Systematic and Applied Microbiology 42 (2019) 126000

Contents lists available at ScienceDirect

Systematic and Applied Microbiology

journal homepage: http://www.elsevier.com/locate/syapm

Distinct relationships between fluorescence *in situ* hybridization and 16S rRNA gene- and amplicon-based sequencing data of bacterioplankton lineages

Insa Bakenhus[a], Bernd Wemheuer[b], Pinar Akyol[b], Helge-Ansgar Giebel[a], Leon Dlugosch[a], Rolf Daniel[b], Meinhard Simon[a,*]

[a] *Institute for Chemistry and Biology of the Marine Environment, University of Oldenburg, Carl von Ossietzky Str. 9-11, D-26129 Oldenburg, Germany*
[b] *Department of Genomic and Applied Microbiology and Göttingen Genomics Laboratory, Institute of Microbiology and Genetics, Georg-August University of Göttingen, Grisebachstr. 8, D-37077 Göttingen, Germany*

ARTICLE INFO

Article history:
Received 6 April 2019
Received in revised form 21 June 2019
Accepted 28 June 2019

Keywords:
Bacteria
CARD-FISH
16S rRNA amplicon sequencing
North Sea
Mesocosm

ABSTRACT

Catalyzed reporter deposition fluorescence *in situ* hybridization (CARD-FISH) and amplicon sequencing of the total (16S rRNA gene) and potentially active (16S rRNA transcripts), community are the major state of the art approaches for assessing the composition of bacterial communities in marine pelagic and other ecosystems. However, CARD-FISH and amplicon sequencing methods have not yet been directly compared to assess the composition of bacterioplankton communities. Therefore, these approaches were used to study the composition of bacterial communities in two North Sea seawater mesocosm experiments supplemented with diatom-derived organic matter (OM). All approaches revealed *Proteobacteria* and *Bacteroidetes* as major components of the bacterioplankton communities. The *Roseobacter* group and its RCA cluster, as well as *Bacteroidetes* and *Gammaproteobacteria*, responded most strongly to OM addition, whereas the SAR11 clade responded in only one of the two mesocosms. A correlation analysis showed that CARD-FISH and amplicon sequencing data of the SAR11 clade and the *Roseobacter* group, together with its RCA cluster, were highly significantly correlated, whereas *Bacteroidetes* did not yield any significant correlation and *Gammaproteobacteria* was only correlated with the potentially active fraction. However, subgroups of these phylogenetic groups, the SAR92 clade, the genera *Pseudoalteromonas* and *Polaribacter*, exhibited significant correlations in one of the two mesocosms. Correlations of CARD-FISH with amplicon sequencing data from the total and potentially active fractions of these lineages exhibited distinct differences. The study showed that CARD-FISH and amplicon sequencing data of distinct bacterioplankton groups and especially the phylogenetic lineages at a higher taxonomic level were correlated but reflected different aspects of their growth dynamics.

Introduction

Heterotrophic bacterioplankton communities are essential structural elements in aquatic systems and play critical roles in biogeochemical cycling of matter and elements such as carbon, nitrogen and sulfur [5,13,43]. These communities are highly diverse due to their various functional roles in organic matter processing and biogeochemical cycles, and the distinct responses of functional groups to varying substrate supply or different environmental factors [40,43,66]. Therefore, the identification of relevant bacterial phylogenetic lineages and their metabolic states is essential for understanding the relationships between bacterioplankton community composition and ecosystem functions. Fluorescence *in situ* hybridization (FISH) and high-throughput sequencing technologies of specific target regions of the 16S rRNA gene are the two major approaches currently used to assess the composition of bacterioplankton communities [2,17,23,49,61,66]. Both approaches have provided detailed insights into the composition and dynamics of bacterioplankton communities in various marine systems [6,24,41,58,72].

FISH and its state of the art improvement with a higher sensitivity, catalyzed reporter deposition (CARD)-FISH [50], allow the relative quantification of specific bacterial phylogenetic lineages on a single cell basis but also the absolute quantification when combined with cell number enumeration, for example, by flow cytometry or epifluorescence microscopy. CARD-FISH probes can target larger or narrower phylogenetic groups depending on the specificities of the probes applied. Limitations are 16S rRNA

* Corresponding author.
E-mail address: m.simon@icbm.de (M. Simon).

https://doi.org/10.1016/j.syapm.2019.06.005
0723-2020/

sequence variabilities that cause mismatches in phylogenetically distant bacterial groups (false-positives) or prevent probe binding in subgroups (false negatives) [2]. Other limitations are that only rather abundant (*i.e.* >0.5% of total cells) and known lineages can be targeted. In addition, the number of probes applied is limited due to time-consuming sample processing. Even though other FISH-techniques are also applied, such as for NanoSIMS [FISHSIMS; 46] and gene or phage detection [geneFISH; 1,44], CARD-FISH is still the most widely applied method to target and quantify single prokaryotic cells in aquatic microbial ecology.

A higher phylogenetic resolution of bacterioplankton community composition than by CARD-FISH is achieved by amplicon sequencing of variable regions of the 16S rRNA gene [17,61]. This approach provides detailed insights into the community composition and can identify simultaneously abundant and, depending on the sequencing depth, rare taxa of the community. It includes a PCR amplification step of the 16S rRNA gene fragment, which can lead to omission, distortion and/or misrepresentation of members of a bacterial community [29,36,52]. Furthermore, primer bias may lead to overrepresentation of distinct phylogenetic lineages, whereas others may be underrepresented [4,33,49,69]. However, this approach allows only a relative quantification of the different taxa and cannot provide absolute numbers due to variations of the copy number of rRNA operons in genomes of different bacterial phylogenetic groups [14]. Nevertheless, efforts to correct for this bias have been made [3,31].

A variation of this approach allows the composition of the active fraction of the bacterial community to be assessed (*i.e.* sequencing of 16S rRNA transcripts). The abundance of 16S rRNA transcripts is a general indicator of protein synthesis and thus of metabolically active bacteria. There is evidence that even in pure bacterial cultures cellular RNA concentrations are not directly proportional to protein synthesis and the growth rate but represent a general indicator of the physiological state, unless the amount of DNA per cell is assessed as well [10]. However, this approach provides additional insights into the dynamics of growth responses of the active bacteria as a fraction of the total bacterial community to biogeochemical and environmental factors over space and time [14,15,30,51,71,73]. It has been argued that this approach is biased because there is no clear-cut relationship of the cellular rRNA content to growth in several bacterial taxa, and degradation of rRNA in bacteria varies considerably when they exit metabolically active states [10]. These authors proposed to use the term potentially active bacteria when the amplicon sequencing approach targets the cDNA. Both approaches have been compared in several field studies [14,15,30,51,71,73] and with *Candidatus* Pelagibacter ubique of the SAR11 clade [35]. These results showed distinct relationships of the rRNA/rDNA ratio for different bacterial taxa.

Despite the wide use of CARD-FISH and 16S rRNA amplicon sequencing in microbial ecology, to the best of our knowledge, neither approach has yet been directly compared to assess the composition of bacterioplankton communities, even when they were applied in the same study [9,34,53,66]. The only direct comparison has been performed with pitcher plants [65]. To interpret the data of either approach better, it is important to know how the relative abundances of different bacterial taxa compare when they are assessed by CARD-FISH and amplicon sequencing, targeting either the 16S rRNA gene or transcript. Are the results of both approaches correlated and are the correlations different for individual bacterial taxa and larger phylogenetic groups?

To address this question, two mesocosm experiments were set up during a cruise in the North Sea in early summer and growth of the bacterioplankton communities was stimulated by adding diatom-derived organic matter (OM). To examine how the approaches mentioned above detected the response of the community composition to the added diatom-derived OM, the study used CARD-FISH probes of different specificity to target distinct taxa, as well as larger phylogenetic groups and amplicon sequencing of the V3–V4 region of the 16S rRNA gene and transcripts.

Table 1
Position, hydrographic conditions, chlorophyll *a*, bacterial abundance and biomass production at a 3 m depth at the stations where the mesocosm experiments were carried out.

	Mesocosm 1	Mesocosm 2
Longitude/latitude	54° 55.2′ N, 6° 49.9′ E	57° 00′ N, 4° 00′ E
Temperature (°C)	14.0	12.0
Salinity	34.085	35.071
Chlorophyll *a* ($\mu g\,L^{-1}$)	0.62	0.13
Bacterial abundance (10^6 cells mL^{-1})	0.93	1.31
Bacterial biomass production ($ng\,C\,L^{-1}\,h^{-1}$)	121.3	144.6

Materials and methods

Experimental design, sample collection and processing

The mesocosm experiments (M1 and M2) were carried out during a cruise with R/V Heincke in the North Sea between 24 May and 5 June 2014, M1 in the southern region (54° 55.2′ N, 6° 49.9′ E) and M2 in the northern region (57° 00′ N, 4° 00′ E; Table 1).

Seawater was collected at a depth of 3 m using a conductivity–temperature–depth (CTD) rosette profiler equipped with 5-L Niskin bottles. The unfiltered samples were filled into 20-L acid-rinsed Nalgene bottles. A total of 500 mL of a diatom-derived OM solution, obtained by autoclaving axenically grown cultures of the diatoms *Thalassiosira rotula* (CCMP1647) and *Leptocylindrus danicus* (CCMP470), was added without any further treatment at days 0 and 4 with three replicates. This volume was equivalent to a final concentration of approximately 200 μM dissolved organic carbon (DOC). Two bottles with no OM-addition served as a control. Bottles were incubated on board the ship for 10 days at *in situ* temperature in the dark and subsampled every other day for microbial cell numbers, CARD-FISH and amplicon sequencing of the potentially active bacteria (16S rRNA) and total bacterial community (16S rRNA gene). There was no pre-filtration of any sample subjected to subsequent analyses. Samples of 1000 mL for bacterioplankton community analyses were harvested on 0.2-μm polycarbonate filters (47 mm diameter, Whatman Nuclepore, GE Healthcare, Munich, Germany) and stored at −80 °C until further analysis. Chlorophyll *a* and bacterial biomass production by incorporation of ^{14}C-leucine were determined at the stations, as described previously [6].

