Distribution and function of marine *Bacteroidetes*

Dissertation

Zur Erlangung des Grades eines

Doktors der Naturwissenschaften

– Dr. rer. nat. –

Dem Fachbereich Biologie/Chemie der

Universität Bremen

vorgelegt von

Christin M. Bennke

Bremen

Februar 2014

Die vorliegende Arbeit wurde in der Zeit von Oktober 2010 bis Februar 2014 am Max-Planck-Institut für marine Mikrobiologie in Bremen angefertigt.

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

1. Prüfer: PD Dr. Bernhard M. Fuchs
2. Prüfer: Prof. Dr. Ulrich Fischer

Tag des Promotionskolloquiums: 27. März 2014

Content

Summary

Members of the phylum *Bacteroidetes* play a pivotal role in degrading organic matter and appear everywhere in marine and freshwater systems, from coastal to open ocean, from polar to equatorial, from surface waters down to the deep sea as well as in association with aggregates and with phytoplankton blooms. The studies described in this thesis elaborate on the distribution and function of marine *Bacteroidetes*. Specifically their association with spring phytoplankton blooms, substrate association by direct surface attachment and their genetic capability of degrading high molecular weight organic matter and in particular polysaccharides were examined.

The *Bacteroidetes* distribution and community structure were analyzed at a temporal scale, by investigating the responses of distinct bacteroidetal clades during and after spring phytoplankton blooms of four consecutive years at the coastal station Helgoland Roads. It could be shown by automated microscopic cell counting that shortly after the chlorophyll *a* maximum concentration *Bacteroidetes* increased to more than 50% of the total bacterioplankton community during spring seasons. The *Bacteroidetes* community comprised only a few dominant genera, which accounted together for more than half of the *Bacteroidetes*. Each year a distinct succession pattern of the clades *Ulvibacter, Formosa* A, and *Polaribacter* was observed with relative abundances of single clades with up to 20%. Furthermore, members of the *Bacteroidetes* inhabited not only the free-living fraction, but they were also found attached to diatoms. Although a quantification of attached *Bacteroidetes* was difficult, qualitative observations were made. For example members of this phylum attach frequently to the diatom *Chaetoceros* spp., which is commonly blooming in spring at Helgoland Roads. The clades *Polaribacter* and *Formosa* A were identified as dominating among those *Chaetoceros*-associated *Bacteroidetes*. In contrast, *Ulvibacter* was not found attached to *Chaetoceros*, but to *Asterionella* spp., another diatom genus occurring in spring blooms.

Since members of *Bacteroidetes* are the first in responding to algal blooms and attached even to distinct diatom species, we investigated their genetic potential to degrade algal derived organic matter. In particular we searched for the presence of polysaccharide utilization loci (PULs) in fosmids retrieved from two contrasting provinces of the North Atlantic Ocean. In total 14 PULs were identified, six on fosmids from the northern station and eight on fosmids from the southern station. Among those PULs one seems to be involved in xylan degradation and four were identified as potential laminarin degradation PULs. Interestingly, GHs were identified which had been assumed to be unique among terrestrial *Flavobacteria*, suggesting a higher capability of open ocean *Bacteroidetes* clades for organic matter degradation than previously anticipated.

Zusammenfassung

Bacteroidetes sind dafür bekannt, ubiquitär verbreitet und wesentlich am Abbau von organischem Material beteiligt zu sein. Vertreter des Phylums sind sowohl in marinen und Süßwassersystemen als auch auch an den Küsten bis hin zum offenen Ozean, von polaren bis hin zu äquatorialen Bereichen, vom Oberflächenwasser bis in die Tiefsee sowie in Assoziation mit Aggregaten und Phytoplanktonblüten zu finden.

Die in dieser Arbeit beschriebenen Studien gehen näher auf die Verteilung und Funktion von marinen *Bacteroidetes* ein. Besonders wurden deren Reaktion auf die Phytoplanktonblüten im Frühjahr, deren Substratpräferenzen zur direkten Oberflächenanhaftung an Algen und deren genetische Fähigkeit, hochmolekulares, organisches Material in Form von Polysacchariden abzubauen, untersucht.

Die Zusammensetzung und Verbreitung der *Bacteroidetes*-Gemeinschaft wurde über einen Zeitraum von vier Jahren am Standort Helgoland „Kabeltonne" untersucht, wobei besonders das Auftreten von bestimmten Untergruppen während und nach der Frühjahrsblüte analysiert wurde. Es konnte mittels automatischer Mikroskopie gezeigt werden, dass im Frühjahr kurz nach dem Chlorophyll-*a*-Maximum mehr als 50% des Bakterioplanktons aus *Bacteroidetes* bestehen. Die *Bacteroidetes* Gemeinschaft wird von wenigen einzelnen Untergruppen dominiert, welche zusammen jedoch mehr als die Hälfte der *Bacteroidetes* ausmachen. So konnte überraschenderweise in jedem der untersuchten Jahre eine definierte Sukzession von *Ulvibacter*, *Formosa* A und *Polaribacter* nachgewiesen werden. Jede dieser Untergruppen besaß eine relative Abundanz von bis zu 20%. Weiterhin konnte gezeigt werden, dass *Bacteroidetes* nicht nur in der freilebenden Fraktion, sondern auch angeheftet an Diatomeen zu finden waren. Die Quantifizierung von angehefteten *Bacteroidetes* erwies sich als schwierig und deshalb wurden nur qualitativ Auswertungen durchgeführt. Die Gruppen *Polaribacter* und *Formosa* A wurden hauptsächlich an *Chaetoceros,* eine vor Helgoland regelmäßig blühende Diatomeengattung, assoziiert gefunden. Im Gegensatz dazu heftete sich *Ulvibacter* präferentiell an eine andere Diatomeengattung an, nämlich *Asterionella.*

Da Vertreter der *Bacteroidetes* offensichtlich stark auf Algensubstrate reagieren, wurde deren genetisches Potenzial, Algenpolymere abzubauen, untersucht. Wir fanden unter mehr als 200 Fosmiden 14 sogenannte „Polysaccharide Utilization Loci" (PULs), welche von zwei unterschiedlichen Stationen des Nordatlantischen Ozeans stammen, sechs von der nördlichen und acht von der südlichen Station. Unter diesen PUL Fosmiden wurde eins identifiziert, welches möglicherweise am Xylanabbau beteiligt ist und vier weitere besitzen die genetische Fähigkeit, Laminarin abzubauen. Interessanterweise wurden einige Glycosylhydrolasen identifiziert, von denen angenommen wird, dass sie nur in Flavobakterien terrestrischer Herkunft vorkommen. Dies deutet daraufhin, dass *Bacteroidetes*-Gruppen aus dem offenen Ozean eine höhere Fähigkeit besitzen organisches Material abzubauen als ursprünglich angenommen.

Chapter 1: Introduction

1.1 Microbial diversity: "*everything is everywhere*, but *the environment selects*" (BAAS BECKING 1934)

Following the statement of BAAS BECKING (1934) many scientists studied microbial biodiversity and biogeographical patterns (e.g. FUHRMAN ET AL. 1993, STALEY & GOSINK 1999, GIOVANNONI & STNGL 2005, POMMIER ET AL. 2005, JOHNSON ET AL. 2006, SCHATTENHOFER ET AL. 2009, GÓMEZ-PEREIRA ET AL. 2010). Microorganisms, in general, are everywhere on Earth, inhabiting diverse environments, e.g. aquatic habitats (freshwater and marine), soils and sediments, living in symbiotic relationships or being pathogenic. They are referred to as the "unseen majority" and their total number is estimated to range between 4 – 6 x 10^{30} cells. Microorganism inhabiting the world's oceans account for 1.2 x 10^{29} cells. They are outnumbered (3.5 x 10^{30} cells) by those occupying the ocean floor and subsurface (WHITMAN ET AL. 1998). In fact, oceans, including their coastal realms, harbor a vast diversity of different organisms ranging in size from several meters (e.g. whales can exceed 30 m) to micrometers (picoplankton in the range of 2 – 0.2 μm). Although, microorganisms have a small cell size, their total biomass exceeds the combined mass of fish and zooplankton in the oceans (POMEROY ET AL. 2007). All these different organisms interact in complex food webs.

Although, aquatic microorganisms are everywhere their community composition differs greatly. For example, there are clades having a preference for marine environments, such as *Alphaproteobacteria, Gammaproteobacteria* or *Bacteroidetes*, while others like *Betaproteobacteria* or *Actinobacteria* are prominent members of the freshwater bacterioplankton (GLÖCKNER ET AL. 1999, NEWTON ET AL. 2011).

1.2 Main marine bacterioplankton groups

In marine habitats the majority of microbes fall into a few clades of *Bacteria* and *Archaea* (Figure 1.1; GIOVANNONI & RAPPÉ 2000). Two phyla – *Proteobacteria* and *Bacteroidetes* – dominate the heterotrophic picoplankton. Among the *Proteobacteria* – *Alphaproteobacteria* and *Gammaproteobacteria* are the dominant clades (KIRCHMAN 2002).

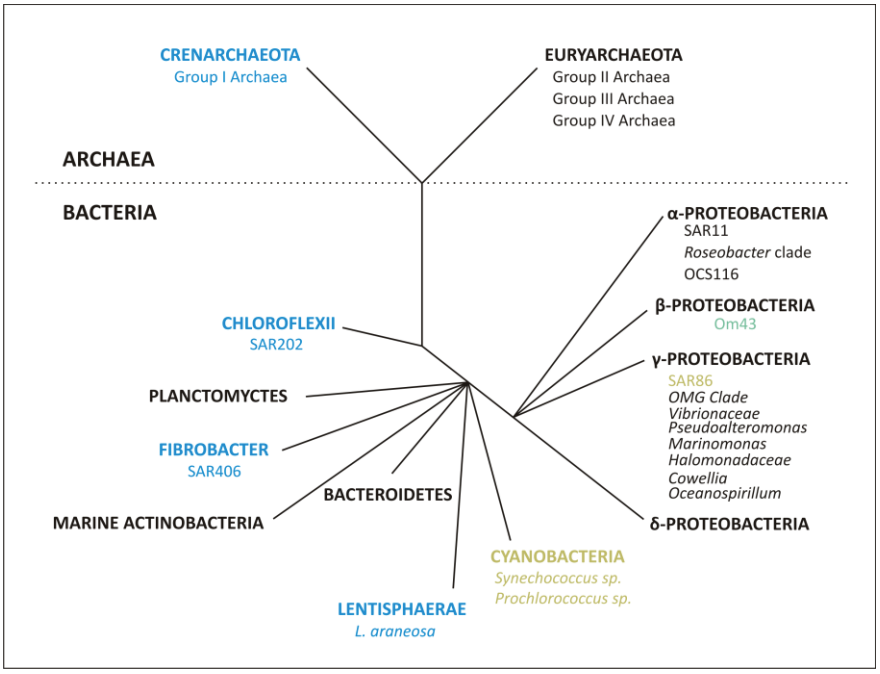


Figure 1.1: Overview of major marine picoplankton clades. Blue letters indicate groups which occur in mesopelagic and surface waters in polar winter; gold indicates bacteria belonging to the photic zone; green indicates clades which are associated with coastal ocean ecosystems and black indicates those groups which are present everywhere in the seawater (modified after GIOVANNONI & STINGLE 2005).

1.2.1 *Alphaproteobacteria*

Members of the *Alphaproteobacteria* are ubiquitously distributed (NEWTON ET AL. 2011), but found in higher abundances in marine than in lacustrine habitats (GLÖCKNER ET AL. 1999, ZWART ET AL. 2002). For example, the SAR11 clade represents the most abundant clade in the ocean´s surface layers, comprising up to 50% of all picoplankton cells (MORRIS ET AL. 2002). In freshwater systems they are also highly abundant, but consist mainly of the LD12 lineage (BAHR ET AL. 1996, ZWART ET AL. 1998, ZWART ET AL. 2002, SALCHER ET AL. 2011, HEINRICH ET AL. 2013). Another alphaproteobacterial clade, *Roseobacter*, is less abundant, but widley distributed in marine environments (BUCHAN ET AL. 2005). Some *Roseobacter* linages are found everywhere, in many different marine habitats such as the DC5-80-3 cluser (BUCHAN ET AL. 2005), others are found in association with eukaryotic marine organisms (RUIZ-PONTE ET AL. 1998) and even others like the DG1128 or NAC11-7 cluster were identified in close relationship to macroalgae and phytoplankton blooms (SCHÄFER ET AL. 2002).

1.2.2 *Gammaproteobacteria*

The *Gammaproteobacteria* are highly abundant in marine systems, often with preferences for coastal environments (RAPPÉ ET AL. 1997, EILERS ET AL. 2001). Members of this class are considered to degrade rapidly the more easily accessible fraction of organic matter (BRETTAR ET AL. 2006). For example, they were identified as one of the bacterioplankton clades taking up glucose in late winter and during a spring phytoplankton bloom in the German Bight of the North Sea (ALONSO ET AL. 2006). It was shown by automated ribosomal intergenic spacer analysis (ARISA) and denaturing gradient gel electrophoresis (DGGE) that about one fourth of the bacterial community in the German Bight was assigned to *Gammaproteobacteria*. Moreover, specific clades were found in strong association with a spring phytoplankton bloom (SAPP ET AL. 2007).

1.2.3 *Bacteroidetes*

The second phylum, which dominates the heterotrophic picoplankton, is *Bacteroidetes*. This phylum is one of the most abundant (especially in coastal areas) clades of *Bacteria*, omnipresent in the world's oceans and also in freshwater habitats (GLÖCKNER ET AL. 1999, SIMON ET AL. 1999, ALONSO ET AL. 2007, ZEDER ET AL. 2009, GÓMEZ-PEREIRA ET AL. 2010). The importance of *Bacteroidetes* in heterotrophic degradation of organic matter and their presence in algal blooms as well as colonizing particles have been shown in many studies (e.g. DELONG ET AL. 1993, SIMON ET AL. 1999, BARBEYRON ET AL. 2001, PINHASSI ET AL. 2004, WOEBKEN ET AL. 2007, SAPP ET AL. 2007, PEDROTTI ET AL. 2009, TADA ET AL. 2011, THOMAS ET AL. 2011) and will be further analysed in this thesis.

1.3 The phylum *Bacteroidetes*

The phylum *Bacteroidetes* is also known as *Cytophaga-Flavobacterium-Bacteroides*-cluster (WOESE 1987). The very same phylum has been referred in literature also as *Bacteroides-Cytophaga-Flexibacter* group (HUGENHOLTZ ET AL. 1998) as well as *Cytophaga-Flavobacteria* cluster (KIRCHMAN 2002) and other variations thereof. Members of this phylum are heterotrophic, gram-negative bacteria which live mainly aerobic in different environments or anaerobic as part of e.g. the gut flora (KIRCHMAN 2002). Some of them represent important members of the intestinal flora (XU ET AL. 2003) and others may be pathogenic (SMITH ET AL. 2006) or live in soils (EILERS ET AL. 2012). However, the majority is present in aquatic habitats e.g. freshwater (GLÖCKNER ET AL. 1999, ZEDER ET AL. 2009), coastal (CRUMP ET AL. 1999, EILERS ET AL. 2001, KIRCHMAN ET AL. 2005, ALONSO ET AL. 2007) and open ocean waters (SCHATTENHOFER ET AL. 2009, GÓMEZ-PEREIRA ET AL. 2010, DÍEZ-VIVES ET AL. 2014), sediments (ISHII ET AL. 2004), hydrothermal vents (SIEVERT ET AL. 2000, KORMAS ET AL. 2006), polar regions (BANO & HOLLIBAUGH 2002, ABELL & BOWMAN 2005, GRZYMSKI ET AL. 2006, MALMSTROM ET AL. 2007), deep-sea

(SCHAUER ET AL. 2010, RUFF ET AL. in rev) and many are associated with marine phytoplankton (SIMON ET AL. 1999, RIEMAN ET AL. 2000, PINHASSI ET AL. 2004, FANDINO ET AL. 2005).

The phylum *Bacteroidetes* consists of four classes (Figure 1.2): the class *Bacteroidia*, *Flavobacteria*, *Sphingobacteria* and *Cytophagia*. The latter was recently ranked as independent class based on 16S rRNA sequence identities (LUDWIG ET AL. 2008). Most members of the class *Flavobacteria* represent the marine clade and only some from *Sphingobacteria* and *Cytophagia*. Members of the class *Bacteroidia* are mainly gut-associated.

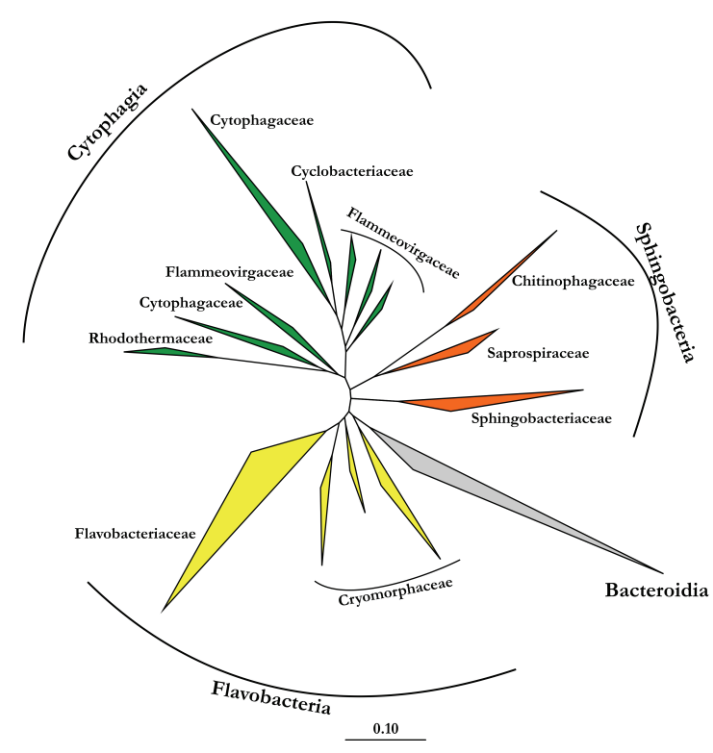


Figure 1.2: Overview of a 16S rRNA gene based phylogenetic tree of the *Bacteroidetes* phylum including classes (color coded) and some families. Phylogenetic tree was adapted from the "All-Species Living Tree" (YARZA et al. 2008), release 111 (February 2013). Tree reconstruction was based on the maximum likelihood algorithm RAxML.

1.3.1 The marine clade of *Flavobacteria*

The class *Flavobacteria* represents a very broad lineage and contains, e.g., the families *Flavobacteriaceae*, *Cryomorphaceae* and *Blattabacteriaceae* (LUDWIG ET AL. 2008). The family *Blattabacteriaceae* consists of non-motile intracellular symbionts of cockroach and termite species (CLARK & KAMBHAMPATI 2003, BERNARDET 2010). Among the family *Cryomorphaceae* only a few members have been isolated from seawater, but most members of the family *Flavobacteriaceae* are from marine origin.

The family *Flavobacteriaceae* is much large and in the latest version of 'Bergey's Manual of Systematic Bacteriology' (editorial deadline June 2006) consists of more than 61 genera, some of which contain up to 40 valid species (BERNARDET 2010). The latest release of 'The All-Species Living Tree Project' (LTP; release February 2013; YARZA ET AL. 2008, MUNOZ ET AL. 2011) comprised 391 species in 81 genera. Members of the *Flavobacteriaceae* seem to be abundant in marine, mostly polar habitats (SIMON ET AL. 1999, KIRCHMAN ET AL. 2003, BANO & HOLLIBAUGH 2002, ABEL & BOWMAN 2005, BOWMAN ET AL. 2006, GÓMEZ-PEREIRA ET AL. 2010), playing an important role in organic matter degradation (BARBEYRON ET AL. 2001, HUMPHRY ET AL. 2001, ELIFANTZ ET AL. 2005, DESCAMPS ET AL. 2006, THOMAS ET AL. 2011). Various species, organized in up to 40 genera, have been isolated from marine habitats and most genera (26) consist of species with a free-living or saprophytic lifestyle. Six out of 40 genera are solely saprophytic and 8 prefer being free-living (BERNARDET 2010). Several *Flavobacteriaceae* members were found in association with marine phytoplankton (SIMON ET AL. 1999, FANDINO ET AL. 2001, NICOLAS ET AL. 2004, PINHASSI ET AL. 2004, FANDINO ET AL. 2005, GÓMEZ-PEREIRA ET AL. 2010) and diatoms (GROSSART ET AL. 2005, SAPP ET AL. 2007) or detritus (ABEL & BOWMAN 2005, PINHASSI ET AL. 2005). More recently, members of this family were isolated from marine macroalgae, such as *Formosa algae* which was isolated from the brown algae *Fucus evanescens* (IVANOVA ET AL. 2004) or *Formosa agariphila* from the green algae *Acrosiphonia sonderi* (NEDASHKOVSKAYA ET AL. 2006). *Ulvibacter litoralis*

(NEDASHKOVSKAYA ET AL. 2004) were isolated for example from the green algae *Ulva fenestrate*. This family comprises of species inhabiting diverse ecological niches and even within one genera, species can show contrasting characteristics (BERNARDET & NAKAGAWA 2006).

1.4 Biogeography of *Bacteroidetes*

Most picoplankton groups, including *Bacteroidetes*, show seasonal (FUHRMAN ET AL. 2006) and latitudinal (FUHRMAN ET AL. 2008) patterns in their distribution. For example, DÍEZ-VIVES ET AL. (2014) recently analyzed the *Bacteroidetes* community of the northwestern Mediterranean Sea at temporal and spatial scale. They found a marked seasonality of the *Bacteroidetes* community, comprising a winter-spring cluster and a summer-fall cluster. Whereas, members of the *Polaribacter* clade were higher abundant during winter and spring, members of the *Owenweeksia* clade revealed higher numbers in summer and fall. However, the overall *Bacteroidetes* abundance was higher in spring and decreased towards winter. Additionally, they analysed the *Bacteroidetes* distribution along an inshore – offshore transect and revealed that the *Bacteroidetes* abundance was similar in all surface samples, but decreased with depth. This was also shown for other sampling sites such as the Atlantic Ocean (SCHATTENHOFER ET AL. 2009, GÓMEZ-PEREIRA ET AL. 2010).

1.4.1 Distribution pattern in the North Atlantic Ocean

GÓMEZ-PEREIRA ET AL. (2010) investigated the *Bacteroidetes* community of the North Atlantic Ocean between the East Greenland current and the North Atlantic subtropical gyre crossing four biogeographical provinces. They showed that *Bacteroidetes* have a preference for provinces with high chlorophyll *a* concentrations and nutrient-rich waters, where their highest numbers of about 20% were obtained. At the North Atlantic subtropical gyre, which is characterized by low chlorophyll *a* concentrations and nutrient depletion, *Bacteriodetes* accounted for only 7% of the total bacterioplankton community. Additional depth profiles revealed that their abundances decreased with depth. However, the *Bacteroidetes* abundances were still higher in the nutrient-rich Boreal Polar Province at 170 m depth than in the surface waters of the oligotrophic subtropical gyre region (Figure 1.3A), indicating a strong relationship to zones with high nutrients and chlorophyll *a* concentrations.

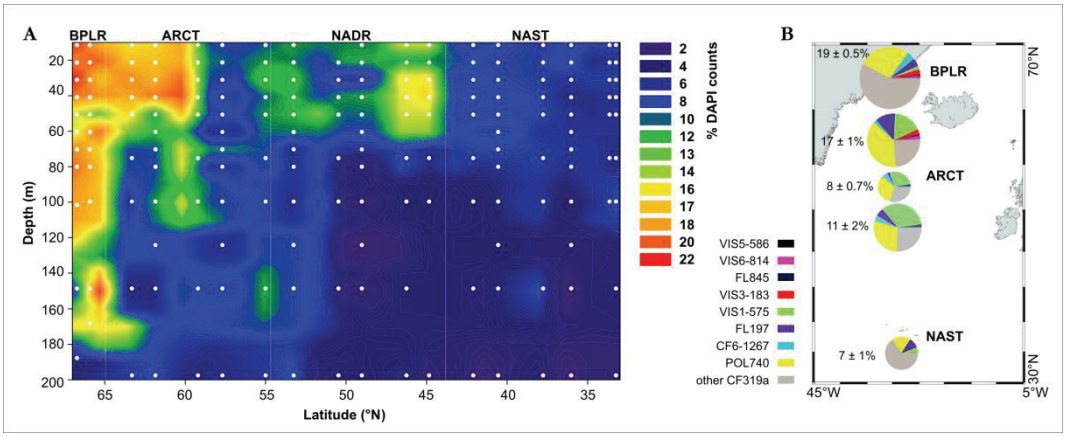


Figure 1.3: Biogeographical distribution of Bacteroidetes in the North Atlantic Ocean adapted from GÓMEZ-PEREIRA ET AL. 2010. A) Latitudinal depth distribution of *Bacteroidetes*, *Bacteroidetes*, detected by probe CF319a. Dots represent sampling depths. B) Relative contribution of each *Bacteroidetes* clade to the total CF319a. Area of the pie charts represent the relative abundance of CF319a (percentage given to the left).

A distinct distribution pattern of *Bacteroidetes* clades were obtained across the transect. For example, the *Polaribacter* clade was identified as the most abundant group of all analysed clades and was present at all stations, but had its highest absolute abundances in the Boreal Polar and Arctic Province. The NS5 clade, consisting of two subclades – the DE2 and VIS1 subclade – showed a distinct spatial preference. Whereas members of the NS5-DE2 clade accounted for a substantial fraction of the *Bacteroidetes* in the Boreal Polar Province, the NS5-VIS1 clade increased significantly in the Arctic Region (Figure 1.3B).

Two fosmid-based metagenome libraries were constructed from surface water samples retrieved from the Boreal Polar and North Atlantic Subtropical Gyre region to provide first insights into the genomic potential of yet uncultured *Bacteroidetes* inhabiting these oceanic provinces. The first 76 fosmids revealed a larger potential for polysaccharide degradation and cell surface attachment to algal cells in the nutrient and phytoplankton-rich Boreal Polar Region, than in the oligotrophic Subtropical Region (GÓMEZ-PEREIRA ET AL. 2012). Within the present study additional 155 fosmids from both regions will be analysed regarding their genetic potential in organic matter degradation.

1.4.2 Distribution pattern in coastal areas

In other cold, nutrient-rich and high chlorophyll *a* systems, such as the Antarctic Ocean, members of the *Bacteroidetes* accounted for up to 72% in the marginal ice zone, while a *Phaeocystis* bloom occurred (SIMON ET AL. 1999). Clearly, *Phaeocystis*, surrounded by a substrate-rich mucus layer, favored the blooming of psychrophilic *Bacteroidetes* clades. Furthermore, members of this phylum were also detected in high numbers in lakes during phytoplankton blooms (EILER ET AL. 2007, ZEDER ET AL. 2009). EILERS ET AL. (2001) identified more than 50% of *Bacteroidetes* during a spring phytoplankton bloom in the German Bight of the North Sea. Members of this phylum were also present during summer and autumn

phytoplankton blooms with up to 40% of the total picoplankton (ALONSO ET AL. 2007). All these findings suggest a close connection of *Bacteroidetes* to phytoplankton blooms. However, bacterial response to phytoplankton blooms were almost exclusively studied in microcosm/mesocosm experiments (RIEMANN ET AL. 2000, PINHASSI ET AL. 2004, RINTA-KANTO ET AL. 2012). Only a few studies investigated phytoplankton blooms *in situ* (FANDINO ET AL. 2001, LAU ET AL. 2007, TADA ET AL. 2011), but they had a limited resolution in time and biodiversity.

In 2009 the MIMAS (Microbial Interactions in Marine Systems) Project was established to examine the diversity and function of marine bacterioplankton in coastal environments. By coupling the diversity to the metabolic potential, the adaptation of marine bacteria to a changing environment was explored. Phytoplankton blooms are a good example for a changing environment, since during their scenescence different algal primary products become available, thus serving as a food source for specific bacterioplankton clades (TEELING ET AL. 2012).

1.5 Phytoplankton Blooms

Many bacteria show a strong association with phytoplankton, especially with diatoms (GROSSART ET AL. 2005, SAPP ET AL. 2007, GÄRDES ET AL. 2011, AMIN ET AL. 2012). Consequently, phytoplankton blooms have an impact on the composition of the heterotrophic bacterioplankton community and in particular on members of the *Bacteroidetes*. Annually recurring phytoplankton blooms are a phenomenon in coastal zones of higher latitudes (GATTUSO ET AL. 1998). There the phytoplankton is responsible for around 19% of the net oceanic primary production, albeit these zones make up less than 7% of the ocean surfaces (FIELD ET AL. 1998). About half of this net primary production is remineralized by heterotrophic bacteria in a process called the 'microbial loop' (AZAM 1998).

Phytoplankton blooms, dominated by diatoms, typically occur in spring in well-mixed coastal areas such as the North Sea, when the turbidity (resulted from sediment resuspension and river plumes) drops below a certain threshold and subsequently light penetrates deeper layers (IRIATE & PURDIE 2004, TIAN ET AL. 2011). When conditions like light penetration and availability of nutrients and in particular silicate in the mixed water layers are favorable, diatoms and other phytoplankton members start to bloom (SMETACEK 1999, WILTSHIRE ET AL. 2008, LOHMANN & WILTSHIRE ET AL. 2012).

During their growth, phytoplankton fixes carbon, but also releases a major fraction of the photoassimilated carbon in form of dissolved organic matter (DOM), which can be taken up by heterotrophic bacteria (AZAM ET AL. 1983). Generally, phytoplankton blooms are both bottom-up and top-down controlled, which means that the termination of blooms can either be affected by nutrient depletion (bottom-up) in particular silicate (SMETACEK 1999, WILTSHIRE ET AL. 2008) or by other factors (top-down), like grazing or viral lysis (BEARE ET AL. 2002, WILTSHIRE ET AL. 2010). Then the phytoplankton biomass is partly converted into dissolved and colloidal matter through cell lysis, and subsequently converted back to particulate organice matter by uptake and cell growth of heterotrophic prokaryotes and thereby, not completely lost from transfer to higher trophic levels (AZAM ET AL. 1983).

1.6 Study areas

The main part of this thesis investigated the distribution and function of marine *Bacteroidetes* on samples collected during spring phytoplankton blooms at the coastal station Helgoland Roads, located in the German Bight, North Sea. The second part of this thesis investigated samples retrieved from the North Atlantic Ocean and is focused on metagenome samples to unveil the genetic potential of *Bacteroidetes* in two contrasting biogeographical provinces.

1.6.1 The North Sea and Helgoland Roads

The North Sea is a marginal sea and is located between the British Islands, The Netherlands, Germany, Scandinavia, Belgium and France. It is characterized as a typical semi-enclosed continental shelf sea, mainly influenced by tides, winds and circulation effects by the inflowing water masses. The German Bight is the southeastern bight of the North Sea and is affected by high saline water masses from the north as direct influx from the North Atlantic Ocean and from the southwest as indirect inflow via the English Channel, as well as from the east by low saline coastal water masses either from the Baltic Sea or from the river discharges (Figure 1.4; GREVE ET AL. 1996, VISSER ET AL. 1996, TIAN ET AL. 2011).

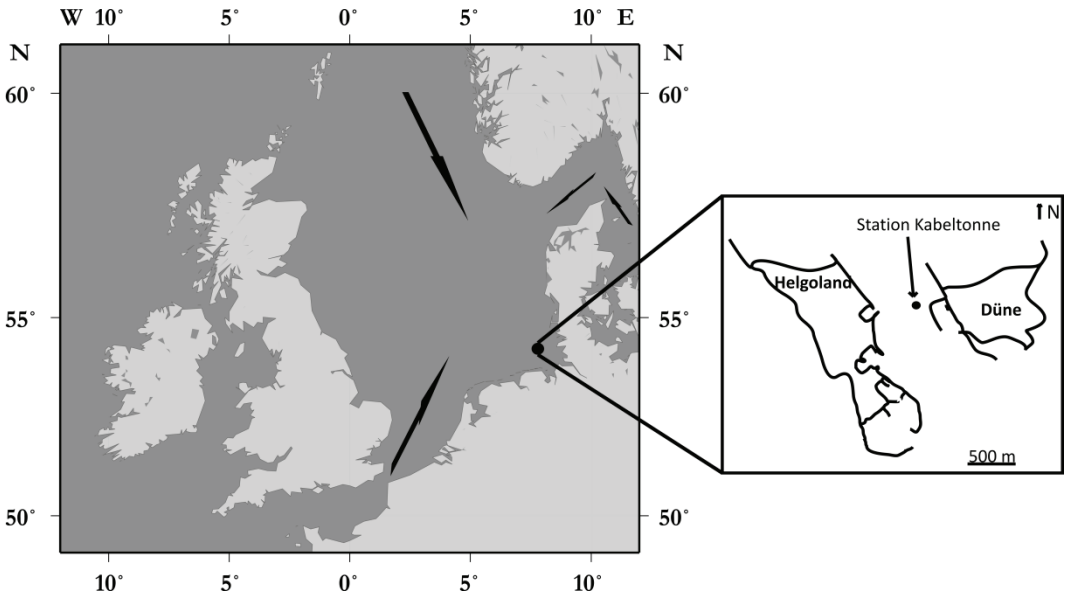


Figure 1.4: Location of Helgoland Island in the North Sea. Arrows indicate influence of different water masses. Map was created with an online map creation tool and modified manually (http://woodshole.er.usgs.gov/mapit/).

Helgoland is a small island in the North Sea about 60 km off the German coast line [54°11.30'N; 07°54.00'E] and is located at the border of the river plumes of Elbe, Weser and Eider and the seasonal thermal front within the complex mixing regime of the German Bight (GERLACH 1995, BECKER ET AL. 1999, RAABE & WILTSHIRE 2009, WILTSHIRE ET AL. 2010). The marine sampling site "Kabeltonne" at Helgoland Roads has been established as a long-term ecological research station

since 1962 (FRANKE ET AL. 2004). Since then daily measurements of physicochemical parameters such as sea surface temperature, salinity, turbidity via Secci depth, chlorophyll *a* and nutrients, as well as biological parameters like phytoplankton and zooplankton composition have been recorded (WILTSHIRE ET AL. 2008).

During the last 40 years a substantial increase of 1.67 °C in sea surface temperature was recorded at Helgoland Roads (WILTSHIRE ET AL. 2010) and the temperature can range from winter to summer between -2 °C and 22 °C (WHILTSHIRE & MANLY 2004). The salinity varies between 24 and 34 PSU (RAABE & WILTSHIRE 2009) and is strongly depending on the prevailing regime, either coastal or marine (DIPPNER 1993). High and low salinity events predominate in winter as well as in spring. Low salinity events result from an accumulation of riverine water near the coast and a subsequent nutrient transport towards Helgoland Roads. Those events are characterized by a decrease of 2 PSU within one day and simultaneously with an increase in silicate and nitrate concentrations (WILTSHIRE ET AL. 2010).

1.6.2 The North Atlantic Ocean

The Atlantic Oceans is the second largest Ocean on Earth and is separated by the Equator into the North Atlantic and the South Atlantic and by the Mid-Atlantic-Ridge into an eastern and western part (GERLACH 1994). It has the largest latitudinal extension than any other ocean of about 21,000 km and reaches north from the Bering Strait over the North Pole and Arctic ice sheet down south to the Antarctic continent (TOMCZAK & GODFREY 1994). Two large gyres, as part of the circular current system occur between 10° N/S and 40° N/S and two smaller gyres appear between 50° N/S to 60° N/S close to the poles as part of the subpolar current system. The subtropical gyre circulates clockwise, while the subpolar gyres moves counter-clockwise. The ocean currents are mainly wind-driven and therefore not stable at one place over years (SCHMITZ & MCCARTNEY 1993).

The subtropical gyre of the North Atlantic Ocean is restricted to the west by the Gulf Stream and to the east by the Canary Current, as well as to the north by the North Atlantic and Azores Current and to the south it is bounded by the North Equatorial Current (LUMBY 1932, SCHMITZ & MCCARTNEY 1993). The subpolar gyre is located between the East Greenland and Labrador Current as well as the Irminger and North Atlantic Current.

The North Atlantic Ocean and its circulation system is one of the best studied and observed systems of the world's ocean (SCHMITZ & MCCARTNEY 1993, ROSSBY 1996). Three main features make the North Atlantic Ocean unique:

1) Along the East Greenland coastline the East Greenland Current carries cold and low saline but nutrient rich waters southwards (BERSCH 1995).
2) The Gulf Stream and its extension currents – the North Atlantic Current (NAC) as well as the North Atlantic Drift Current – carries more warm water masses from the tropics into higher latitudes than any other ocean current and is therefore responsible for the moderate and mild climate in Scandinavia and Western Europe comparing to areas at the same latitude. If the NAC and its extensions decrease in heat and water transport the European climate would change dramatically (KRAUSS 1986, ARHAN 1990, ROSSBY 1996).
3) North of Greenland and Iceland dense and cold water masses sink below the surface and take part in the global ocean water circulation system (LUMBY 1932, SCHMITZ & MCCARTNEY 1993).

Since the first half of the 18th century the distribution of ice and mainly driftwood along the east Greenland coast and throughout the northern part of the North Atlantic was known. Since that time oceans were always explored to observe the global ocean circulation system. Nowadays scientists have further interest in the oceanic communities especially in microbial relationships and their activities in biogeochemical cycles as well in degradation of organic matter.

On this account the VISION cruise MSM03/01 was undertaken to investigate the diversity, structure and function of the microbial community including bacterio- and phytoplankton as well as virioplankton in September 2006 (21-30.09.2006). Samples were taken throughout the North Atlantic Ocean from the Arctic Circle, close to the east Greenland coast, alongside the 30° W meridian towards the Azore Islands (Figure 1.5). Thereby different biogeographical provinces – the Boreal Polar (BPLR), Arctic (ARCT), North Atlantic Drift (NADR) and North Atlantic Subtropical Gyre (NAST) Province – were passed. The bacterioplankton and especially the *Bacteriodetes* community, regarding subgroup distribution, was analysed by GÓMEZ-PEREIRA ET AL. (2010). Additionally, metagenome libraries from two contrasting stations (S3 and S18) were constructed to provide first insights into the genomic potential of yet uncultured *Bacteroidetes* inhabiting these oceanic provinces.

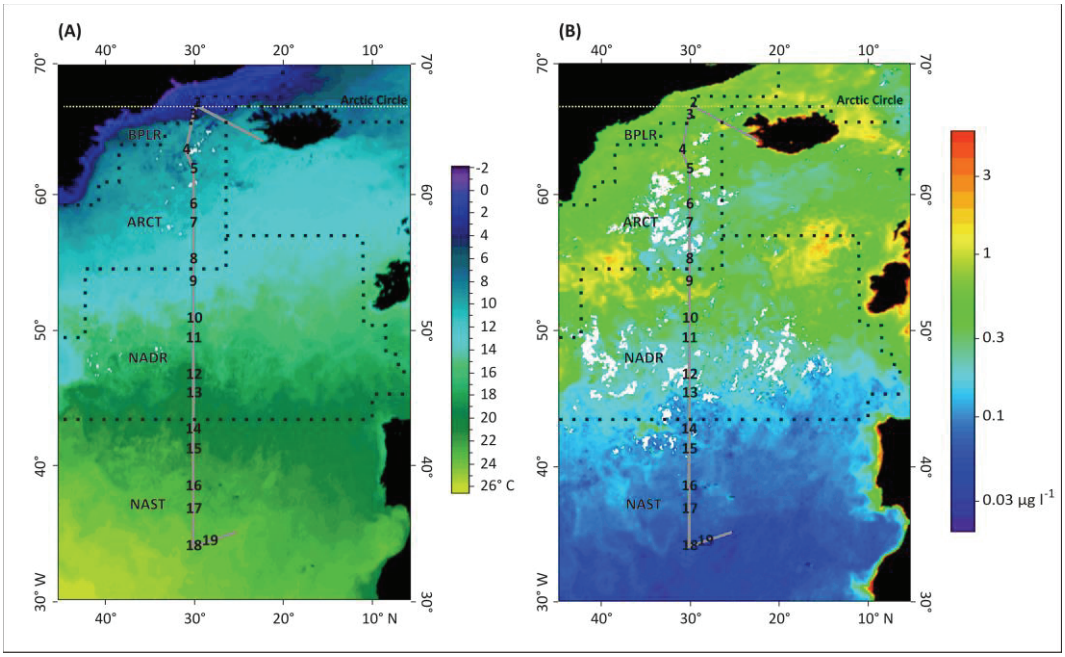


Figure 1.5: Overview of the VISION cruise track in the North Atlantic Ocean from the Arctic Circle towards the Azores Islands including their respective biogeographical provinces defined after LONGHURST (1998): the Boreal Polar Province (BPLR), Arctic Province (ARCT), the North Atlantic Drift Province (NADR) and the North Atlantic Subtropical Gyre Province (NAST) which are indicated by the dotted lines. (A) The monthly average of the sea surface temperature in September 2006; (B) The monthly average of the chlorophyll a concentration in September 2006 [provided by the AquaModis satellite (http://oceancolor.gsfc.nasa.gov; requested 7.05.2010)]. Missing data is shown as white.

The station 3 (S3) is located at the border of the Boreal Polar Province (BPLR) and is characterized by seas surface temperatures of 0.7 °C and salinity of 33 PSU at the surface and 34.7 PSU at 150 m depth. The chlorophyll a concentration was 0.7 mg m^{-3} and silicate had a concentration of 2.9 μM. Nitrate, nitrite and ammonia accounted together for 3.3 μM and phosphate was 0.4 μM. Station 18 (S18) is located in the North Atlantic Subtropical Gyre region and there the sea surface temperature was 23.5 °C. The salinity was 36 PSU and the chlorophyll a concentration was below 0.1 mg m^{-3}. All inorganic nutrients were below 0.3 μM, revealing a nutrient depletion at that station (GÓMEZ-PEREIRA ET AL. 2012).

1.7 Genes encoding the degradation and transport of polysaccharides

Many studies reported on the role of *Bacteroidetes* in organic matter degradation and cultured strains are known to be capable of degrading, e.g. cellulose, chitin and proteins (COTTRELL & KIRCHMAN 2000, COTTRELL ET AL. 2005, TEELING ET AL. 2012, FERNÁNDEZ-GÓMEZ ET AL. 2013). Recently, it has been shown, that *Flavobacteria* strains affiliated with the genera *Polaribacter* and *Formosa* were able to utilize a number of mono-, di-, tri- and polysaccharides, such as cellulose, including carboxymethylcellulose (which has been shown only for a few *Bacteroidetes* species), laminarin and xylan (HAHNKE 2013, PP. 197-198), corroborating the findings of ARNOSTI ET AL. 2012. They identified *Bacteroidetes* in the North Atlantic Ocean hydrolyzing enzymatically high molecular weight substrates. The hydrolysis of different polysaccharides, like laminarin, xylan, fucoidan and chondroitin sulfate, were measured and *Bacteroidetes* likely contributed to this hydrolysis, representing between 15 and 26% of the post-incubation community (ARNOSTI ET AL. 2012).

