Spatiotemporal dynamics of the bacterial community in the German Bight

Dissertation

Zur Erlangung der Würde des

Doktors der Naturwissenschaften

- Dr. rer.nat. –

Dem Fachbereich Biologie/Chemie der

Universität Bremen

vorgelegt von

Judith Lucas

Bremen

September 2015

Die vorliegende Arbeit wurde in der Zeit von Januar 2012 bis September 2015 an der Biologischen Anstalt Helgoland, Alfred-Wegener-Institut Helmholtz Zentrum für Polar- und Meeresforschung angefertigt.

1. Gutachter: Prof. Dr. Rudolf Amann
2. Gutachter: Dr. Gunnar Gerdts

1. Prüfer: Prof. Dr. Ulrich Fischer
2. Prüfer: Dr. Antje Wichels

Tag des Promotionskolloquiums: 6. November 2015

"Unendlich groß ist die Rolle des unendlich Kleinen in der Natur"

Louis Pasteur um 1880

TABLE OF CONTENTS

INTRODUCTION

Microbial biogeography in marine systems

Although invisible to the naked eye microorganisms have been found to be ubiquitous in any conceivable niche on earth. From polar regions to temperate and tropical latitudes they inhabit terrestrial and aquatic environments, are associated with other organisms and also capable of coping with extreme environmental conditions. They do not only exist in incredible large numbers (4-6 x 10^{30} cells on earth) but also constitute an enormous pool of carbon, phosphorus and nitrogen (Whitman *et al.* 1998). In combination with their inexhaustible metabolic and physiological versatility this makes them essential to all biogeochemical cycling processes found on earth. Microorganisms provide what is called ecosystem services (breaking down complex substrates into small molecules) and have been associated with climate change (Bardgett *et al.* 2008) and even mass extinction events (Baune & Böttcher 2010) thus, greatly influencing and shaping our world.

Microorganisms in marine systems have been studied since more than 100 years. But it was only in the 1980s when the importance and dominance of prokaryotes in the oceans in terms of productivity, abundance and biomass has been recognized and proven in ever more complex concepts of marine microbial food webs (Pomeroy 1974; Azam *et al.* 1983; Sherr & Sherr 1988). The advent of molecular biological tools then allowed to assess the enormous taxonomic diversity, and characterize environmental microbial communities on a phylogenetic level for the first time (Giovannoni *et al.* 1990). Today, modern high-throughput DNA sequencing techniques allow for comprehensive, high resolution studies on large temporal and spatial scales and substantially enhanced research in the field of marine microbial biogeography. Biogeography describes the distribution patterns of communities in terms of richness, diversity - that is the number of different taxa in a community and their relative abundances – and community composition, across space and time (Martiny *et al.* 2006). However, biogeographical studies are not restricted to the description of diversity patterns in different habitats, but also aim to unveil the environmental parameters and processes that create and sustain these patterns, and to understand the functional roles of individual taxa within communities. Enhanced knowledge on diversity patterns and respective environmental driving forces of naturally occurring microbial assemblages will greatly contribute to ecosystem modeling and thus, help to predict and estimate potential changes in these patterns and its impact on ecosystem functioning.

The biodiversity of natural microbial assemblages has been studied intensively since it was recognized that they are vital to the function of all ecosystems on earth. For example, in a global study on marine bacterioplankton diversity and community composition using 16S rRNA clone libraries, Pommier *et al*. (2007) found many bacterial taxa to be endemic to one location, thus exhibiting pronounced biogeographic patterns. They additionally identified only few cosmopolitan taxa that were highly abundant at all locations. In another global study Fuhrman *et al.* (2008) demonstrated a latitudinal gradient in bacterial richness in oceanic surface waters decreasing from the tropics towards the poles. Thus, the existence of marine microbial biogeographic patterns is now well established (Hanson *et al.* 2012).

A huge variety of environmental factors strongly influences these biogeographic patterns. Along a salinity gradient, Fortunato & Crump (2011) found a shift in bacterial community composition leading to distinct populations in fresh-, brackish- and marine waters. Bacterial community variation at the San Pedro Times Series station was demonstrated to be driven by changes in temperature and nutrient concentrations (Chow *et al.* 2013). Likewise many other studies reported shifts in community composition in relation to varying oxygen concentrations (Ganesh *et al.* 2014), day length (Gilbert *et al.* 2012), organic pollutants (Störmer *et al.* 2013) or phytoplankton blooms (Teeling *et al.* 2012). Concerning the variety of oceanic environments that have been investigated in all these studies, it appears obvious that variation in marine bacterial communities is regulated by different sets of environmental parameters in the individual habitats.

Community variation on spatial and temporal scales

Marine habitats represent continuous, highly connected environments, where changes in bacterial communities are complex and not only triggered by different environmental factors but also comprise temporal (succession) and spatial (dispersion) components. Investigations on the temporal variability of bacterial communities comprise different time scales ranging from hours, days and weeks to seasonal and multiannual studies. Rink et al. (2008) observed relatively stable communities of particle-attached and free-living bacteria during a tidal cycle in the southern North Sea. Similar results were obtained by (Riemann & Middelboe 2002) over the range of days to months in Danish coastal environments. In contrast, Hewson *et al.* (2006) demonstrated shifts in richness and diversity in the course of a few days in the Gulf of Mexico, the North Pacific and the West Tropical Atlantic. Likewise, short-term succession as response to a phytoplankton bloom was demonstrated

in the range of days to weeks in a comprehensive metagenomic and -proteomic study (Teeling *et al.* 2012). Thus, it becomes clear that short-term variations do not follow universally valid rules but differ in individual oceanic habitats. Indeed, short-term variation of microbial communities is suggested to depend on the stability of the abiotic and biotic environment that can change rapidly due to changes in weather, or induced by biotic interactions. The extent to which these variations occur, probably reflects the different stability of different environments in terms of abiotic and biotic factors (Hedges 1992).

Seasonal variation among communities has been shown in long-term studies on the order of months to years in the Baltic Sea (Andersson *et al.* 2009), the English Channel (Gilbert *et al.* 2009), the German North Sea (Gerdts *et al.* 2004), the Mediterranean Sea (Ghiglione *et al.* 2007), the Sargasso Sea (Morris *et al.* 2005) and the Pacific Ocean (Cram *et al.* 2015). Multi-annual studies also found the bacterioplankton communities to reoccur in annual patterns that are predictable from environmental conditions (Fuhrman *et al.* 2006) and showed simultaneously that their structure and the dominance of specific species are maintained on a weekly to monthly scale (Brown 2005; Morris *et al.* 2005; Fuhrman *et al.* 2006).

Spatial variability in bacterial communities has been explored in horizontal direction on millimeter scales (Long & Azam 2001), along centimeters (Martiny *et al.* 2011), kilometers (Fortunato & Crump 2011) and global scales (Nemergut *et al.* 2011), as well as in vertical direction observing distinct depth profiles (Herlemann *et al.* 2011; Fortunato *et al.* 2013). There is evidence that microbial communities exhibit a huge variability on a scale of few kilometers, when strong environmental gradients are present. For instance, Riemann and Middleboe (2002) reported changes in community composition over a transect of approximately 35 km crossing frontal waters between the Baltic and the North Sea. In contrast, other studies revealed bacterial communities to be relatively similar in cohesive water masses of a few to 50 km in diameter assuming that this is the scale on which mixing occurs (Hewson *et al.* 2006).

Summed up, shifts in bacterial community composition have been demonstrated on different temporal and spatial scales in various oceanic environments. However, spatiotemporal analyses comprising both components simultaneously are still rare. The above explained variations on temporal and spatial scales are influenced by a variety of abiotic and biotic factors such as salinity, temperature, pH, nutrients, day length, grazing pressure, viral lysis, species interactions and many more. Additionally, stochastic

mechanisms like mutation, drift and dispersal are contributing to the generation of biogeographic patterns (Hanson *et al.* 2012). However, the extent to which the unique parameters and processes influence bacterial community variation still remains unclear.

Dissolved organic matter - bacteria interactions

Dissolved organic matter (DOM) provides a fundamental energy and nutrient source for marine heterotrophic bacteria. It is comprised of a huge variety of different compounds (e.g. amino acids, carbohydrates, lignin, black carbon etc.), among which dissolved organic carbon (DOC), dissolved organic nitrogen (DON) and dissolved organic phosphorus (DOP) are considered to be key compounds (Kirchman 2008). There are a number of sources that release DOM into the marine environment, but most DOM originates from direct excretion during phytoplankton growth, production by grazers during predation, release via viral and bacterial cell lysis, or solubilization of particles and bacterial transformation (Carlson 2002). A major part of the DOM is taken up, degraded and remineralized by heterotrophic bacteria, that channel energy and nutrients to higher trophic levels via the microbial loop (e.g. Pomeroy 1974; Azam *et al.* 1983). The relationship of heterotrophic bacteria and DOM is influenced by the complex composition and the various supply sources of DOM that make it "probably one of the major factors that help to maintain a high diversity of prokaryote communities in the oceans" (Nagata 2008).

Various studies have demonstrated relationships between specific DOM compounds and bacterial community structure. For instance, experimental setups showed changes in the bacterial community composition in response to enrichment with protein (Pinhassi *et al.* 1999), high molecular weight-DOM (McCarren *et al.* 2010), or dissolved free amino acids (Sarmento *et al.* 2013a). Seasonal shifts of bulk DOC concentration and bacterial activity have been demonstrated *in situ* over an annual cycle (Sintes *et al.* 2010). Others suggest, that different phylogenetic groups of bacteria tend to exploit different organic resources (e.g. Cottrell & Kirchman 2000; Elifantz *et al.* 2005). Influences of enhanced organic matter supply due to extracellular release by phytoplankton have been studied intensively and strong associations of certain bacterial groups with phytoplankton have been demonstrated (e.g. Sintes *et al.* 2010; Amin *et al.* 2012; Teeling *et al.* 2012). However, all these observations are restricted to selected compound classes of DOM. Only recently, the development of ultrahigh resolution Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) facilitated the characterization of DOM compounds on molecular level and thus, provides detailed insight into DOM compositions of different

sources. Few studies began to link the molecular DOM composition with changes in bacterial community structure (Osterholz *et al.* 2014; Medeiros *et al.* 2015; Seidel *et al.* 2015), but so far no studies related specific taxa with individual DOM molecules. Investigations on bacteria-DOM interactions on high taxonomic and analytic resolution will contribute substantially to understand, how bacterial communities are regulated by organic matter supply and in turn how bacteria shape the organic matter composition.

Study area: The North Sea, German Bight and Helgoland Roads

The greater North Sea stretching from 51 °N to 62 °N between the European mainland and the Atlantic Ocean (Fig. 1) is a typical semi-enclosed continental shelf sea with a surface area of approximately 750.000 km². It is entirely surrounded by seven industrial countries (Great Britain, France, Belgium, The Netherlands, Germany, Denmark and Norway), exposing it to huge anthropogenic influences such as industrial shipping, fisheries, tourism, use of energy resources and recreation (MUMM 2000). Topographically, it can be divided into the shallow southern North Sea (up to 54 °N) that includes the German Bight and exhibits depths on average < 40m, the northern North Sea with average depths between 40 m and 200 m and the Norwegian Trench with a maximum depth of 750 m (Howarth 2001). Since the North Sea is located in the temperate latitudes it is exposed to pronounced seasonal dynamics. The average temperature ranges from 6 °C in winter times to 17 °C during the summer months. Due to the large annual river input of 300 km³ and an additional freshwater input of 470 km³ per year via the Baltic Sea (MUMM 2000), salinity varies between 32-35 in northern offshore regions and 15-25 in estuaries (OSPAR 2000). Prevailing westerly winds on the north-west European shelf, tides and density gradients caused by freshwater input through rivers and the Baltic Sea, lead to an intense anticlockwise long-term water circulation along the coast (Sündermann & Pohlmann 2011) (Fig. 1). In some regions, mainly the northern part, stratification occurs during summer months due to heat

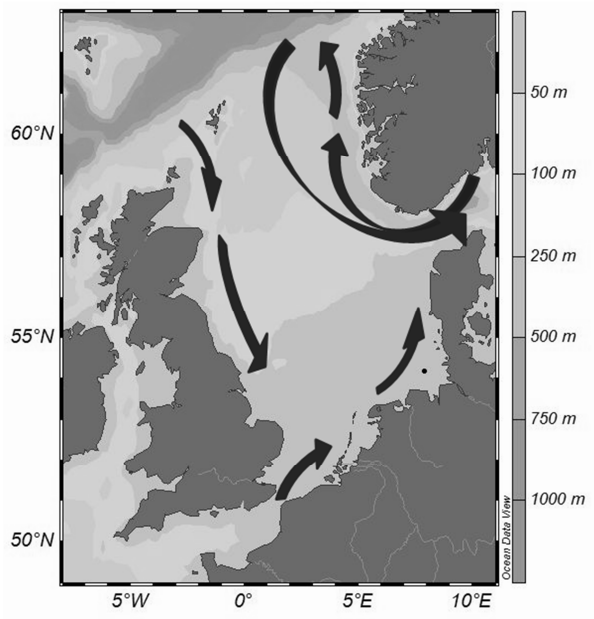


FIGURE 1: Location and topography of the North Sea. Arrows indicate the simplified main water circulation in the North Sea. Map was created with Ocean Data View (R. Schlitzer, http://odv.awi.de).

input. In contrast the water column in the southern North Sea and in particular in the German Bight tends to be well mixed throughout the year due to shallow depths and strong tides (Howarth 2001).

The relatively shallow German Bight (Fig. 2) is located in the south-eastern part of the North Sea, adjacent to the world's largest tidal flat area and enclosed by the Netherlands and Germany to the south and Denmark and Germany to the east. Water currents in the German Bight are predominantly influenced by tides, wind forces and the inflow of freshwater from rivers (Howarth 2001). The mixing of fresh and marine water typically leads to high spatial variability with respect to environmental parameters such as temperature, salinity, pH and organic loads (Atlas & Bartha 1987). High nutrient input from the rivers Elbe and Weser also substantially raises the productivity in the German Bight up to 430 g C m^{-2} a^{-1} (Rick *et al.* 2006; Heath & Beare 2008). Changes in the afore mentioned parameters, the biota (phyto-, zoo- and bacterioplankton) and current patterns have been continuously monitored for more than five decades around Helgoland Island in the German Bight (54°11.3' N, 7°54.0' E), known as the Helgoland Roads time series (Wiltshire *et al.* 2008) (Fig. 2). Based on this comprehensive long-term data set, Wiltshire *et al.* (2010) demonstrated a continuous increase in water surface temperature and salinity since 1962 for instance. The bacterioplankton community at Helgoland Roads is very well studied under various aspects using a wide range of different molecular methods such as the fingerprint methods RISA, DGGE and 16S rRNA gene tag sequencing. Numerous studies focused on temporal variation patterns and describe seasonal bacterial community dynamics on different time scales covering several months to multiple years, as well as the response to short-term events like phytoplankton blooms (e.g. Gerdts *et al.* 2004; Sapp *et al.* 2007; Teeling *et al.* 2012). Others investigated spatial patterns and compared bacterial communities at pelagic offshore and coastal inshore sites in the German Bight (e.g. Rink *et al.* 2011). However all these investigations, focused on either

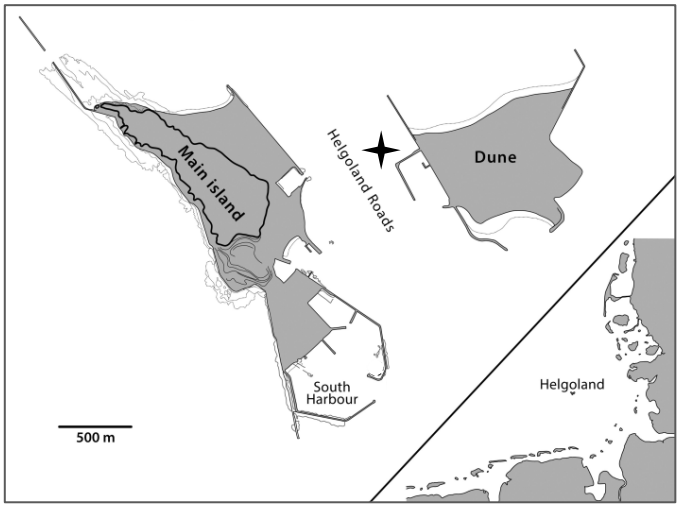


FIGURE 2: Map of Helgoland and its location in the German Bight (bottom right corner). Star indicates the sampling site of the Helgoland Roads time series (54°11.3' N, 7°54.0' E). Map modified after (Beermann 2014).

temporal or spatial patterns and were done with either limited temporal (Teeling *et al.* 2012) or taxonomic resolution (Gerdts *et al.* 2004) and thus, are lacking to uncover the complexity and diversity of the microbial community that has been described for other oceanic sites by high throughput sequencing techniques (e.g. Fuhrman *et al.* 2006; Gilbert *et al.* 2012).

Methodological approaches

The discovery of the small-subunit (16S) ribosomal gene as universal phylogenetic marker in the 1980s (Woese 1987) initialized the development of a broad range of new molecular biological tools that greatly increased the potential to investigate microbial community structure and taxonomic composition on different phylogenetic levels (e.g. kingdom, class, species).

Fingerprint methods like terminal restriction fragment length analysis (T-RFLP), amplified ribosomal DNA restriction analysis (ARDRA), denaturing gradient gel electrophoresis (DGGE) and automated ribosomal intergenic spacer analysis (ARISA) allow for the observation of the whole community diversity. Most of these fingerprinting methods rely on the comparison of the 16S rDNA. In contrast, ARISA uses the intergenic spacer (IGS) region implemented between the small (16S) and large (23S) subunits of the rRNA gene operon. This region demonstrates a pronounced species-specific length polymorphism, allowing for higher resolving community analysis than DGGE for instance (Okubo & Sugiyama 2009). ARISA was used in this thesis to explore the spatiotemporal diversity of the surface water bacterial community in the German Bight.

During the last decade new high-throughput sequencing technologies have been developed and replaced the conservative Sanger sequencing. Further improvements in read length and decreasing sequencing costs make “next generation sequencing” platforms like “Illumina MiSeq” accessible to comprehensive microbial ecological studies. Such high-throughput techniques enable the processing of a large number of samples in parallel. To accomplish this, individual barcode sequences (tags) are added to each sample. These tagged samples are then pooled and analyzed in a single sequencing run. Sequence data can be distinguished and sorted afterwards, based on the assigned tags. To date approximately 25 million paired end reads with 300 basepairs (bp) length can be accomplished during a single Illumina run. This massive amount of data is currently used to document the vast unexplored biodiversity in marine environments in more detail. During the course of this thesis, 16S rRNA tag sequencing was used to provide detailed fine-scale description of

daily changes in community composition and throughout an annual cycle and to further analyze the relationships of single OTUs with phytoplankton blooms and dissolved organic matter molecules.

Dissolved organic matter (DOM) comprises a fundamental energy source for marine microbes and it is hypothesized that the DOM composition influences the bacterial community composition and vice versa. Only recently, first insights into the highly complex composition of DOM, consisting of millions of different molecules - most of them in very low concentration (Dittmar & Paeng 2009) – have been gained. Coupling of electrospray ionization (ESI) with ultrahigh resolution Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) allows to determine molecular formulae and elemental composition of DOM (Stenson *et al.* 2003). FT-ICR-MS has successfully been used in combination with fingerprint techniques (DGGE) to analyze bacterial-DOM interactions (Osterholz *et al.* 2014; Osterholz *et al.* 2015). In the course of this thesis it is – for the first time- applied in conjunction with 16S rRNA tag sequencing of the bacterioplankton community to unravel DOM-bacterioplankton relationships on short time scales.

RESEARCH AIMS

The scope of this thesis was to comprehensively describe the variability of the bacterioplankton community in the German Bight at relevant temporal and spatial scales and to identify environmental parameters that most likely drive these variation patterns.
Microbial community variation comprises simultaneous changes on a broad range of temporal and spatial scales. However, most of the studies conducted on the bacterioplankton community of the German Bight only focused on individual aspects of this variation and were done with limited taxonomic resolution. Therefore, a thorough understanding of the bacterioplankton community assembly mechanisms in the German Bight is still lacking. The current thesis aimed at contributing to a profound picture of the bacterioplankton community variation by elaborating on the following topics, using high resolution methods at both, temporal and spatial dimensions:

I. Long-term variability of the bacterioplankton community

Previous studies in the German Bight and particularly at Helgoland Roads (54°11.3' N, 7°54.0' E) demonstrated seasonality of bacterial communities (e.g. Gerdts *et al.* 2004; Stevens *et al.* 2005; Sapp *et al.* 2007; Boer *et al.* 2009). Based on this evidenced seasonality, we focused on unravelling which bacterial taxa were dominant during the different seasons and whether the bacterioplankton community composition (BCC) is changing constantly throughout the year, or if stable communities displace each other due to abrupt environmental changes.

II. Short-term variability of the bacterioplankton community

The taxonomic and functional succession of bacterioplankton communities during phytoplankton blooms (i.e. enhanced organic matter supply) has been studied intensively at Helgoland Roads (Teeling et al. 2012). Chapter II of this thesis aimed at gaining more detailed insights into the response of the bacterial community to dissolved organic matter (DOM) supply. A central goal was to test the hypothesis that the BCC co-varies with the molecular DOM composition. Furthermore, we aimed at investigating, if specific bacterial taxa were favoured by particular DOM compounds.

III. Spatiotemporal variation of the bacterioplankton community

Although seasonal dynamics of the bacterioplankton community and differences between coastal and offshore communities have been studied in the German Bight (Sapp et al. 2007; Rink et al. 2011), knowledge on spatiotemporal variation concerning gradients from coastal to pelagic offshore sites does not exist. Hence, this thesis aimed at describing the spatiotemporal variation of the free-living and particle-attached bacterioplankton community along a transect reaching from the Elbe estuary towards the open North Sea over a period of one year, and to investigate the relative importance of temporal and spatial variation. A second goal was to reveal environmental factors that potentially drive the spatiotemporal variation

OUTLINE

The present thesis consists of a general introduction, three chapters representing one manuscript each and a general discussion.

Manuscript I (published in *FEMS* Microbiology Ecology)

Lucas J, Wichels A, Teeling H, Chafee M, Scharfe M and Gerdts G (2015) **Annual dynamics of North Sea bacterioplankton: Seasonal variability superimposes short-term variation.** *FEMS Microbiology Ecology* DOI:10.1093/femsec/fiv099

This manuscript describes the seasonal variation of the bacterioplankton community in relation to relevant environmental parameters. The main outcome is that short-term bacterioplankton successions in response to phytoplankton blooms are indirectly affected by temperature, which is a major niche-defining factor in the German Bight. Furthermore, results suggest an annual recurrence and resilience of few main taxa. The laboratory investigations were carried out by Judith Lucas. Calculation of hydrographic currents was carried out by Mirco Scharfe. 16S rRNA gene tag sequencing was performed at the U.S. Department of Energy Joint Genome Institute (JGI, Walnut Creek, CA, USA) and sequence processing was done by Meghan Chafee under the guidance of Hanno Teeling. Environmental data were kindly provided by Karen H. Wiltshire. Statistics, evaluation and manuscript writing was carried out by Judith Lucas under the guidance of Antje Wichels and Gunnar Gerdts.

Manuscript II (submitted to Frontiers in Microbiology)

Lucas J, Koester I, Wichels A, Niggemann J, Dittmar T, Callies U, Wiltshire KH and Gerdts G **Short-term dynamics of North Sea bacterioplankton-dissolved organic matter interactions on molecular level.**

In this manuscript short-term bacterioplankton-DOM interaction was demonstrated in strong correlations between specific bacterial taxa and particular DOM molecules thus, suggesting potential specialization of bacteria on particular DOM molecules. Sampling and laboratory investigations were accomplished by Irina Köster, Judith Lucas and Jutta Niggemann. Environmental data were provided by Karen H. Wiltshire. Hydrographic backmodelling was calculated by Ulrich Callies. 16S rRNA gene tag sequencing was done at LGC Genomics GmbH (Berlin, Germany). Analysis of DOM compositions was performed by Irina Köster under the guidance of Jutta Niggemann and Thorsten Dittmar.

Analysis of sequencing data was done by Judith Lucas. Planning, statistical analysis, evaluation and writing were carried out by Judith Lucas under the guidance of Antje Wichels, and Gunnar Gerdts.

Manuscript III (submitted to Microbial Ecology)

Lucas J, Wichels A and Gerdts G **Spatiotemporal variation of the bacterioplankton community in the German Bight: From estuarine to offshore regions**

This manuscript demonstrates the deconvolution of temporal and spatial bacterioplankton variability in the German Bight. It is shown that the spatial variation of the bacterial community is defined by pronounced hydrographic current conditions, separating coastal and oceanic populations. This spatial variation is overwhelmed by a strong temporal variation that is triggered by temperature as a main driving force throughout the German Bight. Sampling was carried out by Judith Lucas. The planning, laboratory investigations, statistical analysis and manuscript writing were accomplished by Judith Lucas with the assistance of Antje Wichels and Gunnar Gerdts.

CHAPTER I

Annual dynamics of North Sea bacterioplankton: Seasonal variability superimposes short-term variation

Judith Lucas[1*], Antje Wichels[1], Hanno Teeling[2], Meghan Chafee[2], Mirco Scharfe[1] and Gunnar Gerdts[1]

[1]Alfred-Wegener-Institute Helmholtz-Center for Polar and Marine Research, Biological Station Helgoland, Kurpromenade 201, 27498 Helgoland, Germany

[2]Max Planck Institute for Marine Microbiology, Celsiusstraße 1, 28359 Bremen, Germany

[*]Judith Lucas, Alfred-Wegener-Institute Helmholtz-Center for Polar and Marine Research, Biological Station Helgoland, Kurpromenade 201, 27498 Helgoland, Germany; Phone: +49 4725 819 3200, Fax: +49 4725 819 3283, Email: judith.lucas@awi.de

Abstract

The dynamics of coastal marine microbial communities are driven by seasonally changing abiotic and biotic factors as well as by rapidly occurring short-term changes such as river fresh-water influxes or phytoplankton blooms. We examined the variability of the free-living bacterioplankton at Helgoland Roads (German Bight, North Sea) over a period of one year with high temporal and taxonomic resolution to reveal variation patterns and main influencing factors. 16S rRNA gene tag sequencing of the bacterioplankton community hints at annual recurrence and resilience of few main taxa belonging to *Alphaproteobacteria, Betaproteobacteria, Flavobacteriia, Acidimicrobiia* and *Thermoplasmata*. Multiple regression analyses with various environmental factors revealed changes in water current patterns and resulting phytoplankton blooms as the main driving factors for short-term variation and temperature as the overlying factor for seasonal variation. Comparison of bacterioplankton successions during spring and summer phytoplankton blooms revealed the same dominating *Flavobacteriia* operational taxonomic units (OTUs) but shifts in *Roseobacter* related OTUs (*Alphaproteobacteria*) and SAR92 clade members (*Gammaproteobacteria*). Network analysis suggests that during spring and summer phytoplankton blooms temperature-dependent guilds are formed. In conclusion, our data imply that short-term bacterioplankton successions in response to phytoplankton blooms are indirectly affected by temperature, which is a major niche-defining factor in the German Bight.

Introduction

Studies that examined the bacterioplankton response to changing environmental factors in diverse marine environments identified a range of oceanographic, physico-chemical and biotic factors that influence variations in bacterioplankton community composition (BCC) (Fuhrman *et al.* 2006; Gilbert *et al.* 2009; Fortunato *et al.* 2012). In particular dissolved and particulate organic matter released by phytoplankton strongly shapes the BCC, however, it has been suggested that in highly dynamic systems such as estuaries of continental shelf seas, the influence of primary producers on microbial dynamics is less important compared with that of abiotic factors (Kirchman *et al.* 2005; Teira *et al.* 2008). Concerning the variability of environmental conditions in different oceanic regions one would expect individual combinations of environmental factors that are driving changes in the bacterial community composition at different sites.

The North Sea is a semi-enclosed continental shelf sea. Especially its southeastern region, the German Bight, is highly influenced by the runoff from the rivers Elbe and Weser and thereby constantly supplied with nutrients, making it a very productive area. The mixing of fresh and marine water typically leads to high spatial variability with respect to environmental parameters such as temperature, salinity, pH and organic loads (Atlas & Bartha 1987). Changes in these parameters, the biota and current patterns have been continuously monitored for more than five decades around Helgoland Island in the German Bight (54°11.3' N, 7°54.0' E), known as the Helgoland Roads time series (Wiltshire *et al.* 2008). This comprehensive long-term data set makes Helgoland Roads an optimal study site to investigate how environmental parameters shape bacterioplankton communities in coastal oceanic environments. The BCC in the German Bight has been well described, particularly during spring phytoplankton blooms using different molecular biological approaches like DGGE, RISA (Sapp *et al.* 2007), CARD-FISH (Alderkamp *et al.* 2006) and 16S rRNA gene tag sequencing (Teeling *et al.* 2012; Wemheuer *et al.* 2014). Previous studies at Helgoland Roads also demonstrated seasonality of bacterioplankton communities driven by different environmental factors or phytoplankton abundances. However, these investigations were done with either limited temporal (Gilbert *et al.* 2009) or taxonomic resolution (Gerdts *et al.* 2004) and thus, are lacking to uncover the complexity and diversity of the microbial community that has been described for other oceanic sites by high throughput sequencing techniques (Fuhrman *et al.* 2006; Gilbert *et al.* 2012).

In this study we examined the bacterioplankton community at Helgoland Roads at both, high temporal and taxonomic resolution. To unravel whether the BCC is changing

constantly throughout the year or if stable communities displace each other due to abrupt environmental changes, we assessed the BCC at Helgoland Roads on a weekly basis over a period of one year. To further elucidate these changes in community structure and define the succession of distinct dominating key taxa in different seasons and phytoplankton blooms, we used 16S rRNA gene tag sequencing of the free-living bacterioplankton fraction (0.2-3 µm). Multivariate statistics and network analyses were applied to determine which environmental parameters exert the strongest influences on the bacterial community and thus, shape the ecological niches that the defined key players occupy. The combination of high temporal and taxonomic resolution methods allowed a detailed understanding of possible controls of the BCC at Helgoland Roads and can serve as basis for future functional approaches.

Materials & Methods

Sample collection and environmental parameter measurements

A total of 42 surface seawater samples (1m depth) were collected weekly from 1 March 2012 to 28 February 2013 at Helgoland Roads (North Sea, Germany, 54°11.3'N, 7°54.0'E). The sampling site is located approximately 60 km off the German coastline. Total water depth varies between 7 and 10 m depending on the tides. Environmental data including dissolved organic carbon (DOC), dissolved inorganic nitrogen (DIN = NO_2^- + NO_3^- + NH_4^+), silicate (SiO_2), phosphate (PO_4^{3-}), salinity (S), water temperature (T), chlorophyll *a* (Chl *a*) and counts of phytoplankton groups (diatoms, dinoflagellates, flagellates, ciliates) were obtained in parallel as part of the Helgoland Roads time series (Wiltshire *et al.* 2008). Flagellate cell counts included also counts for heterotrophic nanoflagellates.

Hydrodynamic variability in the German Bight was assessed using current velocity fields from the model BSHcmod (Dick *et al.* 2001) operated by the Bundesamt für Seeschiffahrt und Hydrographie (BSH). First, current velocities of high temporal resolution (15 minutes) were averaged to obtain weekly mean vector (i.e. u, v) fields for the period March 2012 to February 2013. Second, Empirical Orthogonal Function (EOF) analysis (von Storch & Zwiers 1999) was applied to identify dominant modes of spatially coherent variability in these current patterns. These EOFs reflect anomaly patterns with regard to the mean current conditions for the selected period with the first EOF covering the highest amount of explained variance in the simulated transport fields. The explained variance of the two leading EOFs is more than 85% (Fig. S2). The corresponding principal components PC1 and PC2 (Fig. S3) provide information about the sign and the amplitude of the EOFs as a function of time. The vector fields shown by the EOFs represent weighting factors (loadings) which are used for mapping each weekly mean current field to one data point of the corresponding principal component time series. See Callies and Scharfe (2014) for a comparable analysis based on the decadal scale.

Sample preparation and DNA extraction

500 ml of each sample were subjected to fractionating filtration using 10, 3 and 0.2 µm pore size polycarbonate membrane filters (Millipore, Schwalbach, Germany), to separate particle-attached bacteria (3-10 µm) from free-living bacteria (0.2-3 µm). For cell counts, 4 ml of each filtrate obtained by filtration through 3 µm pore size filters were fixed with formaldehyde (1% [w/v] final concentration) for 1 h at room temperature and subsequently

stored at -80 °C until further processing. Cell counts were determined as described in Krause *et al.* (2012) using an Accuri C6 flow cytometer (BD Accuri Cytometers, Ann Arbor, MI, USA). The threshold on FL1-H was set to 700. DNA of free-living bacteria was extracted from filters as described previously (Sapp *et al.* 2007). Briefly, cells were lysed using lysozyme/SDS, DNA was obtained by phenol-chloroform extraction and subsequent isopropanol precipitation. DNA concentration per sample was quantified using the Invitrogen (Carlsbad, CA, USA) Quant-iT™ PicoGreen ® dsDNA Reagent as per manufacturer's instructions.

16S rRNA V4 amplicon sequencing

16S rRNA gene tag sequencing was performed at the U.S. Department of Energy Joint Genome Institute (JGI, Walnut Creek, CA, USA). Community DNA samples were sent to JGI in a 96-well plate for generation of 16S V4 rRNA amplicon libraries for Illumina sequencing. Sample preparation was performed on a PerkinElmer (Waltham, MA, USA) Sciclone NGS G3 Liquid Handling Workstation capable of processing 96 plate-based samples in parallel, utilizing 5 PRIME (Gaithersburg, MD 20878, USA) HotMasterMix amplification kit and custom amplification primers targeting the V4 region of the 16S rRNA gene using 515F (5' GTGCCAGCMGCCGCGGTAA 3') and 806R (5' GGACTACHVGGGTWTCTAAT 3') (Caporaso *et al.* 2011). Primers also contained the Illumina adapter sequence and a unique barcode index. PCR reactions were set up in 75 µl total with 1x HotMasterMix (5 PRIME) with final concentrations of 0.4 µg μl^{-1} BSA and 0.2 µM of each primer. This total volume was split into triplicate 25 µl reactions for independent amplification and then pooled to reduce PCR bias. Prepared amplicon libraries were normalized and multiplexed into a single pool of amplicons per plate. 16S V4 rRNA amplicon library pools were quantified using the KAPA Biosystems (Wilmington, MA, USA) next-generation sequencing library qPCR kit and run on a Roche (San Francisco, CA, USA) LightCycler 480 real-time PCR instrument. The quantified pool was loaded on an Illumina (San Diego, CA, USA) MiSeq sequencer using 2x250 bp chemistry. The 16S rRNA gene tag sequences are available from the DOE-JGI website GOLD database (Project ID: Gp0056779) as part of the community sequencing project COGITO (Coastal Genomic & Taxonomic Observatory).

Raw paired-end reads were merged and filtered using scripts from illumina-utils (https://github.com/meren/illumina-utils) to retain only those sequences without mismatches in the overlapping region. These high-quality tags were processed through the

SILVAngs pipeline (Quast *et al.* 2013). Sequences were dereplicated at 100% identity and then clustered within each individual sample at 98% similarity to reduce computational demands for classification. Representative sequences from operational taxonomic unit clusters (OTUs) were classified up to genus level against the SILVA v115 database using BLAST as described by Ionescu *et al.* (2012). Genus-level classifications were used in the final abundance matrix for downstream analyses. Each classification contained the sum of all sequences represented by OTUs with the same taxonomic path. For the purposes of this study we were not interested in diversity calculated at the level of 98% clustered OTUs but rather used BLAST identities as our operational taxonomic unit. From this point on, we define these taxa as OTUs for simplicity. Therefore, in our study, OTU refers to a unique taxonomy and not a cluster of sequences defined by percent similarity. Eukaryotic, chloroplast and mitochondria-derived OTUs were removed from the resulting OTU matrix. To account for variation in total bacterial abundance over the year (Tab. S1) OTU abundances were weighted by multiplying the relative OTU abundances with total bacterial cell counts according to Andersson *et al.* (2009). We were interested in analyzing which environmental parameters drive the ecologically most important microbial taxa. Since the most abundant microbial taxa are also thought to be the most active ones, contributing the most to biomass production and are most important in fluxes of dissolved organic matter (Cottrell & David 2003; Zhang *et al.* 2006) we decided to omit the 'rare biosphere' and focus on OTUs with an annual average relative abundance $\geq 0.1\%$. This 'trimmed data set' was used for further analyses.

Statistical analyses

To reveal patterns in bacterial community composition, principal coordinates analysis (PCoA) of all samples was carried out using Hellinger distance (D17; Legendre & Legendre 1998), which uses square root transformed relative abundances of sequence read numbers for distance matrix calculation. Analyses were carried out with the Primer v.6 software package (PRIMER-E, UK).

The relationship between environmental parameters and bacterial community structure was statistically analysed in SigmaPlot (Systat, Version 11). Multiple stepwise forward regressions were calculated using above mentioned PCoA scores of the first two PCoA axes as dependent variables and all measured environmental parameters as independent variables. Since temperature and DIN were highly correlated (R=-0.803) this suggests a large shared contribution to the model. To disentangle unique and shared contributions, we

individually regressed DIN against temperature and replaced original DIN values with the residuals of that regression. Multiple regression analysis was then accomplished with all parameters and replaced DIN values. Only variables that significantly ($p<0.05$) added to the prediction of the dependent variables were used to build the multiple regression model. Residual analyses of the regression models were carried out to investigate the difference between observed and predicted scores in detail.

Correlations between all environmental parameters were determined using Spearman rank order correlations applying a significance level of $p<0.05$. Additionally, correlations between relative abundances of all OTUs and scores of the first two PCoA axes were calculated. To visualize the relationship between OTUs and PCoA axes, OTUs that were statistically significantly correlated ($p<0.05$) with one, or both PCoA axes were used to perform interaction network analysis using Cytoscape version 3.2.0 (Shannon *et al.* 2003).

Results

Environmental conditions at sampling site

Concurrent with water sampling, physicochemical parameters were recorded and current components were calculated (Fig. S1, S2, S3 and Tab. S1). Spearman rank order correlation analysis revealed statistically significant correlations (p<0.05) between abiotic parameters (Tab. S2) but only few had particularly high correlation coefficients (R>0.6) such as DIN and temperature (R=-0.803), DOC and salinity (R=0.633) and DIN and SiO_2 (R=0.634) (Tab. S2). Two Chl *a* peaks were measured in April and August and were referred to as spring and summer phytoplankton blooms, respectively (Fig. 1). The spring bloom was dominated by a combination of dinoflagellates and flagellates, whereas the summer bloom seemed to be more diverse and was characterised by high diatom, ciliate, flagellate and dino-flagellate cell numbers (Fig. 1).

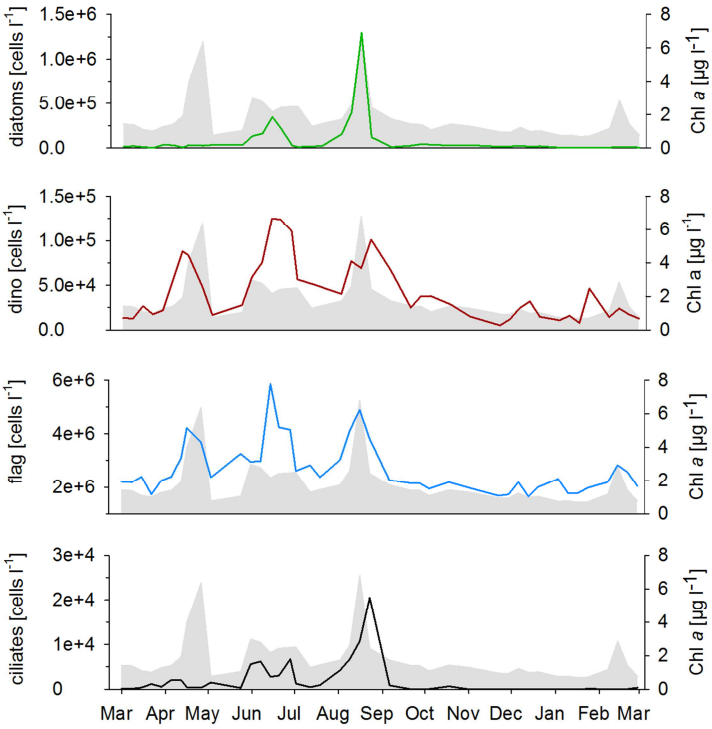


FIGURE 1: Phytoplankton abundances recorded from 01. March 2012 – 28. February 2013 at Helgoland Roads. Cell numbers per litre are depicted for diatoms, dinoflagellates, flagellates and ciliates. Chlorophyll *a* concentrations ranging from 0.74-6.8 µg l^{-1} are depicted as grey area.

Bacterial community structure in relation to environmental parameters

The bacterioplankton community at Helgoland Roads was assessed by 16S rRNA gene tag sequencing at weekly time intervals over a period of one year. After exclusion of eukaryote and organelle sequences and OTUs with an average relative abundance ≤0.1% of the total community, 4,739,551 high quality sequences were retained that represented 116 different OTUs. Principal coordinates analysis (PCoA) of the bacterial community revealed a strong seasonal pattern (Fig. 2). A spring cluster was followed by rapidly changing bacterial community compositions with high week to week variance (rapid change phase), which

passed into a relatively stable summer community. At the end of August a second rapid change phase occurred, which led to a stable autumn community. During winter, the community structure changed slowly and returned to the previous year's spring community. Rapid changes of the community structure followed the first principal coordinate axis (PCoA1), whereas the more gradual changes within the summer cluster and from the autumn to the spring community occurred along the second PCoA axis (PCoA2). Multiple regression analyses (MRA) with all environmental parameters as independent and scores of each PCoA axis as dependent variables revealed that PCoA1 was best explained by a combination of salinity, SiO_2, flagellates and PC2 (Fig. 3A and Tab. S3), whereas the regression model for PCoA2 was significantly influenced by temperature, SiO_2, DIN, flagellates and PC1 (Fig. 3B and Tab. S3). The fits of both models were statistically significant (p<0.001). Residuals depicted for the models (Fig. 3A and B) displayed the largest differences between calculated and predicted PCoA scores from June until August.

