested by 2-way crossed ANOSIM. ANOSIM statistics were based on the same sampling site-similarity matrix of Bray-Curtis coefficients as used for the nMDS and were computed with 999 permutations.

To identify which TRF formed a strong gradient along the latitude gradient and into the depth, we used principal components analysis (PCA), transforming the variable space (the RFI of each TRF of each sampling site) into its orthogonal principal components. Afterward, the eigenvectors of the TRF and principal component scores of the sampling sites were visualized. Similarity percentage analysis (SIMPER) was used to explore the significance of TRFs in water bodies. A significant TRF was defined as one with (1) an average RFI within the represented water body at least 2-fold greater than that in the other water bodies (relative abundance), (2) a ratio higher than one between the contribution to the average Bray-Curtis dissimilarity (average dissimilarity) between all pairs of sampling sites (1 within the represented oceanic provinces and 1 outside) and the standard deviation (SD) of those contributions (average contribution to difference) (3) and an RFI of ≥4% in at least 1 sampling site (absolute abundance). Corresponding peaks of the significant representative TRF were again inspected manually in the original T-RFLP pattern with GenMapper to confirm that the analyzed fluorescence signal was unaffected by neighbouring TRFs. Finally, the biogeographies of the TRFs were visualized in Ocean Data View (v.3.4.2, AWI) (Schlitzer 2002).

In silico prediction of the fragment size

16S rRNA gene sequences were retrieved with the ARB program (Ludwig et al. 2004) from the SILVA database (rel102ref, 391 167 bacterial sequences) (Pruesse et al. 2007) by targeting both T-RFLP primers with 0 to 2 mismatches. Four sets of sequences were generated: (1) 135 761 sequences of all phyla, (2) 87 of 361 sequences of *Synechococcus*, (3) 382 of 944 sequences of *Prochlorococcus* and (4) 233 of 885 sequences of a *Bacteroidetes*-specific clone library

retrieved from Gómez-Pereira et al. (2010). These sequences were trimmed to the T-RFLP amplicon size. The script TRFragCalc — written in MATLAB (v.2.9.0.529 R2009b, The MathWorks) and available on enquiry and at www.mpi-bremen.de/en/Software _-_TRFragCalc.html — was applied to imported sequences to identify the restriction recognition site and to calculate the resulting T-RFLP fragments. Starting with an *in silico* fragment, e.g. iTRF_128nt for *Synechococcus*, we investigated the distribution of TRFs in a range of ±5 nucleotides, e.g. TRF_123nt to TRF_133nt. This window of 10 nucleotides was necessary because an absolute determination of the length of TRFs with capillary electrophoresis is currently not possible (Bruland et al. 1999, Hahn et al. 2001, Olejniczak et al. 2005).

RESULTS

Oceanographic changes and the bacterial diversity determined by T-RFLP

We sampled the North Atlantic Ocean at depths between 20 m and 500 m along the 30° W meridian from the productive cold Greenland current (66° 39′ N) across the cold north and warm south of the North Atlantic Current to the oligotrophic central Atlantic Ocean (34° 24′ N) (Fig. S1). T-RFLP is technically limited to maximal 860 TRFs. We observed a bacterial diversity of 467 terminal restriction fragments (TRFs) in all samples (γ diversity), with a 25% quantile, median, and 75% quantile of 58, 86 and 105 TRFs per sample (α diversity) respectively (Fig. S3 in Supplement 1). The Shannon diversity index (Shannon 1948) based on the relative abundance of the TRFs was large in the epipelagic zone, with a high diversity north and south of the North Atlantic Drift (Fig. 1).

The fragment patterns of the sampling sites were constrained by environmental conditions (salinity, conductivity, temperature and the concentration of dissolved oxygen, ammonium and nitrate) in a unimodal model. The CCA covered one third of the total variance (inertia = 2.07), reflected by a mean squared contingency coefficient of the constrained axes of 31% (inertia = 0.63). The CCA revealed a distribution of bacterial populations along a latitudinal gradient and with water depth (Fig. 2). Oxygen, temperature and salinity coincided with the first dimension, a proxy for latitude, whereas ammonium, nitrate and phosphate levels were proxies for the water depth. Bacterial communities from the