The degradation of organic material, mainly polysaccharides, is depending on the presence and expression of specific hydrolytic enzymes (WARREN 1996, ARNOSTI 2003). However, the initial breakdown of macromolecules, mainly catalyzed by extracellular enzymes, to smaller subunits (< 600 Da) is an important step within the degradation process (MCBRIDE ET AL. 2009). In the case of polysaccharides this degradation is often linked to the presence of TonB-dependent transporters (BRAUN & HERRMANN 1993, KOEBNIK 2005, SCHAUER ET AL. 2008). Such integrated systems were firstly described in the gut bacterium *Bacteroides thetaiotaomicron*. There, a set of genes encoding for binding, hydrolysis and transport of starch was named starch utilization system (sus) comprising of eight genes – *susA, susB, susC, susD, susE, susF, susG* and *sus*R (ANDERSON & SALYERS 1989, SHIPMAN ET AL. 2000). The *sus*-system in *B. thetaiotaomicron* represents a prototypic system and derivatives of this system (*sus*-like systems) have been identified in many other *Bacteroidetes* genomes, not only in gut symbionts, but also in terrestrial and marine strains (BAUER ET AL. 2006, MARTENS ET AL. 2009, MCBRIDE ET AL. 2009, FERNÁNDEZ-GÓMEZ ET AL. 2013) and metagenomes (WILLIAMS ET AL. 2013). Such systems are not restricted to starch, but also encode degradation of other polysaccharides (MARTENS ET AL. 2011). The so called *sus*-like systems (Figure 1.6) consist of an outer membrane associated transport protein homologue to *susC*, also known as TonB-dependent receptor (SCHAUER ET AL. 2008), and an outer membrane associated lipoprotein homologue to *susD* (REEVES ET AL. 1997, SHIPMAN ET AL. 2000, CHO & SALYERS 2001, BJURSELL ET AL. 2006, MARTENS ET AL. 2011) which is unique for *Bacteroidetes* (THOMAS ET AL. 2011). TonB-dependent receptors belong to a group of outer membrane-spanning β-barrel proteins which are energized via the TonB-ExbBD complex using energy from the proton motive force to transport macromolecules into the periplasma, where they will be further hydrolyzed and then transported into the cytoplasma (KOEBNIK 2005, SCHAUER ET AL. 2008). The *susD* monomer contains an oligosaccharide-binding pocket, which interacts with up to three individual glucose units (KOROPATKIN ET AL. 2008). The *susC* and *susD* homologs are clustering together with other genes

encoding carbohydrate degradation such as glycosyl hydrolases (GH), carbohydrate esterases (CE), carbohydrate binding modules (CBM) and polysaccharide lyases (PL) and are referred to as polysaccharide utilization loci (PULs, MARTENS ET AL. 2011).

For the group of carbohydrate active enzymes (CAZymes) there is a dedicated CAZy database (CANTAREL ET AL. 2009, LOMBARD ET AL. 2014). Sulfatases can additionally be part of PULs as it was shown by GÓMEZ-PEREIRA ET AL. (2012) and others (BAUER ET AL. 2006, THOMAS ET AL. 2011, TEELING ET AL. 2012, MANN ET AL. 2013).

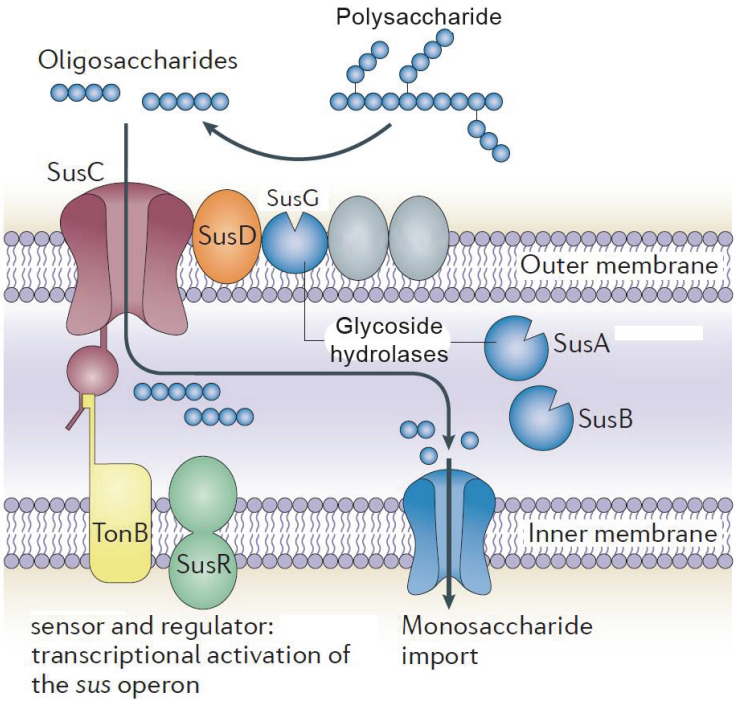


Figure 1.6: The starch utilization system (*sus*). The *susD* protein is responsible for the specific substrate binding and the glycoside hydrolases located at the outer membrane decompose the polysaccharides into oligosaccharides. The oligosaccharides are then transported into the periplasm by an ATP driven TonB-dependent transporter (*susC*) and are further hydrolyzed into monosaccharides by periplasmic glycoside hydrolases (scheme was adapted from KOROPATKIN ET AL. 2012).

1.8 Aims

This doctoral thesis contributes to the understanding of the ecological role of *Bacteroidetes* in marine environments and elaborates on their distribution and function at different marine sites. Specifically, I am interested in the response of *Bacteroidetes* to phytoplankton spring blooms, their substrate association by direct surface attachment and their genetic capability of degrading high molecular weight organic matter. Therefore, the full cycle rRNA approach was implemented to study bacteriodetal diversity and ecology independent of cultivation-based methods (AMANN ET AL. 1995), adressing the fundamental questions in ecology: „Who is out there?"; "How many of a kind are there?" and "What is their function?". Consequently, one part of this thesis concentrates on method development and their evaluation. A second part concentrates on the analysis of *Bacteroidetes* diversity and abundance during spring blooms at the coastal station Helgoland Roads. The third part aims at the understanding of functional traits of *Bacteroidetes* to gain further insights in their ecological niches.

I. <u>Method Development</u>

The well-established method CARD-FISH is frequently used to address the 'Who?' and 'How many?', but this results often in tedious manual microscopic counting, which is time consuming and quite subjective. Therefore, automated cell enumeration becomes more and more important, since it is less time consuming and less prone to possible human errors. In **chapter 2** we tested a newly developed automatic cell enumeration system extensively to apply it as a high-throughput enumeration system for chapter 3 and 4, as well as to use it in a pilot study on board a research vessel, obtaining site specific cell counts and community composition.

In a second study (**chapter 5**) we used microscopy to visualize, (a) if specific *Bacteroidetes* clades can be assigned in their association with phytoplankton to certain ecological niches and (b) if *Bacteroidetes* cells excrete *in situ* extracellular polymeric substances in form of capsules or slime as a basis for their attachment. A prerequisite of this study was to modify the CARD-FISH protocol and to combine it with fluorescent lectin binding analysis (FLBA).

II. Abundance and Diversity

This section addressed the question whether bacteria, and more specifically *Bacteroidetes*, are benefit from dying algae? Therefore, the main objective in **chapter 3** was the identification and quantification of specific *Bacteroidetes* clades responding to the spring phytoplankton bloom at Helgoland Roads in 2009. We postulated that the abundance and distribution pattern of distinct clades is highly bottom-up controlled by the availability of a range of algal substrates of different origin. Following up on this hypothesis in **chapter 4** we therefore investigated the following questions: (a) Is each spring phytoplankton bloom causing a distinct bacterial succession similar to that of 2009? (b) Can additional blooming *Bacteroidetes* clades be detected in 2010 – 2012 which had no maxima in 2009? (c) Are the maxima of specific bacteroidetal clades correlated with the bloom or senescence of specific phytoplankton groups? (d) Does the presence or absence of distinct phytoplankton groups or species impact certain *Bacteroidetes* clades? And finally, (e) how does the Bacteroidetes community evolve during the remaining seasons?

III. Function

The aim of the study presented in **chapter 6** was to gain insights into the degradation potential of *Bacteroidetes* from two contrasting oceanic provinces in the North Atlantic Ocean. This study is a follow up study of GÓMEZ-PEREIRA ET AL. (2012; Appendix B) in which it was postulated that the phytoplankton-rich Boreal Polar Region has a larger potential for polysaccharide degradation than the oligotrophic North Atlantic Subtropical Region. Specifically, (a) we investigated in total 155 fosmids from both regions in their genomic content to see whether this hypothesis is supported by the extended metagenome dataset. Furthermore, we (b) characterized several novel bacteroidetal PULs.

1.9 Methods

The present study was based on the full cycle rRNA approach used as a phylogeny-based toolbox for cultivation-independent studies of microbial diversity and ecology (AMANN ET AL. 1995). After comparative sequence analysis of full length 16S rRNA gene clone libraries, new probes were designed and evaluated (see Appendix A). Subsequently, CARD-FISH was used for the quantification of bacteroidetal clades and enumerated with the newly developed automatic counting system (based on ZEDER & PERNTHALER 2009), which is described in Chapter 2. A brief overview of methods used in this study will be given here, while detailed information is provided in the manuscript chapters.

Chapter 2: Evaluation of a high-throughput automatic cell enumeration system.

The main methods used here were DAPI staining and CARD-FISH according to THIELE ET AL. 2011, then cell enumeration was done manually and automatically, using the newly developed software ACMEtool, which is described in detail within the chapter.

Chapter 3: Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom.

The main methods used here were CARD-FISH according to THIELE ET AL. 2011 and cell enumeration was done manually and automatically, using the newly developed software ACMEtool. Additional methods used by co-authors are described in detail in the supplementary information of the manuscript.

Chapter 4: Response patterns of distinct bacteroidetal clades to spring phytoplankton blooms.

The main methods used here were CARD-FISH according to THIELE ET AL. 2011, modifications are reported in the Material & Method section of the manuscript and cell enumeration was done automatically, using the newly developed software ACMEtool. A full length 16S rRNA gene clone library was constructed in house and sequenced by ABI Sanger sequencing technique.

Chapter 5: Mapping glycoconjugate-mediated interactions of marine *Bacteroidetes* with diatoms.

Within this study a protocol for combining CARD-FISH and fluorescent lectin binding analysis (FLBA) were developed and high resolution microscopy was done using the confocal laser scanning microscope Leica TCS SP5X.

Chapter 6: Characterisation of Polysaccharide Utilization Loci (PULs) in marine *Bacteroidetes.*

Fosmids were sequenced at Genoscope National Sequencing Center, France, using the 454-sequencing plattform. The fosmid annotation was done using RAST (AZIZ ET AL. 2008). The detailed carbohydrate-active-enzyme and peptidase annotation were done using an in-house annotation pipeline according to HUANG 2013, P. 22. Comparison of PULs was done using the genome alignment tool MAUVE (DARLING ET AL. 2010).

1.10 Manuscripts

I. Evaluation of a high-throughput automatic cell enumeration system

Christin M. Bennke, Greta Reintjes, Andreas Ellrott, Jörg Wulf, Michael Zeder, Rudolf Amann & Bernhard M. Fuchs

Manuscript in preparation.

Contribution to the manuscript: C. M. B. performed all experiments and data analysis; conceived and wrote manuscript, on board section was done by G. R. and A. E.

II. Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom

Hanno Teeling, Bernhard M. Fuchs, Dörte Becher, Christine Klockow,Antje Gardebrecht, **Christin M. Bennke**, Mariette Kassabgy, Sixing Huang, Alexander J. Mann, Jost Waldmann, Marc Weber, Anna Klindworth, Andreas Otto, Jana Lange, Jörg Bernhardt, Christine Reinsch, Michael Hecker, Jörg Peplies, Frank D. Bockelmann, Ulrich Callies, Gunnar Gerdts, Antje Wichels, Karen H. Wiltshire, Frank Oliver Glöckner, Thomas Schweder & Rudolf Amann

Published 2012 in *Science*, 336: 608-611

Contribution to the manuscript: C. M. B. *Bacteroidetes* specific probe design and testing, CARD-FISH and *in situ* abundance of *Bacteroidetes*, assistance in manuscript writing.

III. Response patterns of distinct bacteroidetal clades to spring phytoplankton blooms

Christin M. Bennke, Bernhard M. Fuchs, Gunnar Gerdts, Karen H. Wiltshire & Rudolf Amann

Manuscript in preparation.

Contribution to the manuscript: C. M. B. performed all experiments and data analysis, conceived and wrote manuscript. Samples were taken by Biologische Anstalt Helgoland (BAH). Phytoplankton counts, as well as physicochemical paramters were provided by BAH.

IV. Mapping glycoconjugate-mediated interactions of marine Bacteroidetes with diatoms

Christin M. Bennke, Thomas R. Neu, Bernhard M. Fuchs & Rudolf Amann

Published 2013 in *Systematic and Applied Microbiology*, 36 (6): 417-425

Contribution to the manuscript: C. M. B. developed idea, designed research and project outline together with B. M. F. and T. R. N.; C. M. B. performed CARD-FISH and lectin staining, as well as microscopy together with T. R. N.; C. M. B. conceived and wrote manuscript.

V. Characterisation of Polysaccharide Utilization Loci (PULs) in marine Bacteroidetes

Christin M. Bennke, Sixing Huang, Margarete Schüler, Valerie Barbe, Hanno Teeling, Bernhard M. Fuchs & Rudolf Amann

Manuscript in preparation.

Contribution to the manuscript: C. M. B. performed data analysis together with S. H., conceived and wrote manuscript.

Chapter 2: Manuscript I

Evaluation of a high-throughput automatic cell enumeration system

Christin M. Bennke, Greta Reintjes, Andreas Ellrott, Jörg Wulf, Michael Zeder, Rudolf Amann & Bernhard M. Fuchs

Manuscript in preparation

Evaluation of a high-throughput automatic cell enumeration system

Christin M. Bennke[1], Greta Reintjes[1], Andreas Ellrott[1], Jörg Wulf[1], Michael Zeder[2], Rudolf Amann[1] & Bernhard M. Fuchs[1]

[1]Max-Planck-Institute for Marine Microbiology, Celsiusstr. 1, 28359 Bremen, Germany

[2]Roche Diagnostics International AG, Forrenstr. 2, 6343 Rotkreuz, Switzerland

Email addresses and telephone numbers of all authors:

C. M. Bennke	cbennke@mpi-bremen.de	+49 421 2028 545
G. Reintjes	greintje@mpi-bremen.de	+49 421 2028 943
A. Ellrott	aellrott@mpi-bremen.de	+49 421 2028 943
J. Wulf	jwulf@mpi-bremen.de	+49 421 2028 940
M. Zeder	mzeder@technobiology.ch	
R. Amann	ramann@mpi-bremen.de	+49 421 2028 930
B. M. Fuchs	bfuchs@mpi-bremen.de	+49 421 2028 935

Corresponding author:

PD Dr. Bernhard M. Fuchs

Max-Planck-Institute for Marine Microbiology

Dept. of Molecular Ecology

Celsiusstr. 1

28359 Bremen

Phone: +49 421 2028 935

Fax: +49 421 2028 790

Abstract

Over the last two decades automated cell enumeration became more and more important, since it is less time consuming, less susceptible to possible human errors, and it consists of a better documentation, images and counting reports, than manual direct microscopic cell enumeration. Within this study we tested a newly developed automatic image acquisition and subsequently cell enumeration system. Manually, direct microscopic obtained cell counts were consistently equivalent to cell counts using the automated enumeration program ACMEtool 2.0. Microscopic images were acquired fully automatically and user independent, which is considerably faster than taken microscopic images manually. Further on, the analysis time of TCC and microbial communities decreases considerable. Manual counting is notably more time and labor intensive and in addition the automated method can process a higher amount of FOVs and consequently analyzes more cells. The probability of obtaining a true picture of the community increases with the number of FOVs analyzed. The automation of these processes also eliminates human error within the counting. We were further able to test this automatic microscope and cell enumeration system on board a research vessel with varying wind and wave conditions. The on-board fully automated microscope enabled precise and site specific evaluation of the microbial community allowing for targeted sampling efforts, which is both, time and money, saving.

Introduction

The determination of microbial cell numbers is fundamental to microbial ecology (PERNTHALER & AMANN 2005, KEPNER & PRATT 1994). It starts with a staining of cells with nucleic acid specific fluorescent dyes, e.g. DAPI [4',6-diamidino-2-phenylindole] (PORTER & FEIG 1980) or acridine orange (HOBBIE ET AL. 1977). Subsequently cells are counted by either flow cytometry (ZUBKOV ET AL. 1998, HAMMES & EGLI 2010) or fluorescence microscopy (BLOEM 1995). Total cell counts (TCC) serve also as a basic parameter for calculating the relative abundances of specific microbial groups determined by fluorescence *in situ* hybridization (FISH) or catalyzed reporter deposition - fluorescence *in situ* hybridization (CARD-FISH) (AMANN ET AL. 1995, PERNTHALER ET AL. 2002).

Determination of TCC from water samples by fluorescence microscopy generally involves filtration of sample onto polycarbonate membrane filters (HOBBIE ET AL. 1977). This has two advantages. First, the defined pore size of the membrane filter, which is often 0.2 μm in diameter, allows for a size fractionation. Subsequently particles, smaller than bacteria (viruses or debris), which can disturb the fluorescence signal will be removed. Second, filtering a defined volume onto a known area (membrane filter) allows the calculation of cell abundances per volume (HOBBIE ET AL. 1977). The accuracy and precision of cell enumeration is important when comparing values, stating significant differences or extrapolation how the bacterial community evolved over a certain time frame. Direct manual microscopic counting of bacterial cells is time consuming and always relies on the accuracy of the person counting. Therefore automated cell enumeration becomes more and more important, since it is less time consuming and less prone to possible human errors.

Over the last two decades scientists have continuously developed tools or add-ons for commercially available programs to enumerate microbial cell numbers automatically (BLOEM ET AL. 1995, SINGLETON ET AL. 2001, PERNTHALER ET AL. 2003, ABRÀMOFF ET AL. 2004, SELINUMMI ET AL. 2005, THIEL & BLAUT 2005, ZHOU ET AL. 2007, DAIMS 2009). However, most of them deal only with the post image processing and the image acquisition was done manually.

The system we are presenting includes automated image acquisition using fully motorized microscopes. Although PERNTHALER ET AL. (2003) already used autonomous image acquisition the determination of the position where images are taken depended on the exact position of the filter section on the object slide and the microscope stage position. Therefore, they created a blueprint for the minimum required size of the filter section, their exact position on the object slides and the position of the microscope stage. This system was adapted, by recognizing shape and size of the filter sections independent of their position on the object slide (ZEDER ET AL. 2011) and incorporated it as a macro into the AxioVision 4.8.2 software package from Zeiss (Carl Zeiss MicroImaging GmbH, Göttingen, Germany). In a next step the number of field of views (FOV) per samples are chosen independently of their position on the sample ensuring that the image acquisition is independent of user's interference and completely random. We tested this system extensively under stable laboratory conditions and subsequently modified the procedure for use on board a research vessel. The automated image acquisition and the focusing routine were modified so that the microscope system can sustain ship movements. We tested this system on-board in an attempt to enhance the scientific capability by allowing for continuous enumeration of total cell counts and the bacterial community composition.

Materials and Methods

Sampling for initial testing experiments

From January 1st to May 31st 2011 sea water samples from 1 m below sea surface were collected twice a week with the research vessel Ade at station "Kabeltonne", Helgoland Roads, North Sea (54°11'30 N, 7°54'00 E). To the seawater samples 37% formaldehyde solution (Sigma-Aldrich, Taufkirchen, Germany) were added to a final concentration of 1%. Subsequently, the samples were fixed for 1 h at room temperature. Volumes of 10 ml were then filtered onto 0.2 μm pore-sized polycarbonate membrane filters with a diameter of 47 mm (Whatman 7060-4702). Filtration was performed by using a vacuum pump (vacuum below 200 mbar). After drying filters were stored at -20 °C until further analysis.

DAPI-staining

Filters were cut and subsequently mounted on glass slides using a glycerin-PBS mounting solution (Citifluor [Citiflour Ltd., London, UK] & VectaShield [Vector laboratories Inc., Burlingame, CA, USA]) containing the nucleic acid dye DAPI (4',6-diamidino-2-phenylindole; Sigma-Aldrich, Steinheim, Germany) at a final concentration of 1 μg ml^{-1}. This was done to avoid cell loss by washing. Then samples were analyzed using the automated microscope ZEISS Axio Imager.Z2 (Carl Zeiss MicroImaging GmbH, Göttingen, Germany) including the software package AxioVision 4.8.2.

Automated focusing and image acquisition

Images of filter sections were acquired with the fully automated working microscope AxioImager.Z2 possessing a cooled CCD camera (AxioCam MRm; Carl Zeiss MicroImaging GmbH, Göttingen, Germany) and a Colibri.2 LED light source (Carl Zeiss MicroImaging GmbH, Göttingen, Germany) with three light-emitting diodes (UV-emitting LED: 365+/-4.5 nm for DAPI, blue-emitting LED: 470+/-14 nm for the tyramide Alexa488, red-emitting LED: 590+/-17.5 nm for the tyramide

Alexa594), combined with the HE-62 multifilter module (Carl Zeiss MicroImaging GmbH, Göttingen, Germany). This high efficient multifilter consists of a triple emission filter TBP 425(+/-25) 527(+/-27) LP615 including a triple beam splitter of TFT 395/495/610.

Within the AxioVision software two macros (SamLoc and MPISYS) were installed. At first with SamLoc (ZEDER ET AL. 2011) an overview image of the objects slides containing the filter sections was recorded (Figure 1). The motorized microscope stage (Physik Instrumente (PI) GmbH & Co.KG, Karlsruhe, Germany) can hold up to eight object slides. During this procedure the object slides were scanned in bright-field illumination with the 1x objective (Carl Zeiss MicroImaging GmbH, Göttingen, Germany). Subsequently, coordinates for image acquisition were chosen and a field of view (FOV) coordinate list was generated within the sample localization software (ZEDER ET AL. 2011), which was then imported into the second macro MPISYS (ZEDER unpublished). Per sample at least 55 coordinates with an interspace of minimum 250 µm were determined, which seems to be adequate for planktonic samples. However, this needs to be newly evaluated for samples from other habitats. Within the MPISYS macro so called 'channels' were defined according to the fluorescent dye used (e.g. DAPI, Alexa488 and Alexa594).

In general, the number of channels is user-defined. For each channel the exposure time (constant or measured), focusing and number of z-stacks, compensating filter unevenness, are selected (ZEDER ET AL. 2010). Notably, for each channel only one LED source and exposure time, either constant or varying, can be selected. If user requests two different exposure times for one LED, two channels have to be defined selecting the same LED.

Focusing was done in the first acquired channel and then set as fixed focus position for the other channels. The focusing routine was adapted from ZEDER & PERNTHALER (2009). This was a nested focusing in bright-field and fluorescence illumination, as well as the acquisition of z-stacks and application of an "extended depth of field" - algorithm, which compensates for FOV unevenness (ZEDER & PERNTHALER 2009). A z-stack of seven layers per FOV and exposure time was recorded and subsequently an EDF file (extended depth of focus) was created, for this the extended depth of field algorithm (ZEDER & PERNTHALER 2009) was used for the calculation. Images were acquired using a 63x/1.4 plan oil apochromatic objective (Carl Zeiss MicroImaging GmbH, Göttingen, Germany). The EDF file was saved in 'tif' format and was loaded into the ACMEtool program (Automated Cell Measuring and Enumeration tool, www.technobiology.ch; based on ZEDER ET AL. 2010).

ACMEtool

Cell enumeration was done using the newly developed ACMEtool 2.0 (based on ZEDER ET AL. 2010). First, images taken with MPISYS were manually checked and low quality images were excluded from further analysis. Images with over- or underexposed parts, areas out of focus (unevenness) or with too many aggregates, phytoplankton cells, debris and particles are defined as low quality images and were deselected. In a second step, a so-called metafile was calculated from the remaining high quality images for a faster image processing. This metafile contained the coordinates of all recognized objects/cells of each high quality image and a dataset where parameter values, e.g. object area, circularity, mean grey value and signal-to-background ratio were described.

For each channel the aforementioned parameters for object/cell detection have to be defined above a certain threshold and applied to the set and subset definitions and subsequently manually verified. Generally, DAPI is used as the reference channel in detecting the cells. For enumerating specific bacterial clades

every hybridized signal needs to possess a FISH and a DAPI signal to count as a valid object/cell. The percentage overlap between such two objects in different channels can also be defined. Finally, after optimization of the parameters, cells were counted automatically by the program and cell numbers were exported in a tab delimited report file. Furthermore, ACMEtool provides a summary file, in which per sample the number of analyzed FOV and total counted cells were given. For a detailed overview, a FOV report is provided, in which the numbers of counted cells per FOV are reported. This report gives a detailed overview of the cell distribution on the filter.

Sampling and sample preparation on board

Planktonic sea water samples were taken during the Atlantic Meridional Transect (AMT) 22, upon the research vessel RRV James Cook (Southampton, UK, to Punta Arenas, Chile, October 10th to November 24th 2012). Samples were taken using a Sea Bird CTD (Sea Bird Electronics Inc., USA) which was deployed twice daily at predawn and solar noon intervals. A total of 50 stations were sampled at 20 m water depth. From each sample 100 ml of seawater was fixed with 1% formaldehyde for 1 h at RT. Subsequently, triplicate 20 ml subsamples were filtered using a vacuum pump (200 mbar) onto 47 mm polycarbonate filters with a 0.2 μm pore size. These filters were then stored at -20 °C until further analysis.

All stations were analyzed using CARD-FISH. The filters were processed according to THIELE ET AL. (2011). After the CARD-FISH process the filters were counter stained with DAPI and mounted using a Citiflour-VectorShield (4:1) mounting solution. Cell enumeration and community analysis were done directly on-board using a fully motorized Axioplan 2 microscope (Carl Zeiss MicroImaging GmbH, Göttingen, Germany) for automatic image acquisition.

Automated image acquisition on-board

An Axioplan 2 microscope was equipped with a 63x/1.4 oil plan apochromatic objective lens, a four slide scanning stage (Märzhäuser, Wetzlar, Germany), LED epifluorescence illumination 365 nm, 470 (465-485) nm, and 590/20 nm (KSL 70, Rapp OptoElectronic, Wedel, Germany), a multi-band optical filter with beam splitter HC365+495+610, and emission filter HC425+527+LP615 (F73-395, and F72-622, AHF Analysentechnik, Tübingen, Germany), and a cooled CCD camera (Orca C4742-95-12NR, Hamamatsu Photonics, Hamamatsu City, Japan). Acquisition of overview images where done via webcam and SamLoc software (ZEDER ET AL. 2011). The original control macro 'MPISYS' has been modified for the different hardware requirements, and were extended with new functionality, e.g., QEDF and FQEDF, which are explained in the result section. The macro was using the AxioVision 4.8.2 software package.

Results

Fixed vs. measured exposure times

In the first set of experiments we investigated whether constant or varying exposure times will result in images suitable for automated cell detection and counting. Therefore, three channels for DAPI imaging were defined, two channels acquired images using constant exposure times (DAPI-25ms and DAPI-50ms) and one with a varying exposure time (DAPI-auto). The latter was determined by the auto-exposure function for every single FOV of the AxioVision software and was in the range between 45 and 275 ms. In total 44 samples were analyzed by using the three different DAPI channels. From each sample 55 FOVs were recorded and after manual image quality control on average 30 ± 10 FOVs per sample were further processed, revealing more than 50% high quality images.

Both fixed exposure time channels resulted in similar cell numbers per FOV (t-test: $p = 0.685$) but differed greatly to the DAPI-auto channel ($p = 0.006$). The

average exposure time of the DAPI-auto channel was 74 ± 22 ms, with a minimum and maximum exposure time of 49 ms and 255 ms, respectively (Figure 2). Too high exposure times resulted in elevated background fluorescence, resulting in low signal to background ratios. Additionally, at higher exposure times two closely located cells on the filter appear merged and were consequently counted as one by the algorithm. On the other hand if the exposure time was too low (e.g. 25 ms) cell signals dropped below the threshold and subsequently they were not counted. Therefore the image acquisition for cell counting was done using a fixed exposure time of 50 ms. Although, auto-measured exposure times will result in similar numbers of cells per volume as fixed exposure times, the conditions of image acquisition differs for each FOV. The advantage of using a fixed exposure time is not only time saving, but also every single FOV from all samples will be acquired under the same conditions. The optimum acquisition method should be reevaluated for samples from different habitats.

Manually vs. automatic cell counts

Any automated cell counting routine needs to be evaluated against manual counting, therefore a subset of 22 samples was enumerated both manually and automatically. Enumeration of TCC and a specific bacteria group (*Bacteroidetes*) was done after CARD-FISH with the oligonucleotide probe CF319a (MANZ ET AL. 1996), assumed to detect most marine *Bacteroidetes* (AMANN & FUCHS 2008). Direct microscopic counts yielded in similar cell numbers obtained with the automated system. Regression analysis of the obtained manual and automatic counts revealed r^2 higher than 0.97 for all analysis and student`s two tailed test provided p-values greater than 0.05 (Figure 3). However, the coefficient of variance (cv) is higher for manual counting (19 ± 7%) than for automatic counting (12 ± 3%). For both counting methods outliers exist, but they are higher for manual (35%) than for automatic counting (24%). The automated cell counting is superior to manually enumeration, but requires manual evaluation for new types of environments.

Minimum FOV requirement for automatic enumeration

In literature the number of DAPI stained cells counted varies largely, e.g. JONES & SIMON (1975) described that 200 to 400 cells or 10 FOV counted is reliable. Others counted 50 FOVs or at least 1000 DAPI stained cells (PORTER & FEIG 1980). Recently an image analysis method was introduced, where 12 FOV per sample were analyzed (SEO ET AL. 2010). We tested, if analyzing a minimum of 5 FOV, which in general corresponds to approximately 300 DAPI stained cells for marine surface water samples, will result in a different number of cells per volume than compared with analyzing a higher number of FOV. Although, there existed no statistically significant differences ($p = 0.726$) in counting 5 (300 cells) or 25 FOVs (~1500 DAPI stained cells), for a better representation of the sample at least 10 FOVs should be analyzed, which is in good agreement with SEO ET AL. (2010). Counting less than 300 DAPI stained cells could lead to an overestimation of up to 10% in total numbers. In general, counting at least 1000 DAPI stained cells (PORTER & FEIG 1980), as a standardized method for direct fluorescence counting, should also be applied to automatic counting. Analyzing 25 FOV automatically is much faster than counting 25 FOV manually.

Modifications in focusing routine for on-board usage

As shown beforehand, automated image acquisition has been well established inside a laboratory environment. In order to use such system on board a research ship, we had to face one major problem for microscopy on sea: The ship environment is constantly in motion. Mass acceleration of the microscope components result in torsion of the instrument, which leads to unstable focus conditions. Especially automated microscope systems with their heavy motorized components, e.g., multi slide position stage, are affected.

One way of solving the problem would be to actively eliminate the ships movement at the microscope system, but as the amplitude of ship movement can reach several meters of height we thought of alternative ways. Thus, our idea for on

board automated image acquisition was driven by a simple rule: Be quick! After automated focus adjustment all multi-channel, multi-layer images should be acquired before the focus is lost.

In this respect we developed the method quick stack with extended depth of focus (QEDF). In comparison to the original stack image acquisition with extended depth of focus (EDF), where positioning and image acquisition is done subsequently for every single image of the stack (ZEDER ET AL. 2011), QEDF uses a live video stream approach. A video stream is recorded, while the microscopic stage moves from first to last stack position. The image frames of the live stream, which fit the needed stack positions best, were taken for the following EDF calculation.

Furthermore, the time consuming EDF calculations stayed on hold for post-processing, after all channels are acquired. Using QEDF it had been possible to acquire multi-channel, multi-layer images on board, but the yield for good quality images was low. Most often the automated imaging was still too slow. Focus drift was already too much at the start of image acquisition.

The average performance of the classical laboratory stacking and EDF method used on board yielded only 10% high quality images per sample, with a range from 0% to 25%. In contrast it yielded in more than 50% high quality images when applied in the laboratory at stable conditions. For statistical reasons, at least 10 FOV per sample should be analyzed. By using the classical stacking and EDF method this criteria was fulfilled on board in only 10% of the cases (n =32).

QEDF applies faster stacking of images than the classical EDF stacking. This increased the total number of FOV taken per sample. Although the average QEDF performance yielded 13% (4-26%), slightly more than EDF, the absolute number of high quality images was higher. In 59% of all cases (n = 21) more than 10 high quality images were obtained per sample by using the QEDF algorithm.

In order to increase the amount of high quality images we combined the QEDF stacking algorithm with in-depth focusing and this was called focused QEDF (FQEDF). The live stream image acquisition was done over a distance of 3 times the required stack size. The focus position of the live stream is calculated and represents the center of the required stack. Around this position image frames are taken and the FQEDF file was calculated, similar to QEDF and EDF.

The FQEDF stacking and imaging algorithm allowed for on-board automated multi-channel, multi-layer image acquisition yielding in a suitable amount of high quality images. On average the FQEDF stacking and imaging algorithm yielded 24% high quality images, ranging from 11% up to 53%. The statistical criteria for 10 or more high quality images per sample and channel were achieved in 97% of all cases (n = 58).

On-board evaluation

Water samples from 50 stations were processed with CARD-FISH using a mix of six oligonucleotide probes specific for the SAR11 clade (Table 1) and were analyzed with the on-board automatic microscope and cell enumeration system. To evaluate the automated on-board image acquisition and cell enumeration system a subset was manually counted on-board. The manually obtained TCC and SAR11 specific counts were in good agreement with the automatic counts. Regression analysis revealed r^2 higher than 0.90 for all analysis and student`s two tailed test provided p-values greater than 0.05 (DAPI: 0.679 and SAR11: 0.317), which means that there exists no differences between the on-board manual and automatic cell enumeration (Figure 4).

Discussion

Within this study we tested a newly developed automatic image acquisition and subsequently cell enumeration system under stable laboratory conditions and applied it to a field study on board a research vessel.

Manually, direct microscopic obtained cell counts were consistently equivalent to cell counts using the automated enumeration program ACMEtool 2.0. Microscopic images were acquired using a fully automated system which is user independent and considerably faster than manual image acquisition. For example, image acquisition for one sample consisting of 55 FOVs and two fluorescent channels (DAPI and FISH) will take approximately 15 min. This allows image acquisition of up to 50 samples overnight (approx. 12 h), enabling high-throughput image acquisition. The subsequent image analysis, consisting of image quality control, metafile creation and cell enumeration is sample number depending, but takes approximately 30 min per sample. Compared to manual microscopic counting, which is more time and labor intensive, the investigation time using the automated system decreased considerably and higher amounts of FOVs are processed. Consequently more cells are determined, revealing a better statistically significance and reflection of the sample. The automation of these processes also eliminates human error within the counting. Although, the program does most steps automatically, manual verification and adjustment is still required.

For using this autonomous counting system on board a research vessel the image acquisition routine had to be adjusted to sustain ship movements. The on-board fully automated microscope enabled precise and site specific evaluation of the microbial community allowing for targeted sampling efforts, which is both, time and money saving.

This system has the potential to become a powerful tool for analyzing microbial abundances and community composition shortly after sampling, similar to flow cytometry. Until now ship-based flow cytometry is the only method gaining real-time results of the microbial community. However, it provides only counts of the total microbial community or for a specific microbial group such as *Prochloroccocus* and *Synechococcus*, since it takes advantages of their natural autofluorescent pigments (JAMESON ET AL. 2010). Yet, determining specific non-fluorescent bacterial communities other techniques like FISH have to be used. Although WALLNER ET AL. (1993) optimized the identification of microorganisms by FISH in combination with flow cytometry, standard FISH has its limitations. Most often bacteria from aquatic environments are small, slow growing and often possess low numbers of ribosomes (AMANN ET AL. 1995, AMANN & FUCHS 2008). Subsequently, CARD-FISH was introduced (PERNTHALER ET AL. 2002). Even, the combination of CARD-FISH and flow cytometry has been successfully applied by GERDTS & LUEDKE (2006), it still has its limitations, e.g.: well dispersed cell suspension, the requirement of liquid samples (SHAPIRO 2000), insufficient permeabilization and it requires many centrifuge steps, where cell material can get lost (HAMMES & EGLI 2010). Moreover, with flow cytometry and CARD-FISH, only a limited amount of samples can be handled in parallel, while CARD-FISH on filters has no limitations in the amount of samples analyzed (SCHATTENHOFER ET AL. 2009, GÓMEZ-PEREIRA ET AL. 2010). Despite these obstacles, CARD-FISH were developed to improve the signal intensity of microorganisms for microscopic enumeration (PERNTHALER ET AL. 2002). Therefore the ship-based autonomous image acquisition and enumeration system represents a good alternative to flow cytometry and besides this, microscopy is way more sensitive than flow cytometry, not only in terms of optics, but also in detecting the fluorescent signals. For example, the detector of the flow cytomter works in the millisecond range, whereas the exposure time for automated microscopy image acquisition an range between milliseconds and seconds, if needed.

Applying this system to soil and sediment samples is still in development. Meanwhile, BIZIC-IONESCU and coworkers applied this autonomous cell enumeration system successfully to aggregate samples (BIZIC-IONESCU ET AL. in revision). It is an advantage that both systems, the autonomous image acquisition and cell enumeration with ACMEtool, can be used separately. Using ACMEtool alone high quality confocal laser scanning or even super resolution structured illumination micrographs can be analyzed for enumerating e.g. diatom-associated bacteria. ACMEtool requires only a grey-scale 8-bit 'tif' image and subsequently a metafile with object coordinates and other descriptors are created.

Acknowledgement

This study was supported by the Max-Planck-Society, Helmholtz Society and Natural Environment Research Council (NERC) Southampton. We thank the Captain and Crew of the 'Ade' for help with sampling at Helgoland and we would like to thank the Captain and Crew of the RRS James Cook for help with sampling in the open ocean. We greatly appreciate the substantial contribution of Dr. Martha Schattenhofer and Marcus Blohs. The German Federal Ministry of Education and Research (BMBF) supported this study by funding the Microbial Interactions in Marine Systems project (MIMAS, project 03F0480A, http://mimas-project.de).

References

1. ABRÀMOFF, M. D., MAGELHAES, P. J. & RAM, S. J. (2004): Image processing with ImageJ. *Biophotonics International* **11**, 36-42.

2. AMANN, R. & FUCHS, B. M. (2008): Single-cell identification in microbial communities by improved fluorescence *in situ* hybridization techniques. *Nature Rev. Microbiol.* **6**, 339-348.

3. AMANN, R. I., LUDWIG, W. & SCHLEIFER, K. H. (1995): Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiol. Rev.* **59**, 143-169.

4. BIŽIĆ-IONESCU, M., ZEDER, M., IONESCU, D., ORLIĆ, S., FUCHS, B.M. & GROSSART, H.-P. (2014): Comparison of bacterial communities on limnic vs. coastal marine particles reveals profound differences in colonization. *ISMJ* in review.

5. BLOEM, J. (1995): Fluorescent staining of microbes for total direct counts. *Molecular Microbial Ecology Manual* **4.1.8**, 1-12.

6. BLOEM, J., VENINGA, M. & SHEPHERD, J. (1995): Fully automatic determination of soil bacterium numbers, cell volumes, and frequencies of dividing cells by confocal laser scanning microscopy and image analysis. *Appl. Environ. Microbiol.* **61**, 926-936.

7. DAIMS, H. (2009): Use of fluorescence in situ hybridization and the *daime* image analysis program for the cultivation-independent quantification of microorganisms in environmental and medical samples. *Cold Spring Harb. Protoc.* 2009: pdb.prot5253. doi:10.1101/pdb.prot5253.

8. FUCHS, B. M., ZUBKOV, M. V., SAHM, K., BURKILL, P. H. & AMANN, R. (2000): Changes in community composition during dilution cultures of marine bacterioplankton as assessed by flow cytometric and molecular biological techniques. *Environ Microbiol.* **2**, 191-201.

9. GERDTS, G. & LUEDKE, G. (2006): FISH and chips: Marine bacterial communities analyzed by flow cytometry based on microfluidics. *J. Microbiol. Methods* **64**, 232-240.

10. GÓMEZ-PEREIRA, P. R., FUCHS, B. M., ALONSO, C., OLIVER, M. J., VAN BEUSEKOM, J. E. E. & AMANN R. (2010): Distinct flavobacterial communities in contrasting water masses of the North Atlantic Ocean. *ISME J.* **4**, 472-487.

11. HAMMES, F. & EGLI, T. (2010): Cytometric methods for measuring bacteria in water: advantages, pitfalls and applications. *Anal. Bioanal. Chem.* **397**, 1083-1095.

12. HOBBIE, J. E., DALEY, R. J. & JASPER, S. (1977): Use of nuclepore filters for counting bacteria by fluorescence microscopy. *Appl. Environ. Microbiol.* **33**, 1225-1228.

13. JAMESON, E., JOINT, I., MANN, N. H. & MÜHLING, M. (2010): Detailed analysis of the microdiversity of *Prochlorococcus* populations along a north-south Atlantic Ocean transect. *Environ. Microbiol.* **12,** 156-171.

14. JONES, J. G. & SIMON, B. M. (1975): An investigation of errors in direct counts of aquatic bacteria by epifluorescence microscopy, with reference to a new method for dyeing membrane filters. *J. Appl. Bacteriol.* **39,** 1-13.

15. KEPNER, R. L. & PRATT, J. R. (1994): Use of fluorochromes for direct enumeration of total bacteria in environmental samples: past and present. *Microbiol. Rev.* **58**, 603-615.

16. MANZ, W., AMANN, R., LUDWIG, W., VANCANNEYT, M., SCHLEIFER, K.H. (1996): Application of a suite of 16S rRNA-specific oligonucleotide probes designed to investigate bacteria of the phylum *Cytophaga-Flavobacter-Bacteroides* in the natural environment. *Microbiol.* **142**, 1097-1106.

17. MORRIS, R. M., RAPPE, M. S., CONNON, S. A., VERGIN, K. L., SIEBOLD, W. A., CARLSON, C. A. & GIOVANNONI, S. J. (2002): SAR11 clade dominates ocean surface bacterioplankton communities. *Nature* **420**, 806-810.

18. PERNTHALER, A., PERNTHALER, J., AMANN, R. (2002): Fluorescence *in situ* hybridization and catalyzed reporter deposition for the identification of marine bacteria. *Appl. Environ. Microbiol.* **68**, 3094-3101.

19. PERNTHALER, J. & AMANN, R. (2005): Fate of heterotrophic microbes in pelagic habitats: focus on populations. *Microbiol. Mol. Biol. Rev.* **69**, 440-461.

20. PERNTHALER, J., PERNTHALER, A. & AMANN, R. (2003): Automated enumeration of groups of marine picoplankton after fluorescence *in situ* hybridization. *Appl. Environ. Microbiol.* **69**, 2631-2637.

21. PORTER, F. G. & FEIG, Y. S. (1980): The use of DAPI for identifying and counting aquatic microflora. *Limnol. Oceanogr.* **25**, 943-948.

22. RAPPE, M. S., CONNON, S. A., VERGIN, K. L. & GIOVANNONI, S. J. (2002): Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. *Nature* **418**, 630-633.

23. SCHATTENHOFER, M., FUCHS, B. M., AMANN, R., ZUBKOV, M. V., TARRAN, G. A. & PERNTHALER, J. (2009): Latitudinal distribution of prokaryotic picoplankton populations in the Atlantic Ocean. *Environ. Microbiol.* **11**, 2078-2093.

24. SELINUMMI, J., SEPPÄLÄ, J., YLI-HARJA, O. & PUHAKKA, J. A. (2005): Software for quantification of labeled bacteria from digital microscope images by automated image analysis. *BioTechniques* **39**, 859-863.

25. SEO, E.-Y.,AHN, T.-S. & ZO, Y.-G. (2010): Agreement, precision, and accuracy of epifluorescence microscopy methods for enumeration of total bacterial numbers. *Appl. Environ. Microbiol.* **76**, 1981-1991.

26. SHAPIRO, H. M. (2000): Microbial analysis at the single-cell level: tasks and techniques. *J. Microbiol. Methods* **42**, 3-16.

27. SINGLETON, S., CAHILL, J. G., WATSON, G. K., ALLISON, C., CUMMINS, D., THURNHEER, T., GUGGENHEIM, B. & GMÜR, R. (2001): A fully automated microscope bacterial enumeration system for studies of oral microbial ecology. *J. Immunoassay Immunochem.* **22**, 253-274.

28. THIEL, R. & BLAUT, M. (2005): An improved method for the automated enumeration of fluorescently labelled bacteria in human faeces. *J. Microbiol. Methods* **61**, 369-379.

29. THIELE, S., FUCHS, B., AMANN, R. (2011): Identification of microorganisms using the ribosomal RNA approach and fluorescence *in situ* hybridization. In: WILDERER P. (Eds.), Treatise on Water Science, *Oxford: Academic Press*, pp. 171-189.

30. WALLNER, G., AMANN, R. & BEISKER, W. (1993): Optimizing fluorescent in situ hybridisation with rRNA-targeted oligonucleotide probes for flow cytometric identification of microorganisms. *Cytom.* **14**, 136-143.