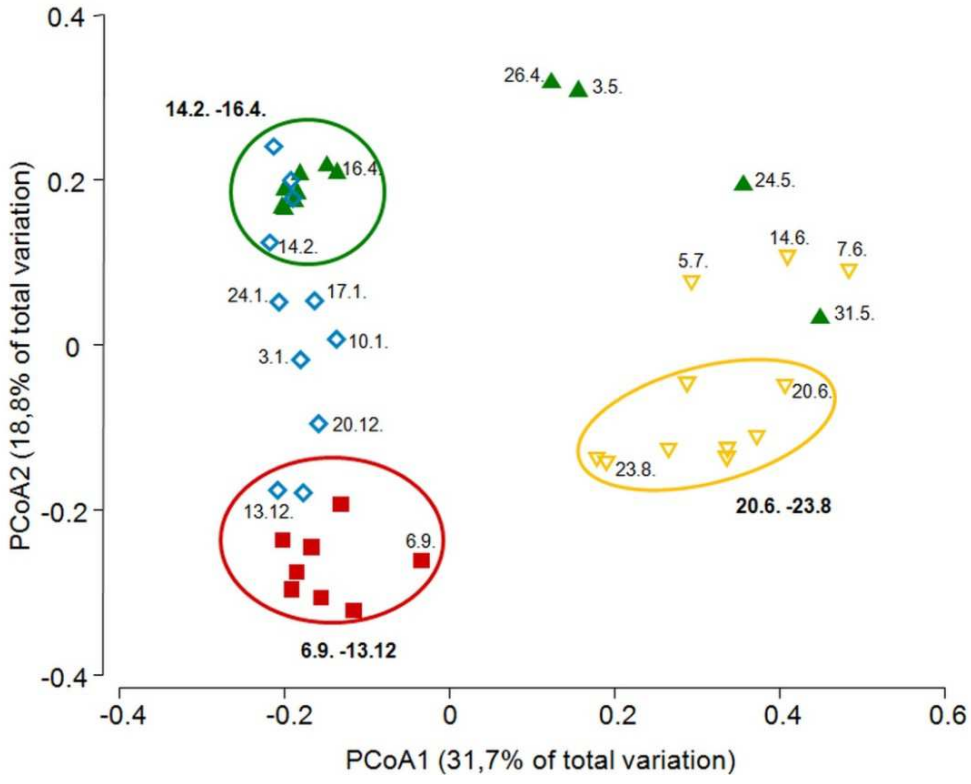
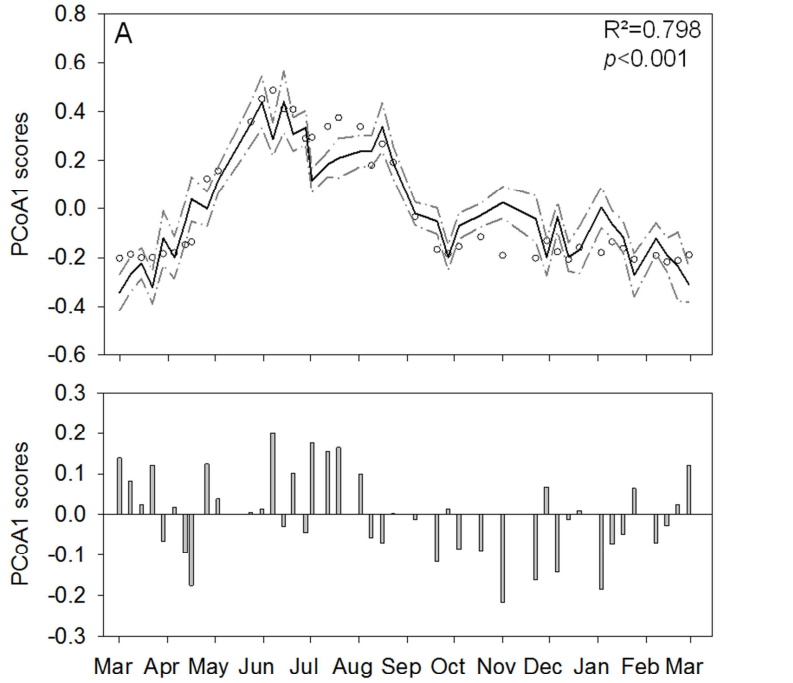
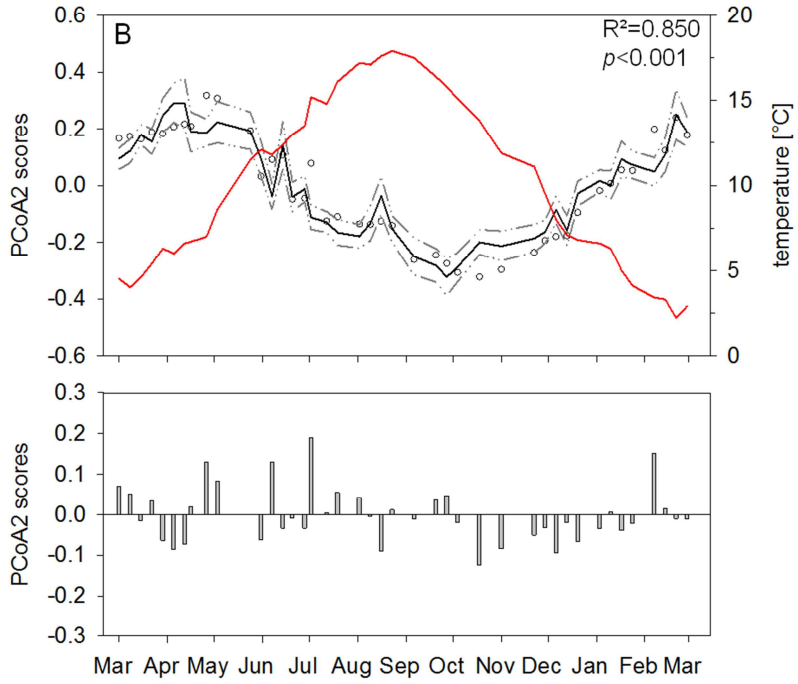


FIGURE 2: Principal coordinates analysis of bacterial communities using Hellinger distance. Symbols represent bacterial communities at sampling dates. Symbols are colour coded according to season. Seasons are defined according to meteorological definition. Green: spring (1 March-31 May); yellow: summer (1 June-31 August); red: autumn (1 September-30 November); blue: winter (1 December-29 February). Spring, summer and autumn clusters are indicated by ellipses on a distance level of 0.42. Numbers represent sampling dates and time frames that are covered by the clusters respectively. Within each cluster the first and last sampling dates are given.



PCoA1=4.38-(0.13xS)-(0.03x SiO_2)+(7.6x10^{-8}xflag)-(0.04xPC2)



PCoA2=0.43-(0.04xT)-(0.02x SiO_2)+(0.02xDIN)+(4.2x10^{-8}xflag)-(0.04xPC1)

FIGURE 3: Stepwise forward multiple linear regression analyses. A: prediction for PCoA1. B: prediction for PCoA2. Open circles: calculated PCoA scores; black line: predicted PCoA scores; dashed lines: 95% confidence interval; red line in B: temperature. R^2, p values and formulae are displayed for the models. Only parameters with $p<0.05$ were considered for model building. Corresponding residuals for models are displayed as bar charts below the models. S: salinity; T: temperature; SiO_2: silicate ; DIN: dissolved inorganic nitrogen (DIN = NO_2^- + NO_3^- + NH_4^+); Chl *a*: chlorophyll *a*; PC1, PC2: Principal components of EOFs; flag: flagellates.

Community composition and succession during phytoplankton blooms

The bacterial community at Helgoland Roads (Fig. S4) was dominated by *Proteobacteria* (mainly *Alpha-, and Gammaproteobacteria*) with an annual mean of 60.1% of the trimmed tag data. *Bacteroidetes* were represented almost exclusively by *Flavobacteriia* and accounted for 24.7% of the trimmed tag data. Other phyla that were present throughout the year and reached relatively high abundances were *Actinobacteria* (5.3%) and *Euryarchaeota* (5.2%).

A seasonal succession of six main OTUs that dominated the community (i.e. had highest relative abundance of all OTUs in a particular sample) at defined periods in time, i.e. at least two weeks in a row, was observed (Fig. 4). On average these six OTUs represented 27.4% of the trimmed tag data. In addition to the seasonal succession pattern, periods of rapid shifts of these dominating groups were observed during the spring and summer blooms (Fig. 4). Spring bloom dominating OTUs belonged to the *Alphaproteobacteria, Gammaproteobacteria* and *Flavobacteriia*. In comparison, higher diversity was found during the summer bloom with dominating OTUs belonging to the *Alpha-* and *Gammaproteobacteria*, *Flavobacteriia*, *Acidimicrobiia* (OCS155 marine group) and *Thermoplasmata* (Marine Group II; *Euryarchaeota*).

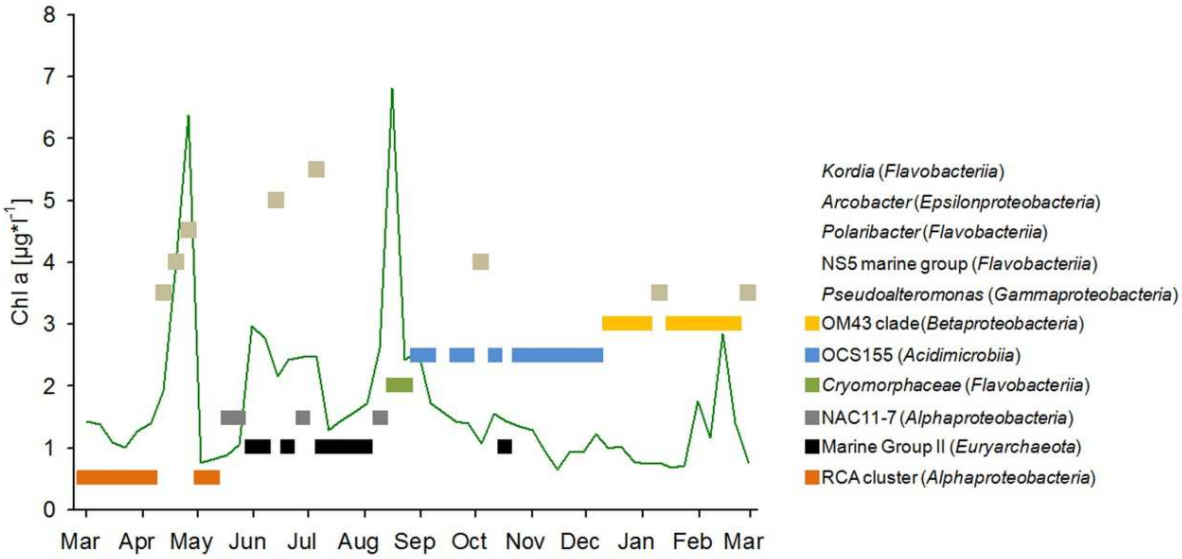


FIGURE 4: Annual succession of dominating OTUs. OTUs that exhibited highest relative abundance of all OTUs in a particular sample were considered to dominate the community at that time point. OTUs that dominated defined periods (at least two weeks in a row) are colour-coded. OTUs that dominated the community in single samples only were kept in beige. Green line: Chlorophyll *a* concentration.

We observed swift successions of distinct OTUs during the spring and summer bloom phases and examined the dominant OTUs within the prominent classes of *Alpha-, Gammaproteobacteria* and *Flavobacteriia* (Fig. 5). During the spring bloom, *Alphaproteobacteria* (*Roseobacter* related DC5-80-3 lineage, referred to as *Roseobacter* clade affiliated (RCA) cluster and NAC11-7 lineage) increased in relative abundance upon bloom decay (Fig. 5A). *Gammaproteobacteria* showed a succession with *Pseudoalteromonas* peaking twice in abundance, once early in the bloom phase and again late after the bloom phase when the Chl *a* concentration began to increase again. During the Chl *a* maximum the SAR86 clade dominated the *Gammaproteobacteri*a, whereas the SAR92 clade responded more to bloom decay (Fig. 5B). *Flavobacteriia* were dominated by the NS5 marine group in the early bloom phase. *Polaribacter* increased simultaneously with increasing Chl *a* concentration and peaked during the Chl *a* maximum. In contrast a *Cryomorphaceae* cluster exhibited higher relative abundances after the bloom decay (Fig. 5C).

During the summer bloom we observed a different succession pattern within the prominent classes (Fig. 5D-F). Succession of *Alphaproteobacteria* was led by *Roseobacter* NAC11-7 members during the early bloom and the Chl *a* maximum. Upon algal decay the relative abundance of SAR116 clade increased slightly, whereas the OCT lineage decreased in relative abundance (Fig. 5D). Within the *Gammaproteobacteria,* OM60/NOR5 clade members responded to the early bloom. As the bloom commenced, members of the ZD0405 clade increased in relative abundance and SAR86 clade members dominated during the late bloom phase (Fig. 5E). Succession of *Flavobacteriia* was again governed by *Cryomorphaceae* until the bloom began to collapse and NS5 marine group members notably increased in relative abundance during bloom decay.

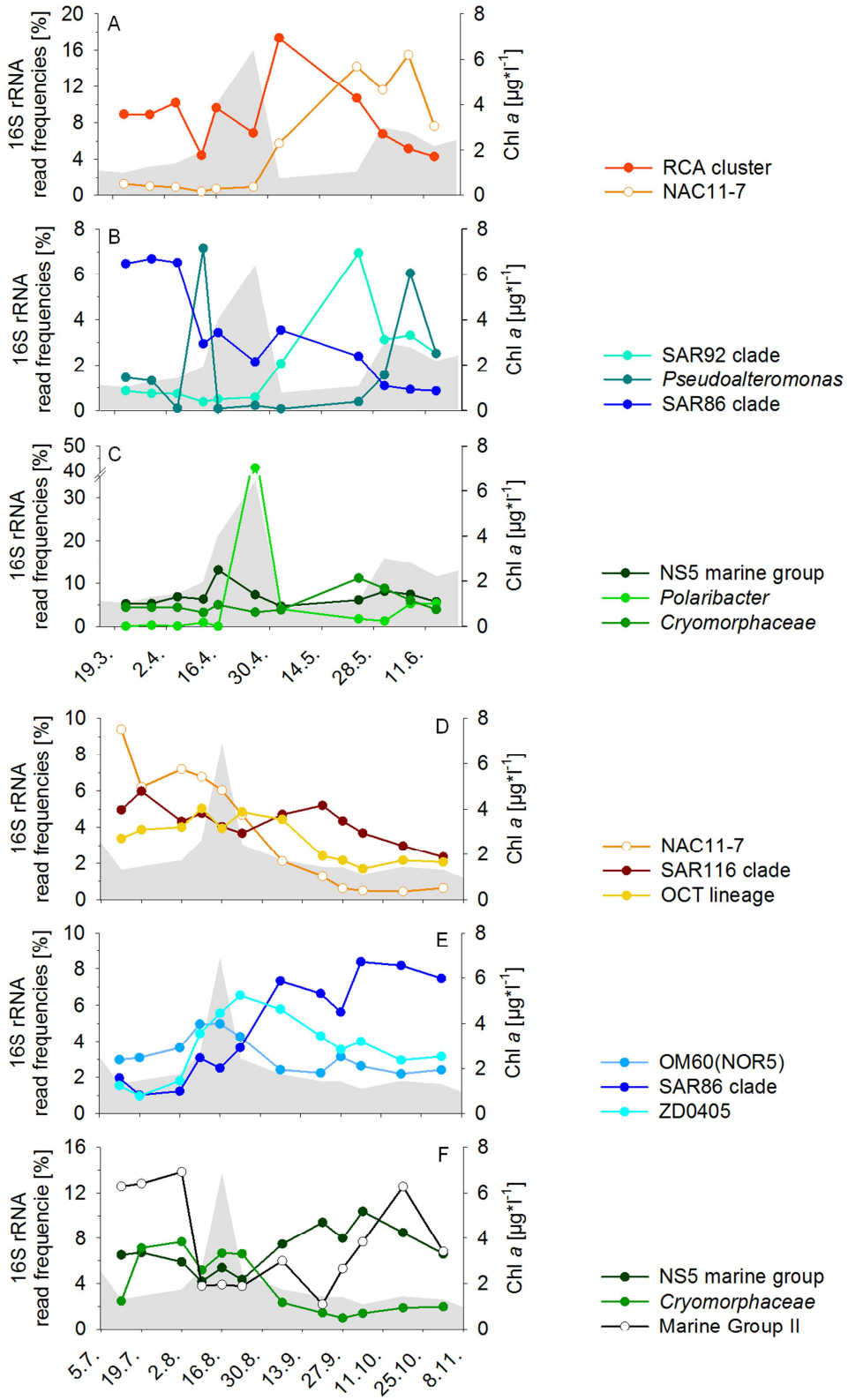


FIGURE 5: Short-term succession of dominant OTUs within the prominent classes of *Alphaproteobacteria*, *Gammaproteobacteria* and *Flavobacteriia* during the spring bloom (A-C) and summer bloom (D-F). A: *Alphaproteobacteria*; B: *Gammaproteobacteria*; C: *Flavobacteriia*; D: *Alphaproteobacteria*, E*: Gammaproteobacteria*, F*: Flavobacteriia* and *Thermoplasmata* (Marine Group II). Chlorophyll *a* concentration is indicated as grey area.

Clustering of OTUs according to response to environmental parameters

To elucidate how single OTUs were influenced by environmental conditions, we selected OTUs that were significantly (p<0.05) correlated with one or both PCoA axes, assuming that these contributed significantly to the explained variation in community structure. These OTUs were subjected to network analysis and sorted according to their response to phytoplankton blooms (Fig. 6). We observed distinct correlation patterns for the different groups. Most interestingly, OTUs that responded to the spring bloom were positively correlated with PCoA2 (i.e. temperature, see Fig. 3B), whereas OTUs responding to the summer bloom were negatively correlated with PCoA2. In both groups strong positive correlations with PCoA1 (i.e. phytoplankton, see Fig. 3A) were observed. OTUs without a response to any of the bloom phases were in general negatively correlated with PCoA2.

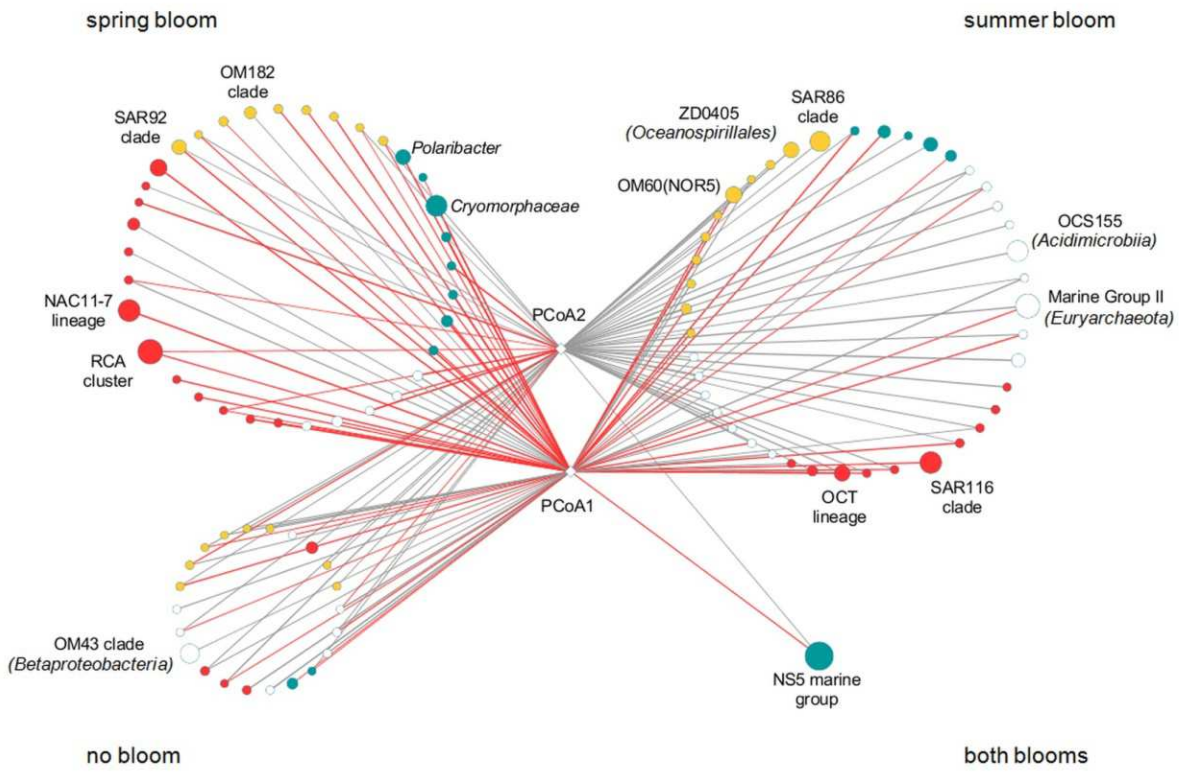


FIGURE 6: Interaction network analysis of OTUs that were significantly correlated (p<0.05) with PCoA axes. Positive correlations are indicated in red, negative correlations in grey. Line width is set proportional to correlation strength. Mean annual OTU relative abundance is set proportional to node size. OTUs belonging to *Alphaproteobacteria, Gammaproteobacteria* or *Flavobacteriia* are color-coded, remaining OTUs are kept in white. Red nodes: *Alphaproteobacteria*; yellow nodes: *Gammaproteobacteria*, blue nodes: *Flavobacteriia*. OTUs were split into groups: i) spring bloom: OTUs with average abundance during spring bloom > annual average abundance, ii) summer bloom: OTUs with average abundance during summer bloom > annual average abundance, iii) both blooms: OTUs with average abundance during spring and summer bloom > annual average abundance and iv) no bloom: OTUs with average abundance during blooms ≤ average annual abundance.

Discussion

Seasonal variation in North Sea bacterioplankton

We observed a pronounced seasonal pattern, which agrees with other studies on the seasonality of bacterioplankton communities (Fuhrman et al. 2006; Fortunato et al. 2012; Gilbert et al. 2012). The seasonal variation in this study is largely driven by temperature as revealed by multiple regression analysis and reflected in the shape of our model which is inversely proportional to the measured temperature curve. This temperature dependency is also supported by other authors who describe temperature as a main determining driver of community composition (Pommier et al. 2007; Gilbert et al. 2009; Chow et al. 2013).

The seasonal variation in community structure is governed by a few OTUs from different taxonomic classes, several of which have been identified as dominant community members in previous studies. Within the *Alphaproteobacteria*, especially the *Roseobacter* RCA cluster and NAC11-7 lineage were highly abundant with up to 18% and 15% of the trimmed tag data, respectively. This high abundances are in line with reports by Selje *et al.* (2004), Giebel *et al.* (2011) and Teeling *et al.* (2012) who reported similar high abundances of these taxa in the North Sea during several years. Using 16S rRNA gene sequencing, Wemheuer *et al.* (2014) determined the *Betaproteobacteria* OM43 clade as a major bacterioplankton OTU in the North Sea. Our study also identified OM43 as a prominent clade, particularly during winter when up to ~13% of the trimmed tag data affiliated with this clade. The *Actinobacteria* related OCS155 marine group is one of the five most abundant and persistent OTUs identified over 10 years of a long-term study in the Southern California Bight (Chow et al. 2013). In congruence with that study we found OCS155 marine group sequences at an annual average abundance of ~4% with maximum abundances of 13.8% during autumn. During summer the Marine Group II (*Euryarchaeota*) became a dominant group, exhibiting relative abundances of up to 15.6%. Pernthaler *et al.* (2002) even determined that this group accounted for >30% of the total picoplankton in North Sea surface waters during spring and summer using polyFISH. Taken together these findings demonstrate resilience of a few bacterial core taxa as was also reported for the English Channel by Caporaso *et al.* (2012). Multiple regression analyses also suggest that the seasonal succession of dominating bacterial OTUs reflects successions of their corresponding niche optima, which in this study are mainly defined by temperature. This temperature-driven succession of core taxa is interrupted during short-term events such as phytoplankton blooms. We suggest that enhanced substrate supply

during these blooms favours taxa capable of a feast-and famine lifestyle resulting in short-term peaks of these taxa.

Short-term variation during spring and summer blooms

Short-term succession is not only driven by phytoplankton but also influenced by changes in hydrographic currents. Hydrographic conditions at Helgoland Roads are governed by an inflow of marine waters from the north-west off the island (Fig. S2 B). This current pattern is related to positive amplitudes of PC2 (Fig. S3). Shortly before both phytoplankton blooms in April and August PC2 exhibited negative amplitudes which indicate a reversed current pattern and thus an inflow of nutrient-rich coastal waters that boosted the phytoplankton blooms at Helgoland Roads. A similar situation has been observed at Helgoland Roads during a spring phytoplankton bloom in 2009 (Teeling et al. 2012). However, the effect of coastal water inflow during summer observed in this study seemed not to be as strong as in spring. Phytoplankton is generally considered as the dominant source of bioavailable DOM in ocean surface waters (Hedges 1992), and heterotrophic bacteria strongly rely on this DOM (Baines & Pace 1991). In conjunction with the above-mentioned overall temperature dependence of the North Sea bacterioplankton, a comparison of bacterioplankton assemblages during spring and summer blooms is particularly interesting. This study is investigating the free-living fraction of the bacterial community only. It is noteworthy that in marine coastal environments and especially during phytoplankton blooms a large fraction of the bacterial community may be attached to particles (e.g. Simon et al. 2002). Lots of studies examined particle-attached and free-living communities in different aquatic environments and found that the free-living bacteria are often more abundant (Ghiglione et al. 2007), but particle-attached communities are more active (e.g. Crump & Baross 2000; Ghiglione et al. 2007). However, comparison of the community composition of free-living and particle-attached bacteria using high-throughput 16S rRNA gene sequencing methods revealed minor differences between both fractions. Campbell and Kirchman (2013) for instance reported that the free-living and particle attached bacteria along a salinity gradient clustered together and shared similar abundances of most bacteria groups. This is also supported by Ortega-Retuerta *et al.* (2013) who found that the community composition of both fractions is similar especially at higher oceanic salinities.

In contrast to Teeling *et al.* (2012) who observed a diatom-dominated spring bloom in 2009, the spring bloom during our study was dominated by dinoflagellates and flagellates.

Diatoms reached their maximum abundance during the summer bloom in August. Spring and summer blooms were dominated by *Alphaproteobacteria*, *Gammaproteobacteria* and *Flavobacteriia.* These classes have been consistently found to dominate bloom-associated bacterial communities as reviewed by Buchan *et al.* (2014). However, we did observe differences between the two blooms at higher taxonomic resolution.

During the spring bloom, dominating OTUs within the *Alphaproteobacteria* included the RCA cluster and NAC11-7 lineage. The RCA cluster tag sequences exhibited higher relative abundances than the NAC11-7 lineage during the beginning of the bloom and the Chl *a* maximum. In response to bloom decay, relative abundance of the RCA cluster tag sequences increased from 5% to 17% of the trimmed tag data. This is consistent with Giebel *et al.* (2011) who reported relative abundances of the RCA cluster of 15% during a phytoplankton bloom in the southern North Sea via quantitative PCR. The NAC11-7 lineage exhibited an even stronger response to algal decay, increasing in read frequency from ~1% to 15% and took over dominance following the RCA cluster after the bloom decay. Conversely, Teeling *et al.* (2012) observed a succession of *Roseobacter* clade members, with the NAC11-7 lineage dominating the early bloom phase and the RCA cluster dominating the late bloom phase. Dominating spring bloom *Flavobacteriia* were *Polaribacter*, NS5 marine group members and a *Cryomorphaceae* related cluster, all of which are known to react to phytoplankton blooms where they are likely involved in biopolymer degradation (Lau *et al.* 2005; Gómez-Pereira *et al.* 2012; Teeling *et al.* 2012; Xing *et al.* 2014). *Alteromonadales* (SAR92 clade, *Pseudoalteromonas*) and SAR86 clade members were the dominating *Gammaproteobacteria* during the spring bloom. Consistent with our study, SAR92 phylotypes have been demonstrated to react to a phytoplankton bloom decay during spring in 2009 at Helgoland Roads (Teeling *et al.* 2012). *Pseudoalteromonas* phylotypes are well known to produce exo-proteases that enable them to degrade complex algae-derived organic matter (Holmström & Kjelleberg 1999; Lee e*t al.* 2000; Ivanova *et al.* 2002; Vázquez *et al.* 2008). The ability to rapidly react to enhanced substrate supply during phytoplankton blooms is reflected in the short-term peaks of *Pseudoalteromonas* during the spring bloom in our study.

During the diatom-dominated summer bloom, the most abundant *Alphaproteobacteria* tag sequences affiliated with the *Roseobacter* clade NAC11-7 and OCT lineages as well as the SAR116 clade. *Roseobacter* clade members are found to associate with phytoplankton blooms and are particularly important for the degradation of dimethylsulfonylpropionate (DMSP), an abundant algal osmolyte (Buchan et al. 2005). The SAR116 clade is an

ubiquitous marine bacterioplankton lineage (Giovannoni & Rappé 2000). The first cultivated SAR116 strain was shown to possess the dmdA gene, responsible for DMSP demethylation (Oh et al. 2010) indicating possible association with phytoplankton blooms. *Flavobacteriia* during the summer bloom were dominated by the NS5 marine group and *Cryomorphaceae*. Both of these clades were also abundant during the spring bloom when nutrient concentrations were higher as compared to the summer bloom. This suggests that members of these clades can cope with a broad range of nutrient concentrations as well as DOM from different phytoplankton species. Summer bloom *Gammaproteobacteria* were dominated by the NOR5 lineage, ZD0405 (*Oceanospirillales*) and the SAR86 clade. The NOR5 lineage has been found to be able to cope with both, nutrient poor and nutrient rich conditions and to occur in pronounced association with phytoplankton blooms (Eilers et al. 2001; Yan et al. 2009). However in this study the NOR5 lineage becomes dominating during the summer bloom only. A similar situation was observed for the SAR86 clade that dominated during both blooms but exhibited much higher relative abundances during the summer bloom. Concerning that both, the NOR5 lineage and the SAR86 clade were positively correlated with temperature this hints at the importance of temperature as an influencing factor for the response of bacterial OTUs to phytoplankton blooms and thus, points to its potential as a main niche buildner.

Comparison of spring and summer blooms revealed similar successions on class level, with *Alphaproteobacteria* dominating the early bloom phase, *Flavobacteriia* increasing in relative abundances as the bloom commences and *Gammaproteobacteria* increasing as the bloom decays. The same succession of bacteria classes was reported for the 2009 spring bloom at Helgoland Roads (Teeling et al. 2012). However, the relative abundances of *Alphaproteobacteria* and *Flavobacteriia* were much higher during the spring bloom (32.9 and 30.4%) compared to the summer bloom (27.2 and 20.8%), when *Gammaproteobacteria* increased strongly in relative abundances to ~48% and even dominated the whole trimmed community during the summer bloom decay. From the spring to the summer bloom, the temperature increased by about 8.8 °C, while the proportion of *Flavobacteriia* was lower during the summer bloom as compared to the spring bloom. This is also supported by our network analysis, which revealed that all significant correlations of *Flavobacteriia* with PCoA2 (i.e. temperature) were negative. This agrees with Tada *et al.* (2013) who stated that growth of *Bacteroidetes* is positively influenced by the quantity and quality of organic matter concentrations, but their contribution to organic matter cycling is larger at colder conditions. We additionally found

an increase of low abundance OTUs; most noticeably we found the *Thermoplasmata* related Marine Group II (*Euryarchaeota*) as a dominating group. We hypothesize that the capability of the Marine Group II to positively respond to phytoplankton blooms is triggered by temperature. This is supported by measurements of the consumption of proteins and lipids during a spring bloom in the north-western Pacific which indicated a potential interaction between diatoms and members of the Marine Group II (Iverson et al. 2012). Additionally, the Marine Group II is known to have a cosmopolitan distribution in marine surface waters and to be abundant during summer months (Pernthaler et al. 2002; Herfort et al. 2007).

Multiple regression models exhibited especially large residuals during summer (Fig. 3), when ciliates and flagellates exhibited pronounced peaks in abundance. There is evidence that the community structure of pelagic bacterial assemblages can be shaped by size-selective protistan predation, which might lead to profound shifts in community composition as reviewed in Pernthaler (2005). Although heterotrophic nanoflagellates smaller than 5 µm account for about 80% of total bacterivory (Unrein et al. 2007), the relative importance of grazing by ciliates seems to be especially high in coastal and estuarine systems (Sherr & Sherr 1987; Simek et al. 2000). It is noteworthy that in this study heterotrophic nanoflagellate cell numbers are included in "flagellate" cell numbers. Thus the potential impact of these grazers is already considered in the regression model. However, ciliate cell numbers did not contribute significantly to the regression model but might explain the large difference between observed and predicted values.

Temperature as major constraint for ecological niches

Growth and activity of heterotrophic bacteria are fuelled by enhanced DOM supply as found during phytoplankton blooms. Bacterial metabolic processes, such as the decomposition of organic matter, are also enhanced by increasing temperature (Pomeroy & Wiebe 2001; Kirchman *et al.* 2009). Thus, increasing water temperatures lead to a tighter coupling of phyto- and bacterioplankton as shown in Hoppe *et al.* (2008) and Wohlers-Zöllner *et al.,* (2012), but it may also result in shifts of bacterial community composition due to different temperature optima of distinct bacterial taxa. We assume that the enhanced supply of DOM by phytoplankton results in successional patterns of taxa that have different niches with respect to organic matter decomposition. This kind of nutrient partitioning was shown during a comprehensive metagenomic and metaproteomic study on the 2009 spring phytoplankton bloom at Helgoland Roads (Teeling et al. 2012). However, in our study we observed succession of different dominant OTUs during the spring and summer blooms. Network analysis revealed group formation of these OTUs, exhibiting specific correlation patterns with temperature. This and the finding that all dominant OTUs found during bloom successions were present throughout the whole year suggest that there is a resident pool of bacterial taxa. This is supported by Caporaso *et al.* (2012) who demonstrated in a comprehensive taxonomic survey in the English Channel, that the vast majority of taxa identified are always present in differing proportions that are predictable. We assume that the different OTUs we found during spring and summer blooms have redundant functional capacities, but are favoured either during spring or summer blooms, based on their ecological niche affiliations, which again seems to be largely defined by temperature. This notion is for instance supported by the dominance of Marine group II *Euryarchaeota* and NOR5/OM60 clade members only during the summer bloom. According to Yan *et al.* (2009) and Pernthaler *et al*. (2002) both taxa exhibit especially high abundances (>30% and up to 13% of total picoplankton communities, respectively) during summer and autumn. Von Scheibner *et al.* (2014) conducted a mesocosm experiment with incubation of natural plankton communities from the Baltic Sea during a phytoplankton bloom at in situ and increased temperatures. They reported an influence of both, the phytoplankton bloom phase and temperature on the bacterial community composition. They found that bacterial communities incubated at warmer temperatures were enriched by additional taxa compared to the communities at lower temperatures. Other studies also suggest that the influence of physico-chemical factors (e.g. day length, salinity, temperature and nutrients) on the microbial diversity in highly dynamic systems

such as estuaries of continental shelf seas is more important than biotic factors (Kirchman et al. 2005; Teira et al. 2008; Gilbert et al. 2012; Sintes et al. 2013).

Although our results suggest a direct impact of temperature on the bacterial community structure at Helgoland Roads we must consider additional linkage to other factors not examined during this study. For example, different taxa dominated the phytoplankton during spring and summer blooms. The release of distinct organic matter in dissolved and particulate form by different phytoplankton species and the differences in the capacity of heterotrophic bacterial populations to consume these differing substrates stimulate discussion about the influence of the phytoplankton composition on changes in bacterial community compositions (e.g. Pinhassi et al. 2004; Rooney-Varga et al. 2005; Sarmento & Gasol 2012; Becker et al. 2014). The spring bloom in this study was dominated by a combination of dinoflagellates and flagellates. In contrast to our study Teeling *et al.* (2012) investigated a diatom dominated spring bloom in 2009 at the same sampling site. Although the dominating phytoplankton groups differ between the two studies, similar dominating bacterial taxa (NAC11-7 lineage, RCA cluster, *Polaribacter*, SAR92 clade) have been found during the blooms. This might support the notion that heterotrophic bacteria react to the general substrate supply during phytoplankton blooms independently of the phytoplankton composition. However, since we cannot provide detailed data on the phytoplankton species composition, assumptions on the coupling of specific OTUs and the phytoplankton composition would be speculative. Another important factor that shapes the bacterioplankton community composition is the cell lysis by viruses. Viruses are well known to primarily affect the largest, most rapidly growing bacterial populations and by this suppress particular bacterial species (Thingstad 2000). The released DOM again favours the surviving bacterial species. All of the above mentioned abiotic and biotic factors also exhibit interactions and thus, affect each other. Therefore, changes in bacterial community composition are likely controlled by complex combinations of these factors rather than by single parameters.

Nonetheless, the temperature signal that was captured by our statistical analyses was significant, as was the influence of phytoplankton blooms. Hence, both of these factors exert a major influence on the bacterioplankton community at Helgoland Roads in the North Sea. We found a pronounced seasonal pattern and indications that this pattern might be annually recurring, which however needs to be evidenced with studies that span multiple years. The pronounced formation of temperature-dependent guilds during spring and summer phytoplankton blooms lets us conclude that short-term bacterial succession in

response to phytoplankton blooms is indirectly affected by temperature as a major factor for the formation of ecological niches, resulting in distinct bacterial communities during colder spring bloom phases and warmer summer bloom phases. For future analyses, access to representative strains of relevant bacterial clades is needed for comprehensive examination of functional capacities under defined experimental conditions.

Acknowledgements

We would like to thank the crew members of the research vessel Aade, Kristine Carstens, Silvia Peters and Matthias Friebe for their technical assistance. We are grateful for kind provision of environmental data by Prof. Dr. Karen Wiltshire. We thank the Bundesamt für Seeschiffahrt und Hydrographie (BSH) in Hamburg for providing the BSHcmod current data. We also appreciate the valuable criticism by two anonymous reviewers. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported under contract no. DE-AC02-05CH11231.

Supplementary material

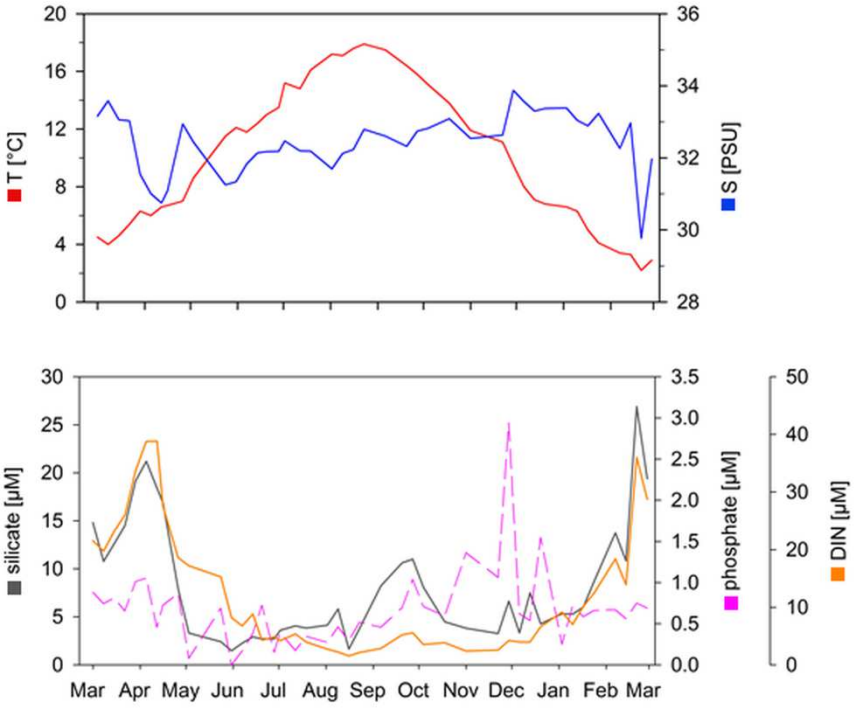


FIGURE S1: Environmental parameters recorded from 01. March 2012 – 28. February 2013 at Helgoland Roads. T: temperature, S: salinity, DIN: dissolved inorganic nitrogen.

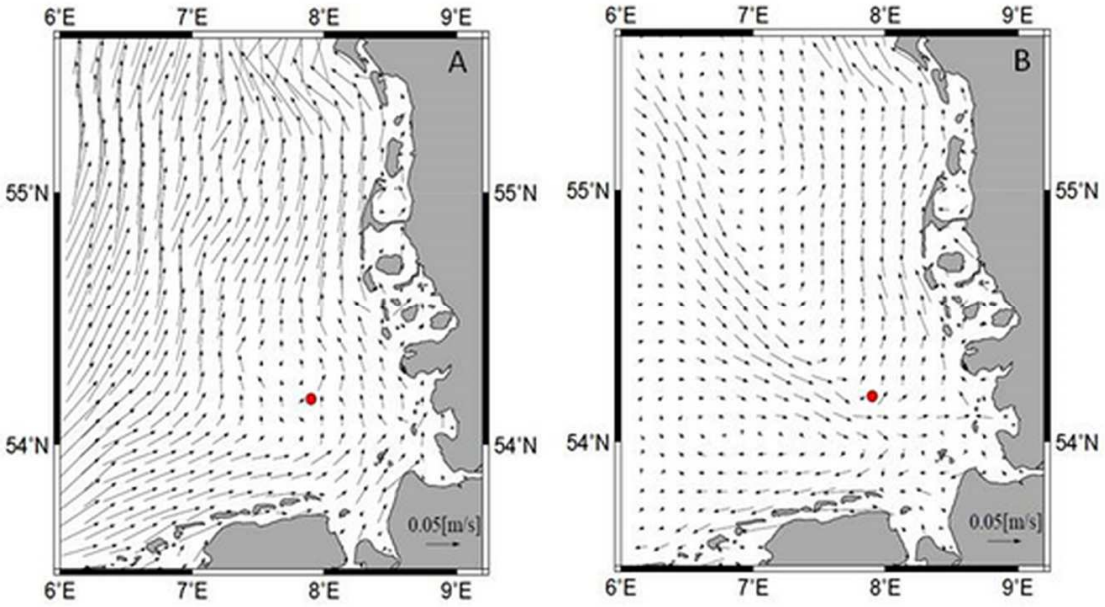


FIGURE S2: Vector fields of current anomalies (EOF pattern) in the German Bight within the period March 2012 – March 2013. Explained variances are 73.4 % for the first (A) and 12.2% for the second (B) EOF. Red dot: Helgoland.

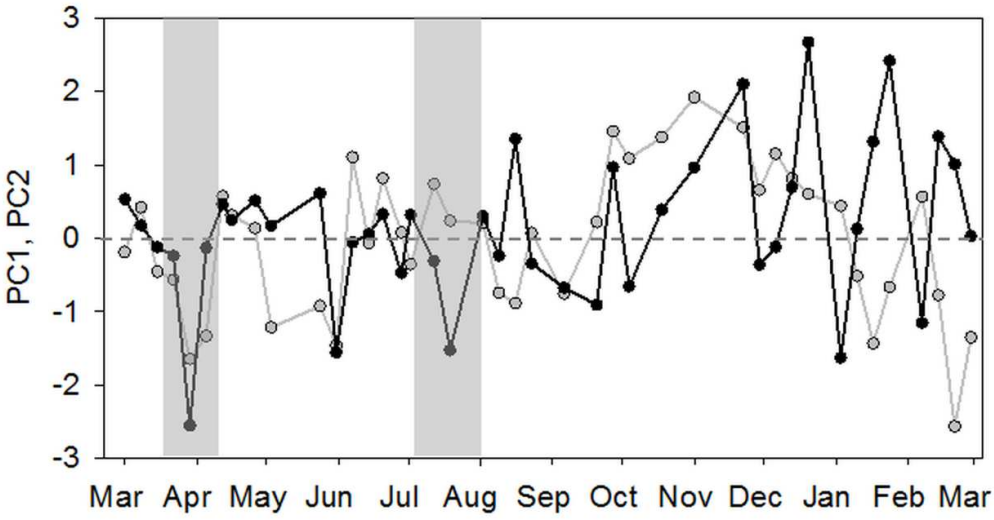


FIGURE S3: Principal components (PCs) corresponding to the EOF pattern shown in Figure S2. Grey areas indicate negative deviations of PC2 shortly before the spring and summer phytoplankton blooms.

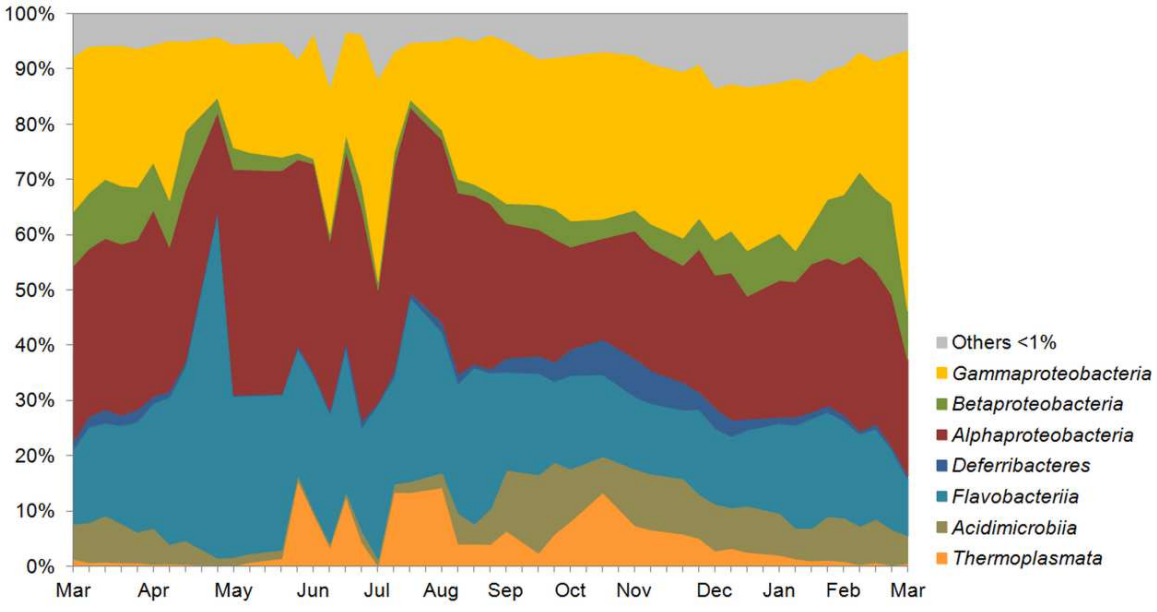


FIGURE S4: Relative abundances of dominating bacteria classes at Helgoland Roads 01. March 2012-28. February 2013. Relative abundances are given in % of total tag sequence data.