Bacterial abundance

Bacterial abundance was enumerated on board by counting cells stained with SYBR Green I (Invitrogen, Unden, The Netherlands) using a BD Accuri C6 cytometer (BD Biosciences, Franklin Lakes, NJ, USA) and the internal fluidics calibration of the device, as described previously [47]. Cells were distinguished based on green and red fluorescence, as well as forward scatter populations of low (LNA) and high nucleic acid content (HNA) [6].

Catalyzed reporter deposition (CARD-FISH)

Ten mL seawater subsamples were fixed with formaldehyde (final concentration 1% v/v) for 1 h at room temperature and filtered onto 0.2 μm polycarbonate membrane filters (47 mm diameter, Whatman Nuclepore, GE Healthcare, Munich, Germany). Filters were stored at −20 °C until further analysis. The CARD-FISH pro-

cedure was performed according to Pernthaler et al. [50] with modifications, as described previously [6]. Hybridizations were carried out with horseradish peroxidase-labelled oligonucleotide probes (Biomers, Ulm, Germany) targeting *Gammaproteobacteria* (GAM42a) [39], the SAR92 cluster (SAR92-627) [62], the SAR86 cluster (SAR86-1245) [74], *Pseudoalteromonas* (PSA184) [22], *Bacteroidetes* (CF319a) [40], *Polaribacter* (POL740) [38], the *Roseobacter* group (ROS536) [12], the RCA cluster (RCA996) [6], and the SAR11 clade (SAR11-441R) [55]. Formamide concentrations of 35% were used for all probes, except for SAR11-441R (45% formamide). Hybridizations were carried out at 35 °C overnight (probe SAR11-441R) or at 46 °C for 2 h (all other probes). Filters were embedded in 4:1 Vectashield H-1000: Citifluor antifading reagent (Vector Laboratories, Burlingame, CA/Citifluor, UK) containing 1 μg mL^{-1} DAPI and arranged on microscope slides. Hybridized and DAPI-stained cells were visualized semi-automatically using an AxioImager.Z2m epifluorescence microscope including the software package AxioVision 4.8.2.0 (Carl Zeiss, Germany) and enumerated using ACMEtool3 (M. Zeder; www.technobiology.ch). Absolute cell numbers were calculated from relative CARD-FISH abundances and bacterial cell numbers were determined by flow cytometry.

Nucleic acid extraction and sequencing

Total DNA and RNA were extracted from the filters using acidic phenol and further purified as described previously [59]. Residual DNA was removed by DNAse treatment and the absence was confirmed by PCR according to Schneider et al. [59] using the primer S-D-Bact-0341-b-S-17 (5′–CCT ACG GGN GGC WGC AG-3) [33] and reverse primer S-D-Bact-0785-a-A-21 (5′-GAC TAC HVG GGT ATC TAA TCC-3′) [33]. DNA-free RNA was converted to cDNA by SuperScript III (Thermo Fisher Scientific, USA) reverse transcription, as described by Wemheuer et al. [70], using the reverse primer S-D-Bact-0785-a-A-21 (5′-GAC TAC HVG GGT ATC TAA TCC-3′) [33]. To assess community structures, 16S rRNA amplicon libraries were generated from DNA and cDNA by PCR using Phusion® Polymerase, as described by Wemheuer et al. [70], with the forward primer S-D-Bact-0341-b-S-17 (see above) and the reverse primer S-D-Bact-0785-a-A-21 (see above) using Illumina Nextera adapters for sequencing. Each sample was subjected to three independent amplifications and pooled in equal amounts. Amplification products were purified using NucleoMag NGS Clean-up and Size Select (Macherey-Nagel, Dueren, Germany), quantified with a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, USA) and barcoded using the Nextera XT-Index kit (Illumina, San Diego, USA) and the Kapa HiFi HotStart polymerase (Kapa Biosystems, USA). Sequencing was performed in the Göttingen Genomics Laboratory on an Illumina MiSeq System (paired end 2 × 300 bp) using the MiSeq Reagent Kit v3 (Illumina). Sequence data were deposited in the sequence read archive (SRA) of the National Center for Biotechnology Information (NCBI) under the accession number SRP151847 (BioProject SRP151847).

Processing and analysis of sequence data

Obtained 16S rRNA data sets were processed as described by Pohlner et al. [53]. In brief, the sequencing data were initially quality-filtered with Trimmomatic version 0.36 [11]. Low quality reads were truncated if the quality dropped below 15 in a sliding window of 4 bp. Subsequently, all reads shorter than 100 bp and orphan reads were removed. Remaining sequences were merged, quality-filtered and further processed with USEARCH version 10.0.240 [20]. Filtering included the removal of reads shorter than 400 bp or longer than 500 bp, as well as the removal of low-quality reads (expected error >1) and reads with more than one ambiguous base. Processed sequences of all samples were concatenated to one file and subsequently dereplicated into unique sequences. These sequences were denoised with the unoise3 algorithm implemented in USEARCH. The UCHIME *de novo* algorithm was used to remove chimeric sequences during denoising. Subsequently, remaining chimeric sequences were removed using the UCHIME reference algorithm [21] in high_confidence mode with the SILVA SSU Ref NR 99 132 database [54] as a reference data set. Filtered sequences were mapped on the final set of denoised, unique sequences (zOTUs) to determine the occurrence and abundance of each zOTU in every sample. To assign taxonomy, zOTU sequences were classified by BLASTn (version 2.6.0+) against the SILVA database [54] with an e-value threshold of 1e-20. All non-bacterial zOTUs were removed based on their taxonomic classification in the respective database.

Comparison of probe and primer match

In order to examine how probes and primers targeted the different bacterial phylogenetic groups, the coverage of the major target lineages was tested by the TestProbe (https://www.arb-silva.de/search/testprobe/) and TestPrime software tools (https://www.arb-silva.de/search/testprime/) using SILVA SSU RefNR 132. Group-specific probes were used for the coverage analysis *via* TestProbe, except for *Flavobacteria*, where the probe for *Bacteroidetes* (CF319a) was used. Both the applied primers [33] and more recently introduced ones [4,48] were included.

Statistical evaluation

Statistical analyses were performed using R software version 3.4.3 [56]. Log-response ratios were calculated to assess treatment effects on bacterial lineages. They represented the log-transformed ratio of the absolute abundance of CARD-FISH positive cells to diatom-derived OM additions relative to those of the untreated controls. Slopes of linear regression models were calculated between CARD-FISH and amplicon sequence abundance data. Analysis of variance (ANOVA) was used to identify treatment effects on bacterial HNA content. Homogeneity of variances was evaluated by a Fligner-Killeen test [19] prior to the analysis.

Results

Initial hydrographic conditions and microbial features at the stations of mesocosm experiments M1 and M2 were rather similar (Table 1).

Bacterial abundance

In both experiments, the added diatom-derived OM resulted in an increase in bacterial cell numbers, whereas they remained almost constant in the controls (Fig. 1). The triplicates showed very similar temporal patterns. In M1, initial cell numbers of 1×10^6 mL^{-1} increased by a factor of 3 and peaked at days 2 and 6, two days after the addition of diatom-derived OM, with a continuous decline thereafter to almost identical numbers compared to the start. The decline in cell numbers at day 4 and after day 6 was presumably due to intense protozoan grazing as many flagellates with ingested bacteria were observed microscopically in the samples at these time points. In M2, initial cell numbers of 2×10^6 mL^{-1} doubled on day 2 and declined until day 8 to the initial numbers but this was followed by another increase at day 10 (Fig. 1).

In experiment M1, HNA cells greatly dominated with an overall mean of 80.9 ± 8.0% and the highest proportion of 91 ± 1.8% at day 2 (Supplementary Table S1). There were no significant differences in the proportions of HNA and LNA cells in the treatment and control