31. ZEDER, M., & PERNTHALER, J. (2009): Multi-spot live-image autofocusing for high-throughput microscopy of fluorescently stained bacteria. *Cytom. Part A.* **75**, 781-788.

32. ZEDER, M., ELLROTT, A., AMANN, R. (2011): Automated sample area definition for high-throughput microscopy. *Cytom. Part A* **79**, 306-310.

33. ZEDER, M., KOHLER, E. & PERNTHALER, J. (2010): Automated quality assessment of autonomously acquired microscopic images of fluorescently stained bacteria. *Cytom. Part A.* **77**, 76-85.

34. ZHOU, Z., NOËLLE PONS, M., RASKIN, L. & ZILLES, J. L. (2007): Automated image analysis for quantitative fluorescence *in situ* hybridization with environmental samples. *Appl. Environ. Microbiol.* **73**, 2956-2962.

35. ZUBKOV, M. V., SLEIGH, M. A., TARRAN, G. A., BURKIL, P. H. & LEAKEY, R. J. G. (1998): Picoplanktonic community structure on an Atlantic transect from 50° N to 50° S. *Deep Sea Res. I* **45**, 1339-1355.

Table 1: Oligonucleotide probes used in this study. All SAR11 specific probes were applied together as a mix. Abbreviations: H = unlabeled helper oligonucleotide; FA = formamide concentration (v/v) in the hybridization buffer.

Probe name	Target group	Probe sequence (5'→3')	Length (nt)	FA (%)	Reference
CF319a	*Bacteroidetes*	TGGTCCGTGTCTCAGTAC	18	35	MANZ ET AL. 1996
SAR11-152R		ATTAGCACAAGTTTCCYCGTGT	22	25	RAPPÈ ET AL. 2002
SAR11-441R		TACAGTCATTTTCTTCCCCGAC	22	25	RAPPÈ ET AL. 2002
SAR11-441R(modif)		TACCGTCATTTTCTTCCCCGAC	22	25	AFTER RAPPÈ ET AL. 2002
SAR11-487(modif)	SAR11 clade	CGGACCTTCTTATTCGGG	18	25	AFTER MORRIS ET AL. 2002
SAR11-487-H3		CGGCTGCTGGCACGAAGTTAGC	22		
SAR11-542R		TCCGAACTACGCTAGGTC	18	25	RAPPÈ ET AL. 2002
SAR11-732R		GTCAGTAATGATCCAGAAAGYTG	22	25	RAPPÈ ET AL. 2002

Figure 1: Flow chart of automatic image acquisition. At first, (1.) the position of object slides have to be selected, then (2.) overview image will be acquired with the 1 x objective. In a next step (3.) samples have to be labeled and number of FOVs is defined. Then (4.) channels are defined and (5.) image acquisition can start. When acquisition is done, (6.) images are imported into the ACMEtool software. Within ACMEtool, at first image quality control is done and with the remaining high quality images the metafile is calculated, which is then analyzed for cell enumeration.

Figure 2: Graph depicts the number of cells per FOV obtained with different exposure time settings, while image acquisition. Two constant exposure times (blue: DAPI-25ms and red: DAPI-50ms) and one varying (black: DAPI-auto) were selected. Outliers are red circled.

Figure 3: Manual versus automatic counts. Cell enumeration was done after CARD-FISH with the oligonucleotide probe CF319a. A: DAPI obtained counts and B: CF319a obtained counts in cells ml^{-1}. Regression coefficient (r^2) and formula is depicted at the lower right site of each graph.

Figure 4: Manual versus automatic on board obtained counts. Cell enumeration was done after CARD-FISH with the oligonucleotide probe SAR11. A: DAPI counts per volume and B: SAR11-specific counts per volume. Regression coefficient (r^2) and formula is depicted at the lower right site of each graph.

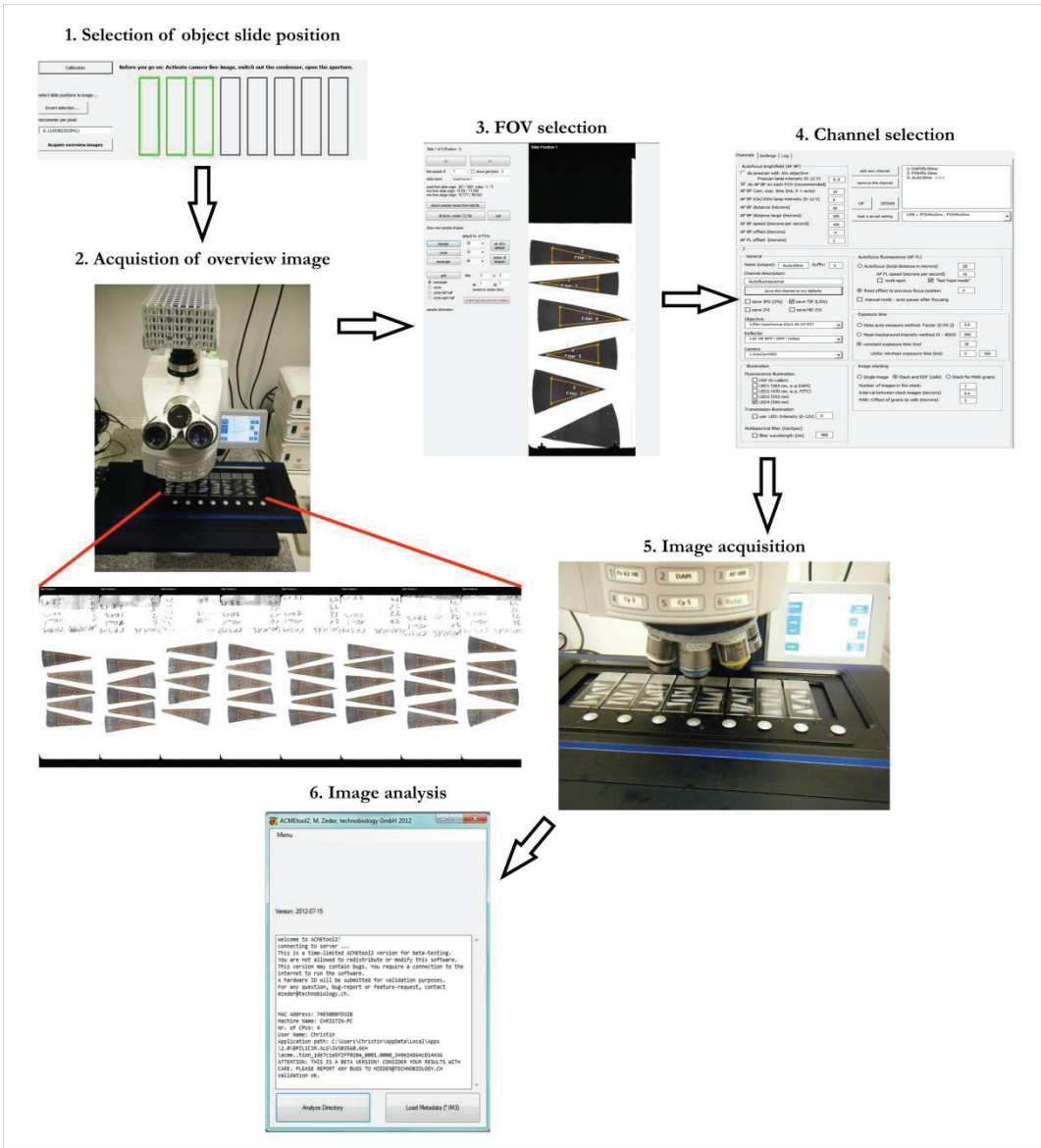

1. Selection of object slide position
2. Acquistion of overview image
3. FOV selection
4. Channel selection
5. Image acquisition
6. Image analysis
Analyze Directory
Load Metadata (*.IM3)

Figure 1

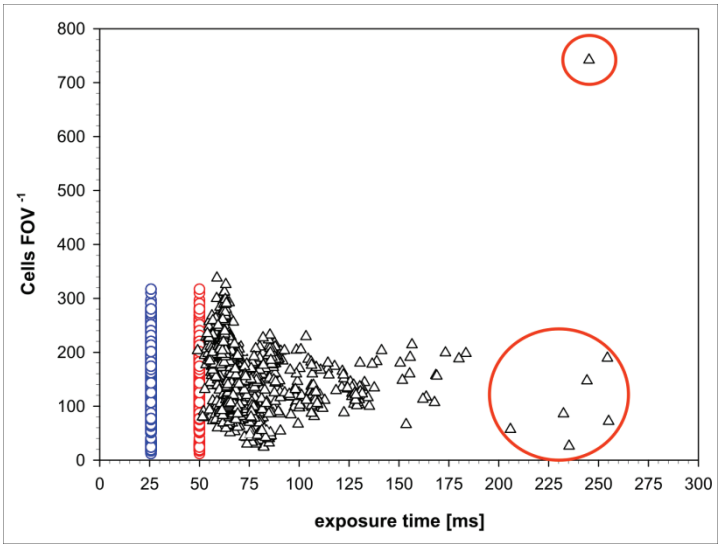


Figure 2

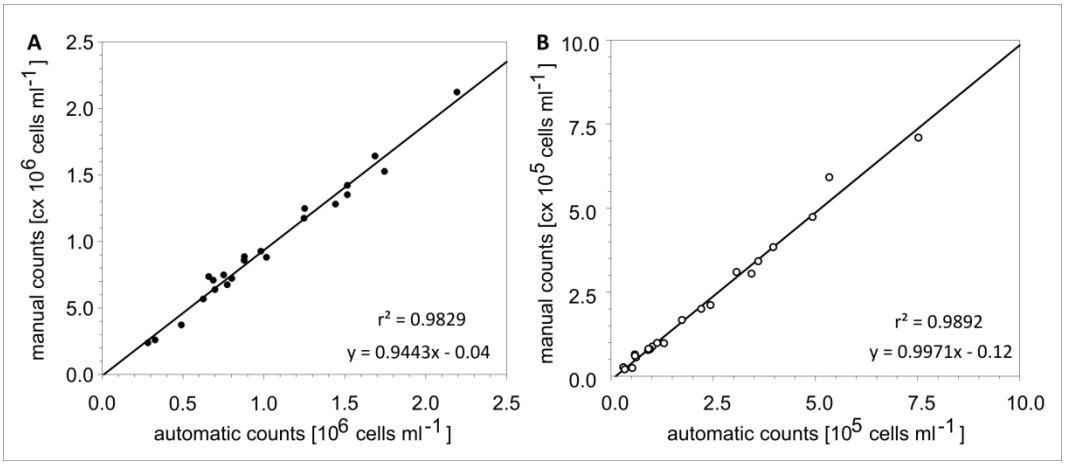


Figure 3

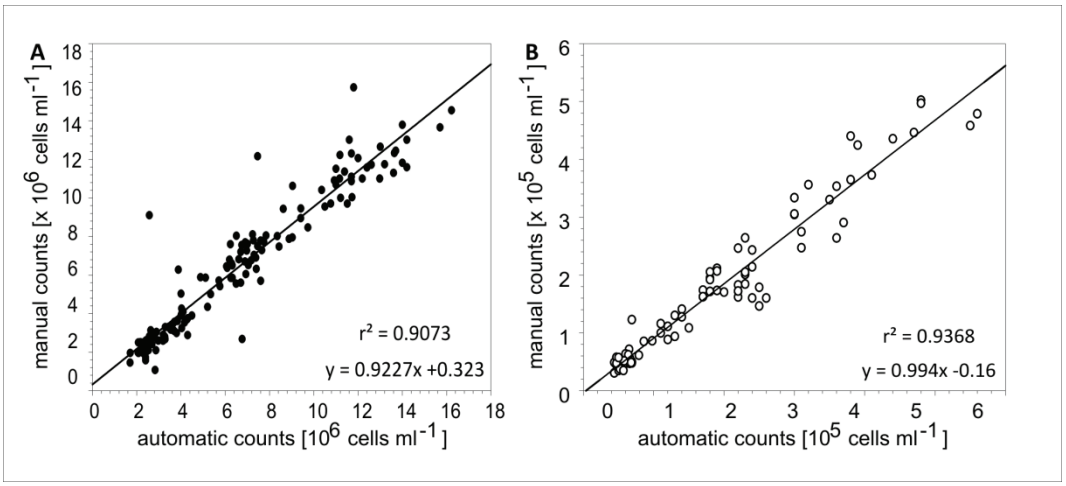


Figure 4

Supplementary material

Evaluation of a high-throughput automatic cell enumeration system

Christin M. Bennke, Greta Reintjes, Andreas Ellrott, Jörg Wulf, Michael Zeder, Rudolf Amann & Bernhard M. Fuchs

Table S1a: On-board performance test for the EDF stack image acquisition with extended depth of focus calculation. Performance test data sets are randomly selected subsets from the experiments done on-board RRS James Cook during AMT 22.

Stack image acquisition with subsequent **EDF** calculation						
	n	**Station**	**Channel**	**Number of FOV tried to image**	**Number of high quality images obtained**	**Percent high quality images obtained**
	1	Stat2-S	DAPI	55	2	*4*
	2		FISH	55	5	*9*
	3	Stat2-D	DAPI	55	0	***0***
	4		FISH	55	6	*11*
	5	Stat3-S	DAPI	55	8	*15*
	6		FISH	55	14	***25***
	7	Stat3-D	DAPI	55	11	*20*
	8		FISH	55	8	*15*
	9	Stat4-S	DAPI	55	2	*4*
	10		FISH	55	4	*7*
	11	Stat4-DCM	DAPI	55	10	*18*
	12		FISH	55	5	*9*
	13	Stat4-DCM	DAPI	91	11	*12*
	14		FISH	91	12	*13*
	15	Stat2-20m	DAPI	91	6	*7*
	16		FISH	91	12	*13*
	17	Stat2	DAPI	55	3	*5*
	18		FISH	55	3	*5*
	19	Stat3	DAPI	55	5	*9*
	20		FISH	55	4	*7*
	21	Stat3	DAPI	55	7	*13*
	22		FISH	55	2	*4*
	23	Stat5	DAPI	55	4	*7*
	24		FISH	55	2	*4*
	25	Stat2	DAPI	55	7	*13*
	26		FISH	55	3	*5*
	27	Stat3	DAPI	55	3	*5*
	28		FISH	55	8	*15*
	29	Stat3	DAPI	55	3	*5*
	30		FISH	55	3	*5*
	31	Stat5	DAPI	55	7	*13*
	32		FISH	55	3	*5*
			Average:	***60***	***6***	
			Total:	***1904***	***183***	
	Performance yield [Percent high quality images of total]:					***9.6***
	Number tests with >= 10 high quality images:					***6***
	Percent tests with >= 10 high quality images of n:					***18.8***

Table S1b: On-board performance test for the QEDF stack image acquisition with extended depth of focus calculation. Performance test data sets are randomly selected subsets from the experiments done on-board RRS James Cook during AMT 22.

	n	Station	Channel	Number of FOV tried to image	Number of high quality images obtained	Percent high quality images obtained
QEDF	1	Stat6-150m	DAPI	91	4	*4*
	2	Stat7-150m	DAPI	91	6	*7*
	3		FISH	91	5	*5*
	4	Stat8-150m	DAPI	91	7	*8*
	5		FISH	91	7	*8*
	6	Stat9-150m	DAPI	91	8	*9*
	7		FISH	91	14	*15*
	8	Stat10-20m	DAPI	91	14	*15*
	9		FISH	91	10	*11*
	10	Stat10-DCM	DAPI	91	9	*10*
	11		FISH	91	20	*22*
	12	Stat11-20m	DAPI	91	17	*19*
	13		FISH	91	21	*23*
	14	Stat11-DCM	DAPI	91	13	*14*
	15		FISH	91	10	*11*
	16	Stat11-150m	DAPI	91	16	*18*
	17		FISH	91	11	*12*
	18	Stat12-20m	DAPI	91	11	*12*
	19		FISH	91	7	*8*
	20	Stat4-150m	DAPI	91	24	*26*
	21		FISH	91	16	*18*
			Average:	***91***	***12***	
			Total:	***1911***	***250***	
	Performance yield [Percent high quality images of total]:					***13.1***
	Number tests with >= 10 high quality images:					***13***
	Percent tests with >= 10 high quality images of n:					***61.9***

Table S1c: On-board performance test for the FQEDF stack image acquisition with extended depth of focus calculation. Performance test data sets are randomly selected subsets from the experiments done on-board RRS James Cook during AMT 22.

	n	Station	Channel	Number of FOV tried to image	Number of high quality images obtained	Percent high quality images obtained
FQEDF	1	Stat2-20m	DAPI	78	10	13
	2		FISH	78	21	27
	3	Stat2-DCM	DAPI	78	31	40
	4		FISH	78	35	45
	5	Stat3-DCM	DAPI	78	32	41
	6		FISH	78	41	53
	7	Stat2-150m	DAPI	65	18	28
	8		FISH	65	20	31
	9	Stat5-20m	DAPI	23	9	39
	10		FISH	23	12	52
	11	Stat3-20m	DAPI	91	24	26
	12		FISH	91	33	36
	13	Stat3-150m	DAPI	91	30	33
	14		FISH	91	22	24
	15	Stat20-D	DAPI	91	11	12
	16		FISH	91	11	12
	17	Sta55-D	DAPI	91	18	20
	18		FISH	91	15	16
	19	Stat56-20m	DAPI	60	10	17
	20		FISH	60	18	30
	21	Stat56-DCM	DAPI	91	12	13
	22		FISH	91	16	18
	23	Stat56-D	DAPI	70	11	16
	24		FISH	70	15	21
	25	Stat57-20m	DAPI	91	10	11
	26		FISH	91	13	14
	27	Stat57-DCM	DAPI	91	20	22
	28		FISH	91	22	24
	29	Stat57-D	DAPI	59	12	20
	30		FISH	59	13	22
	31	Stat58-20m	DAPI	91	22	24
	32		FISH	91	34	37
	33	Stat58-DCM	DAPI	73	8	11
	34		FISH	73	18	25
	35	Stat58-DCM	DAPI	62	26	42
	36		FISH	62	18	29
	37	St59-20m	DAPI	91	11	12
	38		FISH	91	17	19
	39	Stat59-DCM	DAPI	91	20	22
	40		FISH	91	34	37
	41	Stat59-D	DAPI	78	16	21
	42		FISH	78	14	18

	43	Stat60-20m	DAPI	91	11	12
	44		FISH	91	19	21
	45	Stat60-DCM	DAPI	91	15	16
	46		FISH	91	24	26
	47	Stat60-D	DAPI	91	16	18
	48		FISH	91	18	20
	49	Stat4-20m	DAPI	78	13	17
	50		FISH	78	17	22
	51	Stat4-DCM	DAPI	48	18	38
	52		FISH	48	19	40
	53	Stat5-DCM	DAPI	50	13	26
	54		FISH	50	12	24
	55	Stat6-D	DAPI	91	19	21
	56		FISH	91	37	41
	57	Stat22-150m	DAPI	91	16	18
	58		FISH	91	29	32
			Average:	***78***	***19***	
			Total:	***4530***	***1099***	
	Performance yield [Percent high quality images of total]:					***24.3***
	Number tests with >= 10 high quality images:					***56***
	Percent tests with >= 10 high quality images of n:					***96.6***

Chapter 3: Manuscript II

Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom

Hanno Teeling, Bernhard M. Fuchs, Dörte Becher, Christine Klockow,Antje Gardebrecht, Christin M. Bennke, Mariette Kassabgy, Sixing Huang, Alexander J. Mann, Jost Waldmann, Marc Weber, Anna Klindworth, Andreas Otto, Jana Lange, Jörg Bernhardt, Christine Reinsch, Michael Hecker, Jörg Peplies, Frank D. Bockelmann, Ulrich Callies, Gunnar Gerdts, Antje Wichels, Karen H. Wiltshire, Frank Oliver Glöckner, Thomas Schweder & Rudolf Amann

Published 2012 in *Science*, 336: 608-611

Abstract

Phytoplankton blooms characterize temperate ocean margin zones in spring. We investigated the bacterioplankton response to a diatom bloom in the North Sea and observed a dynamic succession of populations at genus-level resolution. Taxonomically distinct expressions of carbohydrate-active enzymes (transporters; in particular, TonB-dependent transporters) and phosphate acquisition strategies were found, indicating that distinct populations of *Bacteroidetes, Gammaproteobacteria*, and *Alphaproteobacteria* are specialized for successive decomposition of algal-derived organic matter. Our results suggest that algal substrate availability provided a series of ecological niches in which specialized populations could bloom. This reveals how planktonic species, despite their seemingly homogeneous habitat, can evade extinction by direct competition.

DOI: 10.1126/science.1218344

Chapter 4: Manuscript III

Response patterns of distinct bacteroidetal clades to spring phytoplankton blooms

Christin M. Bennke, Bernhard M. Fuchs, Gunnar Gerdts, Karen H. Wiltshire & Rudolf Amann

This manuscript will be part of a larger publication.

Response patterns of distinct bacteroidetal clades to spring phytoplankton blooms

Christin M. Bennke[1], Bernhard M. Fuchs[1], Gunnar Gerdts[2], Karen H. Wiltshire[2] & Rudolf Amann[1]

[1]Max Planck Institute for Marine Microbiology, Celsiusstr. 1, 28359 Bremen, Germany

[2]Alfred Wegener Institute, Kurpromenade, Building C, 27498 Helgoland, Germany

Email addresses and telephone numbers of all authors:

C. M. Bennke	cbennke@mpi-bremen.de	+49 421 2028 545
B. M. Fuchs	bfuchs@mpi-bremen.de	+49 421 2028 935
G. Gerdts	Gunnar.Gerdts@awi.de	+49 4725 819 3245
K. H. Wiltshire	Karen.Wiltshire@awi.de	+49 4725 819 3238
R. Amann	ramann@mpi-bremen.de	+49 421 2028 930

Abstract

Phytoplankton blooms are an annual recurring phenomenon in the German Bight, North Sea. In a previous study it was shown, that members of the phylum *Bacteroidetes* responded first to the available substrates during phytoplankton senescence. Mainly three bacteroidetal clades, *Ulvibacter, Formosa* A and *Polaribacter* dominated the bloom, representing abundances as high as 20% each. With the present study we were able to show that these clades recurred in the following spring seasons 2010-2012, making up almost 50% of the *Bacteroidetes* community. Additionally, we were able to identify another blooming, yet uncharacterized *Bacteroidetes* clade, the *Owenweeksia* clade and together with *Ulvibacter* and *Polaribacter*, they represent around 60% of the *Bacteroidetes* community in spring 2010. Taken all clades together, they resolved almost the entire *Bacteroidetes* community during the spring seasons. However, the distinct succession pattern seen in spring 2009 was unrivalled among all examined spring seasons so far. It was postulated that this succession is bottom-up controlled by the availability of different algal primary products which can derive from different phytoplankton clades, triggering the blooming of specific bacteria. However, half of the *Bacteroidetes* clades showed a random correlation with specific diatom species, but were significantly positive correlated with diatoms in general. We could observe coherence between the declining abundances of diatoms and some *Bacteroidetes* clades. For example, in spring 2012 the diatom numbers were low, only a quarter of the previous year, and the clades *Ulvibacter* as well as *Formosa* B were barely detected that year, they represented at most 0.5% of the bacterioplankton. Although, most of the bacteroidetal clades were present during the entire year, their highest abundance was in spring.

Key words: *Bacteroidetes*, phytoplankton blooms, diatoms, succession, recurrence, seasonality

Introduction

Phytoplankton blooms are annually recurring phenomena in coastal zones of higher latitudes (GATTUSO ET AL. 1998). There the phytoplankton is responsible for around 19% of the net oceanic primary production albeit these zones make up less than 7% of the ocean surfaces (FIELD ET AL. 1998). About half of this net primary production is remineralized by heterotrophic bacteria in a process called the microbial loop (AZAM 1998). SMITH ET AL. (1995) showed that bacterial cell numbers and their hydrolytic enzyme activity increased significantly during phytoplankton senescence. So far, bacterial responses to phytoplankton blooms were mostly studied in mesocosm experiments. In these experiments members of the *Bacteroidetes*, *Gammaproteobacteria* and *Alphaproteobacteria* were detected (RIEMANN ET AL. 2000, SCHÄFER ET AL. 2001, PINHASSI ET AL. 2004, TADA ET AL. 2011, RINTA-KANTO ET AL. 2012).

In 2009 TEELING and coworkers observed a bacterial succession triggered by the spring phytoplankton bloom at Helgoland Roads (German Bight, North Sea; TEELING ET AL. 2012). The spring season at Helgoland Roads is characterized by increasing temperature and solar radiance (TIAN ET AL. 2011). Furthermore, it is influenced by different water masses, consisting either of marine waters from the central North Sea or of estuarine waters from river discharges (GREVE ET AL. 1996, VISSER ET AL. 1996, TIAN ET AL. 2011). Tidal driven currents as well as changing wind regimes are responsible for the steady exchange of water masses between the Wadden Sea and the North Sea (STANEVA ET AL. 2009) which determine whether at Helgoland marine or coastal characteristics prevail (GREVE ET AL. 1996, WILTSHIRE ET AL. 2010). At well-mixed coastal sites such as Helgoland Roads the spring bloom will start when the turbidity (resulting from sediment resuspension and river plumes) drops below a certain threshold and subsequently light can penetrate into deeper layers (IRIATE & PURDIE 2004, TIAN ET AL. 2011). When light penetration and availability of nutrients, especially silicate, in the mixed water layers are favorable

diatoms and other phytoplankton start to bloom (SMETACEK 1999, WILTSHIRE ET AL. 2008, LOHMANN & WILTSHIRE 2012). For the termination of phytoplankton spring blooms both bottom-up controls like nutrient and in particular silicate depletion (SMETACEK 1999, WILTSHIRE ET AL. 2008) as well as top-down controlling factors like protozoan grazing or viral lysis need to be considered (BEARE ET AL. 2002, WILTSHIRE ET AL. 2010). Dissolved and particulate organic matter released from the algae serve as substrates for bacterioplankton cells provoking a successional bloom of distinct bacterial clades (TEELING ET AL. 2012).

In spring 2009 the bacterioplankton numbers increased from around 0.5 to 3.5 x 10^6 cells ml^{-1} shortly after the chlorophyll *a* maximum was reached. *Bacteroidetes* responded first to the available substrates from dying diatoms and increased rapidly in cell numbers while SAR11 remained constant. The *Gammaproteobacteria* increased as well but with an offset of about a week to the *Bacteroidetes* and they did not reach the same high cell numbers. Whereas *Bacteroidetes* made up more than 50% of all cells, *Gammaproteobacteria* contributed a maximum of about 35% to the total community. Both groups together accounted for more than three-quarter of the bacterioplankton community in spring 2009. The major subgroups of *Bacteroidetes* were identified as *Ulvibacter, Polaribacter* and the *Formosa* clade A with peak abundances as high as 20%. Minor subgroups as the NS5-DE2 clade, the NS9 clade and a clade belonging to the class *Cytophagia* showed abundances below 5% each (TEELING ET AL. 2012). Taken together all clades represented on average more than 62% of the *Bacteroidetes* community during the spring season in 2009.

TEELING ET AL. (2012) had postulated that the appearance and succession of specific *Bacteroidetes* clades was determined by the substrates released from phytoplankton. In this study we therefore investigated the following questions: (I) Is each spring phytoplankton bloom causing a distinct bacterial succession similar to that of 2009? (II) Are the bacteroidetal clades *Polaribacter*, *Formosa* A, *Ulvibacter*, NS5-DE2, NS9 and *Cytophagia*, which had maxima during the spring diatom bloom in

2009, also recurring in the spring seasons 2010-2012? If yes, are they as abundant as in 2009? (III) Can additional blooming *Bacteroidetes* clades be detected in 2010-2012 which had no maxima in 2009? (IV) Are the maxima of specific bacteroidetal clades correlated with the bloom or senescence of specific phytoplankton groups? And finally, (V) is the presence or absence of distinct phytoplankton groups or species correlated with certain *Bacteroidetes* clades?

In a side project we also investigated whether the same or other bacteroidetal clades also appeared and even dominated during and after summer and autumn phytoplankton blooms? To this end, we additionally quantified the cell numbers of bacteroidetal clades for the entire years of 2009 and 2010.

The present study was based on the full cycle rRNA approach. Based on the comparative sequence analysis of full length 16S rRNA gene clone libraries new probes were designed and evaluated. Subsequently, CARD-FISH was used for the quantification of bacteroidetal clades during the spring seasons of 2009 to 2012 as well as for the entire years of 2009 and 2010. Cells were enumerated with the newly developed automatic counting system (based on ZEDER &PERNTHALER 2009).

Material and Methods

Sampling and sample preparation

Samples were taken 1 m below the sea surface with the research vessel Ade at station "Kabeltonne", Helgoland Roads, North Sea (54°11'30 N, 7°54'00 E) during the spring seasons (March 1st - May 31st) in 2009 to 2012. In 2009 samples were taken twice a week and from 2010 on weekdays. Additionally in 2009 and 2010 samples were harvested for the remainder of the year once every week. To the seawater samples 37% formaldehyde solution (Sigma-Aldrich, Taufkirchen, Germany) was added in a final concentration of 1%. Subsequently, the samples were fixed for 1 h at room temperature. Volumes of 10 ml (low volume filters) and 100 ml (high volume

filters) were then filtered in triplicates onto 0.2 μm pore-sized polycarbonate membrane filters with a diameter of 47 mm (Whatman 7060-4702). Filtration was performed by using a vacuum pump (pressure below 200 mbar). After drying filters were stored at -20 °C until further analysis.

Physicochemical parameters like temperature, salinity, chlorophyll *a*, silicate, phosphate, nitrate, nitrite and ammonia as well the phytoplankton abundances and species composition were determined every working day by the Biologische Anstalt Helgoland (BAH), using the methods described in WILTSHIRE ET AL. (2010). Only the seven most abundant phytoplankton species in spring were considered for the present study.

Bacterial biomass samples for diversity analysis

For bacterial biomass extraction 100 l seawater were sampled and pre-filtered through 10 μm pore-sized 142 mm polycarbonate membrane filter (type TCTP, Millipore, Eschborn, Germany). Subsequently, the collected filtrate was filtered through 3 μm pore-sized 142 mm polycarbonate membrane filter (type TSTP, Millipore, Eschborn, Germany). This was done to remove most eukaryotes before bacteria were collected on 0.2 μm pore-sized 142 mm polycarbonate membrane filters (type GPWP, Millipore, Eschborn, Germany). Filters were immediately frozen and kept at -80 °C until further processing.

DNA extraction and 16S rRNA gene libraries

Total nucleic acids were extracted from a section of the 0.2 μm pore-size biomass filter, equivalent to about 25 l surface water, according to ZHOU ET AL. (1996). The 16S rRNA genes were amplified by polymerase chain reaction (PCR) using the general bacterial primers GM3F 5'-AGA GTT TGA TCM TGG C-3' and GM4R 5'-TAC CTT GTT ACG ACT T-3' (MUYZER ET AL. 1995). The PCR reaction [0.3 mg ml-1 BSA, 250 μM total dNTPs, 0.5 μM each primer, 1 x Taq buffer and 0.2 U of Master Taq Polymerase (5 PRIME GmbH, Hamburg, Germany), 1 μg μl^{-1}

template] was carried out in ten replicates using a Mastercycler (Eppendorf) with an initial denaturation step at 95 °C for 5 min, followed by 26 cycles of 1 min denaturation at 95 °C, 1 min annealing at 44 °C, and 2 min elongation at 72 °C. There was a final extension step of 1 h at 60 °C to promote a better poly-A tailing.

The PCR products were purified using the QIAquick PCR purification kit (QIAGEN, Hilden, Germany). The clean PCR product was ligated using the TOPO TA cloning kit (Invitrogen, Karlsruhe, Germany) and cloned into high-efficiency chemically competent *E. coli* top 10 cells (Invitrogen, Karlsruhe, Germany). The library was screened for the presence of the inserts using the vector primers M13F (5'-GTA AAA CGA CGG CCA G-3') and M13R (5'-CGA GAA ACA GCT ATG AC-3'). The PCR reaction and cycling condition was similar as above, but with 35 cycles and a final extension step of 10 min at 72 °C. The annealing temperature was set to 55 °C. Afterwards the PCR products were purified using the Nucleo Fast 96 PCR kit (Macherey-Nagel, Dueren, Germany) and directly sequenced using in-house pipeline applying Sanger sequencing technology. The 16S rRNA gene sequences of this study were deposited in GeneBank under the accession numbers XXX.

Phylogenetic reconstruction

Retrieved 16S rRNA sequences were analyzed with the software package ARB (LUDWIG ET AL. 2004) using Silva database release 115, August 2013 (PRUESSE ET AL. 2007). The phylogenetic reconstruction was done according to PEPLIES ET AL., (2008) and operational taxonomic units (OTU) were calculated using ‚mothur' (SCHLOSS ET AL. 2009) based on the distance matrix calculated in ARB. Phylogenetic tree reconstruction was done in ARB using the maximum likelihood algorithm RaxML and a *Bacteroidetes* specific filter by base frequency.

Probe design

Probe design including competitors were done using the ARB probe design tool on basis of the SILVA NR 104 database release October 2010 (PRUESSE ET AL. 2007) in accordance to HUGENHOLTZ ET AL. 2001. Helpers were designed according to FUCHS ET AL. 2000. Horseradish peroxidase labeled oligonucleotide probes, unlabeled competitors and helpers were provided as lyophilized powder by Biomers (Biomers.net GmbH, Ulm, Germany). Oligonucleotides, competitors and helpers were dissolved in nuclease free water to a concentration of 8.42 pmol μl^{-1}. Probes, competitors and helpers used in this study are listed in Table S1.

Determination of total cell counts

Total cell counts were determined on "low volume filters" of a diameter of 4.7 cm on which 10 ml of surface water were directly filtered with low vacuum. Subsequently, filters were cut and subsequently mounted on glass slides using a glycerin-PBS mounting solution (Citifluor (Citifluor Ltd., London, UK) & VectaShield (Vector Laboratories Inc., Burlingame, CA, USA)) containing the nucleic acid specific dye DAPI (4',6-diamidino-2-phenylindole). This was necessary to avoid cell loss by washing. Then samples were analyzed using the automated microscope ZEISS Axio Imager.Z2 (Carl ZEISS MicroImaging GmbH) including the software package AxioVision 4.8.2 (Carl ZEISS MicroImaging GmbH) in conjunction with the macros SamLoc (ZEDER ET AL. 2011) and MPISYS (based on ZEDER & PERNTHALER 2009). For automated cell enumeration the software package ACMEtool 2.0 (based on ZEDER ET AL. 2010) was used (BENNKE ET AL. IN PREP.)

To test for even distribution of cells on filters one set of triplicates for each volume were cut in to equal sized sections and 12 of them from each filter were DAPI stained and subsequently counted. Total cell counts (TCC) from both filter sets resulted in similar abundances for each volume. The p-values of the variance analysis (ANOVA) for low ($p = 0.360$) and high volume ($p = 0.188$) filter triplicates were higher than 0.05 revealing no significant differences for the filter triplicates. The

filter sections from each filter revealed also no significant difference ($p = 536$, $n = 9$). However, the comparison of TCC revealed that on 100 ml filters the numbers determined were only about half of those on 10 ml.

Determination of absolute abundances of specific bacteroidetal clades

Absolute abundances of specific groups of bacteria were calculated from the proportion of hybridized cells and the total cell numbers (TCC) determined on 10 ml filters. The latter were determined from untreated filters which were only stained with DAPI. For highly abundant clades low volume (10 ml) filters were used for CARD-FISH, and for low abundant clades high volume filters (100 ml) were hybridized and quantified. Although, TCC retrieved from low volume filters (10 ml) were almost double as high as TCC recovered from high volume filters (100 ml), the proportion of CARD-FISH positive cells of both analysis revealed no significant difference ($p = 0.336$). Those differences in TCC might be due to overlapping cells on the filter. High volume filters were therefore used only to enumerate the frequency of low abundant clades, and their absolute numbers were calculated by multiplying this frequency with the TCC determined on 10 ml filters.

Finally, it was tested whether tides might influence the bacterioplankton numbers and community composition, since Helgoland Roads is part of the intertidal zone. Therefore, samples were taken twice a day at station Ferry Box always at high and low tide additionally to those taken every morning at station Kabeltonne. Both stations are roughly 500 m separated from each other. It turned out that the difference in cell numbers between the samples were small, and, when random sampling variability was excluded, there was no statistically significant difference ($p=0.963$). Although the numbers are not identical between low and high tide or to station Kabeltonne they showed the same trend. Only at one time point all three samples differed about 22% in absolute numbers (Figure S3). However, this case was most likely due to a change in water mass since a drop in salinity (from 31 to 29.7 PSU) was observed during that day.

CARD-FISH

Catalyzed reporter deposition – fluorescence *in situ* hybridization (CARD-FISH) was performed according to THIELE ET AL. (2011) with modifications. Hybridization was done in a modified glass humidity chamber and preheated at 50 °C prior to hybridization to equilibrate the thick glass to the temperature. The chamber was lined with a tissue paper soaked with 6 ml of a formamide-water-mix. The formamide concentration of the mix was the same as in the hybridization buffer. The filters were placed face-up into Petri-dishes and each section was carefully covered with 100 µl of the hybridization-probe-mix [0.9 M NaCl, 20 mM Tris-HCl pH 8.0, 0.02% SDS, 1% blocking reagent, 15% dextran sulfate (Sigma-Aldrich, Steinheim, Germany), 35% formamide, 0.0842 pmol μl^{-1} probe (Biomers, Ulm, Germany)]. The Petri-dishes were placed carefully in the pre-warmed glass humidity chamber and incubated in the hybridization oven at 46 °C for 2.5 hours. Washing was done in a buffer containing 20 mM Tris-HCl (pH 8.0), 5 mM EDTA (pH 8.0), 0.07 M NaCl, and 0.01% SDS for 15 min at 48 °C in a water bath and for additional 10 min in 1xPBS [0.14 M NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4] at room temperature. Tyramide signal amplification was performed at 37°C for 30 min in a glass humidity chamber containing 6 ml water soaked tissue paper. The custom made tyramide Alexa488 (PERNTHALER ET AL. 2001) was added to the amplification buffer [2 M NaCl, 15% dextran sulfate (Sigma-Aldrich, Steinheim, Germany), 0.1% blocking reagent, 1xPBS, 0.15% H_2O_2, 1 $\mu g\ \mu l^{-1}$ tyramide Alexa488]. Filters were washed in 1xPBS for 15 min and 2 min in 96% ethanol. After drying filters were subsequently mounted with Citifluor (Citifluor Ltd., London, UK) and VectaShield (Vector Laboratories Inc., Burlingame, CA, USA) containing 1µg ml^{-1} DAPI, similar as for total cell counts. For cell enumeration the automated approach was used, similar as for total cell counts.

Statistical analysis

Correlations between *Bacteroidetes* clade abundances and environmental variables, like physicochemical parameters (chlorophyll *a*, temperature, salinity, silicate, phosphate, nitrate, nitrite and ammonia) and phytoplankton abundances (classes: diatoms, dinoflagellates, coccolithophorids, slicoflagellates, flagellates, ciliates, greenalgae; species: *Mediophyxis helysia*, *Thalassiosira nordenskioeldii*, *Chaetoceros debilis* and *Chaetoceros minimus*, *Rhizosolenia styliformis*, *Chattonella* and *Phaeocystis*) were done using Spearman rank correlation test (Table S2a-c). *Bacteroidetes* and phytoplankton numbers were transformed to log-scale for better comparison. Linear regression (Table S3) were done using stepwise forward regression model by using the log transformed *Bacteroidetes* and phytoplankton abundances (*Mediophyxis helysia, Thalassiosira nordenskioeldii, Chaetoceros debilis and minimus, Rhizosolenia styliformis, Chattonella, Phaeocystis*, silicoflagellates, coccolithophorids and greenalgae). All statistical analyses were performed using the software Sigma-Plot 12 (SYSTAT, Santa Clara, CA, USA).

Results

Description of the bacterioplankton and Bacteroidetes community in spring

In late winter the total cell numbers of the bacterioplankton were between 5-7 x 10^5 cells ml^{-1}, except for 2012 when the numbers exceeded 1.2 x 10^6 cells ml^{-1} in the beginning of March. After the chlorophyll *a* maximum the bacterioplankton numbers increased between three and seven times (Figure 1A). With 0.6 x 10^6 cells ml^{-1} the average annual TCC were lowest in spring 2010 compared to the other years where the average TCC ranged between 1.2 and 1.4 x 10^6 cells ml^{-1}.

Every spring season was characterized by a strong decrease in relative abundances of SAR11, while *Bacteroidetes* and *Gammaproteobacteria* increased (Figure 1B). The *Gammaproteobacteria* either increased simultaneously with *Bacteroidetes* (2011

and 2012) or in an offset of approximately three weeks (2010). In 2009 and 2010 they contributed up to 34 ± 1%, in 2011 and 2012 only up to 15 ± 3% (Figure 1B). Since *Bacteroidetes* responded always first to the senescence of the phytoplankton bloom reaching maxima between 35% (2012) and 63% (2010) of the total community, this study was focusing on this phylum.

For spring 2009 it had been shown that the bacteroidetal clades *Ulvibacter*, *Formosa* A, *Polaribacter*, NS5-DE2, NS9 and *Cytophagia* accounted on average for 62% of the *Bacteroidetes* community, with *Ulvibacter*, *Formosa* A and *Polaribacter* being the dominant clades (TEELING ET AL. 2012). The three clades made up more than half of the *Bacteroidetes* community in spring 2009 and shortly after the chlorophyll *a* maximum, they even accounted for nearly all *Bacteroidetes*. By analysis of the 16S rRNA clone library of spring 2009 (TEELING ET AL. 2012) an additional abundant *Bacteroidetes* clade could be identified and this represented 6% of the clone library and were affiliated with the NS3 clade. All retrieved NS3 sequences were more than 99% identical in their 16S rRNA sequence, but showed less than 97% 16S rRNA identity to sequences of the remaining NS3 clade, except for clones retrieved from coastal Fjord waters at Norway (NEWBOLD ET AL. 2012). The latter were more than 99% identical to the Helgoland NS3 clones. The newly designed NS3a-840 oligonucleotide probe detects besides the Helgoland clones almost the entire clade and covers a minimum target group identity of 90%. In spring 2009 NS3 showed a pronounced peak approximately three weeks after chlorophyll *a* maximum making up 9% of the bacterioplankton community. Later NS3 decreased rapidly in numbers and by the end of spring they represented only 0.2% (3×10^3 cells ml^{-1}).

The same *Bacteroidetes* clades present in spring 2009 recurred in spring 2010, but accounted on average only for 56% of the *Bacteroidetes* community, whereof the three dominant clades *Polaribacter*, *Formosa* A and *Ulvibacter* made up already 46% (figure 2A). In contrast to 2009 where those three clades accounted on some days for the total *Bacteroidetes* community, in 2010 at most three quarters were resolved. To

further characterize the missing *Bacteroidetes* clades a 16S rRNA full length clone library was constructed from samples retrieved at April, 8th 2010 (figure 2B). Interestingly, about one-third of the clones (33%) were affiliated with a flavobacterial clade called *Owenweeksia*. This clade was named after the type strain *Owenweeksia hongkongensis* (LAU ET AL. 2005). Although, our clones shared only 88% 16S rRNA identity with the type strain we will stick with the name *Owenweeksia*. All retrieved sequences were scattered throughout the *Owenweeksia* cluster consisting of 9 OTUs (99% similarity cutoff), whereby one OTU was 99% identical to clones (1.5%) retrieved in spring 2009. The oligonucleotide probe VIS6-814 (GÓMEZ-PEREIRA ET AL. 2010) targeted all Helgoland clones, but additional helpers had to be designed to enhance the probe signal intensity. Hybridization with the probe specific for this clade revealed that in spring 2010 *Owenweeksia* was detected *in situ* with relative abundances of up to 20%, simultaneously with *Polaribacter*. Both clades together made up almost 50% of the entire bacterioplankton community and more than 80% of the *Bacteroidetes* at that time point (figure 2C). Applying the same oligonucleotide probe to the samples of spring 2009 we showed that *Owenweeksia* had also been present in spring 2009 in a relative abundance up to 11%.