surface water were well separated from bacterial communities from water deeper than 80 m. The border between the mixed layer and stratified deeper waters was 80 to 100 m depth (Fig. S1). The ANOSIM test showed that differences between epipelagic and upper mesopelagic bacterial communities were significant (global R = 0.79, p < 0.001). A hierachical clustering at 48% similarity defined 3 groups of samples: (1) an epipelagic BPLR-ARCT cluster, (2) an epipelagic NADR-NAST cluster and (3) one common cluster of upper mesopelagic bacterial communities (Fig. S4 in Supplement 1). Samples from the Longhurstian provinces BPLR, ARCT, NADR and NAST formed individual groups in the CCA (Fig. 2). The 7 smaller clusters of surface water (Fig. 1) coincided with the water bodies defined by Gómez-Pereira et al. (2010). Within the provinces, bacterial communities of the water bodies ARCT1 (Stns 4 to 6) and ARCT2 (Stns 7 and 8) were significantly different (R = 0.27, p < 0.001). The NADR (Stns 10 to 13) showed a high variation, reflecting the dynamic environment of the ocean current. The geographical distance between Stns 10 and 11 was smaller than between Stns 11 and 12. In contrast, the dissimilarity of the bacterial community between Stns 10 and 11 was large compared to the dissimilarity between Stns 11 and 12 (Fig. 2). Stns 16 to 18 of the water bodies NAST1 and NAST2 could not be distinguished on the basis of the T-RFLP pattern. Stns 14 and 16 belonged to the water body NAST1, but the bacterial community of Stn 16 clustered with the communities of the water body NAST2, and the community of Stn 14 was separated from the community of Stn 16. The significant separation of bacterial communities in the BPLR, ARCT1 and ARC2 water bodies, compared to the more similar sampling sites in the NADR and NAST province, was reflected in the ANOSIM results. Overall, the differences between bacterial communities of the water bodies were significant (global R = 0.60, p < 0.001). Pairwise tests of epipelagic water bodies showed a significant different diversity in the water bodies BPLR1 and NAST2, ARCT2 and NADR1, and NADR2 and NAST1 (Table S2 in Supplement 2). Water bodies within the same province (e.g. NAST1 and NAST2) had a larger shared bacterial community.

To confirm the clustering of bacterial communities with water bodies along the latitudinal gradient and an independence from the CCA method (which uses χ^2 distances), the similarity between the bacterial communities of individual sampling sites (β diversity) was calculated with the Bray-Curtis similarity (based
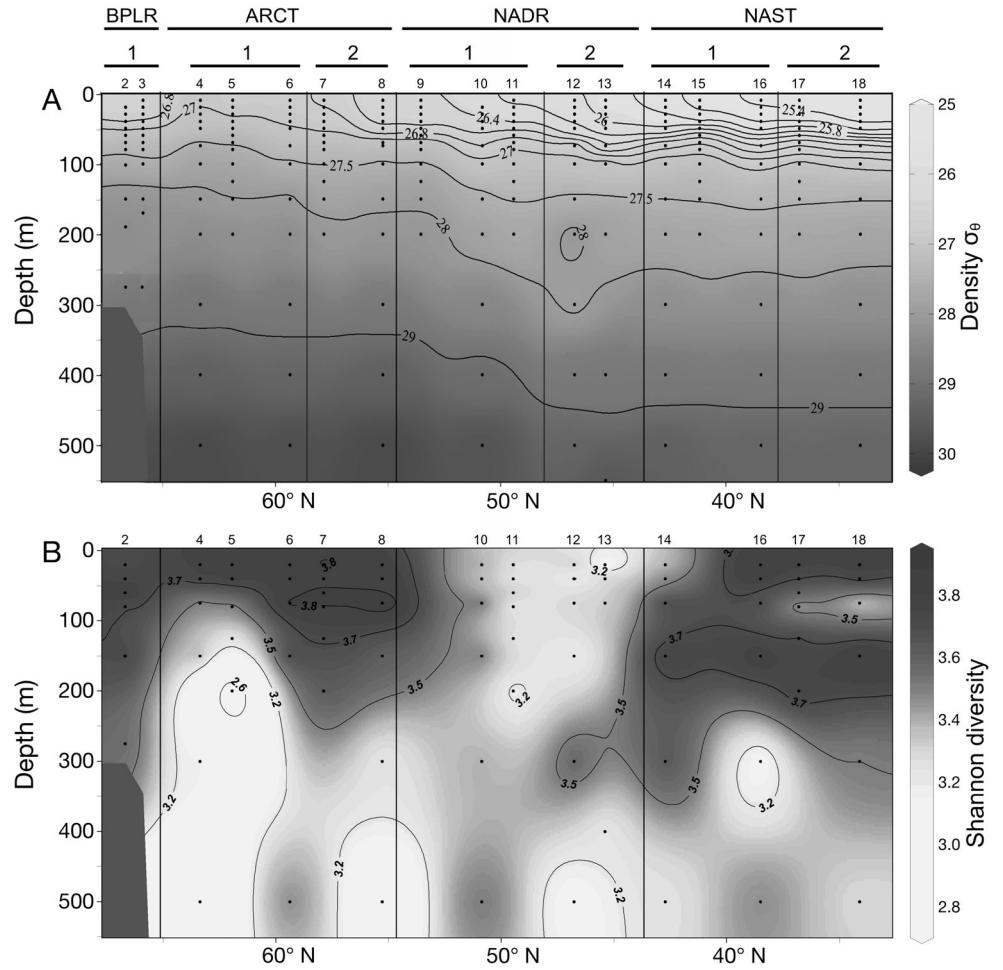
Fig. 1. (A) Density σ_θ and (B) Shannon diversity index of bacterial 16S rRNA T-RFLP profiles in the North Atlantic Ocean. Water was sampled from the East Greenland Current (BPLR, between Greenland and Iceland) through the areas north (ARCT) and south (NADR) of the Gulf Stream to the North Atlantic Subtropical Gyre (NAST, south of the Azores). The Longhurst provinces and water bodies (see Table 1 for abbreviations) are shown for orientation
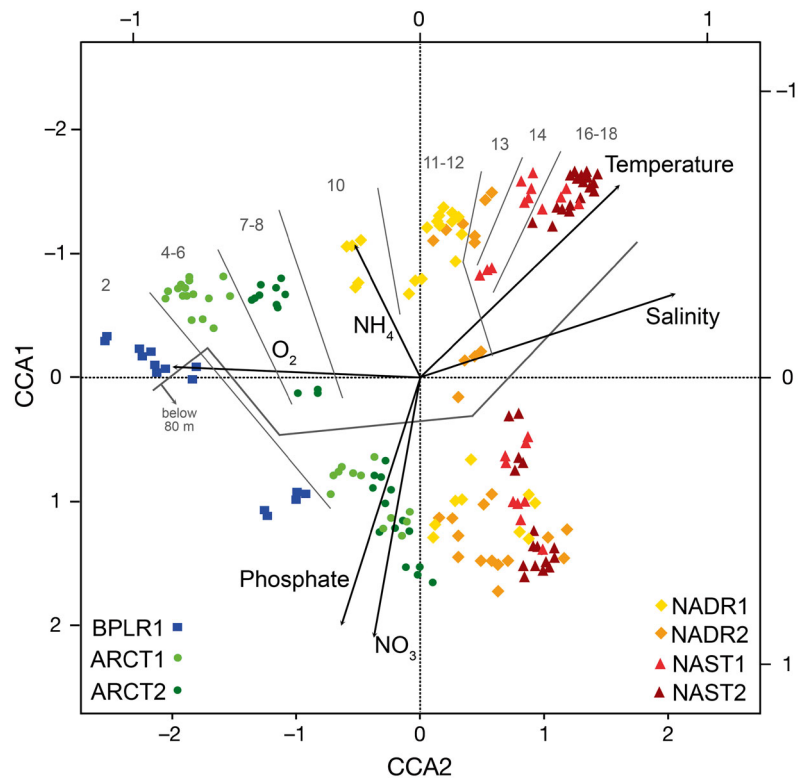