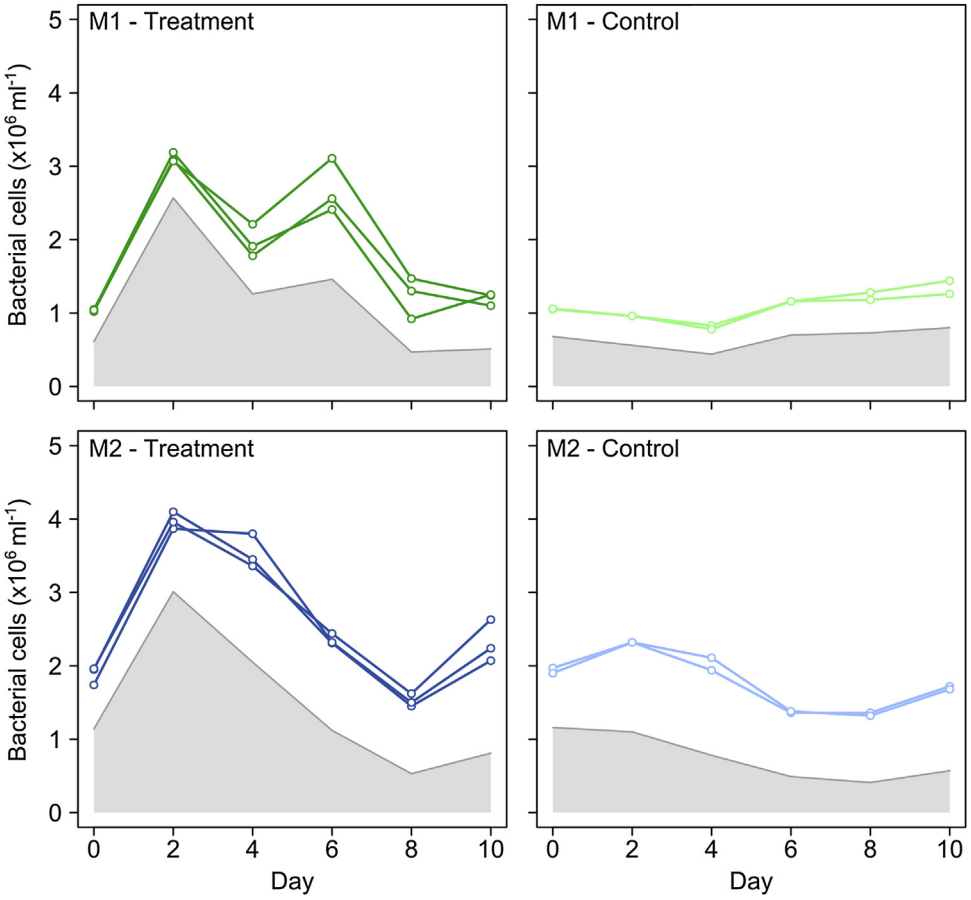

Fig. 1. Bacterial cell numbers and total numbers of CARD-FISH positive cells (gray areas) in the treatments and controls of mesocosm experiments 1 and 2.

groups. In experiment M2, HNA cells constituted on average only 53 ± 9.6% of total cells (Supplementary Table S1). However, in this experiment, the proportion of HNA was significantly higher in the treatment group than in the control group (ANOVA; $p < 0.001$).

Bacterial community composition (CARD-FISH)

Bacteroidetes, *Gammaproteobacteria* and the *Roseobacter* group dominated the initial community in M1, whereas SAR11 greatly dominated in the M2 equivalent (Fig. 2). The bacterial community composition changed over time in both mesocosm experiments and specific target groups responded differently to diatom-derived OM addition. The four target groups together accounted for 35–83% of all DAPI-stained cells (Fig. 1). The proportion of CARD-FISH identifiable cells covaried with total bacterial cell numbers in both experiments (Fig. 2A, B). *Bacteroidetes* and *Gammaproteobacteria* increased very similarly in both experiments during the first two days and decreased thereafter (Fig. 2). In M1, *Bacteroidetes* and *Gammaproteobacteria* constituted 7–27% and 13–26%, respectively, and in M2 the respective proportions of these groups were 8–14% and 4–11%. The *Roseobacter* group was most responsive of all target groups to diatom-derived OM addition in both experiments and accounted for 2–23% and 2–19% in M1 and M2, respectively. In M2, the SAR11 clade dominated throughout the experiment with proportions of 43% at day 0 but it continuously declined to 20% towards the end. This clade, however, did not respond to diatom-derived OM addition in the M1 experiment and fractions of this clade remained below 20%. A refined analysis of subgroups of the *Roseobacter* group and *Bacteroidetes* showed that the RCA (*Roseobacter* clade affiliated) cluster accounted for 0.8–13% of the total bacterial cells in both experiments, which represented 29–61% of all roseobacters, and the genus *Polaribacter* was up to 8% of total bacterial cells and up to 58% of *Bacteroidetes* (Fig. 2C, D). The SAR92 cluster represented an average of 20% of total *Gammaproteobacteria* in M1, corresponding to 2–7% of total bacterial cells (Supplementary Fig. S1). Due to low abundances of SAR92, this cluster was not analyzed in detail in M2.

Log response ratios underscored that in both mesocosm experiments the *Roseobacter* group was most responsive to the diatom-derived OM additions, exhibiting the highest ratios until day 6 (Supplementary Fig. S2). In M1, *Bacteroidetes* ranked second, whereas in M2 this group and *Gammaproteobacteria* exhibited similar ratios. The SAR11 clade was least responsive in both experiments.

Bacterial community composition (amplicon sequencing data)

The composition of the total and active bacterioplankton community in both mesocosm experiments was assessed by amplicon sequencing. After quality control, removal of chimera and singletons, a total of 1,881,610 bacterial sequences were obtained (DNA: 904,730, RNA: 976,880).

Alpha- and *Gammaproteobacteria* and *Bacteroidetes* constituted 96% of the total (DNA) and 98% of the potentially active (RNA) bacterial communities of both experiments (Fig. 3). *Alphaproteobacteria* accounted for 45% of the total and 52% of the potentially active bacterioplankton community with relatively low proportions of SAR11 and lower ones in the active group than in the total community (Supplementary Table S2). More than 80% of the SAR11 sequences were affiliated to the surface I subclade and the rest to the surface II and IV subclades (Supplementary Table S2). The SAR116 clade also constituted relatively low proportions but responded to the diatom-derived OM additions until day 4 in M1 and day

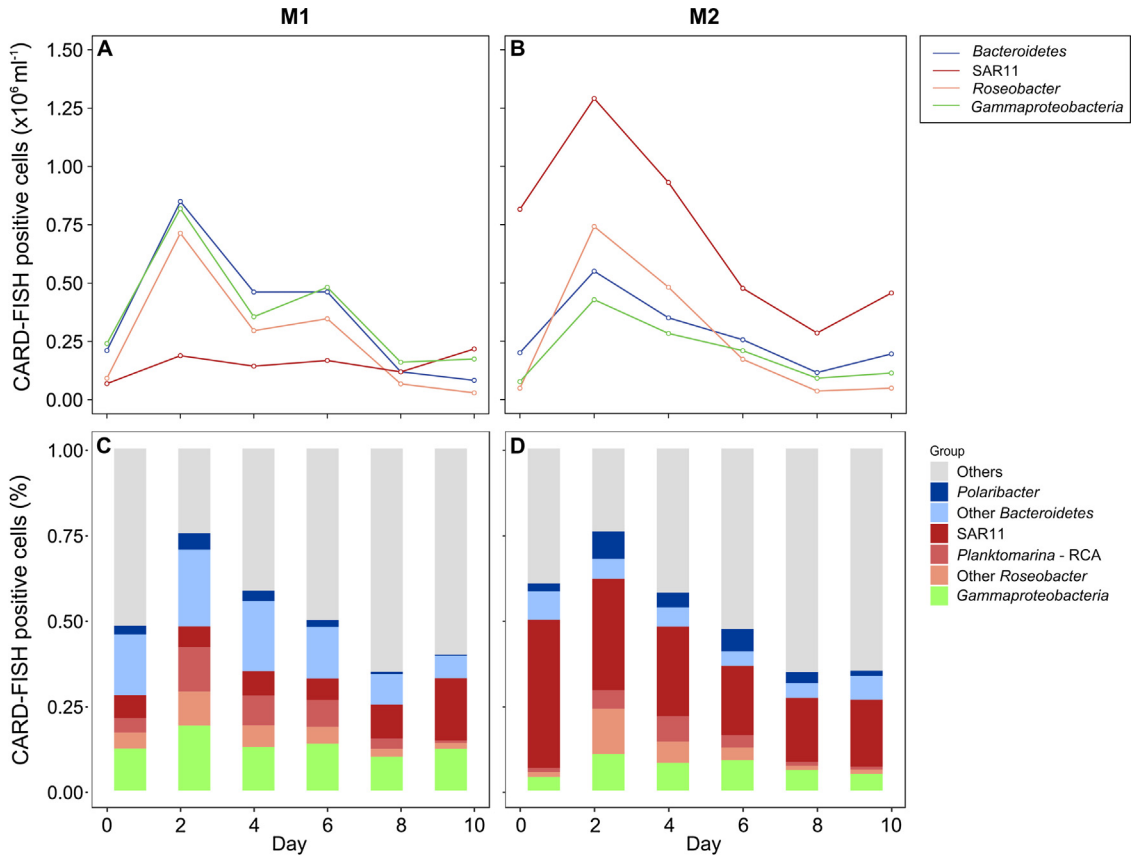

Fig. 2. Community structure of the bacterioplankton, as determined by CARD-FISH in the mesocosms. (A, B) Absolute cell numbers in M1 and M2. (C, D) Percentages of DAPI-stained cells detected by group-specific probes in M1 and M2 (percentage of total DAPI-stained cells).