High abundant clades

We found that not all clades being present in high abundances of above 10% in spring 2009 reached similar high numbers in the following spring seasons 2010-2012 (Figure 3A). *Ulvibacter*, for example, which had reached in spring 2009 relative abundances of 20% of the total community, corresponding to more than 50% of the *Bacteroidetes* community, was less frequent in the following years. In spring 2010 their maximum was only 14% and appeared together with *Polaribacter* and *Owenweeksia* (Figure 3A). That year, *Ulvibacter* reached cell numbers of 6.4×10^4 cells ml^{-1} at most, which was almost three times lower than the maximum of 1.7×10^5 cells ml^{-1} in spring 2009. In 2011 their absolute numbers were similar as in 2010 (6.5×10^4 cells ml^{-1}), yet they represented not more than 6%. In spring 2012

Ulvibacter was detected by the end of the season with maximum numbers of 1.7 x 10^4 cells ml^{-1} representing 1.6% of the total community and 4.8% of the *Bacteroidetes* (Figure 3A).

Formosa A represented the second clade in spring 2009 forming a pronounced peak of 23% after *Ulvibacter*, but in spring 2010 they contributed at most with 5% towards the end of the spring season. Their total maximum numbers (0.6 x 10^5 cells ml^{-1}) were about six times lower than in spring 2009 (3.5 x 10^5 cells ml^{-1}). In contrast in spring 2011 *Formosa* A appeared again after *Ulvibacter* reaching absolute numbers of 3.2 x 10^5 cells ml^{-1}, similar to 2009. However, they appeared simultaneously with *Polaribacter* reaching abundances higher than *Polaribacter* (17% vs. 14%). In spring 2012 *Formosa* A appeared again after *Polaribacter*, similar as in spring 2010, reaching at most 8% (1.9 x 10^5 cells ml^{-1}, Figure 3A).

As already mentioned above, *Polaribacter* were identified as the third clade in spring 2009, representing up to 27% of the total community. This clade was also present with more than 25% in spring 2010 and 2012, whereas in spring 2011 they represented only 14%. In spring 2010, they formed a pronounced peak simultaneously with *Ulvibacter* and the newly identified clade *Owenweeksia*. The relative abundances of *Polaribacter* in spring 2009 and 2010 were similar but their absolute numbers (1.3 x 10^5 cells ml^{-1}) were about a sixth of that in spring 2009 (7.8 x 10^5 cells ml^{-1}), similar to what was observed for *Formosa* A. Although in spring 2011 the relative abundances was with 14% lower as compared to the other spring seasons, the absolute numbers were with 2.4 x 10^5 cells ml^{-1} actually twice as high as in the spring maximum of 2010. In spring 2012 *Polaribacter* reached relative peak abundances of 25% and represented 4.6 x 10^5 cells ml^{-1}, which is almost twice as high as the previous spring and 3.5 times higher as in spring 2010 (Figure 3A).

The *Owenweeksia* clade seemed to follow the distribution pattern of *Polaribacter*. During the spring season of 2009 until 2011 both clades appeared simultaneously but had different abundances. Whereas in spring 2009 *Owenweeksia*

accounted on average for less than half of the *Polaribacter* abundance (0.6 x 10^5 cells ml^{-1} vs. 1.6 x 10^5 cells ml^{-1}), in spring 2010 both clades appeared simultaneously with similar abundances of 1.0-1.4 x 10^5 cells ml^{-1}. Together they represented almost 50% of the entire bacterioplankton community and more than 80% of the *Bacteroidetes* in early April 2010. That year *Owenweeksia* reached their highest relative abundances of all four spring seasons. However, their absolute numbers were highest in spring 2009 (1.9 x 10^5 cells ml^{-1}) and lowest in spring 2011 (0.7 x 10^5 cells ml^{-1}). In 2011 they also made up only 3% of the bacterioplankton community. In spring 2012 the simultaneous occurrence of *Owenweeksia* and *Polaribacter* stopped and *Owenweeksia* reached abundances of 11% between the second and third *Polaribacter* peak (Figure 2), accounting for 1.2 x 10^5 cells ml^{-1} (Figure 3A).

Low abundant clades

Clades which were present in abundances below 10% in spring 2009 remained also low in the following spring seasons 2010-2012 (Figure 3B). Some clades also revealed a multimodal distribution, like the NS9 clade. Always, by the end of the winter period NS9 showed abundances around 1%, but entering spring they decreased to less than 0.1%. With increasing chlorophyll *a* concentration NS9 increased as well and formed a pronounced peak with relative abundances of up to 3% shortly after the chlorophyll *a* maximum. By the end of the spring season NS9 formed an additional peak with relative abundances of up to 2.5%. This was consistent for the spring seasons 2009 until 2011. In contrast NS9 were detected with less than 1% during spring 2012. Also their absolute numbers were lowest in spring 2012 with about 2.3 x 10^4 cells ml^{-1} and highest in spring 2009 when they reached up to 7.7 x 10^4 cells ml^{-1} (Figure 3B).

The NS5-DE2 subclade seemed to have a similar distribution pattern as NS9. However, in transition from winter to spring NS5-DE2 represented less than 0.1% of the total community in all four spring seasons. Spring 2011 seems to resemble the distribution pattern of spring 2009 where this clade increased to its highest

abundances of 5% representing about 6 x 10^4 cells ml^{-1} shortly after the chlorophyll *a* maximum. In spring 2010 several peaks appeared with maximum abundances of 2.2% representing absolute numbers between 0.9-1.3 x 10^4 cells ml^{-1}. Here, the first increase in abundances appeared while chlorophyll *a* increased as well. In contrast, in spring 2012 NS5-DE2 showed a peak of 2.2% about a month after the chlorophyll *a* maximum (Figure 3B).

The NS3 group recurred every spring season and formed pronounced peaks (Figure 3). However, they never reached abundances as in spring 2009, when they showed absolute numbers of 2.7 x 10^5 cells ml^{-1}. For this clade it also seems that the distribution pattern of spring 2009 is resembled in spring 2011 (Figure 3B) where they made up 6% of the total bacterioplankton community representing 1.0 x 10^5 cells ml^{-1}. In spring 2011, NS3 formed a peak simultaneously with *Polaribacter* and *Formosa* A similar as in spring 2009. During the spring seasons of 2010 and 2012 NS3 were one order of magnitude lower in absolute numbers (2.3-3 x 10^4 cells ml^{-1}) and represented between 2.1% (2012) and 3.4% (2010). In spring 2012 NS3 had the lowest cell numbers of all four spring seasons, similar as *Ulvibacter* and NS9 (Figure 3).

This was also found for the second *Formosa* clade B (Figure 3B). They also had their lowest cell numbers (4.5 x 10^3 cells ml^{-1}) in spring 2012. Usually, every spring except for 2012 the *Formosa* clade B increased in their abundances within a few days after the chlorophyll *a* maximum (Figure 3). They formed a pronounced peak shortly before the *Formosa* clade A and then went below the detection limit. In spring 2009 and 2010 this clade represented 4% of the entire bacterioplankton community and in 2011 they reached only 2%. However, their absolute numbers were lower in 2010 (2.2 x 10^4 cells ml^{-1}) than in 2011 (2.8 x 10^4 cells ml^{-1}). In spring 2012 this clade was barely detected (0.2% at most), similar as *Ulvibacter,* NS3 and NS9 (Figure 3).

The *Marinoscillum* clade of the family *Flammeovirgaceae* within the class *Cytophagia* formed a pronounced peak in spring 2009, simultaneous with *Polaribacter*,

Formosa A, *Owenweeksia* and the NS3 clade, reaching a maximum of 9.8 x 10^4 cells ml^{-1} and representing 3% of the entire community (Figure 3). In spring 2010 and 2011 this clade was detected at most with 1.7 or 1.6 x 10^4 cells ml^{-1}, respectively < 1%). In spring 2012 *Marinoscillum* reached relative abundances similar to those of 2009, yet their absolute numbers were with at most 4.1 x 10^4 cells ml^{-1} lower than in 2009 (Figure 3B).

Does the phytoplankton community influence the Bacteroidetes community?

The phytoplankton bloom of the spring seasons 2009 until 2011 were highly diatom dominated with maxima of 2.3 x 10^6, 1.4 x 10^6 and 1.0 x 10^6 cells l^{-1}, respectively, whereas in spring 2012 the diatom numbers were on average only a sixth (2.5 x 10^5 cells l^{-1}) of the previous spring seasons. Not only had the blooms become weaker every year, also the diatom composition changed. In 2009 the diatom community consisted on average of 90% Centrales and 10% Pennales, whereas in spring 2010 about two-thirds Centrales and one-third Pennales were present. In spring 2011 about half of the diatom community was represented by Centrales and in 2012 only one-third belonged to Centrales. Those changes within the diatom community were most likely due to *Mediopyxis helysia*, a pennate diatom newly identified at Helgoland (KRABERG ET AL. 2012).

In addition to diatoms other phytoplankton members were present, for example the raphidophycean flagellate *Chattonella,* which was present with peak abundances between 4-8 x 10^5 cells l^{-1} in the spring seasons 2009, 2011 and 2012. This group was below the detection limit in spring 2010 and in spring 2012 they had their highest abundances (8.2 x 10^5 cells l^{-1}) where diatoms were low 2.5 x 10^5 cells l^{-1}. The genus *Phaeocystis* of the class *Coccolithophyceae* was present already in the beginning of the spring season with numbers exceeding 1 x 10^5 cells l^{-1}, except for spring 2012 when *Phaeocystis* reached those abundances only in April. Although *Phaeocystis* were numerically more abundant than diatoms, their contribution to the biomass is lower than that of diatoms. In order to check whether phytoplankton composition and

abundances were correlated with clade abundances of *Bacteroidetes* we performed stepwise forward regression excluding physicochemical parameters.

Spearman rank correlations (Table S2a-c) revealed only limited dependencies. *Ulvibacter* were significantly positive correlated with diatoms in general ($r^2 = 0.515$), and showed an even stronger correlation with Centrales ($r^2 = 0.661$). Correlating the five most abundant diatoms with this clade they showed a positive correlation with both *Chaetoceros* species *C. debilis* ($r^2 = 0.479$), *C. minimus* ($r^2 = 0.388$), *T. nordenskioeldii* ($r^2 = 0.365$) and *M. helysia* ($r^2 = 0.361$). NS5-DE2 were also positive correlated with *M. helysia* ($r^2 = 0.404$), *C. debilis* ($r^2 = 0.386$) and *C. minimus* ($r^2 = 0.352$). However, their correlation with diatoms ($r^2 = 0.353$) and subsequently Centrales ($r^2 = 0.498$) were weaker than those of *Ulvibacter*. *Cytophagia*, *Polaribacter* and NS3 were positive correlated with *Chattonella* ($r^2 = 0.425$, 0.380, 0.345, respectively). Most of the correlations were supported by linear regression analysis, which was performed by using the bacteroidetal clade as the dependent variable and the phytoplankton groups as independent variables. Half of the clades were randomly correlated with the phytoplankton groups, only the presence of *Ulvibacter*, both *Formosa* clades, NS5-DE2 and NS9 can partly be explained (Table S3).

Recurrence within the same year?

Besides the identification of the bacteroidetal clades in subsequent spring seasons we were also interested how these clade abundances developed during the year. Therefore, the entire years of 2009 and 2010 were examined with the same probe set. We were able to identify clades which were present solely in spring, others were present until summer and yet others throughout the entire year. *Formosa* B was the only clade exclusively present in spring (Figure 4A). Although, the NS3 and *Cytophagia* clade were detected throughout the entire year of 2009 and 2010, they only showed a pronounced peak in spring (Figure 4A). However, the NS3 clade reappeared with maximum abundances of 3.9×10^4 cells ml^{-1} and formed a peak of

1.4% in midsummer of 2009 (Figure 4A). For the remaining year both clades were present with less than 1.4 x 10^4 cells ml^{-1} (<0.5%).

Ulvibacter was present until midsummer with numbers of up to 4.6 x 10^4 cells ml^{-1} in both years, equivalent to at most 2.2% in summer 2009, and 3.3% in summer 2010. Similar as *Ulvibacter* the NS5-DE2 clade was present until midsummer and formed a pronounced peak in early summer 2009 with 3.5 x 10^4 cells ml^{-1} which represented 3.6% (Figure 4B). In summer 2010 NS5-DE2 were present with less than 1.4 x 10^4 cells ml^{-1}, similar as NS3 and *Cytophagia*. *Ulvibacter* and NS5-DE2 were below the detection limit for the remaining year.

In contrast the clades *Formosa* A, *Polaribacter*, *Owenweeksia* and NS9 were present throughout the year (Figure 4C). *Formosa* A reappeared in summer and autumn 2009 as well as in summer 2010. In summer 2009 this clade revealed two peaks, one in early summer with peak abundances of 2.1 x 10^5 cells ml^{-1} representing 14% and the other during midsummer (1.3 x 10^5 cells ml^{-1}). During the remaining summer *Formosa* A were present with less than 1%, but reappeared in autumn 2009 with peak abundances of 1.0 x 10^5 cells ml^{-1} representing less than 4%. In 2010 *Formosa* A formed a pronounced peak with 2.4 x 10^5 cells ml^{-1} representing 10% of the total community and almost 50% of the *Bacteroidetes* community in late summer. Then they declined rapidly in numbers and went below the detection limit until the end of December 2010 where they suddenly made up 1.5% (1.3 x 10^4 cells ml^{-1}).

Members of the *Polaribacter* clade were present throughout both years with absolute numbers greater than 1 x 10^4 cells ml^{-1} and formed several peaks during summer and autumn. In the beginning of summer 2009 *Polaribacter* formed a pronounced peak representing 15% of the total bacterioplankton community and 35% of the *Bacteroidetes*. But their absolute abundances were only one fifth of the abundance as in spring. In July 2009 this clade formed again three successive peaks with relative abundances up to 10%. Although their relative abundances were similar, their absolute numbers decreased from 2.6 to 0.8 x 10^5 cells ml^{-1}. *Polaribacter* cell

numbers recovered in the beginning of autumn where they formed a pronounced peak with abundances of 1.4 x 10^5 cells ml^{-1} representing 6% of the bacterioplankton community and 31% of the *Bacteroidetes* community. In the middle of autumn 2009 *Polaribacter* revealed another peak with 0.7 x 10^5 cells ml^{-1} representing 5% of the total bacterioplankton and 17% of *Bacteroidetes*. In summer 2010 this clade was present with peak abundances of 8% and absolute numbers up to 1.3 x 10^5 cells ml^{-1}. Then their numbers decreased and in the beginning of autumn 2010 they formed again a peak with 2.3 x 10^5 cells ml^{-1} representing 7% of the total bacterioplankton community and 26% of the *Bacteroidetes* slightly higher as in autumn 2009.

Owenweeksia also formed several peaks of 10% and 6% during summer 2009 representing between 9.4 to 5.3 x 10^4 cells ml^{-1}. Then they declined until their cell numbers recovered by the beginning of autumn where they formed again a pronounced peak of 1.9 x 10^5 cells ml^{-1} representing 10% of the total community and 22% of the *Bacteroidetes*. During summer 2010 *Owenweeksia* revealed a peak of 6% representing 7.9 x 10^4 cells ml^{-1} similar as in summer 2009. Then their cell numbers declined and remained below 2% for the remaining year.

The NS9 was the only low abundant clade which was present throughout the year forming several peaks (Figure 4D). In summer 2009 this clade showed two successive peaks with similar relative abundances of up to 3% as in spring. However, their absolute numbers were about three times lower than in spring (7.7 x 10^4 cells ml^{-1} vs. 2.8 x 10^4 cells ml^{-1}). After the second peak in summer 2009 their numbers decreased below 1%. Then NS9 recovered in mid-autumn and showed a peak with absolute abundances of 1.7 x 10^4 cells ml^{-1} representing 1.2% of the total bacterioplankton community. In summer 2010 NS9 showed peak abundances of 3.5 x 10^4 cells ml^{-1} slightly higher than in summer 2009, but their relative numbers were lower (1.8%) in summer 2010. In contrast to 2009, there was no autumn peak of NS9 in 2010.

In a 16S rRNA gene library, constructed from September 1st, 2009 bacterioplankton sample, sequences of the VIS1 subclade of NS5 was retrieved with 3% frequency. This clade made up 3% of the clone library and *in situ* they represented 1.3% of the total bacterioplankton community on the same day as the clone library. In mid-autumn they reappeared with 2.1%. This clade was below the detection limit in spring 2009 and slightly increased during summer. In contrast in 2010 NS5-VIS1 was barely detected in spring (<0.5%) then increased in abundances in the beginning of summer and formed a pronounced peak with 9.6 x 10^4 cells ml^{-1} representing 5% of the total bacterioplankton in midsummer. Then their numbers declined and in transition from summer to autumn 2010 VIS1 recovered and showed a pronounced peak with 1.5 x 10^5 cells ml^{-1} making up again 5% of the total community. During autumn VIS1 remained below 2% and increased in abundances again to 5% by the end of December 2010. Here, they accounted for almost 50% of the *Bacteroidetes* community, whereas in summer and autumn that year they represented less than 20% of the *Bacteroidetes* community.

Discussion

TEELING and coworkers had hypothesized that the succession of specific clades of *Bacteroidetes* and *Gammaproteobacteria* after the spring diatom bloom in the North Sea is determined by the massive release of algal polysaccharides (TEELING ET AL. 2012). In the present study we tested this hypothesis and followed the distribution pattern of ten dominant *Bacteroidetes* clades over two entire years and three spring seasons at Helgoland. We in particular investigated the recurrence of specific clades and whether the composition and timing of phytoplankton would trigger the rise of specific bacteroidetal clades.

All *Bacteroidetes* clades, except for the NS5-VIS1 clade, recurred every spring albeit not all of them to the extent as they did in 2009. The NS5-VIS1 clade represents most likely a winter clade, since it made up almost 50% of the *Bacteroidetes*

community in December 2010. Furthermore, they bloomed (between 5 and 7%) when diatoms were mainly absent. Interestingly, both NS5 clades, the NS5-VIS1 and NS5-DE2, seem to possess opposing seasonal patterns. This is supported by Spearman rank correlations, where NS5-DE2 is negatively correlated to NS5-VIS1 (r^2 = -0.400). Whereas DE2 was present in spring, VIS1 was barely detected. This suggests different ecological niches for the two clades. Similarly in the North Atlantic Ocean (GÓMEZ-PEREIRA ET AL. 2010, 2012) NS5-VIS1 appeared in higher abundances in the Arctic Province of the North Atlantic Ocean (GÓMEZ-PEREIRA ET AL. 2010), whereas NS5-DE2 was found in the Boreal Polar region only (GÓMEZ-PEREIRA ET AL. 2012).

It is notable that the cumulative abundances of the clades monitored agreed in general well with the counts obtained with probe CF319a (MANZ ET AL. 1996), assumed to detect most marine *Bacteroidetes* (AMANN & FUCHS 2008). Only for some time points mainly in the beginning and end of the spring seasons the fraction of uncharacterized *Bacteroidetes* was between 11 and 25% of the total bacterioplankton. In this study we could show that a major gap occurring in spring 2010 was due to the clade *Owenweeksia*. Members of this clade were also identified in surface waters of the North Atlantic Ocean and the oligonucleotide probe designed within that study (GÓMEZ-PEREIRA ET AL. 2010) also targets the Helgoland clade. Including this clade for community analysis, also in spring 2010 more than 90% of the entire *Bacteroidetes* community could be assigned to specific clades.

Whereas the clade recurrence was pronounced, the clade peak abundances and succession patterns were different from year to year. In 2009 three clades were dominating with relative abundances well above 20%. *Ulvibacter* was followed by members of the *Formosa* clade A and subsequently *Polaribacter* dominated (TEELING ET AL. 2012). All three clades were also abundant in the following three spring seasons, yet, .in 2010 *Ulvibacter* appeared simultaneously with *Polaribacter* and *Formosa* A was the last of the three clades to bloom. In spring 2011, again *Ulvibacter* and

Polaribacter appeared first and after their breakdown the latter recovered and appeared for a second time simultaneously with *Formosa* A. In spring 2012 *Ulvibacter* was barely detected and *Polaribacter* bloomed before *Formosa* A. This distribution pattern can either be due to the presence of several subclades, detected by the same probe, or from the release of different substrates over time from the senescent phytoplankton, which triggered different subclades. Alternatively, a competition for substrates between different *Bacteroidetes* clades could also play a role. For example, we were able to identify two different *Polaribacter* OTUs in spring 2009 and 2010. These OTUs shared less than 98% 16S rRNA identity to each other indicating that two different *Polaribacter* subclades were present in the two consecutive years. Since both subclades are detected by the same *Polaribacter*-specific oligonucleotide probe POL740 (MALMSTROM ET AL. 2007) subsequent peaks could be due to different *Polaribacter* populations. Based on published and yet unpublished genome data (GÓMEZ-CONSARNAU ET AL. 2007, GONZÁLEZ ET AL. 2008, XING ET AL. in prep.) it is conceivable that these populations have a quite different potential in degrading organic matter. HAHNKE & HARDER (2013) were able to cultivate 375 flavobacterial strains from different coastal sites of the German Bight, yielding in seven novel genera and 42 novel species as well 37 validated species (HAHNKE & HARDER 2013). Among those 42 novel species five were affiliated with *Polaribacter* (HAHNKE & HARDER 2013, HAHNKE ET AL. in rev.). All *Polaribacter* strains were isolated during the phytoplankton bloom in spring 2010, indicating the presence of minimum two *Polaribacter* subclades in spring 2010. Two of these *Polaribacter* strains were sequenced and they have a genome size of 3 Mbp and 3.9 Mbp, suggesting different lifestyles and most likely they are adapted towards different niches (XING ET AL. in prep).

In spring 2010 the *Owenweeksia* clade peaked simultaneously with *Polaribacter* reaching similar abundances making up almost 50% of the entire bacterioplankton community. A spearman rank correlation revealed a positive correlation ($r^2 = 0.580$) between both clades. According to the niche-exclusion theory (WHITTAKER ET AL.

1973) the two clades should have different niches. It has recently been shown that *Polaribacter* is attaching to diatoms (BENNKE ET AL. 2013). In contrast, *Owenweeksia* was mainly found attached to aggregates (BAKENHUS unpublished). The type strain *Owenweeksia hongkongensis* was isolated from Hongkong coastal sea waters (LAU ET AL. 2005), but since it shares less than 88% of 16S rRNA sequence identity with the *Owenweeksia* sequences retrieved at Helgoland little can be predicted on its metabolic function.

TEELING ET AL. (2012) postulated that the abundance and distribution pattern of distinct clades is highly bottom-up controlled by the availability of different algal primary products, which can derive from different phytoplankton. In spring 2009 the algal biomass was highly diatom dominated and different diatom species were present. Thus, we postulated that distinct diatom species trigger the blooming of specific bacteria. However, it seems that the extended dataset generated in this study for four spring seasons does not support this hypothesis. It looks as if the *Bacteroidetes* clades bloom irrespectable of the diatom composition itself. More specifically, we could find no hints of a direct relationship of a specific diatom species or genus with the blooming of a specific bacteroidetal clade, but we found strong correlations between the diatoms in general and specific bacteroidetal clades. For example *Ulvibacter* revealed only a slight positive correlation with *Thalassiosira* ($r^2 = 0.37$) but the correlation was more significant with Centrales ($r^2 = 0.66$) indicating that *Thalassiosira* alone might not be responsible for the dominance of the *Ulvibacter* clade. Furthermore, the strong correlation with diatoms might explain why the annually decreasing diatom biomass reflected the as well decreasing bacterioplankton numbers and especially abundances of *Ulvibacter* and *Formosa* B. Both clades were barely detected in spring 2012, where the diatom numbers collapsed. In other words, diatoms can be seen as "substrate bags", which are highly similar in chemical composition, regardless of their taxonomic affiliation, and thus provide similar substrates to the bacteria.

Besides diatoms, the raphidophycean flagellate *Chattonella* also produces a glycocalyx consisting of sulfated and non-sulfated complex carbohydrates functioning together with a neutral carbohydrate-protein complex consisting of 1,2-glycol groups and α-D-glucosyl and α-D-mannosyl residues (YOKOTE & HONJO 1984, ODA ET AL. 1998). This glycocalyx can be stained with the fluorescent labeled lectin ConA (Concanavalin A), which has a carbohydrate binding specificity towards α-glucose, α-mannose and α-N-acetylglucosamine (OKAMOTO ET AL. 2000). Furthermore, certain *Chattonella* species like *Chattonella marina* are well known of causing noxious red tides and affecting fish highly (OKAMOTO ET AL. 2000). Among our four year spring dataset a positive Spearman rank correlation was shown between *Chattonella* and *Polaribacter* ($r^2 = 0.380$). In spring 2010, where *Chattonella* was barely detected, Polaribacter showed its lowest absolute numbers of all four spring seasons. During the other spring seasons *Polaribacter* bloomed always shortly after *Chattonella*. This was also observed during summer and autumn 2009 as well in summer 2010.

This study revealed a high correlation of *Bacteroidetes* and diatoms throughout the year. Diatoms are also blooming in summer and autumn, though often not as pronounced as in spring. Hence, during their senescence the same substrates may become available. This would result in the occurrence of the same *Bacteroidetes* clades blooming in spring. However, we identified only some clades (*Polaribacter, Formosa* A, *Owenweeksia* and NS9) blooming in summer and autumn, but their abundances were not comparable with the spring ones. Others, like *Cytophagia* and NS3, were present but did not bloom. Most likely, grazing hindered or prevented the blooming of those *Bacteroidetes* clades, as it was shown for the Southern Ocean (THIELE ET AL. 2012).

In addition, already GERDTS ET AL. (2004) had shown a distinct seasonality of the bacterioplankton community at station Kabeltonne. With the present study we were able to provide further detail for *Bacteroidetes*. For example *Formosa* B was identified as "spring clade", similar as *Cytophagia* and NS3. Although, both were

present above 0.1% during summer and autumn, they bloomed only in spring (Figure 5C). *Ulvibacter* and NS5-DE2 were present until mid-summer and showed another peak of 2.2 x 10^4 and 3.5 x 10^4 cells ml^{-1}, respectively. However, their abundances were absolute abundances were much lower than in spring. NS5-DE2 reached only half of their spring abundances and Ulvibacter only one fifth. For identification of additional blooming, yet uncharacterized *Bacteroidetes* 16S rRNA clone libraries are required. First indications are given by the autumn 16S rRNA clone library from September 1st 2009. In this clone library sequences were retrieved which cluster with *Formosa* A and NS9, yet are not detected by established *Formosa* A and NS9 oligonucleotide probes.

Outlook

Since, the fraction of uncharacterized *Bacteroidetes* was greater in summer and autumn than in spring, further probe design is required for resolving the *Bacteroidetes* community also in the other seasons. This study provides further evidence that the occurrence and composition of the *Bacteroidetes* clades depends on substrates released by phytoplankton. Until we understand the specifics of the compounds (fractions of the organic matter) within the phytoplankton or derived from phytoplankton, we can only describe observations and speculate on dependencies. Thus is affected by changing environmental conditions and regime shifts. In addition, the grazing community develops during summer and autumn, making this system much more complex. Furthermore, analyzing only four spring seasons and two entire years might represent a too short time period to reveal real connections between certain phytoplankton members and *Bacteroidetes* or other bacterioplankton clades. Helgoland Roads, for example has a history of more than 50 years in recording physicochemical parameters and phytoplankton composition, after analyzing about 40 years, an increase in sea surface temperature of 1.67 °C was discovered (WILTSHIRE ET AL. 2008). This does not mean that 40 years have to be analyzed in terms of bacterial

community successions. But further analysis are needed, including chemical investigations, e.g. carbohydrate composition and incorporating other disciplines like metagenomics, metaproteomics and cultivation-dependent methods we might be able to detect causal relationships. With this dataset we were able to show a higher diversity and dynamic of the succeeding flavobacterial clades after the main phytoplankton bloom compared to the pre-blooming situation unveiling rapid shifts of the *Bacteroidetes* community within a few days.

Acknowledgement

This study was supported by the Max-Planck-Society and the Helmholtz Society. We thank the Captain and Crew of the 'Ade' for help with sampling. We are also grateful to Dr. Antje Wichels, Hilke Döpke, Kristine Carstens and Silvia Peters from the Biologische Anstalt Helgoland for help in sample proceeding and fruitful discussions. We greatly appreciate the substantial contribution of Jörg Wulf, Katy Hoffmann and Marcus Blohs. The German Federal Ministry of Education and Research (BMBF) supported this study by funding the Microbial Interactions in Marine Systems project (MIMAS, project 03F0480A, http://mimas-project.de).

References

1. AMANN, R. & FUCHS, B. M. (2008): Single-cell identification in microbial communities by improved fluorescence *in situ* hybridization techniques. Nature Reviews Microbiology 6, 339-348.

2. AZAM, F. (1998): Microbial control of oceanic carbon flux: The plot thickens. *Sience* **280**, 694-696.

3. BAKENHUS, I. (2014): Spezifische Assoziation mariner Bakterien mit Diatomeen in der Frühjahrsblüte der Nordsee. Master Thesis, University Oldenburg, pp. 62.

4. BEARE, D. J., BATTEN, S., EDWARDS, M. & REID, D. G. (2002): 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 Res.* **48**, 29-49.

5. BENNKE, C. M., NEU, T. R., FUCHS, B. M. & AMANN, R. (2013): Mapping glycoconjugate-mediated interactions of marine *Bacteroidetes* with diatoms. *Syst. Appl. Microbiol.* **36**, 417-425.

6. BENNKE, C. M., REINTJES, G., ELLROTT, A., WULF, J., ZEDER, M., AMANN, R. & FUCHS, B. M. (2014): Evaluation of a high-throughput automatic enumeration system. *IN PREPARATION*

7. DIPPNER, J.W. (1993): A frontal-resolving model for the German Bight. *Cont. Shelf Res.* **13**, 49-66.

8. FIELD, C. B., BEHRENFELD, M. J., RANDERSON, J. T. & FALKOWSKI, P. (1998): Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. *Sience* **281**, 237-240.

9. FUCHS, B. M., GLÖCKNER, F. O., WULF, J. & AMANN, R. (2000): Unlabeled helper oligonucleotides increase the in situ accessibility to 16S rRNA of fluorescently labeled oligonucleotide probes. *Appl. Environm. Microbiol.* **66**, 3603-3607.

10. GATTUSO, J.-P., FRANKIGNOULLE, M. & WOLLAST, R. (1998): Carbon and carbonate metabolism in coastal aquatic ecosystems. Ann. Rev, Ecol. Syst. 29, 405-435.

11. GERDTS, G., WICHELS, A., DÖPKE, H., KLINGS, K.-W., GUNKEL, W. & SCHUTT, C. (2004): 40-year long-term study of microbial parameters near Helgoland (German Bight, North Sea): Historical view and future perspectives. *Helgol. Mar. Res.* **58**, 230-242.

12. GÓMEZ-CONSARNAU, L., GONZÁLEZ, J. M., COLL-LLADÓ, M., GOURDON, P., PASCHER, T., NEUTZE, R., PEDRÓS-ALIÓ, C. & PINHASSI, J. (2007): Light stimulates growth of proteorhodopsin-containing marine *Flavobacteria*. *Nature* **11**, 210-213.

13. GÓMEZ-PEREIRA, P. R., ALONSO, C., OLIVER, M., VAN BEUSEKOM, J., FUCHS, B. M. (2010) Distinct flavobacterial communities in contrasting water masses of the North Atlantic Ocean. *ISME J.* **4**, 472-487.

14. GÓMEZ-PEREIRA, P. R., SCHÜLER, M., FUCHS, B. M., BENNKE, C., TEELING, H., WALDMANN, J., RICHTER, M., BARBE, V., BATAILLE, E., GLÖCKNER, F. O., AMANN, R. (2012) Genomic content of uncultured *Bacteroidetes* from contrasting oceanic provinces in the North Atlantic Ocean. *Environ. Microbiol.* **14**, 52-66.

15. GONZÁLEZ, J. M., FERNÁNDEZ-GÓMEZ, B., FERNÀNDEZ-GUERRA, A., GÓMEZ-CONSARNAU, L., SÁNCHEZ, O., COLL-LLADÓ, M., DEL CAMPO, J., RODRÍGUEZ-MARTÍNEZ, R., ALONSO-SÁEZ, L., LATASA, M., PAULSEN, I., NEDASHKOVSKAYA, O., LEKUNBERRI, I., PINHASSI, J. & PEDRÓS-ALIÓ, C. (2008): Genome analysis of the proteorhodopsin containing marine bacterium *Polaribacter* sp. MED152 (*Flavobacteria*). *PNAS* **105**, 8724-8729.

16. GRAHAM, M. H. (2003): Conforming multicollinearity in ecological multiple regression. *Ecology* **84**, 2809-2815.

17. GREVE, W., REINERS, F. & NAST, J. (1996): Biocoenotic changes of the zooplankton in the German Bight: the possible effects of eutrophication and climate. *ICES J. Mar. Sci.* **53**, 951-956.

18. HAHNKE, R. & HARDER, J. (2013): Phylogenetic diversity of *Flavobacteria* isolated from the North Sea on solid media. *Syst. Appl. Microbiol.* **36**, 497-504.

19. HAHNKE, R., BENNKE, C. M., FUCHS, B. M., MANN, A. J., TEELING, H., AMANN, R. & HARDER, J. (2014): Dilution cultivation of marine heterotrophic bacteria benefiting from a coastal diatom bloom. *Environ. Microbiol.* submitted

20. IRIARTE, A. & PURDIE, D. (2004): Factors controlling the timing of major spring bloom events in an UK south coast estuary. *Estuarine Coastal Shelf Sci.* **61**, 679-690.

21. HUGENHOLTZ, P., TYSON, G. W. & BLACKALL, L. L. (2001): Design and evaluation of 16S rRNA-targeted oligonucleotide probes for fluorescence in situ hybridisation. *In Methods in Molecular Microbiology*, vol. 176. Steroid Receptor Methods: Protocols and Assays (ed. B. A. Lieberman), pp. 29-42. Humana Press Inc.

22. LUDWIG, W., STRUNK, O., WESTRAM, R., RICHTER, L., MEIER, H., YADHUKUMAR, BUCHNER, A., LAI, T., STEPPI, S., JOBB, G., FORSTER, W., BRETTSKE, I., GERBER, S., GINHART, A. W., GROSS, O., GRUMANN, S., HERMANN, S., JOST, R., KONIG, A., LISS, T., LUSSMANN, R., MAY, M., NONHOFF, B., REICHEL, B., STREHLOW, R., STAMATAKIS, A., STUCKMANN, N., VILBIG, A., LENKE, M., LUDWIG, T., BODE, A. & SCHLEIFER, K.-H. (2004): ARB: a software environment for sequence data. *Nucleic Acids Res.* **32,** 1363-1371.

23. KRABERG, A. C., CARSTENS, K., PETERS, S., TILLY, K. & WILTSHIRE, K. H. (2012): The diatom *Mediopyxis helysia* Kühn, Hargreaves & Halliger 2006 at Helgoland Roads: a success story? *Helgol. Mar. Res.* **66**, 463-468.

24. LAU, K. W. K., NG, C. Y. M., REN, J., LAU, S. C. L., QIAN, P.-Y., WONG, P.-K., LAU, T. C. & WU, M. (2005): *Owenweeksia* hongkongensis gen. nov., sp. nov., a novel marine bacterium of the phylum '*Bacteroidetes*'. *Int. J. Syst. Evol. Microbiol.* **55**, 1051-1057.

25. LOHMANN, G. & WILTSHIRE, K. H. (2012): Winter atmospheric circulation signature for the timing of the spring bloom of diatoms in the North Sea. *Mar. Biol.* **159**, 2573-2581.

26. MALMSTROM, R., R., STRAZA, T., R. A., COTTRELL, M., T. & KIRCHMAN, D., L. (2007): Diversity, abundance, and biomass production of bacterial groups in the western Arctic Ocean. *Aquat. Microb. Ecol.* **47,** 45-55.

27. MANZ, W., AMANN, R., LUDWIG, W., VANCANNEYT, M. & SCHLEIFER, K.-H. (1996): Application of a suite of 16S rRNA-specific oligonucleotide probes designed to investigate bacteria of the phylum *Cytophaga-Flavobacter-Bacteroides* in the natural environment. *Microbiology* **142**, 1097-1106.

28. MUYZER, G., TESKE, A., WIRSEN, C. & JANNASCH, H. (1995): Phylogenetic relationships of *Thiomicrospira* species and their identification in deep-sea hydrothermal vent samples by denaturing gradient gel electrophoresis of 16S rDNA fragments. *Arch. Microbiol.* **164**, 165-172.

29. NEWBOLD, L. K., OLIVER, A. E., BOOTH, T., TIWARI, B., DESANTIS, T., MAGUIRE, M., ANDERSEN, G., VAN DER GAST, C. J. & WHITELEY, A. S. (2012): The response of marine picoplankton to ocean acidification. *Environ. Microbiol.* **14**, 2293-2307.

30. ODA, T., NAKAMURA, N., OKAMOTO, T., ISHIMATSU, A. & MURAMATSU, T. (1998): Lectin-induced enhancement of superoxide anion production by red tide phytoplankton. *Mar. Biol.* **131**, 383-390.

31. OKAMOTO,T., KIM, D., ODA, T., MATSUOKA, K., ISHIMATSU, A.& MURAMATSU, T. (2000): Concanavalin A-induced discharge of glycocalyx of raphidophycean flagellates, *Chattonella* marina and *Heterosigma akashiwo*. *Biosci. Biotechnol. Biochem*. 8, 1767-1770.

32. PEPLIES, J., KOTTMANN, R., LUDWIG, W. & GLÖCKNER, F. O. (2008): A standard operating procedure for phylogenetic inference (SOPPI) using (rRNA) marker genes. *Syst. Appl. Microbiol.* **31**, 251-257.

33. PERNTHALER, J., GLÖCKNER, F. O., SCHÖNHUBER, W. & AMANN, R. (2001): Fluorescence in situ hybridization with rRNA-targeted oligonucleotide probes. *In Methods in Microbiology: Marine Microbiology*, vol. 30 (ed. J. Paul), pp. 666. Academic Press Ltd.

34. PINHASSI, J., SALA, M. M., HAVSKUM, H., PETERS, F., GUADAYOL, O., MALITS, A. & MARRASE, C. (2004): Changes in bacterioplankon composition under different phytoplankton regimes. *Appl. Environ. Microbiol.* **70**, 6753-6766.

35. PRUESSE, E., QUAST, C., KNITTEL, K., FUCHS, B. M., LUDWIG, W., PEPLIES, J. & GLÖCKNER, F. O. (2007): SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.* **35**, 7188-7196.

36. RIEMANN, L., STEWARD, G. F. & AZAM, F. (2000): Dynamics of bacterial community composition and activity during during a mesocosm diatom bloom. *Appl. Environ. Microbiol.* **66**, 578-587.

37. RINTA-KANTO, J. M., SUN, S., SHARMA, S., KIENE, R. P. & MORAN, M. A. (2012): Bacterial community transcription patterns during a marine phytoplankton bloom. *Environ. Microbiol.* **14**, 228-239.

38. SCHÄFER, H., BERNARD, L., COURTIES, C., LEBARON, P., SERVAIS, P., PUKALL, R., STACKEBRANDT, E., TROUSSELLIER, M., GUINDULAIN, T., VIVES-REGO, J. & MUYZER, G. (2001): Microbial community dynamics in Mediterranean nutrient-enriched seawater mesocosms: changes in the genetic diversity of bacterial populations. *FEMS Microbiol. Ecol.* **34**, 243-253.

39. SCHLOSS, P. D., WESTCOTT, S. L., RYABIN, T., HARTMANN, M., HOLLISTER, E. B., LESNIEWSKI, R. A., OAKLEY, B. B., PARKS, D. H., ROBINSON, C. J., SAHL, J. W., STRES, B., THALLINGER, G. G., VAN HORN, D. J. & WEBER, C. F. (2009): Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Appl. Environ. Microbiol.* **75**, 7537-7541.

40. SMETACEK, V. (1999): Diatoms and the ocean carbon cycle. *Protist* **150**, 25-32.

41. SMITH, D. C., STEWARD, G. F., LONG, R. A. & AZAM, F. (1995): Bacterial mediation of carbon fluxes during a diatom bloom in a mesocosm. *Deep-Sea Res. Pt II* **42**, 75-97.

42. STANEVA, J., STANEV, E., WOLFF, J.-O., BADEWIEN, T., REUTER, R., FLEMMING, B., BARTHOLOMÄ, A. & BOLDING, K. (2009): Hydrodynamics and sediment dynamics in the German Bight. A focus on observations and numerical modelling in the east Frisian Wadden Sea. *Cont. Shelf Res.* **29**, 302-319.

43. TADA, Y., TANIGUCHI, A., NAGAO, I., MIKI, T., UEMATSU, M., TSUDA, A. & HAMASAKI, K. (2011): Differing growth responses of major phylogenetic groups of marine bacteria to natural phytoplankton blooms in the western North Pacific Ocean. *Appl. Environ. Microbiol.* **77**, 4055-4065.

44. TEELING, H., FUCHS, B. M., BECHER, D., KLOCKOW, C., GARDEBRECHT, A., BENNKE, C. M., KASSABGY, M., HUANG, S. X., MANN, A. J., WALDMANN, J., WEBER, M., KLINDWORTH, A., OTTO, A., LANGE, J., BERNHARDT, J., REINSCH, C., HECKER, M., PEPLIES, J., BOCKELMANN, F. D., CALLIES, U., GERDTS, G., WICHELS, A., WILTSHIRE, K.H., GLOCKNER, F. O., SCHWEDER, T., AMANN, R. (2012) Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* **336**, 608-611.

45. THIELE, S., FUCHS, B., AMANN, R. (2011): Identification of microorganisms using the ribosomal RNA approach and fluorescence *in situ* hybridization. In: WILDERER P. (Eds.), Treatise on Water Science, Oxford: Academic Press., pp. 171-189.

46. THIELE, S., FUCHS, B. M., RAMAIAH, N. & AMANN, R. (2012): Microbial community response during the iron fertilization experiment LOHAFEX. *Appl. Environ. Microbiol.* **78**, 8803-8812.

47. TIAN, T., SU, J., FLÖSER, G., WILTSHIRE, K. & WIRTZ, K. (2011): Factors controlling the onset of spring blooms in the German Bight 2002-2005: Light, wind and stratification. *Cont. Shelf Res.* **31**, 1140-1148.

48. VISSER, M., BATTEN, S., BECKER, G., BOT, P., COLIJN, F. & DAMM, P. (1996): Time series analysis of monthly mean data of temperature, salinity, nutrients, suspended matter, phyto- and zooplankton at eight locations on the Northwest European shelf. *Ocean Dyn.* **48**, 299-323.

49. WILTSHIRE, K.H., MALZAHN, A.M., WIRTZ, K., GREVE, W., JANISCH, S., MANGELSDORF, P., MANLY, B.F.J. & BOERSMA, M. (2008): Resilience of North Sea phytoplankton spring bloom dynamics: An analysis of long-term data at Helgoland Roads. *Limnol. Oceanogr.* **53**, 1294-1302.

50. WILTSHIRE, K.H., KRABERG, A., BARTSCH, I., BOERSMA, M., FANKE, H.-D., FREUND, J., GEBÜHR, C., GERDTS, G., STOCKMANN, K. & WICHELS A. (2010): Helgoland Roads, North Sea: 45 Years of Change. *Estuaries and Coasts* **33**, 295-310.

51. WHITTAKER, R. H., LEVIN, S. A. & ROOT, R. B. (1973): Niche, habitat, and ecotype. *Am. Nat.* **107**, 321-338.

52. XING, P., HAHNKE, R. L., UNFRIED, F., MARKERT, S., HUANG, S., MANN, A. J., WERNER, J., FUCHS, B. M., HARDER, J., SCHWEDER, T., AMANN, R. I., GLÖCKNER, F. O. & TEELING, H. (2014): Niche separation of two Polaribacter strains isolated from the German Bight of the North Sea during a spring diatom bloom. In preparation form *Environm. Microbiol.*

53. YOKOTE, M. & HONJO, T. (1985): Morphological and histochemicai demonstration of a glycocalyx on the cell surface of Chattonella antiqua, a 'naked flagellate'. *Experientia* **41**, 1143-1145.