Fig. 2. Canonical correspondence analysis (CCA) of T-RFLP profiles and constrained environmental parameters. The CCA presentation of the differences in the TRF pattern revealed a clustering of bacterial communities along a latitudinal gradient (BPLR, square; ARCT, dot; NADR, diamond; NAST, triangle; see Table 1 for abbreviations of water bodies) and a separation in either epipelagic or upper mesopelagic origin (pycnocline at 50 to 100 m depth). The depth gradient is represented by the amount of ammonium, nitrate and phosphate, whereas the latitude gradient is represented by the amount of oxygen, temperature and salinity, because including the depth and latitude into the calculation would ultimately change the 2-dimensional visualization (abundances were constrained to the environmental data). Numbers represent sampling stations. Secondary x- and y-axes indicate eigenvalues of the environmental parameters

on relative abundances) and the Sørensen index (based on presence/absence). The nMDS of both indices revealed a distribution of bacterial populations along the 4 provinces and with water depth (Fig. S4), comparable to the results of the CCA. This supports our hypothesis of a change in bacterial communities with water bodies (β diversity). However, the discrimination of bacterial communities in different water bodies was more pronounced when relative abundances were considered.

Altogether, the applied nonparametric statistical analyses demonstrated the presence of individual bacterial communities in the different water bodies.

Characteristic TRFs for individual oceanic provinces

Differences between bacterial communities present in water bodies were traced to individual TRFs with SIMPER (Table 1). Among the abundant TRFs, only TRF_58nt and TRF_152nt were detected in all stations below 80 m water depth, whereas 40 TRFs varied in their presence. In the north, TRF_203nt and TRF_259nt were characteristic of the BPLR. ARCT1 contained a significant population of the TRFs 125nt and 605nt, ARCT2 contained a significant population of TRF_158nt, TRF_193nt and TRF_201nt, and NADR2

Table 1. Abundance and significance of terminal restriction fragments (TRFs) that were representative of water bodies (BPLR: Boreal Polar; ARCT: Atlantic Arctic; NADR: North Atlantic Drift; NAST: North Atlantic Suptropical Gyre). Depicted are the relative fluorescence intensity (RFI) maximum and the associated sampling site (station, depth), the average RFI in the water bodies (BPLR, ARCT1/2, NADR1/2 and NAST1/2), the average RFI of TRFs in the represented oceanic province (In, shown in **bold**) and in all other provinces (Out) and the mean dissimilarity (Diss/SD) from SIMPER above the pycnoline. TRF_152nt and TRF_241nt are reported for the mixed layer (0–80 m) and the deeper water (80–500 m). Additionally, the same information is given for the sum of TRF (sum) that are representative of water bodies