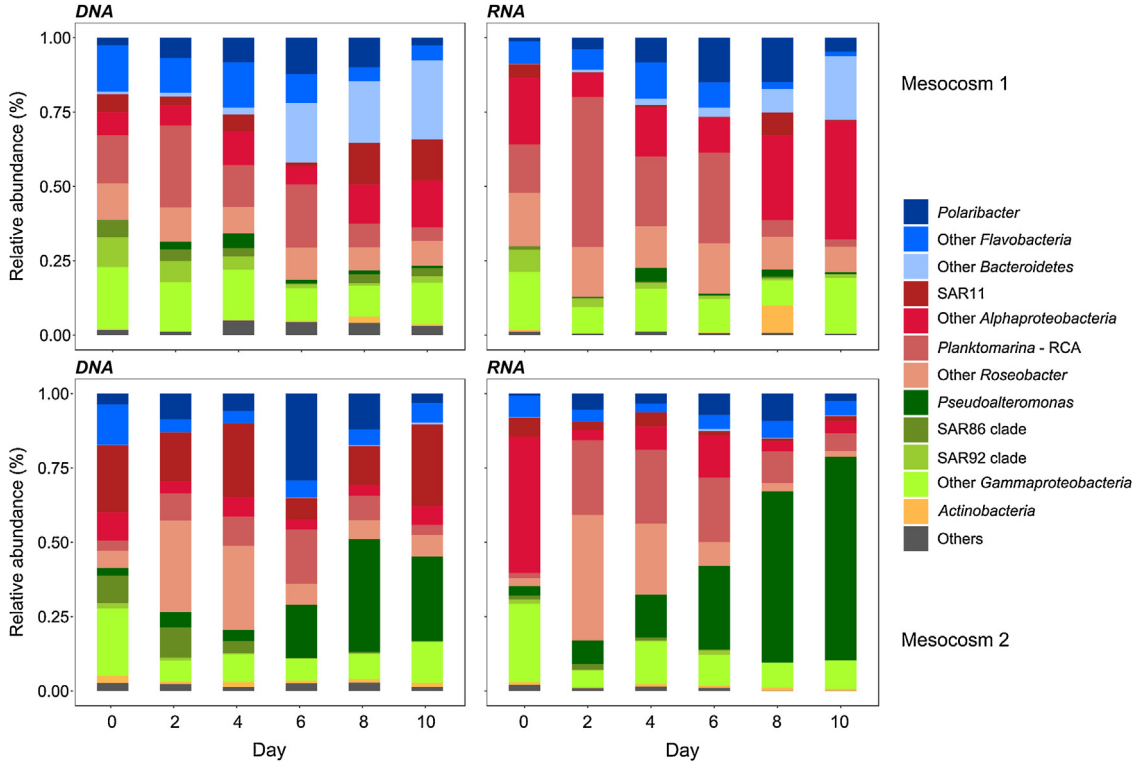

Fig. 3. Relative abundances of total (DNA) and active (RNA) major bacterial lineages in mesocosm experiments 1 and 2 determined by 16S rRNA gene and transcript sequencing.

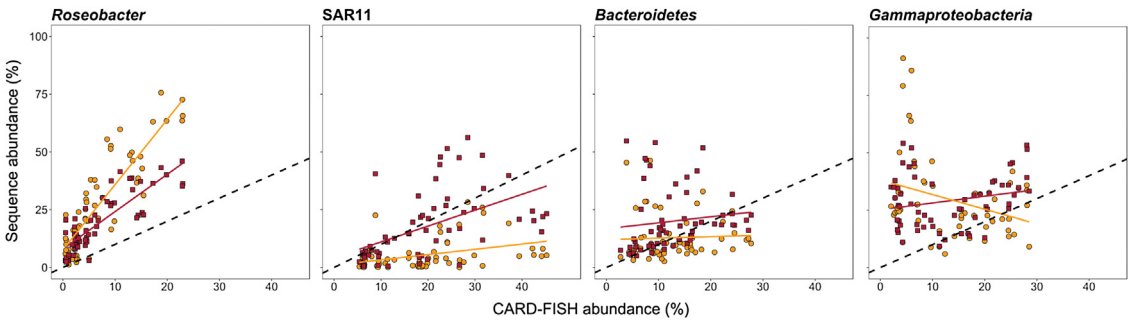

Fig. 4. Regression analysis of the relative abundances of the *Roseobacter* group, SAR11 clade, *Bacteroidetes* and *Gammaproteobacteria* assessed by CARD-FISH *versus* their relative abundance assessed by amplicon sequencing. The dashed line indicates the 1:1 reference relationship.

6 in M2 (Supplementary Table S2). *Caulobacterales* peaked in the active community in M1 towards the end of the experiment, reaching 34.5% of the active community (Supplementary Table S2). The *Roseobacter* group, in contrast, constituted approximately 55% of this proteobacterial class. Its RCA cluster was most responsive of all groups to diatom-derived OM addition, increasing its proportion from day 0 to 2 from 15% to 50% of the active community in M1 and from 2% to 25% in M2 (Fig. 3). Other roseobacters also constituted high proportions during distinct phases of the experiments, such as members of the genus *Ascidiaceihabitans*, close relatives of *Planktotalea frisia*, isolated from and abundant in the North Sea [27,28], which constituted 8% of the total and 12% of the potentially active bacterial communities at days 0 and 2 in M1, as well as members of the genus *Amylibacter* that constituted 24–26% of the total and 18–35% of the potentially active bacterial community at days 2 and 4 in M2 (Supplementary Table S2). *Bacteroidetes*, including its *Polaribacter* cluster, also responded to diatom-derived OM addition but their proportion in the total remained consistently higher than in the potentially active community, with up to 40% of the total and 26% of the potentially active bacterial community at day 6, respectively (Fig. 3, Supplementary Table S2). Towards the end of M2, at days 8 and 10, the *Gammaproteobacteria* genus *Pseudoalteromonas* increased its fraction greatly and reached 28% and 68% of the total and potentially active bacterioplankton, respectively (Fig. 3, Supplementary Table S2).

The relative growth dynamics (*i.e.* increasing and decreasing proportions) of the most important phylogenetic lineages of the total and active bacterial communities are summarized and illustrated in Supplementary Fig. S3.

Table 2
Linear regression statistics (slope, r^2) and significance levels for CARD-FISH-data *versus* amplicon sequencing data of the 16S rRNA gene (DNA) and transcripts (rRNA) of bacterial target groups obtained during two mesocosm experiments (M1, M2) in the North Sea. For graphical illustration see Supplementary Figs. S1 and S4–S6. ns: not significant.

Target group	Slope DNA	Slope rRNA	r^2 DNA	r^2 rRNA	p DNA	p rRNA
M1 + M2						
SAR11	0.68	0.23	0.26	0.12	***	**
Roseobacter	1.62	2.82	0.71	0.76	***	***
RCA	1.70	3.28	0.57	0.74	***	***
Bacteroidetes	0.25	0.06	0.00	−0.2	ns	ns
Polaribacter	1.04	0.80	0.1	0.15	ns	ns
Gammaproteobacteria	0.29	−0.69	0.03	0.09	ns	*
M1						
SAR11	0.89	0.03	0.24	−0.03	**	ns
Roseobacter	1.28	2.32	0.58	0.63	***	***
RCA	1.76	3.36	0.66	0.76	***	***
Bacteroidetes	−0.55	−0.54	0.03	0.08	ns	ns
Polaribacter	−1.07	2.71	−0.03	0.06	ns	ns
Gammaproteobacteria	1.80	0.86	0.63	0.17	***	*
SAR92	1.98	1.11	0.81	0.72	***	***
M2						
SAR11	0.11	0.19	−0.03	0.01	ns	ns
Roseobacter	2.04	3.26	0.83	0.86	***	***
RCA	1.31	3.27	0.26	0.62	**	***
Bacteroidetes	0.39	−0.19	−0.02	−0.03	ns	ns
Polaribacter	2.10	0.67	0.38	0.16	***	*
Gammaproteobacteria	−1.13	−1.19	0.04	−0.01	ns	ns
Pseudoalteromonas	5.32	10.16	0.58	0.62	***	***

* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

Comparison of CARD-FISH and sequencing data

In order to compare directly the CARD-FISH and amplicon sequencing data of the total and potentially active bacterioplankton community, those phylogenetic groups from the amplicon sequencing data sets were selected that matched the groups targeted by CARD-FISH, and a linear regression analysis was carried out by pooling the data of the triplicates of the treatments. This analysis yielded highly significant positive correlations for CARD-FISH and the total as well as potentially active communities of the pooled data from both experiments for the *Roseobacter* group, its RCA cluster and the SAR11 clade (Fig. 4, Table 2, Supplementary Figs. S4 and S5). No significant correlation was found between CARD-FISH and amplicon sequencing data of total *Bacteroidetes* (Fig. S6, Table 2) but there was between CARD-FISH and amplicon sequencing data from the total and potentially active *Polaribacter* cluster (Table 2, Supplementary Fig. S6). For *Gammaproteobacteria*, significant correlations existed only for the entire data set of the potentially active fraction and CARD-FISH, and for experiment M1 (Table 2, Supplementary Fig. S1). For the SAR92 clade and the *Pseudoalteromonas* genus, highly significant correlations were obtained between the data sets of experiments M1 and M2, respectively (Table 2, Supplementary Fig. S1).

The slopes of the linear regressions of the *Roseobacter* group, its RCA cluster, the SAR11 clade and *Bacteroidetes* differed strikingly (Table 2). The slope of the linear regression between the CARD-FISH and amplicon sequencing data of total *Roseobacter* and RCA data was 1.6- and 1.7-fold, and that of the respective data with the potentially active fraction 2.8- and 3.3-fold higher than the 1:1 reference line. On the other hand, the slope of the linear regression between CARD-FISH and amplicon sequencing data of the total SAR11 clade was only 0.68, whereas for the potentially active SAR11 clade it was only 0.23. For the *Polaribacter* genus, correlations between CARD-FISH and the total as well as potentially active fractions were rather similar and close to the 1:1 reference line. Respective correlations of *Bacteroidetes* and *Gammaproteobacteria* and their subgroups, or only for the data from one of the experiments, did not yield such clear cut results (Supplementary Figs. S1 and S5).