Table S1: Physico-chemical and phytoplankton data recorded from 01. March 2012 – 28. February 2013 at Helgoland Roads. S: salinity; T: temperature [°C]; DOC: dissolved organic carbon [µg l^{-1}]; SiO_2: silicate [µmol l^{-1}]; PO_4^{3-}: phosphate [µmol l^{-1}]; DIN: dissolved inorganic nitrogen (DIN = NO_2^- + NO_3^- + NH_4^+) [µmol l^{-1}]; Chl *a*: chlorophyll *a* [µg l^{-1}]; PC1, PC2: Principal components of EOFs; bac: bacteria cell counts [cells ml^{-1}]; diatom [cells l^{-1}]; dino: dinoflagellates [cells l^{-1}]; flag: flagellates [cells l^{-1}]; cili: ciliates [cells l^{-1}].

Sample	S	T	DOC	SiO_2	PO_4^{3-}	DIN	Chl *a*	PC1	PC2	Bac $*10^5$	Diatom $*10^4$	Dino $*10^4$	Flag $*10^6$	Cili $*10^3$
01.03.12	33.2	4.5	122.9	14.80	0.88	21.5	1.4	-0.189	0.532	1.09	1.24	1.34	2.21	0.06
08.03.12	33.6	4	126.5	10.78	0.74	19.7	1.4	0.420	0.175	1.82	2.11	1.29	2.18	0.1
15.03.12	33.1	4.6	117.6	12.63	0.81	23.2	1.1	-0.454	0.124	2.19	1.01	2.67	2.40	0.36
22.03.12	33.0	5.4	158.5	14.51	0.66	26.0	1.0	-0.565	0.240	2.54	0.17	1.76	1.72	1.16
29.03.12	31.5	6.3	139.7	19.16	1.01	33.7	1.3	-1.646	2.551	2.46	3.44	2.20	2.25	0.54
05.04.12	31.0	6	149.0	21.19	1.05	38.8	1.4	-1.330	0.128	4.24	3.08	5.54	2.39	2.04
12.04.12	30.7	6.6	136.0	18.42	0.46	38.8	1.9	0.572	0.463	5.16	0.54	8.80	3.10	2.02
16.04.12	31.1	6.7	128.5	16.87	0.72	27.6	4.0	0.316	0.247	5.35	3.24	8.37	4.23	0.42
26.04.12	32.9	7	113.0	7.94	0.87	18.6	6.4	0.139	0.515	8.11	2.5	4.80	3.68	0.32
03.05.12	32.4	8.6	120.1	3.31	0.08	17.2	0.8	-1.217	0.171	13.1	3.46	1.68	2.37	1.48
24.05.12	31.3	11.5	154.2	2.42	0.69	15.3	1.1	-0.922	0.612	6.13	3.36	2.80	3.25	0.28
31.05.12	31.3	12.1	146.4	1.46	0.00	8.2	3.0	-1.459	1.558	6.11	13.3	5.94	2.96	5.57
07.06.12	31.8	11.8	125.5	2.27	0.17	6.8	2.8	1.104	0.055	5.85	16.1	7.51	2.98	6.18
14.06.12	32.1	12.4	147.7	2.92	0.38	8.8	2.2	-0.068	0.056	7.75	34.6	12.5	5.85	2.76
20.06.12	32.2	13	123.8	2.71	0.72	4.3	2.4	0.818	0.326	8.38	22.2	12.4	4.24	3.04
28.06.12	32.2	13.5	179.8	2.70	0.15	4.9	2.5	0.078	0.471	11	2.74	11.0	4.15	6.68
05.07.12	32.5	15.2	139.7	3.60	0.39	4.2	2.5	-0.353	0.319	8.44	1.06	5.65	2.61	1.24
12.07.12	32.2	14.8	147.7	4.07	0.17	5.4	1.3	0.737	0.312	5.96	1.83	5.16	2.82	0.44
19.07.12	32.2	16.1	122.1	3.83	0.35	4.0	1.5	0.238	1.523	6.99	2.09	4.79	2.37	0.92
02.08.12	31.7	17.2	136.7	4.14	0.27	2.7	1.7	0.207	0.301	9.06	15.5	4.03	3.02	4.20
09.08.12	32.1	17.1	153.6	5.82	0.46	2.2	2.6	-0.744	0.237	10.3	39.9	7.69	4.12	6.72
16.08.12	32.2	17.6	150.5	1.63	0.29	1.6	6.8	-0.885	1.355	15.7	130	6.90	4.89	10.7
23.08.12	32.8	17.9	131.5	3.77	0.52	2.1	2.4	0.072	0.346	11.8	12.0	10.1	3.81	20.5
06.09.12	32.6	17.5	294.8	8.22	0.46	2.8	1.7	-0.756	0.673	7.35	0.99	6.62	2.27	0.84
20.09.12	32.3	16.4	126.9	10.63	0.70	5.2	1.4	0.221	0.907	7.92	2.65	2.50	2.15	0.006
27.09.12	32.7	15.8	132.0	11.01	1.04	5.6	1.4	1.457	0.966	7.92	4.25	3.75	2.15	0
04.10.12	32.8	15.1	113.3	8.10	0.71	3.5	1.1	1.088	0.656	10.4	3.62	3.80	1.94	0.004
18.10.12	33.1	13.8	107.0	4.49	0.59	3.8	1.4	1.375	0.387	6.52	2.76	2.86	2.19	0.63
01.11.12	32.5	11.9	108.7	3.82	1.36	2.4	1.3	1.919	0.963	8.36	3.13	1.50	1.97	0
22.11.12	32.6	11.1	111.7	3.26	1.06	2.6	1.0	1.508	2.100	5.11	1.39	0.52	1.68	0
29.11.12	33.9	9.5	107.1	6.64	2.94	4.2	0.9	0.658	0.360	4.01	1.81	1.21	1.72	0
06.12.12	33.6	8	111.0	3.33	0.62	4.0	1.2	1.152	0.118	2.38	2.23	2.47	2.19	0
13.12.12	33.3	7.1	114.2	7.50	0.54	4.0	1.0	0.814	0.695	3.11	1.65	3.21	1.65	0
20.12.12	33.4	6.8	98.4	4.27	1.55	6.6	1.0	0.603	2.670	3.67	1.94	1.48	2.00	0
03.01.13	33.4	6.6	126.6	5.33	0.25	9.2	0.7	0.440	1.627	1.78	0.33	1.09	2.31	0
10.01.13	33.0	6.3	111.4	5.27	0.72	7.0	0.8	-0.515	0.125	1.70	0.16	1.60	1.76	0
17.01.13	32.9	5	105.4	6.02	0.59	10.2	0.7	-1.433	1.317	1.72	0.23	0.8	1.77	0
24.01.13	33.2	4.1	90.2	8.67	0.66	12.4	0.7	-0.667	2.417	1.92	0.32	4.64	1.97	0.008
07.02.13	32.3	3.4	119.0	13.75	0.67	18.4	1.2	0.566	1.153	1.85	0.53	1.43	2.19	0.002
14.02.13	33.0	3.3	101.3	10.82	0.56	13.9	2.9	-0.777	1.388	3.22	1.06	2.42	2.83	0
21.02.13	29.8	2.2	148.7	26.90	0.75	36.0	1.4	-2.563	1.009	2.64	1.12	1.74	2.57	0.002
28.02.13	32.0	2.9	128.4	19.39	0.69	28.7	0.8	-1.351	0.030	3.04	0.49	1.28	2.03	0.34

Table S2: Spearman rank correlation of all environmental parameters. Significant correlations ($p<0.05$) are indicated in bold. DOC: dissolved organic carbon; SiO_2: silicate; PO_4^{3-}: phosphate; DIN: dissolved inorganic nitrogen (DIN = NO_2 + NO_3^- + NH_4^+); Chl *a*: chlorophyll *a*; PC1, PC2: Principal components of mean surface currents.

	Salinity	Temperature	DOC	SiO_2	PO_4^{3-}	DIN	Chl *a*	PC1	PC2	Diatoms	Dinoflagellates	Flagellates	Ciliates
Temperature	-0.171												
DOC	**-0.633**	**0.325**											
SiO_2	0.017	**-0.598**	-0.018										
PO_4^{3-}	0.241	**-0.381**	**-0.359**	**0.484**									
DIN	-0.208	**-0.803**	0.085	**0.634**	0.235								
Chl *a*	**-0.430**	**0.449**	**0.397**	-0.212	-0.287	-0.186							
PC1	**0.343**	0.281	**-0.398**	-0.213	0.175	**-0.431**	-0.010						
PC2	0.145	-0.218	**-0.358**	-0.026	0.293	0.013	-0.044	0.076					
Diatoms	**-0.409**	**0.624**	0.298	**-0.480**	-0.151	**-0.383**	**0.552**	0.126	-0.077				
Dinoflagellates	**-0.494**	**0.561**	**0.496**	-0.298	**-0.452**	-0.216	**0.739**	-0.009	-0.144	**0.547**			
Flagellates	**-0.587**	**0.334**	**0.532**	**-0.309**	**-0.446**	0.012	**0.773**	-0.224	-0.084	**0.554**	**0.749**		
Ciliates	**-0.569**	**0.460**	**0.639**	-0.304	**-0.552**	-0.098	**0.612**	-0.290	**-0.307**	**0.537**	**0.757**	**0.724**	
Bacteria	**-0.370**	**0.823**	**0.346**	**-0.505**	**-0.334**	**-0.569**	**0.546**	0.116	-0.059	**0.711**	**0.622**	**0.507**	**0.577**

TABLE S3: Results from multiple regression analyses using stepwise forward selection. Significant values (p<0.05) are indicated in bold. R^2, p and F values for models of PCoA1 and PCoA2 are given. Also displayed are p and F values for all variables as well as R^2, Delta R^2 and step of entering the model for variables in the model. N=42.

	p (model)	R^2 (model)	F (model)	variable	p (var)	F (var)	Step	R^2 (cum)	Delta R^2
PCoA1	**<0.001**	**0.818**	**41.537**	salinity	**<0,001**	**26,152**	**3**	**0,786**	**0,146**
				temperature	0,964	0,00206			
				DOC	0,672	0,182			
				SiO_2	**<0,001**	**74,229**	**2**	**0,639**	**0,220**
				PO_4^{3-}	0,571	0,327			
				DIN	0,338	0,944			
				Chl *a*	0,652	0,206			
				PC1	0,146	2,204			
				PC2	**0,015**	**6,528**	**4**	**0,818**	**0,0321**
				diatoms	0,593	0,290			
				dino	0,282	1,192			
				flagellates	**0,001**	**12,324**	**1**	**0,420**	**0,420**
				ciliates	0.653	0.205			
PCoA2	**<0.001**	**0.868**	**47.336**	salinity	0,756	0,0981			
				temperature	**<0,001**	**15,003**	**4**	**0,837**	**0,0296**
				DOC	0,946	0,00466			
				SiO_2	**<0,001**	**21,346**	**2**	**0,772**	**0,153**
				PO_4^{3-}	0,202	1,690			
				DIN	**<0,001**	**37,654**	**1**	**0,619**	**0,619**
				Chl *a*	0,256	1,330			
				PC1	**0,006**	**8,423**	**3**	**0,807**	**0,0355**
				PC2	0,452	0,578			
				diatoms	0,125	2,462			
				dino	0,843	0,0399			
				flagellates	**0,006**	**8,412**	**5**	**0,868**	**0,0308**
				ciliates	0.785	0.0753			

CHAPTER II

Short-term dynamics of North Sea bacterioplankton-dissolved organic matter interactions on molecular level

Judith Lucas[1*+], Irina Koester[2+], Antje Wichels[1], Jutta Niggemann[2], Thorsten Dittmar[2], Ulrich Callies[3], Karen H. Wiltshire,[1,4] and Gunnar Gerdts[1]

[1] Biological Station Helgoland, Shelf Sea Systems Ecology, Alfred-Wegener-Institute Helmholtz-Center for Polar and Marine Research, Helgoland, Germany

[2] Research Group for Marine Geochemistry (ICBM-MPI Bridging Group), Institute for Chemistry and Biology of the Marine Environment (ICBM), Carl von Ossietzky University Oldenburg, Oldenburg, Germany

[3] Institute of Coastal Research, Modelling for the Assessment of Coastal Systems, Helmholtz-Zentrum Geesthacht Centre for Materials and Coastal Research, Geesthacht, Germany

[4] Wattenmeerstation Sylt, Coastal Ecology, Alfred-Wegener-Institute Helmholtz-Center for Polar and Marine Research, List/Sylt, Germany

[*]Judith Lucas, Alfred-Wegener-Institute Helmholtz-Center for Polar and Marine Research, Biological Station Helgoland, Kurpromenade 201, 27498 Helgoland, Germany; Phone: +49 4725 819 3200, Fax: +49 4725 819 3283, Email: judith.lucas@awi.de

[+] These authors contributed equally to the manuscript

Abstract

Remineralisation and transformation of dissolved organic matter (DOM) by marine microbes shape the DOM composition and thus, have large impact on global carbon and nutrient cycling. However, information on bacterioplankton-DOM interactions on a molecular level is limited. We examined the variation of bacterial community composition at Helgoland Roads (North Sea) in relation to variation of molecular DOM composition and various environmental parameters on short-time scales. Surface water samples were taken daily over a period of twenty days. Bacterial community and molecular DOM composition were assessed via 16S rRNA gene tag sequencing and ultrahigh resolution Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR-MS), respectively. Environmental conditions were driven by a coastal water influx during the first half of the sampling period and the onset of a summer phytoplankton bloom towards the end of the sampling period. These phenomena led to a distinct grouping of bacterial communities and DOM composition which was particularly influenced by total dissolved nitrogen concentration, temperature and salinity, as revealed by distance-based linear regression analyses. Bacterioplankton-DOM interaction was demonstrated in strong correlations between specific bacterial taxa and particular DOM molecules, thus, suggesting potential specialization on particular substrates. We propose that a combination of high resolution techniques, as used in this study, may provide substantial information on substrate generalists and specialists and thus, contribute to prediction of bacterial community composition variation.

Introduction

The global marine net primary production is estimated at 50 Gt C per year (Hedges 1992). Part of this primary production is transferred to the marine dissolved organic matter (DOM) pool, making it one of the largest active carbon pool on earth (700 Gt), containing as much carbon as the Earth's atmospheric CO_2 or all land plant biomass (Hedges 1992). Bacterial consumption and remineralisation of DOM via the microbial loop, transfers energy to higher trophic levels and thus, provides an important base for marine food webs (Azam *et al.* 1983; Azam 1998). Between 30 to >90% of net primary production pass through the so-called labile dissolved organic carbon (DOC) fraction (Ducklow 1999) and face rapid turnover by heterotrophic prokaryotes on a time scale of hours to days. Additionally, semi-labile DOC, which exhibits turnover times of months to years and can be followed as seasonal variability in DOC concentrations, provides support for the microbial loop (Hansell 2013). Part of the DOC resists rapid bacterial degradation and as recalcitrant DOC, comprises a huge carbon pool of ~630 Gt, the largest part residing in the deep oceans. Since bacterial processing of DOM has large impact on global carbon and nutrient cycling, it is of great importance to understand how organic matter-bacteria interactions are controlled.

Different phylogenetic groups of bacteria tend to exploit different organic resources (e.g. Cottrell & Kirchman 2000; Elifantz *et al.* 2005). This resource partitioning implies, that the DOM composition influences the bacterial community composition and vice versa. Most studies that focus on DOM-bacterioplankton interactions are restricted to limited taxonomic resolution of microbial communities and selected compound classes of DOM. For instance, seasonal shifts of bulk DOC concentration and bacterial activity have been demonstrated over an annual cycle (Sintes *et al.* 2010). McCarren *et al.* (2010) examined the genomic and transcriptional response of microbial communities to addition of high molecular weight DOM in microcosms over the course of one day. However, only few studies observed interactions of bacterial communities and molecular DOM composition in marine systems (Osterholz *et al.* 2014; Medeiros *et al.* 2015; Seidel *et al.* 2015) and to our knowledge there is no study investigating the interactions of bacterial community variation and molecular DOM composition on high resolutions and short time-scales such as day to day variation.

This study examines short-term dynamics in bacterial community and DOM composition at Helgoland Roads (North Sea, German Bight) on a daily basis over a period of 20 days including the onset of the summer phytoplankton bloom. We hypothesize that changes in

the bacterial community composition are closely linked to patterns in DOM composition and vice versa. We assessed the bacterial community composition via 16S rRNA gene tag sequencing and the DOM composition via electrospray ionisation (ESI) coupled with Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR-MS). Multiple regression analyses were used to identify environmental parameters that have fundamental impact on the bacterial community and DOM compositions.

Material and Methods

Site description and sampling

From August 6 to 26, 2012 a total of 19 surface water samples were collected daily at Helgoland Roads (North Sea, German Bight, 54°18.31 N, 7°88.97 E). Environmental data including water level, water temperature, salinity, dissolved O_2 and CO_2 concentrations, turbidity, pH, SiO_2, PO_4^{3-}, NO_2^-, NO_3^- and chlorophyll *a* (Chl *a*) were obtained as part of the Helgoland Roads Time Series (Wiltshire *et al.* 2008). Data were measured with a Ferry Box system installed at the eastern pier at Helgoland. Water intake is at about 2-4 m depth, depending on tides. Data are accessible via the open database PANGAEA 2004 (http://www.pangaea.de).

Surface seawater was sampled daily at 1 pm with a thoroughly rinsed bucket from the pier at the inflow site of the Ferry Box. Samples were transferred to glass bottles (cleaned with acidified pH 2, ultrapure water and rinsed with sample water before use), transported to the lab immediately and further processed within one hour at the Biological Station Helgoland.

16S rRNA gene tag sequencing of bacterial communities

500 ml of surface seawater were vacuum filtered through 0.22 µm polycarbonate filters (GTTP, Ø 47 mm, Merck Millipore, USA) to obtain bacterial biomass. Filters were stored at -20°C and further processed within four weeks. DNA extraction was carried out as described in Sapp *et al.* (2007). Briefly, lysozyme and sodium dodecyl sulfate were used for cell lysis followed by extraction with phenol-chloroform-isoamyl alcohol (25:24:1) and precipitation with isopropanol. DNA concentration per sample and purity were determined photometrically using a Tecan Infinite© 200, NanoQuant microplate reader (Tecan, Switzerland).

16S rRNA gene tag sequencing was performed at LGC Genomics GmbH (Berlin, Germany). Community DNA samples were sent to LGC in a 96-well plate for generation of 16S V4 rRNA amplicon libraries for Illumina sequencing. Community DNA was amplified utilising amplification primers targeting the V4 region of the 16S rRNA gene using 515F (5' GTGCCAGCMGCCGCGGTAA 3') and 806R (5' GACTACHVGGGTWTCTAAT 3') (Caporaso *et al.* 2011). Primers also contained the Illumina sequencing adapter sequence and a unique barcode index. Sequencing was done on an Illumina MySeq platform using 2x250 bp chemistry.

Raw paired-end reads were merged using the FLASh 1.2.4 software (http://ccb.jhu.edu/software/FLASH/) and processed through the SILVAngs pipeline

(Quast *et al.* 2013). Sequences were de-replicated at 100% identity and then clustered within each individual sample at 98% similarity. Representative sequences from operational taxonomic unit clusters (OTUs) were classified up to genus level against the SILVA v119 database using BLAST as described by (Ionescu *et al.* 2012). Genus-level classifications were used in the final abundance matrix for downstream analyses. Each classification contained the sum of all sequences represented by OTUs with the same taxonomic path. For the purposes of this study we were not interested in diversity calculated at the level of 98% clustered OTUs but rather used BLAST identities as our operational taxonomic unit. From this point on, we define these taxa as OTUs for simplicity. Eukaryotic, chloroplast and mitochondria-derived OTUs were removed from the resulting OTU matrix. Only OTUs with an average relative abundance $\geq 0.1\%$ were considered for further analysis.

Sequence data were deposited in the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (NCBI) under accession number SRP058371.

Dissolved organic matter (DOM)

For DOM extraction, two liters of each sample were filtered through 2 and 0.7 µm glass fiber filters (GMF and GF/F, Whatman, United Kingdom, combusted at 400°C, 4 h), acidified to pH 2 (HCl 32% p.a., Carl Roth, Germany) and stored at 4 °C in the dark. Aliquots of the acidified 0.7 µm filtrate were sampled for quantification of dissolved organic carbon (DOC) and total dissolved nitrogen (TDN). DOC and TDN concentrations were analyzed by high-temperature catalytic combustion using a TOC-VCPH/CPN Total Organic Carbon Analyzer equipped with an ASI-V autosampler and a TNM-1 module (Shimadzu, Japan). Prior to analysis, the acidified samples were purged with synthetic air to remove dissolved inorganic carbon. L-arginine solutions ranging from 5 to 500 µmol C l^{-1} and 6.6 to 333.3 µmol N l^{-1}, respectively, were used for calibration and Deep Atlantic Seawater reference material (DSR, D. A. Hansell, University of Miami, Florida, USA) was measured during each run to control for instrumental precision and accuracy. Samples were measured in duplicates, average deviation of duplicate analysis was 4.4 % for DOC and 1.6 % for TDN.

DOM was extracted using modified styrene divinyl benzene polymer cartridges (PPL, Agilent, USA) as described in Dittmar *et al.* (2008). Cartridges were rinsed with two cartridge volumes of pH 2 ultrapure water to remove remaining salts, dried with inert pure argon gas and eluted with 6 ml methanol (ULC/MS grade, Biosolve, Netherland) into amber vials. The extract volume was determined by weight. 100 µl of the methanol

extracts were evaporated overnight and re-dissolved in 10 ml ultrapure water at pH 2 for DOC analysis. The extraction efficiency was calculated as percentage DOC amount of the extract on the DOC amount of the original sample.

Mass spectra were obtained using a 15 T Solarix FT-ICR-MS (Bruker Daltonics, USA) equipped with an electrospray ionization source (Bruker Apollo II) applied in negative mode. Methanol extracts were diluted in a 1:1 ratio with ultrapure water to a final concentration of 20 mg C l^{-1}. A total of 500 scans were accumulated per run and mass spectra were evaluated in the range from 150 to 2000 Da. Mass spectra were calibrated with an internal calibration list of known molecular formulae mass peaks (Bruker Daltonics Data Analysis 4.0 SP 3 software package). Mass to charge ratios, peak intensities and resolutions were exported and molecular formulae were assigned to the detected mass peaks with a minimum signal-to-noise ratio of 4, according to Koch *et al*. (2007). Masses were kept for further data analysis when detected in more than two samples. Masses present in less than 20% of the samples were allowed if the S/N ratio was >20 in at least one sample. Additionally, formulae were deleted that contained following combinations of heteroatoms: NSP, N_2S, N_3S, N_4S, N_2P, N_3P, N_4P, NS_2, N_2S_2, N_3S_2, N_4S_2, PS_2. Remaining double assignments were removed. Peak intensities were normalized to the sum of peak intensities of all masses. Masses which are listed as known contaminations including their homologous series and ^{13}C peaks were removed. For each assigned formula the double bond equivalents ($DBE = 1 + ½(2C–H + N + P)$) as a measure for the degree of unsaturation (Koch & Dittmar 2006) and the modified aromaticity index ($AI_{mod} = (1 + C–½O–S–½H) / (C–½O–S–N–P)$) were calculated to assess the presence and abundance of aromatic structures (Koch & Dittmar 2006). Based on elemental ratios, AI_{mod} and heteroatom contents, molecular formulae can be categorized into compound groups (Seidel *et al.* 2014).

Statistical analysis

Principal coordinates analyses (PCoA) were accomplished to reveal patterns in environmental parameters, bacterial community composition (BCC) and DOM composition. Environmental variables were log transformed and normalised prior to analyses. PCoA for environmental parameters was carried out using Euclidean distances. Patterns in BCC were revealed by conducting PCoA of OTU read numbers using Hellinger distance (Legendre & Legendre 1998), which uses square root transformed relative abundances for distance calculation. Patterns in DOM composition were observed based on Bray-Curtis distances, generated from square root transformed mass spectrometric data.

Samples were grouped via non-hierarchical group-average linkage clustering implemented in the non-parametric *k-R clustering* approach of the Primer v.7 software package (PRIMER-E, UK). In this approach, the classic idea of *k-means clustering*, which seeks to minimise within-group sums of squares about k group centroids, is generalised to non-parametric *k-R clustering* which analogously maximises ANOSIM R and thus, allows the application of any resemblance measure desired. Based on the PCoA patterns the desired number of groups was specified *a priori* to k=3 for environmental data and k=2 for 16S rRNA tag sequencing and DOM data. An iterative search then attempts to divide the samples into k groups in such a way that samples with greatest similarities (defined as the average of the pairwise similarities between a sample and all members of the same group) fall into one group. Significance of groups was confirmed using permutational multivariate analysis of variance (PERMANOVA) with fixed factors and 999 permutations at a significance level of p<0.05 (see Tab. S3). Analysis of variance (ANOVA) was applied at a significance level of p<0.05 using Statistica 11 (StatSoft, USA), to test for significant difference of single environmental parameters between groups of samples.

The linear discriminant analysis effect size method (LEfSe) (Segata *et al.* 2011) was used to determine particular bacterial taxa and DOM molecules which were most likely to explain differences between two groups of samples. LEfSe uses the non-parametric factorial Kruskal-Wallis sum-rank test to detect features (OTUs or DOM molecules respectively) with significant differential abundance with respect to the groups of interest. Linear discriminant analysis (LDA) is then used to rank features according to their relative difference (effect size) among groups. Kruskal-Wallis tests were done on a significance level of p<0.05. The threshold on the logarithmic LDA score for discriminative features was set at 2. An implementation of LEfSe including a convenient graphical interface

incorporated in the Galaxy framework (Giardine *et al.* 2005; Blankenberg *et al.* 2010; Goecks *et al.* 2010) is provided online at http://huttenhower.sph.harvard.edu/lefse/.

Correlations between all environmental parameters were determined using Spearman rank order correlation (Statistica 11, StatSoft, USA) to reveal multicollinearities. Based on these correlations, environmental parameters were selected for multiple regression analysis to unravel their relationship with BCC and DOM composition. Multiple regression analyses were performed using distance-based linear modeling (DistLM). DistLM models were build using stepwise selection, adjusted R^2 and applying 999 permutations at a significance level of $p<0.05$. Results were visualized via distance-based redundancy analysis (dbRDA). All multivariate analyses were performed using the Primer v.7 software package (PRIMER-E, UK). To further unravel the relationship of DOM molecules with specific environmental parameters, correlations of DOM molecules with salinity, temperature and DOC were calculated using Pearson product-moment correlation (Statistica 11, StatSoft, USA).

To investigate the relationship between specific OTUs, DOM compounds and environmental parameters, pairwise correlations were calculated with R (R Development Core Team 2014) using Pearson product-moment correlation at a significance level of $p<0.05$. When considering several hypotheses in the same test the problem of multiple statistical inference arises (Holm 1979). If one accounts for this family-wise error rate, e.g. via the Holm-Bonferroni correction (Holm 1979), few of the apparent correlations would remain statistically significant. We compared raw data of OTU relative abundances and molecular formulae intensities to demonstrate that the observed correlations are plausible and consistent, and do not occur in a random fashion (Fig. S2). High correlations ($r\leq-0.9$ or ≥0.9) were visualized in a network constructed using Cytoscape version 3.2.0 (Shannon *et al.* 2003).

Results

Oceanographic conditions at sampling site

Concurrent with water sampling, various physico-chemical parameters and nutrient concentrations were recorded (Tab. S1). Most striking was the high variation in salinity during the sampling period (Fig. 1). During the first week of sampling, salinity decreased from 32.6 to 31.2 on August 12 2012, followed by an increase to 32.9 in the following week. Additionally, total dissolved nitrogen (TDN) concentration and temperature increased over the sampling period (Fig. 1). The Chl *a* concentration increased towards the end of the sampling period, indicating the onset of a summer phytoplankton bloom.

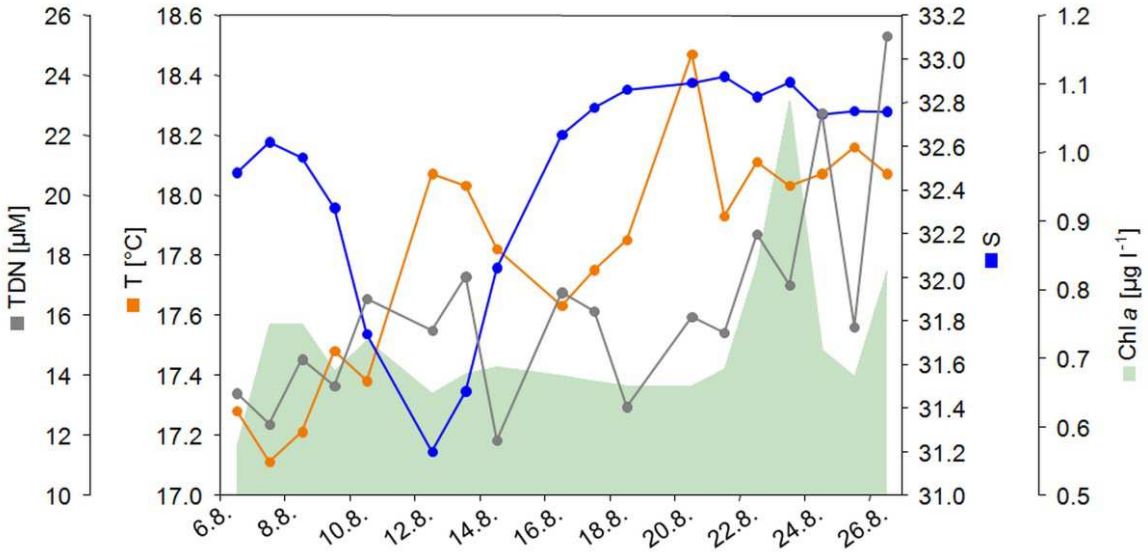


FIGURE 1: Salinity (S), temperature (T), total dissolved nitrogen (TDN) and Chlorophyll *a* (Chl *a*) concentration during the sampling period from August 6 to August 26, 2012.

Spearman rank order correlations revealed strong significant multicollinearity ($p<0.5$, $r>0.6$, Graham 2003) among turbidity and Chl *a* (R=0.890), salinity and pH (R= -0.799), salinity and DOC (R=-0.600), O_2 and CO_2 (R=-0.807), temperature and NO_2^- (R=0.645), NO_3^- and SiO_2 (R=0.607) and depth and pH (R=-0.620) (Tab. S2). The power to detect a significant effect of a predictor on a response variable decreases nonlinearly with increasing multicollinearity (Graham 2003). Therefore we decided to drop collinear variables from further analysis, knowing that this might result in a substantial loss of overall explanatory power. Based on previous studies that uncovered Chl *a* (as proxy for phytoplankton abundance) and salinity as important driving factors for bacterial

community dynamics (Fortunato *et al.* 2012; Lucas *et al.* 2015b) we decided to treat pH, turbidity and CO_2 as functionally less important and excluded those variables from all further analyses.

PCoA of environmental data suggested that samples might cluster in three groups (Fig. 2A), reflecting pronounced changes in environmental conditions during the sampling period. Non-hierarchical *k-R clustering* results however revealed, that the third group was built by a single sample (13.8.). Thus, a separation into two groups (group A and B) appeared more reasonable and the sample 13.8. was added to group A during all following analyses. ANOVA confirmed significant ($p<0.05$) differences between both groups for TDN, temperature, salinity and Chl *a* (Tab. S3). Group A is characterized by lower average temperature (17.6 °C), salinity (32.25), TDN (14.61 µM) and Chl *a* (0.68 µg l^{-1}) concentrations compared to group B where the average values were 18.13 °C, 32.84, 17.98 µM TDN and 0.79 µg l^{-1} Chl *a*.

Bacterial community composition, variation and relation to environmental parameters

A total of 1,720,615 high quality sequences were obtained, clustering into 98 different taxonomically assigned OTUs. During the sampling period the community was mainly composed of *Proteobacteria* (51.2%), *Bacteroidetes* (26.5%) and *Actinobacteria* (5.9%). On class level, *Flavobacteriia* was the predominant group (23.3%), closely followed by *Alphaproteobacteria* (22.8%) and *Gammaproteobacteria* (19.2%). Other highly abundant classes were *Betaproteobacteria* (8.2%) and *Acidimicrobiia* (5.6%).

Prevailing OTUs within the *Flavobacteriia* were the NS5 marine group, *Tenacibaculum* and a *Cryomorphaceae* related cluster (Fig. S1). *Alphaproteobacteria* were dominated by OTUs affiliated with the *Roseobacter* clade (*Candidatus Planktomarina*, NAC11-7 lineage, OCT lineage and *Sulfitobacter*). The prominent G*ammaprotoeobacteria* were *Candidatus Actinomarina*, *Oceanospirillales* related clone ZD0405 and SAR86 clade. However, *Burkholderia* (*Betaproteobacteria*) was found to be the most abundant OTU accounting on average for 5.6% of the total tag sequence data. Exceptionally high abundances of *Burkholderia* occurred at two time points (14.8. and 20.8.) when it reached up to 25% of total tag sequences (Fig. S1).

PCoA and non-hierarchical clustering of bacterial community tag data revealed a separation of samples into two groups (Fig. 2B). DistLM analysis suggested that temperature, TDN, O_2, PO_4^{3-} and SiO_2 significantly influenced this group formation (Fig. 2D, Tab. S4).

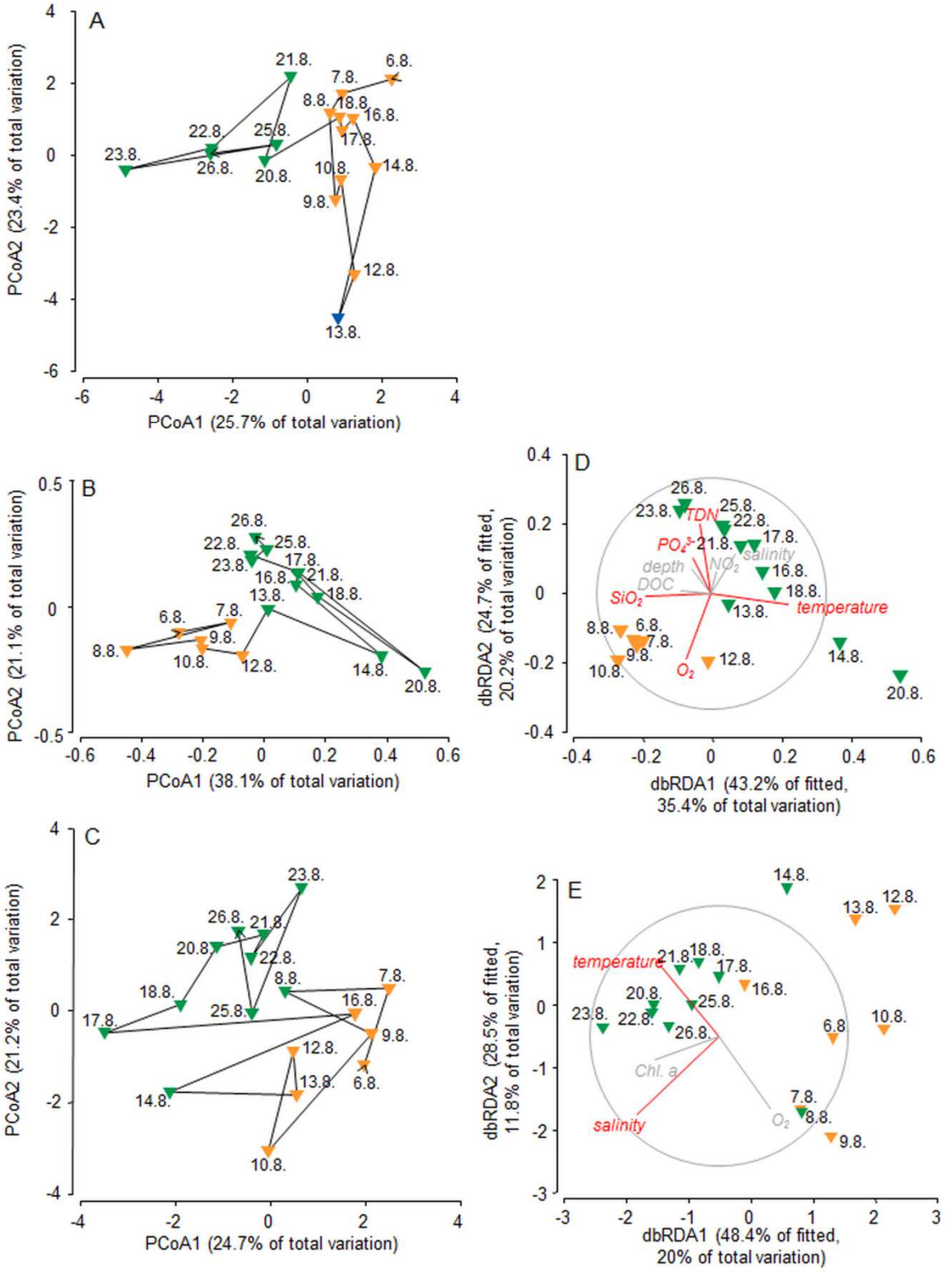


FIGURE 2: Principal coordinates analyses (PCoA) of (A) environmental variables based on Euclidean distance, (B) bacterial communities based on Hellinger distance and (C) molecular DOM composition based on Bray-Curtis similarity. Distance-based redundancy analyses (dbRDA) of (D) bacterial communities and (E) DOM composition, both based on Bray-Curtis similarities. Environmental parameters explaining the variation significantly (p<0.05) are depicted in red, non–significant parameters are depicted in grey. Color code refers to group formation according to *k*-R *Clustering*. Orange: group A, Green: group B, Blue: group C.

To determine which bacterial taxa were most likely contributing to the differences in community composition between the two groups, linear discriminant effect size analysis (LEfSe) was performed (Fig. 3A and 3B). In general, *Alpha- and Gammaproteobacteria* were dominating group A. *Gammaproteobacteria* decreased in relative abundance in group B, whereas *Flavobacteriia* became dominating (Fig. 3A). In particular the *Gammaproteobacteria* OTUs TBZ33 and BPS-CK174 (*Oceanospirillales*), *Chromohalobacter, Idiomarina, Glaciecola* and SAR86 clade, the *Flavobacteriia* related OTUs *Formosa, Fluviicola*, NS2b, NS4 and NS7 marine group, *Defluviicoccus* and *Polaribacter* and the OTU *Burkholderia* (*Betaproteobacteria*) contributed most to the differences (Fig. 3B).

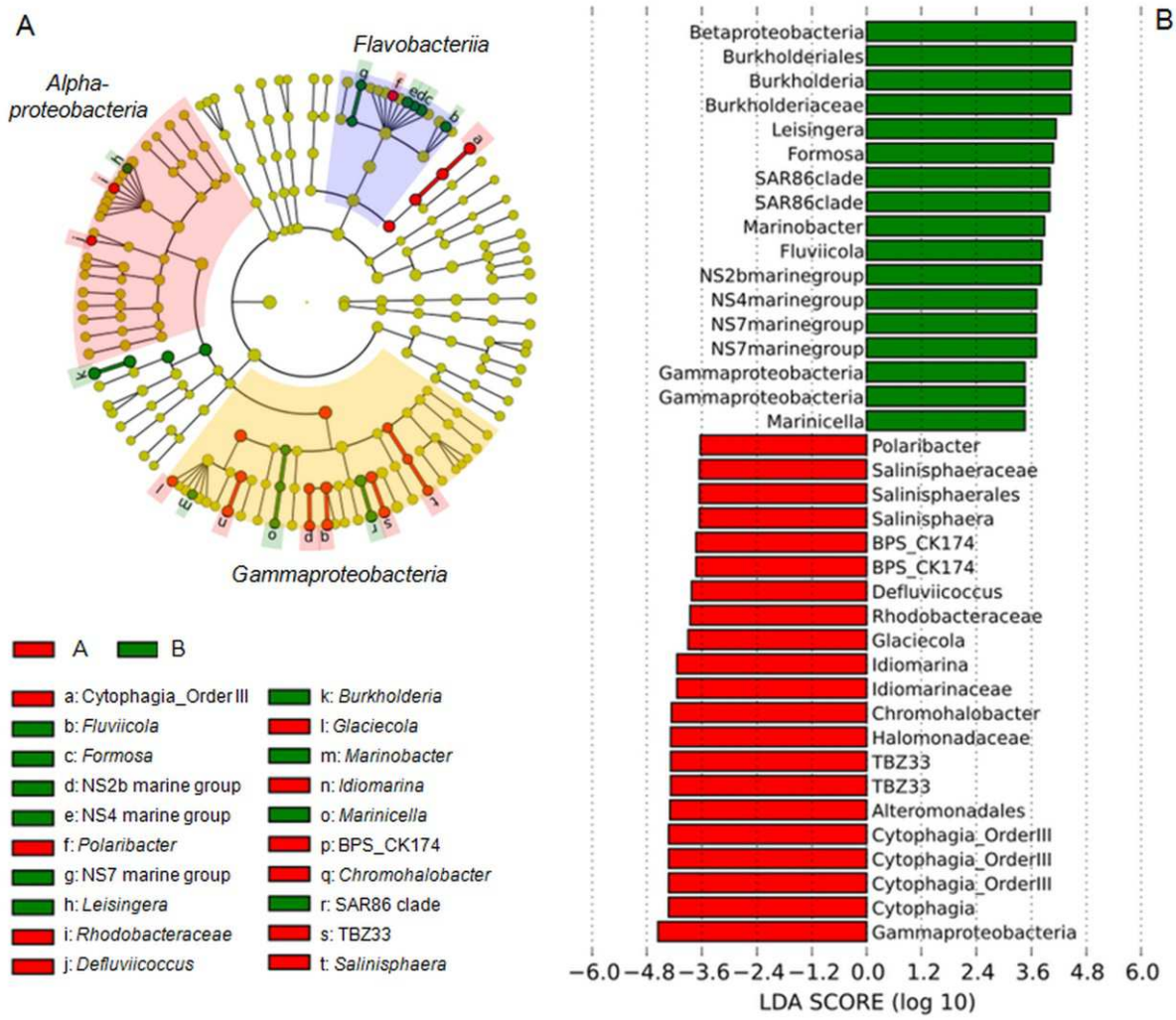


FIGURE 3: Linear discriminant effect size analysis (LEfSe) results on bacterioplankton communities. (A) Taxonomic representation of statistically consistent differences between group A and B. Differences are represented in the colour of the group, where the OTU is most abundant. (B) Histogram of linear discriminant analysis (LDA) scores computed for OTUs, differently abundant in group A and B. LDA scores can be interpreted as the degree of consistent difference in relative abundance between the two groups. The histogram thus identifies which OTUs among all those detected as statistically different explain the greatest difference between group A and B.

DOM composition, variation and relation to environmental parameters

The average solid phase extraction efficiency was 44% (± 3.3%). A total of 4039 molecular formulae were assigned, ranging between 3842 and 3947 formulae per sample (average of all samples: 3892). The identified peaks covered a mass range from 159 to 809 Da with weighted average masses per sample between 370.2 and 385.4 Da (average of all samples: 377.1 Da).

As for bacterial community composition and for environmental parameters, PCoA and non-hierarchical clustering revealed a separation of samples into two groups (Fig. 2C). DistLM analysis identified salinity and temperature as main influencing factors (Fig. 2E, Tab. S4). However, temperature exhibited a significant (p<0.05) correlation with TDN (R=0.508) and NO_2^- (R=0.645) and salinity exhibited a significant correlation with DOC (R=-0.6) and O_2 (R=-0.556), thus, there might be a shared contribution to the explanation of variation in DOM composition. Elemental ratios of assigned molecular formulae provide information on molecular characteristics, which can be visualized in van Krevelen diagrams (Kim *et al.* 2003). Van Krevelen plots of all molecules that were significantly correlated (p<0.05) with either salinity, temperature or DOC revealed the nature of these relationships in more detail (Fig. 4). Molecules that were positively correlated with salinity had higher H/C ratios and were clearly distinguished from molecules that were negatively correlated with salinity and showed lower H/C ratios (Fig. 4A). Molecules that were positively correlated with temperature formed a dense cluster in the center of the van Krevelen diagram, whereas molecules negatively correlated with temperature were more scattered, showed higher H/C ratios and covered a broader range of O/C ratios (Fig. 4B). The distribution of H/C and O/C ratios of molecules significantly correlated (p<0.5) with DOC is depicted in Figure 4C. Molecules that were positively correlated with DOC showed low H/C ratios, while negatively correlated molecules exhibited higher H/C ratios.