| Water body | TRF (nt) | RFI maximum | | | Average abundance in water body (%) | | | | | | | | | Diss/SD |
		RFI (%)	Stn	Depth (m)	BPLR	ARCT 1	2	NADR 1	2	NAST 1	2	Out	In	
BPLR	203	4.3	2	20	**3.4**	0.9	1.3	0.6	0.5	1.5	1.7	1.1	**3.4**	2.2
	259	6.4	2	20	**2.6**	0.1	0.9	0.1	0.1	0.3	0.2	0.2	**2.6**	1.0
	Σ				**6.0**	1.0	2.2	0.7	0.6	1.8	1.9	1.3	**5.9**	1.4
BPLR_ARCT1	202	5.8	5	20	**4.4**	**2.6**	0.7	0.1	0.1	0.0	0.0	0.2	**3.2**	1.7
BPLR_ARCT1_ARCT2	249	8.6	6	75	**6.9**	**5.2**	**5.1**	1.3	0.7	1.2	1.3	1.2	**5.5**	2.4
	461/2	4.7	2	20	**2.5**	**1.8**	**1.5**	0.0	0.0	0.0	0.0	0.1	**1.7**	1.7
	Σ				**9.4**	**7.0**	**6.6**	1.3	0.7	1.2	1.3	1.2	**7.2**	2.5
ARCT1	125	12.3	5	20	1.4	**5.5**	2.5	3.3	0.6	0.4	0.6	1.5	**5.5**	1.5
	605	4.7	6	20	0.9	**2.6**	0.2	0.3	0.3	0.3	0.4	0.3	**2.6**	1.5
	Σ				2.3	**8.1**	2.7	3.5	0.8	0.7	1.0	1.8	**8.1**	1.6
ARCT2	201	4.8	8	20	0.0	0.0	**1.2**	0.5	0.0	0.0	0.0	0.1	**1.2**	0.8
	193	17.1	7	40	0.3	3.1	**6.3**	1.8	2.1	1.3	1.2	2.3	**4.8**	1.4
	158	5.3	7	20	0.0	0.7	**1.8**	1.3	2.3	0.6	0.0	0.8	**1.8**	1.2
	Σ				0.3	3.9	**9.3**	3.6	4.4	1.9	1.2	2.7	**9.9**	1.2
ARCT2_NADR	194	11.9	12	40	0.2	1.1	**4.8**	**4.7**	**6.6**	1.9	0.3	0.9	**5.3**	2.3
NADR	195	8.0	12	40	2.7	3.0	0.7	**4.1**	**5.3**	1.8	1.3	1.9	**4.6**	1.7
NADR2	158	3.5	12	20	0.0	0.7	1.8	1.3	**2.3**	0.6	0.0	0.8	**2.3**	1.9
NAST	183	5.1	16	40	0.1	0.0	0.1	0.1	0.1	**1.8**	**2.0**	0.1	**1.9**	1.3
	207	5.1	17	20	0.0	0.0	0.1	0.1	0.1	**2.6**	**2.8**	0.2	**2.7**	1.6
	242	4.0	16	75	0.2	0.0	0.1	0.3	0.6	**1.8**	**1.8**	0.3	**1.8**	2.4
	Σ				0.3	0.0	0.3	0.5	0.8	**6.2**	**6.5**	0.5	**6.4**	1.9
NADR_NAST	227	5.8	14	75	0.1	0.6	1.1	**2.6**	**2.7**	**3.2**	**2.1**	0.2	**2.6**	1.6
	246	4.8	13	20	0.0	0.3	0.5	**3.0**	**2.7**	**3.5**	**3.9**	0.3	**3.3**	3.1
	189	44.1	18	75	1.0	4.1	2.1	**19.8**	**14.7**	**14.9**	**22.2**	2.8	**18.3**	1.8
	Σ				1.2	4.9	3.7	**25.4**	**20.1**	**21.7**	**28.2**	3.3	**24.3**	2.2
BPLR_ACRT_NADR	217	5.3	10	75	**2.0**	**2.0**	**2.9**	**2.9**	**1.3**	1.2	0.9	0.9	**2.2**	1.4
ARCT_NADR_NAST	204	21.2	12	75	2.1	**4.1**	**4.3**	**8.5**	**11.1**	**7.7**	**6.0**	2.1	**6.6**	1.1
Mixed layer	152	30.1	13	400	8.0	4.0	5.5	3.8	5.1	1.1	0.8	3.8		
Deeper water					**11.2**	**15.6**	**21.8**	**17.9**	**22.2**	**22.2**	**18.6**		**19.3**	2.4
Mixed layer	241	12.9	2	275	2.7	1.5	2.3	0.2	0.2	0.0	0.0	0.9		
Deeper water					**10.8**	**5.3**	**6.5**	**4.2**	**3.5**	**3.3**	**1.7**		**4.5**	1.5