54. ZEDER, M., & PERNTHALER, J. (2009): Multi-spot live-image autofocusing for high-throughput microscopy of fluorescently stained bacteria. *Cytom. Part A.* **75**, 781-788.

55. ZEDER, M., KOHLER, E. & PERNTHALER, J. (2010): Automated quality assessment of autonomously acquired microscopic images of fluorescently stained bacteria. *Cytom. Part A.* **77**, 76-85.

56. ZEDER, M., ELLROTT, A., AMANN, R. (2011): Automated sample area definition for high-throughput microscopy. *Cytom. Part A* **79**, 306-310.

57. ZHOU, J., BRUNS, M. A. &. TIEDJE, J. M. (1996): DNA recovery from soils of diverse composition. *Appl. Environ. Microbiol.* **62**, 316-322.

Figure 1: Cell abundances of the four spring seasons 2009 to 2012.

The upper graphs (A) depict the absolute abundances of the total bacterioplankton (TCC) and the lower (B) the three major bacterioplankton groups: *Bacteroidetes* (CF319a, black line), *Gammaproteobacteria* (Gam42a, black dashed line) and SAR11, as proxy for *Alphaproteobacteria,* (grey dashed line). The represented time frame (Julian day 60 to 150) reflects the entire spring season according to the meteorological definition of spring. Julian days were selected for a better comparison between the years, independently of intercalary years. The grey shaded area in each plot represents the chlorophyll *a* concentration, as a proxy for phytoplankton biomass increase. Colored bars below the graphs represent the distribution pattern of the most dominant diatoms: *Mediopyxis helysia* (green), *Thalassiosira nordenskioeldii* (orange), *Chaetoceros debilis* (dark blue), *Rhizosolenia styliformis* (yellow) and *Chaetoceros minimus* (light blue) as well as *Chattonella* (*Raphidophyceae*, purple). Color gradients within the bars reflect the abundance pattern of the according phytoplankton members, as brighter the color, as lower the abundance and as darker the color as higher the abundance.

Figure 2: Cumulative *Bacteroidetes* abundances of the spring season 2009 and 2010.

This figure is separated into three parts. Graph A represents the cumulative abundances of *Bacteroidetes* clades (bars) present in spring 2009 (adapted from TEELING ET AL. 2012). The black line represents the overall abundance of *Bacteroidetes* (CF319a). The space between the bars and the line represents the uncharacterized *Bacteroidetes* community. Graph B represents the cumulative abundances of the *Bacteroidetes* clades in spring 2010, detected with the same probe set used in 2009. However, only half of the *Bacteroidetes* community was identified and from the sample indicated with the yellow star 16S rRNA clone library was constructed revealing additional *Bacteroidetes* clades. The pie chart reflects the *Bacteroidetes* fraction (80%) of the clone library. Graph C shows the cumulative abundances of all analyzed *Bacteroidetes* clades in spring 2010. For most time points the fraction of uncharacterized *Bacteroidetes* were resolved.

Figure 3: Relative abundances of the high (A) and low (B) abundant *Bacteroidetes* clades during the spring seasons 2009 to 2012.

This figure is structured similarly as Figure 1. The upper graphs (A) represent the relative abundances of the four high abundant *Bacteroidetes* clades: *Ulvibacter* (light blue), *Formosa* A (green), *Polaribacter* (orange) and *Owenweeksia* (purple). The lower graphs (B) represent the relative abundances of the six low abundant *Bacteroidetes* clades: *Formosa* B (black), NS3 (red), NS5/DE2 clade (dark blue), NS5/VIS1 clade (yellow), NS9 (light purple) and *Cytophagia* (brown). Chlorophyll *a* is represented as the grey shaded area in each plot and the bars below the graphs represent the dominant phytoplankton members, similar as in Figure 1. The triangle on top of the 2009 and 2010 plots reflect the day from the 16S rRNA clone library construction. Attention the left y-axis (abundance [%]) has changed between the high and low abundant clades.

Figure 4: Recurrence of *Bacteroidetes* clades within the year 2009 and 2010.

This figure is structured similarly as Figure 1 and 2. Depicted are the relative abundances of the ten *Bacteroidetes* clades during the entire year of 2009 and 2010. The grey shaded area within each plot represents the chlorophyll *a* concentration and the gradient bars below the graphs reflect the abundance and distribution pattern of selected diatoms and other phytoplankton members. The plots represent the seasonal distribution pattern of (A) *Formosa* B (black), NS3 (red) and *Cytophagia* (brown); (B) *Ulvibacter* (light blue) and NS5/DE2 (dark blue); (C) Formosa A (green), *Polaribacter* (orange) and *Owenweeksia* (purple); (D) NS9 (light purple) and NS5/VIS1 (yellow). Note different scaling for relative abundances. Triangles on top of the plots reflect the days from the 16S rRNA clone library construction.

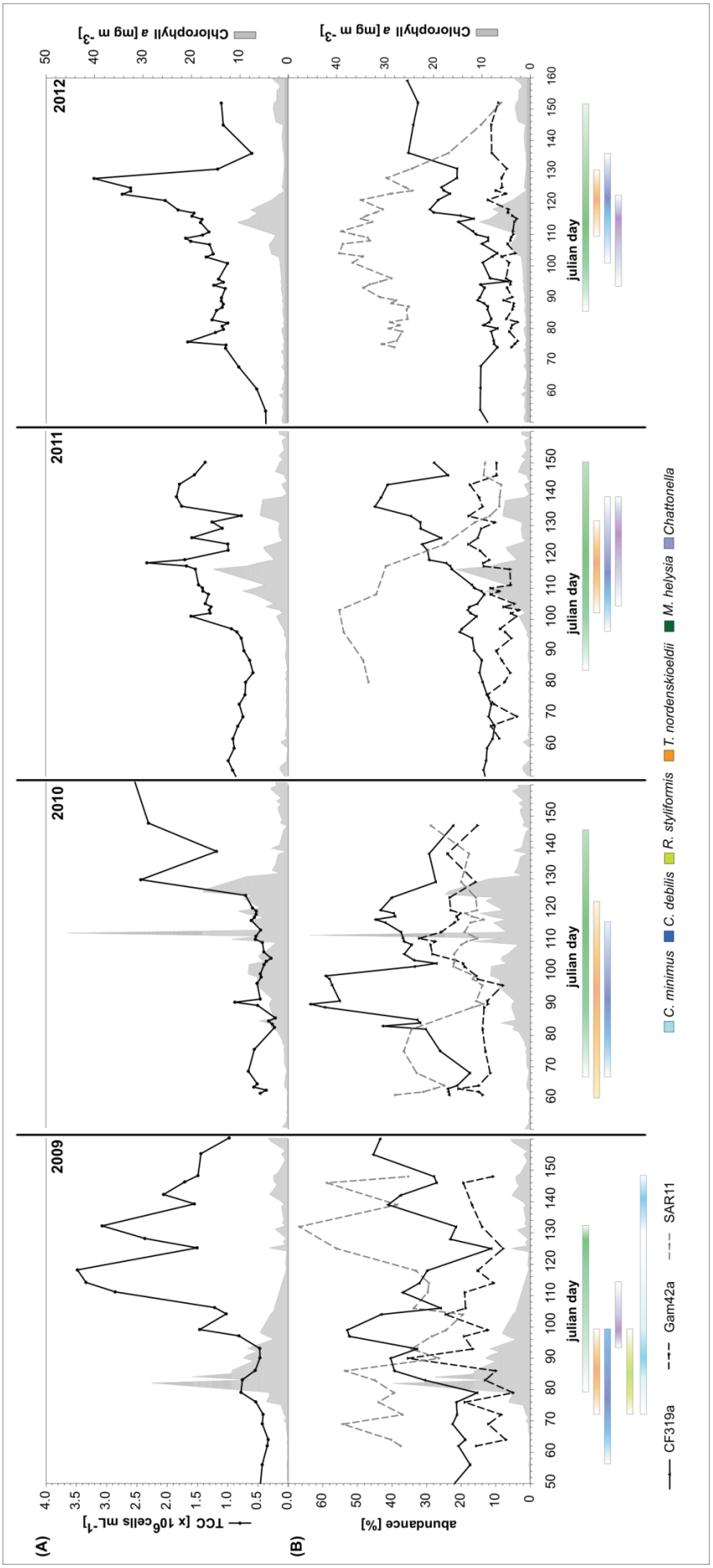

(A)
(B)
2009
2010
2011
2012
TCC [x 10^6cells mL^-1]
Chlorophyll a [mg m^-3]
abundance [%]
julian day
CF319a
Gam42a
SAR11
C. minimus
C. debilis
R. styliformis
T. nordenskioeldii
M. helysia
Chattonella

Figure 1

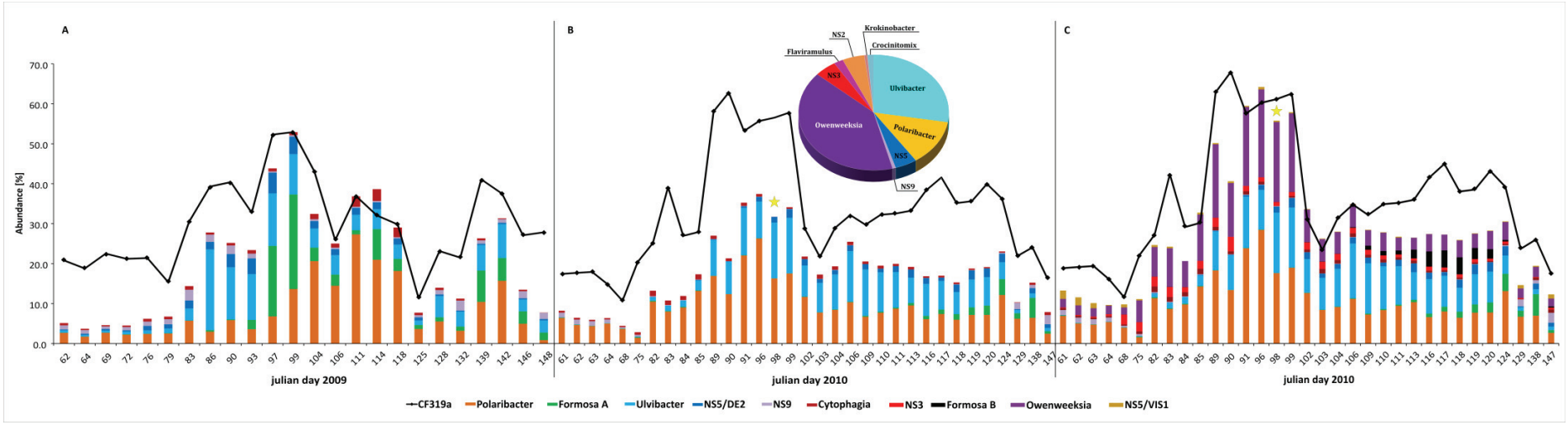

A
B
C
Abundance [%]
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
julian day 2009
julian day 2010
julian day 2010
Krokinobacter
NS2
Crocinitomix
Flaviramulus
NS3
Ulvibacter
Owenweeksia
Polaribacter
NS5
NS9
CF319a
Polaribacter
Formosa A
Ulvibacter
NS5/DE2
NS9
Cytophagia
NS3
Formosa B
Owenweeksia
NS5/VIS1

Figure 2

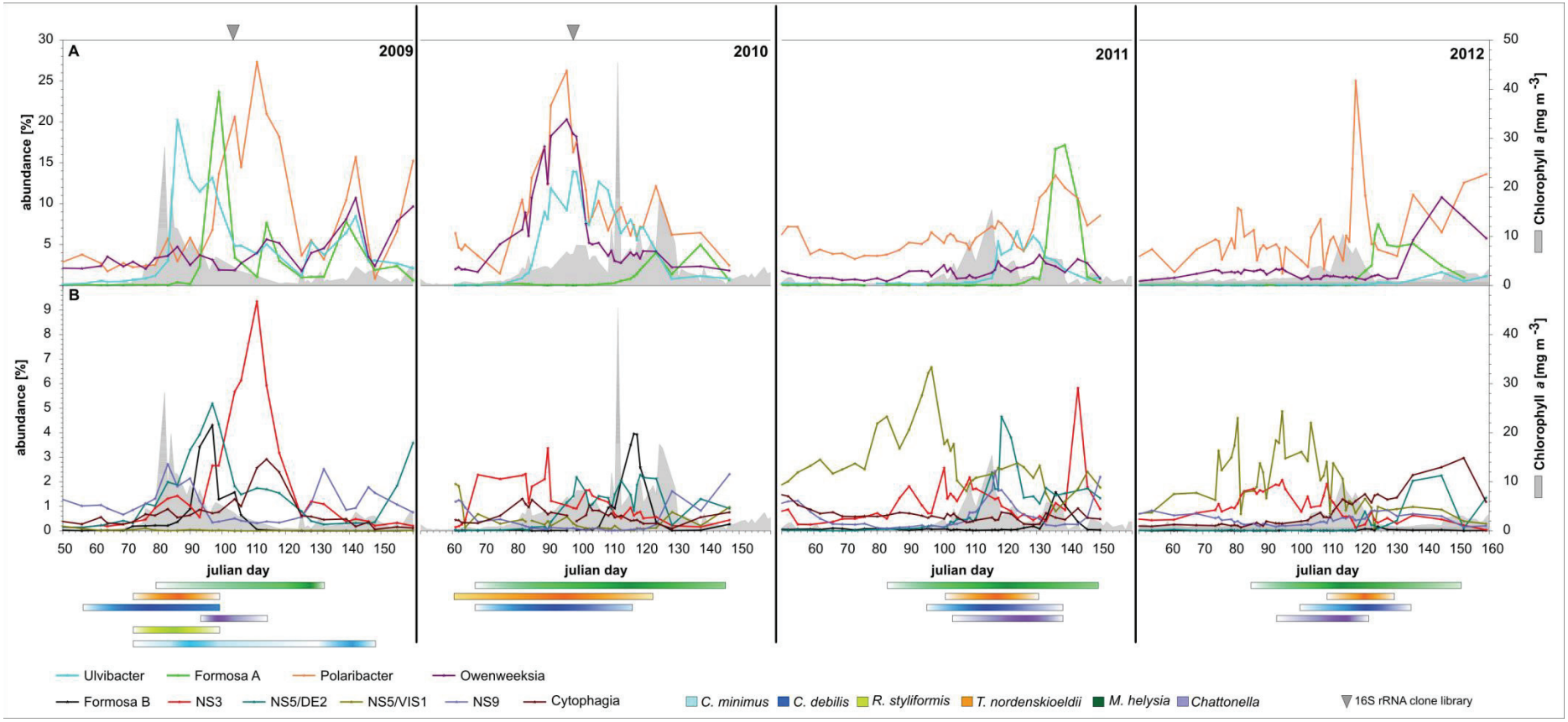

A
2009
2010
2011
2012
abundance [%]
Chlorophyll a [mg m -3]
B
julian day
Ulvibacter
Formosa A
Polaribacter
Owenweeksia
Formosa B
NS3
NS5/DE2
NS5/VIS1
NS9
Cytophagia
C. minimus
C. debilis
R. styliformis
T. nordenskioeldii
M. helysia
Chattonella
16S rRNA clone library

Figure 3

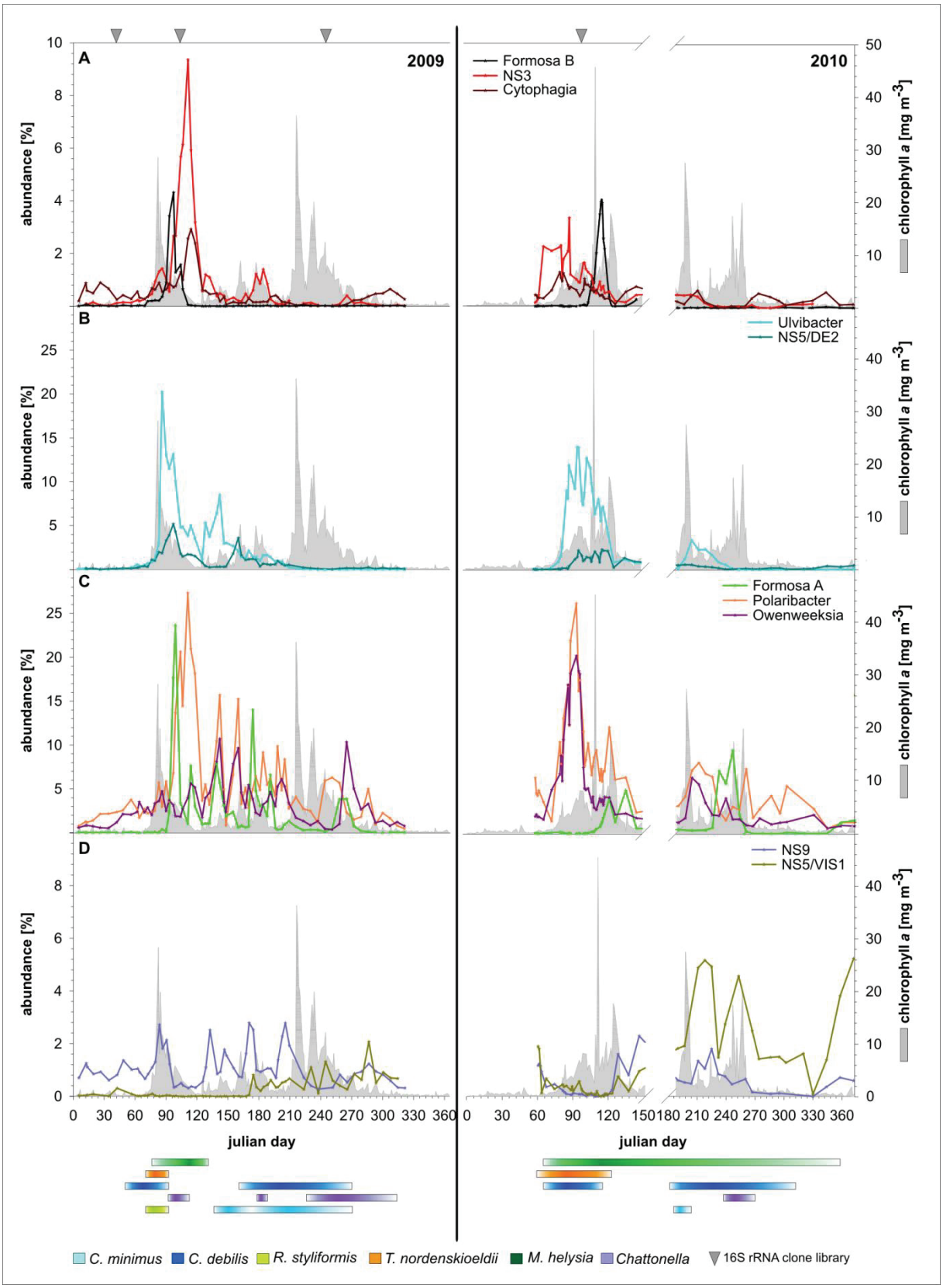

A
2009
2010
Formosa B
NS3
Cytophagia
B
Ulvibacter
NS5/DE2
C
Formosa A
Polaribacter
Owenweeksia
D
NS9
NS5/VIS1
abundance [%]
chlorophyll a [mg m-3]
julian day
C. minimus
C. debilis
R. styliformis
T. nordenskioeldii
M. helysia
Chattonella
16S rRNA clone library

Figure 4

Supplementary material

Response patterns of distinct bacteroidetal clades to spring phytoplankton blooms

Christin M. Bennke, Bernhard M. Fuchs, Gunnar Gerdts, Karen H. Wiltshire & Rudolf Amann

Table S1: *Bacteroidetes* specific oligonucleotide probes, competitors and helpers used in this study. Abbreviations: C = unlabeled competitor oligonucleotide, H = unlabeled helper oligonucleotide; FA = formamide concentration (v/v) in the hybridization buffer. Letters in red show mismatch.

Probe name	Target group	Probe sequence (5'→3')	Length (nt)	FA (%)	Reference
CF319a	*Bacteroidetes*	TGGTCCGTGTCTCAGTAC	18	35	MANZ ET AL. 1996
POL740	*Polaribacter* clade	CCCTCAGCGTCAGTACATACGT	22	35	MALMSTROM ET AL. 2007
FORM181A	*Formosa* clade A	GATGCCACTCTAAGAGAC	18	25	TEELING ET AL. 2012
FORM181A_C		GATGCCACTCTTAGAGAC	18		TEELING ET AL. 2012
FORM181B	*Formosa* clade B	GATGCCACTCTTAGAGAC	18	35	TEELING ET AL. 2012
FORM181B_C		GATGCCACTCTAAGAGAC	18		TEELING ET AL. 2012
ULV995		TCCACGCCTGTCAGACTACA	20	35	TEELING ET AL. 2012
ULV995_C1	*Ulvibacter* clade I	TCCACTCCTGTCAGACTACA	20		TEELING ET AL. 2012
ULV995_C2		TCCACCCCTGTCAGACTACA	20		TEELING ET AL. 2012
VIS6-814		CAGCGAGTGATGATCGTT	18	15	GÓMEZ-PEREIRA ET AL. 2010
VIS6-814_C1	*Owenweeksia* clade	CAGCGAGTGATCATCGTT	18		GÓMEZ-PEREIRA ET AL. 2010
VIS6-814_H1		TACGGCGTGGACTACCAGGGT	21		This study
VIS6-814_H2		CCGCYGACAGTATATCGCCAA	21		This study
NS3a-840		CTTAGCCGCTCAGAACTCAAGG	22	35	This study
NS3a-840_C1		CTTGGCCGCCCAGAACTCAAGG	22		This study
NS3a-840_C2	NS3a clade	CTTGGCCGCCCAGCACTCAAGG	22		This study
NS3a-840_H1		TYCCGAACAGCTAGTATCCATCGTT	25		This study
NS3a-840_H2		CCAGGTGGGATACTTATCACTTTCG	25		This study
NS5/DE2-471		GTAAGTAGGTTTCTTCCTGTAT	22	25	GÓMEZ-PEREIRA ET AL. 2012
NS5/DE2-471_C1	NS5-DE2 clade	GTAAGTAGGTTTCTTCCTGTAGAAAA	26		GÓMEZ-PEREIRA ET AL. 2012
NS5/DE2-471_C2		GTAAGTAGGTTTCTTCCTATAT	22		GÓMEZ-PEREIRA ET AL. 2012
NS5/VIS1-575	NS5-VIS1 clade	CTTAACAAACAGCCTGCGGACC	22	35	GÓMEZ-PEREIRA ET AL. 2010
NS5/VIS1-575_C1		CTTAAAAAACAGCCTGCGGACC	22		GÓMEZ-PEREIRA ET AL. 2010
NS9-664	NS9 clade	ACATGACCTATTCCGCCAACTT	22	35	GÓMEZ-PEREIRA ET AL. 2012
CYT-734	*Marinoscillum* clade	CAGTTTCTGCCTAGTAAG	18	25	GÓMEZ-PEREIRA ET AL. 2012

Table S2a: Spearman rank correlations of specific *Bacteroidetes* clades to phytoplankton groups. Correlation values considered only when p-value below 0.05.

	Diatoms	Pennales	Centrales	Greenalgae	Dinoflagellates	Silicoflagellates	Coccolithophorids	Flagellates	Ciliates
CF319a			0.266		0.456		0.268	0.315	0.195
p-value:			0.002		<0.001		0.002	<0.001	0.024
Polaribacter		-0.203			0.406			0.229	0.173
p-value:		0.019			<0.001			0.008	0.046
Formosa A		-0.314			0.249	-0.266			
p-value:		<0.001			0.005	0.003			
Formosa B	0.228		0.217	0.198	0.270			0.236	
p-value:	0.010		0.014	0.025	0.002			0.007	
Ulvibacter	0.515	0.202	0.661		0.288	-0.291	0.240	0.296	
p-value:	<0.001	0.021	<0.001		<0.001	<0.001	0.060	<0.001	
Owenweeksia	0.212		0.400		0.177				0.178
p-value:	0.014		<0.001		0.042				0.040
NS3a				-0.179	0.443	0.289		0.244	0.221
p-value:				0.042	<0.001	<0.001		0.005	0.012
NS5-DE2	0.353		0.498		0.446	-0.185	0.269	0.344	
p-value:	<0.001		<0.001		<0.001	0.035	0.002	<0.001	
NS5-VIS1	-0.458	-0.202	-0.540			0.463			
p-value:	<0.001	0.021	<0.001			<0.001			
NS9	-0.261	-0.419		-0.392	0.347		0.301	0.233	
p-value:	0.002	<0.001		<0.001	<0.001		<0.001	0.007	
Cytophagia		-0.188		-0.256	0.475			0.366	
p-value:		0.034		0.004	<0.001			<0.001	

Table 2b: Spearman rank correlations of specific *Bacteroidetes* clades to specific phytoplankton groups. Correlation values considered only when p-value below 0.05.

	Mediopyxis helysia	*Chaetoceros debilis*	*Chaetoceros minimus*	*Rhizosolenia styliformis*	*Thalassiosira nordenskioeldii*	*Phaeocystis*	*Chattonella*
CF319a		0.264	0.246			0.341	0.350
p-value:		0.002	0.004			<0.001	<0.001
Polaribacter						0.201	0.380
p-value:						0.021	<0.001
Formosa A			0.264		-0.226		
p-value:			0.003		0.010		
Formosa B	0.302			0.272	0.225	0.283	0.189
p-value:	<0.001			0.002	0.011	0.001	0.033
Ulvibacter	0.361	0.479	0.388		0.365	0.496	
p-value:	<0.001	<0.001	<0.001		<0.001	<0.001	
Owenweeksia		0.228	0.224			0.208	
p-value:		0.008	0.010			0.016	
NS3a					-0.177		0.345
p-value:					0.044		<0.001
NS5-DE2	0.404	0.386	0.352	0.243	0.228	0.481	
p-value:	<0.001	<0.001	<0.001	0.005	0.009	<0.001	
NS5-VIS1		-0.354	-0.441		-0.295	-0.256	0.369
p-value:		<0.001	<0.001		<0.001	0.003	<0.001
NS9	-0.188		0.283	0.231			0.285
p-value:	0.033		<0.001	0.008			<0.001
Cytophagia							0.425
p-value:							<0.001

Table S2c: Spearman rank correlations of specific *Bacteroidetes* clades to physicochemical parameters. Correlation values considered only when p-value below 0.05.

	Temperature	Salinity	Chlorophyll *a*	Silicate	Phosphate	Nitrite	Nitrate	Ammonia
CF319a	0.771		0.334	-0.497	-0.583	-0.487	-0.457	-.0193
p-value:	<0.001		<0.001	<0.001	<0.001	<0.001	<0.001	0.026
Polaribacter	0.629		0.201	-0.238	-0.367	-0.450	-0.310	
p-value:	<0.001		0.022	0.006	<0.001	<0.001	<0.001	
Formosa A	0.617			-0.430	-04.467	-05.12	-0.426	
p-value:	<0.001			<0.001	<0.001	<0.001	<0.001	
Formosa B	0.273	-0.326	0.296	-0.551	-0.425			-0.358
p-value:	0.002	<0.001	<0.001	<0.001	<0.001			<0.001
Ulvibacter	0.451		0.453	-0.861	-0.763		-0.518	-0.446
p-value:	<0.001		<0.001	<0.001	<0.001		<0.001	<0.001
Owenweeksia	0.500		0.251	-0.456	-0.474	-0.401	-0.393	
p-value:	<0.001		0.004	<0.001	<0.001	<0.001	<0.001	
NS3a	0.350	-0.429		0.212			0.218	0.225
p-value:	<0.001	<0.001		0.015			0.012	0.010
NS5-DE2	0.619	-0.226	0.403	-0.801	-0.774	-0.217	-0.519	-0.425
p-value:	<0.001	0.009	<0.001	<0.001	<0.001	0.013	<0.001	<0.001
NS5-VIS1		-0.206		0.578	0.439		0.453	0.431
p-value:		0.018		<0.001	<0.001		<0.001	<0.001
NS9	0.536				-0.329	-0.364	-0.294	
p-value:	<0.001				<0.001	<0.001	<0.001	
Cytophagia	0.558	-0.282			-0.240	-0.305		
p-value:	<0.001	0.001			0.006	<0.001		

Chapter 5: Manuscript IV

Mapping glycoconjugate-mediated interactions of marine *Bacteroidetes* with diatoms

Christin M. Bennke, Thomas R. Neu, Bernhard M. Fuchs & Rudolf Amann

Published 2013 in *Systematic and Applied Microbiology*, 36 (6): 417-425

Abstract

The degradation of diatoms is mainly catalyzed by *Bacteroidetes*. This process is of global relevance for the carbon cycle. We used a combination of catalyzed reporter deposition – fluorescence *in situ* hybridization (CARD-FISH) and fluorescent lectin binding analysis (FLBA) to identify and map glycoconjugates involved in the specific interactions of *Bacteroidetes* and diatoms as well as detritus at the coastal marine site Helgoland Roads (German Bight, North Sea). This study probed both the presence of lectin-specific extracellular polymeric substances (EPS) of *Bacteroidetes* for cell attachment and that of glycoconjugates on diatoms with respect to binding sites for *Bacteroidetes*. Members of the clades *Polaribacter* and *Ulvibacter* were shown to form microcolonies within aggregates for which FLBA indicates the presence of galactose containing slime. *Polaribacter* spp. was shown to bind specifically to the setae of the abundant diatom *Chaetoceros* spp. These setae were stained with fucose-specific lectins. In contrast, *Ulvibacter* spp. attached to diatoms of the genus *Asterionella* which bound, among others, the mannose-specific lectin PSA. The newly developed CARD-FISH/FLBA protocol is limited to those glycoconjugates that persist after the initial CARD-FISH procedure. The differential attachment of bacteroidetal clades to diatoms and their discrete staining by FLBA provide evidence for the essential role that formation and recognition of glycoconjugates play in the interaction of bacteria with phytoplankton.

Chapter 6: Manuscript V

Characterisation of Polysaccharide Utilization Loci (PULs) in marine *Bacteroidetes*

Christin M. Bennke, Sixing Huang, Angélique Gobet, Margarete Schüler, Valerie Barbe, Hanno Teeling, Gurvan Michel, Bernhard M. Fuchs & Rudolf Amann

Manuscript in preparation.

Characterisation of Polysaccharide Utilization Loci (PULs) in marine *Bacteroidetes*

Christin M. Bennke[1], Sixing Huang[2], Angélique Gobet[3], Margarete Schüler[4], Valerie Barbe[5], Hanno Teeling[1], Gurvan Michel[3], Bernhard M. Fuchs[1] & Rudolf Amann[1]

[1]Max-Planck-Institute for Marine Microbiology, Celsiusstr. 1, 28359 Bremen, Germany

[2]Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures, Inhoffstr. 7B, 38124 Braunschweig, Germany

[3]Station Biologique de Roscoff, Place Georges Teissier, 29680 Roscoff, France

[4]Max-Planck-Institute for Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany

[5]Genoscope - Centre National de Séquençage, 2 rue Gaston Crémieux CP5706, 91057 Évery, France

Email addresses and telephone numbers of all authors:

C. M. Bennke	cbennke@mpi-bremen.de	+49 421 2028 545
S. Huang	sih13@dsmz.de	+49 531 2616 381
A. Gobet	agobet@sb-roscoff.fr	+33 298 2923 62
M. Schüler	schueler@biochem.mpg.de	+49 89 8578 2617
V. Barbe	vbarbe@genoscope.cns.fr	+33 160 8725 00
H. Teeling	hteeling@mpi-bremen.de	+49 421 2028 976
G. Michel	gurvan.michel@sb-roscoff.fr	+33 298 2923 30
B. M. Fuchs	bfuchs@mpi-bremen.de	+49 421 2028 935
R. Amann	ramann@mpi-bremen.de	+49 421 2028 930

Abstract

In marine systems *Bacteroidetes* are widespread, appearing everywhere, in coastal and open ocean sites, in tropical and polar waters, even in sediments they are found. It is known that cultured strains are capable of degrading high molecular weight organic matter, such as cellulose, chitin and proteins. In a previous study fosmid-based metagenome libraries were constructed, whereof the genomic content of 76 fosmids, affiliated with *Bacteroidetes*, from two contrasting oceanic provinces of the North Atlantic Ocean were reported. It was shown that fosmids retrieved from the Boreal Polar region of the North Atlantic Ocean had a larger potential for polysaccharide degradation and cell surface attachment to algal cells than the oligotrophic subtropical region. Within the present study we analyzed additional 155 fosmids from both regions to gain further insights in the genetic capabilty of *Bacteroidetes* degrading organic matter and in particular polysaccharides. Among those fosmids a higher frequency of genes encoding peptidases, proteases and other genes related to protein degradation were identified at the southern region, corroborating the inital findings. Yet, the frequencies of genes involved in polysaccharide degradation was no longer significantly different between both provinces. The frequency of genes encoding TBDR, SusD and GHs were comparable, only the frequency of sulfatases were higher at the northern region than at the southern.

Further on we identified 14 fosmids carrying gene islands involved in polysaccharide degradation - so called polysaccharide utilization loci (PUL). Among those PULs one seems to be involved in xylan degradation and four were identified as potential laminarin degradation PULs. Interestingly, within the entire fosmid dataset some GHs were identified, which are supposed to be unique among terrestrial *Flavobacteria*, suggesting a higher capability of organic matter degradation by open ocean *Bacteroidetes* clades than previously anticipated from single open ocean genomes.

Introduction

The phylum *Bacteroidetes* encompasses many marine bacteria which are found in coastal environments (ALONSO ET AL. 2007, TEELING ET AL. 2012) but also in the open ocean (SCHATTENHOFER ET AL. 2009, GÓMEZ-PEREIRA ET AL. 2010). They occur both free-living in the water column, but also attached to particles (DELONG ET AL. 1993, BENNKE ET AL. 2013). Furthermore, many members of *Bacteroidetes* are known to be involved in degrading high molecular weight dissolved organic matter (HMW-DOM) such as polysaccharides, chitin and proteins (COTTRELL & KIRCHMAN 2000, COTTRELL ET AL. 2005, TEELING ET AL. 2012, FERNÁNDEZ-GÓMEZ ET AL. 2013). This was corroborated by the first genome analysis of a marine representative of the class *Flavobacteria* within the phylum *Bacteroidetes*, '*Gramella forsetii*' which revealed a high glycolytic and proteolytic potential since it encoded by high numbers of glycoside hydrolases and peptidases (BAUER ET AL. 2006). Also other so far sequenced marine *Flavobacteria* genomes show adaptations towards organic matter degradation (GÓMEZ-PEREIRA ET AL. 2012). High numbers of glycosyl hydrolases (GHs), carbohydrate binding modules (CBM) and other genes involved in polysaccharide degradation as well as peptidases and proteases have for example been detected in the genomes of for example *Formosa agariphila* (MANN ET AL. 2013), *Polaribacter dokdonensis* MED152 (GONZÁLEZ ET AL. 2008), '*Gramella forsetii*' (BAUER ET AL. 2006) and *Robiginitalea biformata* (OH ET AL. 2009).

In a previous study (GÓMEZ-PEREIRA ET AL. 2012) two fosmid based metagenome libraries from two contrasting provinces of the North Atlantic Ocean were constructed (Table 1, Figure S1) to provide first insights into the genomic potential of yet uncultured *Bacteroidetes* inhabiting these oceanic provinces. One library originated from surface water masses retrieved at station 3 (S3) of the VISION cruise (GÓMEZ-PEREIRA ET AL. 2010) in the Boreal Polar region (BPLR) and contained 35,000 fosmids. The other was constructed from surface water collected at station 18 (S18) close to the Azores in the North Atlantic Subtropical

region (NAST) and had a size of 50,000 fosmids. Within the study of GÓMEZ-PEREIRA ET AL. (2012), both libraries were screened with a PCR assay targeting the 16S rRNA gene with *Bacteroidetes* specific primers (CF319 and CF967). A total of 13 (S3) and 15 (S18) fosmids carrying the 16S rRNA gene were identified and fully sequenced. Subsequently, end-sequencing was done on 16,938 fosmids from S3 and 16,255 fosmids from S18. By comparative analysis of end sequences, which possessed an average length of 623 bp, a taxonomic affiliation was predicted based on cumulative evidence inferred from the gene as well as the sequence level (GÓMEZ-PEREIRA ET AL. 2012), an additional set of 183 fosmids of both stations (S3: 98 and S18: 85) were predicted to be affiliated with *Bacteroidetes* and fully sequenced by Genoscope, France (GÓMEZ-PEREIRA ET AL. 2012). The analysis of the first 76 *Bacteroidetes* fosmids (GÓMEZ-PEREIRA ET AL. 2012) revealed that marine *Bacteroidetes* from both regions are specialized in degrading certain polymers. There were also indications that *Bacteroidetes* clades present in the oligotrophic southern region might have a hydrolytic potential more adapted to protein and peptidoglycan degradation than to that of polysaccharides (GÓMEZ-PEREIRA ET AL. 2012).

In this study we present the results obtained in the analysis of additional 155 fosmids selected based on end-sequences (S3: 82 and S18: 73). We also checked the end-sequenced fosmids to see whether they would extend existing fosmids, with the main goal of completing so called polysaccharide utilization loci (PUL). These loci were firstly described in the gut symbiont *Bacteroides thetaiotaomicron* and consist of genes involved in uptake, hydrolysis and transport of carbohydrates. PULs often contain genes related to the starch utilization system (SusA-SusG and SusR; ANDERSON & SALYERS 1989, SHIPMAN ET AL. 2000). However, they are not restricted to starch but encode also degradation other polysaccharides (MARTENS ET AL. 2011). The so called Sus-like systems consists of an outer membrane associated transport protein homologue to SusC also known as TonB-dependent receptors (SCHAUER ET AL. 2008) and an outer membrane associated lipoprotein homologue to SusD

(REEVES ET AL. 1997, SHIPMAN ET AL. 2000, CHO & SALYERS 2001, BJURSELL ET AL. 2006, MARTENS ET AL. 2011) which is unique for *Bacteroidetes* (THOMAS ET AL. 2011). Within PULs *susC* and *susD* homologs are clustering together with other genes encoding carbohydrate degradation such as glycosyl hydrolases (GH), carbohydrate esterases (CE), carbohydrate binding modules (CBM) and polysaccharide lyases (PL). For this group of carbohydrate active enzymes (CAZy) there is a dedicated database (CANTAREL ET AL. 2009, LOMBARD ET AL. 2014). Sulfatases can additionally be part of PULs as it was shown by GÓMEZ-PEREIRA ET AL. (2012) and others (BAUER ET AL. 2006, THOMAS ET AL. 2011, TEELING ET AL. 2012, MANN ET AL. 2013).

The aim of the present study was to analyze this extended fosmid dataset of two contrasting oceanic provinces of the North Atlantic Ocean with a special focus on testing the hypothesis that differing oceanic provinces select for *Bacteroidetes* clades with different PULs (GÓMEZ-PEREIRA ET AL. 2012).

Results and Discussion

We here report two times more high-quality finished fosmid sequences as in the study of GÓMEZ-PEREIRA ET AL. (2012). Thereby, we enlarge the dataset from 76 to 231 fosmids consisting of 131 fosmids retrieved from S3 of a total length of 4.6 Mbp and 100 fosmids from S18 amounting to 4.0 Mbp.

Phylogenetic affiliation

These fosmids were analysed with the newly developed meta-tool 'Taxometer' (WALDMANN ET AL. in prep.) for their taxonomic affiliation. At the station S3 and S18 97% and 93%, respectively, of the fosmid sequences were assigned to *Bacteroidetes*. This indicated that the error rate of selecting *Bacteroidetes* fosmids based on the 1.5 kb Sanger read end-sequences (both ends combined) was with 5% in the same range as with a PCR-based screening (GÓMEZ-PEREIRA ET AL. 2012). The non-*Bacteroidetes* fosmids were affiliated with the PVC-cluster and

Proteobacteria. Within the *Bacteroidetes* 95% of the fosmids were affiliated with the class *Flavobacteria* at S3 and 94% at S18, all of those predicted originate from members of the family *Flavobacteriaceae*. The meta-tool Taxometer further indicated that at S3 38% and 6% of the fosmids were affiliated with the genera *Polaribacter* and *Gramella*, respectively, whereas at S18 only 7% belonged to *Polaribacter* and 2% to *Gramella*. In contrast, at S18 26% and 20% of the fosmids were affiliated with the genera *Dokdonia* and *Leeuwenhoekiella*, respectively, which were much less frequent (<6%) at S3. Fosmids associated with the genus *Robiginitalea* were found in higher numbers at S18 (7.5%) than at S3 (2.5%) and fosmids affiliated with the genus *Flavobacterium* showed similar abundances (12%) at both stations (Figure 1). The taxonomic distribution of the *Bacteroidetes* fosmids is nicely in line with community composition patterns of *Bacteroidetes* obtained in the same samples by catalyzed reporter deposition fluorescence *in situ* hybridization (GÓMEZ-PEREIRA ET AL. 2010).

Genomic content of the Bacteroidetes fosmids

The S3 and S18 fosmid sets had similar frequencies of glycosyl hydrolases (GH; S3: 15.5 ORFs per Mbp, S18: 15.9 ORFs per Mbp), transporters (ABC and TBDR (S3: 4.2 ORFs per Mbp, S18: 4.2 ORFs per Mbp) as well as SusD genes (S3: 2 ORFs per Mbp, S18: 2.2 ORFs per Mbp), but the sulfatase gene frequencies (Table 2, Figure 2) was with 7.7 ORFs per Mbp at S3 twice as high as at S18 (3.5 ORFs per Mbp). In contrast, the S18 fosmid dataset showed a higher gene frequency of peptidases (60 ORFs Mbp^{-1}) than S3 where 50 ORFs Mbp^{-1} were assigned to peptidases. Overall the frequency of genes encoding proteolysis is about 3.5 times higher than that glycolytic activities. This is in good agreement with FERNÁNDEZ-GÓMEZ ET AL. (2013).

A main goal of this study was to investigate whether the composition of GH differed between S3 and S18 (Figure 2). GH family 3 and 16 occurred in similar numbers at both stations (on average 2.5 and 1.4 ORFs Mbp^{-1}, respectively). Both

GH families are predicted to be involved in laminarin degradation (CANTAREL ET AL. 2009, KUMAGAI & OJJMA 2010, KABISCH ET AL. 2014). The ß-1,3/ß-1,6-linked glucans laminarin as well as chrysolaminarin are abundant storage compounds of diatoms (BEATTIE ET AL. 1961). Members of the GH16 were frequently found among marine but also in human gut *Bacteroidetes* (THOMAS ET AL. 2011).

The family GH92, which includes alpha-mannosidase (CANTAREL ET AL. 2009, LOMBARD ET AL. 2014), was found more frequently on fosmids retrieved from S3 (1.8 ORFs Mbp^{-1}) than on fosmids from S18 (0.3 ORFs Mbp^{-1}). In contrast GH74, which is responsible of hydrolyzing ß-1,4-linkages of many different glucans and especially xyloglucans (CANTAREL ET AL. 2009, LOMBARD ET AL. 2014), was found in higher frequencies at S18 (2 ORFs Mbp^{-1}) than at S3 (0.7 ORFs Mbp^{-1}). Nine GHs were exclusively found at S3 (GH2, GH9, GH26, GH29, GH30, GH36, GH93, GH99 and GH106) and five at S18 (GH5, GH31, GH37, GH65 and GH72). Those GHs found only in the S3 fosmids contain among others ß-galactosidases, ß-mannosidases (GH2), endo-ß-1,3/ß-1,4-glucanases, lichenase-laminarinases, ß-glucosidases (GH9), endo-ß-1,4-mannanases, ß-1,3-xylanases (GH26), exo-α-1,3/α-1,4-fucosidases (GH29), endo-ß-1,4-xylanases (GH30), α-galactosidases (GH36), exo-α-L-1,5-arabinase (GH93), endo-α-1,2-mannosidases (GH99) and α-L-rhamnosidases. The GH families found only in the S18 fosmids hydrolyse mainly α-linkages. This suggests that *Bacteroidetes* occurring at S3 have a greater potential degrading algal derived organic matter than at S18, supporting the initial findings of GÓMEZ-PEREIRA ET AL. (2012).