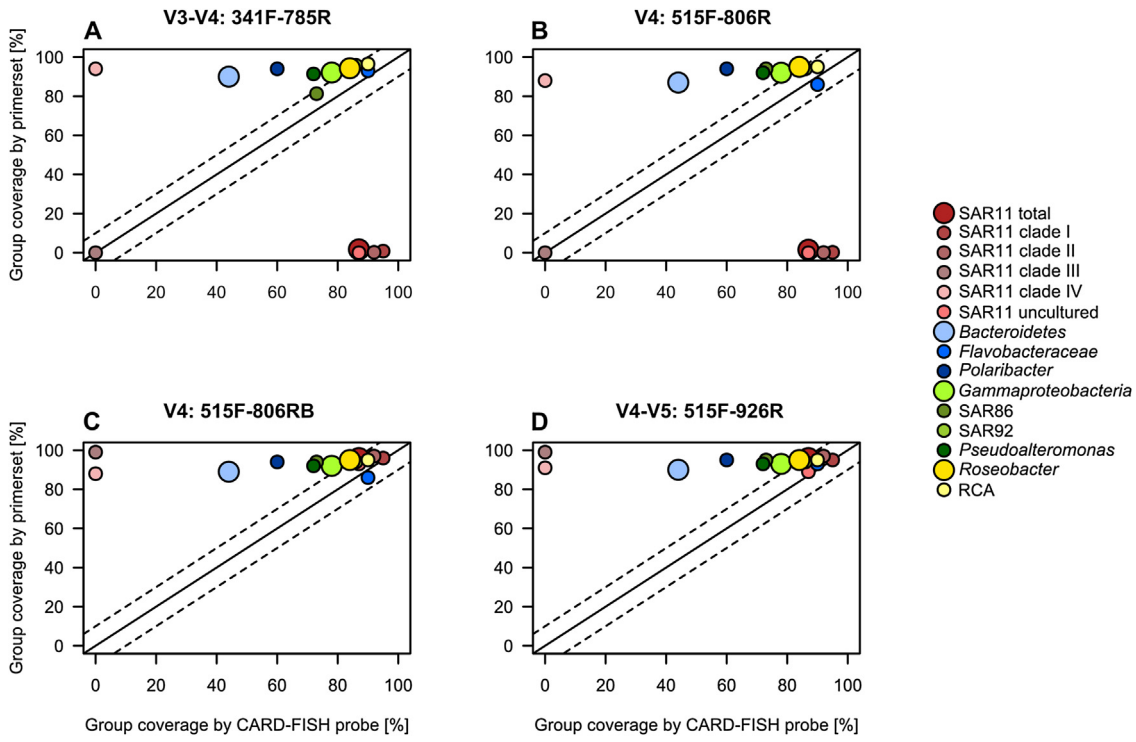

Fig. 5. Coverage of different phylogenetic lineages by primer sets targeting variable regions of the 16S rRNA gene *versus* coverage of different phylogenetic lineages by probes for FISH analyses. Target regions were analyzed by TestPrime and TestProbe. A: primer set targeting the V3–V4 region [33]; B: primer set targeting the V4 region [4]; C: modified primer set targeting the V4 region [4]; D: primer set targeting the V4–V5 region [49]. Target regions of FISH-probes are the same in all panels.

Discussion

Assessing the composition of bacterial communities in aquatic systems by CARD-FISH and amplicon sequencing of the 16S rRNA gene are state of the art approaches, although they have different emphasis regarding absolute and relative quantification, the resolution of the community analysis and relating the data to cell abundances [2,17,49,61,66]. Targeting not only the 16S rRNA gene but also its transcripts provides additional insights into the metabolically or potentially active members of bacterial communities [10,15,51,71,73]. Despite the very wide use of the CARD-FISH and amplicon sequencing approaches, a direct comparison has not been carried out for aquatic systems but it is important to interpret how these approaches reflect the significance of bacterial phylogenetic lineages in organic matter cycling in marine pelagic systems and in a broader ecological context.

To compare these approaches, two mesocosm experiments supplemented with diatom-derived OM were set up to assess the bacterioplankton community composition. The aim of these experiments was to induce changes in the community composition, which implied differential growth responses of the different phylogenetic taxa. As the comparison of these approaches was the main aim of the experiments, the discussion is focused on this comparison and any potential biases.

Potential biases for comparing CARD-FISH and 16S rRNA amplicon sequencing approaches

A major issue for this comparison was the specificity and coverage of the primers applied for amplifying the 16S rRNA gene fragment and the probes of the various target groups. A search with TestProbe revealed that the probes applied targeted 87% of the SAR11 clade, 44% of *Bacteroidetes* but 90% of *Flavobacteriaceae*, the predominant family in marine pelagic systems of this phylum [66]. Furthermore, the probes targeted 78% of *Gammaproteobacteria* and 65% of *Rhodobacteraceae* but 82% of the *Roseobacter* group, the major group of this family (Fig. 5). Looking at a more refined taxonomic level, the SAR11 probe targeted >90% of the sequences of clades I and II but not those of clades III and IV. The *Polaribacter* probe targeted 60% of the sequences of this genus, the SAR86 and SAR92 probes 73% and 86% of the sequences of these clades, respectively, and the RCA probe 90% of those in this cluster (Fig. 5). Hence, the probes applied covered to a large extent the main phylogenetic bacterial groups to be expected in the North Sea and in the experiments [66]. A similar search with TestPrime revealed that the primer set applied, targeting the V3–V4 region of the 16S rRNA gene [33], covered <2% of all SAR11 sequences and <1% of clades I–III but 94% of clade IV and at least 90% of all other target groups except the SAR86 clade that was covered by 81% (Fig. 5A). However, if one mismatch is allowed, the coverage of this primer set increased to 97% of the SAR11 surface I clade [33]. A primer set targeting the V4 region of the 16S rRNA gene introduced by Apprill et al. [4] yielded rather similar proportions of the target groups (Fig. 5B). However, these authors and another study introduced new primer sets targeting the V4 and the V4–V5 regions of the 16S rRNA gene to overcome this bias [4,49]. These new primer sets were generated at critical positions and covered 96% of the SAR11 clade and >91% of all other target groups (Fig. 5C, D). This search analysis showed that the primer set applied strongly discriminated against the major SAR11 subclades occurring in the North Sea [66], whereas all other phylogenetic lineages were covered to a large extent. It is puzzling though that the majority of the SAR11 sequences detected affiliated with the surface I subclade. This finding can be explained by the fact that, if one mismatch is allowed, the coverage of this primer set of the SAR11 surface I clade greatly increased [33] and is in line with other reports of high proportions of the SAR11 surface I clade in the North Sea [33,66].

Other biases regarding CARD-FISH may concern false positive target sequences but this did not appear to be a major issue in the context of the current study as it concerns lineages of *Bacteroidetes*,

excluding *Flavobacteriales* and *Flexibacteriales* [2]. Differential DNA and RNA extraction efficiencies of various phylogenetic groups may have affected the sequencing results of the 16S rRNA gene and rRNA but systematic studies on the groups in the study are not available. Another potential bias regarding the 16S rRNA gene may be that the genomes of different phylogenetic groups vary in the copy number of the rRNA operon. Efforts to correct for this bias have been proposed [3,31] but it is difficult to correct for lineages for which reference genomes are not available. The genomes of the major groups of *Alphaproteobacteria* relevant for the study, the SAR11 clade and the pelagic members of the *Roseobacter* group, encompass 1–2 rRNA operons [7,8,26,68]. Genomes of *Flavobacteriaceae*, the major family of *Bacteroidetes* in marine waters, and the *Gammaproteobacteria* genera *Alteromonas* and *Pseudoalteromonas*, encompass up to four and five rRNA operons, respectively [15,18,23,37]. This suggests that amplicon sequencing of the 16S rRNA gene should provide rather similar proportions for the SAR clade and pelagic roseobacters as cell-based approaches but may differ in cell-based abundances of *Bacteroidetes* and *Gammaproteobacteria*, when enhanced numbers of the rRNA operon dominate these bacterial groups.

Growth of bacterial target groups in the mesocosms and comparison of both approaches

The four bacterial target groups responded to diatom-derived OM addition but the growth dynamics differed among the groups and between both mesocosm experiments. Based on the absolute numbers obtained by CARD-FISH abundances combined with bacterial cell numbers, the *Roseobacter* group was most responsive, followed by *Bacteroidetes* and *Gammaproteobacteria*, whereas the SAR11 clade showed only weak responses. These results comply with the generally important role of these groups in organic matter processing during phytoplankton blooms [13,66]. The relative abundances of the major sublineages of *Alphaproteobacteria*, SAR11 and the *Roseobacter* group, assessed by both approaches, were highly significantly linearly correlated, albeit in different ways. For the SAR11 clade, amplicon sequencing data of the 16S rRNA gene yielded lower proportions than CARD-FISH data, and even the transcripts were only 23% of the CARD-FISH data. As the SAR11 clade has only one rRNA operon [26], a ratio of 1.0 would be expected between amplicon sequencing and CARD-FISH data in the ideal case when other lineages also had only one rRNA operon. The lower value may be due to the discrimination against the SAR11 subclade surface I of the primer set applied (see above). The even lower ratio of the potentially active fraction of the SAR11 clade over the CARD-FISH data may also be partially affected by this primer bias, although it also indicates that cells of the SAR11 clade contained low amounts of rRNA and thus rather inactive cells. Previous studies have already shown that the SAR11 clade usually exhibits rRNA/rDNA ratios <1 [15,16,30,72] and lower proportions of cell-proliferating cells than that of the total bacterioplankton community [7,63,64], as well as in metatranscriptomic relative to metagenomic reads [25,68]. In a study with *Candidatus* Pelagibacter ubique of the SAR11 clade, Lankiewicz et al. [35] found an inverse correlation between the rRNA/rDNA ratio and a growth rate below 0.5 per day, complying with the findings of the current study. Hence the results of the correlation analysis of the CARD-FISH with the amplicon sequencing data of the total SAR11 clade, as well as with the potentially active fraction of this clade, were consistent with previous findings and underlined that CARD-FISH data were more in line with amplicon sequencing data of the total SAR11 clade than with that of its potentially active fraction.