contained a significant population of the TRF_158nt. The sampling sites in the NADR province had a high abundance of TRF_195nt, and in the NAST province, TRF_183nt, TRF_207nt and TRF_242nt were abundant. The analysis revealed a number of TRFs significant for 2 adjacent provinces. TRFs 204nt and 217nt were less abundant in the north and in the south, respectively. Water samples below 80 m were characterized by TRF_152nt and TRF_241nt (Table 1). The PCA was used to identify TRFs forming a strong increasing RFI along the latitude or depth gradients (Fig. S5 in Supplement 1). The first principal component distinguished between the northern (BPLR and ARCT) and the southern provinces (NADR and NAST) and revealed a strong influence of latitude (55.8% of the total variation). The second principal component covered 13.6% of the total variance and
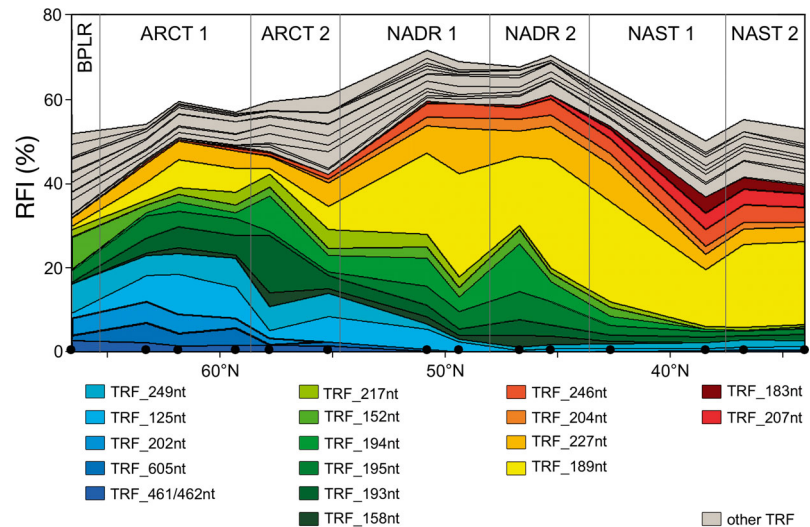

Fig. 3. Regionality of terminal restriction fragments (TRFs) between sea surface and 80 m depth, represented by their mean (of triplicates) relative fluorescence intensity (RFI). Shown are TRFs with a RFI of ≥4% in at least 1 water sample. Sampling sites along the transect through the northern (BPLR, ARCT) and southern (NADR, NAST) North Atlantic Ocean provinces (see Table 1 for abbreviations) are represented by black dots

distinguished communities above and below 80 m. The TRF_189nt had the largest eigenvector parallel to the first principal component. The second principal component had major contributions from TRF_125nt, TRF_152nt, TRF_193nt and TRF_204nt.

Thus, different water bodies along the North Atlantic Ocean transect were well characterized by individual TRFs, a proxy for bacterial taxa (Fig. 3).

Assignment of terminal restriction fragments to bacterial taxa

With TRFragCalc, written in MATLAB for the present study, we assigned *in silico* terminal restriction fragments to cyanobacteria and compared the result with fluorescence detection of cyanobacteria by flow cytometry (Fig. 4a,c,e) and with measured TRFs (Fig. 4b,d,f). *In silico* terminal restriction fragments (iTRFs) of *Synechococcus* 16S rRNA genes had 73, 128, 190 and 205 nucleotides. The iTRF_128nt originated from 246 sequences of *Synechococcus* Clade I and 3 sequences of *Synechococcus* Clade III (overall, 286 cyanobacteria in 327 sequences). TRF abundance pattern and the distribution determined by flow cytometry of *Synechococcus* cell counts indicated concurrence with the abundance of TRF_125nt, with a Pearson correlation coefficient ρ_P of 0.89. Within the diversity of 135 761 bacterial 16S rRNA sequences present in the dataset, the absence of other iTRF in

the range 123 to 133 nt originating from marine bacteria also supported an assignment of *Synechococcus* iTRF_128nt to the observed TRF_125nt.

Ecotypes of *Prochlorococcus* differ in their 16S rRNA genes (Rocap et al. 2002). We found 2 iTRFs: the iTRF_190nt originated from 16S rRNA gene sequences of *Prochlorococcus* strains that were adapted to high light; low light adapted *Prochlorococcus* were represented by iTRF_205nt. In the T-RFLP profiles, TRF_189nt and TRF_204nt showed in the subtropical province a distribution which matched the distribution of high light (ρ_P = 0.91) and low light (ρ_P = 0.46) adapted *Prochlorococcus* (Fig. 4c–f).