Polysaccharide utilization loci

Within the 155 yet unpublished fosmids we identified in total 14 (S3: 6 and S18: 8) containing PULs (Figure 3 and 4). Each PUL contained a predicted TonB dependent receptor (TBDR) coupled with a SusD-like gene adjacent to glycosyl hydrolase and sulfatase genes.

PULs at S3

At S3 fosmids with PULs were affiliated with the genera *Polaribacter*, *Gramella*, *Flavobacterium*, *Psychroflexus*, *Leeuwenhoekiella* and *Cellulophaga*. The *Polaribacter* fosmid (S3-C51_14YA06) contained at least two sulfatase genes within the PUL. This was similar to the already published *Polaribacter* fosmid S3-860 (GÓMEZ-PEREIRA ET AL. 2012) from the same station. Within the new set of fosmids we identified another fosmid (S3-21YE11) which was enriched in sulfatase genes (5), but did not possess TBDR and SusD-like genes. This fosmid was affiliated with the clade *Flavobacterium*.

The second *Flavobacterium* fosmid (S3-C45) contained genes predicted as GH3 and GH16 downstream of the SusD and TBDR genes. This gene arrangement is present throughout the *Bacteroidetes* phylum and seems to be involved in laminarin degradation (HAHNKE 2013, KABISCH ET AL. 2014). We identified a second fosmid (S3-C33) affiliated with the *Gramella* clade which showed the same gene arrangement as the *Flavobacterium* fosmid. Both fosmids consisted of the TBDR-SusD tandem adjacent to two GH16 genes, which was followed by a GH3 gene and another GH16. Comparing both PULs, they seem to be conserved in their gene arrangement. However, the first GH16 adjacent to the SusD gene on the *Flavobacterium* fosmid (S3-C45) is different to the first GH16 on the Gramella fosmid (S3-C33), whereas both GH3 seemed to be more similar (Figure 5D). The GH16 family represents a broad family encompassing enzymes cleaving β-1,3 or β-1,4 glycosidic bonds in a variety of glucans and galactans. This family possesses enzymes with various substrate specificities such as agar, κ-carrageenan, xyloglucan, β-1,3-glucans and β-1,3-galactans (LOMBARD ET AL. 2014). Therefore, those GH16 found on the *Gramella* and *Flavobacterium* fosmid might encode different enzyme activities.

The remaining fosmids were affiliated with the genera *Leeuwenhoekiella* (S3-27YE12_C102), *Psychroflexus* (S3-C42) and *Cellulophaga* (S3-C99). Both, *Psychroflexus* and *Cellulophaga* fosmids contained two GH92 genes, which likely encode exo-acting

α-mannosidase. However, the GH92 on the *Psychroflexus* fosmid (S3-C42) was located upstream of the TBDR and SusD genes adjacent to a predicted sodium-dependent mannose transporter. In contrast the GH92 on the *Cellulophaga* fosmid (S3-C99) was located downstream of the TBDR-SusD tandem. On the *Leeuwenhoekiella* fosmid (S3-27YE12_C102) the TBDR-SusD tandem was identified at the 5' end. We were able to prolong this fosmid yielding an insert length of 64.5 kb. However, we were only able to prolong this fosmid in direction of the 3' end. There, most of the predicted genes were identified as house-keeping genes and hypothetical proteins, except for two additional sulfatase genes which were found adjacent to a GH106 predicted to act as an α-L-rhamnosidase. On the same fosmid another TBDR was identified, but not in conjunction with a SusD gene. This TBDR was shorter in their amino acid sequence and is most likely not involved in carbohydrate uptake, but may function instead as carrier protein for vitamin B12 uptake (KOEBNIK 2005, SCHAUER ET AL. 2008).

PULs at S18

At S8 we identified eight fosmids with PULs (Figure 4). These fosmids were affiliated with the family *Flavobacteriaceae* except of one which was predicted to be affiliated with intestinal genus *Bacteroides* (S18-C59). This fosmid was enriched in sulfatases and contained many different GHs as well as a second TBDR gene and a predicted maltose transporter. The second TBDR gene was again shorter in their amino acid sequence, comparable with the second TBDR on the *Leeuwenhoekiella* fosmid (S3-27YE12_C102) from S3. However, in contrast to the second TBDR on the *Leeuwenhoekiella* fosmid from S3, the second TBDR on the *Bacteroides* fosmid was neighbored by cazymes (GH20, CBM9 and GH5). Most likely this second TBDR is still involved in carbohydrate uptake, even in the absence of the SusD gene. The predicted maltose transporter was localized adjacent to two GH65 genes. The GH65 family contains hydrolases acting on α-glycosidic bonds and possesses mainly phosphorylases, like maltose phosphorylases (CANTAREL ET AL. 2009, LOMBARD ET

AL. 2014). Both, the maltose transporter and the adjacent GH65 were found upstream to the TBDR-SusD tandem. More glycosyl hydrolases from different families (GH5, GH20, GH43, GH78 and GH109) were identified downstream to the SusD and TBDR system. The GH43 was found adjacent to a sulfatase and this glycosyl hydrolase family represents many enzymes involved in xylan degradation (DODD & CANN 2009), e.g. α-L-arabinofuranosidases, endo-α-L-arabinases and β-D-xylosidases. Another glycosyl hydrolases belonged to the family 109 and was localized between two sulfatase genes. This GH was predicted to act as α-N-acytylgalactosaminidases. The GH20, GH5 and GH78 were identified towards the 5' end. GH78 is characterized as α-L-rhamnosidase and GH20 as well as GH5 represent both broader families for which activities are difficult to predict. GH5 was formerly characterized as Cellulase family A (HENRISSAT ET AL. 1989) and this family incorporates a variety of specific enzymes possessing exo- and endo-activities, e.g. endo- and exoglucanases, endo- and exomannases as well as β-glucosidases and β-mannosidases. The GH20 contains mainly exo-acting N-acetylglucosaminidases and N-acetygalacosaminidases (CANTAREL ET AL. 2009, LOMBARD ET AL. 2014). Most of the glycosyl hydrolase families consist of many enzymes with different substrate specificities (HENRISSAT 1991, HENRISSAT & BAIROCH 1996, DAVIES ET AL. 2008).

Two fosmids, one was affiliated with the clade *Flavobacterium* (S18-C30) and the other with *Robiginetalea* (S18-C31), contained PULs potentially encoding laminarin degradation. A second *Flavobacterium* fosmid (S18-34YL06) possessed many sulfatases and GHs and seemed to be involved in xylan degradation. Upstream of the SusD-TBDR-system four GHs were identified belonging to the families 3, 5 and 10. Between two GH3 genes a carbohydrate esterase family 6 was identified as well. This is a predicted acetyl-xylan esterase, removing acetyl residues from the xylan backbone (MARGOLLES-CLARK ET AL. 1996). The GH3 family consists of many enzymes possessing different substrate specificities and often these enzymes are bifunctional, for example an enzyme having α-L-arabinofuranosidase and β-D-xylopyranosidase

activity (LEE ET AL. 2003). This GH family contains also enzymes with xylan 1,4-β-xylosidase activity (LOMBARD ET AL. 2014). The GH families 5 or 10 possess enzymes hydrolyzing the β-1,4 glycosidic bonds within the xylan backbone generating short chain oligomers (CANTAREL ET AL. 2009, LOMBARD ET AL. 2014). Downstream of the SusD and TBDR system again GH10 were identified as well as GH43 possessing α-L-arabinofuranosidase activity and cleaving of arabinofuranose side chains (DODD & CANN 2009). Adjacent to the GH43 and GH10 a xyloside transporter (xynT) was identified, which could be involved in transporting xyloside into the cytoplasma, where it then is further transformed into xylose and further into xylulose and xylulose-5-phosphate. Further downstream we identified the genes xylA and xylB. Both are involved in converting xylose in xylulose (xylA - xylose-isomerase) and further phosphorylating xylulose (xylB - xylulose-kinase) which than enters as xylulose-5-phosphate the pentose-phosphate-pathway and will further metabolized. Our findings are in good agreement with the results of ARNOSTI ET AL. (2012), who showed with fluorescein-labeled polysaccharides not only the potential for xylan hydrolysis at S18 , but also potential for laminarin degradation at S3 and S18 (ARNOSTI ET AL. 2012). We also identified two potentially laminarin PULs at both stations.

On the *Flavobacterium* fosmid S18-C63 two TBDR-SusD tandem systems were identified. The one at the 5' end is most likely involved in carbohydrate uptake, since the adjacent genes were predicted as GH13 and GH3. The GH13 family contains enzymes acting on α-glucoside linkages, like α-amylases and pullulanases (LOMBARD ET AL. 2014). The other TBDR-SusD unit at the 3' end seemed to be involved in protein uptake, since the neighboring gene was identified as a peptidase.

Of three more fosmids containing TBDR-SusD units two were affiliated with the genus *Dokdonia* (S18-C68 & S18-5a4YN02_C90) and the third one (S18-C52) to *Capnocytophaga*. All three fosmids have in common that the genes neighboring TDBR-SusD were only identified as hypothetical proteins. We can only speculate

that the respective gene products are also involved in carbohydrate degradation potentially indicating the presence of yet unknown enzymes involved in organic matter degradation.

Comparative analysis of PULs

To see whether the genes of the PUL-carrying fosmids were unique or if homology can be identified we compared them among the stations by using the alignment program MAUVE (DARLING ET AL. 2010). The PUL fosmids revealed several locally collinear blocks (LCB), not only within the fosmids of S3 or S18, but also between both stations. For example a region between 12.89-15.82 kbp of the *Flavobacterium* fosmid S3-21YE11 was highly similar to a region between 51.28-54.32 kbp of the *Leeuwenhoekiella* fosmid S3-27YE12_C102 (Figure 5A). Within this region two ORFs were identified and annotated as L-arabinose isomerase and sulfatase. The similarity profile of the LCB indicates a high level of conservation between both fosmids. The same *Flavobacterium* fosmid contained another region between 23.54-33.45 kbp which seemed to be homologue to another region between 12.00-22.81 kbp of the *Leeuwenhoekiella* fosmid. However, the *Leeuwenhoekiella* fosmid has an insertion between 19.54 and 20.10 kbp. The homologue region consisted of seven ORFs and were annotated as L-rhamnose permease, L-lactate dehydrogenase, rhamnulokinase, hypothetical protein, rhamnulose-1-phosphate aldolase, fructose-1,6-bisphosphatase and as transcriptional regulator of rhamnose utilization.

The *Psychroflexus* fosmid S3-C42 included a region between 26.60-28.87 kbp which was highly similar to a region between 13.11-15.36 kbp of the *Cellulophaga* fosmid (S3-C99). This region was annotated as a GH92 and from the similarity profile of the LCB a high level of conservation between both fosmids can be assumed (Figure 5B).

Comparing the PUL fosmids among the provinces it was shown that also conserved regions on the fosmids between the stations existed. For example the GH3 on the *Polaribacter* fosmids S3-C51_15YA06 were highly similar to the GH3 on the *Flavobacterium* fosmids S18-34YL06. Interestingly, this GH3 was neighbored by three sulfatase genes on both fosmids, but MAUVE identified no homology between the sulfatase genes, only with the GH3 (Figure 5C).

Four fosmids related to laminarin degradation, two from each station, showed a higher level of conservation for the PUL region (Figure 5D). However, on the *Robiginitalea* fosmid S18-C31 another GH16 was inserted after the SusD gene, which was highly similar to the GH16 adjacent to the SusD of the *Flavobacterium* fosmid S3-C45. These findings are in good agreement with HAHNKE (2013, p. 229) who showed that PULs potentially involved in laminarin degradation are conserved among *Bacteroidetes*.

Gliding ability

Besides the identification of PULs, the combined fosmid dataset was screened for gliding ability, since *Bacteroidetes* are known to move by gliding (HALL-STOODLEY ET AL. 2004, MCBRIDE & ZHU 2012). We identified four fosmids at each station which carried gliding genes predicted as gldB, gldC, gldF, gldG and gldH. GldB has been predicted being membrane associated and involved in production of glycoproteins. The gldC is not directly required for gliding motility but enhances colony spreading (HUNNICUTT & MCBRIDE 2000). The gldF and gldG are required for cell movement (HUNNICUTT ET AL. 2002). The gldH represents a lipoprotein which is also involved in cell movement (MCBRIDE ET AL. 2003).

For S3 two fosmids were affiliated with the genus *Flavobacterium* and genes annotated as gldH and gldG were identified at the 5' end of the fosmid sequence. Two other gldH genes were identified on fosmids affiliated with *Gramella* and

Polaribacter. The closest relatives of theses fosmids comprise all of the full set of 14 gliding genes (MCBRIDE 2004, BAUER ET AL. 2006, GONZÁLEZ ET AL. 2008).

Also at S18 one fosmid, affiliated with *Flavobacterium*, was carrying the gldH gene. Besides the gldH gene other gliding genes (gldB, gldC and gldF) were identified on several fosmids from that station. Two fosmids were affiliated with *Dokdonia* and one of them carried the genes gldB and gldC located adjacent to each other. The gldC gene was also identified on another *Dokdonia* fosmid. On the genomes of *Dokdonia donghaensis* MED134 (GÓMEZ-CONSARNAU ET AL. 2007), *Polaribacter* MED152 (GONZÁLEZ ET AL. 2008) and '*Gramella forsetii*' (BAUER ET AL. 2006) all genes involved in gliding were identified supporting our findings. Furthermore, on three fosmids at S18 and one at S3 genes annotated as putative adhesion lipoproteins were identified. Those adhesion lipoproteins were postulated to be involved in cell-surface contact (SATO ET AL. 2010; NAKANE ET AL. 2013).

Interestingly, KOLTON ET AL. (2013) reported that GH78 and GH106, responsible for rhamnogalacturonan utilization, were only found in terrestrial genomes, which suggests an adaption of terrestrial *Bacteroidetes* strains to plant-related carbohydrate metabolism. Rhamnogalacturonan is exclusively found as part of the terrestrial plant hemicelluloses. Within the entire fosmid set we identified gene frequencies of GH78 for both stations around 0.25 ORFs per Mbp and GH106 was solely identified at the northern region (S3) possessing a frequency of 0.7 ORFs per Mbp. Additionally, KOLTON ET AL. (2013) found even more CAZy domains which are supposed to be unique for terrestrial *Bacteroidetes* strains such as GH10, GH43 or CBM6. Although the gene frequencies were not remarkably high, we identified other CAZymes which are unique among terrestrial *Bacteroidetes* clades. For example, GH43 was found at both stations with a gene frequency of 1.1 ORFs per Mbp for S3 and 1.2 ORFs per Mbp for S18. The GH43 is predicted to be involved in xylan degradation. Xylan is not only found in terrestrial plants as part of hemicelluloses, but also in marine algae in form of cell covering (OKUDA 2002) and in some diatoms

as cell wall component (WUSTMAN ET AL. 1998. MURRAY ET AL. 2007). Our findings of enzymes involved in xylan degradation corroborates the results of ARNOSTI ET AL. (2012), where they found enzymatic activity for xylan degradation in both provinces. The identification of GHs classified as unique for terrestrial plant degradation, such as GH43, GH78, and GH106, in our fosmid dataset is in contrast to KOLTON ET AL. (2013) findings. However, it is believed that the early land plants evolved from green algae (SØRENSEN ET AL. 2011). Most likely, the input on such material is higher at the northern region than at the southern, since S3 is located at the lower border of the East Greenland Current which transports cold, low saline, but nutrient- and phytoplankton-rich waters southwards (BERSCH 1995).

Conclusion

The present study is extending the study of GÓMEZ-PEREIRA ET AL. (2012) and communicates the sequence-based analysis of 155 fosmids from two different oceanic provinces. Fosmid-based metagenomics was chosen since it provides contiguous DNA fragments of an average length of 35-40 kb revealing not only about 1% of the genome of a marine *Bacteroidetes*, but also offering a good chance to find operons or larger genomic loci in comparison to high-throughput sequencing technologies, where only short genomic fragments, e.g. in a size of a single gene, are generated, fosmids allow for in-depth analysis of longer DNA fragments including PULs.

Based on the analysis of 76 fosmids, GÓMEZ-PEREIRA ET AL. (2012) had found a trend that bacteroidetal fosmids from the boreal station S3 were enriched in genes encoding polysaccharide degradation whereas in those from the subtropical gyre station S18 proteinases were more frequent. Our analysis of the extended fosmid dataset of now 231 fosmids does not fully support these earlier findings. There is still a higher frequency of genes encoding peptidases, proteases and other

genes related to protein degradation at station 18, yet the frequencies of genes involved in polysaccharide degradation was no longer significantly different between S3 and S18. The frequency of genes encoding TBDR, SusD and GHs were comparable, only the frequency of sulfatases was higher at S3 than at S18. This is in good agreement with the findings of ARNOSTI ET AL. (2012) who measured at the same stations the potential of microbes inhabiting the large-particle fraction (>10 μm fraction) to enzymatically hydrolyze high molecular weight substrates. This study had indicated higher potential rates of degradation for the sulfated polysaccharides chondroitin and fucoidan at S3 than S18. As it stands now, 4 of 131 fosmids from S3 have colocalizations of GHs and sulfatases. This includes fosmids S3-860 and S3_DL_C5 already identified by GÓMEZ-PEREIRA ET AL. 2012 and fosmids S3-C51_15YA06 and S3-21YE11 first reported here. Among the 100 bacteroidetal fosmids analyzed from S18 only S18-C59 and S18-34YL06 showed this co-localization. Even with this numerical extension of the study of GOMEZ-PEREIRA ET AL. 2012 these low numbers do still not provide good statistical support.

Our data indicate that, although the overall number of GHs was similar for both stations, their composition differed slightly. S18 contained additional GHs hydrolyzing carbohydrates characterized mainly by α-linkages. Alpha-linked carbohydrates such as α-1,4-glucans are often produced by bacteria as a storage compound during their stationary growth phase (PREISS 1984, FIELD ET AL. 1998). This suggests a potential lifestyle for *Bacteroidetes* occurring at S18 in feeding not only on algal derived organic matter but also on bacterial derived α-glucans.

The alignment tool MAUVE revealed a moderate conserved sequential gene arrangement of TBDR-SusD systems adjacent to GH16, GH3, GH16 between the provinces. This typical gene sequence was identified in numerous members of the *Bacteroidetes* independent of their origin (KABISCH ET AL. 2014). It will be interesting in future studies, how ubiquitously distributed such a characteristic PUL is and it

needs to be explored in metatranscriptomics or metaproteomics studies to test, under which circumstances these PULs are expressed.

This fosmid-based metagenome dataset adds to our increasing knowledge on the importance of PULs for polysaccharide utilization generated by genome (BAUER ET AL. 2006, GÓMEZ-CONSARNAU ET AL. 2007, GONZÁLEZ ET AL. 2008) as well as metagenome analysis (TEELING ET AL. 2012, WILLIAMS ET AL. 2013, GEORGES ET AL. 2014). So far, we only reached the tip of the iceberg and further work has to be done unveiling the enzymology and transport mechanisms involved in marine organic matter degradation.

Material and Methods

Study site

Samples were taken in the North Atlantic Ocean during the VISION cruise MSM03/01 on board of the research vessel Maria S. Merian in September 2006 (Figure S1, Table 1). Fosmid-based metagenomic libraries were constructed from samples of two contrasting oceanic provinces. Samples from S3 were collected in the Boreal Polar (BPLR) Province (65°52.64' N, 29°56.54' W) and samples from S18 in the North Atlantic Subtropical Province (34°04.43' N, 30°00.09' W), according to LONGHURST (1998). The accompanying environmental parameters are summarized in Table 1 and have been described in GÓMEZ-PEREIRA ET AL. (2010). Fosmid library construction and selection procedure were described in details in GÓMEZ-PEREIRA ET AL. (2012).

Selection for fosmid prolongation

In a next step the entire dataset of the fully sequenced fosmids were mapped onto the remaining end-sequenced fosmids to see whether they can be prolonged. This mapping yielded in additional 83 fosmids for S3 and 47 fosmids for S18.

Considering only those end-sequenced fosmids for full sequencing which yielded a mapping similarity of more than 94.5% the number of fosmids were reduced to 43 fosmids for S3 and 27 for S18. Since we were only interested in prolonging fosmids containing PULs, we selected 11 fosmids for S3 and 9 for S18. Sequencing of this additional fosmid set was conducted at LGC Genomics (LGC Genomics GmbH, Berlin, Germany) using 454-pyrosequencing technique.

Assembly

Expecting that the second new fosmid dataset will prolong existing fosmids all sequences were pooled within each station and assembled again with SeqMan (Lasergene 8). Default setting was used, except single sequence contigs were added. The assembly quality was checked via the strategy view option within the program. Sequences were exported as FASTA files for annotation.

Gene prediction and annotation

For the published fosmid dataset the in-house pipeline for gene prediction and annotation was used. When retrieving the remaining fosmids we uploaded the entire dataset to RAST and used their gene prediction and annotation pipeline (AZIZ ET AL. 2008). To compare the published dataset with the new one, at first we re-evaluated the gene prediction and annotation from the in-house pipeline with the RAST prediction and annotation. Small scale differences occurred in the gene prediction, but the annotations were as good as the annotations yielded with RAST. Annotation files were downloaded in a 'Genbank' file format and incorporated in the in-house pipeline for taxonomic classification and detailed CAZyme as well as peptidase annotation. The protein fosmid sequences were searched against the CAZy database (CANTAREL ET AL. 2009) and against the MEROPS 9.8 database (RAWLINGS ET AL. 2012). For every protein, the best hit with an e-value below E-15 was kept.

PUL identification and visualization

The PULs were identified and analyzed with a customized Python script (Huang 2013, pp. 22-23). In short, annotations from various tools (BLAST, Pfam and CAZy) were summarized for each ORF in a contig. The whole contig was selected for detailed analysis, if genes involved in polysaccharide degradation were present. Functions considered to be polysaccharide degradation related are CAZymes, membrane transporters, such as the TonB components and ABC transporters, numerous substrate-binding modules. Fosmids containing PULs were visualized using CorelDRAW® X4 (CorelDRAW Graphic Suits X4, Corel GmbH, Munich, Germany).

PUL comparison

The fosmids caring PULs from both stations were compared by using the MAUVE alignment tool (DARLING ET AL. 2010). This comparison was done to identify conserved regions between the fosmids independent of their origin. When an alignment has been computed with the Progressive Mauve algorithm, regions conserved between the fosmids were identified as locally collinear blocks (LCB). Conserved LCBs are recognized by similar coloring of the LCB outline. Each of these block outlines surrounds a region of the fosmid sequence which aligned to a part of another fosmid sequence, and is most likely homologous and internally free from genomic rearrangement (DARLING ET AL. 2010). Regions outside the LCBs lack detectable homology among fosmids input. Within each LCB a similarity profile is drawn. The height of the similarity profile corresponds to the average level of conservation in the distinct region. Areas that are completely white within the LCB were not aligned and probably contain sequence elements specific to the particular fosmid.

Acknowlegements

We thank the Captain and Crew of the FS Maria S. Merian (cruise MSM03/01) for their support during the cruise. We greatly appreciate the contribution of Dr. Richard Hahnke with the genome alignment tool MAUVE. This study was funded by the Max-Planck-Society and the NoE Marine Genomics Europe.

References

1. ALONSO, C., WARNECKE, F., AMANN, R., PERNTHALER, J. (2007): High local and global diversity of *Flavobacteria* in marine plankton. *Environ. Microbiol.* **9**, 1253-1266.

2. ANDERSON, K. L. & SALYERS, A. A. (1989): Genetic evidence that outermembrane binding of starch is required for starch utilization by *Bacteroides thetaiotaomicron*. *J. Bacteriol.* **171,** 3199-3204.

3. ARNOSTI, C., FUCHS, B. M., AMANN, R. & PASSOW, U. (2012): Contrasting extracellular enzyme activities of particle-associated bacteria from distinct provinces of the North Atlantic Ocean. *Front Microbiol.* **3**, 1-9.

4. AZIZ, R. K., BARTELS, D., BEST, A. A., DEJONGH, M., DISZ, T., EDWARDS, R. A., FORMSMA, K., GERDES, S., GLASS, E. M., KUBAL, M., MEYER, F., OLSEN, G. J., OLSON, R., OSTERMAN, A. L., OVERBEEK, R. A., MCNEIL, L. K., PAARMANN, D., PACZIAN, T., PARRELLO, B., PUSCH, G. D., REICH, C., STEVENS, R., VASSIEVA, O., VONSTEIN, V., WILKE, A. & ZAGNITKO, O. (2008): The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* **9,** published online

5. BAUER , M., KUBE, M., TEELING, H., RICHTER, M., LOMBARDOT, T., ALLERS, E., WÜRDEMANN, C. A., QUAST, C., KUHL, H., KNAUST, F., WOEBKEN, D., BISCHOF, K., MUSSMANN, M., CHOUDHURI, J. V., MEYER, F., REINHARDT, R., AMANN, R. I. & GLÖCKNER, F. O. (2006): Whole genome analysis of the marine *Bacteroidetes* '*Gramella forsetii*' reveals adaptations to degradation of polymeric organic matter. *Environ Microbiol.* **8,** 2201-2213.

6. BEATTI A., HIRST, E. L. & PERCIVAL, E. (1961): Studies on the metabolism of the *Chrysophyceae*. Comparative structural investigations on leucosin (chrysolaminarin) separated from diatoms and laminarin from the brown algae. *Biochem. J.* **79**, 531-537.

7. BENNKE, C. M., NEU, T. R., FUCHS, B. M. & AMANN, R. (2013): Mapping glycoconjugate-mediated interactions of marine *Bacteroidetes* with diatoms. *Syst. Appl. Microbiol.* **36**, 417-425.

8. BERSCH, M. (1995): On the circulation of the northeastern North Atlantic. Deeps Sea Res. I **42**, 1583-1607.

9. BJURSELL, M. K., MARTENS, E. C. & GORDON, J. I. (2006): Functional genomic and metabolic studies of the adaptations of a prominent adult human gut symbiont, *Bacteroides thetaiotaomicron*, to the suckling period. *J. Biol. Chem.* **281**, 36269-36279.

10. CANTAREL, B. L., COUTINHO, P. M., RANCUREL, C., BERNARD, T., LOMBARD, V., & HENRISSAT, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.* **37,** D233-D238.

11. CHO, K. H., & SALYERS, A. A. (2001): Biochemical analysis of interactions between outer membrane proteins that contribute to starch utilization by *Bacteroides thetaiotaomicron*. *J Bacteriol.* **183,** 7224-7230.

12. COTTRELL, M. T. & KIRCHMAN, D. L. (2000): Natural assemblages of marine proteobacteria and members of the *Cytophaga-Flavobacter* cluster consuming low- and high-molecular-weight dissolved organic matter. *Appl. Environ. Microbiol.* **66,** 1692-1697.

13. COTTRELL, M. T., YU, L. Y. & KIRCHMAN, D. L. (2005): Sequence and expression analyses of Cytophaga-like hydrolases in a Western arctic metagenomic library and the Sargasso seat. Appl. Environ. Microbiol. 71, 8506-8513.

14. DARLING, A. C. E., MAU, B., BLATTNER, F. R. & PERNA, N. T. (2010): progressive Mauve: Multiple Genome Alignment with gene gain, loss, and rearrangement. *PLoS One*. **5**, e11147.

15. DAVIES, G. J. & SINNOTT, M. L. (2008): Sorting the diverse: the sequence-based classifications of carbohydrate-active enzymes. *Biochem. J.* (only online available). DOI: 10.1042/BJ20080382

16. DELONG, E.F., FRANKS, D.G., ALLDREDGE, A.L. (1993) Phylogenetic diversity of aggregate-attached vs. free-living marine bacterial assemblages. Limnol. Oceanogr. 38, 924-934DODD, D. & CANN, I. K. O. (2009): Enzymatic deconstruction of xylan for biofuel production. *Glob. Change Biol. Bioenergy*. **1**, 2-17.

17. DODD, D. & CANN, I. K. O. (2009): Enzymatic deconstruction of xylan for biofuel production. *Glob. Change Biol. Bioenergy*. **1**, 2-17.

18. FERNÁNDEZ-GÓMEZ, B., RICHTER, M., SCHÜLER, M., PINHASSI, J., ACINAS, S. G., GONZÁLEZ, J. M. & PEDRÓS-ALIÓ, C. (2013): Ecology of marine *Bacteroidetes*: a comparative genomics approach. *ISME J*. **7**, 1026-1037.

19. FIELD, C. B., BEHRENFELD, M. J., RANDERSON, J. T. & FALKOWSKI, P. (1998): Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. *Sience* **281**, 237-240.

20. GEORGES, A. A., EL-SWAIS, H., CRAIG, S. E., LI, W. K. W. & WALSH, D. A. (2014): Metaproteomic analysis of a winter to spring succession in coastal northwest Atlantic Ocean microbial plankton. *ISME J.,* published online

21. GÓMEZ-CONSARNAU, L., GONZÁLEZ, J. M., COLL-LLADÓ, M., GOURDON, P., PASCHER, T., NEUTZE, R., PEDRÓS-ALIÓ, C. & PINHASSI, J. (2007): Light stimulates growth of proteorhodopsin-containing marine *Flavobacteria*. *Nature* **11**, 210-213.

22. GÓMEZ-PEREIRA, P. R., ALONSO, C., OLIVER, M., VAN BEUSEKOM, J., FUCHS, B. M. (2010) Distinct flavobacterial communities in contrasting water masses of the North Atlantic Ocean. *ISME J*. **4**, 472-487.

23. GÓMEZ-PEREIRA, P. R., SCHÜLER, M., FUCHS, B. M., BENNKE, C., TEELING, H., WALDMANN, J., RICHTER, M., BARBE, V., BATAILLE, E., GLÖCKNER, F. O., AMANN, R. (2012) Genomic content of uncultured *Bacteroidetes* from contrasting oceanic provinces in the North Atlantic Ocean. *Environ. Microbiol.* **14**, 52-66.

24. GONZÁLEZ, J. M., FERNÁNDEZ-GÓMEZ, B., FERNÀNDEZ-GUERRA, A., GÓMEZ-CONSARNAU, L., SÁNCHEZ, O., COLL-LLADÓ, M., DEL CAMPO, J., RODRÍGUEZ-

MARTÍNEZ, R., ALONSO-SÁEZ, L., LATASA, M., PAULSEN, I., NEDASHKOVSKAYA, O., LEKUNBERRI, I., PINHASSI, J. & PEDRÓS-ALIÓ, C. (2008): Genome analysis of the proteorhodopsin containing marine bacterium *Polaribacter* sp. MED152 (*Flavobacteria*). *Proc. Natl. Acad. Sci. USA* **105,** 8724-8729.

25. HAHNKE, R. (2013): Cultivation of *Flavobacteria* and other in situ abundant bacteria from the North Sea. *Doctoral thesis*, University Bremen, pp. 295.

26. HALL-STOODLEY, L., COSTERTON, J. W., & STOODLEY, P. (2004): Bacterial biofilms: from the natural environment to infectious diseases. *Nat. Rev. Microbiol.* **2,** 95-108.

27. HENRISSAT, B., CLAEYSSENS, M., TOMME, P., LEMESLE, L., & MORNON, J. P. (1989): *Cellulase families revealed by hydrophobic cluster analysis.* Gene. **81**, 83-95.

28. HENRISSAT, B. (1991): A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **280,** 309-316.

29. HENRISSAT, B. & BAIROCH, A. (1996): Updating the sequence-based classification of glycosyl hydrolases. *Biochem. J.* **316**, 695-696.

30. HUNNICUTT, D. W. & MCBRIDE, M. J. (2000): Cloning and Characterization of the *Flavobacterium johnsoniae* Gliding-Motility Genes *gldB* and *gldC*. *J. Bacteriol.* **182,** 911-918.

31. HUNNICUTT, D. W., KEMPF, M. J. & MCBRIDE, M. J. (2002): Mutations inFlavobacterium johnsoniae gldF and gldG disrupt gliding motility and interfere with membrane localization of gldA. *J. Bacteriol.* **184**, 2370-2378.

32. KABISCH, A., OTTO, A. , KÖNIG, S., BECHER, D., ALBRECHT, D., SCHÜLER, M., TEELING, H., AMANN, R. & SCHWEDER, T. (2014): Functional characterization of polysaccharide utilization loci in the marine *Bacteroidetes* '*Gramella forsetii*' KT0803. *ISME J.* doi:10.1038/ismej.2014.4

33. KOEBNIK, R. (2005): TonB-dependent trans-envelope signaling: the exception or the rule? *Trends Microbiol.* **13**, 343-347.

34. KOLTON, M., SELA, N., ELAD, Y. & CYTRYN, E. (2013): Comparative Genomic Analysis indicates that niche adaption of terrestrial *Flavobacteria* is strongly linked to plant glycan metabolism. *PloS one*, **e76704**.

35. KUMAGAI, Y. & OJIMA, T. (2010): Isolation and characterization of two types of beta-1,3-glucanases from the common sea hare *Aplysia kurodai*. *Comp.Biochem. Physiol.* **155,** 138-144.

36. LEE, R. C., HRMOVA, M., BURTON, R. A., LAHNSTEIN, J. & FINCHER, G. B. (2003): Bifunctional family 3 glycoside hydrolases from barley with alpha -L-arabinofuranosidase and beta -D-xylosidase activity. Characterization, primary structures, and COOH-terminal processing. *J. Biol. Chem.* **278,** 5377-87.

37. LOMBARD, V., RAMULU, H. G., DRULA, E., COUTINHO, P. M. & HENRISSAT, B. (2014): The carbohydrate-active enzymes database. *Nucleic Acids Res.* **42**, D490-D495.

38. LONGHURST, A. (1998). *Ecological Geography of the Sea.* Academic Press.

39. MANN, A. J., HAHNKE, R. L., HUANG, S., WERNER, J., XING, P., BARBEYRON, T., HUETTEL, B., STÜBER, K., REINHARDT, R., HARDER, J., GLÖCKNER, F. O., AMANN, R. I. & TEELING, H. (2013): The genome of the algae-associated marine *Flavobacterium Formosa agariphila* KMM 3901T reveals a broad potential for the degradation of algal polysaccharides. *Appl. Environ. Microbiol.* **79**, 6813-6822.

40. MARGOLLES-CLARK, E., TENKANEN, M., SÖDERLAND, H. & PENTTILÄ, M. (1996): Acetyl xylan esterase from *Trichoderma reesei* contains an active site serine and a cellulose binding domain. *Eur. J. Biochem.* **237,** 553-560.

41. MARTENS, E. C., LOWE, E. C., CHIANG, H., PUDLO, N. A., WU, M., MCNULTY, N. P., ABBOTT, D. W., HENRISSAT, B., GILBERT, H. J, BOLAM, D. N. & GORDON, J. I. (2011): Recognition and Degradation of Plant Cell Wall Polysaccharides by Two Human Gut Symbionts. *PLoS Biol.* **9**, e1001221.

42. MCBRIDE, M. J. & ZHU, Y. (2012): Gliding Motility and Por Secretion System Genes Are Widespread among Members of the Phylum *Bacteroidetes*. *J. Bacteriol.* ***195,*** 270-278.

43. MCBRIDE, M. J. (2004): *Cytophaga Flavobacterium* gliding motility. *J. Mol. Microbiol. Biotechnol.* **7,** 63-71.

44. MCBRIDE, M. J., BRAUN, T. F. & BRUST, J. L. (2003): *Flavobacterium johnsoniae* GldH is a lipoprotein that is required for gliding motility and chitin utilization. *J. Bacteriol.* **185**, 6648-6657.

45. MURRAY, A. E., ARNOSTI, C., DE LA ROCHA, C. L., GROSSART, H.-P. & PASSOW, U. (2007): Microbial dynamics in autotrophic and heterotrophic seawater mesocosms. II. Bacterioplankton community structure and hydrolytic enzyme activities. *Aquat. Microb. Ecol.* **49**, 123-141.

46. NAKANE, D., SATO, K., WADA, H., MCBRIDE, M. J. & NAKAYAMA, K. (2013): Helical flow of surface protein required for bacterial gliding motility. *PNAS* **110**, 11145-11150.

47. OH, H. M., GIOVANNONI, S. J., LEE, K., FERRIERA, S., JOHNSON, J. & CHO, J. C. (2009): Complete genome sequence of *Robiginitalea biformata* HTCC2501. *J. Bacteriol.* **191**, 7144-7145.

48. OKUDA, K. (2002): Structure and phylogeny of cell coverings. J. Plant. Res. 115, 283-288.

49. PREISS, J. (1984): Bacterial glycogen-synthesis and it structure. Ann. Rev. Microbiol. 38, 419-458.

50. RAWLINGS, N.D., BARRETT, A. J., & BATEMAN, A. (2012): MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids. Res.* **40**, D343-D350.

51. REEVES, A. R., WANG, G. R. & SLAYERS, A. A. (1997): Characterization of four outer membrane proteins that play a role in utilization of starch by Bacteroides thetaiotaomicron. J. Bactetiol. 179, 643-649.

52. SATO, K., NAITO, M., YUKUTAKE, H., HIRAKAWA, H., SHOJI, M., MCBRIDE, M. J., RHODES, R. G. & NAKAYAMA, K. (2010): A protein secretion system linked to bacteroidete gleiding motility and pathogenesis. *PNAS* **107**, 276-281.

53. SCHATTENHOFER, M., FUCHS, B. M., AMANN, R., ZUBKOV, M. V., TARRAN, G. A. & PERNTHALER, J. (2009): Latitudinal distribution of prokaryotic picoplankton populations in the Atlantic Ocean. *Environ. Microbiol.* **11**, 2078-2093.

54. SCHAUER, K., RODIONOV, D. A., DE REUSE H. (2008): New substrates for TonB-dependent transport: do we only see the 'tip of the iceberg'? *Trends Biochem. Sci.* **33**, 330-338.

55. SHIPMAN, J. A., BERLEMAN, J. E., & SALYERS, A. A. (2000): Characterization of four outer membrane proteins involved in binding starch to the cell surface of *Bacteroides thetaiotaomicron*. *J. Bacteriol.* **182,** 5365-5372.

56. SØRENSEN, I., PETTOLINO, F. A., BACIC, A., RALPH, J. LU, F., O'NEILL, M. A. FEI, Z., ROSE, J. K. C., DOMOZYCH, D. S. & WILLATS, W. G. T. (2011): The charophycean green algae provide insights into the early origins of plant cell walls. Plant. J. 68, 201-211.

57. TEELING, H., FUCHS, B. M., BECHER, D., KLOCKOW, C., GARDEBRECHT, A., BENNKE, C. M., KASSABGY, M., HUANG, S. X., MANN, A. J., WALDMANN, J., WEBER, M., KLINDWORTH, A., OTTO, A., LANGE, J., BERNHARDT, J., REINSCH, C., HECKER, M., PEPLIES, J., BOCKELMANN, F. D., CALLIES, U., GERDTS, G., WICHELS, A., WILTSHIRE, K.H., GLOCKNER, F. O., SCHWEDER, T., AMANN, R. (2012) Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* **336**, 608-611.

58. THOMAS, F., BARBEYRON, T., TONON, T., GÉNICOT, S., CZJZEK, M. & MICHEL, G. (2011): Characterization of the first alginolytic operons in a marine bacterium: from their emergence in marine *Flavobacteriia* to their independent transfers to marine *Proteobacteria* and human gut *Bacteroides. Environ. Microbiol.* **14,** 2379-2394.

59. WALDMANN J., TEELING H., HUANG S., MANN A., WERNER J., WEBER M., GLÖCKNER, F. O. (2014): Taxometer: an approach for robust taxonomic classification of metagenomic fragments. [in prep.]

60. WILLIAMS, T. J., WILKINS, D., LONG, E., EVANS, F., DEMAERE, M. Z., RAFTERY, M. J. & CAVICCHIOLI, R. (2013): The role of planktonic *Flavobacteria* in processing algal organic matter in coastal East Antarctica revealed using metagenomics and metaproteomics. *Environ. Microbiol.* **15**, 1302-1317.

61. WUSTMAN, B. A., LIND, J., WETHERBEE, R. & GRETZ, M. R. (1998): Extracellular matrix assembly in diatoms (Bacillariophyceae) III. Organization of Fucoglucuronogalactans within the Adhesive Stalks of *Achnanthes longipes*. Plant. Physiol. 116, 1431-1441.

Table 1: Depicted are the oceanic provinces including their coordinates, date and time (UTC) samples were taken, the physicochemical parameters and *Bacteroidetes* relative abundance (GÓMEZ-PEREIRA ET AL. 2010) of the stations from which the metagenomic libraries were constructed. Filtered volumes, library size and number of screened and sequenced clones taken from GÓMEZ-PEREIRA ET AL. (2012).

	S3	**S18**
Oceanic Province	BPLR	NAST
Latitude	65°52.64'N	34°04.43'N
Longitude	29°56.54'W	30°00.09'W
Date	22.09.2006	30.09.2006
Time [UTC]	7 pm	8 am
Temperature [°C]	0.7	23.5
Salinity [PSU]	33	36.4
Dissolved oxygen [ml l^{-1}]	7.7	4.7
Phosphate [µM]	0.4	0.01
Nitrate [µM]	2.9	0.03
Nitrite [µM]	0.13	0.03
Ammonia [µM]	0.29	0.26
Silicate [µM]	2.9	0.28
Chlorophyll *a* [mg m^{-3}]	0.7	<0.1
Bacteroidetes [%]	19	7
Volume filtered [l]	90	87
Fosmid library size [no. of clones]	35000	50000
Fosmids screened for rrna (CF319/CF967)	15500	27600
rrna-fosmids sequenced	13	15
Fosmids end-sequenced	16938	16255
Predicted bacteroidetal fosmids	240	148
Sequenced predicted bacteroidetal fosmids	131	100

Table 2: Depicted are the gene frequencies [number of ORFs per Mbp] of genes involved in organic matter degradation. ABC: ABC-transporter, TBDR: TonB-dependent receptor, SusD: starch utilization system, CAZymes: carbohydrate active enzymes, GH: glycosyl hydrolases, CBM: carbohydrate binding module, CE: carbohydrate esterase, PL: polysaccharide lyase and GT: glycosyl transferase. Sum of GH, CBM, CE, PL and GT equals number of CAZymes.

	S3	S18
ABC	4.4	4.2
TBDR	4.2	4.2
SusD	2.0	2.2
CAZymes	41.8	41.8
GH	15.5	15.9
CBM	2.9	4.7
CE	7.4	6.2
PL	0.2	0.5
GT	15.8	14.4
Sulfatases	7.7	3.5
Peptidases	50.3	59.7

Figure 1: Site specific overview of the taxonomic affiliation on genus level of the fosmid-based metagenome libraries. Results obtained with 'taxometer' (WALDMANN ET AL. in prep.).

Figure 2: Site specific comparison of *Bacteroidetes* metagenomes regarding the gene frequency of specific functional group of genes (number of genes per Mbp). (A) comparison of ABC-transporters, TonB-dependent receptors (TBDR), SusD genes and sulfatases; (B) comparison of cazymes: glycosyl hydrolases (GH), carbohydrate binding modules (CBM), carbohydrate esterases (CE), polysaccharide lyases (PL) and glycosyl transferases (GT); (C) detailed comparison of GH families and (D) of peptidases.

Figure 3: Genomic content of fosmids containing PULs from station 3. Name of fosmid and ist taxonomic affiliation are given and number in brackets represents fosmid length. PUL genes or associated PUL genes are coloured according to legend.

Figure 4: Genomic content of fosmids containing PULs from station 18. Name of fosmid and ist taxonomic affiliation are given and number in brackets represents fosmid length. PUL genes or associated PUL genes are coloured according to legend.