In contrast to the SAR11 clade, for the *Roseobacter* group and its major component in temperate seas, the RCA cluster [60,68], amplicon sequencing data of the total and potentially active group were much higher than CARD-FISH data, implying that the latter approach greatly underestimated the biogeochemical significance of this bacterial group. Applying CARD-FISH data, the *Roseobacter* group typically yields 5–10% of the total bacterioplankton community [6,45,63,64]. Considering the correlation, it was established that these numbers translated into proportions of 15–30% of the potentially active bacterioplankton community, emphasizing the important role this bacterial group plays in organic matter cycling in marine pelagic systems. In a pure culture study with *Ruegeria pomeroyi* of the *Roseobacter* group, a positive correlation between the rRNA/rDNA ratio and a growth rate of up to 2.6 per day was established [35], in agreement with our findings at a community level. These findings indicated that for the *Roseobacter* group and its major pelagic component, the RCA cluster, CARD-FISH and 16S rRNA amplicon sequence data were correlated but the latter approach yielded higher proportions, in particular considering the potentially active fraction.

It may appear puzzling that the rDNA/CARD-FISH ratio of the *Roseobacter* group and its RCA cluster was >1, despite the fact that the genome of *Planktomarina temperata* RCA, the type strain of this cluster, encompasses only one rRNA operon [67]. As the genomes of other pelagic roseobacters encompass either one or two rRNA operons [7,8,47], this ratio of the entire *Roseobacter* community may reflect this fact, except in the case of the RCA cluster. For the latter cluster, a ratio >1 may indicate that its members continuously encompass more than 1 rRNA operon per cell because of continuous DNA replication. Oligotrophic, slowly growing bacteria, such as mycobacteria, and not only fast-growing bacteria such as *Escherichia coli*, continuously amplify their DNA by multi-fork replication [67]. This observation makes it plausible that even a bacterium with only one rRNA operon, such as *P. temperata*, may contain more rRNA operons, leading to a ratio of rDNA/CARD-FISH data >1. The fact that the rRNA/rDNA ratio and also the rRNA/CARD-FISH ratio exceeded 1 reflects the very high metabolic activity of the RCA cluster, as already reported in a previous study [6].

For *Bacteroidetes*, no significant correlation existed between CARD-FISH and amplicon sequencing data, neither for the total nor for the active community, and the data varied greatly. In addition, for *Gammaproteobacteria* there was no significant correlation for the entire data set and experiment M2 but there was for M1. The main reason presumably is that both groups encompassed different phylogenetic subgroups, which responded differently to the diatom-derived OM additions. Consistent with this assumption was the fact that for narrower subgroups of both phylogenetic groups significant linear correlations existed, such as for the genus *Polaribacter* in experiment M2, the SAR92 clade in M1 and the genus *Pseudoalteromonas* in M2. In these mesocosms, these subgroups constituted substantial proportions of *Bacteroidetes* and *Gammaproteobacteria*. The fact that for the SAR92 clade in experiment M1 the slope of the data set with the total population (16S rDNA) was higher than that with the potentially active fraction, and respective correlations of the *Pseudoalteromonas* genus in experiment M2 exhibited opposite features, further illustrated that for total *Gammaproteobacteria* no correlation existed for the entire data set. This analysis thus explains why more correlations were found between CARD-FISH and 16S amplicon sequence data for phylogenetic lineages on a higher taxonomic level than for large phylogenetic groups. The varying numbers of rRNA operons in the genomes of different phylogenetic lineages may also affect these correlations. Depending on the activity status of the cells and the number of rRNA operons CARD-FISH may reflect more the total or the active fraction of a given phylogenetic lineage, although CARD-FISH data alone did not allow any conclusion to be made for the activity status. In addition to (sub)group-specific genomic variances in 16S rRNA operons, the compositional nature of amplicon data *per se* may alter the overall relative abundance

of certain groups if the relative proportions of others increase or decrease.

rRNA content of potentially active bacteria and growth status of marine bacteria

A critical review on the significance of rRNA-based amplicon sequencing data presents evidence that no clear-cut information exists that a high cellular amount of rRNA reflects high rates of metabolic activities [10]. Therefore, it was proposed to call bacteria with a high cellular rRNA content not active but potentially active. Furthermore, the review presented evidence from compiled studies that the cellular rRNA content does not directly correlate with bacterial growth rates. Dormant cells (*e.g.* in soil) have also been found to contain enhanced amounts of rRNA because it is not degraded. Most of the studies that serve as a basis for the arguments of this critical review [10] are based on pure cultures and soil. In our experiments and marine bacterioplankton communities in general, it is highly improbable that dormant cells contain enhanced amounts of rRNA. Protozoan grazing, which is size-dependent, would eliminate large bacteria rapidly in suspension but not when bacteria could escape in a structured soil and root habitat [57]. Hence, dormant or less active bacteria in marine pelagic systems are smaller than active bacteria and are often considered to be LNA bacteria [42]. However, this does not rule out that the LNA bacterial fraction would also contain even fairly active bacteria [35,42].

Information on a positive correlation between the cellular rRNA content and bacterial growth rate is ambiguous. Lankiewicz et al. [35], however, showed that the growth rate of various copiotrophic marine bacteria was positively correlated to the rRNA/rDNA ratio, although oligotrophic bacteria, such as SAR11 and SAR92, exhibited much lower ratios and growth rates. There is some evidence that the two alphaproteobacterial groups that in our study showed contrasting correlations between CARD-FISH and amplicon sequencing data, SAR11 and the *Roseobacter* group, also exhibited strikingly different growth rates. In a recent study, Bakenhus et al. [6] estimated that *in situ* growth rates of the SAR11 clade were more than 20-fold lower than those of the *Roseobacter* group. Despite the high growth rate the latter group exhibited low abundances (<10% of total bacteria) presumably because of high grazing pressure by protozoans.

One of the conclusions from these studies, and in agreement with a recent review on bacterial growth rates in the ocean [32], is that in marine pelagic systems, where grazing removes large bacteria, slow growing and dormant bacteria are small and, as such, the arguments of Blazewicz et al. [10] presumably do not hold true. This implies that rRNA/rDNA ratios or marine bacteria do reflect their activity status unless the respective bacteria have a high number of rRNA operons.

Conclusion

The mesocosm experiments showed that CARD-FISH and amplicon sequencing data of the major alphaproteobacterial marine bacterioplankton groups, the *Roseobacter* group and SAR11, were highly significantly correlated, but differently correlated regarding the total and active communities. Other major bacterioplankton groups, such as *Bacteroidetes* and *Gammaproteobacteria*, did not exhibit such correlations. In contrast, the respective data of subgroups at the genus level did show correlation, although only in data subsets. These results are the first that have directly compared both state of the art approaches to assess the composition of bacterioplankton communities. They shed new light on how CARD-FISH and amplicon sequencing data of the total and potentially active fraction of the major bacterioplankton groups and subgroups are related. The results indicated that CARD-FISH and amplicon sequencing data reflected different aspects of how the different groups contribute to the communities regarding their quantity and activity, as well as their relevance in organic matter degradation, in particular regarding the different roles of the *Roseobacter* group and the SAR11 clade in marine pelagic systems. In order to obtain a better insight into the roles of other members of bacterial communities, future work needs to apply this comparative approach to other bacterial lineages in marine and other ecosystems.

Acknowledgments

We thank the scientists and the crew aboard R/V Heincke for their excellent support during the cruise (AWI-HE425), and Bernhard Fuchs for critical comments on an earlier version of this manuscript. This work was supported by Deutsche Forschungsgemeinschaft (DFG) within the Collaborative Research Centre *Roseobacter* (TRR51).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.syapm.2019.06.005.

References

[1] Allers, E., Moraru, C., Duhaime, M.B., Beneze, E., Solonenko, N., Barrero-Canosa, J., Amann, R., Sullivan, M.B. (2013) Single-cell and population level viral infection dynamics revealed by phageFISH, a method to visualize intracellular and free viruses. Environ. Microbiol. 15, 2306–2318, http://dx.doi.org/10.1111/1462-2920.12100.

[2] Amann, R., Fuchs, B.M. (2008) Single-cell identification in microbial communities by improved fluorescence *in situ* hybridization techniques. Nat. Rev. Microbiol. 6, 339–348, http://dx.doi.org/10.1038/nrmicro1888.

[3] Angly, F.E., Dennis, P.G., Skarshewski, A., Vanwonterghem, I., Hugenholtz, P., Tyson, G.W. (2014) CopyRighter: a rapid tool for improving the accuracy of microbial community profiles through lineage-specific gene copy number correction. Microbiome 2, 1–13, http://dx.doi.org/10.1186/2049-2618-2-11.

[4] Apprill, A., Mcnally, S., Parsons, R., Weber, L. (2015) Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. Aquat. Microb. Ecol. 75, 129–137, http://dx.doi.org/10.3354/ame01753.

[5] Azam, F., Malfatti, F. (2007) Microbial structuring of marine ecosystems. Nat. Rev. Microbiol. 5, 782–791, http://dx.doi.org/10.1038/nrmicro1747.

[6] Bakenhus, I., Dlugosch, L., Billerbeck, S., Giebel, H.A., Milke, F., Simon, M. (2017) Composition of total and cell-proliferating bacterioplankton community in early summer in the North Sea — roseobacters are the most active component. Front. Microbiol. 8, 1–14, http://dx.doi.org/10.3389/fmicb.2017.01771.

[7] Bakenhus, I., Voget, S., Poehlein, A., Brinkhoff, T., Daniel, R., Simon, M. (2018) Genome sequence of *Planktotalea frisia* type strain (SH6-1T), a representative of the *Roseobacter* group isolated from the North Sea during a phytoplankton bloom. Stand. Genomic Sci. 13, 1–9, http://dx.doi.org/10.1186/s40793-018-0311-5.

[8] Billerbeck, S., Wemheuer, B., Voget, S., Poehlein, A., Giebel, H.A., Brinkhoff, T., Gram, L., Jeffrey, W.H., Daniel, R., Simon, M. (2016) Biogeography and environmental genomics of the *Roseobacter*-affiliated pelagic CHAB-I-5 lineage. Nat. Microbiol. 1, 16063, http://dx.doi.org/10.1038/nmicrobiol.2016.63.