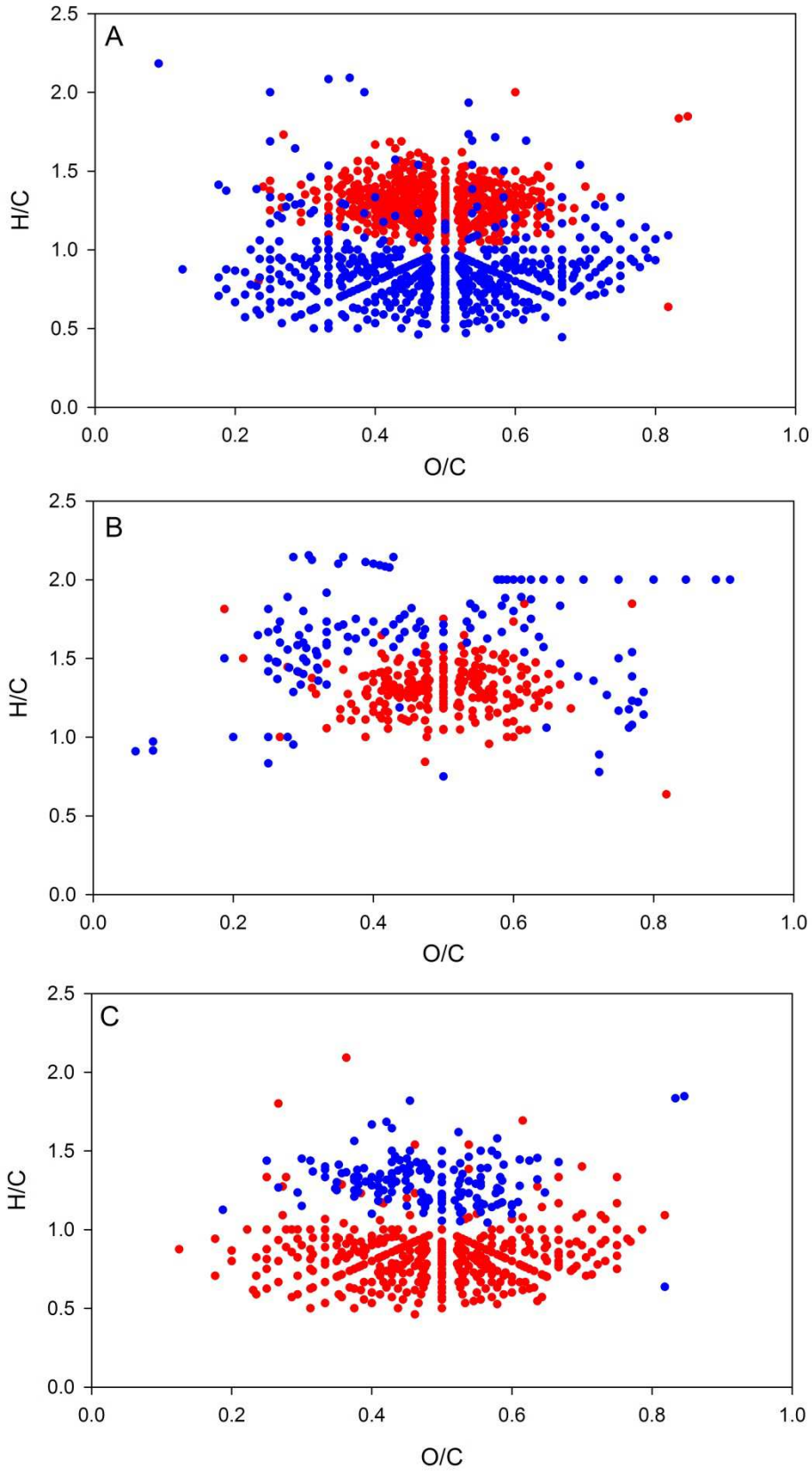


FIGURE 4: Van Krevelen plots of DOM molecular formulae with relative intensity correlating significantly (p<0.05) with salinity (A), temperature (B) and dissolved organic carbon concentration (DOC) (C). Molecular formulae showing positive correlations are depicted in red; formulae with negative correlations are shown in blue.

LEfSe analysis identified few molecules that were significantly contributing to the differences in DOM composition between the two groups (Fig. 5). Those molecules belonged mainly to the category of highly unsaturated compounds (AI_{mod} ≤0.5 and H/C <1.5) that increased in relative abundance from 80.9% in group A to 82.1% in group B and unsaturated aliphatics (2.0> H/C ≥1.5) that decreased slightly from group A (8.9%) compared to group B (8.1%).

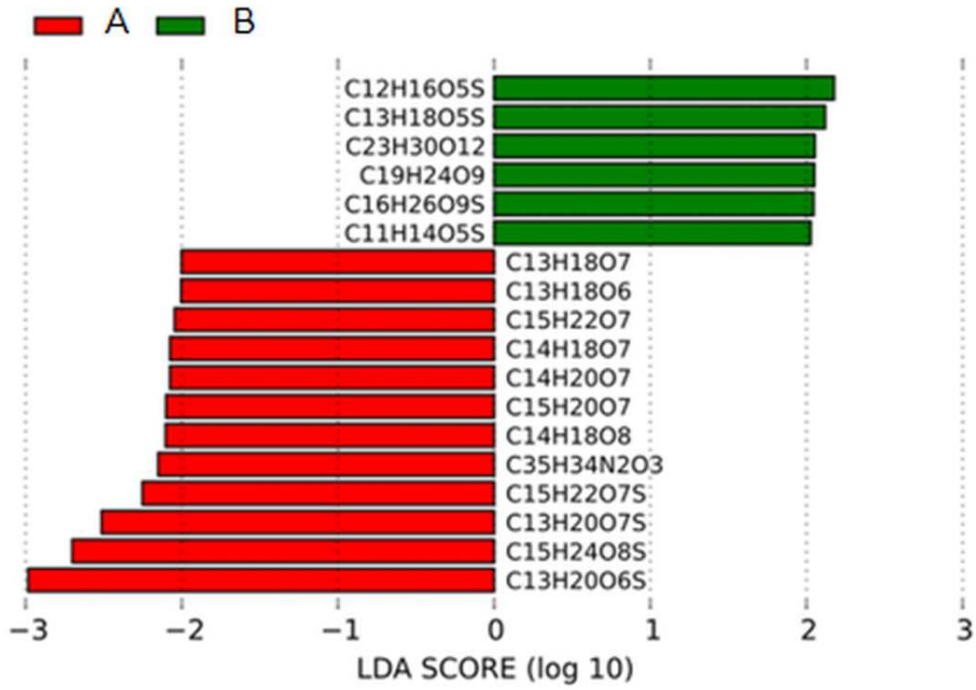


FIGURE 5: Linear discriminant effect size analysis (LEfSe) results on DOM molecular formulae. Histograms of linear discriminant analysis (LDA) scores computed for DOM molecules, differently abundant in group A and B. LDA scores can be interpreted as the degree of consistent difference in relative abundance between the two groups. The histogram thus identifies which DOM formulae among all those detected as statistically different explain the greatest difference between group A and B.

Linking bacterial communities with molecular DOM composition

Although similar patterns in bacterial community structure and molecular DOM composition have been observed via PCoA, statistical analysis (PRIMER-E; RELATE subroutine, data not shown) failed to confirm co-variation of both; i.e. among-sample relationships within the sequence data set differed from that within the DOM data set. Nevertheless, strong correlations between single OTUs and DOM molecules were detected. The majority of significant correlations (p<0.05) exhibited correlation coefficients in the range of 0.5-0.6 (Tab. 1). As the coefficient increased, the number of significant correlations decreased to 56 with R≥0.9 of which 51 were exhibited between DOM molecules and OTUs. These strong correlations were formed between only seven OTUs and 36 DOM molecules (Fig. 6 and Fig. S2). Five OTUs belonged to the *Gammaproteobacteria*, one OTU to the *Alphaproteobacteria* and one to the *Cytophagia.* Most of the DOM compounds belonged to unsaturated aliphatics (2.0> H/C ≥1.5) or saturated fatty acids (H/C≥ 2.0 or O/C ≥0.9). A group of seven distinct DOM compounds exhibited strong correlations (R≥0.9) with more than one OTU, whereas the remaining DOM compounds were correlated with either*, Defluviicoccus, Idiomarina* or *Glaciecola*. Strong correlations of *Defluviicoccus* were restricted to unsaturated aliphatics, whereas strong correlations of *Idiomarina* occurred almost exclusively with saturated fatty acids. All OTUs exhibiting strong correlations also belonged to the ones contributing most to the differences between groups A and B as revealed by LEfSe analyses (Fig. 3 and 5).

TABLE 1: Pivot-table for spearman rank order correlations between DOM molecules and environmental parameters, DOM molecules and OTUs and OTUs and environmental parameters. Correlations are sorted according to correlation strength. Numbers of correlations are given. Numbers in brackets refer to percentage of correlations on total correlations within the corresponding group. Env: environmental parameter, neg: negative correlations, pos: positive correlations.

Coefficient	< 0.5	0.5 - 0.6	0.6 - 0.7	0.7 - 0.8	0.8 - 0.9	0.9 - 1	Total
DOM and Env	789 (14.3)	2796 (50.5)	1296 (23.4)	517 (9.3)	130 (2.3)	5 (0.1)	**5533**
neg	462	1603	769	271	96	5	
pos	327	1193	527	246	34	0	
OTUs and DOM	4248 (16.6)	14910 (58.2)	5087 (19.9)	1126 (4.4)	183 (0.7)	51 (0.2)	**25605**
neg	2262	7912	2522	466	32	1	
pos	1986	6998	2565	660	151	50	
OTUs and Env	17 (10.1)	84 (50)	42 (25)	19 (11.3)	6 (3.6)	0 (0)	**168**
neg	7	34	18	12	3	0	
pos	10	50	24	7	3	0	
Sum	**5054**	**17790**	**6425**	**1662**	**319**	**56**	**31306**

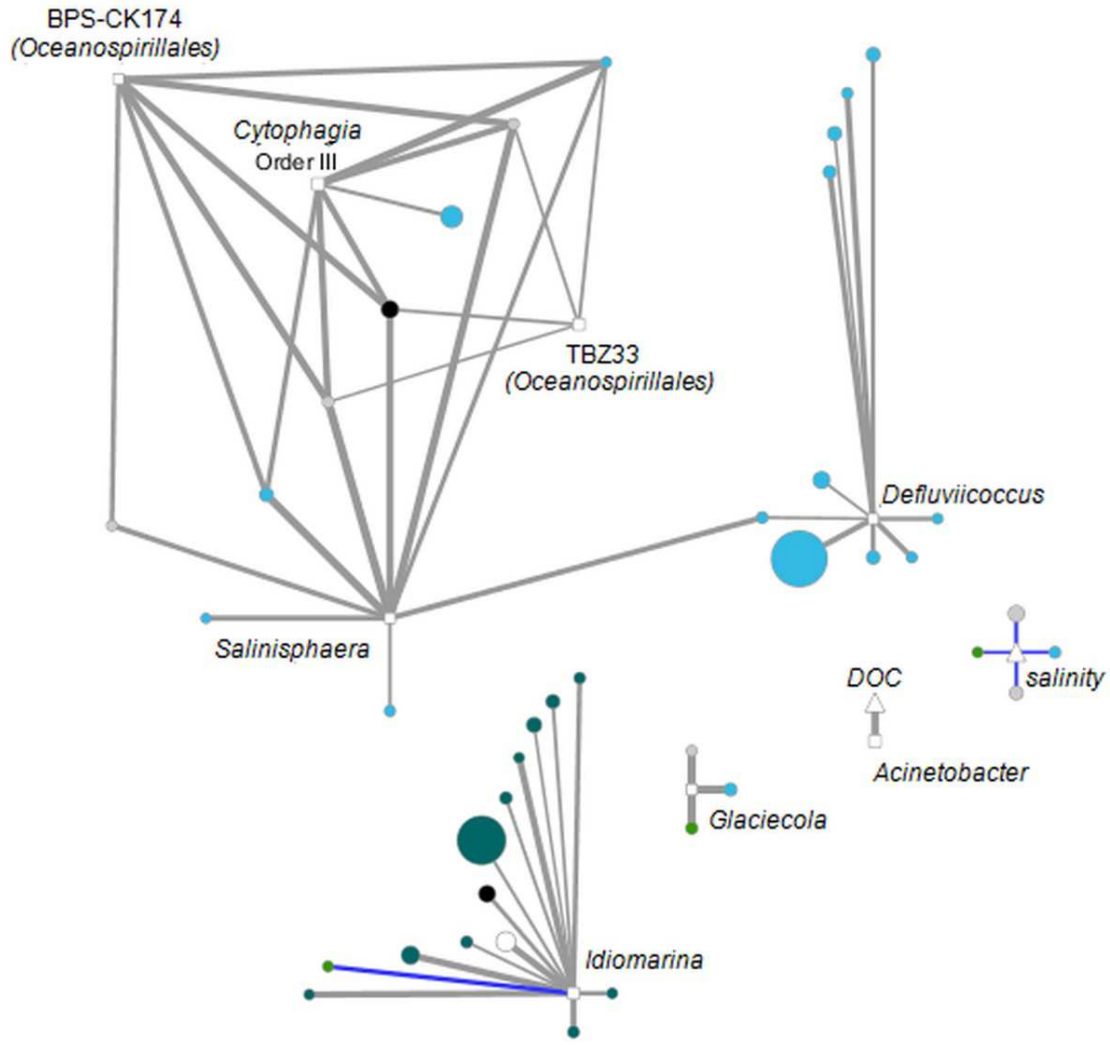


FIGURE 6: Interaction network analysis of OTUs (squares), DOM molecules (circles) and environmental parameters (triangles) that were significantly correlated (p<0.05) with R≥0.9 or R≤-0.9 (Tab. 1). Positive correlations are indicated in grey, negative correlations in blue. Line width is set proportional to correlation strength. Average OTU relative abundance and DOM molecules abundances are set proportional to node size. Nodes are coloured according to DOM category. Blue: unsaturated aliphatics, Petrol: saturated fatty acids, Grey: polyphenols, Black: peptides, Green: highly unsaturated compounds, White: unspecified.

Discussion

Impact of environmental conditions on BCC

Salinity dynamics at Helgoland Roads are mainly controlled by hydrological and meteorological forces and by the huge runoff from the rivers Elbe and Weser (Atlas & Bartha 1987). Long-term studies of oceanographic environmental parameters at Helgoland Roads reported mean annual salinities ranging between 31 and 33 (Raabe & Wiltshire 2009). Transport of central North Sea water towards Helgoland Roads results in high salinities, whereas coastal water influx is related to lower salinities (Wiltshire *et al.* 2010). Low salinity events at Helgoland Roads presumably occur during winter months, especially in February, when the Elbe discharge is particularly high (Raabe & Wiltshire 2009). During this study conducted in August 2012, the recorded salinity values exhibited a salinity shift of ~1.5 within four days which is exceptional during usually more stable hydrographic conditions in summer and points to a strong short-term influence of coastal water masses.

The intermittent change of water masses during our study is confirmed by the results of tracer particle backtracking. Trajectories were simulated with PELETS-2D (Callies *et al.* 2011) based on pre-calculated near surface current velocity fields from the hydrodynamic model BSHcmod (Dick *et al.* 2001) operated by the Federal Maritime and Hydrographic Agency of Germany (Bundesamt für Seeschifffahrt und Hydrographie, BSH). Model results (Fig. 7) help delineate regions of origin by analyzing the percentages of particle trajectories that crossed certain grid cells over the previous three weeks. We organized grid cells in a cobweb like structure centered at Helgoland Roads in order to take account of uncertainty increasing with distance. Figure 7 clearly documents an event with strong advection from the inner German Bight (near the Elbe estuary) towards Helgoland during the period around the 12th of August (Fig. 7, right bottom panel) when salinity values at Helgoland were found to drop substantially. According to model simulations, this inshore origin of water masses did not exist (or was at least much less pronounced) at both the beginning and the end of the sampling period.

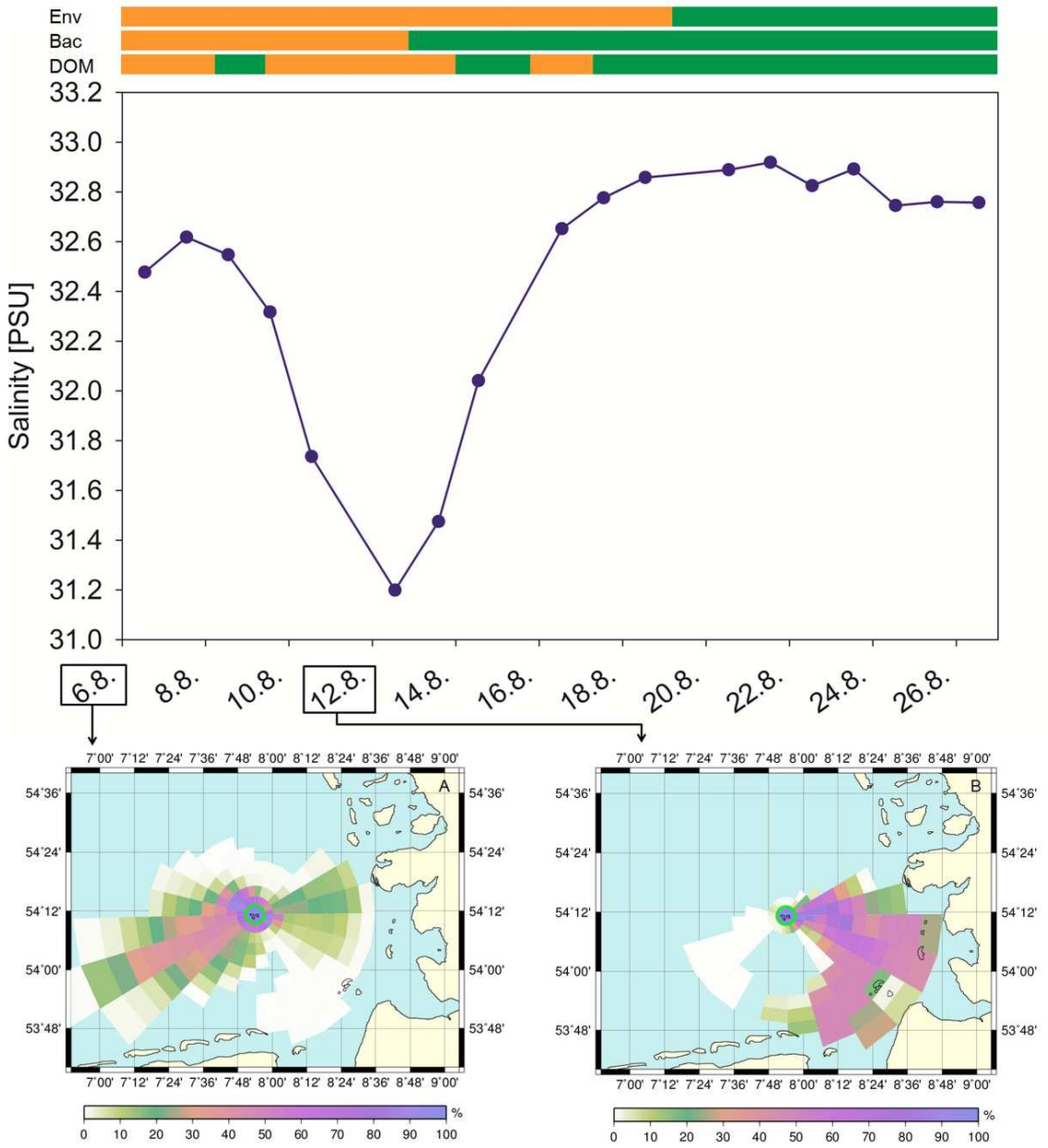


FIGURE 7: Histories of water bodies observed at Helgoland Roads. Based on pre-calculated near surface current velocities from the hydrodynamic model BSHcmod, 500 passive tracer particles were tracked backward in time. Referring to a cobweb like grid structure, the color scale reflects the percentage of particle trajectories that touched a given grid cell at any time within the past three weeks. Bars on top depict group formation for environmental parameters (Env), bacterial communities (Bac) and DOM. Orange: group A, Green: group B.

Statistical analysis on 16S rRNA tag data suggested a separation of samples into two groups that is particularly influenced by temperature, TDN, O_2, $PO_4^{3-}$and SiO_2. Multicollinearity of parameters describing ecological conditions might lead to biased interpretation of linear regression models (Graham 2003). Correlation analysis confirmed multicollinearity of salinity with O_2 and DOC, which hints at a shared contribution of these

parameters. In conjunction with the above mentioned coastal water inflow this leads to the assumption that salinity can be interpreted as a proxy for different water bodies with differing environmental conditions.

TDN is composed of dissolved inorganic nitrogen (DIN) and dissolved organic nitrogen (DON). DON comprises a large pool of fixed nitrogen in most aquatic systems, accounting for as much as 40-70% of the TDN pool in surface seawater (Bronk 2002). It consists of labile, rapidly overturning proteins, amino polysaccharides, urea and nucleic acids, but also includes more refractory compounds like humic acids (e.g. Bronk *et al.* 2007), most of which derive from primary producers but also from bacterial cell wall material (McCarthy *et al.* 1998). TDN concentrations measured during this study increased towards the end of the sampling period and reached highest concentrations shortly after the Chl *a* maximum. Thus, the increase in TDN concentration might reflect the permanent release of DOM by phytoplankton during its growth phase and additional release due to grazing or viral lysis that may affect the termination of the bloom (Beare *et al.* 2002; Wiltshire *et al.* 2010). However, it has to be noticed that DON concentration was not measured thus, this interpretation is speculative and needs to be evidenced.

In general, the bacterial community observed in our study was dominated by *Flavobacteriia, Alpha-* and *Gammaproteobacteria*. These classes have been consistently found to dominate bloom-associated bacterial communities as reviewed by Buchan et al. (2014). Also, the most abundant OTUs found during this study are common members of the North Sea bacterial community during phytoplankton blooms (e.g. Giebel *et al.* 2011; Teeling *et al.* 2012; Wemheuer *et al.* 2014). Comparison of bacterial communities of group A, with low average TDN concentration (14.61 µM), and group B ,exhibiting higher TDN concentrations (19.98 µM), revealed that group B is characterized by higher relative abundance of *Flavobacteriia* (*Formosa, Fluviicola*, NS2b, NS4 and NS7 marine group) which are well known to be active in biopolymer degradation and reacting to phytoplankton blooms (e.g. Teeling *et al.* 2012; Buchan *et al.* 2014; Lucas *et al.* 2015b). *Formosa* for instance has been found to be among the first taxa responding to a phytoplankton bloom and dominating the bacterial community at Helgoland Roads (Teeling *et al.* 2012). On the other hand, a strong increase of *Betaproteobacteria* relative abundances from 3.2% in group A, to 10.9% in group B was observed. This increase was mainly due to higher relative abundances of *Burkholderia*, which also contributed strongly to the differences between group A and B, and exhibited distinct short-term peaks during the onset of the summer bloom. The genus *Burkholderia* comprises more than 60 species

which are mainly associated with plants and saprophytes but have also been found in water and include human, animal and plant pathogens (Compant *et al.* 2008; Estrada-de los Santos *et al.* 2013). *Burkholderia* has not been identified as typical marine genus and information about its occurrence in oceanic environments is scarce. But it is known, that complex estuarine bacterial communities can include groups recruited from different sources such as atmospheric deposition (Jones *et al.* 2008), freshwater discharge and unique estuarine habitats (Crump *et al.* 2004) or sediments (Comte *et al.* 2014), and might be related to runoff from land. Since it is known that *Betaproteobacteria* are prevalent in freshwater environments (e.g. Bouvier & del Giorgio 2002; Cottrell & David 2003), decrease with increasing salinity (Kirchman *et al.* 2005) and thus, are rarely found in marine environments, we assume that *Burkholderia* has been passively transported with coastal waters. We propose that its increasing relative abundances reflect the short-term impact of coastal water inflows on the bacterial community composition at Helgoland Roads.

Impact of environmental conditions on molecular DOM composition

Variability in the DOM composition was mainly driven by salinity and temperature as revealed by DistLM (Fig. 2E). The molecules that were positively correlated with salinity exhibited higher H/C ratios and were clearly separated from molecules that were negatively correlated with salinity and showed lower H/C values (Fig. 4A). In general, marine DOM has higher H/C ratios, is more aliphatic and contains higher proportions of carbohydrates, amino acids and lipids, whereas terrestrial DOM is more aromatic, contains carboxyl and hydroxyl functionalities and has lower H/C ratios (Sleighter & Hatcher 2008; Medeiros *et al.* 2015; Seidel *et al.* 2015). Similar observations by Kim *et al.* (2003) and Koch *et al.* (2005) support the assumption that the molecules positively correlated with salinity are associated with marine DOM and those negatively correlated are associated with terrigenous DOM. This interpretation is also supported by our findings that on average unsaturated aliphatics ($2.0 > H/C \geq 1.5$) were most abundant during the coastal water inflow. Thus, we conclude that differences between the observed groups in this study can be partly explained by different water masses and thus, origins of DOM.

Furthermore, we found that molecules negatively correlated with temperature had higher H/C ratios than positively correlated molecules (Fig. 4B). Higher H/C ratios indicate higher saturation which is characteristic for compounds that are rapidly degradable, e.g. fatty acids and proteinaceuos material. As temperature increased towards the end of the sampling period, molecules with high H/C ratio decreased. This observation could be

explained by a scenario where microbial activity had increased with rising temperatures and as response to enhanced organic matter supply released by phytoplankton. Due to enhanced metabolism, the microbial community may have consumed more DOM (Pomeroy & Wiebe 2001), which could have preferentially diminished the pool of labile DOM and thus, diminished the amount of molecules with high H/C ratio. The notion of enhanced microbial activity is also supported by the H/C ratios of molecules correlated with DOC. The simultaneous increase of compounds with low H/C ratios and DOC concentration (reflected in the positive correlation of these compounds with DOC) support the scenario of preferential consumption of rapidly degradable compounds with high H/C ratios leading to an increasing relative abundance of molecules with low H/C ratios.

Relation between bacterial community and molecular DOM composition

Linkage of bacterial relative abundances with DOM data revealed evidence for dependency of specific OTUs on particular DOM molecules. Especially *Gammaproteobacteria* showed strong positive correlations ($R \geq 0.9$) with unsaturated aliphatics and saturated fatty acids (Fig. 6), all of which showed decreased relative abundances in group B compared to group A. *Gammaproteobacteria* are known to be typical marine bacteria, thus, one would expect increasing relative abundances in group B, which exhibits higher salinity than group A, which was strongly influenced by a coastal freshwater inflow. A possible explanation for the observed predominance of *Bacteroidetes* in group B might be the onset of the summer phytoplankton bloom that occurred in group B and might have supported enhanced growth of *Bacteroidetes* by providing complex organic compounds. Thus, *Gammaproteobacteria* relative abundances might have decreased due to increasing *Bacteroidetes* abundances.

Network analysis revealed few DOM compounds that were highly correlated ($R \geq 0.9$) with different bacteria taxa and thus, seem to serve as a general substrate. On the other hand we observed strong correlations of *Defluviicoccus* (*Alphaproteobacteria*) with unsaturated aliphatics as the only substrate category (Fig. 6), which might indicate that *Defluviicoccus* specialized on selected DOM compounds that are not as intensively consumed by other taxa and which might be important for defining its ecological niche. *Defluviicoccus* spp. is typically found in wastewater treatment plants (e.g. Nobu *et al.* 2014), and capable of taking up a narrow range of substrates such as acetate, propionate, pyruvate and glucose (Burow *et al.* 2007) which supports our findings of strong correlations with unsaturated aliphatics. Furthermore, the simultaneously decreasing relative abundances of unsaturated aliphatics and *Defluviicoccus* in group B (0.1%) compared to group A (0.3%) suggest that

this OTU is more frequently found in coastal waters and was passively transported to Helgoland waters via coastal water inflow. Another example for specialization on specific substrate classes is *Idiomarina* (*Gammaproteobacteria*), which is almost exclusively strongly correlated with specific saturated fatty acids (Fig. 6). *Idiomarina* spp. have been isolated from various marine environments such as deep-sea waters, hydrothermal vents, sediments, and reef building corals (e.g. Ivanova *et al.* 2000; Donachie *et al.* 2003; Chen *et al.* 2012; Zhang *et al.* 2012), but also from estuarine environments like Baltic Sea surface water (Brettar *et al.* 2003). *Idiomarina* relative abundances were decreasing from group A (3.1%) to group B (0.3%) during our study as did relative abundances of saturated fatty acids (group A: 0.6%, group B: 0.2%). This again suggests the assumption of a short-term influence of coastal water inflows on the bacterial community at Helgoland Roads. The notion of specific bacterial taxa specializing on selected DOM molecules is supported by previous studies that also suggested coordinated resource partitioning by bacterial specialists leading to a defined temporal succession of bacteria taxa (e.g. McCarren *et al.* 2010; Teeling *et al.* 2012). However, these studies analyzed the transcriptional responses of microbial communities to high-molecular-weight DOM amendment or enhanced substrate supply by phytoplankton blooms. Here, we demonstrate the possibility to link single bacterial taxa to specific DOM formulae rather than just molecule categories. Even though FT-ICR-MS has, as any analytical technique, a defined analytical window, it is an important step to further unravel the specific microorganisms and metabolic pathways responsible for the degradation and transformation of DOM in the oceans.

Although we were able to relate single OTUs with specific DOM molecules, a direct general relationship between bacterial community structure and DOM composition could not be demonstrated. One possible explanation might be that freshly produced labile DOM that is accessible for microorganisms is rapidly turned over by the bacterial community as shown in several studies (Kirchman *et al.* 1991; Amon & Benner 1996; Weiss & Simon 1999). Thus, the pool of labile DOM compounds that could have a significant influence on bacterial community structure might not be detectable with the methods applied in this study. An instantaneous degradation of introduced fresh DOM by bacterioplankton is also proposed for arctic fjords (Svalbard), which results in a predominance of the prevailing semi-refractory and refractory DOM pool Osterholz *et al.* (2014).

Furthermore, methodological limitations could have led to lacking evidence of a relation between the bacterial community structure and DOM composition. A previous study by Flerus *et al.* (2012) suggested that colloidal material and low molecular weight DOM

(<250 Da) can be lost during solid phase extraction (SPE) as used in this study. Thus, labile DOM that might have an influence on bacterial community structure, escapes the analytical window. The low extraction efficiencies observed during our study might indicate that a substantial fraction of DOM was not extracted. Although the efficiencies were in the range described for marine samples (Dittmar *et al.* 2008), they were considerably lower than in more recent studies (Rossel *et al.* 2013; Osterholz *et al.* 2014). Despite these methodological limitations inherent to any analytical method, we identified significant variations in DOM composition and successfully linked them to environmental conditions and bacterial community composition.

To our knowledge this is the first time that dynamics of bacterial community composition and molecular DOM composition have been documented on high temporal and analytical resolution. We conclude that the bacterial community is highly influenced by freshly produced, labile DOM pulses as derived from phytoplankton blooms. Rapid transformation of labile DOM might lead to selective relative enrichment of more refractory DOM and thus, hamper the detection of interdependencies between bacterial community structure and DOM composition. High resolution techniques like 16S rRNA tag sequencing and FT-ICR-MS provide substantial information on substrate generalists and specialists and may help to predict changes in bacterial community composition. To further unravel the relationship between bacteria and molecular DOM composition it has to be considered that metabolic capabilities are not restricted to specific phylogenetic groups. Thus, for future analyses we suggest combining FT-ICR-MS analyses of DOM with functional approaches of bacterial communities rather than phylogenetic description.

Acknowledgements

We would like to thank Kristine Carstens, Matthias Friebe, Ina Ulber and Katrin Klaproth for their technical assistance. We are grateful for calculation of correlations between DOM, OTUs and environmental factors with R by Stephan Frickenhaus. We gratefully acknowledge the provision of BSHcmod current velocity fields by the Federal Maritime and Hydrographic Agency of Germany (Bundesamt für Seeschifffahrt und Hydrographie, BSH) in Hamburg.

Supplementary material

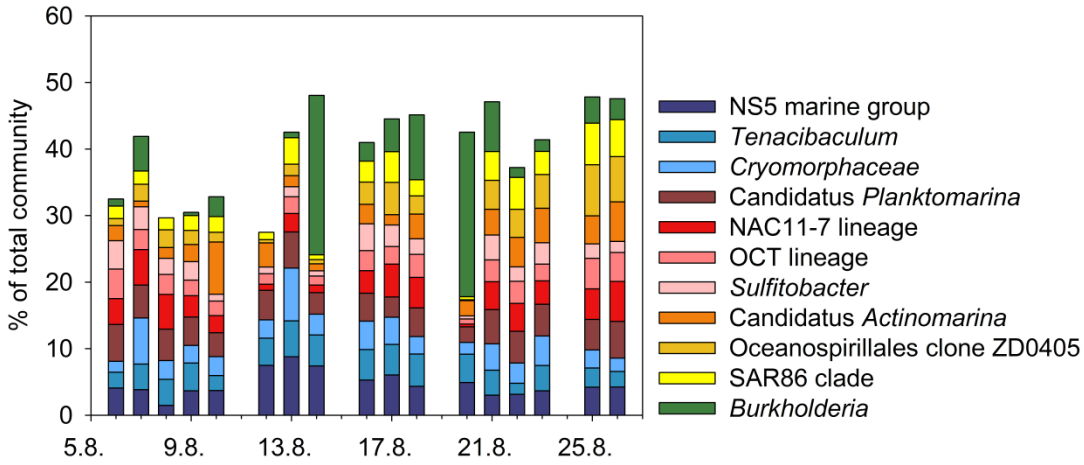


FIGURE S1: Contribution of the 10 most abundant OTUs to the total community given in percentage. Additionally *Burkholderia* abundances are depicted to visualize strong short-term peaks in their relative abundances.

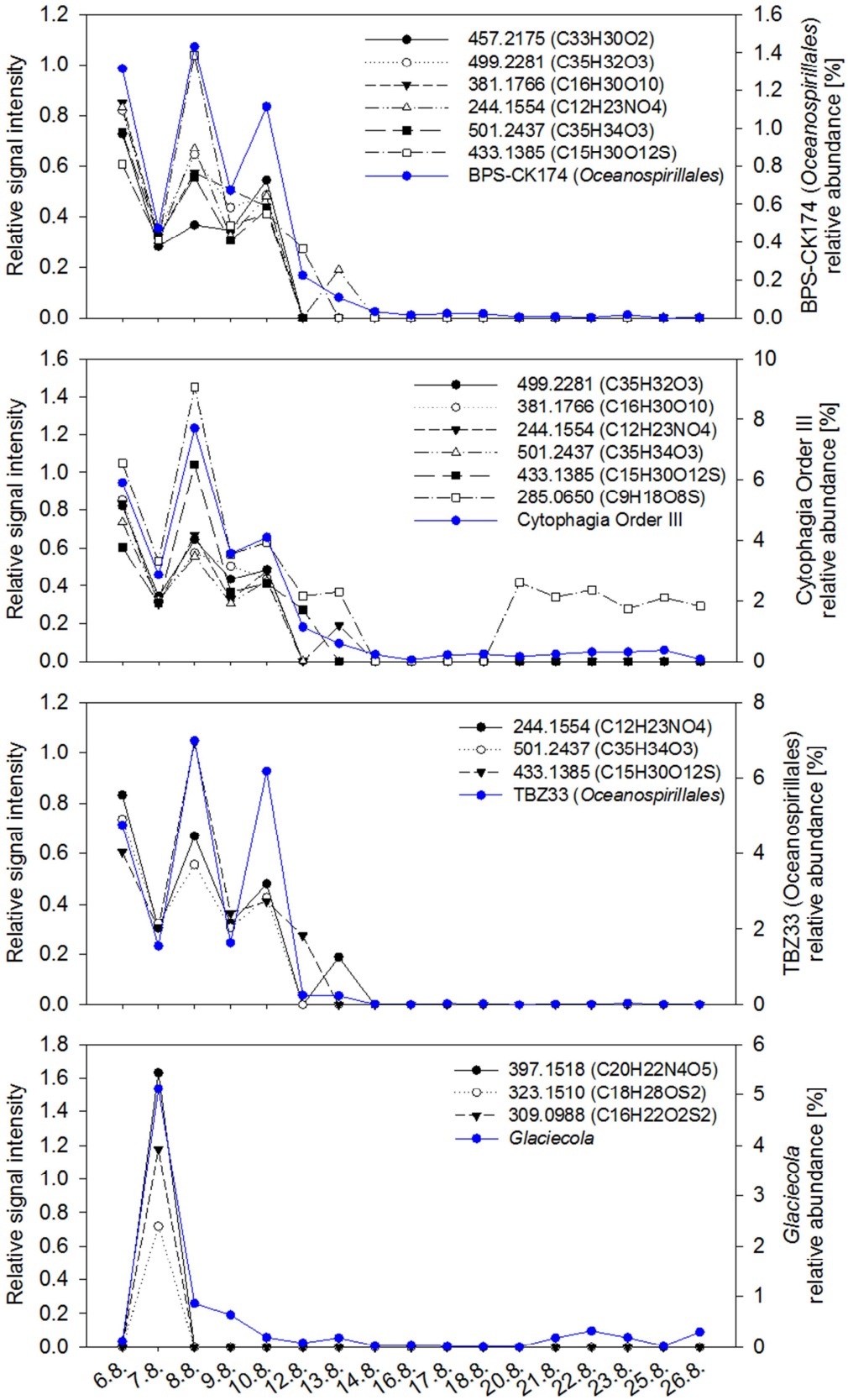


Figure S2-1: Relationship of specific bacterial OTUs with particular DOM molecules. Relative signal intensities and relative abundances of DOM molecules and OTUs that were highly correlated (R>0.9) are depicted for the whole sampling period. The curves of the OTUs and the correlated DOM molecules exhibit similar behaviour during the time course. Thus, artificial correlations that might occur due to the size of the data set can be excluded.

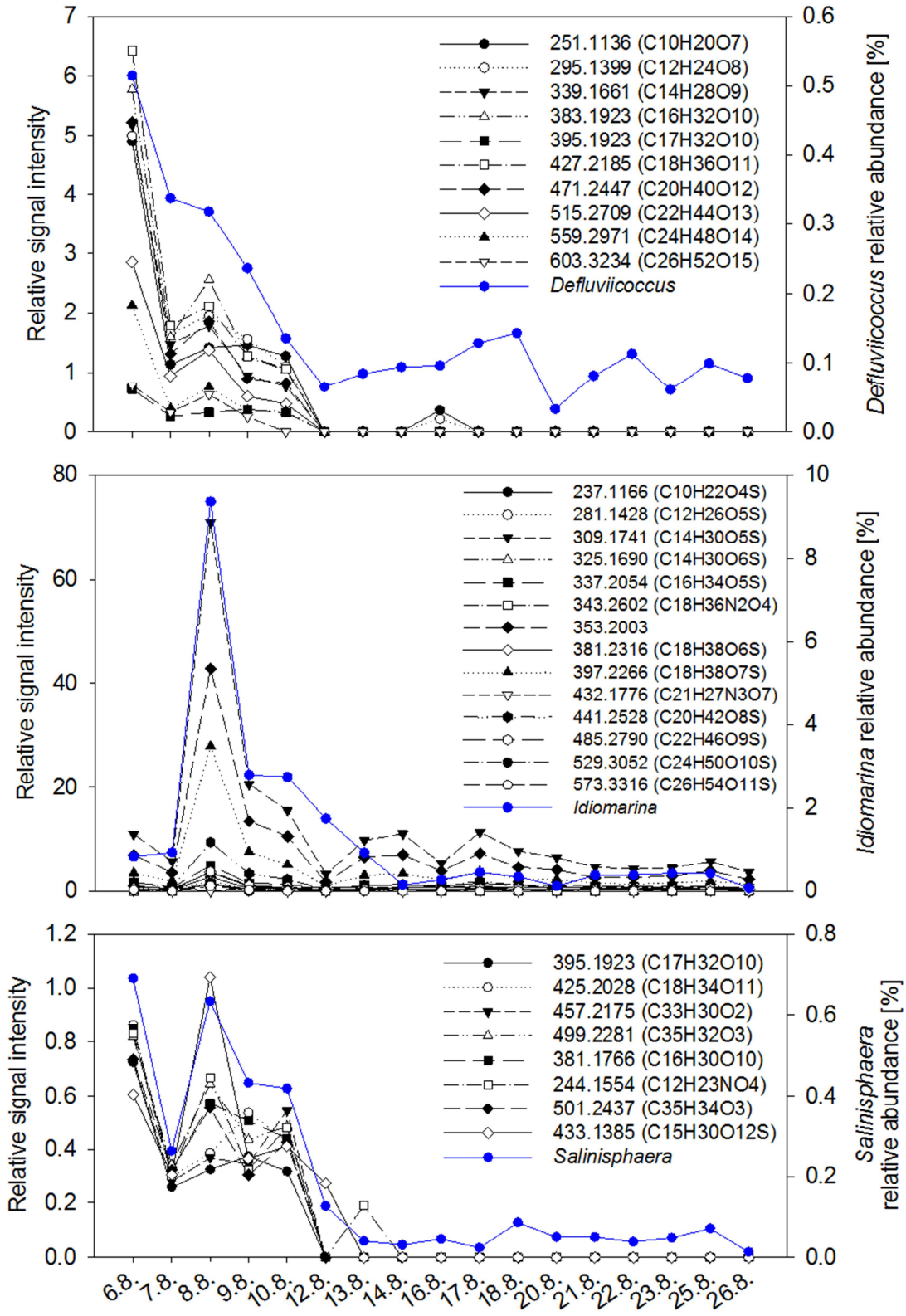


FIGURE S2-2: Relationship of specific bacterial OTUs with particular DOM molecules. Relative signal intensities and relative abundances of DOM molecules and OTUs that were highly correlated (R>0.9) are depicted for the whole sampling period. The curves of the OTUs and the correlated DOM molecules exhibit similar behaviour during the time course. Thus, artificial correlations that might occur due to the size of the data set can be excluded.

TABLE S1: Environmental parameters measured during August 6 till August 26 2012. DOC: dissolved organic carbon; TDN: total dissolved nitrogen; depth: water depth; T: temperature; S: salinity; O_2: oxygen; SiO_2: silicate; PO_4^{3-}: phosphate; NO_2^-: nitrite; NO_3^-: nitrate; Chl *a*: Chlorophyll *a*; CO_2: carbon dioxid, SPE extr. eff.: Solid phase extraction efficiency.