In a similar coincidence study, we used population quantification by *in situ* hybridization (Gómez-Pereira et al. 2010) for *Flavobacteria*. The iTRF_461nt and iTRF_462nt originated from *Polaribacter* only, and iTRF_464nt originated from *Polaribacter* and the groups NS4 and NS2b. The T-RFLP pattern of TRF_461nt and TRF_462nt concurred in the northern provinces (Fig. S6 in Supplement 1). The *Flavobacteriaceae* VIS4 group was solely responsible for iTRF_604nt. The T-RFLP pattern of TRF_605nt coincided well with the VIS4 population in *in situ* hybridization cell counts of 2% of all DAPI stained cells at Stns 4 and 6 (Fig. S6). The *Flavobacteriaceae* group DE2 contributed exclusively to iTRF_606nt. DE2 was abundant in ARCT and NAST according to the T-RFLP pattern of TRF_607nt and to *in situ* hybridization cell counts. The *Flavobacteriaceae* group
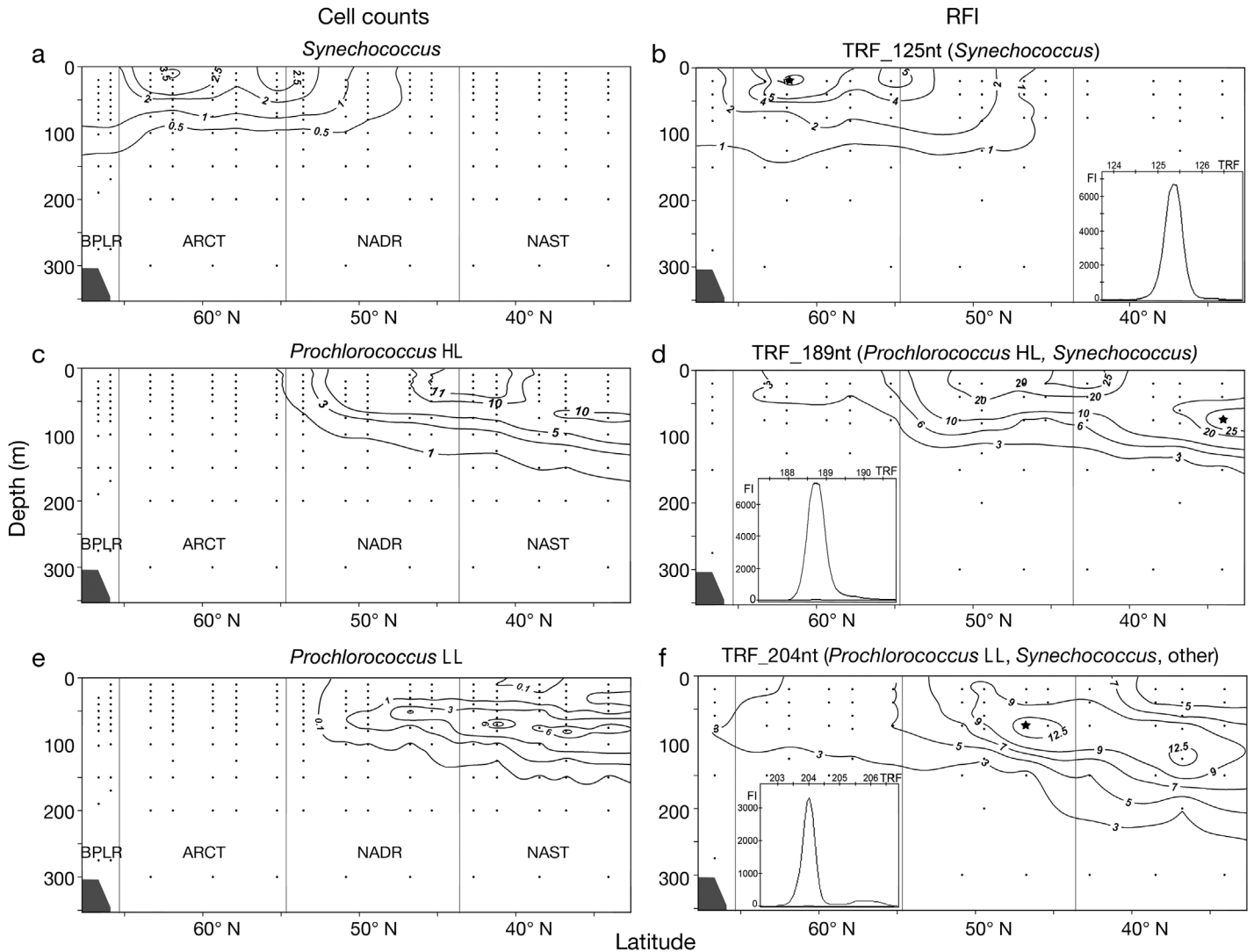
Fig. 4. Cyanobacterial populations analyzed (a,c,e) by flow cytometry and (b,d,f) by their mean (of triplicates) relative fluorescence intensity (RFI) of affiliating terminal restriction fragments (TRFs). Flow cytometry detected (a) *Synechococcus* more abundant in the mesotrophic region of the northern North Atlantic Drift (NADR) and (c,e) *Prochlorococcus* in the oligotrophic central Atlantic gyre. The peak abundance of (c) the high light (HL) adapted *Prochlorococcus* ecotype was at 40 m, whereas (e) the low light (LL) adapted *Prochlorococcus* ecotype was observed at deeper water layers. The pattern of TRF corresponded to the related cyanobacterial populations: (b) TRF_125nt affiliated to *Synechococcus*, (d) TRF_189nt affiliated to the HL-adapted *Prochlorococcus* ecotype and *Synechococcus*, and (f) TRF_204nt affiliated to the LL-adapted *Prochlorococcus* ecotype, *Synechococcus* and other taxa. A star indicates the sampling site with the largest RFI, and the inset shows the corresponding fluorescence intensity (FI) in the T-RFLP profile. Contour lines indicate the relative abundance of cell counts (% total cell counts) or TRFs (% RFI). Black dots indicate sampling sites

DE2 gave also iTRFs in the range of 817 to 825 nucleotides. The TRF_820nt was found only in surface waters in the NAST province, but its abundance was below 1% relative fluorescence intensity.

The *in silico* analyses did not allow a clear assignment of other TRFs that were representative of oceanic provinces (Table 1), indicating the low taxonomic resolution of the T-RFLP method. In several cases, however, assignments to *Alphaproteobacteria* as well as *Gammaproteobacteria* were possible (Fig. S7 in Supplement 1).