[9] Bižić-Ionescu, M., Zeder, M., Ionescu, D., Orlić, S., Fuchs, B.M., Grossart, H.P., Amann, R. (2015) Comparison of bacterial communities on limnic versus coastal marine particles reveals profound differences in colonization. Environ. Microbiol. 17, 3500–3514, http://dx.doi.org/10.1111/1462-2920.12466.

[10] Blazewicz, S.J., Barnard, R.L., Daly, R.A., Firestone, M.K. (2013) Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. ISME J. 7, 2061–2068, http://dx.doi.org/10.1038/ismej.2013.102.

[11] Bolger, A.M., Lohse, M., Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120, http://dx.doi.org/10.1093/bioinformatics/btu170.

[12] Brinkmeyer, R., Rappé, M., Gallacher, S., Medlin, L. (2000) Development of clade- (*Roseobacter* and *Alteromonas*) and taxon-specific oligonucleotide probes to study interactions between toxic dinoflagellates and their associated bacteria. Eur. J. Phycol. 35, 315–329, http://dx.doi.org/10.1080/09670260010001735921.

[13] Buchan, A., LeCleir, G.R., Gulvik, C.A., González, J.M. (2014) Master recyclers: features and functions of bacteria associated with phytoplankton blooms. Nat. Rev. Microbiol. 12, 686–698, http://dx.doi.org/10.1038/nrmicro3326.

[14] Campbell, B.J., Yu, L., Straza, T.R.A., Kirchman, D.L. (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.

[15] Campbell, B.J., Yu, L., Heidelberg, J.F., Kirchman, D.L. (2011) Activity of abundant and rare bacteria in a coastal ocean. Proc. Natl. Acad. Sci. 108, 12776–12781, http://dx.doi.org/10.1073/pnas.1101405108.
[16] Campbell, B.J., Kirchman, D.L. (2013) Bacterial diversity, community structure and potential growth rates along an estuarine salinity gradient. ISME J. 7, 210–220, http://dx.doi.org/10.1038/ismej.2012.93.
[17] Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S.M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J.A., Smith, G., Knight, R. (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. 6, 1621–1624, http://dx.doi.org/10.1038/ismej.2012.8.
[18] Choe, H., See, S.H., Kim, S.G., Park, D.S., Nasir, A., Kim, K.M. (2016) Complete genome of *Pseudoalteromonas phenolica* KCTC 12086(T) (= O-BC30(T)), a marine bacterium producing polybrominated aromatic compounds. J. Biotechnol. 218, 23–24, http://dx.doi.org/10.1016/j.jbiotec.2015.11.028.
[19] Conover, W.J., Johnson, M.E., Johnson, M.M. (1981) A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data. Technometrics 23, 351–361, http://dx.doi.org/10.1080/00401706.1981.10487680.
[20] Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26, 2460–2461.
[21] Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R. (2011) UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27, 2194–2200, http://dx.doi.org/10.1093/bioinformatics/btr381.
[22] Eilers, H., Pernthaler, J., Glöckner, F.O., Amann, R. (2000) Culturability and *in situ* abundance of pelagic *Bacteria* from the North Sea. Appl. Environ. Microbiol. 66, 3044–3051, http://dx.doi.org/10.1128/AEM.66.7.3044-3051.2000.
[23] Fernández-Gómez, B., Richter, M., Schüler, M., Pinhassi, J., Acinas, S.G., González, J.M., Pedrós-Alió, C. (2013) Ecology of marine *Bacteroidetes*: a comparative genomics approach. ISME J. 7, 1026–1037, http://dx.doi.org/10.1038/ismej.2012.169.
[24] Ghiglione, J.-F., Galand, P.E., Pommier, T., Pedros-Alio, C., Maas, E.W., Bakker, K., Bertilson, S., Kirchman, D.L., Lovejoy, C., Yager, P.L., Murray, A.E. (2012) Pole-to-pole biogeography of surface and deep marine bacterial communities. Proc. Natl. Acad. Sci. 109, 17633–17638, http://dx.doi.org/10.1073/pnas.1208160109.
[25] Gifford, S.M., Sharma, S., Moran, M.A. (2014) Linking activity and function to ecosystem dynamics in a coastal bacterioplankton community. Front. Microbiol. 5, 1–12, http://dx.doi.org/10.3389/fmicb.2014.00185.
[26] Giovannoni, S.J., Tripp, H.J., Givan, S., Podar, M., Vergin, K.L., Baptista, D., Bibbs, L., Eads, J., Richardson, T.H., Noordewier, M., Rappé, M.S., Short, J.M., Carrington, J.C., Mathur, E.J. (2005) Genetics: genome streamlining in a cosmopolitan oceanic bacterium. Science 309, 1242–1245, http://dx.doi.org/10.1126/science.1114057.
[27] Hahnke, S., Tindall, B.J., Schumann, P., Sperling, M., Brinkhoff, T., Simon, M. (2012) *Planktotalea frisia* gen. nov., sp. nov., isolated from the southern North Sea. Int. J. Syst. Evol. Microbiol. 62, 1619–1624, http://dx.doi.org/10.1099/ijs.0.033563-0.
[28] Hahnke, S., Sperling, M., Langer, T., Wichels, A., Gerdts, G., Beardsley, C., Brinkhoff, T., Simon, M. (2013) Distinct seasonal growth patterns of the bacterium *Planktotalea frisia* in the North Sea and specific interaction with phytoplankton algae. FEMS Microbiol. Ecol. 86, 185–199, http://dx.doi.org/10.1111/1574-6941.12151.
[29] Hong, S., Bunge, J., Leslin, C., Jeon, S., Epstein, S.S. (2009) Polymerase chain reaction primers miss half of rRNA microbial diversity. ISME J. 3, 1365–1373, http://dx.doi.org/10.1038/ismej.2009.89.
[30] Hunt, D.E., Lin, Y., Church, M.J., Karl, D.M., Tringe, S.G., Izzo, L.K., Johnson, Z.I. (2013) Relationship between abundance and specific activity of bacterioplankton in open ocean surface waters. Appl. Environ. Microbiol. 79, 177–184, http://dx.doi.org/10.1128/AEM.02155-12.
[31] Kembel, S.W., Wu, M., Eisen, J.A., Green, J.L. (2012) Incorporating 16S gene copy number information improves estimates of microbial diversity and abundance. PLoS Comput. Biol. 8, 16–18, http://dx.doi.org/10.1371/journal.pcbi.1002743.
[32] Kirchman, D.L. (2016) Growth rates of microbes in the oceans. Ann. Rev. Mar. Sci. 8, 285–309, http://dx.doi.org/10.1146/annurev-marine-122414-033938.
[33] Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., Glöckner, F.O. (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res. 41, e1, http://dx.doi.org/10.1093/nar/gks808.
[34] Klindworth, A., Mann, A.J., Huang, S., Wichels, A., Quast, C., Waldmann, J., Teeling, H., Glöckner, F.O. (2014) Diversity and activity of marine bacterioplankton during a diatom bloom in the North Sea assessed by total RNA and pyrotag sequencing. Mar. Genom. 18, 185–192, http://dx.doi.org/10.1016/j.margen.2014.08.007.
[35] Lankiewicz, T.S., Cottrell, M.T., Kirchman, D.L. (2016) Growth rates and rRNA content of four marine bacteria in pure cultures and in the Delaware estuary. ISME J. 10, 823–832, http://dx.doi.org/10.1038/ismej.2015.156.
[36] Lee, C.K., Herbold, C.W., Polson, S.W., Wommack, K.E., Williamson, S.J., McDonald, I.R., Cary, S.C. (2012) Groundtruthing next-gen sequencing for microbial ecology-biases and errors in community structure estimates from PCR amplicon pyrosequencing. PLoS One 7, 15–17, http://dx.doi.org/10.1371/journal.pone.0044224.
[37] López-Pérez, M., Gonzaga, A., Martin-Cuadrado, A.B., López-García, P., Rodriguez-Valera, F., Kimes, N.E. (2013) Intra- and intergenomic variation of ribosomal RNA operons in concurrent *Alteromonas macleodii* strains. Microb. Ecol. 65, 720–730, http://dx.doi.org/10.1007/s00248-012-0153-4.
[38] Malmstrom, R., Straza, T., Cottrell, M., Kirchman, D. (2007) Diversity, abundance, and biomass production of bacterial groups in the western Arctic Ocean. Aquat. Microb. Ecol. 47, 45–55, http://dx.doi.org/10.3354/ame047045.
[39] Manz, W., Amann, R., Ludwig, W., Wagner, M., Schleifer, K.H. (1992) Phylogenetic oligodeoxynucleotide probes for the major subclasses of *Proteobacteria*: problems and solutions. Syst. Appl. Microbiol. 15, 593–600, http://dx.doi.org/10.1016/S0723-2020(11)80121-9.
[40] 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.
[41] Milici, M., Tomasch, J., Wos-Oxley, M.L., Decelle, J., Jáuregui, R., Wang, H., Deng, Z.L., Plumeier, I., Giebel, H.A., Badewien, T.H., Wurst, M., Pieper, D.H., Simon, M., Wagner-Döbler, I. (2016) Bacterioplankton biogeography of the Atlantic ocean: a case study of the distance-decay relationship. Front. Microbiol. 7, 1–15, http://dx.doi.org/10.3389/fmicb.2016.00590.
[42] Morán, X.A.G., Ducklow, H.W., Erickson, M. (2011) Single-cell physiological structure and growth rates of heterotrophic bacteria in a temperate estuary (Waquoit Bay, Massachusetts). Limnol. Oceanogr. 56, 37–48, http://dx.doi.org/10.4319/lo.2011.56.1.0037.
[43] Moran, M.A. (2015) The global ocean microbiome. Science 350, http://dx.doi.org/10.1126/science.aac8455, aac8455.
[44] Moraru, C., Lam, P., Fuchs, B.M., Kuypers, M.M.M., Amann, R. (2010) GeneFISH — an *in situ* technique for linking gene presence and cell identity in environmental microorganisms. Environ. Microbiol. 12, 3057–3073, http://dx.doi.org/10.1111/j.1462-2920.2010.02281.x.
[45] Morris, R.M., Frazar, C.D., Carlson, C.A. (2012) Basin-scale patterns in the abundance of SAR11 subclades, marine *Actinobacteria* (OM1), members of the *Roseobacter* clade and OCS116 in the South Atlantic. Environ. Microbiol. 14, 1133–1144, http://dx.doi.org/10.1111/j.1462-2920.2011.02694.x.
[46] Musat, N., Foster, R., Vagner, T., Adam, B., Kuypers, M.M.M. (2012) Detecting metabolic activities in single cells, with emphasis on nanoSIMS. FEMS Microbiol. Rev. 36, 486–511, http://dx.doi.org/10.1111/j.1574-6976.2011.00303.x.
[47] Newton, R.J., Griffin, L.E., Bowles, K.M., Meile, C., Gifford, S., Givens, C.E., Howard, E.C., King, E., Oakley, C.A., Reisch, C.R., Rinta-Kanto, J.M., Sharma, S., Sun, S., Varaljay, V., Vila-Costa, M., Westrich, J.R., Moran, M.A. (2010) Genome characteristics of a generalist marine bacterial lineage. ISME J. 4, 784–798, http://dx.doi.org/10.1038/ismej.2009.150.
[48] Osterholz, H., Niggemann, J., Giebel, H.A., Simon, M., Dittmar, T. (2015) Inefficient microbial production of refractory dissolved organic matter in the ocean. Nat. Commun. 6, 7422, http://dx.doi.org/10.1038/ncomms8422.
[49] Parada, A.E., Needham, D.M., Fuhrman, J.A. (2016) Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. Environ. Microbiol. 18, 1403–1414, http://dx.doi.org/10.1111/1462-2920.13023.
[50] Pernthaler, A., Pernthaler, J., Amann, R. (2002) Fluorescence *in situ* hybridization and catalyzed reporter deposition for the identification of marine bacteria. Appl. Environ. Microbiol. 68, 3094–3101, http://dx.doi.org/10.1128/AEM.68.6.3094.
[51] Petersen, J. (2011) Phylogeny and compatibility: plasmid classification in the genomics era. Arch. Microbiol. 193, 313–321, http://dx.doi.org/10.1007/s00203-011-0686-9.
[52] Pinto, A.J., 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.
[53] Pohlner, M., Degenhardt, J., von Hoyningen-Huene, A.J.E., Wemheuer, B., Erlmann, N., Schnetger, B., Badewien, T.H., Engelen, B. (2017) The biogeographical distribution of benthic roseobacter group members along a Pacific transect is structured by nutrient availability within the sediments and primary production in different oceanic provinces. Front. Microbiol. 8, 1–14, http://dx.doi.org/10.3389/fmicb.2017.02550.
[54] Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O. (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 41, 590–596, http://dx.doi.org/10.1093/nar/gks1219.
[55] Rappé, M.S., Connon, S.A., Vergin, K.L., Giovannoni, S.J. (2002) Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. Nature 418, 630–633, http://dx.doi.org/10.1038/nature00917.
[56] R Core Team, 2017 R: a Language and Environment for Statistical Computing, R foundation for Statistical Computing, Vienna, Austria https://www.R-project.org/.
[57] Rønn, R., Mccaig, A.E., Griffiths, B.S., Prosser, J.I. (2002) Impact of protozoan grazing on bacterial community structure in soil microcosms. Appl. Environ. Microbiol. 68, 6094–6105, http://dx.doi.org/10.1128/AEM.68.12.6094.
[58] Schattenhofer, M., Fuchs, B.M., Amann, R., Zubkov, M.V., Tarran, G.A., 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.
[59] Schneider, D., Wemheuer, F., Pfeiffer, B., Wemheuer, B. (2017) Extraction of total DNA and RNA from marine filter samples and generation of a cDNA as universal template for marker gene studies. In: Streit, W.R., Daniel, R. (Eds.), Metagenomics: Methods and Protocols, 2nd ed., Springer-Verlag, New York, NY, US, pp. 13–22.