Sample	DOC [µM]	TDN [µM]	Depth [m]	T [°C]	S	O_2 [mg l^{-1}]	turbidity	ph	SiO_2 [µM]	PO_4^{3-} [µM]	NO_2^- [µM]	NO_3^- [µM]	Chl *a* [µg l^{-1}]	CO_2 [µg l^{-1}]	SPE extr. eff.
6.8.	114.13	13.38	3.93	17.3	32.5	7.81	0.01	8.28	4.29	0.00	0.31	2.03	0.57	366.6	51.50
7.8.	122.27	12.36	4.09	17.1	32.6	8.20	0.03	8.30	4.64	0.00	0.26	1.56	0.75	342.9	43.07
8.8.	119.7	14.52	3.68	17.2	32.5	8.38	0.03	8.31	5.21	0.00	0.3	1.75	0.75	337.6	46.57
9.8.	117.02	13.62	3.34	17.5	32.3	9.19	0.02	8.36	6.23	0.70	0.28	1.57	0.68	285.3	45.22
10.8.	130.41	16.53	2.93	17.4	31.7	8.54	0.03	8.34	4.56	0.00	0.29	1.59	0.73	314.4	48.78
11.8.	170.63	48.63	2.36	17.5	32.0	8.30	0.02	8.32	4.86	0.42	0.38	1.78	0.66	329.6	42.20
12.8.	131.57	15.47	1.79	18.1	31.2	8.54	0.02	8.37	4.14	0.56	0.36	1.67	0.65	302.1	43.57
13.8.	150.44	17.27	1.55	18.0	31.5	8.32	0.02	8.38	0.00	0.60	0.38	1.19	0.68	304.1	39.78
14.8.	130.73	11.81	1.65	17.8	32.0	7.29	0.02	8.31	2.46	0.00	0.36	1.51	0.69	365.2	48.32
16.8.	121.35	16.74	2.52	17.6	32.7	7.46	0.02	8.32	3.52	0.00	0.28	1.51	0.67	364.5	45.55
17.8.	120.87	16.11	2.89	17.8	32.8	7.38	0.02	8.31	1.21	0.38	0.28	1.46	0.67	377.0	41.59
18.8.	108.67	12.93	3.17	17.9	32.9	7.26	0.02	8.30	1.84	0.40	0.34	1.55	0.66	382.2	45.01
20.8.	116.08	15.93	3.93	18.5	32.9	8.19	0.02	8.17	1.28	0.00	0.34	1.23	0.66	306.5	42.08
21.8.	114.22	15.39	4.02	17.9	32.9	7.34	0.02	8.11	2.37	0.36	0.3	2.16	0.68	359.2	43.87
22.8.	122.9	18.68	4.2	18.1	32.8	7.99	0.04	8.16	3.47	0.38	0.29	1.16	0.84	317.5	42.04
23.8.	119.77	16.99	3.94	18.0	32.9	7.99	0.08	8.16	2.92	0.58	0.36	1.26	1.08	310.1	40.57
24.8.	261.04	22.72	3.49	18.1	32.7	8.17	0.03	8.17	4.60	0.40	0.37	1.50	0.71	305.0	19.86
25.8.	123.62	15.60	3.18	18.2	32.8	8.02	0.02	8.16	4.72	0.35	0.42	2.39	0.67	319.9	42.03
26.8.	121.87	25.29	2.62	18.1	32.8	8.15	0.04	8.15	4.58	0.00	0.42	2.01	0.83	331.0	39.71

TABLE S2: Spearman rank order correlations of environmental parameters. Significant correlations (p<0.05) are indicated in bold.

	DOC [µM]	TDN [µM]	Depth [m]	T [°C]	S	O_2 [mg l^{-1}]	turbidity	ph	SiO_2 [µM]	PO_4^{3-} [µM]	NO_2^- [µM]	NO_3^- [µM]	Chl *a* [µg l^{-1}]
TDN [µM]	0.309												
depth [m]	**-0.517**	-0.118											
T [°C]	0.209	**0.508**	-0.068										
S	**-0.600**	0.169	**0.597**	0.362									
O_2 [mg l^{-1}]	0.350	0.162	-0.039	-0.087	**-0.556**								
turbidity	0.219	0.467	0.316	0.066	0.211	0.250							
ph	0.409	-0.139	**-0.620**	-0.381	**-0.799**	0.460	-0.297						
SiO_2 [µM]	0.032	-0.199	0.206	-0.367	-0.314	**0.536**	0.255	0.007					
PO_4^{3-} [µM]	0.048	0.163	-0.063	0.292	0.030	0.118	-0.052	0.232	-0.251				
NO_2^- [µM]	0.270	0.220	-0.379	**0.645**	-0.037	-0.048	-0.018	-0.224	-0.156	0.119			
NO_3^- [µM]	-0.184	-0.338	0.047	-0.214	-0.116	0.099	-0.203	-0.231	**0.607**	-0.285	0.159		
Chl *a* [µg l^{-1}]	0.215	0.337	0.310	-0.076	0.158	0.136	**0.890**	-0.298	0.271	-0.072	-0.069	-0.158	
CO_2 [µg l^{-1}]	-0.385	-0.382	0.013	-0.382	0.289	**-0.807**	-0.256	-0.269	-0.208	-0.421	-0.211	0.166	-0.175

TABLE S3: Tests of significance of group formation. PERMANOVA pair-wise comparison of groups of environmental parameters, bacterial community composition and DOM molecules. ANOVA comparison of environmental parameters of group A and B. Significant results (p<0.05) are highlighted in bold. Res: Residual, df: degrees of freedom, SS: sums of squares, MS: mean squares, perms: number of unique permutations per comparison.

PERMANOVA

	Comparison	df	SS	MS	Pseudo-F	p	perms
Evs	A vs B	1	46.377	46.377	4.1276	**0.001**	964
	Res	15	168.54	11.236			
	Total	16	214.91				
Bacteria	A vs B	1	0.62761	0.62761	5.6982	**0.002**	961
	Res	15	1.6521	0.11014			
	Total	16	2.2797				
DOM	A vs B	1	33.928	33.928	3.8429	**0.001**	968
	Res	15	132.43	8.8286			
	Total	16	166.36				

ANOVA

Parameter		df	SS	MS	F	p
TDN	Intercept	1	4124.321	4124.321	573.2954	**0.000000**
	group	1	44.041	44.041	6.1219	**0.025782**
	error	15	107.911	7.194		
depth	Intercept	1	164.8180	164.8180	247.5661	**0.000000**
	group	1	2.3685	2.3685	3.5576	0.078790
	error	15	9.9863	0.6658		
temperature	Intercept	1	4956.129	4956.129	58630.48	**0.000000**
	group	1	1.080	1.080	12.78	**0.002767**
	error	15	1.268	0.085		
salinity	Intercept	1	16446.08	16446.08	78933.82	**0.000000**
	group	1	1.38	1.38	6.60	**0.021357**
	error	15	3.13	0.21		
O_2	Intercept	1	991.4368	991.4368	3293.673	**0.000000**
	group	1	0.0294	0.0294	0.098	0.759084
	error	15	4.5152	0.3010		
SiO_2	Intercept	1	173.6063	173.6063	59.33325	**0.000001**
	group	1	0.2244	0.2244	0.07668	0.785629
	error	15	43.8893	2.9260		
PO_4^{3-}	Intercept	1	1.043070	1.043070	14.19164	**0.001864**
	group	1	0.005705	0.005705	0.07762	0.784355
	error	15	1.102483	0.073499		
NO_2^-	Intercept	1	1.730985	1.730985	802.1696	**0.000000**
	group	1	0.006938	0.006938	3.2151	0.093148
	error	15	0.032368	0.002158		
NO_3^-	Intercept	1	41.83353	41.83353	329.0224	**0.000000**
	group	1	0.05661	0.05661	0.4453	0.514724
	error	15	1.90717	0.12714		
Chl *a*	Intercept	1	8.426253	8.426253	827.5221	**0.000000**
	group	1	0.048503	0.048503	4.7633	**0.045384**
	error	15	0.152738	0.010183		
DOC	Intercept	1	231197.8	231197.8	2598.226	**0.000000**
	group	1	80.2	80.2	0.901	0.357596
	error	15	1334.7	89.0		

TABLE S4: Distance-based linear models of the bacterial community structure based on Hellingers distance and DOM composition based on Bray-Curtis similarity. Variables contributing significantly ($p<0.05$) to the model are indicated in bold. Prop: Proportion of variability explained by the respective variable; Adj R^2: adjusted R^2; Cumul: Cumulative proportion of explained variability.

	Marginal test				**Sequential test**					
	Variable	**Pseudo-F**	**P**	**Prop.**	**Variable**	**Adj R^2**	**Pseudo-F**	**P**	**Prop.**	**Cumul.**
Bacteria	DOC	0.77736	0.571	0.04927	**Temperature**	0.16715	4.211	**0.001**	0.2192	0.2192
	TDN	1.5771	0.158	0.095135	**TDN**	0.24541	2.5559	**0.027**	0.12054	0.33974
	depth	1.0742	0.359	0.066829	**O_2**	0.32617	2.6778	**0.003**	0.11278	0.45251
	Temperature	4.211	**0.002**	0.2192	**PO_4^{3-}**	0.40081	2.6194	**0.006**	0.098095	0.55061
	Salinity	1.8696	0.084	0.11083	**SiO_2**	0.45396	2.168	**0.025**	0.073989	0.6246
	O_2	2.7004	**0.023**	0.15256	Salinity	0.48977	1.7721	0.073	0.056511	0.68111
	SiO_2	2.5982	**0.027**	0.14764	DOC	0.5405	2.1039	0.041	0.060421	0.74153
	PO_4^{3-}	0.97841	0.432	0.061233	NO_2^-	0.54699	1.1289	0.326	0.031963	0.77349
	NO_2^-	1.0389	0.382	0.064777	depth	0.58531	1.7394	0.101	0.045081	0.81857
	NO_3^-	1.0129	0.392	0.063253						
	Chl *a*	1.1226	0.323	0.069628						
DOM	DOC	1.6038	0.087	0.096595	**Salinity**	0.1114	3.0059	**0.003**	0.16694	0.16694
	TDN	0.76048	0.717	0.048252	**Temperature**	0.16698	2.0007	**0.048**	0.10416	0.2711
	depth	1.979	**0.036**	0.11656	O_2	0.20057	1.5883	0.115	0.07936	0.35046
	Temperature	1.9142	0.054	0.11317	Chl *a*	0.21755	1.282	0.267	0.062695	0.41316
	Salinity	3.0059	**0.001**	0.16694						
	O_2	1.6155	0.103	0.097229						
	SiO_2	1.3137	0.242	0.080527						
	PO_4^{3-}	0.64816	0.79	0.041421						
	NO_2^-	0.79863	0.633	0.050551						
	NO_3^-	0.76982	0.664	0.048816						
	Chl *a*	1.6885	0.064	0.10118						

CHAPTER III

Spatiotemporal variation of the bacterioplankton community in the German Bight: From estuarine to offshore regions

Judith Lucas*, Antje Wichels and Gunnar Gerdts

Alfred-Wegener-Institute Helmholtz-Center for Polar and Marine Research
Biological Station Helgoland, Kurpromenade 201, 27498 Helgoland, Germany
Microbial ecology group

*Judith Lucas, Alfred-Wegener-Institute Helmholtz-Center for Polar and Marine Research, Biological Station Helgoland, Kurpromenade 201, 27498 Helgoland, Germany, Phone: +49 4725 819 3200, Fax: +49 4725 819 3283, email: judith.lucas@awi.de

Abstract

Marine microbial biogeography has been studied intensively; however few studies address community variation across temporal and spatial scales simultaneously so far. Here we present a yearlong study investigating the dynamics of free-living and particle-attached bacterioplankton community across a 100 km transect in the German Bight reaching from the Elbe estuary towards the open North Sea. Community composition was assessed using automated ribosomal intergenic spacer analysis (ARISA) and linked to environmental parameters applying multivariate statistical techniques. Results suggest that the spatial variation of the bacterioplankton community is defined by hydrographic current conditions, which separate the inner German Bight from the open North Sea and lead to pronounced differences in the coastal and oceanic bacterioplankton community. However this spatial variation is overwhelmed by a strong temporal variation which is triggered by temperature as the main driving force throughout the whole transect. Variation in the free-living community was predominantly driven by temperature, whereas the particle-attached community exhibited stronger spatial variation patterns.

Introduction

Marine microbes are the most abundant organisms on earth (Whitman *et al.* 1998), capable of thriving in all oceanic habitats and thus, constitute an enormous biodiversity. Due to their inexhaustible metabolic and physiological versatility they are substantial key players in every biogeochemical cycle and thus, are fundamental to ecosystem functioning. Hence, unveiling the mechanisms that regulate and maintain this diversity, microbial community assembly, distribution and variation is of fundamental interest in marine ecology. The existence of microbial biogeographic patterns is well established and it has been studied extensively in aquatic systems during the past few decades on various spatial scales (e.g. Long & Azam 2001; Fuhrman *et al.* 2008; Martiny *et al.* 2011; Lear *et al.* 2014). A common understanding is that bacterial community similarity is decreasing with increasing geographic distance referred to as "distance-decay" relationship. These spatial variations are often linked to dispersal limitation and shifts in physico-chemical environmental factors (Hanson *et al.* 2012) that exhibit strong gradients. Among these environmental factors, temperature and salinity seem to have largest influence on global bacterial community structure and richness (Lozupone & Knight 2007; Fuhrman *et al.* 2008). On the other hand microbial communities on microscales (Long & Azam 2001), within estuaries (Wang *et al.* 2015) and along transects of up to 2000 kilometres (Fortunato & Crump 2011; Herlemann *et al.* 2011) varied in response to organic matter distribution, salinity, temperature, depth, nutrient concentrations and suspended particles for instance. Furthermore, the temporal variation has been extensively studied in various aquatic environments. Seasonal shifts in bacterial community composition (BCC) are substantially driven by changes in temperature and nutrient concentrations (Andersson *et al.* 2009; Gilbert *et al.* 2009; Lucas *et al.* 2015b). Multi-annual studies revealed that recurrence of bacterial community structure is predictable from ocean environmental conditions such as temperature and day length for instance (Fuhrman *et al.* 2006; Gilbert *et al.* 2012). However, marine habitats represent continuous, highly connected environments, where changes in bacterial communities are complex and triggered by temporal and spatial components simultaneously. So far, only few studies consider both components and describe spatiotemporal variation patterns in oceanic environments (Morris *et al.* 2005; Hewson *et al.* 2006; Fortunato *et al.* 2012).

The German Bight, located in the south-eastern part of the North Sea, is a relatively shallow (10-40 m) temperate, semi-enclosed continental shelf sea. Water currents in the German Bight are predominantly influenced by tides, wind forces and freshwater inflow

from the rivers Elbe and Weser (Howarth 2001). Mixing of marine and freshwater typically leads to pronounced salinity and temperature gradients. Additionally, high loads of dissolved and particulate organic matter are introduced from intertidal flats and Elbe and Weser rivers (Lübben *et al.* 2009). The environmental conditions in this highly dynamic ecosystem have been continuously monitored since 1962 around the Island of Helgoland in the German Bight (54°11.3' N, 7°54.0' E), known as the Helgoland Roads time series (Wiltshire *et al.* 2008). The herein recorded data include physico-chemical parameters such as temperature, salinity, Secchi-depth, and concentrations of dissolved inorganic nutrients (phosphate, nitrate, nitrite, ammonium, silicate), as well as biological parameters such as qualitative and quantitative data on phyto-, zoo- and bacterioplankton.

The bacterioplankton community at Helgoland Roads has been in-deep studied under temporal aspects using a wide range of different microbiological and molecular methods. Seasonal variation was demonstrated on different time scales covering several months to multiple years using fingerprint methods like RISA, DGGE and 16S rRNA gene tag sequencing (Gerdts *et al.* 2004; Sapp *et al.* 2007; Lucas *et al.* 2015b). The authors linked variation in community composition with various environmental parameters and revealed temperature and phytoplankton abundance as main driving forces. Short-term variation of the bacterioplankton community at Helgoland Roads during a spring phytoplankton bloom was analyzed in the frame of a comprehensive metagenomic and –proteomic study (Teeling *et al.* 2012). Additionally, day to day variation was linked to variation in the molecular composition of dissolved organic matter (DOM) to investigate bacteria-DOM interactions (Lucas *et al.* 2015a). Although temporal aspects have been well studied, spatial variation patterns in the German Bight have rarely been examined. One study by Rink *et al.* (2011) compared bacterial communities at pelagic offshore and coastal inshore sites in the German Bight, in relation to suspended particulate matter and phytoplankton composition. However, conditions at Helgoland Roads are assumed to be influenced by the large-scale hydrographic regime in the German Bight (Raabe & Wiltshire 2009; Stockmann *et al.* 2010), thus, observed changes in the bacterial community are complex and comprise both temporal (succession) and spatial (dispersion) components. A single study by Selje & Simon (2003) observed spatiotemporal dynamics of the community composition in the salinity gradient along the Weser and the Weser estuary. Nonetheless, this study only considered nearshore sites and thus, knowledge on spatiotemporal variation on gradients from coast to offshore in the German Bight does not exist.

In this study the spatiotemporal variation of bacterioplankton community in the German Bight was analyzed by automated ribosomal intergenic spacer analysis (ARISA) and multivariate statistical techniques. To integrate the temporally well studied community variation at Helgoland Roads into a spatial context within the German Bight, the surface water community was sampled on a monthly basis over a period of one year along two transects, from the Elbe estuary towards the open North Sea. We aimed at disentangling the temporal and spatial patterns in community variation and focused on the identification of relevant environmental parameters that drive these variation patterns. Furthermore, we tried to uncover differences in the regulation of community assembly of the free-living and particle-attached bacteria.

Material and Methods

Sampling and measurements of environmental parameters

Water samples were obtained monthly at 15 stations along two transects on board the research vessel Uthörn from March 2012 to February 2013 (Fig. 1). The P8 transect starts at Helgoland Island, located in the inner German Bight (54°18.31 N, 7°88.97 E), heads in a north-western direction from Helgoland Island and covers approximately 46 km. The second transect reaches from Helgoland Island to the Elbe estuary at the German coast and is referred to as Elbe transect. Taken together, both transects span a distance of approximately 100 km. At all stations, surface water was collected at 1 m depth using 5 L Niskin bottles attached to a CTD (SST-CTD90, Sea&Sun Technology, Germany). Temperature, salinity, dissolved oxygen (DO), Chlorophyll *a* (Chl *a*), turbidity and colored dissolved organic matter (cDOM) were recorded simultaneously. For determination of dissolved organic carbon (DOC) concentrations, 20 ml of each sample were filtered through 0.7 µm glass fiber filters (GF/F Whatman, UK) into precombusted glass vials (400°C, 5h), acidified to pH 2 (HCl 32% p.a., Carl Roth, Germany) and stored at 4°C in the dark. DOC concentrations were measured by high-temperature catalytic combustion using a TOC-VCPH/CPN total organic carbon analyzer (Shimadzu, Japan). The Deep Sea Reference Standard from the Consensus Reference Material Project (CRM; http://yyy.rsmas.miami.edu/groups/biogeochem/CRM.html) was used to determine the precision and accuracy of the measured concentrations in each run.

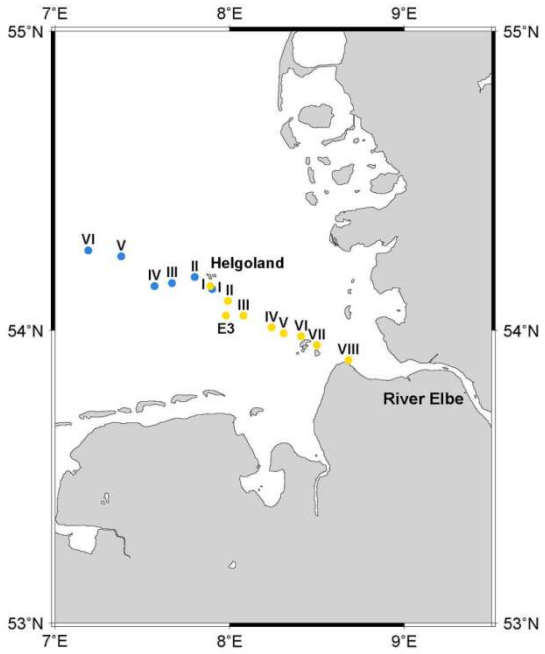


FIGURE 1: Sampling sites along the two transects Elbe and P8 reaching from the Elbe estuary (German coastline) towards the open North Sea. Circles denote sampling sites along the transects Elbe (yellow) and P8 (blue). Map was created with Ocean Data View (R. Schlitzer, http://odv.awi.de).

Bacterial community analysis

500 ml of each sample were subjected to sequential filtration through 10, 3 and 0.2 µm pore size polycarbonate filters (Millipore, Germany) to separate particle-attached from free-living bacteria. Filters with bacterial biomass were stored at -20°C until further processing. DNA extraction from the 3 µm and 0.2 µm filters was done as described previously (Sapp *et al.* 2007). Briefly lysozyme and sodium dodecyl sulphate were used for cell lysis followed by extraction with phenol-chloroform-isoamylalcohol (25:24:1) and precipitation with isopropanol. DNA concentration per sample and purity were measured in duplicates using a Tecan Infinite M200 NanoQuant microplate reader (Tecan, Switzerland).

Automated ribosomal intergenic spacer analysis (ARISA) was performed as described in Krause *et al.* (2012) with slight modifications. Extracted DNA was amplified with forward primer L-D-Bact-132-a-A-18 (*5'-CCGGGTTTCCCCATTCGG-3'*) and reverse primer S-D-Bact-1522-b-S-20 (*5'-TGCGGCTGGATCCCCTCCTT-3'*)**,** the latter labelled with an infrared dye (Ranjard *et al.* 2000). PCRs were performed in volumes of 25 µl containing 5 ng template DNA. PCR products were diluted (1:5) with autoclaved ultrapure water. Diluted PCR products were then mixed with an equal volume of formamide containing loading buffer and 0.25 µl were separated in 5.5% polyacrylamide gels at 1500 V for 14h on a LI-COR 4300 DNA Analyzer. A 50-1500 bp size standard was run as a size reference on each gel (all materials: LI-COR Bioscience, USA).

Gels were analysed using the Bionumerics 5.10 software (Applied Maths, Belgium). Bands with intensities lower than 2% of the maximum value of the respective lane and bands smaller than 262 bp were neglected. Binning to band classes was performed according to Kovacs *et al.* (2010). Each band class is referred to as an ARISA operational taxonomic unit (OTU). Peak intensities of ARISA OTUs were translated to binary data reflecting the presence or absence of the respective OTU.

Statistical analyses

To reveal spatial and temporal patterns in environmental conditions along the sampled transects, principal component analysis (PCA) was accomplished for the environmental parameters. Parameters were normalized prior to analyses. To test for statistically significant variance among environmental parameters along the two transects, permutational multivariate analysis (PERMANOVA) was performed based on Euclidean distances at a significance level of $p<0.05$.

Accordingly, principal coordinates analyses (PCoA) were performed with the ARISA fingerprint data of the free-living and particle–attached bacterial communities separately, based on the Jaccard index. To test for statistically significant variance among the free-living and particle-attached bacterial communities and for differences among the community along the two transects respectively, PERMANOVA was accomplished at a significance level of $p<0.05$. Tests of significant differences in the within-group dispersion among groups were accomplished by performing tests of homogeneity of dispersions (PERMDISP) at a significance level of $p<0.05$. To identify axes separating the *a priori* groups (here P8 and Elbe transects), canonical analysis of principle coordinates (CAP) was performed. In contrast to unconstrained ordination techniques (e.g. PCA, PCoA) which maximize total variation among the samples, CAP tries to identify axes that separate samples into *a priori* defined groups in such a way that group differences are maximised (Anderson & Willis 2003). Analyses were performed using Primer v.7 and the PERMANOVA add on software package (both PRIMER-E, UK). Spatiotemporal visualization of PCAs and PCoA scores was accomplished using Surfer 12 (GoldenSoftware, USA). Contour plots were created by using the point kriging method to generate the interpolated grid.

Spearman rank order correlations of environmental parameters were calculated at a significance level of $p<0.05$ to identify potential collinearity. To examine the relationship between the variation in the bacterial community and environmental parameters, stepwise forward multiple regression analyses (MRA) were conducted, using the PCoA scores of the first three PCoA axes as dependent variables and PCA scores of the first three PCA axes as independent variables referred to as principial component regression (e.g. Hotelling 1957). Spearman rank order correlation and MRA were carried out using Statistica 11 (StatSoft, USA).

Results

Spatiotemporal variation in environmental conditions

For the recorded environmental parameters different variations across space and time were observed (Fig. S1). Salinity exhibited a pronounced spatial gradient along the two transects increasing from 15.8 (annual average) at the estuarine sampling site closest to the coastline (Elbe VIII) to 33.7 at the sampling site furthest offshore (P8 VI). Similar but reversed patterns were observed for DOC which decreased on average from 341.3 µM in the estuary to 96.8 µM offshore, turbidity (41.5 to 1.1 FTU), cDOM (26.6 to 2.5) and Chl *a* (5.1 to 0.9 µg l^{-1}). The slope of the observed gradient for the above mentioned parameters was relatively constant from sites P8 VI to Elbe V but became noticeably steep at the estuarine sites Elbe VI –Elbe VIII. In contrast, temperature was relatively stable along both transects, but varied temporally with lowest average values in February (2.8 °C) and highest average values in September (18.1 °C). Dissolved oxygen exhibited a similar but reversed pattern as temperature and lowest average values were measured in September (7.2 mg l^{-1}); highest average values were measured in February (11.1 mg l^{-1}). Although temperature and dissolved oxygen varied slightly from estuarine to offshore sites, they are referred to as temporal parameters as the variation on temporal scale was more pronounced. Opposed variation of salinity and DOC, turbidity, cDOM and Chl *a* and temperature and dissolved oxygen is also reflected in Spearman correlation coefficients (Tab. 1). Salinity showed high significant, negative correlations with DOC(R=-0.97), turbidity (R=-0.82), cDOM (R=-0.95) and Chl *a* (R=-0.69). Temperature was significantly correlated with dissolved oxygen (R=-0.92).

Principal component analyses of the environmental parameters revealed distinct spatiotemporal patterns in sample variation (Fig. 2). The first axis (PC 1) explained 56.6% of the observed variability and is mainly defined by a combination of salinity, DOC, turbidity and cDOM, contributing with roughly equally weighted coefficients to PC 1 (Fig. 2A). The second axis (PC 2) was predominantly defined by the large coefficients of temperature and DO and explained 30.1% of the observed variation (Fig. 2B). The third axis (PC 3) still explained 9.3 % of the variation. Chl *a* contributed with a remarkable high coefficient to this axis (Fig. 2C). In general variation was higher at sites along the Elbe transect compared to sites along the P8 transect as reflected in the larger range of PCoA scores covered by samples along the Elbe transect compared to the P8 transect (Fig. 2).

TABLE 1: Spearman rank order correlations of environmental parameters. Significant correlations ($p<0.05$) are indicated in bold. DO: dissolved oxygen, Chl *a*: Chlorophyll *a*, cDOM: colored dissolved organic matter, DOC: dissolved organic carbon.

	Temperature	Salinity	DO	Chl *a*	Turbidity	cDOM
Salinity	-0.06					
DO	**-0.92**	-0.14				
Chl *a*	**0.47**	**-0.69**	**-0.26**			
Turbidity	-0.12	**-0.82**	**0.23**	**0.50**		
cDOM	-0.14	**-0.95**	**0.33**	**0.54**	**0.88**	
DOC	0.14	**-0.97**	0.04	**0.71**	**0.81**	**0.92**

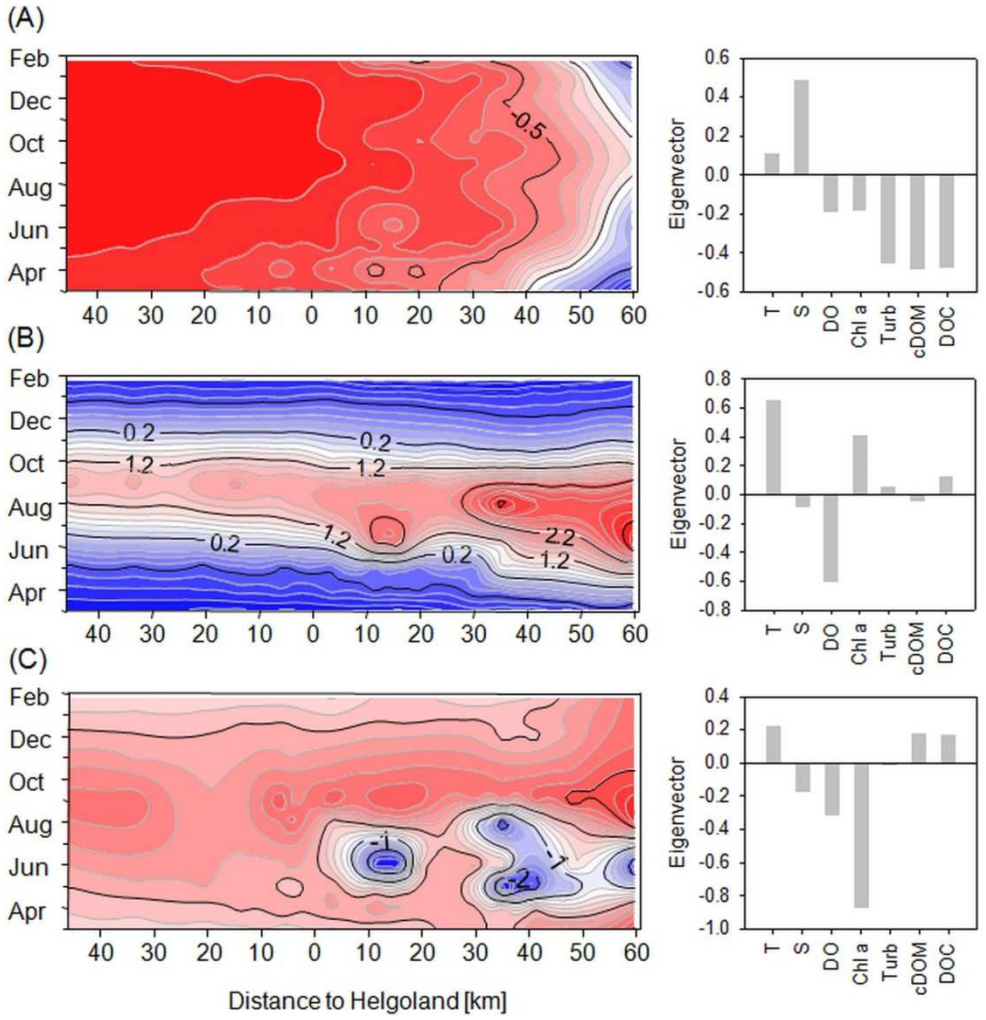


FIGURE 2: Principal component analyses (PCA) of measured environmental parameters. PCA scores of the first PCA axis (A), second PCA axis (B) and third PCA axis (C) are depicted in contour plots. The horizontal axes depicts the distance [km] of sampling sites to Helgoland Island which was set to 0 km. Increasing distance to the left represents the sampling sites along the P8 transect, increasing distance to the right represents sampling sites along the Elbe transect. The vertical axis refers to the sampling date; color code reflects PCA scores of respective samples with blue colors indicating lower scores and red colors indicating higher scores. Coefficients of the environmental parameters in the linear combinations defining the respective PCA axes are given next to the contour plots. T: temperature, S: salinity, DO, dissolved oxygen, Chl *a*: Chlorophyll *a*, Turb: turbidity, cDOM: colored dissolved organic matter, DOC: dissolved organic matter.

Spatiotemporal variation of bacterial community composition and relevant driving forces

PERMANOVA revealed significant (p<0.05) differences between the free-living and particle-attached bacterial community (Fig. S2 and Tab. S1). To further elucidate patterns of variation within the free-living and particle-attached community, multivariate statistical analyses were accomplished separately for each fraction. To identify which environmental parameters are most likely to drive the variation, MRA were conducted with PCoA axes as dependent variables. To account for multicollinearity as revealed by high significant correlations (|r|>0.7) of environmental parameters (Tab. 1) and to avoid erroneous MRA we replaced the original environmental data by scores of the PCA axes as explanatory variables (also referred to as latent variables), according to the "Principial component regression" approach (Hotelling 1957; Dormann *et al.* 2013). Since PCA axes are orthogonal (i.e. perfectly uncorrelated) multicollinearity was completely removed by this approach.

The free-living bacterial community exhibited a pronounced spatiotemporal pattern when taking the scores of the first PCoA axis into account (Fig. 3), explaining 15.1% of the total variation. The pattern of PCoA 1 scores in summer was clearly different from that in spring and winter along both transects (Fig. 3A). The spatiotemporal variation of PCoA 1 scores is significantly (p<0.05) and well described (R^2_{adj}=0.72) by the regression model. Among the explaining latent variables, PC 2 showed highest relative contribution to the prediction of PCoA 1 (Fig. 3A) as reflected in the high standardized regression coefficient (b*=0.73). PC 2 is mainly characterized by temperature as reflected in the high coefficient of temperature in the linear combination defining PC 2 (Fig. 2B). PC 1 and PC 3 contribute less (b*=0.32 and b*=0.29, respectively), but still significantly to the prediction of the variation pattern of PCoA 1. The second PCoA axis (PCoA 2) explains 11.6% of the total variation (Fig. 3B). The PCoA 2 pattern of the P8 transect north-west off Helgoland was homogenous throughout the whole year, whereas pronounced variation was observed along the Elbe transect. Here, the pattern in late spring and summer (May to August) was similar to that of the P8 transect, but differed strongly from that in spring and autumn along the Elbe transect. The spatiotemporal pattern of PCoA 2 is significantly (p<0.05) described by a combination of PC 1 (characterized by salinity, DOC, turbidity, cDOM) and PC 3 (mainly defined by Chl *a*) (Fig. 3B). Considering the low coefficient of determination (R^2_{adj}=0.20), the variation of PCoA 2 scores is explained rather poorly. The third axis (PCoA 3) explains 9.1% of the total variation and demonstrates high similarity of the

pattern along both transects during late spring and summer (May to July) (Fig. 3C). However, the pattern during spring and summer exhibited pronounced differences when compared to the patterns of autumn and winter. PC 1 and PC 2 contributed significantly with comparable amounts to the prediction of PCoA 3 (b*=-0.38 and b*=0.41). PC 3 contributed to a less extend to the prediction (b*=-0.18) (Fig. 3C). As for PCoA 2, the MRA model for PCoA 3 exhibited a low coefficient of determination (R^2_{adj}=0.3).

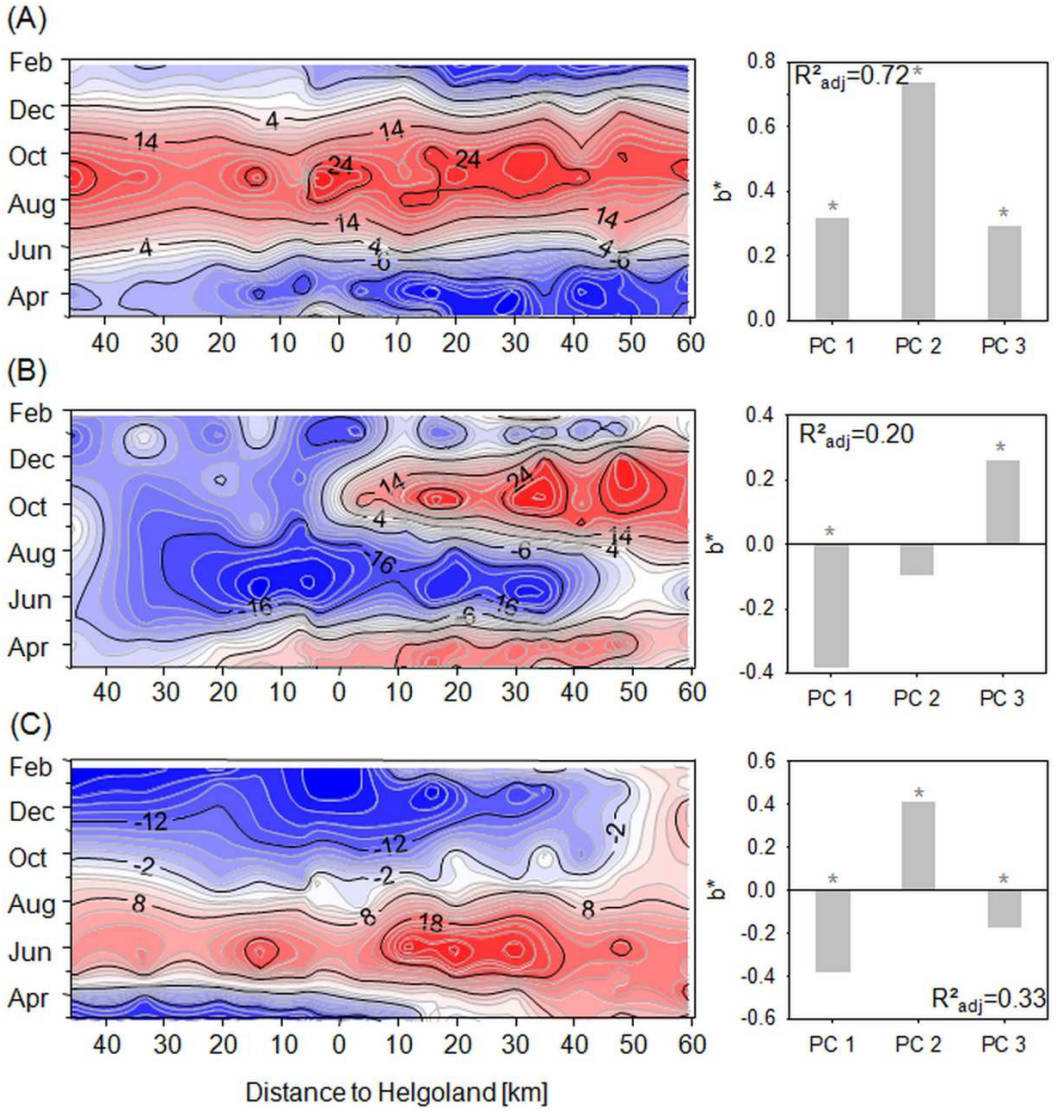


FIGURE 3: Principal coordinates analysis (PCoA) of ARISA OTUs of the free-living fraction, based on Jaccard index. PCoA scores of the first PCoA axis (A), second PCoA axis (B) and third PCoA axis (C) are depicted in contour plots. The horizontal axes depicts the distance [km] of sampling sites to Helgoland Island which was set to 0 km. Increasing distance to the left represents the sampling sites along the P8 transect, increasing distance to the right represents sampling sites along the Elbe transect. The vertical axis refers to the sampling date; color code reflects PCoA scores of respective samples with blue colors indicating lower scores and red colors indicating higher scores. Standardized regression coefficients (b*) of PCA axes of MRA using scores of PCoA axes as dependent and scores of PCA axis as independent variables are depicted next to the corresponding contour plots. MRA were done at a significance level of $p<0.05$, R^2_{adj} values are given. Asterisks indicate significance of regression coefficient.

Figure 4 depicts the spatiotemporal variation of the respective PCoAs of the particle-attached bacterial community. PCoA 1 explains 16% of the total variation, PCoA 2 explains 10.4% and PCoA 3 explains 7.3% (Fig. 4). For PCoA 1 the spatiotemporal pattern of the entire P8 transect was homogenous throughout the year, whereas the pattern along the Elbe transect was clearly more variable (Fig. 4A). In late spring and summer (May to August) the pattern along the Elbe transect was similar to that of the P8 transect, but varied from the early spring, autumn and winter patterns along the Elbe transect. The variation pattern of PCoA 1 seems to be mainly predicted by PC 1 (mainly defined by salinity, DOC, turbidity, cDOM) with b*=0.49 and to a less but still considerable amount by PC 2 (temperature) with b*=0.28 (Fig. 4A). The variation pattern of PCoA 2 along time and space is depicted in Figure 4B. Here, the pattern in spring (April to June) was similar along both transects, but clearly differed from the patterns in summer, autumn and winter. Variation of PCoA 2 is solely explained by PC 3 (mainly defined Chl *a*), contributing with a standardized regression coefficient of b*=0.38 to the prediction of PCoA 2 (Fig. 4B). Variation of PCoA 3 (Fig. 4C) reveals general differences between summer and winter. Focusing on the variation during summer (June to August), it becomes clear that the pattern along the Elbe transect and sites P8 I to III was particularly similar. Spatiotemporal patterns of PCoA 3 are significantly explained by PC 2 (defined by temperature) to a large extend (b*=0.67) (Fig. 4C). PC 1 and PC 3 also contributed significantly but to a much lower extend (b*=-0.27 and b*=0.16) to the prediction of PCoA 3.

The variation patterns of the particle-attached bacterial community are generally less well described than the patterns in the free-living community, which is reflected in the low R^2_{adj} values (R^2_{adj}=0.31 for PCoA 1, R^2_{adj}=0.14 for PCoA 2 and R^2_{adj}=0.54 for PCoA 3).

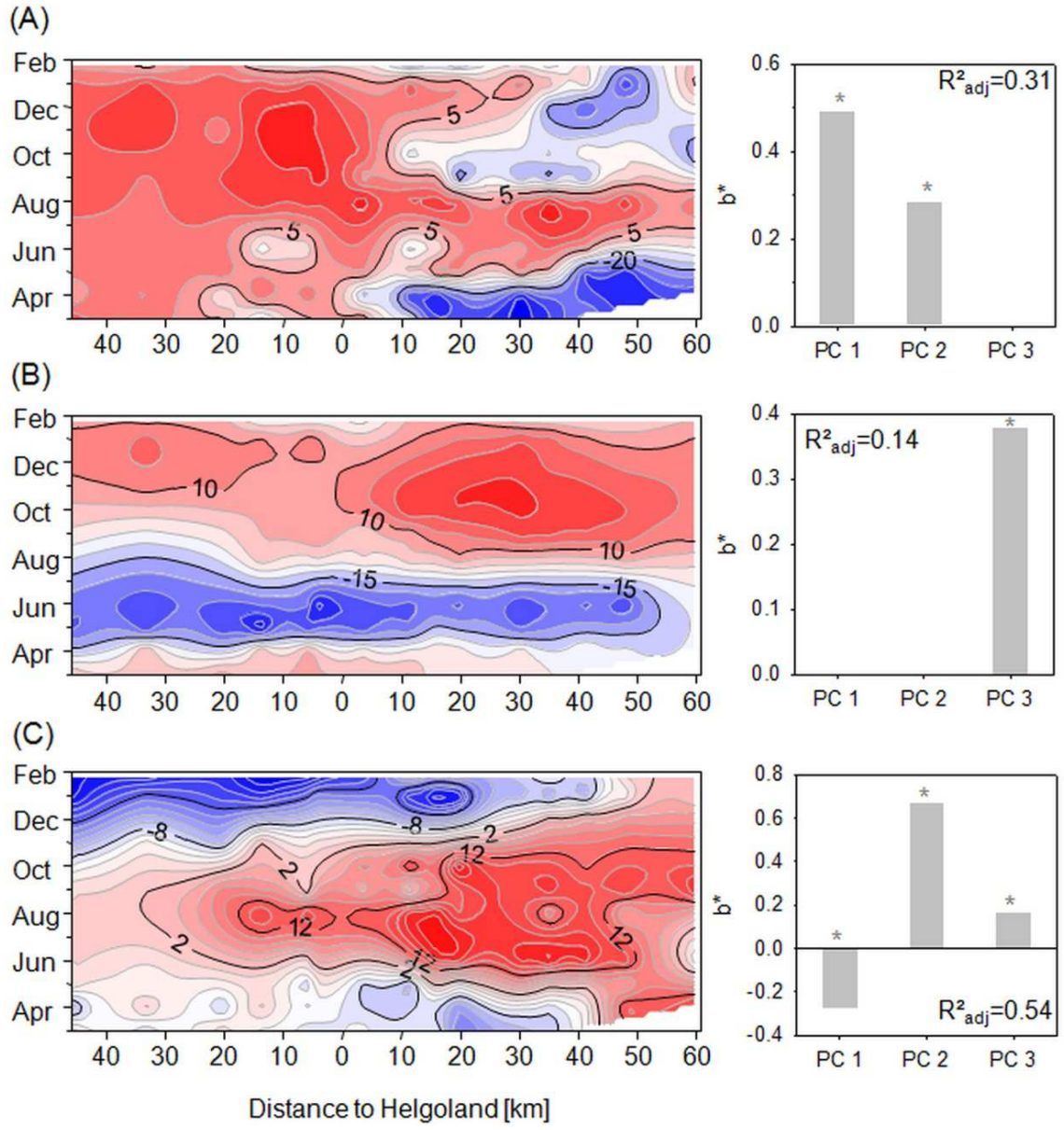


FIGURE 4: Principal coordinates analysis (PCoA) of ARISA OTUs of the particle-attached fraction, based on Jaccard index. PCoA scores of the first PCoA axis (A), second PCoA axis (B) and third PCoA axis (C) are depicted in contour plots. The horizontal axes depicts the distance [km] of sampling sites to Helgoland Island which was set to 0 km. Increasing distance to the left represents sampling sites along the P8 transect, increasing distance to the right represents sampling sites along the Elbe transect. The vertical axis refers to the sampling date; color code reflects PCoA scores of respective samples with blue colors indicating lower scores and red colors indicating higher scores. Standardized regression coefficients (b*) of PCA axes of MRA using scores of PCoA axes as dependent and scores of PCA axis as independent variables are depicted next to the corresponding contour plots. MRA were done at a significance level of $p<0.05$, R^2_{adj} values are given. Asterisks indicate significance of regression coefficient.