Figure 5: Comparison with the genomic alignment tool MAUVE (DARLING ET AL. 2013) of fosmids containing PULs from station 3 and 18. When an alignment has been computed with Progressive Mauve algorithm, regions conserved among the fosmids were identified as locally collinear blocks (LCB) and consist of similar block outline colors and homologous regions within the LCBs are colored the same. Each of these block outlines surrounds a region of the fosmid sequence which aligned to a part of another fosmid sequence. Regions outside the LCBs lack detectable homology among fosmid input. Within each LCB a similarity profile is drawn and the height of the profile corresponds to the average level of conservation in the distinct region. Areas that are completely white within the LCB were not aligned and probably contain sequence elements specific to the particular fosmid. Fosmid

annotations are shown below the LCBs. Graphs A-D show conserved regions between (A) *Flavobacterium* fosmid S3-21YE11 and *Leeuwenhoekiella* fosmid S3-27YE12_C102 from station 3; (B) *Psychroflexus* fosmid S3-C42 and *Cellulophaga* fosmid S3-C99 also from station 3; (C) *Polaribacter* fosmid S3-C51_15YA06 from station 3 and the *Flavobacterium* fosmid S18-34YL06 from station 18 and (D) between four fosmids, two from each station, the *Flavobacterium* fosmid S18-C30 and the *Robiginitalea* fosmid S18-C31 as well th *Flavobacterium* fosmid S3-C45 and *Gramella* fosmid S3-C33. Homologous regions are conected by lines consisting of the same color as the region.

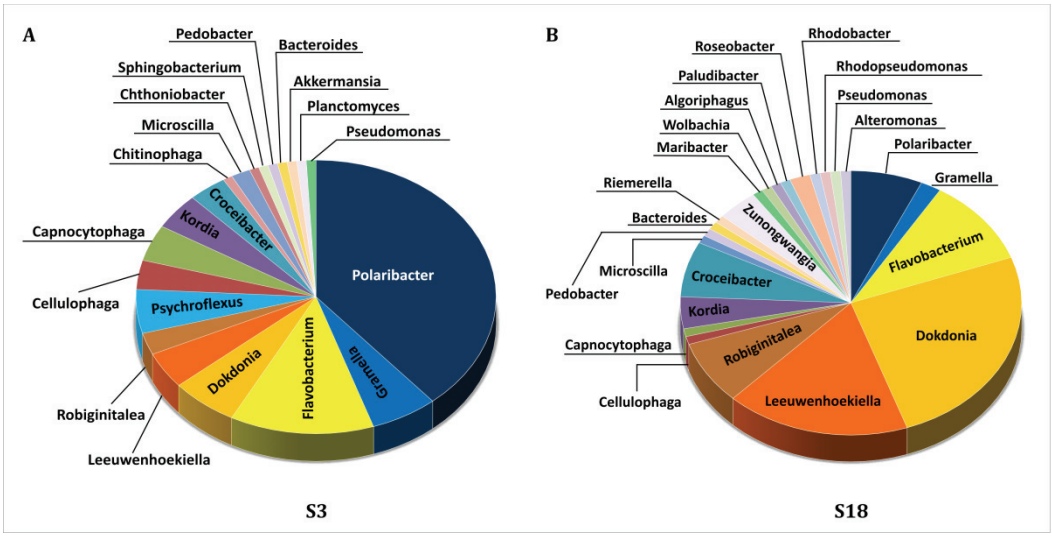


Figure 1

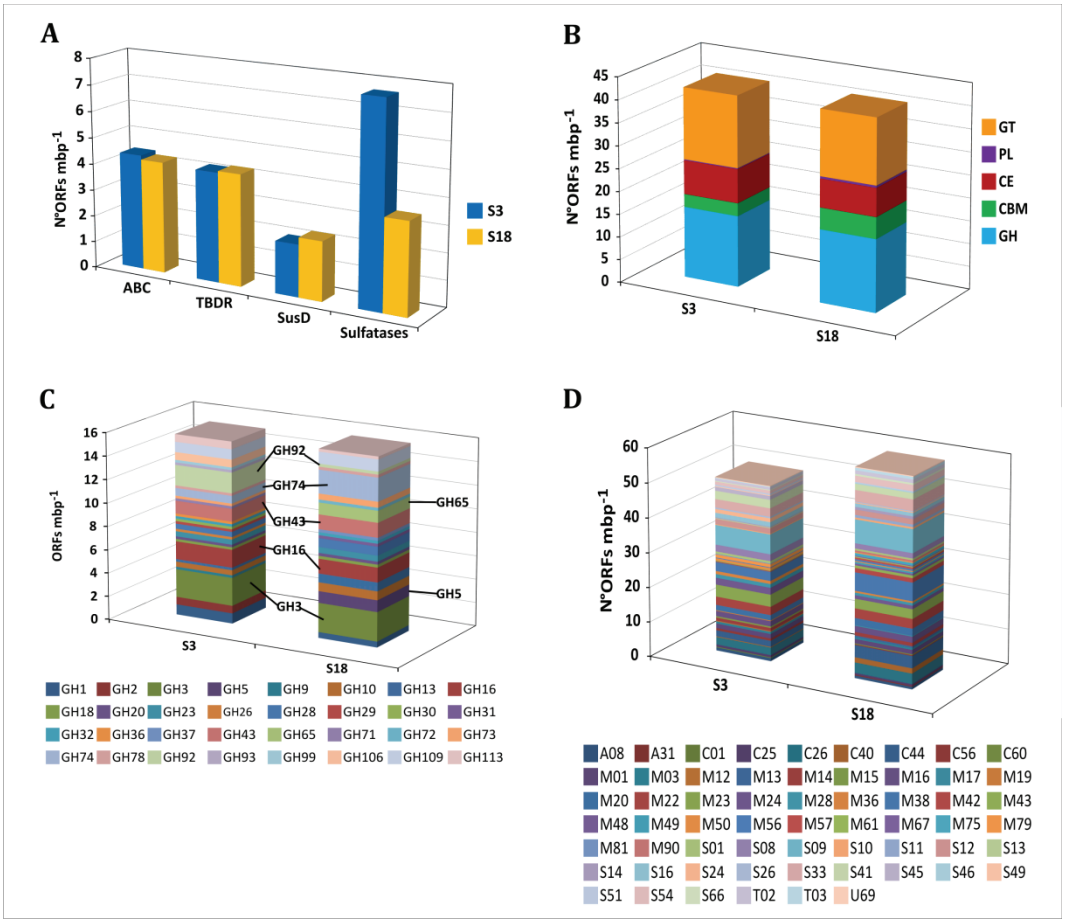


Figure 2

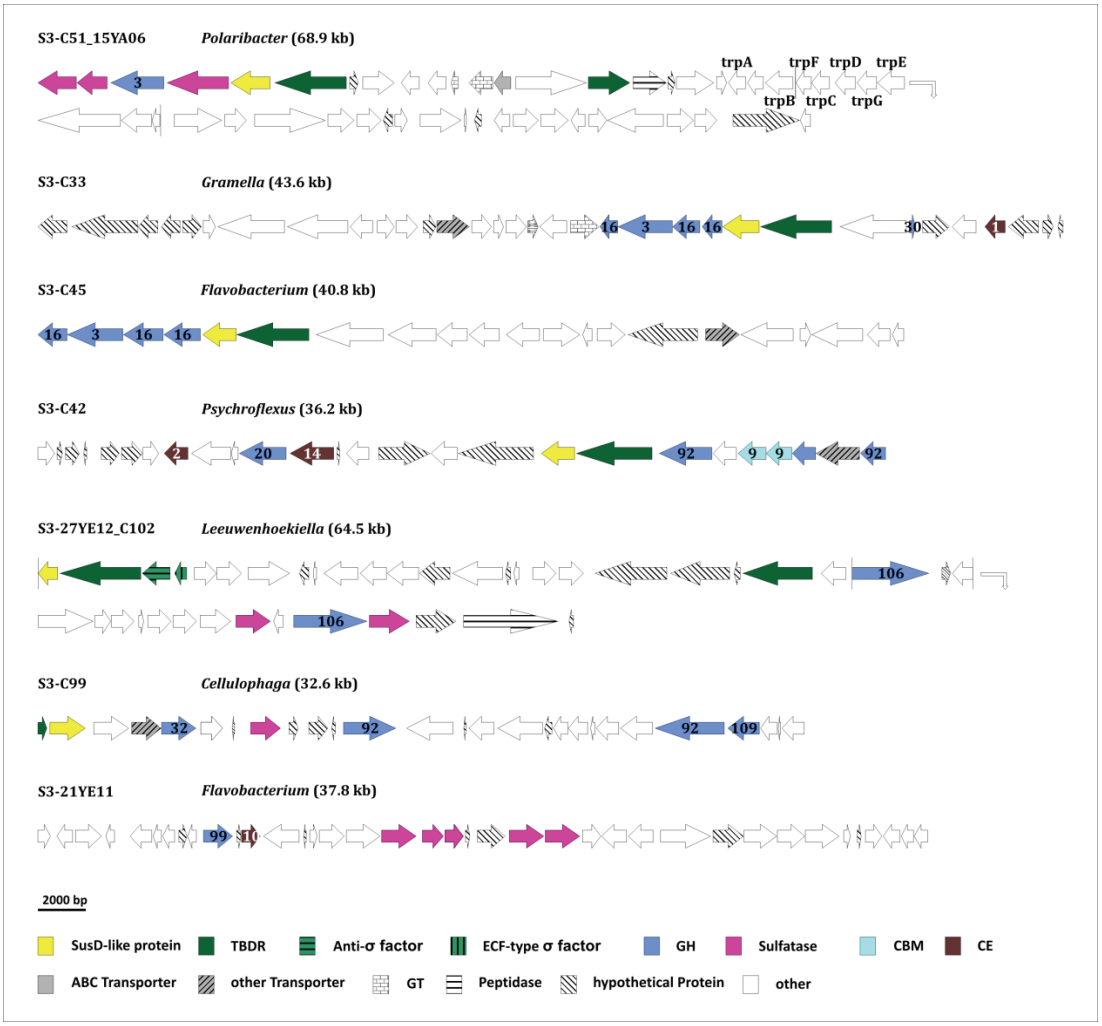

S3-C51_15YA06 Polaribacter (68.9 kb)
trpA trpF trpD trpE
trpB trpC trpG
S3-C33 Gramella (43.6 kb)
S3-C45 Flavobacterium (40.8 kb)
S3-C42 Psychroflexus (36.2 kb)
S3-27YE12_C102 Leeuwenhoekiella (64.5 kb)
S3-C99 Cellulophaga (32.6 kb)
S3-21YE11 Flavobacterium (37.8 kb)
2000 bp
SusD-like protein
TBDR
Anti-σ factor
ECF-type σ factor
GH
Sulfatase
CBM
CE
ABC Transporter
other Transporter
GT
Peptidase
hypothetical Protein
other

Figure 3

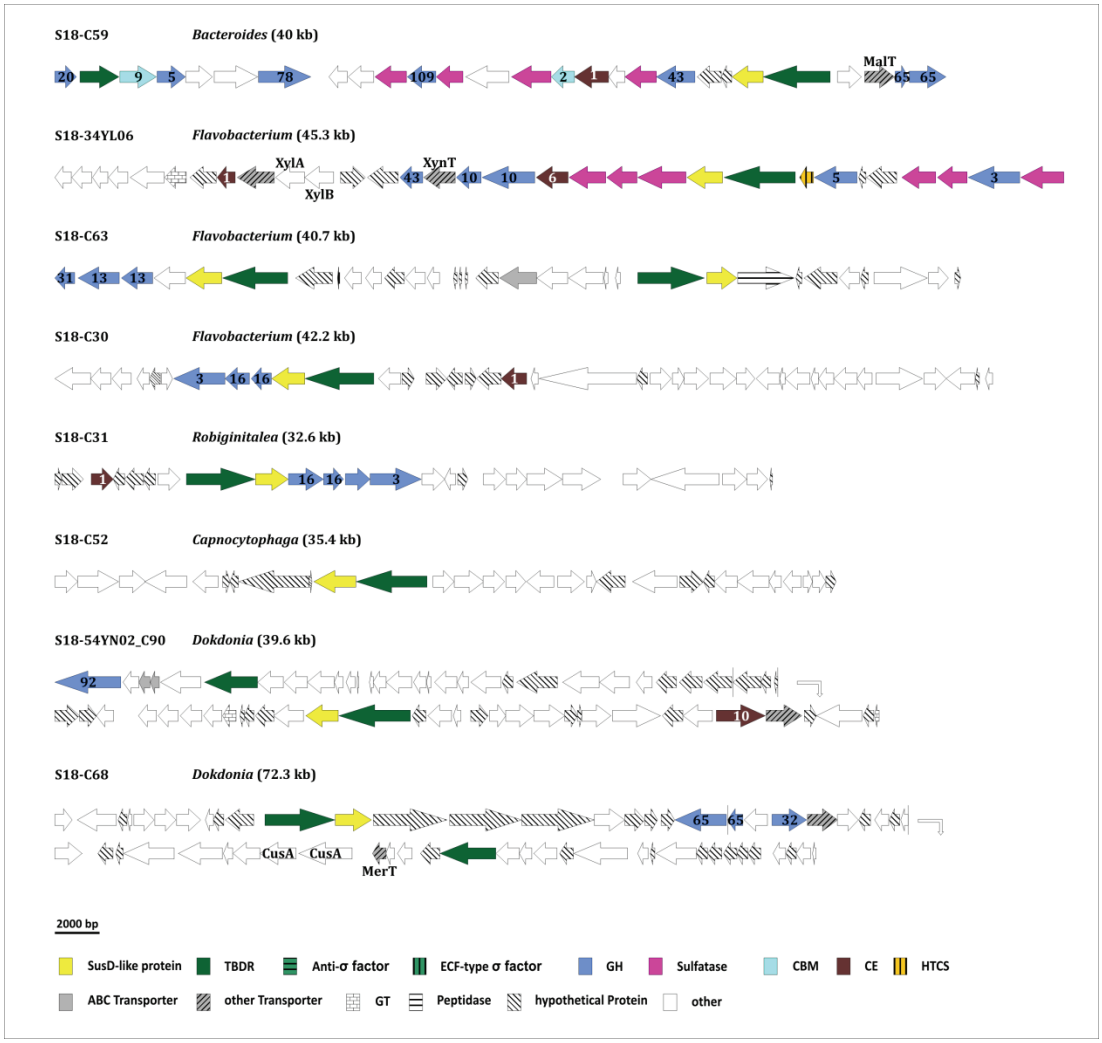

S18-C59 Bacteroides (40 kb)
20 9 5 78 109 2 43 MalT 65 65
S18-34YL06 Flavobacterium (45.3 kb)
XylA XylB XynT 43 10 10 6 5 3
S18-C63 Flavobacterium (40.7 kb)
31 13 13
S18-C30 Flavobacterium (42.2 kb)
3 16 16
S18-C31 Robiginitalea (32.6 kb)
16 16 3
S18-C52 Capnocytophaga (35.4 kb)
S18-54YN02_C90 Dokdonia (39.6 kb)
92
10
S18-C68 Dokdonia (72.3 kb)
65 65 32
CusA CusA MerT
2000 bp
SusD-like protein
TBDR
Anti-σ factor
ECF-type σ factor
GH
Sulfatase
CBM
CE
HTCS
ABC Transporter
other Transporter
GT
Peptidase
hypothetical Protein
other

Figure 4

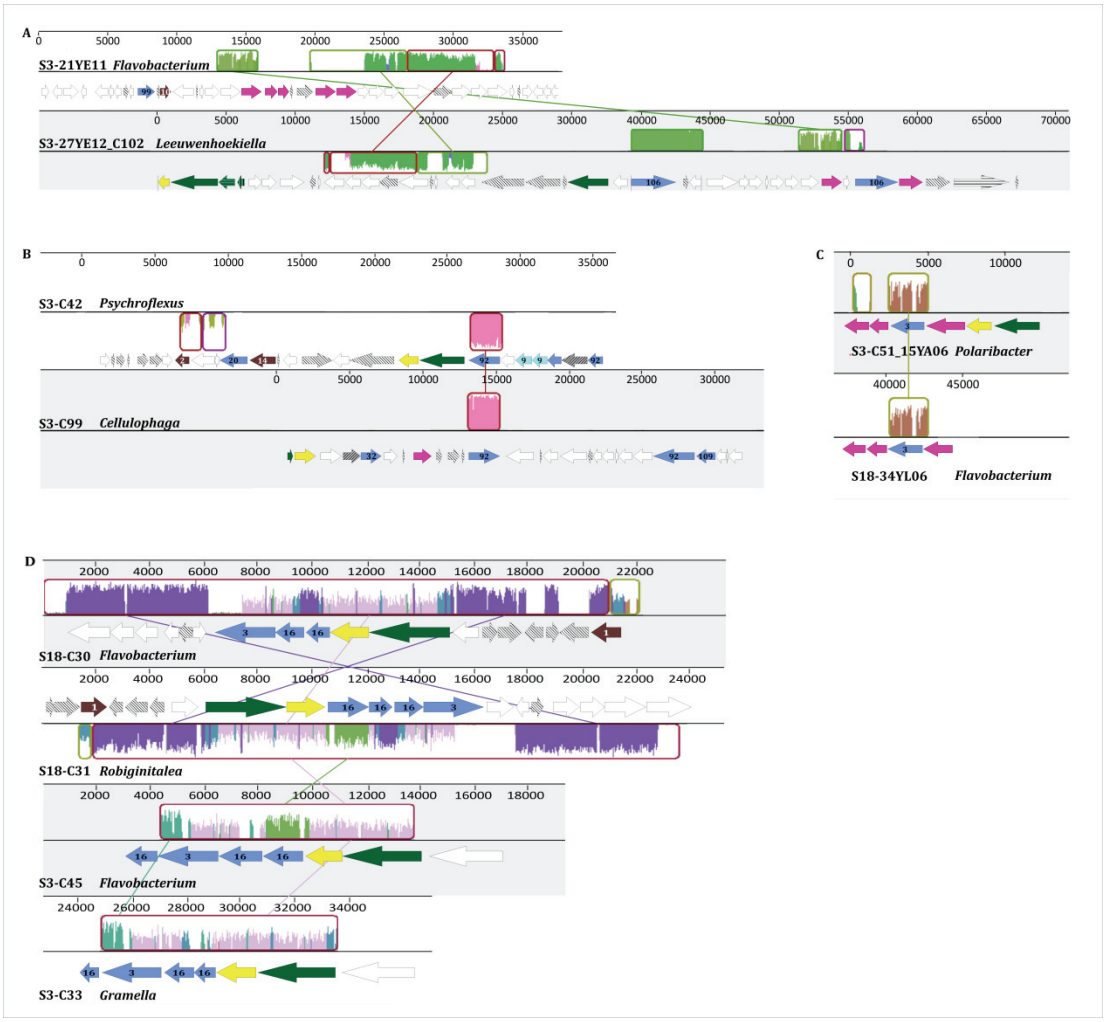

A
S3-21YE11 Flavobacterium
S3-27YE12_C102 Leeuwenhoekiella
B
S3-C42 Psychroflexus
S3-C99 Cellulophaga
C
S3-C51_15YA06 Polaribacter
S18-34YL06 Flavobacterium
D
S18-C30 Flavobacterium
S18-C31 Robiginitalea
S3-C45 Flavobacterium
S3-C33 Gramella

Figure 5

Supplementary material

Characterisation of Polysaccharide Utilization Loci (PULs) in marine *Bacteroidetes*

Christin M. Bennke, Sixing Huang, Angélique Gobet, Margarete Schüler, Valerie Barbe, Hanno Teeling, Gurvan Michel, Bernhard M. Fuchs & Rudolf Amann

Figure S1: Overview of the VISION cruise track (MSM03/01) in the North Atlantic Ocean from 21. - 30. September 2006. Map displays the investigation area from the Artic Circle towards the Azores Islands including their respective biogeographical provinces defined after Longhurst (1998): the Boreal Polar Province (BPLR), Arctic Province (ARCT), The North Atlantic Drift Province (NADR) and the North Atlantic Subtropical Province (NAST) which are indicated by the dotted lines.

A: monthly average of the sea surface temperature

B: monthly average of the chlorophyll a concentration [provided by the quaModis satellite (http://oceancolor.gsfc.nasa.gov; requested May 7th 2010)]. Missing data shown as white.

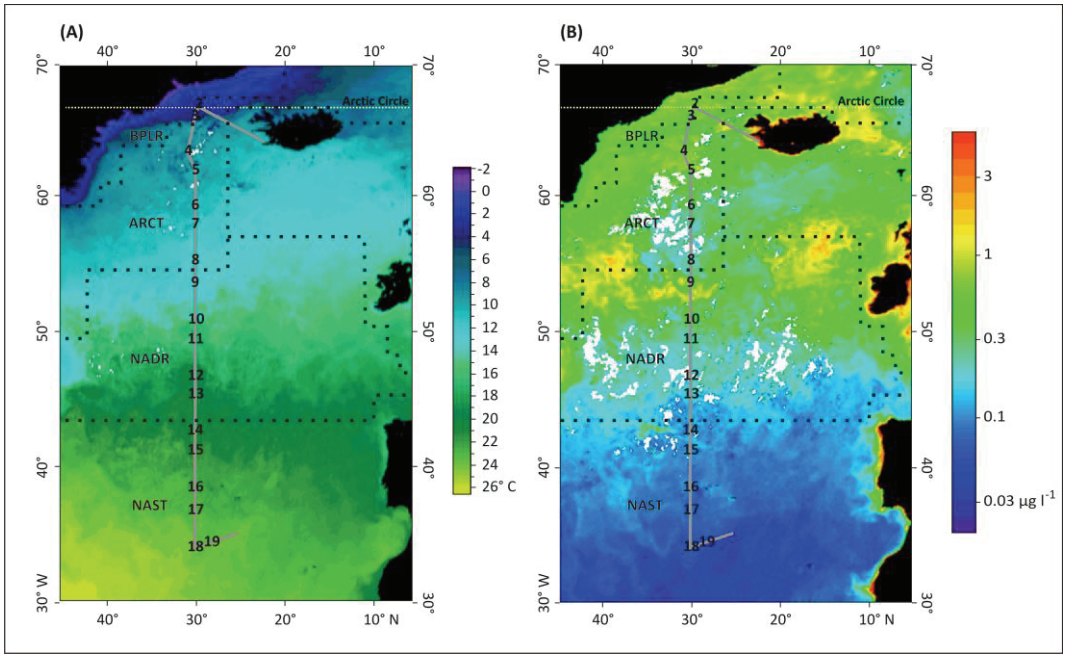

(A)
(B)
Arctic Circle
BPLR
ARCT
NADR
NAST
40°
30°
20°
10°
10° N
70°
60°
50°
40°
30° W
-2
0
2
4
6
8
10
12
14
16
18
20
22
24
26° C
3
1
0.3
0.1
0.03 µg l-1

Figure S1

Chapter 7: Discussion and Outlook

7.1 General Discussion

This doctoral thesis aims at improving the understanding of the ecology of *Bacteroidetes* in marine environments. Therefore, I investigated their diversity, abundance and composition at the coastal site Helgoland Roads and their genetic capability for organic matter degradation at two contrasting sites in the North Atlantic Ocean.

The fundamental questions of ecology "who?", "how many?" and "what are they doing?" mark essential steps in the characterization of ecosystems. This is not only valid in macrobiology, such as botany or zoology, but also applies to microbiology. Initially, total organismal abundance is determined which is usually done in microbiology by staining cells with nucleic acid specific fluorescent dyes like DAPI [4',6-diamidino-2-phenylindole] (PORTER & FEIG 1980) or acridine orange (HOBBIE ET AL. 1977) and subsequent counting by flow cytometry (ZUBKOV ET AL. 1998) or fluorescence microscopy (BLOEM 1995). This seems to be a trivial task since it is routinely done in many laboratories. However, manual microscopic quantification is very time consuming and always relies on the accuracy of the person counting. In literature there are many suggestions of how many cells have to be counted in order to obtain reliable results. For example, in the 1970s scientists counted only 200 to 400 cells (e.g. JONES & SIMON 1975), which changed in the

1980s, where PORTER & FEIG (1980) reported that at least 1000 DAPI stained cells have to be counted to gain statistically sound results. This rule is still valid. Today computer-based cell enumeration becomes more and more important, since it is well documented, less time consuming and independent of human errors. Accurate, precise and reproducible determination of bacterial cell numbers in aquatic samples is of high importance to understand the ecological role of bacteria in their environment (PERNTHALER & AMANN 2005, KEPNER & PRATT 1994).

7.1.1 High-throughput cell enumeration

The first aim of this study was, therefore, to test a newly developed automatic cell counting system (based on ZEDER & PERNTHALER 2009), and apply it to the enumeration of total, bacterial and specifically *Bacteroidetes* abundances during four spring blooms and over two entire years. In the evaluation of counting routines changes had to be implemented which also resulted in an update of the software. The compatibility of these updates was tested to ensure that the results obtained with the different software versions are reliable (Figure 7.1). The current version contains functions to easier and more precisely defined detection parameters, additional marker tools for manual inspection and alternative segmentation modes, which were implemented to obtain cell numbers of e.g. nanoflagellates (THIELE 2013) and bacteria colonizing aggregates (BIŽIĆ –IONESCU ET AL. 2014, in rev.).

Initially it was thought that the quality check of images taken by the automated microscope stage could be done automatically as well, but most of the samples retrieved from Helgoland contained aggregates, debris and phytoplankton cells. This made it difficult for the image quality algorithm to select images of good quality. Most of the images containing uneven or out-of-focus parts were correctly discarded, but images containing over- or underexposed areas, due to aggregates or phytoplankton cells still remained and manual de-selection of cells marked for counting was needed. Therefore, manual inspection was done directly after the image

acquisition, which gave a better reflection of the sample, e.g. where cells are located, if many not probe-specific signals, like particles or picoeukaryotes were present. As a standard procedure for automatic counting, two fluorescent channels for image acquisition were used, one for DAPI and one for FISH to obtain community specific counts. However, images are taken in grey-scale and cells or objects possessing autofluorescences are possibly captured by the illumination and filter used for the FISH channel. Consequently, signals derived from autofluorescence e.g. cyanobacteria, particles, debris or picoeukaryotes, cannot be distinguished from the FISH positive probe signal. Therefore, a third channel, detecting autofluorescence was defined and subsequently probe-specific and autofluorescent signals could be reliably discriminated, resulting in more reliable counts.

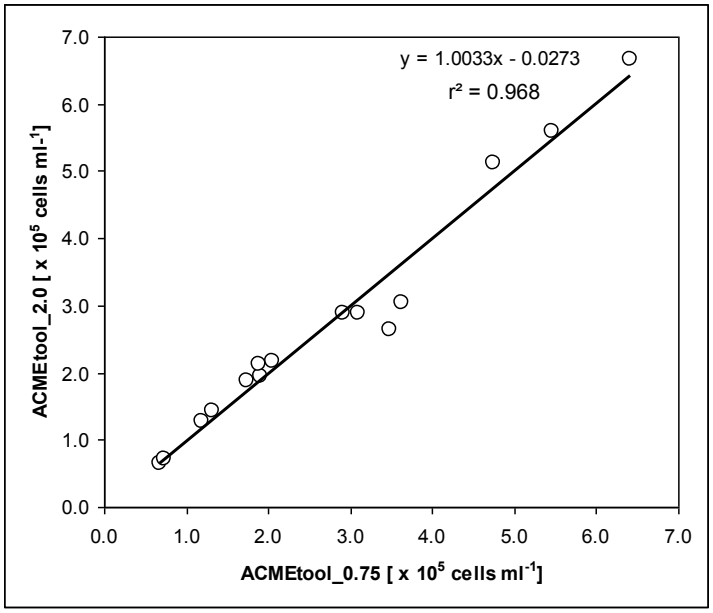


Figure 7.1: Regression analysis of the dataset analyzed with both available ACMEtool versions. Student´s two tailed test revealed a p-value of 0.827 excluding the possibility that false detection parameters were chosen.

7.1.2 Why studying coastal spring blooms?

The second part of my thesis concentrated on the "who is out there?" question. My task was to analyze a pleothora of surface water samples collected during spring phytoplankton blooms at the long term ecological research station Helgoland Roads specifically for the diversity and abundance of *Bacteroidetes clades*.

There are mainly three reasons, why spring phytoplankton blooms are a good example to study the bacterial response to phytoplankton blooms. First of all, the onset of the spring bloom is triggered by increasing temperature and solar radiance, and availability of nutrients (Wiltshire et al. 2008, TIAN ET AL. 2011). Due to winter mineralization and water mixing events, organic and inorganic nutrients are not limiting (MAHADEVAN ET AL. 2012). Secondly, the grazer community is not developed yet and consequently the mortality among phytoplankton is low, resulting in a mainly bottom-up regulated ecosystem (WILTSHIRE ET AL. 2009). The zooplankton predators can only proliferate when the phytoplankton is already abundant (e.g. STORM ET AL. 2001, CALBET & LANDRY 2004). Usually, the phytoplankton bloom is terminated not only by grazing, but more often by nutrient depletion in silicate, iron or phosphorus, and viral lysis (SMETACK 1999, BEARE ET AL. 2002, WILTSHIRE ET AL. 2008). This leads to a collapse of the algal population and as a result a massive amount of algal organic matter, mainly in form of polysaccharides, is released and mineralized by heterotrophic bacteria (ALDERKAMP ET AL. 2007). These less complex conditions are special for spring blooms, which make them easier to study than summer and autumn blooms. In these later blooms, grazer communities might already have developed, ready to attack growing phyto- and bacterioplankton (e.g. PERNTHALER 2005, ANDERSON ET AL. 2012). An example for the complexity of summer phytoplankton blooms was revealed by the monitoring of the LOHAFEX iron fertilization experiment. Iron is a limiting factor for phytoplankton growth in high nutrient-low chlorophyll waters typical of southern oceans (Martin 1990) and its effect was studied in many iron fertilization experiments

(e.g. HALL & SAFI 2001, OLIVER ET AL. 2004, THIELE ET AL. 2012). During the LOHAFEX experiment both the phytoplankton and the bacterioplankton community showed only a slight response to the fertilization which was most likely due the presence of a well-established grazer community, mainly consisting of nanoflagellates (Thiele et al. 2012). In other studies it was shown that fast growing bacteria like *Bacteroidetes, Roseobacter* or *Gammaproteobacteria* are strongly affected by grazing (e.g. PERNTHALER 2005).

7.1.3 From PULs to blooms

The third part of this thesis addressed the ecological question: "What is their function?". Here I analyzed fosmid-based metagenomes regarding their genetic capability for polysaccharide binding, hydrolysis and uptake. From previous studies it is known that cultured *Bacteroidetes* are capable of degrading high molecular weight organic matter, such as cellulose, chitin and proteins (see review by KIRCHMAN 2002). Recently, it has been shown that *Bacteroidetes* are involved in degradation of polysaccharide such as laminarin and xylan at different sites of the North Atlantic Ocean (ARNOSTI ET AL. 2012).

The genes found in marine *Bacteroidetes* are homologous to those used by the gut symbiont *B. thetaiotaomicron* for polysaccharide degradation (SHIPMAN ET AL. (2000) which are organized in operon-like structures and named polysaccharide utilization loci (PUL; BIJURSEL ET AL. 2006). Those loci were identified as being present in many other *Bacteroidetes* strains (BAUER ET AL. 2006, MCBRIDE ET AL. 2009, THOMAS ET AL. 2011, MANN ET AL. 2013) and metagenomes (GÓMEZ-PEREIRA ET AL. 2012, WILLIAMS ET AL. 2013). For example the well-studied coastal *Bacteroidetes* strain '*Gramella forsetii*' KT0803 (BAUER ET AL. 2006) possesses a high gene frequency of glycosid hydrolases (GH) and other carbohydrate-active enzymes (CAZymes) as well as high numbers of peptidases and other protein degradation genes (BAUER ET AL. 2006), resulting in a peptidase/GH ratio of 2.2. This ratio points out, which

degradation type, either protein degradation (high ratio) or polysaccharide decomposition (low ratio), is preferred (KOLTON ET AL. 2013). For example, the peptidase/GH ratio of *B. thetaiotaomicron* is 0.5, mirroring that this gut symbiont encodes only half of the numbers of peptidases than of GHs, thus revealing a lifestyle, which is tailored to polysaccharide decomposition. In contrast many open ocean strains, e.g. *Croceibacter atlanticus* has a high peptidase/GH ratio of 6.8 and possesses more than 30 ORFs per Mbp peptidases and other genes related to protein degradation. *Croceibacter atlanticus* was isolated from the Sargasso Sea (Cho & Giovannoni) which comprised very oligotrophic conditions, reflecting that *C. atlanticus* is very adapted to this type of environment. There is no need in possessing many GHs or other genes related to polysaccharide degradation when there are barely some around. In contrast to *B. thetaiotaomicron,* who lives in an environment surrounded by polysaccharides, has a higher need in encoding GH and other polysaccharide degradation genes like SusD and TBDRs. Comparing the peptidase/GH ratio of the North Atlantic Ocean fosmids with other open ocean, coastal and soil flavobacterial strains (Figure 7.2), it became evident that the fosmids fell between the coastal and open ocean strains. This indicates a higher capability of polysaccharide degradation by open ocean *Bacteroidetes* clades than previously anticipated from single open ocean genomes, such as *Polaribacter irgensii* with a ratio of 5.2, *Croceibacter atlanticus* with 6.8 or *Robiginitalea biformata* with 4.3 (Figure 7.2). Those strains are more tailored to protein degradation.

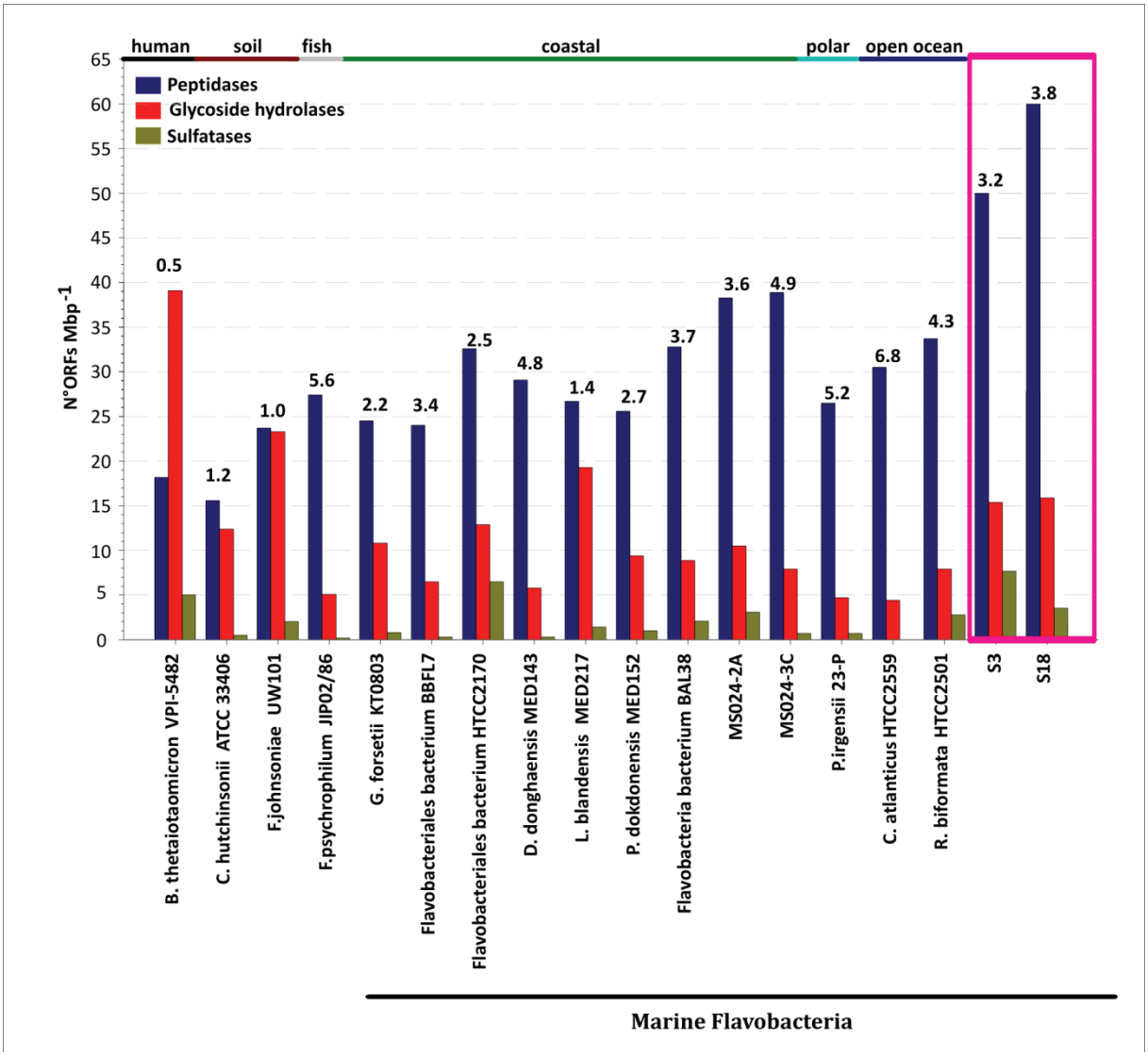


Figure 7.2: Number of glycosyl hydrolases (GH), peptidases and sulfatases per Mbp in Bacteroidetes genomes from different habitats. Based on Pfam profiles searches and CAZy family members (modified after GÓMEZ-PEREIRA ET AL. 2012). Numbers on top of the bars represent peptidase/GH ratio according to (FERNÁNDEZ-GÓMEZ ET AL. 2013, KOLTON ET AL. 2013).

Knowing that *Bacteroidetes* – independent of their origin (coastal, open ocean or soil) – possess those genetic islands – the so called PULs – explains why they respond first to phytoplankton spring blooms, profiting from the decaying phytoplankton. They are equipped with specific sets of genes, involved in carbohydrate binding, hydrolysis and uptake, as well as involved in protein degradation and cell attachment, allowing them to directly access the available substrates during phytoplankton senescence. Most of the *Bacteroidetes* clades are able to degrade more than one polysaccharide, since they possess not only one PUL with single substrate specificity, but several (e.g. BAUER ET AL 2006, WOYKE ET AL. 2009, MANN ET AL. 2013). This is also supported by culture studies, where the growth media of distinct flavobacterial strains were supplied with different carbon sources. Most of the tested strains were able to utilize several polysaccharides, such as laminarin, cellulose, xylan and carrageenan (HAHNKE 2013, P. 197-198). The potential of specific strains degrading biopolymers is most often reflected in their genome size, encoding for different lifestyles (marine vs. terrestrial, planktonic or sedimental, free-living or attached). For example, the genomes of terrestrial strains are mostly larger than those from aquatic origins. The soil inhabiting *Flavobacterium johnsoniae* has a genome size of 6.1 Mbp and encodes for 138 GHs, 9 polysaccharide lyases and 17 carbohydrate esterases, suggesting a metabolism related to carbohydrate degradation (MCBRIDE ET AL. 2009). This was also shown for other terrestrial flavobacterial strains (KOLTON ET AL. 2013), and was recently identified for marine *Flavobacteria*, such as *Formosa agariphila* (MANN ET AL. 2013) or some *Polaribacter* species (TEELING unpublished).

Although, the genome size of coastal *Bacteroidetes* are generally smaller than those of terrestrial origin, coastal *Bacteroidetes* possess also a high number of GHs, TonB-dependent receptors, SusD domains and other genes involved in polysaccharide degradation enabling them to access directly the phytoplankton derived substrates (see chapter 3). This explains their massive increase during

phytoplankton blooms, but their abundances declined almost immediately after the algal derived substrates were utilized by specific bacteroidetal clades, such as *Ulvibacter, Formosa* A and B, *Polaribacter*, as well as *Owenweekisia*. It was observed that some clades like *Polaribacter* and *Formosa* A reoccurred not only later in spring, but also during the summer and autumn phytoplankton blooms (see chapter 4). Either the respective substrates were re-appearing since these blooms also comprised of diatoms or these bacteroidetal clades possess broader substrate specificity, utilizing alternative polysaccharides. However, the presence of subclades might also explain their multimodal distribution.

We identified, for example, two different *Polaribacter* OTUs in spring 2009 and 2010, sharing less than 98% 16S rRNA identity. This indicated the presence of two different *Polaribacter* subclades in the two consecutive years. Since both subclades were detected by the same *Polaribacter*-specific oligonucleotide probe POL740 (MALMSTROM ET AL. 2007) subsequent peaks could be due to different *Polaribacter* populations. Consequently, the presence of two different *Polaribacter* subclades in both spring seasons could possibly explain why this clade appeared as one of the last clades in spring 2009 and one of the first clades in spring 2010. Most likely both species possess a different genetic potential in assessing and degrading organic matter. For example, the metagenome of April 7th 2009 (TEELING ET AL. 2009) revealed a low sulfatase gene level of 0.2% for *Polaribacter* taxobins (HUANG 2013). It could be speculated that the *Polaribacter* clade present in spring 2009 was barely able to degrade sulfated extracellular polymers from the diatom cells. In contrast, the sulfatase gene level was high in *Formosa* taxobins (1%), which would explain why *Formosa* A appears before *Polaribacter* in spring 2009, since they were responsible for the initial breakdown of the diatom cells in removing the sulfated extracellular polymers. For spring 2010 and subsequently 2011 and 2012, it could be possible that the present *Polaribacter* species might be enriched in sulfatases accessing directly the sulfated extracellular polymers, bypassing *Formosa* A. This could explain why

Polaribacter bloomed ahead of *Formosa* A. This, of course, is highly speculative and awaits further metagenome and metaproteom investigations.

Despite the fact, that *Bacteroidetes* possess many PULs with different substrate specificities and gene arrangements, the coupling of the SusD protein to the TonB-dependent receptor (TBDR) at the outer membrane is conserved and seems to be a unique feature of *Bacteroidetes* PULs (THOMAS ET AL. 2011). Other bacterial groups like the gammaproteobacterial clade SAR92 possess also PULs which lack, however, the SusD gene (KAPPELMANN 2013). Either, SusD is involved in substrate binding unique for *Bacteroidetes*, making them highly competitive against others, or SusD is replaced in SAR92 members by a yet unknown protein.

7.1.4 How microscopy can help visualizing ecological niches

Microscopic observations of the large 10 µm fraction revealed the attachment of *Bacteroidetes* to the diatom *Chaetoceros* spp. at the maximum of the phytoplankton bloom in spring 2011 (Figure 7.3). A protocol was developed in which cell identification by CARD-FISH was combined with glycoconjugate recognition by fluorescent lectin binding analysis (FLBA), in order to study *Bacteroidetes*-diatom associations. The combination of CARD-FISH and FLBA provided new insights into the molecular basis of the attachment of specific *Bacteroidetes* clades to planktonic diatoms. It has great potential in understanding habitats *in situ* and possible substrate preferences of attached living bacteria. Therefore, this technique was used to follow the *Bacteroidetes* attachment to *Chaetoceros* spp. in the time frame of three weeks during the spring phytoplankton bloom 2011 (BAKENHUS 2014).

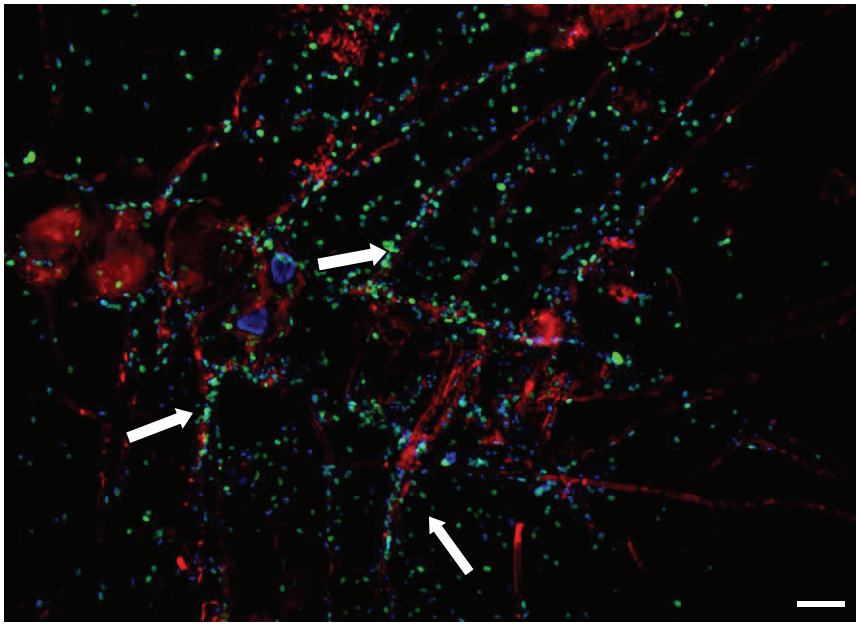


Figure 7.3: SR-SIM micrograph shows the attachment of *Bacteroidetes* (green), detected with probe CF319a (MANZ ET AL. 1996), known to target most marine Bacteroidetes (AMANN & FUCHS 2008) to the diatom *Chaetoceros* spp. and especially to the setae (arrows). Scale bar is 10 µm. green: CF319a-Alexa488; red: AAL-Alexa568, blue: DAPI.