[60] Selje, N., Simon, M., Brinkhoff, T. (2004) A newly discovered *Roseobacter* cluster in temperate and polar oceans. Nature 427, 445–448, http://dx.doi.org/10.1038/nature02272.

[61] Sogin, M.L., Morrison, H.G., Huber, J.A., Mark Welch, D., Huse, S.M., Neal, P.R., Arrieta, J.M., Herndl, G.J. (2006) Microbial diversity in the deep sea and the underexplored "rare biosphere". Proc. Natl. Acad. Sci. U. S. A. 103, 12115–12120, http://dx.doi.org/10.1073/pnas.0605127103.

[62] Stingl, U., Desiderio, R.A., Cho, J.C., Vergin, K.L., Giovannoni, S.J. (2007) The SAR92 clade: an abundant coastal clade of culturable marine bacteria possessing proteorhodopsin. Appl. Environ. Microbiol. 73, 2290–2296, http://dx.doi.org/10.1128/AEM.02559-06.

[63] Tada, Y., Taniguchi, A., Nagao, I., Miki, T., Uematsu, M., Tsuda, A., Hamasaki, K. (2011) Differing growth responses of major phylogenetic groups of marine bacteria to natural phytoplankton blooms in the Western North Pacific Ocean. Appl Environ. Microbiol 77, 4055–4065.

[64] Tada, Y., Makabe, R., Kasamatsu-Takazawa, N., Taniguchi, A., Hamasaki, K. (2013) Growth and distribution patterns of *Roseobacter*/*Rhodobacter*, SAR11, and *Bacteroidetes* lineages in the Southern Ocean. Polar Biol. 36, 691–704, http://dx.doi.org/10.1007/s00300-013-1294-8.

[65] Takeuchi, Y., Chaffron, S., Salcher, M.M., Shimizu-Inatsugi, R., Kobayashi, M.J., Diway, B., von Mering, C., Pernthaler, J., Shimizu, K.K. (2015) Bacterial diversity and composition in the fluid of pitcher plants of the genus *Nepenthes*. Syst. Appl. Microbiol. 38, 330–339, http://dx.doi.org/10.1016/j.syapm.2015.05.006.

[66] Teeling, H., Fuchs, B.M., Bennke, C.M., Krüger, K., Chafee, M., Kappelmann, L., Reintjes, G., Waldmann, J., Quast, C., Glöckner, F.O., Lucas, J., Wichels, A., Gerdts, G., Wiltshire, K.H., Amann, R.I. (2016) Recurring patterns in bacterioplankton dynamics during coastal spring algae blooms. Elife 5, 1–31, http://dx.doi.org/10.7554/eLife.11888.

[67] Trojanowski, D., Hołówka, J., Ginda, K., Jakimowicz, D., Zakrzewska-Czerwińska, J. (2017) Multifork chromosome replication in slow-growing bacteria. Sci. Rep. 7, 1–7, http://dx.doi.org/10.1038/srep43836.

[68] Voget, S., Wemheuer, B., Brinkhoff, T., Vollmers, J., Dietrich, S., Giebel, H.A., Beardsley, C., Sardemann, C., Bakenhus, I., Billerbeck, S., Daniel, R., Simon, M. (2015) Adaptation of an abundant *Roseobacter* RCA organism to pelagic systems revealed by genomic and transcriptomic analyses. ISME J. 9, 371–384, http://dx.doi.org/10.1038/ismej.2014.134.

[69] Wear, E.K., Wilbanks, E.G., Nelson, C.E., Carlson, C.A. (2018) Primer selection impacts specific population abundances but not community dynamics in a monthly time-series 16S rRNA gene amplicon analysis of coastal marine bacterioplankton. Environ. Microbiol. 20, 2709–2726, http://dx.doi.org/10.1111/1462-2920.14091.

[70] Wemheuer, B., Wemheuer, F., Hollensteiner, J., Meyer, F.D., Voget, S., Daniel, R. (2015) The green impact: bacterioplankton response towards a phytoplankton spring bloom in the southern North Sea assessed by comparative metagenomic and metatranscriptomic approaches. Front. Microbiol. 6, 1–13, http://dx.doi.org/10.3389/fmicb.2015.00805.

[71] Wemheuer, B., Wemheuer, F., Meier, D., Billerbeck, S., Giebel, H.-A., Simon, M., Scherber, C., Daniel, R. (2017) Linking compositional and functional predictions to decipher the biogeochemical significance in DFAA turnover of abundant bacterioplankton lineages in the North Sea. Microorganisms 5, 68, http://dx.doi.org/10.3390/microorganisms5040068.

[72] Wietz, M., Gram, L., Jørgensen, 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.

[73] Zhang, Y., Zhao, Z., Dai, M., Jiao, N., Herndl, G.J. (2014) Drivers shaping the diversity and biogeography of total and active bacterial communities in the South China Sea. Mol. Ecol. 23, 2260–2274, http://dx.doi.org/10.1111/mec.12739.

[74] Zubkov, M.V., Fuchs, B.M., Burkill, P.H., Amann, R. (2001) Comparison of cellular and biomass specific activities of dominant bacterioplankton groups in stratified waters of the Celtic Sea. Appl. Environ. Microbiol. 67, 5210–5218, http://dx.doi.org/10.1128/AEM.67.11.5210-5218.2001.