Separation of samples into a priori groups corresponding to the sampled transects

The variation patterns of the free-living and particle-attached bacterial communities point to differences in community composition between the P8 transect and the Elbe transect (Figs. 3B and 4A) and thus, might suggest a separation of samples into the two *a priori* groups, corresponding to the two transects. Indeed, comparison of both transects via PERMANOVA revealed significant (p<0.05) differences in community composition of the free-living and particle-attached bacterial community as well as differences in environmental parameters (Tab. S1). However, PERMDISP also revealed significant differences in dispersion of the respective data sets (Tab. S2). These evidenced heterogeneities in dispersion may affect PERMANOVA results adversely and thus, PERMANOVA results need to be interpreted with caution. Alternatively and as a complementary approach, canonical analysis of principal coordinates (CAP) was used to identify an axis separating the multivariate data cloud and possibly relocating samples allocated falsely to one of the *a priori* groups. Separation of samples along this axis helps to further characterize the *a priori* groups of samples, to visualize differences among them and to assess how distinct these groups are from each other. Separation of samples into the two *a priori* groups (P8 and Elbe) via CAP partly confirmed the significant differences that have been revealed by PERMANOVA (Fig. 5 and Tab. 2). Concerning the free-living bacterial community the separation of the *a priori* groups (Fig. 5A) is supported by the reasonably large correlation value of 0.85 (Tab. 2), which indicates the strength of the association between the multivariate data and the hypothesis of group differences. However, 9.5% of the samples were misclassified, most of which derived mainly from sampling site Elbe I (Tab. 2) in the vicinity of the island of Helgoland. Accordingly, separation of *a priori* groups based on the particle-attached community composition (Fig. 5B), revealed a comparably high correlation value (0.83) and a misclassification error of 10.7% (Tab. 2). The majority of the misclassified samples derived from sampling sites Elbe I and Elbe II. However, separation of samples based on environmental parameters (Fig. 5C) led to a considerably lower correlation value of 0.6 and a higher misclassification error of 18.9% (Tab. 2). Here, most of the misclassified samples belong to sampling sites Elbe I to Elbe III. Noticeably, in January and February samples from sites Elbe IV and V were also misclassified.

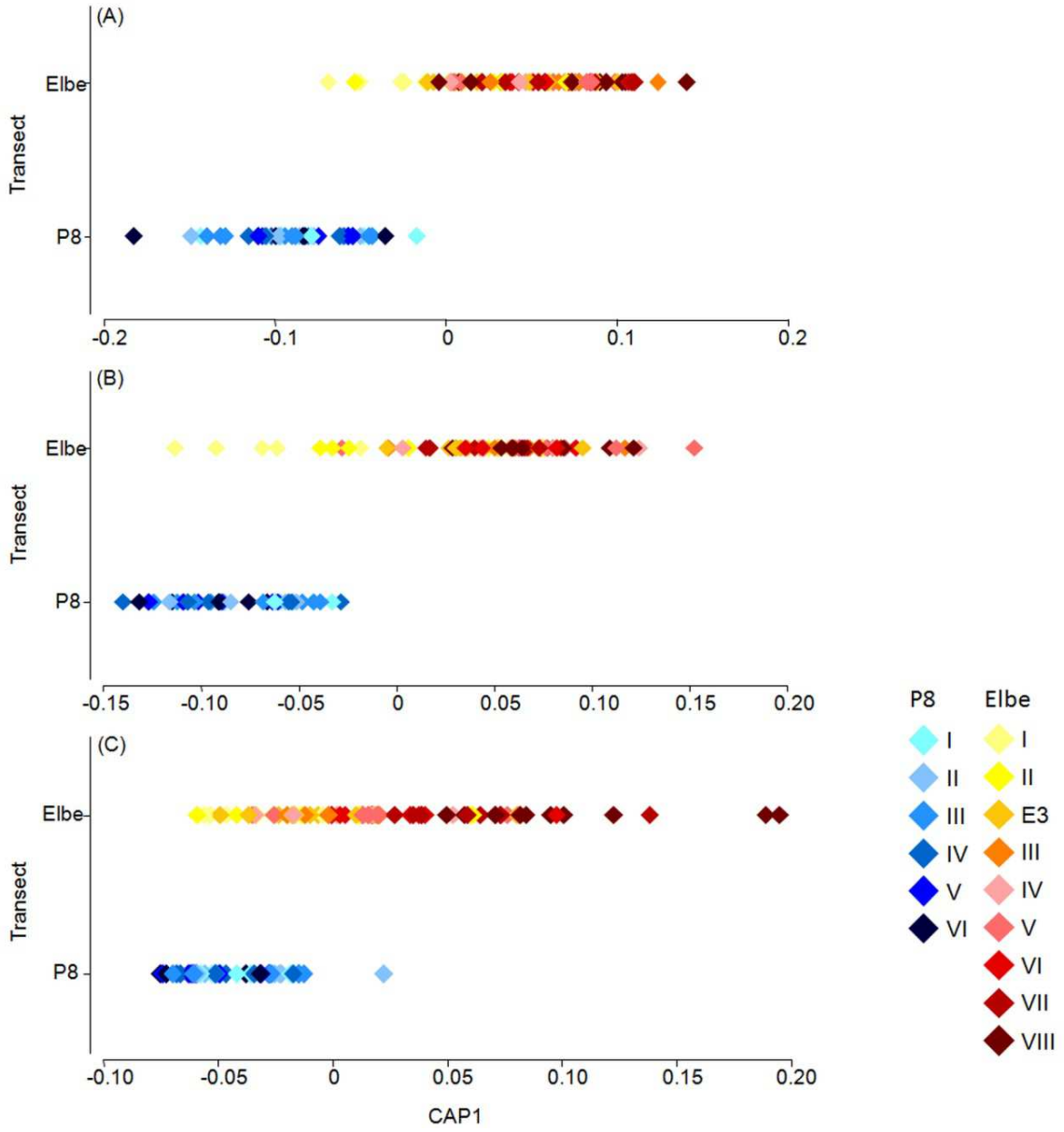


FIGURE 5: Canonical analyses of principal coordinates (CAP). Separation of *a priori* groups (P8 and Elbe) based on (A) free-living bacterial community composition, (B) particle-attached bacterial community composition and (C) environmental parameters. Roman numerals in the legend refer to the corresponding sampling sites along the P8 and Elbe transects.

TABLE 2: Canonical analyses of principal components (CAP) of the free-living and particle-attached bacterioplankton community and of environmental parameters (env). Orig. group: Indicates the a priori group of the respective samples, Class. group: indicates the group classification resulting from CAP analysis.

	free-living	**particle-attached**	**env**
Eigenvalue	1	1	1
Correlation	0.8456	0.8308	0.597
Corr.Sq.	0.715	0.6902	0.3564
Total correct	115/127 (90.6%)	109/122 (89.3%)	103/127 (81.1%)
Miss-classification error	9.50%	10.70%	18.90%

Individual samples that were miss-classified

Sample	**Orig. group**	**Class. group**	**Sample**	**Orig. group**	**Class. group**	**Sample**	**Orig. group**	**Class. group**
P8 I (Sep)	P8	Elbe	P8 IV (Apr)	P8	Elbe	P8 II (Apr)	P8	Elbe
Elbe I (Mar)	Elbe	P8	P8 III (Jun)	P8	Elbe	Elbe I (Mar)	Elbe	P8
Elbe I (Apr)	Elbe	P8	P8 III (Mar)	P8	Elbe	Elbe II (Mar)	Elbe	P8
Elbe I (May)	Elbe	P8	Elbe I (May)	Elbe	P8	Elbe I (May)	Elbe	P8
Elbe II (Jun)	Elbe	P8	Elbe I (Aug)	Elbe	P8	Elbe II (May)	Elbe	P8
Elbe I (Aug)	Elbe	P8	Elbe V (Aug)	Elbe	P8	Elbe III (May)	Elbe	P8
Elbe I (Oct)	Elbe	P8	Elbe II (Jan)	Elbe	P8	Elbe E3 (Aug)	Elbe	P8
Elbe I (Jan)	Elbe	P8	Elbe III (Jan)	Elbe	P8	Elbe I (Sep)	Elbe	P8
Elbe I (Feb)	Elbe	P8	Elbe I (Feb)	Elbe	P8	Elbe I (Oct)	Elbe	P8
Elbe E3 (Sep)	Elbe	P8	Elbe I (Mar)	Elbe	P8	Elbe II (Oct)	Elbe	P8
Elbe II (Jan)	Elbe	P8	Elbe II (Mar)	Elbe	P8	Elbe E3 (Oct)	Elbe	P8
Elbe VIII (Mar)	Elbe	P8	Elbe I (Sep)	Elbe	P8	Elbe III (Oct)	Elbe	P8
			Elbe II (Sep)	Elbe	P8	Elbe I (Jan)	Elbe	P8
						Elbe E3 (Jan)	Elbe	P8
						Elbe III (Jan)	Elbe	P8
						Elbe IV (Jan)	Elbe	P8
						Elbe V (Jan)	Elbe	P8
						Elbe I (Feb)	Elbe	P8
						Elbe III (Mar)	Elbe	P8
						Elbe E3 (Mar)	Elbe	P8
						Elbe II (Sep)	Elbe	P8
						Elbe E3 (Sep)	Elbe	P8
						Elbe II (Jan)	Elbe	P8
						Elbe IV (Feb)	Elbe	P8

Discussion

Although temporal and spatial patterns of bacterial communities have been studied intensively in various marine environments only few studies consider the simultaneous variation in time and space (Morris *et al.* 2005; Hewson *et al.* 2006; Fortunato *et al.* 2012). Here we present a comprehensive spatiotemporal annual survey of the bacterioplankton community along a 100 km transect in the German Bight (North Sea) reaching from brackish waters (Elbe estuary) to offshore sites and link it with environmental parameters.

Helgoland Roads: An oceanographic transition zone

Principal component analysis revealed that the variation in environmental parameters was mainly driven by spatial gradients and to a less extend by temporal patterns. It is known that the coastal waters of the German Bight are strongly influenced by the discharge of Elbe riverine water, which represents the most relevant freshwater source in the German Bight (Callies & Scharfe 2015). This freshwater input is accompanied by high nutrient and particle loads and leads to the observed pronounced gradients of increasing salinity and decreasing concentrations of DOC, cDOM, nutrients and turbidity from the Elbe estuary towards offshore areas. However, the observed gradient was most pronounced at the Elbe transect sites between Helgoland Island and the coastline; environmental conditions at the offshore sites (P8 transect) north-west off Helgoland appeared to be more homogenous which might be due to different influencing water masses. Scharfe (2013) stated that the main water current pattern in the German Bight is characterized by advection of water masses from a western direction into the German Bight, which then moves on in a northern direction. Helgoland Island is located at the eastern boundary of this main current direction. Thus, it might be seen as border, where sampling sites north-west of Helgoland are influenced by water masses following this main current pattern and exhibit oceanic environmental conditions. In contrast, sampling sites south-east of Helgoland are influenced by costal water masses and river Elbe inflow, i.e. coastal conditions with high particle load and nutrient concentrations but low salinity predominate. However, the Helgoland area is occasionally influenced by riverine dominated coastal waters, controlled by hydrological and meteorological forces and river discharge (Stockmann *et al.* 2010), which might result in short-term interference of environmental conditions as demonstrated in Lucas *et al.* (2015a) and Teeling *et al.* (2012). Hence, the classification of water masses around Helgoland Island to either marine or coastal water is not trivial and thus, the Helgoland area might be referred to as an oceanic transition zone between coastal and

central North Sea waters as already suggested by Raabe & Wiltshire (2009). Canonical analyses of principal components of either environmental parameters or bacterial community composition strongly confirm this idea and further localize this transition zone more precisely (Fig. 5) as discussed in the following paragraph. Along the investigated transects a strong gradient in spatial parameters as reflected by salinity, DOC, turbidity and cDOM from the Elbe estuary towards the central North Sea is obvious (Fig. 5C). This gradient is not consistent, but exhibits varying strength in different sections of the transects. It is most pronounced at the estuarine sites Elbe VI – VIII, where environmental conditions show strong, abrupt changes, reflected in the relatively large range in which samples of these sites stretch along the CAP axis (Fig. 5C). Towards Helgoland Island (Elbe IV and V) environmental conditions are changing more gradually and thus, the gradient flattens. Approximately 20 km south-east of Helgoland Island environmental conditions become similar to that of the offshore sampling sites north-west of Helgoland (P8 transect) which is reflected in misclassification of samples of sites Elbe I-III (Tab. 2) and the visible overlap of sites Elbe I-III with sites of the P8 transect (Fig. 5C).

However, the separation of samples based on environmental parameters was not congruent with CAP analysis of the bacterial community (Fig. 5A and B). The classification of samples suggested for the bacterial community lets us assume that a reasonable spatial separation of samples could be achieved by assigning all samples along the P8 transect plus the samples from sampling site Elbe I for the free-living community and samples along P8 plus sites Elbe I and II for the particle-attached community, into one group (referred to as offshore), and the remaining samples along the Elbe transect into a second group (coastal). A possible explanation is that different water masses with differing salinity and related density gradients might lead to dispersal limitation of bacterial populations, which might explain the observed separation of coastal (Elbe transect) and offshore (P8 transect) samples based on the free-living bacterial community composition. A comparable separation of water masses and thereby communities has been also proposed for other coastal-offshore transects (Fortunato & Crump 2011), for deep-water research moorings (Morris *et al.* 2005) and on a global scale (Fuhrman *et al.* 2008).

Free-living and particle-attached bacterial communities are triggered differently

Pronounced spatial patterns of marine bacterial community composition have been described for estuarine areas that exhibit strong salinity gradients (Fortunato & Crump 2011; Wang *et al.* 2015) as well as oceanic water masses with distinct gradients in salinity

or temperature for instance (Fuhrman *et al.* 2008; Herlemann *et al.* 2011). However, as part of a semi-enclosed continental shelf sea the German Bight represents a unique, highly productive coastal environment that is strongly influenced by its intertidal flats, freshwater inflow of rivers and exhibits rather small-scale, highly variable hydrographic properties (Becker *et al.* 1992; Staneva *et al.* 2009). Few studies systematically compared water or sediment bacterial community composition of coastal and oceanic sites in this region (Rink *et al.* 2011; Störmer 2013) hence, knowledge of the spatiotemporal variation of the bacterial community and its driving forces in the German Bight is scarce.

Due to the above mentioned strong freshwater input of the Elbe River and the observed gradients in salinity, DOC, cDOM and turbidity, it could be assumed that the variation of the bacterial community composition changes gradually as well from riverine to marine habitats as shown by other studies on the spatial variability along environmental gradients (e.g. Herlemann *et al.* 2011; Fortunato *et al.* 2012). Surprisingly, variation in the free-living bacterial community was dominated by temporal changes in temperature along both transects, rather than by parameters that exhibit pronounced spatial gradients (salinity, DOC, turbidity, cDOM). Fuhrman *et al.* (2008) defined temperature as the major influencing factor in a global large-scale study on bacterioplankton richness. They stated that temperature strongly affects kinetic mechanisms (rates of reproduction, dispersal, species interaction, adaptive evolution etc.) and thus, has potentially strong influence on the diversity. This is also supported by a recent study on the annual bacterial dynamics at Helgoland Roads (Lucas *et al.* 2015b). The authors suggest that temperature constitutes a major factor for the formation of ecological niches in the German Bight and indirectly affects short-term bacterial succession in response to phytoplankton blooms. This supports the assumption that the variation of the bacterial community along the examined transect in the German Bight is mainly driven by temperature. The strong influence of temperature overlying other environmental factors like salinity, DOC, DOM (as represented by cDOM) and phytoplankton (as represented by Chl *a*) might also point to a relatively broad tolerance of the free-living coastal bacterial community concerning the latter factors. However, it has to be noted that this study is based on binary data, i.e. our diversity analyses only consider the presence or absence of ARISA OTUs. Relative abundances or activity of specific OTUs however, might be triggered by different environmental parameters depending on their respective ecological niches.

Considering the influences of the different environmental parameters, the impact of phytoplankton abundance (represented by Chl *a* concentrations) on the spatiotemporal

free-living community variation in this study is particularly interesting. It is a known fact that bacterioplankton community composition is strongly influenced by enhanced substrate supply during and on decline of phytoplankton blooms and many studies assessed the response of bacterial communities to phytoplankton blooms with regard to different aspects (e.g. Zubkov 2001; Pinhassi *et al.* 2004; Rooney-Varga *et al.* 2005; Rink *et al.* 2007; Teeling *et al.* 2012; Sarmento *et al.* 2013b; Wemheuer *et al.* 2014). Although our data also imply an influence of phytoplankton on the community structure, this influence is only of minor importance since the main contribution of Chl *a* is to the third PCA axis (Fig. 2), which again is of minor importance for the explanation of the variation pattern of the free-living community (Fig. 3). There is a major contribution of Chl *a* to the explanation of the variation pattern of PCoA 2 of the particle-attached community (Fig 4B), but as the variation pattern is explained rather poorly (R^2_{adj}=0.14) this does not point to a pronounced influence of phytoplankton on the community variation. Therefore, we propose that a strong influence of phytoplankton on the bacterioplankton community composition is restricted to short time scales during phytoplankton blooms and is of minor importance for the overall long-term patterns in community composition such as resilience and recurrence. This assumption is supported by an 16S rRNA gene tag sequencing based annual survey on the bacterioplankton community at Helgoland Roads that reported a rapidly changing community composition during phytoplankton blooms which was overwhelmed by temperature-driven seasonal variation (Lucas *et al.* 2015b). However, interdependencies between phyto- and bacterioplankton cannot easily be disentangled since growth of both organism groups rely to some extent on the same environmental triggers (temperature, nutrients) and also interact (via exudates) or compete (nutrients) (Buchan *et al.* 2014).

Despite the strong temporal influence, spatial patterns were also observed for the free-living community which is reflected in the patterns of PCoA 2 (Fig. 3B) albeit this patterns were merely explained by the measured environmental parameters (R^2_{adj}=0.20). Due to this poor relationship of environmental parameters and patterns of PCoA 2, we assume that other factors that were not analyzed during this study might be relevant for interpretation. As already mentioned a varying coastal water inflow to the Helgoland area is assumed which is related to meteorological and hydrodynamic conditions and might result in short-term interference of environmental conditions (Stockmann *et al.* 2010; Scharfe 2013). To relate this varying current pattern in the coastal area with the observed PCoA 2 pattern the hydrodynamic variability in the German Bight was assessed using current velocity fields

from the model BSHcmod (Dick *et al.* 2001) operated by the Federal Maritime and Hydrographic Agency of Germany (Bundesamt für Seeschiffahrt und Hydrographie, BSH) (detailed information see supplementary material; Figs. S2 and S3). Deviations of the main current patterns in the German Bight within the period March 2012-March 2013 are depicted in Figure 6A and 6B.

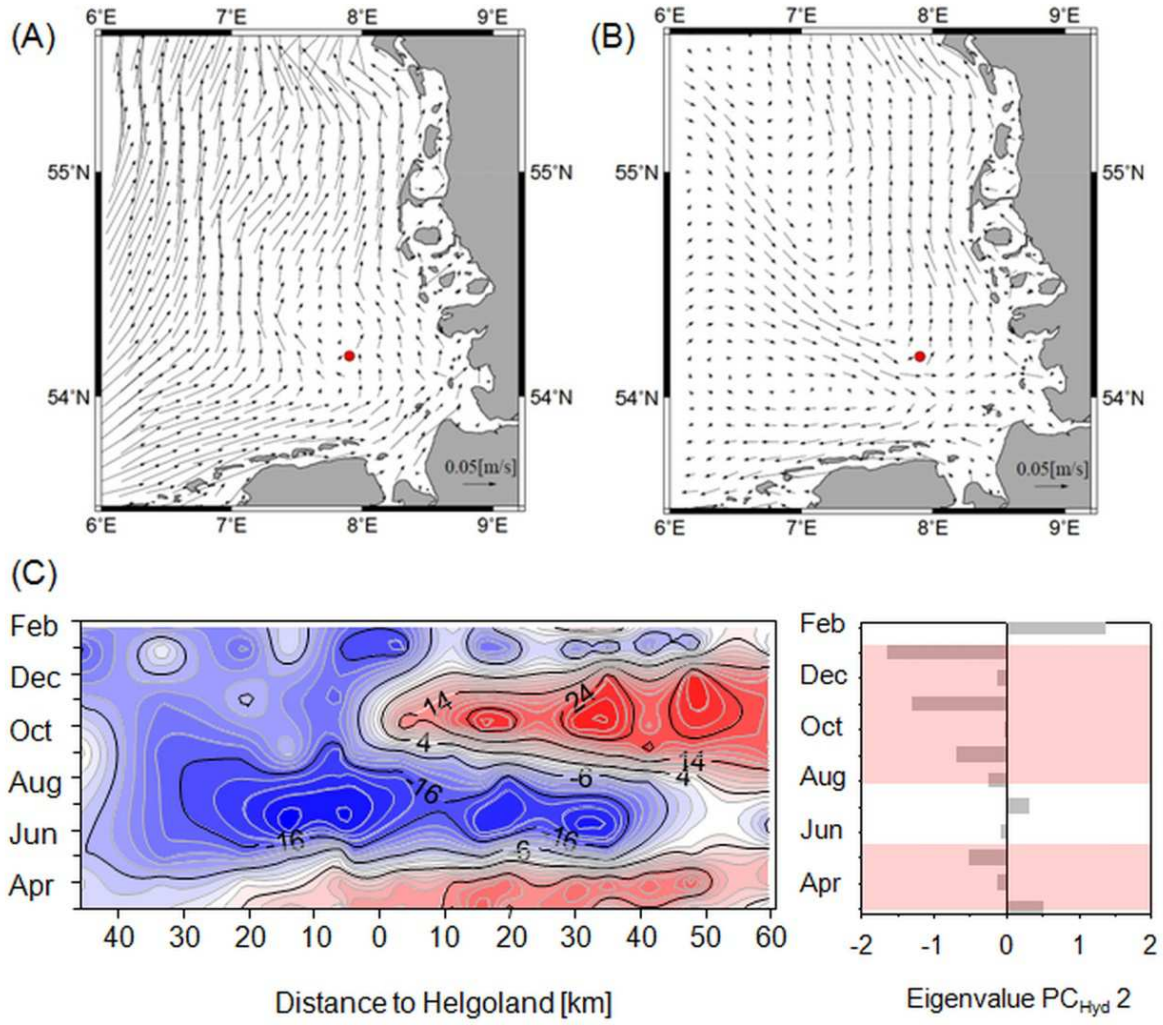


FIGURE 6: Vector fields of current anomalies (EOF pattern) in the German Bight within the period March 2012 – March 2013. Explained variances are 73.4 % for the first (A) and 12. 2% for the second (B) EOF. Red dot: Helgoland. The second principal component (PC_{Hyd} 2) corresponding to the second EOF pattern is compared to the variation of the free-living bacterial community along the second PCoA axis (C). Transparent red boxes mark timeframes in which the bacterial communities of the coastal site are notably different to that of the offshore sites and in which eigenvalues of PC_{Hyd} 2 become negative.

It is obvious that hydrographic conditions at Helgoland Roads are influenced by current anomalies that represent an inflow of open North Sea waters (Fig. 6B). The corresponding time coefficient (PC_{Hyd} 2) of this pattern is compared to the PCoA 2 pattern in Figure 6C. Positive values of the time coefficient reflect the pattern depicted in Figure 6B, negative values reflect the reverse pattern when central North Sea water flows into a northern direction off Helgoland and is replaced by an inflow of coastal water. Since negative values of PC_{Hyd} 2 which are related to enhanced coastal water influence at Helgoland Roads occurred predominantly in spring and autumn, we assume that the observed differences of bacterial communities along the transect can be partly explained by these current patterns.

In contrast to the free-living community, the variation of the particle-attached community was mainly driven by salinity, DOC, cDOM and turbidity, thus following their pronounced spatial patterns (Fig. 4). Temperature dependent variation was not as relevant as for the free-living community, which is evident from the relatively small contribution of PC 2 to the variation of the PCoAs (Fig. 4). However, spatiotemporal patterns of the particle-attached community were poorly explained by MRA analyses, reflected by the small R^2_{adj} values. Thus, interpretation of the variation is difficult and the main driving forces remain unclear.

Although we demonstrated clear patterns in the variability of the bacterial community composition of the free-living community in the German Bight, there are some drawbacks that need to be considered. First, the set of measured environmental variables was rather small and additionally composed of many parameters that exhibited a pronounced multicollinearity. Consideration of additional abiotic and biotic parameters describing top down or bottom up processes in more detail (nutrient availability, predation by grazers and lysis by viruses) might contribute to a better and more detailed explanation of the observed patterns. Second, microbial biogeography is not only driven by deterministic processes such as selection (i.e. adaptation to prevailing environmental conditions) but also by stochastic processes like dispersal and mutation as reported by Hanson et al. (2012). The authors argue that mutation might add noticeably to the compositional variability among different locations in particular, when considering highly variable genetic regions such as the intergenic spacer (IGS) region. However, to our knowledge there are no studies that focus on the effect of mutation on the variation of microbial biogeography. Third ARISA only captures the most dominant species (Sogin *et al.* 2006), therefore missing a huge amount of diversity. Since the relationship between environmental factors and rare or

dominant taxa might be different also according to their lifestyle (generalists vs. specialists), inferences on bacterial community variation based on ARISA fingerprints are limited.

Despite these drawbacks we were able to reveal clear patterns in the spatiotemporal community variation in the German Bight and to unravel possible driving mechanisms. To our knowledge this study is the first systematic investigation of the bacterioplankton community in the German Bight combining both, relatively fine spatial resolution and long-term scales. The results provide relevant new insights into the different driving mechanisms of the variation of the free-living and particle-attached bacterial community composition. We conclude that spatial variation within the German Bight is defined by pronounced hydrographic current conditions that separate the inner German Bight from the central North Sea and thus, may lead to dispersal limitation of the bacterioplankton community and distinct offshore and coastal populations. However, temporal influences are dominating over the spatial variation and seem to play a major role in community assembly. Temporal variation is triggered by temperature as the main driving force throughout the examined transect, and by underlying short-term events like phytoplankton blooms.

Acknowledgements

We would like to thank the crew of the Uthörn, Kristine Carstens, Silvia Peters and Matthias Friebe for technical assistance and help during sampling and DOC measurements. We gratefully acknowledge the provision of BSHcmod current velocity fields by the Federal Maritime and Hydrographic Agency of Germany (Bundesamt für Seeschifffahrt und Hydrographie, BSH, Hamburg) and the calculation of principal components of water currents by Mirco Scharfe.

Supplementary material

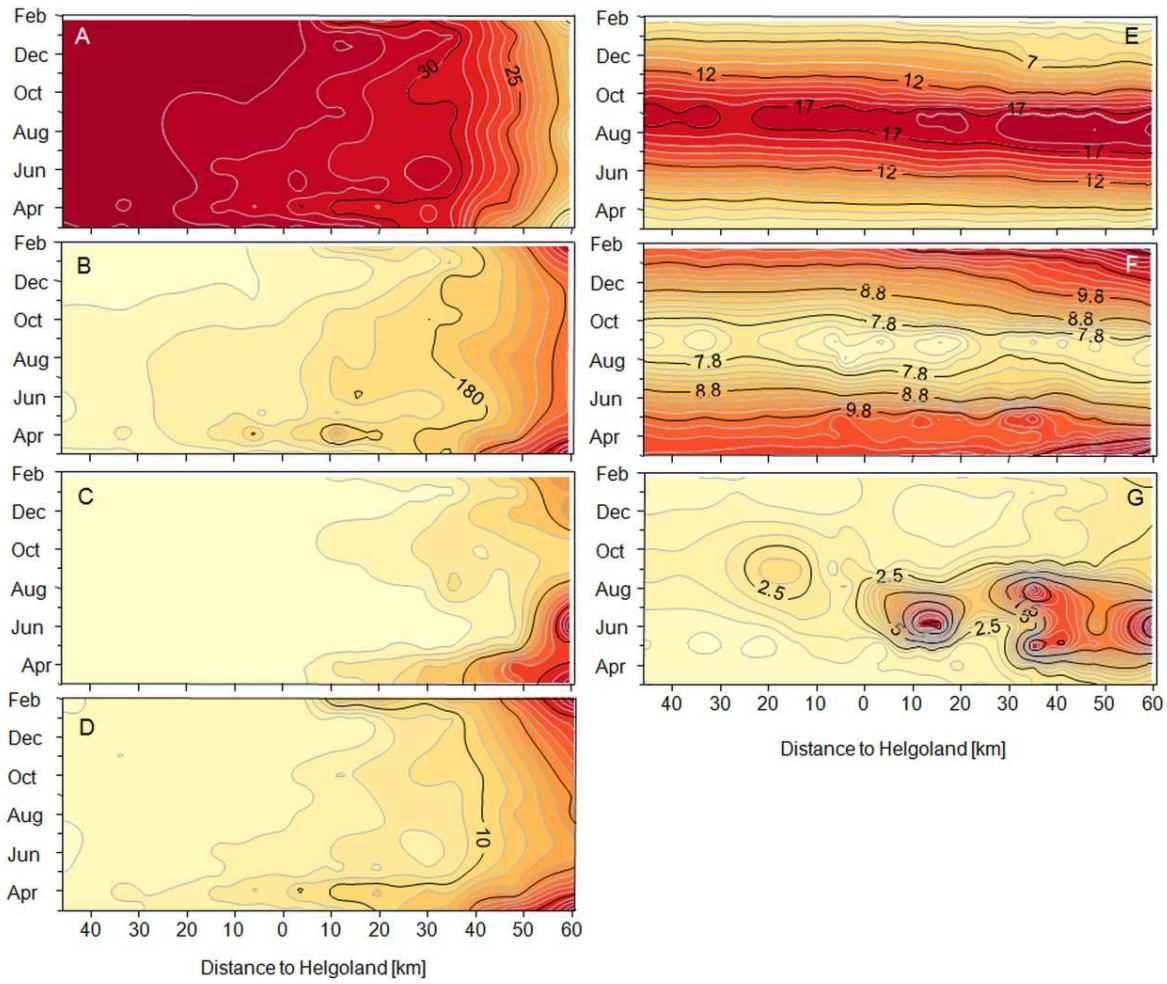


FIGURE S1: Contour plots of all measured environmental parameters. The horizontal axes depicts the distance [km] of sampling sites to Helgoland Island which was set to 0 km. Increasing distance to the left represents sampling sites along the P8 transect, increasing distance to the right represents sampling sites along the Elbe transect. The vertical axis refers to the sampling date; color code reflects measured values with of respective environmental parameters with yellow colours indicating lower values and red colors indicating higher values. A: salinity, B: dissolved organic carbon, C: turbidity, D: colored dissolved organic matter, E: temperature, F: dissolved oxygen, G: Chlorophyll *a*.

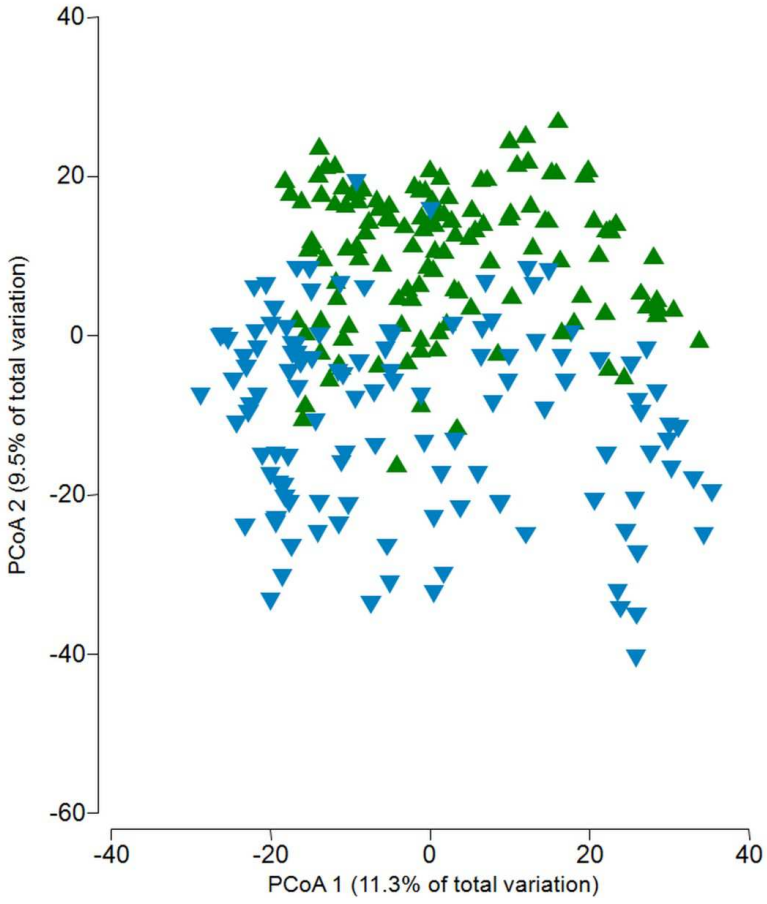


FIGURE S2: Principal component analysis (PCoA) of ARISA OTUs of the free-living and particle-attached fraction based on Jaccard index. Green triangles depict free-living fraction, blue triangles indicate particle-attached fraction.

Assessment of hydrodynamic variability in the German Bight

Hydrodynamic variability in the German Bight was assessed using current velocity fields from the model BSHcmod (Dick *et al.* 2001) operated by the Federal Maritime and Hydrographic Agency of Germany (Bundesamt für Seeschiffahrt und Hydrographie, BSH, Hamburg). First, current velocities of high temporal resolution (15 minutes) were averaged to obtain weekly mean vector (i.e. u, v) fields for the period March 2012 to February 2013. Second, Empirical Orthogonal Function (EOF) analysis (von Storch & Zwiers 1999) was applied to identify dominant modes of spatially coherent variability in these current patterns. These EOFs reflect anomaly patterns with regard to the mean current conditions for the selected period with the first EOF covering the highest amount of explained variance in the simulated transport fields. The explained variance of the two leading EOFs is more than 85% (Fig. S3). The corresponding principal components PC1 and PC2 (Fig. S4) provide information about the sign and the amplitude of the EOFs as a function

of time. The vector fields shown by the EOFs represent weighting factors (loadings) which are used for mapping each weekly mean current field to one data point of the corresponding principal component time series. See Callies & Scharfe (2015) for a comparable analysis based on the decadal scale.

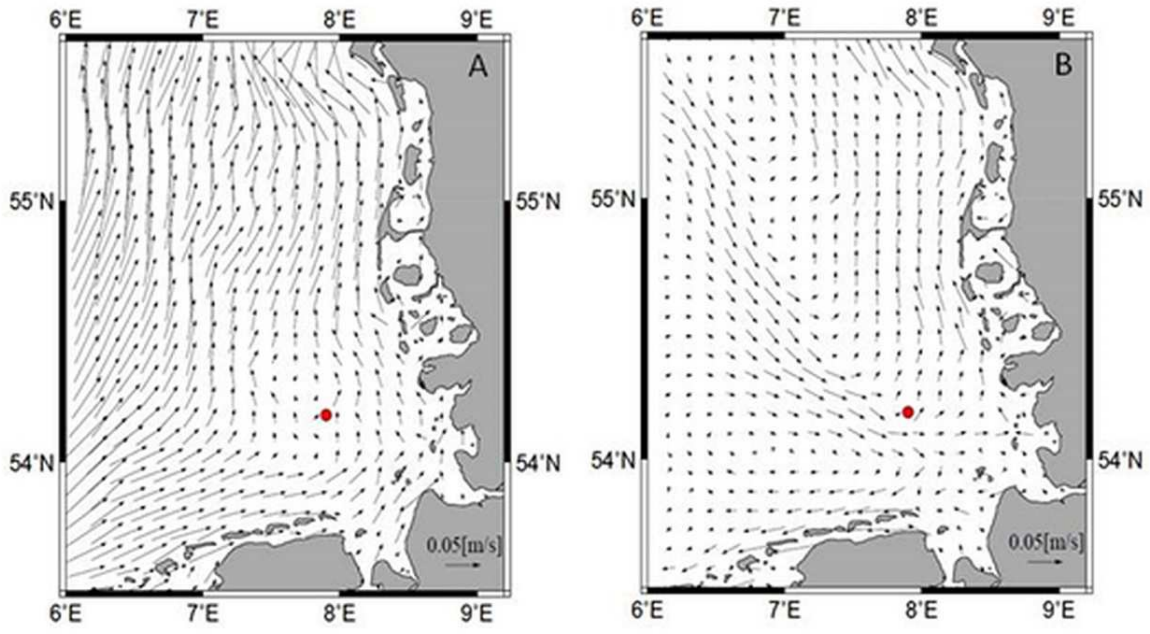


FIGURE S3: Vector fields of current anomalies (EOF pattern) in the German Bight within the period March 2012 – March 2013. Explained variances are 73.4 % for the first (A) and 12.2% for the second (B) EOF. Red dots: Helgoland.

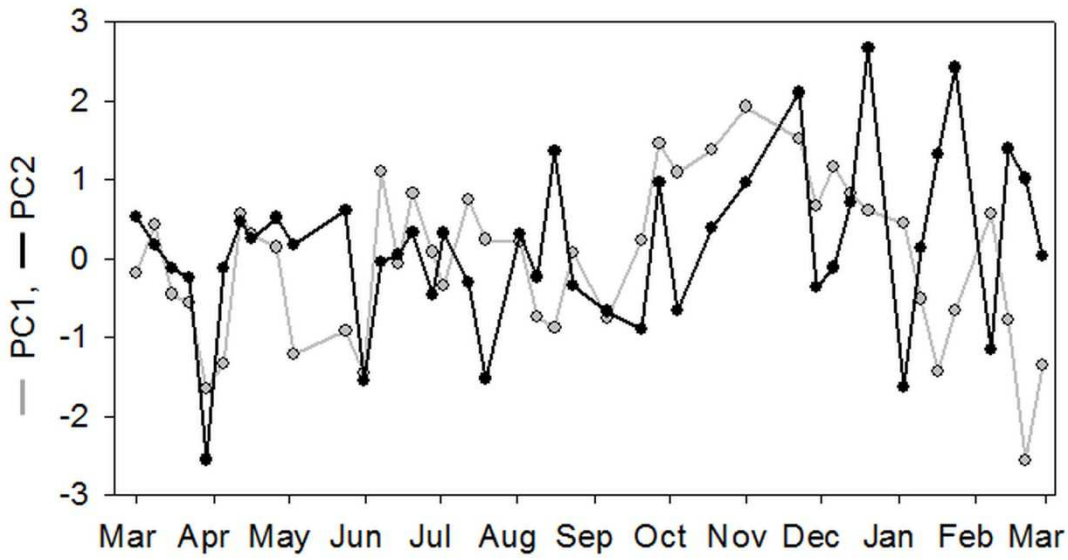


FIGURE S4: Principal components (PCs) corresponding to the EOF pattern shown in Figure S2.

Table S1: Permutational analysis of variance (PERMANOVA). PERMANOVA main test of bacterial community composition was based on Jaccard dissimilarities of ARISA profiles. Main test of environmental parameters was based on Euclidean distances. P-values were obtained using type III sums of squares and 9999 permutations under the full model. df: degrees of freedom, SS: sums of squares, perms: number of unique permutations. All tests were done on a significance level of $p<0.05$; significant values are indicated in bold.

PERMANOVA	Parameter	df	SS	Pseudo-F	*p(perm)*	perms
free-living vs particle-attached	fraction	1	35949	17.13	**0.0001**	9878
	res	247	518490			
	total	248	554440			
P8 vs Elbel free-living	transect	1	11285	6.20	**0.0001**	9904
	res	125	227690			
	total	126	238980			
P8 vs Elbe particle-attached	transect	1	16175	7.37	**0.0001**	9910
	res	120	263340			
	total	121	279510			
P8 vs Elbe environmental	transect	1	148	25.17	**0.0001**	9947
	res	125	734			
	total	126	882			

Table S2: Tests of homogeneity of dispersion (PERMDISP). PERMDISP was performed on the basis of Jaccard dissimilarities of ARISA profiles for the bacterial community and on the basis of Euclidean distances for environmental parameters. P-values were obtained using 9999 permutations and tests were performed on a significance level of $p<0.05$; significant values are indicated in bold. N: Number of samples, Average: average distance to the group centroid on the scale of the chosen resemblance measure, SE: standard error for the distance to the group centroid.

PERMDISP	N	Average	SE	F	*p*(perm)
free-living bacteria	127			15.493	**0.0005**
P8	45	39.592	0.78814		
Elbe	82	43.322	0.55304		
particle-attached bacteria	122			15.611	**0.0002**
P8	44	43.395	0.83585		
Elbe	78	47.592	0.64348		
environmental	127			40.02	**0.0001**
P8	45	1.2133	0.082857		
Elbe	82	2.508	0.14442		

GENERAL DISCUSSION

After recognizing and understanding the fundamental impact of marine microbes on energy fluxes, all known biogeochemical cycles and thus, ecosystem functioning in the early 90s, microbiologists began to unravel biogeographic patterns among a wide range of spatial and temporal scales and various habitats. The existence of these patterns is now well established (Hanson *et al.* 2012), but to predict and estimate potential ecological impacts that arise from changes in these patterns, it is of utmost importance to unveil the mechanisms that are generating and maintaining them. Different underlying processes (selection, drift, dispersal and mutation) that drive microbial biogeographic patterns have been proposed but the majority of studies on bacterial community dynamics report that variation most likely depends on the adaptation of taxa to specific environmental conditions (selection). Specific combinations of main environmental driving forces have been revealed for individual habitat types such as coastal areas and estuaries or pelagic oceanic regions (e.g. Martiny *et al.* 2006; Gilbert *et al.* 2009; Fortunato *et al.* 2012; Chow *et al.* 2013).

The German Bight in the southern North Sea constitutes a highly dynamic unique coastal environment, influenced by the world's largest tidal flat area, a noticeably freshwater influx by rivers and anthropogenic activities, making it a perfect study site to investigate how changes in environmental parameters affect the bacterioplankton community variation. Changes in environmental parameters (temperature, salinity and nutrients), the biota and current patterns have been continuously monitored for more than five decades around Helgoland Island in the German Bight (54°11.3' N, 7°54.0' E), known as the Helgoland Roads time series (Wiltshire et al. 2008). In the frame of this comprehensive long-term data series, the temporal variability of the bacterioplankton community composition at Helgoland Roads has been well described. Previous studies investigated the seasonal variability by using different molecular biological approaches like DGGE, RISA (Sapp et al. 2007), and CARD-FISH (Alderkamp et al. 2006). However, these investigations were done with limited taxonomic and temporal resolution and detailed information on the succession of specific bacterial taxa and the temporal scale on which this succession is taking place is missing. Short-term variation in response to a spring phytoplankton bloom has been studied in a comprehensive metagenomic and –proteomic study (Teeling *et al.* 2012). The authors demonstrated strong dependency of the taxonomic and functional succession of bacterial taxa on changes of the organic matter composition

during the course of the phytoplankton bloom. However, currently new analytical high resolution methods like FT-ICR-MS provide detailed information on the molecular composition of the organic matter. Usage of such methods potentially enhances knowledge on the bacteria-DOM interactions and gives new insights into the regulation of bacterial communities during enhanced substrate supply as expected during phytoplankton blooms.

Marine habitats represent continuous, highly connected environments where changes in bacterial communities are complex and comprise temporal (succession) and spatial (dispersion) components simultaneously. Therefore, spatiotemporal studies are crucial to understand the mechanisms that generate and maintain patterns in microbial community assembly. Understanding these mechanisms will help to estimate the impact of bacterial community variation on the functioning of biogeochemical cycles and thus, eventually contribute to the prediction of potential changes in ecosystem functioning as response to environmental change. Until today, only very few studies focused on both, temporal and spatial microbial community dynamics at the same time (Morris *et al.* 2005; Hewson *et al.* 2006; Fortunato & Crump 2011; Fortunato *et al.* 2012) and thus, knowledge on spatiotemporal patterns – particularly in highly dynamic environments like coastal areas and estuaries – is scarce. As described above, studies on the bacterial community in the German Bight rather focused on the temporal variation and so far, no spatiotemporal analyses of the bacterioplankton community in the German Bight have been conducted.