DISCUSSION

The biogeography of microorganisms has been a topic in microbiology for a long time, but the interest was renewed with the characterization of microbes from the Arctic and Antarctic (Staley & Gosink 1999), and recently, large sequence datasets were obtained from samples from many regions on earth (Zinger et al. 2011). Diversity studies with reference to the studied latitude have shown a pole-to-pole biogeography (Ghiglione et al. 2012), a bipolar distribution (Sul et al.

2013) and a latitudinal diversity gradient (Fuhrman et al. 2008). Baldwin et al. (2005) reported the microbial diversity in a Pacific Ocean pole-to-pole transect between 154° W and 172° E and detected 4 biological provinces: sub-Arctic/Arctic, temperate, tropical and sub-Antarctic/Antarctic. Each province covered a large range of latitudes. The situation is different in the North Atlantic Ocean, which has 4 oceanic provinces present on a relatively small range of latitudes. This pattern results from the North Atlantic Current with a profound influence on the history of water bodies. The sampling line along the 30° W meridian is nearly perpendicular to the North Atlantic Current which transports water with a velocity of 1 m s^{-1} (Clarke et al. 1980). Fluorescence *in situ* hybridizations (Gómez-Pereira et al. 2010) revealed a biogeography of *Polaribacter* with a higher abundance in the northern provinces. Now, our T-RFLP analyses extended this observation to whole bacterial communities in the North Atlantic Ocean consistent with water bodies. The photic pelagial showed large differences between bacterial communities, whereas the mesopelagial had less diverse bacterial communities. Larger changes in environmental parameters above the pycnocline as well as the presence of phototrophic microorganisms might contribute to the greater diversity. Unicellular cyanobacteria affiliated to *Synechococcu*s and *Prochlorococcus* are among the major bacterial populations with a biogeography in the oceans (Li 1994, Liu et al. 1997, Veldhuis et al. 1997). The large genetic diversity within *Synechococcus* and *Prochlorococcus* has lead to the definition of ecotypes for genetically well-defined subgroups (Zwirglmaier et al. 2008). *Prochlorococcus* has HL-adapted and LL-adapted ecotypes (Zubkov et al. 2007, Zwirglmaier et al. 2008, Huang et al. 2012). *In silico* fragment length calculations and the coincidence of the biogeography of the TRF with fluorometric measurements of pigments enabled the assignment of TRFs to *Synechococcus* and to low and high light adapted *Prochlorococcus* ecotypes. This technique showed the distribution of *Prochlorococcus* ecotypes in different water depths.

In summary, the water bodies of oceanic provinces in the North Atlantic Ocean hosted different bacterial communities. Bacterial populations varied along the latitudinal transect, such that individual TRFs could serve as representative proxies of individual oceanic provinces.

LITERATURE CITED

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

Aiken J, Rees N, Hooker S, Holligan P and others (2000) The Atlantic Meridional Transect: overview and synthesis of data. Prog Oceanogr 45:257−312

Arnosti C, Fuchs BM, Amann R, Passow U (2012) Contrasting extracellular enzyme activities of particle-associated bacteria from distinct provinces of the North Atlantic Ocean. Front Microbiol 3:425

Baldwin AJ, Moss JA, Pakulski JD, Catala P, Joux F, Jeffrey WH (2005) Microbial diversity in a Pacific Ocean transect from the Arctic to Antarctic circles. Aquat Microb Ecol 41:91−102

Boström KH, Simu K, Hagström A, Riemann L (2004) Optimization of DNA extraction for quantitative marine bacterioplankton community analysis. Limnol Oceanogr Methods 2:365−373

Bruland O, Almqvist EW, Goldberg YP, Boman H, Hayden MR, Knappskog PM (1999) Accurate determination of the number of CAG repeats in the Huntington disease gene using a sequence-specific internal DNA standard. Clin Genet 55:198−202

Clarke KR (1993) Non-parametric multivariate analysis of changes in community structure. Aust J Ecol 18:117−143

Clarke RA, Hill HW, Reiniger RF, Warren BA (1980) Current system south and east of the Grand Banks of Newfoundland. J Phys Oceanogr 10:25−65

Devred E, Sathyendranath S, Platt T (2007) Delineation of ecological provinces using ocean colour radiometry. Mar Ecol Prog Ser 346:1−13

Dixon P (2003) VEGAN, a package of R functions for community ecology. J Veg Sci 14:927−930

Emery WJ, Meincke J (1986) Global water masses: summary and review. Oceanol Acta 9:383−391

Esaias WE, Iverson RL, Turpie K (2000) Ocean province classification using ocean colour data: observing biological signatures of variations in physical dynamics. Glob Change Biol 6:39−55

Forney LJ, Zhou X, Brown CJ (2004) Molecular microbial ecology: land of the one-eyed king. Curr Opin Microbiol 7:210−220

Fuhrman JA, Steele JA, Hewson I, Schwalbach MS, Brown MV, Green JL, Brown JH (2008) A latitudinal diversity gradient in planktonic marine bacteria. Proc Natl Acad Sci USA 105:7774−7778