This thesis aimed at improving the understanding of bacterioplankton community variation in the German Bight by providing new perceptions on the above highlighted knowledge gaps. In Chapter I, detailed insight into the annual succession of specific bacterial OTUs at Helgoland Roads, gained via 16S rRNA gene based high throughput sequencing and linkage to environmental conditions, is demonstrated. The results complement the findings of previous studies that focused on the seasonality of broad bacterial groups on class level (Sapp *et al.* 2007). Moreover, in Chapter II, linkage of the bacterial community variation at Helgoland Roads with molecular DOM composition on short-term scales (day to day variation) was presented for the first time. This approach is a first step towards the identification of relationships between potential substrate specialists and particular DOM molecules. Regarding bacterial response to enhanced organic substrate supply, this may improve the understanding of taxonomical and functional succession during phytoplankton blooms and thus, enlarge the knowledge already gained during a metagenomic and-proteomic study (Teeling *et al.* 2012). A more widespread view on the bacterioplankton community variation in the German Bight was given in Chapter III. Here, the

spatiotemporal dynamics of bacterioplankton communities in the German Bight and the main regulating forces are described for the first time. Multivariate statistical techniques enabled the deconvolution of spatial and temporal signals.

In the following sections selected issues presented in Chapters I to III are discussed in a general context. In particular, the observed temporal variation of the bacterial community in the German Bight is discussed in the light of different existing time scales in the section "Notes on different temporal scales in microbial biogeography". The results from the spatiotemporal analyses are compared to results from two other spatiotemporal studies in coastal environments. As temperature was identified as major driving force for long-term and spatiotemporal patterns in the German Bight, its ecological relevance as influencing factor is reviewed in a separate section ("The role of temperature as driving force for bacterial community variation"). Moreover, the different methodological approaches used during the course of this thesis (multiple regression analysis, ARISA, 16S rRNA gene tag sequencing) are elucidated with regard to their reliability and implications for the ecological interpretation of the data in the sections "The issue of multicollinearity in microbial ecology" and "ARISA versus 16S rRNA gene tag sequencing". Finally, still existing knowledge gaps are highlighted and suggestions for future research that might fill these gaps are given in the section "Future perspectives".

Notes on different temporal scales in microbial biogeography

The temporal variation of microbial communities has been explored intensively during the last two decades in numerous habitats and on a huge range of scales. Recurrence of community assemblies over multiple years, seasonal shifts of bacterial community composition and short-term responses to phytoplankton blooms are well evidenced for various oceanic environments (e.g. Fuhrman *et al.* 2006; Gilbert *et al.* 2009; Caporaso *et al.* 2012; Teeling *et al.* 2012). However, the mechanisms that shape the microbial assemblages are still not understood properly, as most of the studies revealed different driving forces (temperature, day length, nutrient concentration, etc.) in individual habitats. Martiny *et al.*, (2011) even stated that the identified drivers for betadiversity variation within one ecosystem depend on the scale one is looking at. This scale-dependency of environmental driving forces is supported by the observations on the bacterioplankton community in the German Bight on different temporal scales. During the annual surveys presented in Chapters I and III, it was demonstrated that the seasonal variation of the bacterioplankton community at Helgoland Roads was dominated by temperature as a main

driving force. Short-term variation on weekly and daily scales was influenced by enhanced organic matter supply during phytoplankton blooms and intermittent change of water masses as evidenced in Chapters I and II.

However, the typical average generation times of marine plankton are approximately a day in surface waters (Ducklow 2000), thus, it appears reasonable to expect diel changes in bacterial community on hourly scales. Indeed, Hewson *et al.* (2006) found up to 20 % change per day in bacterial community composition in oligotrophic waters during drifter studies. Mével *et al*. (2008) demonstrated a rapid response of bacterial biomass to changes in chlorophyll *a* concentrations on an hourly time scale in the Mediterranean Sea and bacterioplankton abundance in the subtropical Atlantic Ocean followed a distinct diel periodicity (Kuipers *et al.* 2000).

To account for this temporal fine-scale variability, the bacterioplankton community composition at Helgoland Roads was additionally investigated on an hourly scale over 24 hours in the context of Chapter II. From principle coordinates analysis it is obvious that community composition exhibited high hourly variability (Fig. 1A). However, these rapid changes did not follow any structured patterns and analysis failed to significantly relate the variation to any of the measured environmental parameters. Since the German Bight and waters around Helgoland Roads are highly dynamic in terms of hydrographic conditions a possible relation between bacterial community composition and tidal cycles was assumed (Fig. 1B), but no significant relationship between both was found. Similar observations have been made by Mével *et al.* (2008), who stated that particle-attached bacteria in the Mediterranean Sea followed rapid and sporadic changes at hourly time scales. Thus, it is assumed that the hourly variation in bacterial community composition at Helgoland Roads is either influenced by variables or processes that have not been measured or, more likely results from unpredictable, stochastic processes which include random changes in species diversity that are not driven by environmental conditions. This is appealing, since these fine-scale stochastic processes might underlie every observed broad-scale pattern and thus, can contribute substantially to the variation of bacterial communities. Recent studies propose frameworks and conceptual models for disentangling deterministic from stochastic processes (e.g. Chase & Myers 2011; Dini-Andreotea *et al.* 2015). However, it is still not known to what extend these components contribute to the mechanisms regulating the bacterial biogeography.

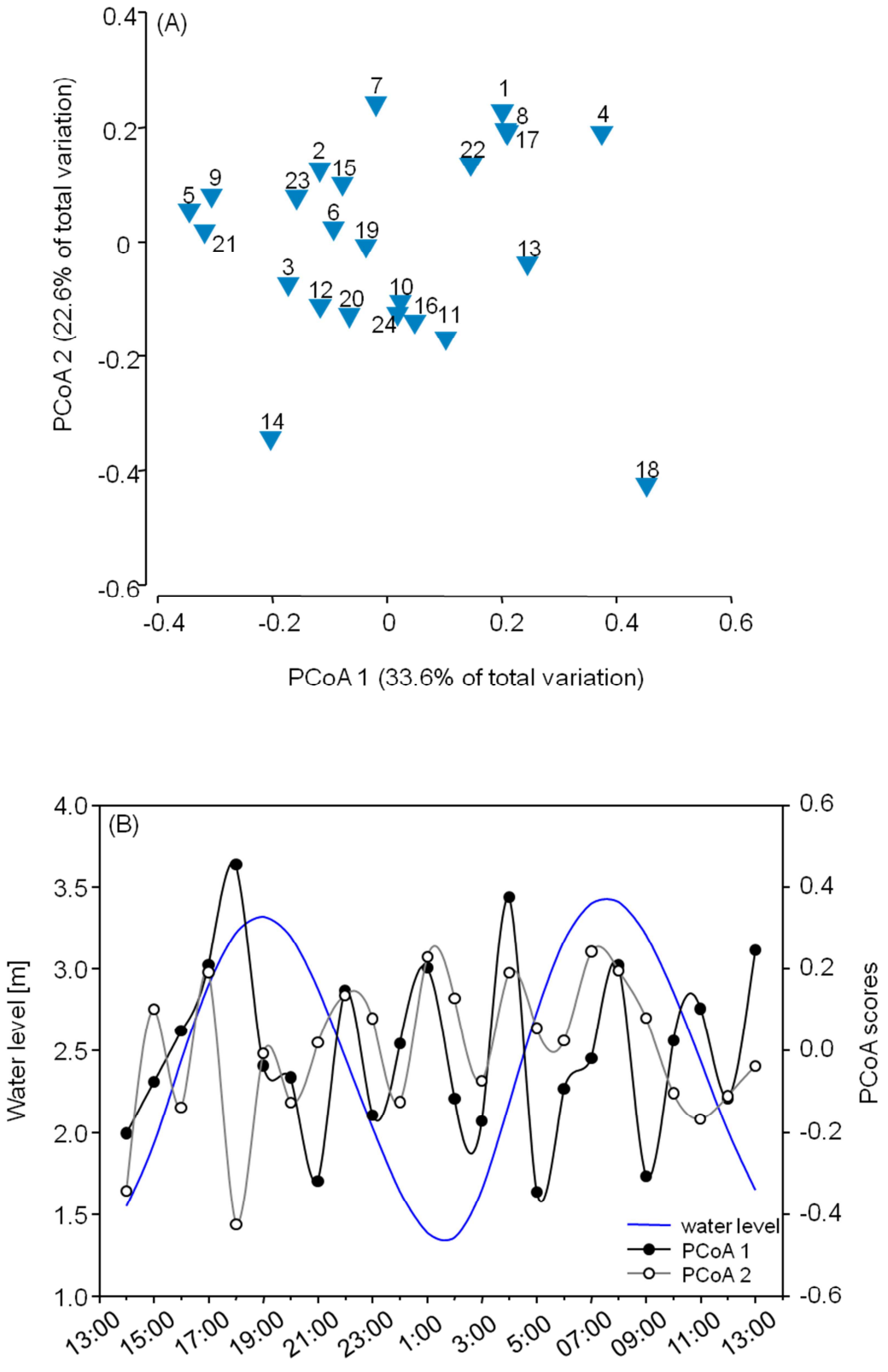


FIGURE 1: Hourly variation of the bacterial community composition at Helgoland Roads. (A) Principal coordinates analysis (PCoA) based on Hellinger distance of 16S rRNA gene tag sequencing data. Each symbol represents the community at a certain sampling time point. Numbers display the time (hour) of day when the community was sampled. (B) Water level changes and PCoA scores of bacterial communities are given over the course of one day on hourly resolution.

Spatiotemporal variation of the bacterioplankton community

Research on microbial biogeography has experienced growing interest during the last two decades and greatly broadened the knowledge on the variation of microbial communities in numerous habitats on a huge range of temporal and spatial scales (Martiny *et al.* 2006; Ramette & Tiedje 2007; Fuhrman *et al.* 2015). However, the majority of biogeographical studies focused on either temporal or spatial community variation. Only very few studies dealing with the community variation patterns in coastal environments combined both components in spatiotemporal approaches (Morris *et al.* 2005; Hewson *et al.* 2006; Fortunato *et al.* 2012). Chapter III of this thesis provides - for the first time - insights into the spatiotemporal variation of the bacterioplankton community in the German Bight. A main outcome of the study is the dominance of seasonal variation patterns over spatial community variation. Furthermore, differences in community composition among coastal and offshore sites, mainly driven by hydrographic currents, were demonstrated.

The results are partly consistent with the findings of Fortunato *et al.* (2012) and Du *et al.* (2013) who also reported a separation of coastal and offshore communities along the Columbia river coastal margin and the coastal South China Sea respectively. Fortunato *et al.* (2012) emphasized that an increase of the sample number from 71 to 300 facilitated a separation of the bacterial community into seven distinct habitats along the river coastal margin. Concerning the comparatively small number of samples (127) obtained during this thesis, it might be assumed that a higher spatial and temporal sampling frequency might also unravel more distinct bacterial habitats within the German Bight. However, in contrast to this thesis, Fortunato *et al.* (2012) and Du *et al.* (2013) suggested that the spatial variability overwhelmed temporal variation patterns. There are different conceivable reasons that might lead to this discrepancy. First, Du *et al.* (2013) assumed that mixing of water masses in the investigated coastal region was partly blocked by the landmass of an island. This scenario might lead to strong dispersal limitation which in turn leads to pronounced spatial variation patterns. In contrast, water masses within the relatively shallow German Bight are very well mixed due to tides and wind forces and thus, theoretically support dispersal processes. Second, different methodological approaches have been applied during this thesis and the mentioned studies. For instance, Fortunato *et al*. (2012) not only focused on horizontal variation patterns, but also investigated vertical profiles of the bacterioplankton community, which might strengthen the observed spatial pattern. The vertical variation in the German Bight however, is assumed to be of minor importance due to the afore-mentioned shallow depth and strong mixing of water masses.

Furthermore, in Chapter III of this thesis free-living and particle-attached bacteria where investigated separately. It is demonstrated, that the free-living bacteria are predominantly influenced by temperature, whereas the particle-attached fraction mainly responds to environmental parameters that exhibit strong spatial gradients such as DOC for instance. Fortunato *et al.* (2012) however, considered the whole bacterial community during their analyses, which did not allow for the deconvolution of variation patterns of the two bacterial fractions. Thus, the variation patterns of the particle-attached fraction might have contributed to the strong spatial signal observed in that study.

Regarding the contrasting results gained during the different spatiotemporal studies it becomes apparent that regulating mechanisms that underlie the variation patterns of bacterial communities are not only influenced by varying environmental conditions, but also strongly depend on the geographic and hydrodynamic conditions. Thus, especially in continuous aquatic environments, hydrodynamic conditions should be taken into account during spatiotemporal future studies.

The role of temperature as driving force for bacterial community variation

The existence of microbial biogeographic patterns is well established and research now focuses on unravelling the mechanisms that generate and regulate these patterns. A huge range of physico-chemical (e.g. temperature, salinity), geochemical (e.g. sediment and elemental composition) and biotic factors (e.g. phytoplankton composition, species interactions) is supposed to potentially regulate bacterial community composition in the world's oceans. Studies on the bacterial community composition in various oceanic habitats identified numerous different combinations of main driving forces such as salinity and depth (Fortunato *et al.* 2012), day length (Gilbert *et al.* 2012), chlorophyll *a*, bacterial production rates and nutrient concentrations (Chow *et al.* 2013), suspended particles (Wang *et al.* 2015) or organic pollutants (Störmer *et al.* 2013), just to mention a few of them. Within Chapters I and III of this thesis the prominent influence of temperature on the seasonal variation of bacterial community structure in the German Bight is demonstrated. The results presented in these Chapters let me assume that temperature predominantly controls the stability of the bacterial community composition, i.e. the annual succession and the proposed recurrence, whereas changes in other factors, such as enhanced DOM supply, result in intensive, rapid, short-term variation of the bacterial community, underlying the temperature-controlled pattern. However, with regard to the various ecological interactions exhibited by different environmental parameters, this assumption

might be oversimplified. Thus, the ecological importance of temperature is described in more detail in the following paragraph.

Ecological studies on the diversity of various terrestrial, freshwater and marine organisms reported a general decline of biodiversity with latitude (Hillebrand 2004 and references therein). Among many different hypotheses, two main processes generating this decline are discussed these days. First, increasing productivity in higher latitudes provides more resources which result in a larger number of ecological niches. More niches support a higher number of specialists and thus, a higher diversity (e.g. Connell & Orias 1964). Second, increasing temperature favours kinetic processes such as reproduction, metabolism, species interactions, adaptation etc. which lead to higher diversity (e.g. Brown *et al.* 2004). Fuhrman *et al.* (2008) stated that many of the species-rich tropical oceanic environments exhibit high temperatures but relatively low productivity, thus, the hypothesis that kinetic mechanisms play a primary role for species diversity patterns might be favoured when considering marine environments. This is supported by studies investigating the variation of marine bacterial communities in marine habitats on latitudinal gradients that indeed found temperature as dominant driving force for this variability (Pommier *et al.* 2007; Fuhrman *et al.* 2008; Yilmaz *et al.* 2012). Temperature was also found to determine the temporal variation of marine bacterial community dynamics (e.g. Fuhrman *et al.* 2006; Gilbert *et al.* 2009). In consistence with the findings of this thesis a multiannual study by Chow *et al.* (2013) also reported a significant influence of temperature on the seasonality of bacterial community structure. However, studies on the short-term variation of marine bacterial communities as response to phytoplankton blooms assume, that the enhanced substrate supply constitutes the dominant driving force for community variation (e.g. Teeling *et al.* 2012; Wemheuer *et al.* 2014). It is known that the bacterial metabolic processes, such as the decomposition of organic matter, are enhanced by increasing temperature (Pomeroy & Wiebe 2001; Kirchman *et al.* 2009). Hoppe *et al*. (2008) and Wohlers-Zöllner *et al.,* (2012) conducted mesocosm experiments with natural sea water containing a spring phytoplankton community. They incubated at *in situ* and increased temperature conditions and showed that, increasing water temperatures lead to a tighter coupling of phyto- and bacterioplankton. Using similar mesocosm experiments, von Scheibner *et al.* (2014) demonstrated an enrichment of bacterial communities with additional taxa when exposed to increased temperatures thus, emphasizing the dominance of temperature as regulating factor for communities over substrate supply. Nevertheless, there is clear evidence that dissolved and particulate

organic matter released by phytoplankton strongly shape the bacterial community composition, but it has been suggested that in highly dynamic systems such as the German Bight, the influence of primary producers on microbial dynamics is less important compared with that of abiotic factors (Kirchman *et al.* 2005; Teira *et al.* 2008). Despite the mentioned facts and studies that demonstrate the dominant role of temperature in bacterial community assembly, it should be noted that temperature tends to exhibit strong correlations with various other parameters that might lead to misinterpretation of results. The problems of such multicollinearities and possible ways to deal with it are discussed in the following paragraph ("The issue of multicollinearity in microbial ecology"). However, when identifying temperature as a leading regulation factor, it should always be kept in mind, that it might serve as a proxy for many other ecological processes such as internal competition, predator-prey relationships, viral lysis etc. These temperature driven processes might serve as 'internal feedback mechanisms' (Fuhrman *et al.* 2015) which maintain the stability of the bacterial communities on the long run.

The annual mean surface water temperature in the German Bight has increased by ~1.7 °C since 1962 (Wiltshire *et al.* 2010) and water temperature is expected to increase further due to global warming. The question is, what implications does global warming – or global change in general - have on marine bacterial community composition and related ecosystem functioning? Some thoughts concerning this question are given in the following paragraph.

As evidenced in this thesis environmental changes lead to changes in bacterial community composition. Research on microbial communities now reaches a status that allows to even predict, how these changes in environmental factors will affect the microbial community composition (Fuhrman *et al.* 2006). However, a fundamental issue in microbial ecology is how variation in bacterial community structure is linked to changes in ecosystem functioning. Contradictory results have been obtained regarding this issue. On the one hand, it has been shown that changes in microbial community composition and diversity may also affect microbial-mediated ecological function in freshwater environments (Langenheder *et al.* 2005; Szabó *et al.* 2007). In contrast, studies that linked temporal and spatial patterns of bacterial community composition with enzyme activities suggest that differing drivers for community structure and function, as well as functional redundancy among different bacterial taxa, lead to a disconnection between microbial community structure and ecological function (Frossard *et al.* 2012; Purahong *et al.* 2014). Hence, the impact of variation in bacterial community composition on its functional properties is still

not properly understood and thus, it remains difficult to estimate how environmental changes will affect microbial-mediated ecosystem functioning in marine systems.

The issue of multicollinearity in microbial ecology

The vast majority of microbial ecological studies, including this study, use multiple regression analyses to describe the regulation of bacterial communities by a set of environmental parameters. But only few studies consider the issue of collinearity - also referred to as multicollinearity, i.e. the explaining environmental parameters are correlated among each other and not independent from each other. Thus, the effects of two collinear parameters on the variable that is to be explained cannot be separated, because there is a shared proportion in their effects. Ignorance of multicollinearity might lead to biased, or at worst, erroneous interpretation of results (Graham 2003).

Multiple regression analysis was frequently used during the course of this thesis. Confounded explaining variables threatened the unbiased statistical and ecological interpretation several times. For instance, in Chapter I the annual variation of the bacterial community composition along two PCoA axes was explained by a multiple regression model. Here, strong correlations of temperature and dissolved organic nitrogen (DIN) were observed. To avoid erroneous multiple regression analysis, shared and unique contributions were disentangled by using residual and sequential regression as proposed by Graham (2003). Based on ecological background knowledge it was assumed that temperature was functionally more important than DIN. Therefore, DIN was regressed against temperature and original DIN values were replaced by the residuals from the regression. Although subsequent regression analysis was unbiased as collinearity was removed, this approach might still lead to erroneous interpretation, since the choice of the variable being functionally more important depends on the researcher's ecological knowledge or literature reports (Graham 2003). A more objective way of dealing with the problem of multicollinearity was chosen in Chapter III. Here, multicollinearity was removed prior to multiple regression analysis by performing a principal component analysis (PCA) with the environmental parameters. The scores of the PCA axes were then used as explaining variables during multiple regression analyses as proposed in Tabachnick and Fidell (1996) and Legendre and Legendre (1998). The big advantage of this approach is that PCA axes are orthogonal, or independent from each other and thus, collinearity is perfectly removed from the data set (Dormann *et al.* 2013). However, the drawback is that results of multiple regression analysis are difficult to interpret, because PCA axes represent a linear

combination of the different variables. Observation of the loadings of the individual variables on the PCA axes however, may help during interpretation.

Other ways of dealing with multicollinearity are reviewed in Graham (2003), Smith *et al.* (2009) and Dormann *et al.* (2013). The easiest way might be, to simply drop collinear variables from analysis, however this might lead to substantial loss of explanatory power (James & McCulloch 1990). Summarizing the mentioned problems Dormann *et al*. (2013) stated that collinearity is a known problem in statistics and it cannot be solved, but there are multiple ways how it can be explored and considered during interpretation of results. In any case, the techniques for reduction of multicollinearity used in this thesis substantially enhanced the reliability of results and their interpretation.

ARISA versus 16S rRNA gene tag sequencing

The composition, richness, diversity and variation of microbial communities are generally described by a vast range of molecular biological tools, among which fingerprinting techniques and sequencing approaches, both of which were used for community description during the course of this thesis, have been commonly used in recent studies. The big advantage of these methods is that they do not require cultivation, but provide genetic diversity measures as they are based on PCR amplification of certain target genes. Fingerprint methods like terminal restriction fragment length polymorphism (tRFLP) or denaturing gradient gel electrophoresis (DGGE) often target the highly conserved 16S rRNA gene which serves as a perfect phylogenetic marker as it is ubiquitous, functionally constant, highly conserved and long enough for detailed analyses of phylogenetic relationships. In contrast, automated ribosomal intergenic spacer analysis (ARISA) uses amplicons of the intergenic spacer (IGS) region, which is embedded between the 16S and 23S rDNA and exhibits a significant heterogeneity in length and nucleotide sequence between different bacterial taxa. In general, ARISA has been found to provide higher resolution of phylotypes than tRFLP and DGGE (Danovaro *et al.* 2006; Okubo & Sugiyama 2009). Further advantages are the quick and easy application, high sample throughput and cost-efficiency. Nevertheless, there has been discussion about the reliability of ARISA based community analysis as it is prone to a number of biases. For instance, it has been noticed that ARISA only targets the most abundant (> 1%) community members (Casamayor *et al.* 2000; Sogin *et al.* 2006). A further common concern, valid for all amplicon dependent approaches, is the erroneous amplification during PCR (e.g. Lueders & Friedrich 2003). Moreover, it appears, that bacterial taxa

might contain more than one 16S operon and thus, possess different IGS sequences or that resulting bands display multiple sequences (Fisher & Triplett 1999; Ranjard *et al.* 2000). Summed up assumptions on ARISA analyses are limited to the detection of broad patterns in community structure. Today, more accurate, detailed descriptions of bacterial richness, community structure, and variation therein are frequently done by using high resolution methods like high throughput sequencing approaches of 16S rRNA genes. However, it has been reported that results on bacterial community structure yielded via ARISA and high throughput sequencing are comparable (Shade *et al.* 2012; Lee 2014).

The short-term dynamics of bacterial communities at Helgoland Roads described via Illumina 16S rRNA gene tag sequencing in Chapter II of this thesis have been additionally assessed using ARISA fingerprints (Fig. 2). PCoA of ARISA fingerprints (Fig. 2A) exhibit a similar pattern and lead to the same group formation when compared to patterns of the whole or trimmed (OTUs >0.1% of the total sequence data) 16S tag sequencing data based on Hellinger distance (Figs. 2B and 2C). Moreover, the explained variations along the first two PCoA axes are comparable (ARISA: 52.1%, whole 16S tag sequencing data: 55%, trimmed 16S tag sequencing data: 59.2%). When using Hellinger distance, the differences for abundant species in the whole data set contribute more to similarity between sites than differences between rare species (Legendre & Legendre 1998). As already mentioned above, ARISA only targets the most abundant taxa (> 1%). Thus, both approaches (PCoA of ARISA fingerprints based on Jaccard index and PCoA of sequencing data based on Hellinger distance) lead to a stronger weighting of abundant taxa which then again leads to similar results during ordination. The fact, that results change only little when using the trimmed data set of 16S tag sequencing demonstrates, that rare taxa have only limited impact on the analysis of the community structure in terms of diversity, i.e. the presence and proportion of different taxa. In contrast they do have massive impact on the community richness. Therefore, PCoA of the 16S tag sequencing data based on Jaccard index, which is a qualitative similarity measure and only accounts for presence or absence of taxa, yields totally different results (Fig. 2D).

Taken together, this comparison highlights that the combination of the method for community description and the chosen resemblance measure might have strong implications on ordination results and thus, need to be selected thoroughly. ARISA allows for a sensitive and reproducible description of community diversity and differences in community structure (Lear *et al.* 2008; Kovacs *et al.* 2010; Meziti *et al.* 2010), whereas

detailed description of bacterial richness need to be performed using sequencing techniques.

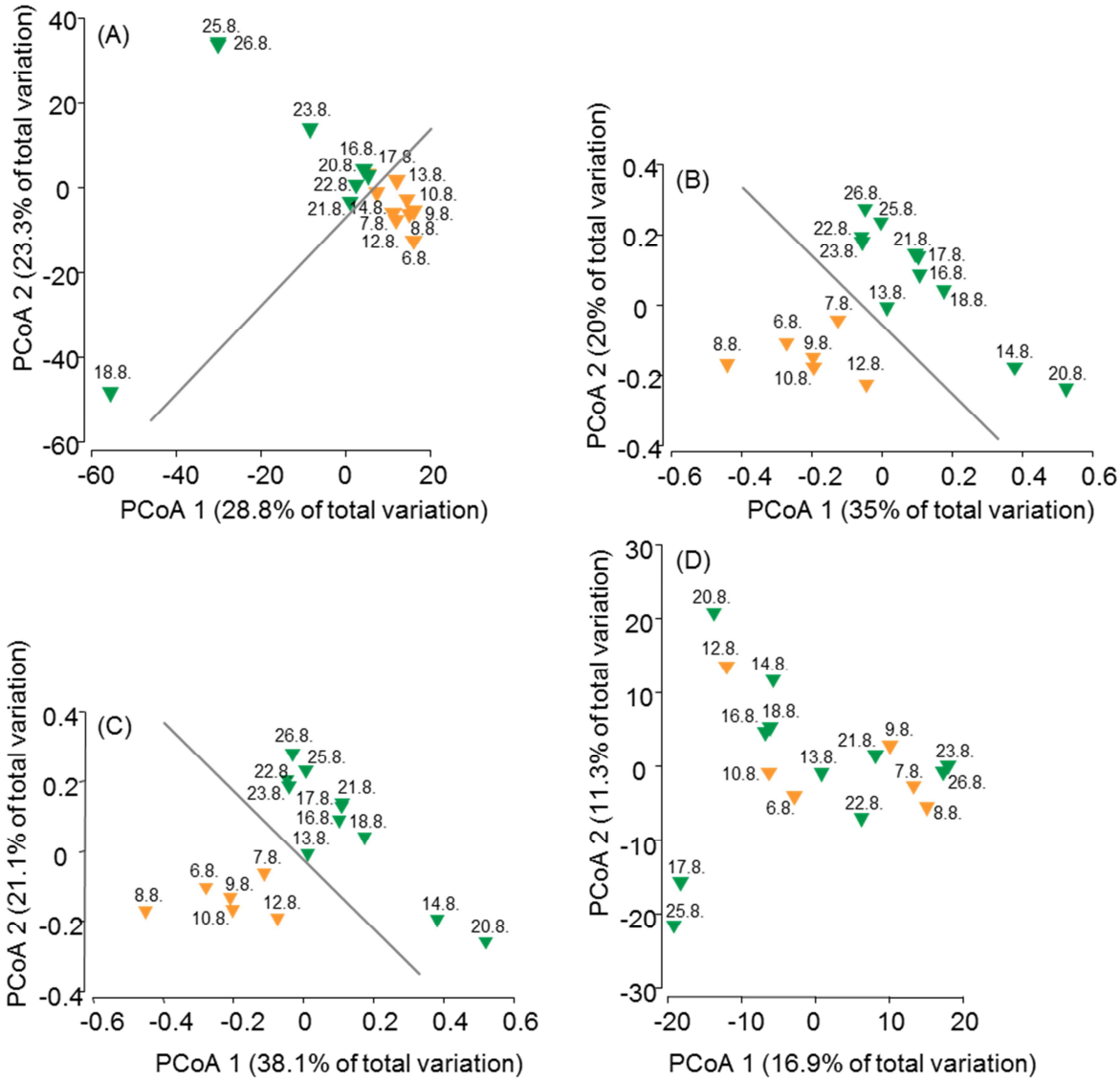


FIGURE 2: Comparison of principal coordinates analyses (PCoA) of ARISA fingerprints and 16S rRNA gene tag sequencing. (A) ARISA fingerprints based on Jaccard index, (B) 16S tag sequencing data based on Hellinger distance for the whole and (C) the "trimmed data set" (only OTUs >1% of total sequence data), (D) PCoA patterns of 16S tag sequencing data based on Jaccard index. Grey lines indicate separation of samples in two distinct groups.

Future perspectives

In its entirety this thesis conveys a nice impression of the spatial and temporal variability and the underlying driving forces of the bacterioplankton community in the German Bight. Chapters I – III comprise a comprehensive descriptive approach which –no doubt- provides a profound basis for hypothesis-driven research and ecosystem modelling.

Chapter I focused on the seasonal variation of the bacterial community at Helgoland Roads. Results obtained during that study hint at annual recurrence of the bacterial community. Multiannual studies on the basis of 16S rRNA sequencing will help to evidence recurrence of the community and also allow for assumptions on the homo- or heterogeneity, i.e. interannual variability, of the environmental conditions at Helgoland Roads. Referring to the investigations of Fuhrman *et al.* (2006), multiannual studies may help to predict the community composition at Helgoland Roads based on environmental conditions.

As discussed before, there is evidence that variation patterns in bacterial community structure and functionality are disconnected and driven differently (e.g. Gutknecht *et al.* 2006; Purahong *et al.* 2014). To estimate implications of environmental changes on ecosystem functioning, future research should consider analysing the active community or investigate the variability among specific functional groups. There is already one study that investigated the active bacterioplankton community in the North Sea during a spring phytoplankton bloom using pyrosequencing-based analysis of 16S rRNA gene amplicons generated from environmental RNA (Wemheuer *et al.* 2014). Another excellent study by Teeling *et al*. (2012) described the taxonomic succession within the bacterioplankton community at Helgoland Roads as response to a spring phytoplankton bloom and related this with the expression of carbohydrate-active enzymes. Thus, there is substantial insight into functional changes of the bacterioplankton community on short time scales. Nevertheless, although important to estimate influences of global changes, such as climate warming or ocean acidification, knowledge on long-term and spatial variability of the communities' functional traits and their regulating forces is still lacking.

During this thesis the focus was set on species sorting, i.e. the environmental conditions determine the patterns in community composition by favouring specific taxa. These analyses were restricted to a number of environmental parameters that barely included biological parameters, such as data on ciliates or flagellates. Data on viruses were not considered at all. However, bacteria are part of a complex marine food web and are linked to protists and viruses via the microbial loop. Top-down processes like the consumption of

bacterial biomass by predators and viral lysis equal the bacterial production (Pernthaler 2005) and a variety of studies proved tight couplings between bacteria, protists and viruses. For example, it has been demonstrated that increased protozoan grazing results in shifts in the phenotypic and genotypic composition of the bacterial assemblage as reviewed in Jürgens and Matz (2002). Viral lysis has been shown to primarily affect the largest, most rapidly growing bacterial populations and by this control the abundances of particular bacterial species (Thingstad 2000). Although short-term influences of viruses and protists on the bacterial abundance and diversity have been well studied in various experiments, research on long-term impacts in natural communities has just begun. For example, Chow *et al.* (2014) demonstrated a significant relation between bacterial, protist and viral communities over a period of three years in Pacific surface waters. Likewise, Boras *et al.* (2015) studied these bacteria-protist-virus interactions in the Mediterranean Sea over two years. But still, knowledge on these long-term interactions in highly productive and dynamic coastal systems like the German Bight is lacking. In conjunction with this, microbial association networks are an interesting tool to display the interactions between several parameters simultaneously and allow for detailed analysis of co-occurring phylotypes, as for example presented in Fuhrman and Steele (2008).

In Chapter III of the current thesis, spatiotemporal analyses of the bacterioplankton community were done using ARISA fingerprints. The analyses suggested a separation of the bacterial community into coastal and offshore communities. To further characterize these coastal and offshore communities, to identify key taxa for the different habitats and to describe the variation patterns of these key taxa, spatiotemporal analyses based on sequencing approaches should be accomplished. The separation of different habitats along the Columbia River coastal margin and the identification of key taxa therein using 16S amplicon pyrosequencing were already demonstrated by Fortunato *et al*. (2012). In Chapter III it was also hypothesized that the spatial variation signal of the bacterial community composition in the German Bight along environmental gradients is not very pronounced because dispersal processes override environmental selection. Therefore, it would be appealing, to study the role of dispersal in the German Bight via experimental setups. An excellent example to do so is given by Lindström & Östman (2011) who studied the role of dispersal in structuring bacterial community composition and function of fresh water bacteria, by using dialysis bags to enable a decoupling between a change in the local environment and dispersal.

Final conclusion

The main purpose of this thesis was the detailed description of the temporal and spatial dynamics of the bacterioplankton community in the German Bight and to further understand, which environmental forces regulate these dynamics. The combination of the high sampling frequency with usage of high resolution techniques like 16S rRNA tag sequencing for the description of the bacterial community, as well as Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) for molecular analysis of dissolved organic matter (DOM), allowed for sensitive observations of bacterial community dynamics and related driving forces. This detailed descriptive approach contributes substantially to the understanding of the mechanisms that drive the bacterioplankton community variation in the German Bight and provides a fundamental basis for hypothesis-driven future research. Since bacteria exhibit rapid lifecycles and respond to environmental changes faster than other large organisms, changes in their biogeographic patterns might serve as indicator for upcoming ecosystem changes. Therefore, the detailed knowledge gained in this study – in conjunction with the knowledge yielded during a comprehensive study on the benthic bacterial community in the German Bight (Störmer 2013) – might be used to predict ecosystem wide responses to environmental changes, such as global warming, ocean acidification and anthropogenic pollution.

SUMMARY

Microbes comprise the most abundant and diverse group of organisms on earth, contribute substantially to every conceivable biogeochemical cycle and thus, are fundamental to ecosystem functioning. Research on marine microbial communities has proven the existence of biogeographic patterns, but the mechanisms that shape the microbial assemblages are still not understood properly, as most of the studies revealed different driving forces for different temporal and spatial scales and individual habitats. However, to predict and estimate potential ecological impacts that arise from microbial community variation, it is of utmost importance to unveil the mechanisms that are generating and maintaining community assembly.

This thesis aimed at providing detailed insight into the spatiotemporal variation of bacterioplankton communities in the German Bight, to deconvolute spatial and temporal signals and to identify the main regulating forces driving these dynamics.

The temporal variation was investigated on different scales, ranging from hours, days and weeks, up to seasonal scales. Additionally, the spatiotemporal biogeography of the bacterioplankton community was assessed on monthly cruises throughout a coastal-offshore transect. Diversity and community composition were assessed via automated ribosomal intergenic spacer analysis (ARISA) and 16S rRNA gene tag sequencing, respectively. Ultrahigh resolution Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) was used, to specifically analyze bacterioplankton-dissolved organic matter (DOM) interactions. Multivariate statistics and network analyses were applied to determine which environmental parameters exert the strongest influences on the bacterial community and thus, are likely to drive the dynamics in community variation.

The bacterioplankton community in the German Bight was found to be highly diverse on both, temporal and spatial scales. The research results lead to the conclusion that the seasonal variation of the bacterioplankton community overrides short-term variation which mainly occurs in response to phytoplankton blooms. This seasonal variation is predominantly driven by temperature and seems to substantially support the stability of the bacterial community in terms of annual succession, recurrence and resilience of few main taxa. Furthermore, this temperature influence leads to the formation of temperature-dependent guilds during spring and summer phytoplankton blooms and thus, is suggested to be a major niche defining factor in the German Bight. Observations on bacterioplankton-

DOM interactions on short time scales imply that the bacterial community is highly influenced by freshly produced, labile DOM pulses as derived from phytoplankton blooms. Strong correlations between specific bacterial taxa and particular molecular DOM formulae provide first hints towards relationships between potential substrate specialists and particular DOM molecules.

The spatiotemporal analyses of the bacterial community variation revealed the predominance of temperature-driven temporal variation over spatial patterns. The spatial variation is defined by pronounced hydrographic current conditions, rather than by environmental gradients. These hydrographic current conditions lead to a clear separation of coastal and offshore communities.

This thesis comprises a detailed descriptive approach which provides a fundamental basis for hypothesis-driven future research. Since bacteria exhibit rapid lifecycles and respond to environmental changes faster than other large organisms, changes in their biogeographic patterns might serve as indicator for upcoming ecosystem changes. Therefore, the detailed knowledge gained in this study might be used to predict ecosystem wide responses to environmental changes, such as global warming, ocean acidification and anthropogenic pollution.

ZUSAMMENFASSUNG

Mikroorganismen sind auf der Erde in allen nur erdenklichen Habitaten zu finden und kommen dort in spezifischen mikrobiellen Gemeinschaften vor. Durch ihre ausgeprägte metabolische Vielfalt übernehmen bakterielle Gemeinschaften Schlüsselrollen in allen bekannten biogeochemischen Stoffkreisläufen und leisten so einen fundamentalen Beitrag zur Aufrechterhaltung bzw. zu Veränderungen von Ökosystemen. Biogeographische Untersuchungen zeigten, dass sich Diversität und Zusammensetzung der Gemeinschaften entlang verschiedener zeitlicher und räumlicher Skalen verändern. Wie und durch welche Faktoren diese Veränderungen gesteuert werden ist aber noch unklar. Um zukünftige Ökosystemänderungen vorhersagen und ihr Ausmaß abschätzen zu können, ist es daher von immenser Bedeutung, die Wechselbeziehungen zwischen bakteriellen Gemeinschaften und möglichen regulierenden abiotischen und biotischen Faktoren, zu untersuchen.
Die vorliegende Arbeit liefert eine detaillierte Beschreibung der zeitlichen und räumlichen Variabilität der Bakterioplankton-Gemeinschaft in der Deutschen Bucht und benennt Umweltfaktoren, welche die Variabilität maßgeblich steuern.

Eine intensive Analyse der zeitlichen Variabilität wurde auf Basis von Beprobungen des Bakterioplanktons an der Helgoländer Reede in unterschiedlichen Zeitabständen (stündlich, täglich, wöchentlich, saisonal) realisiert. Zusätzlich ermöglichte eine monatliche Beprobung der Bakterioplankton-Gemeinschaft entlang eines Transekts vom Elbe Ästuar bis in die offene Nordsee (Pelagial) die gleichzeitige Untersuchung räumlicher und zeitlicher Muster.
Die Zusammensetzung der Bakterioplankton-Gemeinschaft wurde dabei mittels automated ribosomal intergenic spacer analysis (ARISA) und Hochdurchsatz-Sequenzierung des 16S rRNA Gens erfasst und über multivariate statistische Analysen mit gleichzeitig erhobenen Umweltparametern in Beziehung gesetzt. Als weitere hochauflösende Methode diente Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) zur molekularen Analyse des gelösten organischen Materials (DOM) im Wasser.

Die Ergebnisse dieser Arbeit belegen eine ausgeprägte Variabilität des Bakterioplanktons in der Deutschen Bucht. Im Hinblick auf die zeitliche Variabilität konnte gezeigt werden, dass saisonale Veränderungen im Jahresverlauf kurzzeitige Veränderungen als Reaktion auf Phytoplankton-Blüten überlagern. Diese saisonalen Muster werden primär durch Schwankungen in der Temperatur gesteuert. Die Temperatur scheint dabei die jährliche

Wiederkehr und Abfolge bestimmter Gemeinschaftszusammensetzungen und Taxa zu begünstigen. Zudem konnte die Bildung temperaturabhängiger Gilden während unterschiedlicher Phytoplankton-Blüten gezeigt werden, was auf die Bedeutung der Temperatur für die ökologische Nischenbildung in der Deutschen Bucht hinweist. Des Weiteren wurden Phasen hoch dynamischer Veränderungen über kurze Zeiträume im Bereich von Tagen und Wochen vor allem als Reaktion auf Phytoplankton-Blüten beobachtet. Während dieser Phasen nimmt vor allem das frisch produzierte labile gelöste organische Material (DOM) entscheidend Einfluss auf die Zusammensetzung der Gemeinschaft. Hier konnten starke Abhängigkeiten zwischen einzelnen Taxa und spezifischen DOM-Molekülen nachgewiesen werden.
Die raumzeitlichen Untersuchungen entlang des Transekts zeigen, dass die Temperatur gesteuerte, zeitliche Variabilität die Räumliche im gesamten Untersuchungsgebiet überlagert. Räumliche Unterschiede zwischen Gemeinschaften werden hauptsächlich durch hydrodynamische Strömungen gesteuert, was zu einer klaren Trennung von Gemeinschaften in Küstengewässern und im Pelagial der Nordsee führt.

Diese Arbeit bietet umfassende Einblicke in die Diversität und Variabilität der Bakterioplankton-Gemeinschaften in der Deutschen Bucht. Zum ersten Mal wurden räumliche und zeitliche Skalen simultan untersucht und zueinander in Beziehung gesetzt. Die neu gewonnenen Erkenntnisse liefern einen substantiellen Beitrag zum Verständnis der Regulationsmechanismen von Bakterioplankton-Gemeinschaften. Sie bieten eine fundierte Basis für zukünftige Ökosystemmodellierungen, sowie für Voraussagen zu möglichen Ökosystemveränderungen als Reaktion auf potentielle Umweltveränderungen wie z.B. Klimaerwärmung, Ozeanversauerung und Umweltverschmutzung.

REFERENCES

Alderkamp A, Sintes E and Herndl G. Abundance and activity of major groups of prokaryotic plankton in the coastal North Sea during spring and summer. *Aquat Microb Ecol* 2006;**45**: 237-46.

Amin SA, Parker MS and Armbrust EV. Interactions between diatoms and bacteria. *Microbiol Mol Biol R* 2012;**76**: 667-84.

Amon RMW and Benner R. Bacterial utilization of different size classes of dissolved organic matter. *Limnol Oceanogr* 1996;**41**: 41-51.

Anderson MJ and Willis TJ. Canonical analysis of principal coordinates: a useful method of constrained ordination for ecology. *Ecology* 2003;**84**: 511-25.

Andersson AF, Riemann L and Bertilsson S. Pyrosequencing reveals contrasting seasonal dynamics of taxa within Baltic Sea bacterioplankton communities. *ISME J* 2009;**4**: 171-81.

Atlas RM and Bartha R. *Microbial Ecology: Fundamentals and Applications.* Menlo Park, California:Benjamin/Cummings, 1987.

Azam F. Microbial control of oceanic carbon flux: The plot thickens. *Science* 1998;**280**: 694-6.

Azam F, Fenchel T, Field JG, Gray JS, Meyer-Reil LA and Thingstad F. The ecological role of water-column microbes in the sea. *Mar Ecol Prog Ser* 1983;**10**: 257-63.

Baines SB and Pace ML. The production of dissolved organic matter by phytoplankton and its importance to bacteria: Patterns across marine and freshwater systems. *Limnol Oceanogr* 1991;**36**: 1078-90.

Bardgett RD, Freeman C and Ostle NJ. Microbial contributions to climate change through carbon cycle feedbacks. *ISME J* 2008;**2**: 805-14.

Baune C and Böttcher ME. Experimental investigation of sulphur isotope partitioning during outgassing of hydrogen sulphide from diluted aqueous solutions and seawater. *Isot Environ Healt S* 2010;**46**: 444-53.

Beare DJ, Batten S, Edwards M and Reid DG. Prevalence of boreal Atlantic, temperate Atlantic and neritic zooplankton in the North Sea between 1958 and 1998 in relation to temperature, salinity, stratification intensity and Atlantic inflow. *J Sea Research* 2002;**48**: 29-49.

Becker G, Dick S and Dippner J. Hydrography of the German Bight. *Mar Ecol Prog Ser* 1992;**91**: 9-18.

Becker JW, Berube PM, Follett CL, Waterbury JB, Chisholm SW, DeLong EF *et al.* Closely related phytoplankton species produce similar suites of dissolved organic matter. *Front Microbiol* 2014, DOI 10.3389/fmicb.2014.00111.

Beermann J. Spatial and seasonal population dynamics of sympatric Jassa species (Crustacea, Amphipoda). *J Exp Mar Biol Ecol* 2014;**459**: 8-16.

Blankenberg D, Kuster GV, Coraor N, Ananda G, Lazarus R, Mangan M *et al.* Galaxy: A web-based genome analysis tool for experimentalists. *Curr Protoc Mol Biol* 2010;**19**: 1-21.

Boer SI, Hedtkamp SIC, van Beusekom JEE, Fuhrman JA, Boetius A and Ramette A. Time- and sediment depth-related variations in bacterial diversity and community structure in subtidal sands. *ISME J* 2009;**3**: 780-91.

Boras JA, Vaqué D, Maynou F, Sà EL, Weinbauer MG and Sala MM. Factors shaping bacterial phylogenetic and functional diversity in coastal waters of the NW Mediterranean Sea. *Estuar Coast Shelf S* 2015;**154**: 102-10.

Bouvier TC and del Giorgio PA. Compositional changes in free-living bacterial communities along a salinity gradient in two temperate estuaries. *Limnol Oceanogr* 2002;**47**: 453-70.

Brettar I, Christen R and Höfle MG. *Idiomarina baltica* sp. nov., a marine bacterium with a high optimum growth temperature isolated from surface water of the central Baltic Sea. *Int J Syst Evol Micr* 2003;**53**: 407-13.

Bronk DA. Dynamics of DON. In: Hansell DA, Carlson CA (eds.) *Biogeochemistry of Marine Dissolved Organic Matter*. San Diego: Academic Press, 2002, 153-249.

Bronk DA, See JH, Bradley P and Killberg L. DON as a source of bioavailable nitrogen for phytoplankton. *Biogeosciences* 2007;**4**: 283-96.

Brown JH, Gillooly JF, Allen AP, Savage VM and West GB. Toward a metabolic theory of ecology. *Ecology* 2004;**85**: 1771-89.

Brown M. Coupling 16S-ITS rDNA clone libraries and automated ribosomal intergenic spacer analysis to show marine microbial diversity development and application to a time series. *Environ Microbiol* 2005;**7**: 1466-79.

Buchan A, González JM and Moran MA. Overview of the Marine *Roseobacter* Lineage. *Appl Environ Microbiol* 2005;**71**: 5665-77.

Buchan A, LeCleir GR, Gulvik CA and González JM. Master recyclers: features and functions of bacteria associated with phytoplankton blooms. *Nat Rev Microbiol* 2014;**12**: 686-98.

Burow LC, Kong Y, Nielsen JL, Blackall LL and Nielsen PH. Abundance and ecophysiology of *Defluviicoccus* spp., glycogen-accumulating organisms in full-scale wastewater treatment processes. *Microbiology* 2007;**153**: 178-85.

Callies U, Plüß A, Kappenberg J and Kapitza H. Particle tracking in the vicinity of Helgoland, North Sea: a model comparison. *Ocean Dynam* 2011;**61**: 2121-39.

Callies U and Scharfe M. Mean spring conditions at Helgoland Roads, North Sea: Graphical modeling of the influence of hydro-climatic forcing and Elbe River discharge. *J Sea Research* 2015;**101**: 1-11.

Campbell BJ and Kirchman DL. Bacterial diversity, community structure and potential growth rates along an estuarine salinity gradient. *ISME J* 2013;**7**: 210-20.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ *et al.* Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci* 2011;**108**: 4516-22.

Caporaso JG, Paszkiewicz K, Field D, Knight R and Gilbert JA. The Western English Channel contains a persistent microbial seed bank. *ISME J* 2012;**6**: 1089-93.

Carlson C. Production and Removal Processes. In: Hansell DA, Carlson C (eds.) *Biogeochemistry of Marine Dissolved Organic Matter*: Academic Press, 2002, 91-150.

Casamayor EO, Schäfer H, Bañeras L, Pedrós-Alió C and Muyzer G. Identification of and spatio-temporal differences between microbial assemblages from two neighboring sulfurous lakes: comparison by microscopy and denaturing gradient gel electrophoresis. *Appl Environ Microbiol* 2000;**66**: 499-508.

Chase JM and Myers JA. Disentangling the importance of ecological niches from stochastic processes across scales. *Philos T R Soc B* 2011;**366**: 2351-63.

Chen M-H, Sheu S-Y, Chen CA, Wang J-T and Chen W-M. *Idiomarina aquimaris* sp. nov., isolated from the reef-building coral *Isopora palifera*. *Int J Syst Evol Micr* 2012;**62**: 1536-42.

Chow C-ET, Kim DY, Sachdeva R, Caron DA and Fuhrman JA. Top-down controls on bacterial community structure: microbial network analysis of bacteria, T4-like viruses and protists. *ISME J* 2014;**8**: 816-29.

Chow C-ET, Sachdeva R, Cram JA, Steele JA, Needham DM, Patel A *et al.* Temporal variability and coherence of euphotic zone bacterial communities over a decade in the Southern California Bight. *ISME J* 2013;**7**: 2259-73.

Compant S, Nowak J, Coenye T, Clément C and Ait Barka E. Diversity and occurrence of *Burkholderia* spp. in the natural environment. *Fems Microbiol Rev* 2008;**32**: 607-26.

Comte J, Lindstrom ES, Eiler A and Langenheder S. Can marine bacteria be recruited from freshwater sources and the air? *ISME J* 2014;**8**: 2423-30.

Connell JH and Orias E. The ecological regulation of species diversity. *Am Nat* 1964;**98**: 399-414.

Cottrell MT and David KL. Contribution of major bacterial groups to bacterial biomass production (thymidine and leucine incorporation) in the Delaware estuary. *Limnol Oceanogr* 2003;**48**: 168-78.

Cottrell MT and Kirchman DL. Natural assemblages of marine *Proteobacteria* and members of the *Cytophaga-Flavobacter* cluster consuming low- and high-molecular-weight dissolved organic matter. *Appl Environ Microbiol* 2000;**66**: 1692-7.

Cram JA, Chow C-ET, Sachdeva R, Needham DM, Parada AE, Steele JA *et al.* Seasonal and interannual variability of the marine bacterioplankton community throughout the water column over ten years. *ISME J* 2015;**9**: 563-80.

Crump BC and Baross JA. Characterization of the bacterially-active particle fraction in the Columbia River estuary. *Mar Ecol Prog Ser* 2000;**206**: 13-22.

Crump BC, Hopkinson CS, Sogin ML and Hobbie JE. Microbial biogeography along an estuarine salinity gradient: Combined influences of bacterial growth and residence time. *Appl Environ Microbiol* 2004;**70**: 1494-505.

Danovaro R, Luna GM, Dell'Anno A and Pietrangeli B. Comparison of two fingerprinting techniques, terminal restriction fragment length polymorphism and automated ribosomal intergenic spacer analysis, for determination of bacterial diversity in aquatic environments. *Appl Environ Microbiol* 2006;**72**: 5982-9.

Dick S, Kleine E, S.H. M-N, Klein H and Komo H. The operational circulation model of BSH (BSHcmod)-model description and validation*Berichte des Bundesamtes für Seeschifffahrt und Hydrographie* volume 29, 2001.

Dini-Andreotea F, Stegen JC, van Elsasa JD and Falcão Sallesa J. Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proc Natl Acad Sci U S A* 2015;**112**: E1326-E32.

Dittmar T, Koch B, Hertkorn N and Kattner G. A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnol Oceanogr-Meth* 2008;**6**: 230-5.

Dittmar T and Paeng J. A heat-induced molecular signature in marine dissolved organic matter. *Nature Geosci* 2009;**2**: 175-9.

Donachie SP, Hou S, Gregory TS, Malahoff A and Alam M. *Idiomarina loihiensis* sp. nov., a halophilic γ-*Proteobacterium* from the Lō‘ihi submarine volcano, Hawaii. *Int J Syst Evol Micr* 2003;**53**: 1873-9.

Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G *et al.* Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 2013;**36**: 27-46.

Du J, Xiao K, Li L, Ding X, Liu H, Lu Y *et al.* Temporal and spatial diversity of bacterial communities in coastal waters of the South China Sea. *Plos One* 2013;**8**: e66968.

Ducklow HW. The bacterial component of the oceanic euphotic zone. *FEMS Microbiol Ecol* 1999;**30**: 1-10.

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

Eilers H, Pernthaler J, Peplies J, Glöckner FO, Gerdts G and Amann R. Isolation of novel pelagic bacteria from the German Bight and their seasonal contributions to surface picoplankton. *Appl Environ Microbiol* 2001;**67**: 5134-42.

Elifantz H, Malmstrom RR, Cottrell MT and Kirchman DL. Assimilation of polysaccharides and glucose by major bacterial groups in the Delaware Estuary. *Appl Environ Microbiol* 2005;**71**: 7799-805.

Estrada-de los Santos P, Vinuesa P, Martínez-Aguilar L, Hirsch A and Caballero-Mellado J. Phylogenetic analysis of *Burkholderia* species by multilocus sequence analysis. *Curr Microbiol* 2013;**67**: 51-60.

Fisher MM and Triplett EW. Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to freshwater bacterial communities. *Appl Environ Microbiol* 1999;**65**: 4630-6.

Flerus R, Lechtenfeld OJ, Koch BP, McCallister SL, Schmitt-Kopplin P, Benner R *et al.* A molecular perspective on the ageing of marine dissolved organic matter. *Biogeosciences* 2012;**9**: 1935-55.

Fortunato C and Crump B. Bacterioplankton community variation across river to ocean environmental gradients. *Microb Ecol* 2011;**62**: 374-82.

Fortunato CS, Eiler A, Herfort L, Needoba JA, Peterson TD and Crump BC. Determining indicator taxa across spatial and seasonal gradients in the Columbia River coastal margin. *ISME J* 2013;**7**: 1899-911.

Fortunato CS, Herfort L, Zuber P, Baptista AM and Crump BC. Spatial variability overwhelms seasonal patterns in bacterioplankton communities across a river to ocean gradient. *ISME J* 2012;**6**: 554-63.

Frossard A, Gerull L, Mutz M and Gessner MO. Disconnect of microbial structure and function: enzyme activities and bacterial communities in nascent stream corridors. *ISME J* 2012;**6**: 680-91.

Fuhrman JA, Cram JA and Needham DM. Marine microbial community dynamics and their ecological interpretation. *Nat Rev Micro* 2015;**13**: 133-46.

Fuhrman JA, Hewson I, Schwalbach MS, Steele JA, Brown MV and Naeem S. Annually reoccurring bacterial communities are predictable from ocean conditions. *Proc Natl Acad Sci U S A* 2006;**103**: 13104-9.

Fuhrman JA and Steele JA. Community structure of marine bacterioplankton: patterns, networks, and relationships to function. *Aquat Microb Ecol* 2008;**53**: 69-81.

Fuhrman JA, Steele JA, Hewson I, Schwalbach MS, Brown MV, Green JL *et al.* A latitudinal diversity gradient in planktonic marine bacteria. *Proc Natl Acad Sci U S A* 2008;**105**: 7774-8.

Ganesh S, Parris DJ, DeLong EF and Stewart FJ. Metagenomic analysis of size-fractionated picoplankton in a marine oxygen minimum zone. *ISME J* 2014;**8**: 187-211.

Gerdts G, Wichels A, Döpke H, Klings KW, Gunkel W and Schütt C. 40-year long-term study of microbial parameters near Helgoland (German Bight, North Sea): Historical view and future perspectives. *Helgoland Mar Res* 2004;**58**: 230-42.

Ghiglione JF, Mevel G, Pujo-Pay M, Mousseau L, Lebaron P and Goutx M. Diel and seasonal variations in abundance, activity, and community structure of particle-attached and free-living bacteria in NW Mediterranean Sea. *Microb Ecol* 2007;**54**: 217-31.

Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P *et al.* Galaxy: A platform for interactive large-scale genome analysis. *Genome Res* 2005;**15**: 1451-5.

Giebel H-A, Kalhoefer D, Lemke A, Thole S, Gahl-Jannsen R, Simon M *et al.* Distribution of *Roseobacter* RCA and SAR11 lineages in the North Sea and characteristics of an abundant RCA isolate. *ISME J* 2011;**5**: 8-19.

Gilbert J, Field D, Swift P, Newbold L, Oliver A, Smyth T *et al.* The seasonal structure of microbial communities in the Western English Channel. *Environ Microbiol* 2009;**11**: 3132 - 9.

Gilbert JA, Steele JA, Caporaso JG, Steinbruck L, Reeder J, Temperton B *et al.* Defining seasonal marine microbial community dynamics. *ISME J* 2012;**6**: 298-308.

Giovannoni S and Rappé M. Evolution, diversity and molecular ecology of marine prokaryotes. In: Kirchman D (ed.) *Microbial Ecology of the Oceans*. New Jersey, US: Wiley and Sons, 2000, 47-84.

Giovannoni SJ, Britschgi TB, Moyer CL and Field KG. Genetic diversity in Sargasso Sea bacterioplankton. *Nature* 1990;**345**: 60-3.

Goecks J, Nekrutenko A, Taylor J and Team TG. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biology* 2010;**11**: R86.

Gómez-Pereira PR, Schüler M, Fuchs BM, Bennke C, Teeling H, Waldmann J *et al.* Genomic content of uncultured *Bacteroidetes* from contrasting oceanic provinces in the North Atlantic Ocean. *Environ Microbiol* 2012;**14**: 52-66.

Graham MH. Confronting multicollinearity in ecological multiple regression *Ecology* 2003;**84**: 2809-15.

Gutknecht JM, Goodman R and Balser T. Linking soil process and microbial ecology in freshwater wetland ecosystems. *Plant Soil* 2006;**289**: 17-34.

Hansell DA. Recalcitrant dissolved organic carbon fractions. *Annu Rev Mar Sci* 2013;**5**: 421-45.

Hanson CA, Fuhrman JA, Horner-Devine MC and Martiny JBH. Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat Rev Microbiol* 2012;**10**: 497-506.

Heath MR and Beare DJ. New primary production in northwest European shelf seas, 1960-2003. *Mar Ecol Prog Ser* 2008;**363**: 183-203.

Hedges JI. Global biogeochemical cycles: Progress and problems. *Mar Chem* 1992;**39**: 67-93.

Herfort L, Schouten S, Abbas B, Veldhuis MJW, Coolen MJL, Wuchter C *et al.* Variations in spatial and temporal distribution of Archaea in the North Sea in relation to environmental variables. *FEMS Microbiol Ecol* 2007;**62**: 242-57.

Herlemann DPR, Labrenz M, Jurgens K, Bertilsson S, Waniek JJ and Andersson AF. Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME J* 2011;**5**: 1571-9.

Hewson I, Steele JA, Capone DG and Fuhrman JA. Temporal and spatial scales of variation in bacterioplankton assemblages of oligotrophic surface waters. *Marine Ecology-Progress Series* 2006;**311**: 67-77.

Hillebrand H. Strength, slope and variability of marine latitudinal gradients. *Mar Ecol Prog Ser* 2004;**273**: 251-67.

Holm S. A simple sequentially rejective multiple test procedure. *Scand J Stat* 1979;**6**: 65-70.

Holmström C and Kjelleberg S. Marine *Pseudoalteromonas* species are associated with higher organisms and produce biologically active extracellular agents. *FEMS Microbiol Ecol* 1999;**30**: 285-93.

Hoppe H, Breithaupt P, Walther K, Koppe R, Bleck S, Sommer U *et al.* Climate warming in winter affects the coupling between phytoplankton and bacteria during the spring bloom: A mesocosm study. *Aquat Microb Ecol* 2008;**51**: 105-15.

Hotelling H. The relations of the newer multivariate statistical methods to factor analysis. *British Journal of Statistical Psychology* 1957;**10**: 69-79.

Howarth MJ. North Sea Circulation. In: John HS (ed.) *Encyclopedia of Ocean Sciences*. Oxford: Academic Press, 2001, 1912-21.

Ionescu D, Siebert C, Polerecky L, Munwes YY, Lott C, Häusler S *et al.* Microbial and chemical characterization of underwater fresh water springs in the Dead Sea. *Plos One* 2012;**7**: e38319.

Ivanova EP, Bakunina IY, Sawabe T, Hayashi K, Alexeeva YV, Zhukova NV *et al.* Two Species of culturable bacteria associated with degradation of brown algae *Fucus evanescens*. *Microb Ecol* 2002;**43**: 242-9.

Ivanova EP, Romanenko LA, Chun J, Matte MH, Matte GR, Mikhailov VV *et al. Idiomarina* gen. nov., comprising novel indigenous deep-sea bacteria from the Pacific Ocean, including descriptions of two species, *Idiomarina abyssalis* sp. nov. and *Idiomarina zobellii* sp. nov. *Int J Syst Evol Micr* 2000;**50**: 901-7.

Iverson V, Morris RM, Frazar CD, Berthiaume CT, Morales RL and Armbrust EV. Untangling genomes from metagenomes: Revealing an uncultured class of marine *Euryarchaeota*. *Science* 2012;**335**: 587-90.

James FC and McCulloch CE. Multivariate Analysis in Ecology and Systematics: Panacea or Pandora's Box? *Annu Rev Ecol Syst* 1990;**21**: 129-66.

Jones SE, Newton RJ and McMahon KD. Potential for atmospheric deposition of bacteria to influence bacterioplankton communities. *FEMS Microbiol Ecol* 2008;**64**: 388-94.

Jürgens K and Matz C. Predation as a shaping force for the phenotypic and genotypic composition of planktonic bacteria. *Anton Leeuw Int J G* 2002;**81**: 413-31.

Kim S, Kramer RW and Hatcher PG. Graphical method for analysis of ultrahigh-resolution broadband mass spectra of natural organic matter, the Van Krevelen Diagram. *Anal Chem* 2003;**75**: 5336-44.

Kirchman DL. Introduction and overview - Dissolved organic matter. In: Kirchman DL (ed.) *Microbial Ecology of the Oceans*: 2 Edition. New Jersey, USA: Wiley & Sons, 2008, 1-26.

Kirchman DL, Dittel AI, Malmstrom RR and Cottrell MT. Biogeography of major bacterial groups in the Delaware Estuary. *Limnol Oceanogr* 2005;**50**: 1697-706.

Kirchman DL, Moran XAG and Ducklow H. Microbial growth in the polar oceans role of temperature and potential impact of climate change. *Nat Rev Microbiol* 2009;**7**: 451-9.

Kirchman DL, Suzuki Y, Garside C and Ducklow HW. High turnover rates of dissolved organic carbon during a spring phytoplankton bloom. *Nature* 1991;**352**: 612-4.

Koch BP and Dittmar T. From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter. *Rapid Commun Mass Sp* 2006;**20**: 926-32.

Koch BP, Dittmar T, Witt M and Kattner G. Fundamentals of molecular formula assignment to ultrahigh resolution mass data of natural organic matter. *Anal Chem* 2007;**79**: 1758-63.

Koch BP, Witt M, Engbrodt R, Dittmar T and Kattner G. Molecular formulae of marine and terrigenous dissolved organic matter detected by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Geochim Cosmochim Ac* 2005;**69**: 3299-308.

Kovacs A, Yacoby K and Gophna U. A systematic assessment of automated ribosomal intergenic spacer analysis (ARISA) as a tool for estimating bacterial richness. *Res Microbiol* 2010;**161**: 192-7.

Krause E, Wichels A, Giménez L, Lunau M, Schilhabel MB and Gerdts G. Small changes in pH have direct effects on marine bacterial community composition: A microcosm approach. *Plos One* 2012;**7**: e47035.

Kuipers B, Noort GJv, Vosjan J and Herndl GJ. Diel periodicity of bacterioplankton in the euphotic zone of the subtropical Atlantic Ocean. *Mar Ecol Prog Ser* 2000;**201**: 13-25.

Langenheder S, Lindström ES and Tranvik LJ. Weak coupling between community composition and functioning of aquatic bacteria. *Limnol Oceanogr* 2005;**50**: 957-67.

Lau KWK, Ng CYM, Ren J, Lau SCL, Qian P-Y, Wong P-K *et al. Owenweeksia hongkongensis* gen. Nov., sp. Nov., a novel marine bacterium of the phylum '*Bacteroidetes*'. *Int J Syst Evol Micr* 2005;**55**: 1051-7.

Lear G, Anderson MJ, Smith JP, Boxen K and Lewis GD. Spatial and temporal heterogeneity of the bacterial communities in stream epilithic biofilms. *FEMS Microbiol Ecol* 2008;**65**: 463-73.

Lear G, Bellamy J, Case BS, Lee JE and Buckley HL. Fine-scale spatial patterns in bacterial community composition and function within freshwater ponds. *ISME J* 2014;**8**: 1715-26.

Lee JE. Niche and neutral processes in aquatic bacterial communities: Are all things equal? volume PhD: Lincoln University, 2014.

Lee S-O, Kato J, Takiguchi N, Kuroda A, Ikeda T, Mitsutani A *et al.* Involvement of an extracellular protease in algicidal activity of the marine bacterium *Pseudoalteromonas* sp. strain A28. *Appl Environ Microbiol* 2000;**66**: 4334-9.

Legendre P and Legendre L. *Numerical ecology.* San Diego, CA:Elsevier, 1998.

Lindström ES and Östman Ö. The importance of dispersal for bacterial community composition and functioning. *Plos One* 2011;**6**: e25883.

Long RA and Azam F. Microscale patchiness of bacterioplankton assemblage richness in seawater. *Aquat Microb Ecol* 2001;**26**: 103-13.

Lozupone CA and Knight R. Global patterns in bacterial diversity. *Proc Natl Acad Sci* 2007;**104**: 11436-40.

Lübben A, Dellwig O, Koch S, Beck M, Badewien T, Fischer S *et al.* Distributions and characteristics of dissolved organic matter in temperate coastal waters (Southern North Sea). *Ocean Dynam* 2009;**59**: 263-75.

Lucas J, Koester I, Wichels A, Niggemann J, Dittmar T, Callies U *et al.* Short-term dynamics of North Sea bacterioplankton-dissolved organic matter interactions on molecular level. unpublished, 2015a.

Lucas J, Wichels A, Teeling H, Chafee M, Scharfe M and Gerdts G. Annual dynamics of North Sea bacterioplankton: Seasonal variability superimposes short-term variation. *FEMS Microbiol Ecol* 2015b, DOI 10.1093/femsec/fiv099.

Lueders T and Friedrich MW. Evaluation of PCR amplification bias by terminal restriction fragment length polymorphism analysis of small-subunit rRNA and mcrA genes by using defined template mixtures of methanogenic pure cultures and soil DNA extracts. *Appl Environ Microbiol* 2003;**69**: 320-6.

Martiny JBH, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL *et al.* Microbial biogeography: putting microorganisms on the map. *Nat Rev Micro* 2006;**4**: 102-12.

Martiny JBH, Eisen JA, Penn K, Allison SD and Horner-Devine MC. Drivers of bacterial β-diversity depend on spatial scale. *Proc Natl Acad Sci* 2011;**108**: 7850-4.

McCarren J, Becker JW, Repeta DJ, Shi Y, Young CR, Malmstrom RR *et al.* Microbial community transcriptomes reveal microbes and metabolic pathways associated with

dissolved organic matter turnover in the sea. *Proc Natl Acad Sci* 2010;**107**: 16420-7.

McCarthy MD, Hedges JI and Benner R. Major bacterial contribution to marine dissolved organic nitrogen. *Science* 1998;**281**: 231-4.

Medeiros PM, Seidel M, Ward ND, Carpenter EJ, Gomes HR, Niggemann J *et al.* Fate of the Amazon River dissolved organic matter in the tropical Atlantic Ocean. *Global Biogeochem Cy* 2015;**29**: 677-90.

Mével G, Vernet M, Goutx M and Ghiglione JF. Seasonal to hour variation scales in abundance and production of total and particle-attached bacteria in the open NW Mediterranean Sea (0-1000 m). *Biogeosciences* 2008;**5**.

Meziti A, Ramette A, Mente E and Kormas KA. Temporal shifts of the Norway lobster (Nephrops norvegicus) gut bacterial communities. *FEMS Microbiol Ecol* 2010;**74**: 472-84.

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

Management Unit of the North Sea Mathematical Models, Royal Royal Belgian Institute of Natural Sciences. North Sea facts, 2000. available at http://www.mumm.ac.be/EN/NorthSea/facts.php.

Nagata T. Organic matter-bacteria interactions in seawater. In: Kirchman DL (ed.) *Microbial Ecology of the Oceans*: 2 Edition. New Jersey, USA: Wiley & Sons, 2008, 207-42.

Nemergut DR, Costello EK, Hamady M, Lozupone C, Jiang L, Schmidt SK *et al.* Global patterns in the biogeography of bacterial taxa. *Environ Microbiol* 2011;**13**: 135-44.

Nobu MK, Tamaki H, Kubota K and Liu W-T. Metagenomic characterization of 'Candidatus *Defluviicoccus tetraformis* strain TFO71', a tetrad-forming organism, predominant in an anaerobic–aerobic membrane bioreactor with deteriorated biological phosphorus removal. *Environ Microbiol* 2014;**16**: 2739-51.

Oh H-M, Kwon KK, Kang I, Kang SG, Lee J-H, Kim S-J *et al.* Complete genome sequence of "Candidatus *Puniceispirillum marinum*" IMCC1322, a representative of the SAR116 clade in the *Alphaproteobacteria*. *J Bacteriol* 2010;**192**: 3240-1.

Okubo A and Sugiyama S. Comparison of molecular fingerprinting methods for analysis of soil microbial community structure. *Ecol Res* 2009;**24**: 1399-405.

Ortega-Retuerta E, Joux F, Jeffrey WH and Ghiglione JF. Spatial variability of particle-attached and free-living bacterial diversity in surface waters from the Mackenzie River to the Beaufort Sea (Canadian Arctic). *Biogeosciences* 2013;**10**: 2747-59.

OSPAR. Quality Status Report 2000. London: OSPAR Commision 2000, 2000.

Osterholz H, Dittmar T and Niggemann J. Molecular evidence for rapid dissolved organic matter turnover in Arctic fjords. *Mar Chem* 2014;**160**: 1-10.

Osterholz H, Niggemann J, Giebel H-A, Simon M and Dittmar T. Inefficient microbial production of refractory dissolved organic matter in the ocean. *Nat Commun* 2015;**6**.

Pernthaler A, Preston CM, Pernthaler J, DeLong EF and Amann R. Comparison of fluorescently labeled oligonucleotide and polynucleotide probes for the detection of pelagic marine *Bacteria* and *Archaea*. *Appl Environ Microbiol* 2002;**68**: 661-7.

Pernthaler J. Predation on prokaryotes in the water column and its ecological implications. *Nat Rev Micro* 2005;**3**: 537-46.

Pinhassi J, Azam F, Hempha?la J, Long RA, Martinez J, Zweifel UL *et al.* Coupling between bacterioplankton species composition, population dynamics, and organic matter degradation. *Aquat Microb Ecol* 1999;**17**: 13-26.

Pinhassi J, Sala MM, Havskum H, Peters F, Guadayol Ò, Malits A *et al.* Changes in bacterioplankton composition under different phytoplankton regimens. *Appl Environ Microbiol* 2004;**70**: 6753-66.

Pomeroy LR. The ocean's food web, a changing paradigm. *BioScience* 1974;**24**: 499-504.

Pomeroy LR and Wiebe WJ. Temperature and substrates as interactive limiting factors for marine heterotrophic bacteria. *Aquat Microb Ecol* 2001;**23**: 187-204.

Pommier T, Canbäck B, Riemann L, Boström KH, Simu K, Lundberg P *et al.* Global patterns of diversity and community structure in marine bacterioplankton. *Mol Ecol* 2007;**16**: 867-80.

Purahong W, Schloter M, Pecyna MJ, Kapturska D, Däumlich V, Mital S *et al.* Uncoupling of microbial community structure and function in decomposing litter across beech forest ecosystems in Central Europe. *Sci Rep* 2014;**4**.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P *et al.* The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res* 2013;**41**: 590-6.

R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2014.

Raabe T and Wiltshire KH. Quality control and analyses of the long-term nutrient data from Helgoland Roads, North Sea. *J Sea Research* 2009;**61**: 3-16.

Ramette A and Tiedje J. Biogeography: an emerging cornerstone for understanding prokaryotic diversity, ecology, and evolution. *Microb Ecol* 2007;**53**: 197-207.

Ranjard L, Brothier E and Nazaret S. Sequencing bands of ribosomal intergenic spacer analysis fingerprints for characterization and microscale distribution of soil bacterium populations responding to mercury spiking. *Appl Environ Microbiol* 2000;**66**: 5334-9.

Rick H, Rick S, Tillmann U, Brockmann U, Gärtner U, Dürselen C *et al.* Primary productivity in the German Bight (1994–1996). *Estuaries Coasts* 2006;**29**: 4-23.

Riemann L and Middelboe M. Stability of bacterial and viral community compositions in danish coastal waters as depicted by DNA fingerprinting techniques. *Aquat Microb Ecol* 2002;**27**: 219-2321.

Rink B, Grüner N, Brinkhoff T, Ziegelmüller K and Simon M. Regional patterns of bacterial community composition and biogeochemical properties in the southern North Sea. *Aquat Microb Ecol* 2011;**63**: 207-22.

Rink B, Martens T, Fischer D, Lemke A, Grossart H-P, Simon M *et al.* Short-term dynamics of bacterial communities in a tidally affected coastal ecosystem. *FEMS Microbiol Ecol* 2008;**66**: 306-19.

Rink B, Seeberger S, Martens T, Duerselen C-D, Simon M and Brinkhoff T. Effects of phytoplankton bloom in a coastal ecosystem on the composition of bacterial communities. *Aquat Microb Ecol* 2007;**48**: 47-60.

Rooney-Varga JN, Giewat MW, Savin MC, Sood S, LeGresley M and Martin JL. Links between phytoplankton and bacterial community dynamics in a coastal marine environment. *Microb Ecol* 2005;**49**: 163-75.

Rossel PE, Vähätalo AV, Witt M and Dittmar T. Molecular composition of dissolved organic matter from a wetland plant (*Juncus effusus*) after photochemical and microbial decomposition (1.25 yr): Common features with deep sea dissolved organic matter. *Org Geochem* 2013;**60**: 62-71.

Sapp M, Gerdts G, Wiltshire KH and Wichels A. Bacterial community dynamics during the winter-spring transition in the North Sea. *FEMS Microbiol Ecol* 2007;**59**: 622-37.

Sarmento H and Gasol JM. Use of phytoplankton-derived dissolved organic carbon by different types of bacterioplankton. *Environ Microbiol* 2012;**14**: 2348-60.

Sarmento H, Romera-Castillo C, Lindh M, Pinhassi J, Sala MM, Gasol JM *et al.* Phytoplankton species-specific release of dissolved free amino acids and their selective consumption by bacteria. *Limnol Oceanogr* 2013a;**58**: 1123-35.

Sarmento H, Romera-Castillo C, Lindh MV, Pinhassi J, Sala MM, Gasol JM *et al.* Phytoplankton species-specific release of dissolved free amino acids and their selective consumption by bacteria. *Limnol Oceanogr* 2013b;**58**: 1125-35.

Scharfe M. Analysis of biological long-term changes based on hydro-climatic parameters in the southern North Sea (Helgoland): University of Hamburg (In German), 2013, 105.

Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett W *et al.* Metagenomic biomarker discovery and explanation. *Genome Biology* 2011;**12**: R60.

Seidel M, Beck M, Riedel T, Waska H, Suryaputra IGNA, Schnetger B *et al.* Biogeochemistry of dissolved organic matter in an anoxic intertidal creek bank. *Geochim Cosmochim Ac* 2014;**140**: 418-34.

Seidel M, Yager PL, Ward ND, Carpenter EJ, Gomes HR, Krusche AV *et al.* Molecular-level changes of dissolved organic matter along the Amazon River-to-ocean continuum. *Mar Chem* 2015, DOI http://dx.doi.org/10.1016/j.marchem.2015.06.019.

Selje N and Simon M. Composition and dynamics of particle-associated and free-living bacterial communities in the Weser estuary, Germany. *Aquat Microb Ecol* 2003;**30**: 221-37.

Selje N, Simon M and Brinkhoff T. A Newly discovered *Roseobacter* cluster in temperate and polar oceans. *Nature* 2004;**427**: 445-8.

Shade A, Read JS, Youngblut ND, Fierer N, Knight R, Kratz TK *et al.* Lake microbial communities are resilient after a whole-ecosystem disturbance. *ISME J* 2012;**6**: 2153-67.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D *et al.* Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003;**13**: 2498-504.

Sherr EB and Sherr BF. High rates of consumption of bacteria by pelagic ciliates. *Nature* 1987;**325**: 710-1.

Sherr EB and Sherr BF. Role of microbes in pelagic food webs: A revised concept. *Limnol Oceanogr* 1988;**33**: 1225-7.

Simek K, Jürgens K, Nedoma J, Comerma M and Armengol J. Ecological role and bacterial grazing of *Halteria* spp.: Small freshwater oligotrichs as dominant pelagic ciliate bacterivores. *Aquat Microb Ecol* 2000;**22**: 43-56.

Simon M, Grossart H-P, Schweitzer B and Ploug H. Microbial ecology of organic aggregates in aquatic ecosystems. *Aquat Microb Ecol* 2002;**28**: 175-211.

Sintes E, Stoderegger K, Parada V and Herndl G. Seasonal dynamics of dissolved organic matter and microbial activity in the coastal North Sea. *Aquat Microb Ecol* 2010;**60**: 85-95.

Sintes E, Witte H, Stodderegger K, Steiner P and Herndl GJ. Temporal dynamics in the free-living bacterial community composition in the coastal North Sea. *FEMS Microbiol Ecol* 2013;**83**: 413-24.

Sleighter RL and Hatcher PG. Molecular characterization of dissolved organic matter (DOM) along a river to ocean transect of the lower Chesapeake Bay by ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Mar Chem* 2008;**110**: 140-52.

Smith A, Koper N, Francis C and Fahrig L. Confronting collinearity: comparing methods for disentangling the effects of habitat loss and fragmentation. *Landscape Ecol* 2009;**24**: 1271-85.

Sogin ML, Morrison HG, Huber JA, Welch DM, Huse SM, Neal PR *et al.* Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc Natl Acad Sci* 2006;**103**: 12115-20.

Staneva J, Stanev EV, Wolff J-O, Badewien TH, Reuter R, Flemming B *et al.* Hydrodynamics and sediment dynamics in the German Bight. A focus on observations and numerical modelling in the East Frisian Wadden Sea. *Cont Shelf Res* 2009;**29**: 302-19.

Stenson AC, Marshall AG and Cooper WT. Exact masses and chemical formulas of individual Suwannee River fulvic acids from ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectra. *Anal Chem* 2003;**75**: 1275-84.

Stevens H, Brinkhoff T and Simon M. Composition of free-living, aggregate-associated and sediment surface-associated bacterial communities in the German Wadden Sea. *Aquat Microb Ecol* 2005;**38**: 15-30.

Stockmann K, Callies U, Manly BFJ and Wiltshire KH. Long-term model simulation of environmental conditions to identify externally forced signals in biological time series. In: Müller F, Baessler C, Schubert H, Klotz S (eds.) *Long-Term Ecolocical Research: Between Theory and Application*. Dordrecht, Netherlands: Springer, 2010, 155-62.

Störmer R. Benthic bacteria in the German Bight: Characterising community structure and influencing environmental factors volume PhD: University of Hamburg, 2013.

Störmer R, Wichels A and Gerdts G. Geo-Chip analysis reveals reduced functional diversity of the bacterial community at a dumping site for dredged Elbe sediment. *Mar Pollut Bull* 2013;**77**: 113-22.

Sündermann J and Pohlmann T. A brief analysis of North Sea physics. *Oceanologia* 2011;**53**: 663-89.

Szabó KÉ, Itor POB, Bertilsson S, Tranvik L and Eiler A. Importance of rare and abundant populations for the structure and functional potential of freshwater bacterial communities. *Aquat Microb Ecol* 2007;**47**: 1-10.

Tabachnick BG and Fidell LS. *Using Multivariate Statistics.* New York, USA:HarperCollins, 1996.

Tada Y, Makabe R, Kasamatsu-Takazawa N, Taniguchi A and Hamasaki K. Growth and distribution patterns of *Roseobacter/Rhodobacter*, SAR11, and *Bacteroidetes* lineages in the Southern Ocean. *Polar Biol* 2013;**36**: 691-704.

Teeling H, Fuchs BM, Becher D, Klockow C, Gardebrecht A, Bennke CM *et al.* Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* 2012;**336**: 608-11.

Teira E, Gasol JM, Aranguren-Gassis M, Fernández A, González J, Lekunberri I *et al.* Linkages between bacterioplankton community composition, heterotrophic carbon cycling and environmental conditions in a highly dynamic coastal ecosystem. *Environ Microbiol* 2008;**10**: 906-17.

Thingstad TF. Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnol Oceanogr* 2000;**45**: 1320-8.

Unrein F, Massana R, Alonso-Sáez L and Gasol JM. Significant year-round effect of small mixotrophic flagellates on bacterioplankton in an oligotrophic coastal system. *Limnol Oceanogr* 2007;**52**: 456-69.

Vázquez SC, Hernández E and Mac Cormack WP. Extracellular proteases from the antarctic marine *Pseudoalteromonas* sp. P96-47 strain. *Rev Argent Microbiol* 2008;**40**: 63-71.

von Scheibner M, Dörge P, Biermann A, Sommer U, Hoppe H-G and Jürgens K. Impact of warming on phyto-bacterioplankton coupling and bacterial community composition in experimental mesocosms. *Environ Microbiol* 2014;**16**: 718-33.

von Storch H and Zwiers FW. *Statistical analysis in climate research.* Cambridge, UK:Cambridge University Press, 1999.

Wang K, Ye X, Chen H, Zhao Q, Hu C, He J *et al.* Bacterial biogeography in the coastal waters of northern Zhejiang, East China Sea is highly controlled by spatially structured environmental gradients. *Environ Microbiol* 2015, DOI 10.1111/1462-2920.12884.

Weiss M and Simon M. Consumption of labile dissolved organic matter by limnetic bacterioplankton: the relative significance of amino acids and carbohydrates. *Aquat Microb Ecol* 1999;**17**: 1-12.

Wemheuer B, Güllert S, Billerbeck S, Giebel H-A, Voget S, Simon M *et al.* Impact of a phytoplankton bloom on the diversity of the active bacterial community in the southern North Sea as revealed by metatranscriptomic approaches. *FEMS Microbiol Ecol* 2014;**87**: 378-89.

Whitman WB, Coleman DC and Wiebe WJ. Prokaryotes: The unseen majority. *Proc Natl Acad Sci* 1998;**95**: 6578-83.

Wiltshire K, Kraberg A, Bartsch I, Boersma M, Franke H-D, Freund J *et al.* Helgoland Roads, North Sea: 45 Years of Change. *Estuaries Coasts* 2010;**33**: 295-310.

Wiltshire KH, Malzahn AM, Wirtz K, Janisch S, Mangelsdorf P and Manly BFJ. Resilience of North Sea phytoplankton spring bloom dynamics: An analysis of long-term data at Helgoland Roads. *Limnol Oceanogr* 2008: 1294-302.

Woese CR. Bacterial Evolution. *Microbial Rev* 1987;**51**: 221-71.

Wohlers-Zöllner J, Biermann A, Engel A, Dörge P, Lewandowska A, von Scheibner M *et al.* Effects of rising temperature on pelagic biogeochemistry in mesocosm systems: A comparative analysis of the AQUASHIFT Kiel experiments. *Mar Biol* 2012;**159**: 2503-18.

Xing P, Hahnke RL, Unfried F, Markert S, Huang S, Barbeyron T *et al.* Niches of two polysaccharide-degrading *Polaribacter* isolates from the North Sea during a spring diatom bloom. *ISME J* 2014, DOI 10.1038/ismej.2014.225.

Yan S, Fuchs BM, Lenk S, Harder J, Wulf J, Jiao N-Z *et al.* Biogeography and phylogeny of the NOR5/OM60 clade of *Gammaproteobacteria*. *Syst Appl Microbiol* 2009;**32**: 124-39.

Yilmaz P, Iversen MH, Hankeln W, Kottmann R, Quast C and Glöckner FO. Ecological structuring of bacterial and archaeal taxa in surface ocean waters. *FEMS Microbiol Ecol* 2012;**81**: 373-85.

Zhang Y-J, Zhang X-Y, Zhao H-L, Zhou M-Y, Li H-J, Gao Z-M *et al. Idiomarina maris* sp. nov., a marine bacterium isolated from sediment. *Int J Syst Evol Micr* 2012;**62**: 370-5.

Zhang Y, Jiao N, Cottrell MT and Kirchman DL. Contribution of major bacterial groups to bacterial biomass production along a salinity gradient in the South China Sea. *Aquat Microb Ecol* 2006;**43**: 233-41.

Zubkov MV, Fuchs, B.M., Archer, S.D., Kiene, R.P., Amann, R., Burkill, P.H. Linking the composition of bacterioplankton to rapid turnover of dissolved dimethylsulphoniopropionate in an algal bloom in the North Sea. *Environ Microbiol* 2001;**3**: 304-11.

ACKNOWLEDGEMENTS

First of all, I would like to thank my PhD Thesis Committee, namely Prof. Dr. Rudolf Amann, PD Dr. Bernhard Fuchs (Max Planck Institute for Marine Microbiology), Dr. Ulrich Callies (Helmholtz Helmholtz Centre Geesthacht) and Dr. Gunnar Gerdts and Dr. Antje Wichels (Alfred Wegener Institute Helmholtz Centre for Polar Marine Research) for excellent supervision, fruitful discussions and valuable input.
I also would like to thank Prof. Dr. Ulrich Fischer (Max Planck Institute for Marine Microbiology) for his interest in my work and for beeing a member of the examination board.

Many thanks to the Alfred-Wegener-Institute Helmholtz Center for Polar and Marine Research for funding. I am also grateful for financial and scientific support by The International Max Planck Research School of Marine Microbiology (MarMic).

I am deeply indebted to each and every person who encouraged and inspired me during the last three years. In particular I want to express my gratitude to my direct supervisors Antje Wichels and Gunnar Gerdts, for always sharing their scientific knowledge, for guiding me but still letting me find my own way and teaching me so much throughout the years. I will always be thankful for your advice, patience, trust and support.

I would like to thank the crews of the Aade and Uthörn, Hilke Döpke, Sara Dehne and a number of bachelor and master students for sampling support. Special thanks to Kristine Carstens and Silvia Peters, for constant motivation during numerous Uthörn cruises and for being such perfect “ship moms”.

Special thanks to all former and current members of the working group “Microbial Ecology”, for the cheerful atmosphere and many amazing moments. I would like to thank Irina Köster for the valuable contribution to this thesis. I really enjoyed working with you and teaching you. I also would like to thank all my colleagues from the Biological Station Helgoland for all the helping hands, support and friendly words.

I would like to acknowledge all my dear friends who have accompanied me for so many years. Thank you so much for endless motivation and countless phone calls, for always hosting me, believing in me and for keeping me laughing and sane.

I am also hugely grateful for the love and support I have received from my partner. Thank you Flo for listening to various weird DNA and bacteria related thoughts during uncountable hours and for helping me to take my mind off these things as well. I would like to thank you for all your affection and encouragement and for keeping me grounded.

Last but not least I would like to thank my family for the invaluable moral support throughout the years of my studies. Thanks for letting me find and follow my own way and for giving me the willingness, endurance and ambition to follow things through.

Name: Judith Lucas **Ort, Datum:** Helgoland, 08. September 2015

Anschrift: Gätkestraße 510 C, 27498 Helgoland

Erklärung

Hiermit erkläre ich, dass ich die Arbeit mit dem Titel

Spatiotemporal dynamic of the bacterial community in the German Bight

selbstständig und ohne unerlaubte fremde Hilfe verfasst und keine anderen als die hier angegebenen Quellen und Hilfsmittel verwendet habe. Die den genutzten Werken wörtlich oder inhaltlich entnommenen Stellen habe ich als solche kenntlich gemacht.

Ebenfalls erkläre ich hiermit eidesstattlich, dass es sich bei den von mir abgegebenen Arbeiten um 3 identische Exemplare handelt.

(Unterschrift)