Ghiglione JF, Galand PE, Pommier T, Pedrós-Alió C and others (2012) Pole-to-pole biogeography of surface and deep marine bacterial communities. Proc Natl Acad Sci USA 109:17633−17638

Gómez-Pereira PR, Fuchs BM, Alonso C, Oliver MJ, van Beusekom JEE, Amann R (2010) Distinct flavobacterial communities in contrasting water masses of the North Atlantic Ocean. ISME J 4:472−487

Hahn M, Wilhelm J, Pingoud A (2001) Influence of fluorophor dye labels on the migration behavior of polymerase chain reaction-amplified short tandem repeats during denaturing capillary electrophoresis. Electrophoresis 22:2691−2700

Huang S, Wilhelm SW, Harvey HR, Taylor K, Jiao NZ, Chen F (2012) Novel lineages of *Prochlorococcus* and *Synechococcus* in the global oceans. ISME J 6:285−297

Li WKW (1994) Primary production of prochlorophytes, cyanobacteria, and eukaryotic ultraphytoplankton: measurements from flow cytometric sorting. Limnol Oceanogr 39:169–175

Liu HB, Nolla HA, Campbell L (1997) *Prochlorococcus* growth rate and contribution to primary production in the equatorial and subtropical North Pacific Ocean. Aquat Microb Ecol 12:39–47

Longhurst A, Sathyendranath S, Platt T, Caverhill C (1995) An estimate of global primary production in the ocean from satellite radiometer data. J Plankton Res 17: 1245–1271

Ludwig W, Strunk O, Westram R, Richter L and others (2004) ARB: a software environment for sequence data. Nucleic Acids Res 32:1363–1371

Mann CR (1967) Termination of the Gulf Stream and the beginning of the North Atlantic Current. Deep-Sea Res 14:337–359

Muyzer G, Teske A, Wirsen C, Jannasch H (1995) Phylogenetic relationships of *Thiomicrospira* species and their identification in deep-sea hydrothermal vent samples by denaturing gradient gel electrophoresis of 16S rDNA fragments. Arch Microbiol 164:165–172

Olejniczak M, Kozlowski P, Sobczak K, Krzyzosiak WJ (2005) Accurate and sensitive analysis of triplet repeat expansions by capillary electrophoresis. Electrophoresis 26:2198–2207

Oliver MJ, Irwin AJ (2008) Objective global ocean biogeographic provinces. Geophys Res Lett 35:1–6

Platt T, Sathyendranath S (1999) Spatial structure of pelagic ecosystem processes in the global ocean. Ecosystems 2: 384–394

Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig WG, Peplies J, Glöckner FO (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res 35:7188–7196

Ramette A (2009) Quantitative community fingerprinting methods for estimating the abundance of operational taxonomic units in natural microbial communities. Appl Environ Microbiol 75:2495–2505

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

Schattenhofer M, Wulf J, Kostadinov I, Glöckner FO, Zubkov MV, Fuchs BM (2011) Phylogenetic characterization of picoplanktonic populations with high and low nucleic acid content in the North Atlantic Ocean. Syst Appl Microbiol 34:470–475

Schlitzer R (2002) Interactive analysis and visualization of geoscience data with Ocean Data View. Comput Geosci 28:1211–1218

Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech J 27:379–423 and 623–656 (addendum)

Staley JT, Gosink JJ (1999) Poles apart: biodiversity and biogeography of sea ice bacteria. Annu Rev Microbiol 53: 189–215

Stewart RH (2008) Chapter 6. Temperature, salinity, and density. In: Introduction to physical oceanography. Department of Oceanography, Texas A & M University, p 73–102. Available at http://oceanworld.tamu.edu/resources/ocng_textbook/PDF_files/book.pdf

Sul WJ, Oliver TA, Ducklow HW, Amaral-Zettler LA, Sogin ML (2013) Marine bacteria exhibit a bipolar distribution. Proc Natl Acad Sci USA 110:2342–2347

Tarran GA, Heywood JL, Zubkov MV (2006) Latitudinal changes in the standing stocks of nano- and picoeukaryotic phytoplankton in the Atlantic Ocean. Deep-Sea Res II 53:1516–1529

Teeling H, Fuchs BM, Becher D, Klockow C and others (2012) Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. Science 336:608–611

Ter Braak CJF (1986) Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology 67:1167–1179

Veldhuis MJW, Kraay GW, van Bleijswijk JDL, Baars MA (1997) Seasonal and spatial variability in phytoplankton biomass, productivity and growth in the northwestern Indian Ocean: the southwest and northeast monsoon, 1992-1993. Deep-Sea Res I 44:425–449

Zinger L, Amaral-Zettler LA, Fuhrman JA, Horner-Devine MC and others (2011) Global patterns of bacterial beta-diversity in seafloor and seawater ecosystems. PLoS ONE 6:e24570

Zubkov MV, Mary I, Woodward EMS, Warwick PE, Fuchs BM, Scanlan DJ, Burkill PH (2007) Microbial control of phosphate in the nutrient-depleted North Atlantic subtropical gyre. Environ Microbiol 9:2079–2089

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