Members of the *Bacteroidetes* phylum were found attached to some *Chaetoceros* spp. right from the beginning of the analysis, which was one week ahead of the phytoplankton bloom. After the phytoplankton maximum a gradual attachment of SAR92 cells to *Chaetoceros* spp. populated by *Bacteroidetes* was observed (Figure 7.4) (BAKENHUS 2014). These observations would support the hypothesis that the SusD domain is unique among *Bacteroidetes* facilitating the access to complex polymers, leaving them unrivalled and highly competitive. In contrast, SAR92 might not be able to access those complex polymers from the beginning, since they are lacking the SusD domain (KAPPELMANN 2013). Notably, in all four analyzed spring seasons SAR92 (CHEN 2012) always peaked after the first bloom of *Bacteroidetes*, which mainly consisted of *Ulvibacter, Formosa* A and *Polaribacter*.

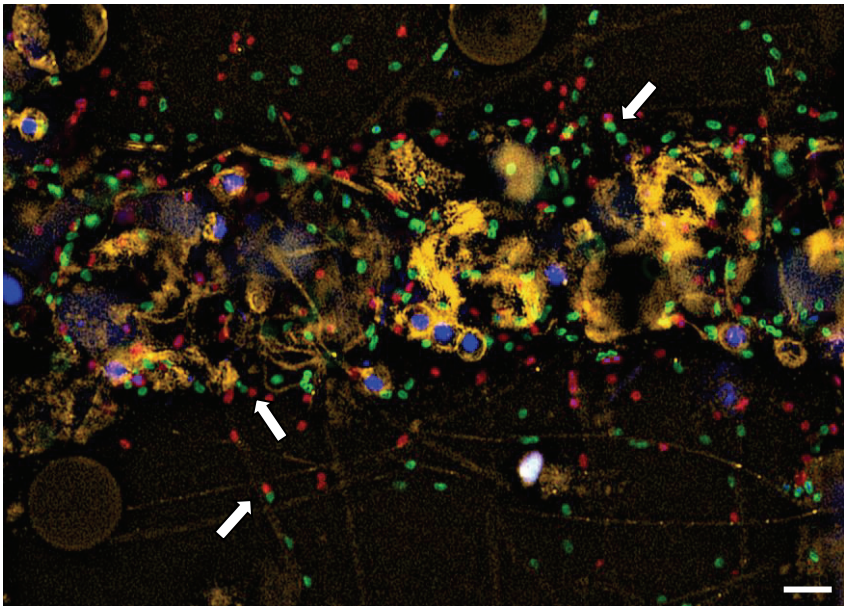


Figure 7.4: SR-SIM micrograph shows the attachment of *Bacteroidetes* (green) and SAR92 (red) to the diatom *Chaetoceros* spp. and especially to the setae. Arrows indicate were close CF319a and SAR92 cells. Scale bar is 10 µm. green: CF319a-Alexa488; red: SAR92-Alexa594, orange: AAL-Alexa561; blue: DAPI.

From our findings it appears that the initial breakdown of those substrates is catalyzed by *Bacteroidetes* and then smaller fractions might be taken up by SAR92. This conclusion is still fairly speculative and requires further investigations. It is further supported by a scanning electron micrograph, where *Bacteroidetes* (gold-labeled) seems to attach to another, yet unknown bacterium (Figure 7.5). In addition, those scanning electron micrographs gave further insights into the specific attachment of *Bacteroidetes* to diatoms and aggregates containing phytoplankton remains (Figure 7.6).

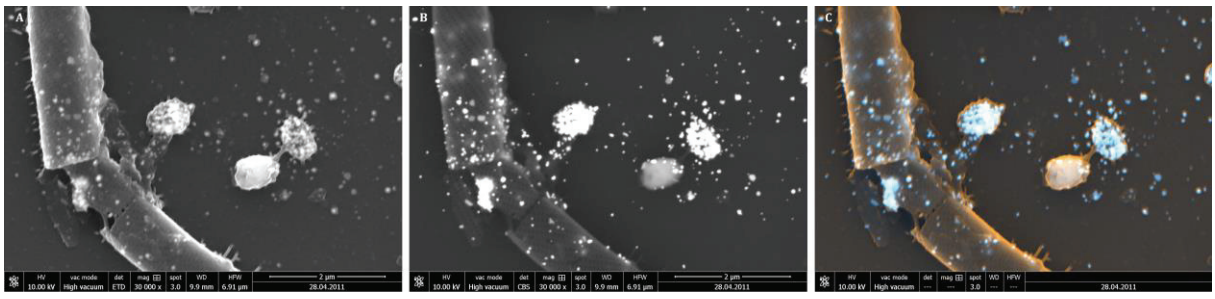


Figure 7.5: SEM micrographs show association of gold labeled Bacteroidetes to phytoplankton remains (setae). A: second electron image (SE, reveals topography), B: backscatter image (BSE, detection of gold label), C: mix of SE and BSE image. Scanning electron microscopy was done with FEI Quanta 250 FEG. Scale bar is 2 µm.

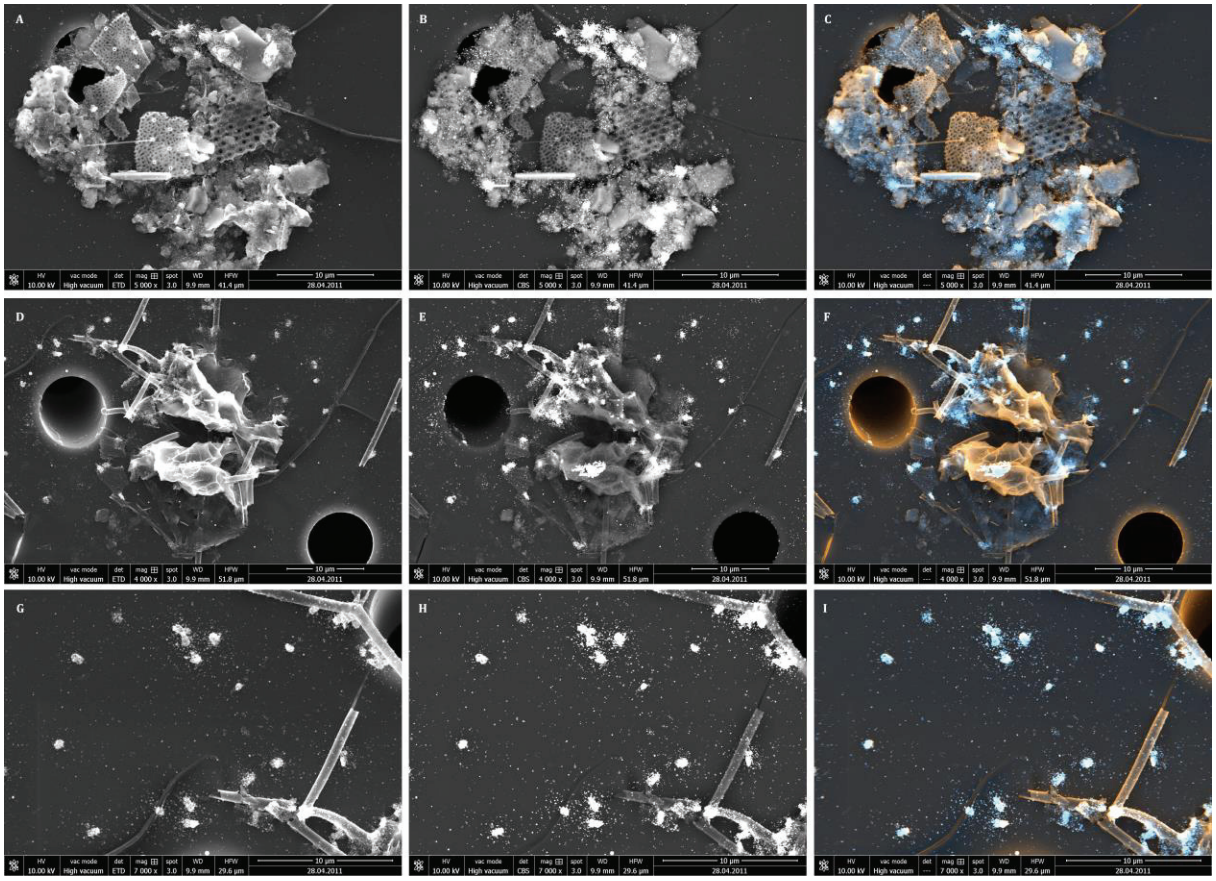


Figure 7.6: SEM micrographs show association of gold labeled *Bacteroidetes* to aggregates (A-C), phytoplankton cells (D-F) and phytoplankton remains (setae, G-I). Images to the left: second electron image (SE, reveals topography), images in the middle: backscatter image (BSE, detection of gold label), images to the right: mix of SE and BSE image. Scanning electron microscopy was done with FEI Quanta 250 FEG. Scale bar is 10 µm.

7.2 Conclusion and Outlook

Bacteroidetes are a very good example for BAAS BECKING'S (1934) statement "*everything is everywhere*, but *the environment selects*". In the marine realm they are found everywhere: from coastal to open ocean, from polar to equatorial latitudes, from surface to abyssal waters as well as in association with aggregates and phytoplankton blooms. It seems that their occurrence is independent of the habitat being sampled, as long as certain food sources such as polysaccharides and proteins are available. *Bacteroidetes* can reach high densities, as shown for the North Sea where they made up more than 50% of the bacterioplankton community (chapter 3 & 4) in four subsequent years in the aftermath of phytoplankton blooms. A research institute studying the involvement of microorganisms in global element cycling should also in the future not neglect such an important group of marine bacteria.

After analyzing the bacteroidetal response to coastal spring diatom blooms in detail, it would be the logical next step to investigate similar open ocean blooms which occur in the higher latitudes. It would be interesting to know, if some of the clades dominating coastal spring diatom blooms at Helgoland would also dominate in oceanic settings. More generally, can such a dynamic succession pattern as found at Helgoland also be demonstrated for open ocean settings? First results of the Atlantic Meridional Transect 21 (AMT21) showed that *Bacteroidetes* made up more than 10% in areas with high chlorophyll *a* concentrations and represented more than 20% in the southern most stations, while a phytoplankton bloom occurred (Figure 7.7). AMT is a multidisciplinary program which undertakes biological, chemical and physical oceanographic research during an annual research cruise between the UK and destinations in the South Atlantic. This program was established in 1995. The AMT21 started on September 29th 2011 in Bristol, UK and ended on October 14th 2011 in Punta Arenas, Chile, passing several biogeographical provinces of the Atlantic Ocean**.** During this cruise 53 stations were sampled to determine the bacterial community. Additionally, a phytoplankton net consisting of four individual

plankton nets with different mash sizes (<180 µm, < 100 µm, <40 µm and < 20 µm) were applied at 31 stations. This was done to identify bacteria which attach to the phytoplankton.

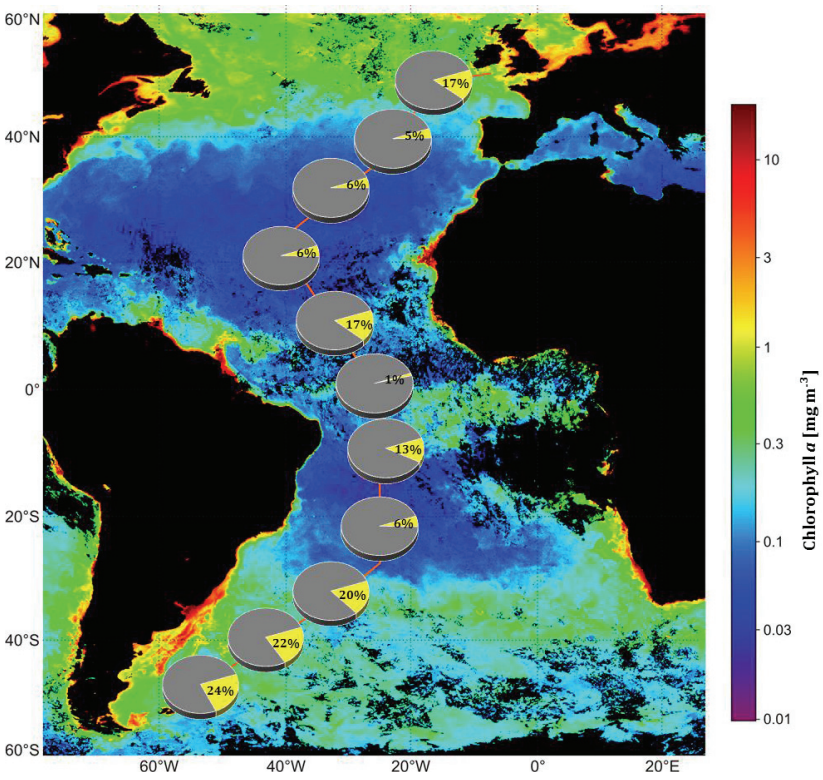


Figure 7.7: Overview of the AMT21 cruise tracke. Pie charts show relative contribution of Bacteroidetes (yellow) to the total bacterioplankton community (grey). Map shows maonthly average of chlorophyll a concentration in September 2011 [provided by AquaModis satelitte (http://oceancolor.gsfc.nasa.gov; requested 18.07.2012).

First analysis of the attached fraction revealed similar results as for the coastal station Helgoland Roads. *Bacteroidetes* were mainly found attached to diatoms, dinoflagellates, aggregates and even to larvae of copepods (Figure 7.8). In future experiments, it would be interesting to know which *Bacteroidetes* clades are involved in this specific attachment. Are they the same members which are commonly found at Helgoland Roads, or do the *Bacteroidetes* community of the South Atlantic Ocean differ from those of the North Sea?

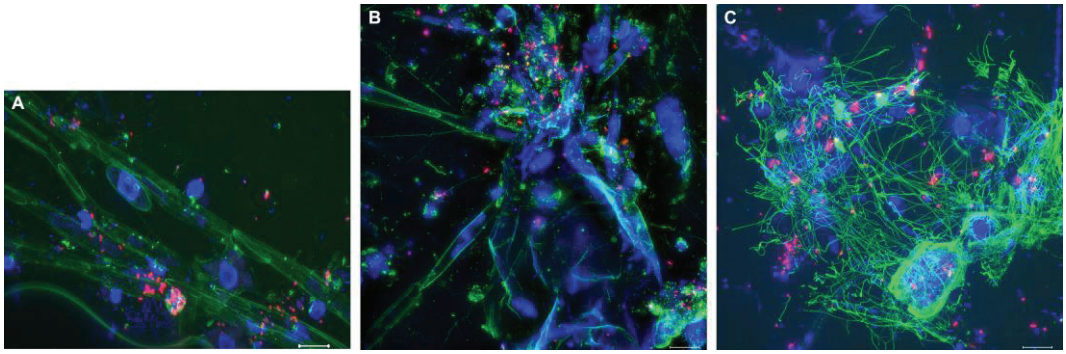


Figure 7.8: SR-SIM micrographs show the attachment of *Bacteroidetes* (red), detected with probe CF319a (MANZ ET AL. 1996), to phyto- and zooplankton in the South Atlantic Ocean. A: Station 76 fraction 20 to 40 µm. B: Station 70 fraction 20 to 40 µm. C: Station 76 fraction 20 to 40 µm. green: lectin AAL-488 used to stain phytoplankton and zooplankton surfaces; red: CF319a-Alexa495, blue: DAPI. Scale bar is 10 µm.

In this thesis I could show that members of the *Polaribacter* clade, one of the dominating *Bacteroidetes* clades in spring at Helgoland, were found attached to the diatom *Chaetoceros* spp. Since diatoms of the same diatom genus are also present in the Southern Atlantic Ocean the hypothesis should be tested that *Polaribacter* also attaches to those *Chaetoceros* species of the South Atlantic Ocean. BACKENHUS (2014) analyzed the attached fraction of the Helgoland phytoplankton bloom in spring 2011 and identified different *Chaetoceros* species, of which some were markedly colonized by *Bacteroidetes* and others were almost sterile. This further supports the perception that the interaction of these bacterial clades with diatom-derived polysaccharides is highly specific. MYKLESTAD ET AL. (1972) analyzed the cell wall and EPS composition of different *Chaetoceros* species and showed that the cell wall composition depended on the species, but consisted mainly of the monomers rhamnose, fucose, galactose and mannose (MYKLESTAD ET AL. 1972, HAUG & MYKLESTAD 1976). Further studies on the chemical composition of diatoms would be a prerequisite for a more detailed investigation of the interaction of their polysaccharides with *Bacteroidetes*.

References

ABELL, G. C. J. & BOWMAN, J. P. (2005): Ecological and biogeographic relationships of class *Flavobacteria* in the Southern Ocean. *FEMS Microbiol. Ecol.* **51**, 265-277.

ALDERKAMP, A.-C., BUMA, A. G. J. & VAN RIJSSEL, M. (2007): The carbohydrates of *Phaeocystis* and their degradation in the microbial food web. *Biogeochem.* **83**, 99-118.

ALONSO, C. & PERNTHALER, J. (2006): *Roseobacter* and SAR11 dominate microbial glucose uptake in coastal North Sea waters. *Environ. Microbiol.* **8**, 2022-2030.

ALONSO, C., WARNECKE, F., AMANN, R. & PERNTHALER, J. (2007). High local and global diversity of *Flavobacteria* in marine plankton. *Environ. Microbiol.* **9**, 1253-1266.

AMANN, R. I., LUDWIG, W. & SCHLEIFER, K. H. (1995): Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiol. Rev.* **59**,143-169.

AMIN, S. A., PARKER, M. S. & ARMBRUST, E. V. (2012): Interactions between diatoms and bacteria. *Microbiol. Mol. Biol. Rev.* **76**, 667.

ANDERSON, K. L. & SALYERS, A. A. (1989): Genetic evidence that outer membrane binding of starch is required for starch utilization by *Bacteroides thetaiotaomicron*. *J. Bacteriol.* **171**, 3199-3204.

ANDERSON, R. WINTER, C. & JÜRGENS, K. (2012): Protist grazing and viral lysis as prokaryotic mortality factors at Baltic Sea oxic-anoxic interfaces. *Mar. Ecol. Prog. Ser.* **467**, 1-14.

ARNOSTI, C. (2003): Microbial extracellular enzymes and their role in dissolved organic matter cycling. In: FINDLAY, S., SINSABAUGH, R. L. (EDS). *Aquatic ecosystems: interactivity of dissolved organic matter.* Academic Press: San Diego, CA. pp 316-342.

ARNOSTI, C., FUCHS, B. M., AMANN, R. & PASSOW, U. (2012): Contrasting extracellular enzyme activities of particle-associated bacteria from distinct provinces of the North Atlantic Ocean. *Front. Microbiol.* **3**, 1-9.

AZAM, F. (1998): Microbial control of oceanic carbon flux: The plot thickens. *Science* **280**, 694-696.

AZAM, F., FENCHEL, T., FIELD, J. G., GRAY, J. S., MEYERREIL, L. A. & THINGSTAD, F. (1983): The ecological role of water-column microbes in the sea. *Mar. Ecol. Prog. Ser.* **10**, 257-263.

AZIZ, R. K., BARTELS, D., BEST, A. A., DEJONGH, M., DISZ, T., EDWARDS, R. A., FORMSMA, K., GERDES, S., GLASS, E. M., KUBAL, M., MEYER, F., OLSEN, G. J., OLSON, R., OSTERMAN, A. L., OVERBEEK, R. A., MCNEIL, L. K., PAARMANN, D., PACZIAN, T., PARRELLO, B., PUSCH, G. D., REICH, C., STEVENS, R., VASSIEVA, O., VONSTEIN, V., WILKE, A. & ZAGNITKO, O. (2008): The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* **9**, published online.

BAAS BECKING, L. G. M. (1931): Gaia of leven en aarde. The Hague, the Netherlands: Nijhoff. Ambtsrede R.U. Leiden (in Dutch).

BAKENHUS, I. (2014): Spezifische Assoziation mariner Bakterien mit Diatomeen in der Frühjahrsblüte der Nordsee. *Master Thesis*, University Oldenburg, pp. 62.

BAHR, M., HOBBIE, J. E. & SOGIN, M. L. (1996): Bacterial diversity in an arctic lake: A freshwater SAR11 cluster. *Aquat. Microb. Ecol.* **11**, 271-277.

BANO, N. & HOLLIBAUGH, J. T. (2002): Phylogenetic composition of bacterioplankton assemblages from the Arctic Ocean. *Appl. Environ. Microbiol.* **68**, 505-518.

BARBEYRON, T., L'HARIDON, S., CORRE, E., KLOAREG, B. & POTIN, P. (2001): *Zobellia galactanovorans* gen. nov., sp. nov., a marine species of *Flavobacteriaceae* isolated from a red alga, and classification of *Cytophaga uliginosa* (ZOBELL AND UPHAM 1944) REICHENBACH 1989 as *Zobellia uliginosa* gen. nov., comb. nov. *Int. J. Syst. Evol. Microbiol.* 51, 985-997.

BAUER, M., KUBE, M., TEELING, H., RICHTER, M., LOMBARDOT, T., ALLERS, E., WÜRDEMANN, C. A., QUAST, C., KUHL, H., KNAUST, F., WOEBKEN, D., BISCHOF, K., MUSSMANN, M., CHOUDHURI, J. V., MEYER, F., REINHARDT, R., AMANN, R. I. & GLÖCKNER, F. O. (2006): Whole genome analysis of the marine *Bacteroidetes 'Gramella forsetii'* reveals adaptations to degradation of polymeric organic matter. *Environ. Microbiol.* **8**, 2201-2213.

BEARE, D. J., BATTEN, S., EDWARDS, M. & REID, D. G. (2002): 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 Res.* **48**, 29-49.

BECKER, G., GIESE, H., ISERT, K., KÖNIG, P., LANGENBERG, H., POHLMANN, T. & SCHRUM, C. (1999): Mesoscale structures, fluxes and water mass variability in the German Bight as exemplified in the KUSTOS-experiments and numerical models. *Ocean Dyn.* **51**, 155-179.

BERNARDET, J.-F. & NAKAGAWA, Y. (2006): Introduction to the Family *Flavobacteriacea*. In: M. DWORKIN, S. FALKOW, E. ROSENBERG, K. H. SCHLEIFER,

AND E. STACKEBRANDT (EDS.): *The Porkaryotes*, third edition, Vol. 7. Springer-Verlag, New York, pp. 550-580.

BERNARDET, J.-F. (2010): The *Bacteroidetes, Spirochaetes, Tenericutes (Mollicutes), Acidobacteria, Fibrobacteres, Fusobacteria, Dictyoglomi, Gemmatimonadetes, Lentisphaerae, Verrucomicrobia, Chlamydiae*, and *Planctomycetes*. In: KRIEG, N.R., STALEY, J.T., BROWN, D.R., HEDLUND, B.P., PASTER, B.J., WARD, N.L. et al (EDS.). *Bergey's Manual of Systematic Bacteriology,* Vol. 4, chapter Class II. *Flavobacteriia* class. nov., Springer, New York, pp. 106-314.

BERSCH, M. (1995): On the circulation of the northeastern North Atlantic. *Deep Sea Res.* I **42**, 1583-1607.

BIŽIĆ-IONESCU, M., ZEDER, M., IONESCU, D., ORLIĆ, S., FUCHS, B. M., GROSSART, H.-P., & AMANN, R. (2014): Comparison of bacterial communities on limnic vs. coastal marine particles reveals profound differences in colonization. *Env. Microbiol.*, in revision.

BJURSELL, M. K., MARTENS, E. C. & GORDON, J. I. (2006): Functional genomic and metabolic studies of the adaptations of a prominent adult human gut symbiont, *Bacteroides thetaiotaomicron*, to the suckling period. *J. Biol. Chem.* **281,** 36269-36279.

BLOEM, J. (1995): Fluorescent staining of microbes for total direct counts. In: AKKERMANS, A.D.L. AND VAN ELSAS, J.D. AND DE BRUIJN, F.J (EDS.). *Molecular Microbial Ecology Manual.* Chapter 4.1.8. Kluwer Academic Publishers, Springer, Heidelberg, pp. 1-12.

BORASTON, A.B., BOLAM, D.N., GILBERT, H.J., DAVIES, G.J. (2004): Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem. J.* 382, 769-781.

BOWMAN, J. P. (2006). The marine clade of the family *Flavobacteriaceae*: The genera A*equorivita, Arenibacter, Cellulophaga, Croceibacter, Formosa, Geldibacter, Gillisia, Maribacter, Mesonia, Muricauda, Polaribacter, Psychroflexus, Psychroserpens, Robignitalea, Salegentibacter, Tenacibaculum, Ulvibacter, Vitellibacter* and *Zobellia*. In: M. DWORKIN, FALKOW, S., ROSENBERG, E., SCHLEIFER, K. AND STACKEBRANDT, E. (**EDS**.). *The Prokaryotes*, Vol. 7, Springer-Verlag, pp. 677-694.

BRAUN, V. & HERRMANN, C.(1993): Evolutionary relationship of uptake systems for biopolymers in *Escherichia coli* - cross-complementation between the TonB ExbB ExbD and the TolA TolQ TolR proteins. *Mol. Microbiol.* **8**, 261-268.

BRETTAR, I., CHRISTEN, R. & HÖFLE, M. G. (2006): *Rheinheimera perlucida* sp. nov., a marine bacterium of the Gammaproteobacteria isolated from surface water of the central Baltic Sea. *Int. J. Sys. Evol. Microbiol.* **56**, 2177-2183.

BUCHAN, A., GONZÁLEZ, J. M. & MORAN, M. A. (2005): Overview of the marine *Roseobacter* lineage. *Appl. Environ. Microbiol.* **71**, 5665-5677.

CALBET, A. & LANDRY, M. R. (2004): Phytoplankton growth, microzooplankton grazing, and carbon cycling in marine systems. *Limnol. Oceanogr.* **49**, 51-57.

CANTAREL, B. L., COUTINHO, P. M., RANCUREL, C., BERNARD, T., LOMBARD, V., & HENRISSAT, B. (2009): The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.* **37**, D233-D238.

CHEN, Y.-L. (2012): Bottom-up control on annually reoccurring bacterial community during spring bloom at station Kabeltonne (Helgoland). *Master Thesis*, University Gent, pp. 70.

CHO, K. H. & SALYERS, A. A. (2001): Biochemical analysis of interactions between outer membrane proteins that contribute to starch utilization by *Bacteroides thetaiotaomicron*. *J. Bacteriol.* **183**, 7224-7230.

CHO, J. C. & GIOVANNONI, S. J. (2003): *Croceibacter atlanticus* gen. nov., sp. nov., a novel marine bacterium in the family *Flavobacteriaceae*. *Syst. Appl. Microbiol.* **26**, 76-83.

CLARK, J.W. & KAMBHAMPATI, S. (2003): Phylogenetic relationships among *Blattabacterium*, endosymbiotic bacteria from the wood roach, *Cryptocercus*. *Mol. Phylogenet. Evol.* **26**, 82-88.

COTTRELL, M. T. & KIRCHMAN, D. L. (2000): Natural assemblages of marine proteobacteria and members of the *Cytophaga-Flavobacter* cluster consuming low- and high-molecular-weight dissolved organic matter. *Appl. Environ. Microbiol.* **66,** 1692-1697.

COTTRELL, M. T., YU, L. Y. & KIRCHMAN, D. L. (2005): Sequence and expression analyses of *Cytophaga*-like hydrolases in a Western Arctic metagenornic library and the Sargasso seat. *Appl. Environ. Microbiol.* **71**, 8506-8513.

CRUMP, B., ARMBRUST, E. & BAROSS, J. (1999): Phylogenetic analysis of particle-attached and free-living bacterial communities in the Columbia River, its estuary, and the adjacent coastal ocean. *Appl. Environ. Microbiol.* **65**, 3192-3204.

DARLING, A. C. E., MAU, B., BLATTNER, F. R. & PERNA, N. T. (2010): progressive Mauve: multiple genome alignment with gene gain, loss, and rearrangement. *PLoS One.* **5**, e11147.

DELONG, E. F., FRANKS, D. G. & ALLDREDGE, A. L. (1993): Phylogenetic diversity of aggregate-attached vs. free-living marine bacterial assemblages. *Limnol. Oceanogr.* **38**, 924-934.

DESCAMPS, V., COLIN, S., LAHAYE, M., JAM, M., RICHARD, C., POTIN, P. BARBEYRON, T., YVIN, J. C. & KLOAREG, B. (2006): Isolation and culture of a marine bacterium degrading the sulfated fucans from marine brown algae. *Mar. Biotechnol.* (NY) **8**, 27-39.

DÍEZ-VIVES, C., GASOL, J. M. & ACINAS, S. G. (2014): Spatial and temporal variability among *Bacteroidetes* populations in the NW Mediterranean Sea. *Syst. Appl. Microbiol.* **37**, 68-78.

DIPPNER, J.W. (1993): A frontal-resolving model for the German Bight. *Cont. Shelf Res.* **13**, 49-66.

EILER, A., & BERTILSSON, S. (2007): *Flavobacteria* blooms in four eutrophic lakes: linking population dynamics of freshwater bacterioplankton to resource availability. *Appl. Environ. Microbiol.* **73**, 3511-3518.

EILERS, H., PERNTHALER, J., PEPLIES, J., GLOECKNER, F. O., GERDTS, G. & AMANN, R. (2001): Isolation of novel pelagic bacteria from the German Bight and their seasonal contributions to surface picoplankton. *Appl. Environ. Microbiol.* **67**, 5134-5142.

EILERS, K. G., DEBENPORT, S., ANDERSON, S. & FIERER, N. (2012): Digging deeper to find unique microbial communities: The strong effect of depth on the structure of bacterial and archaeal communities in soil. *Soil Biol. Biochem.* **50**, 58-65.

ELIFANTZ, H., MALMSTROM, R. R., COTTRELL, M. T. & KIRCHMAN, D. L. (2005): Assimilation of polysaccharides and glucose by major bacterial groups in the Delaware Estuary. *Appl. Environ. Microbiol. 71*, 7799-7805.

FANDINO, L. B., RIEMANN, L., STEWARD, G. F. & AZAM, F. (2005): Population dynamics of *Cytophaga-Flavobacteria* during marine phytoplankton blooms analyzed by real-time quantitative PCR. *Aquat. Microbial. Ecol.* **40**, 251-257.

FANDINO, L. B., RIEMANN, L., STEWARD, G. F., LONG, R. A. & AZAM, F. (2001): Variations in bacterial community structure during a dinoflagellate bloom analyzed by DGGE and 16S rDNA sequencing. *Aquat. Microb. Ecol.* **23**, 119-130.

FERNÁNDEZ-GÓMEZ, B., RICHTER, M., SCHUELER, M., PINHASSI, J., ACINAS, S. G., GONZÁLEZ, J. M. & PEDRÓS-ALIÓ, C. (2013): Ecology of marine *Bacteroidetes:* a comparative genomics approach. *ISME J.* **7**, 1026-1037.

FIELD, C. B., BEHRENFELD, M. J., RANDERSON, J. T. & FALKOWSKI, P. (1998): Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* **281**, 237-240.

FRANKE, H.-D., BUCHHOLZ, F. & WILTSHIRE, K. H. (2004): Ecological long-term research at Helgoland (German Bight, North Sea): retrospect and prospect - an introduction. *Helgol. Mar. Res.* **58**, 223-229.

FUHRMAN, J. A., HEWSON, I., SCHWALBACH, M. S., STEELE, J. A., BROWN, M. V. & NAEEM, S. (2006): Annually reoccurring bacterial communities are predictable from ocean conditions. *PNAS* **103**, 13104-13109.

FUHRMAN, J. A., MCCALLUM, K. & DAVIS, A. A. (1993): Phylogenetic diversity of subsurface marine microbial communities from the Atlantic and Pacific Oceans. *Appl. Environm. Microbiol.* **59**, 1294-1302.

FUHRMAN, J. A., STEELE, J. A., HEWSON, I., SCHWALBACH, M. S., BROWN, M. V., GREEN, J. L. & BROWN, J. H. (2008): A latitudinal diversity gradient in planktonic marine bacteria. *PNAS* **105**, 7774-7778.

GATTUSO, J.-P., FRANKIGNOULLE, M. & WOLLAST, R. (1998): Carbon and carbonate metabolism in coastal aquatic ecosystems. *Ann. Rev. Ecol. Syst.* **29**, 405-435.

GERLACH, S. (1995): North Sea research: where might it go? *Helgol. Mar. Res.* **49**, 703-707.

GERLACH, S. A. (1994): Die "Sieben Meere". In: S. A. GERLACH (ED.), *Spezielle Ökologie: Marine Systeme*, Springer Verlag, pp. 226.

GIOVANNONI, S. J. & RAPPÉ, M. (2000): Evolution, diversity, and molecular ecology of marine prokaryotes. p. 47-84. In: D. L. KIRCHMAN (ED.), *Microbial ecology of the oceans*. John Wiley & Sons, Inc., New York, N.Y.

GIOVANNONI, S. J. & STINGL, U. (2005): Molecular diversity and ecology of microbial plankton. *Nature* **437**, 343-348.

GLÖCKNER, F. O., FUCHS, B. M. & AMANN, R. (1999): Bacterioplankton composition in lakes and oceans: a first comparison based on fluorescence *in situ* hybridization. *Appl. Environ. Microbiol.* **65**, 3721-3726.

GÓMEZ-PEREIRA, P. R., FUCHS, B. M., ALONSO, C., OLIVER, M. J., VAN BEUSEKOM, J. E. E. & AMANN R. (2010): Distinct flavobacterial communities in contrasting water masses of the North Atlantic Ocean. *ISME J.* **4**, 472-487.

GÓMEZ-PEREIRA, P.R., SCHÜLER, M., FUCHS, B.M., BENNKE, C., TEELING, H., WALDMANN, J., RICHTER, M., BARBE, V., BATAILLE, E., GLÖCKNER, F.O., AMANN, R. (2012): Genomic content of uncultured *Bacteroidetes* from contrasting oceanic provinces in the North Atlantic Ocean. *Environ. Microbiol.* **14**, 52-66.

GREVE, W., REINERS, F. & NAST, J. (1996): Biocoenotic changes of the zooplankton in the German Bight: the possible effects of eutrophication and climate. *ICES J. Mar. Sci.* **53**, 951-956.

GROSSART, H.-P., LEVOLD, F., ALLGAIER, M., SIMON, M., BRINKHOFF, T. (2005): Marine diatom species harbour distinct bacterial communities. *Environ. Microbiol.* **7**, 860-873.

GRZYMSKI, J. J., CARTER, B. J. & DELONG, E. F. (2006): Comparative genomics of DNA fragments from six Antarctic marine planktonic bacteria. *Appl. Environ. Microbiol.* **72**, 1532-1541.

HAHNKE, R. L. (2013): Cultivation of *Flavobacteria* and other *in situ* abundant bacteria from the North Sea. *Doctoral Thesis*, University Bremen, pp. 311.

HALL, J. A. & SAFI, K. (2001): The impact of *in situ* Fe fertilisation on the microbial food web in the Southern Ocean. *Deep-Sea Res.* II **48**, 2591-2613.

HAUG, A. & MYKLESTAD, S. (1976): Polysaccharides of marine diatoms with special reference to *Chaetoceros* species. *Mar. Biol.* **34**, 217-222.

HEINRICH, F., EILER, A. & BERTILSSON, S. (2013): Seasonality and environmental control of freshwater SAR11 (LD12) in a temperate lake (Lake Erken, Sweden). *Aquat. Microb. Ecol.* **70**, 33-44.

HERVE, C., ROGOWSKI, A., BLAKE, A.W., MARCUS, S.E., GILBERT, H.J., KNOX, J.P. (2010): Carbohydrate-binding modules promote the enzymatic deconstruction of intact plant cell walls by targeting and proximity effects. *PNAS* **107**, 15293-15298.

HOBBIE, J. E., DALEY, R. J. & JASPER, S. (1977): Use of nuclepore filters for counting bacteria by fluorescence microscopy. *Appl. Environ. Microbiol.* **33**, 1225-1228.

HUANG, S. (2013): Assessing the polysaccharide degradation potential in marine microbial genomes and metagenomes. *Doctoral Thesis*, University Bremen, pp. 233.

HUGENHOLTZ, P., GOEBEL, B. M. & PACE, N. R. (1998): Impact of culture-indipendent studies on the emerging phylogenetic view of bacterial diversity. *J. Bact.* **180**, 4765-4774.

HUMPHRY, D. R., GEORGE, A., BLACK, G. W. & CUMMINGS, S. P. (2001): *Flavobacterium frigidarium* sp. nov., an aerobic, psychrophilic, xylanolytic and laminarinolytic bacterium from Antarctica. *Int. J. Syst. Evol. Microbiol.* **51**, 1235-1243.

IRIARTE, A. & PURDIE, D. (2004): Factors controlling the timing of major spring bloom events in an UK south coast estuary. *Estuar. Coast. Shelf Sci.* **61**, 679-690.

ISHII, K., MUßMANN, M., MACGREGOR, B. J. & AMANN, R. (2004): An improved fluorescence *in situ* hybridization protocol for the identification of bacteria and archaea in marine sediments. *FEMS Microbiol. Ecol.* **50**, 203-212.

IVANOVA, E. P., ALEXEEVA, Y. V., FLAVIER, S., WRIGHT, J. P., ZHUKOVA, N. V., GORSHKOVA, N. M., MIKHAILOV, V. V., NICOLAU, D. V. AND CHRISTEN, R. (2004): *Formosa algae* gen. nov., sp. nov., a novel member of the family *Flavobacteriaceae*. *Int. J. Syst. Evol. Microbiol.* **54**, 705-711.

JOHNSON, Z. I., ZINSER, E. R., COE, A., MCNULTY, N. P., WOODWARD, E. M. S. & CHISHOLM, S. W. (2006): Niche partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental gradients. *Science* **311**, 1737-1740.

JONES, J. G. & SIMON, B. M. (1975): An investigation of errors in direct counts of aquatic bacteria by epifluorescence microscopy, with reference to a new method for dyeing membrane filters. *J. Appl. Bacteriol.* **39,** 1-13.

KAPPELMANN, L. (2013): (Meta-)genomic analysis of the diversity and the carbohydrate degradation potential of the SAR92 clade during a diatom-induced bacterioplankton bloom. *Master Thesis*, University Bremen, pp. 53.

KEPNER, R. L. & PRATT, J. R. (1994): Use of fluorochromes for direct enumeration of total bacteria in environmental samples: past and present. *Microbiol. Rev.* **58**, 603-615.

KIRCHMAN, D. L. (2002): The ecology of *Cytophaga-Flavobacteria* in aquatic environments. *FEMS Microbiol. Ecol.* **39**, 91-100.

KIRCHMAN, D. L., YU, L. & COTTRELL, M. T. (2003): Diversity and abundance of uncultured *Cytophaga*-like bacteria in the Delaware Estuary. *Appl. Environ. Microbiol.* **69**, 6587-6596.

KIRCHMAN, D. L., DITTEL, A. I., MALMSTROM, R. R. & COTTRELL, M. T. (2005): Biogeography of major bacterial groups in the Delaware Estuary. *Limnol. Oceanogr.* **50**, 1697-1706.

KOEBNIK, R. (2005): TonB-dependent trans-envelope signalling: the exception or the rule? *Trends Microbiol.* **13**, 343-347.

KOLTON, M., SELA, N., ELAD, Y. & CYTRYN, E. (2013): Comparative genomic analysis indicates that niche adaption of terrestrial *Flavobacteria* is strongly linked to plant glycan metabolism. *PLoS One*, e76704.

KORMAS, K. A., TIVEY, M. K., VON DAMM, K. & TESKE, A. (2006): Bacterial and archaeal phylotypes associated with distinct mineralogical layers of a white smoker spire from a deep-sea hydrothermal vent site (9 degrees N, East Pacific Rise). *Environ. Microbiol.* **8**, 909-920.

KOROPATKIN, N. M., CAMERON, E. A. & MARTENS, E. C. (2012): How glycan metabolism shapes the human gut microbiota. *Nat. Rev. Microbiol.* 10, 323-335.

KOROPATKIN, N. M., MARTENS, E. C., GORDON, J. I., SMITH, T. J. (2008): Starch catabolism by a prominent human gut symbiont is directed by the recognition of amylose helices. *Structure* **16**, 1105-1115.

KRAUSS, W. (1986). The North Atlantic Current. *J. Geophys. Res.* **91,** 5061-5074.

LAU, W. W. Y., KEIL, R. G. & ARMBRUST, E. V. (2007): Succession and Diel transcriptional response of the glycolate-utilizing component of the bacterial community during a spring phytoplankton bloom. *Appl. Environm. Microbiol.* **73**, 2440-2450.

LOHMANN, G. & WILTSHIRE, K. H. (2012): Winter atmospheric circulation signature for the timing of the spring bloom of diatoms in the North Sea. *Mar. Biol.* **159**, 2573-2581.

LOMBARD, V., RAMULU, H. G., DRULA, E., COUTINHO, P. M. & HENRISSAT, B. (2014): The carbohydrate-active enzymes database. *Nucleic Acids Res.* **42**, D490-D495.

LONGHURST, A. (1998): Ecological Geography of the Sea. Academic Press, New York.

LUDWIG, W., EUZEBY, J. P. & WHITMAN, W. B. (2008): Draft taxonomic outline of the *Bacteroidetes, Planctomycetes, Chlamydiae, Spirochaetes, Fibriobacteres, Fusobacteria, Acidobacteria, Verrucomicrobia, Dichtyoglomi, Gemmatimonadetes.* In: B. HEDLUND, KRIEG, N., LUDWIG, W., PASTER, B., STALEY, J., WARD, N. WHITMAN, W. B. (ED.), *Bergey`s Manual of Systematic Bacteriology*, Vol. 4, Springer Verlag, New York.

LUMBY, J. R. (1932): Current systems of the North Atlantic and the North Sea. *J. Ecol.* **20**, 314-325.

MAHADEVAN, A., D'ASARO, E., LEE, C. & PERRY, M. J. (2012): Eddy-driven stratification initiates North Atlantic spring phytoplankton blooms. *Science* 337, 54-58.

MALMSTROM, R., R., STRAZA, T., R. A., COTTRELL, M., T. & KIRCHMAN, D., L. (2007): Diversity, abundance, and biomass production of bacterial groups in the western Arctic Ocean. *Aquat. Microb. Ecol.* **47**, 45-55.

MANN, A. J., HAHNKE, R. L., HUANG, S., WERNER, J., XING, P., BARBEYRON, T., HUETTEL, B., STÜBER, K., REINHARDT, R., HARDER, J., GLÖCKNER, F. O., AMANN, R. I. & TEELING, H. (2013): The genome of the alga-associated marine Flavobacterium *Formosa agariphila* KMM 3901(T) reveals a broad potential for degradation of algal polysaccharides. *Appl. Environ. Microbiol.* **79**, 6813-6822.

MARTENS, E. C. , LOWE, E. C., CHIANG, H., PUDLO, N. A., MCNULTY, N. P., ABBOTT, D. W., HENRISSAT, B., GILBERT, H. J., BOLAM, D. N. &GORDON, J. I. (2011): Recognition and degradation of plant cell wall polysaccharides by two human gut symbionts. *PLoS Biol.* **9**, e1001221.

MARTENS, E. C., KOROPATKIN, N. M., SMITH, T. J. & GORDON, J. I. (2009): Complex glycan catabolism by the human gut microbiota: The *Bacteroidetes* Sus-like paradigm. *J. Biol. Chem.* **284**, 24673-24677.

MARTIN, J. H. (1990): Glacial-interglacial CO_2 change: The iron hypothesis. *Paleoceanography* **5**, 1-13.

MCBRIDE, M. J., XIE, G., MARTENS, E. C., LAPIDUS, A., HENRISSAT, B., RHODES, R. G., GOLTSMAN, E., WANG, W, XU, J, HUNNICUTT, D. W., STAROSCIK, A. M., HOOVER, T. R., CHENG, Y. Q., STEIN, J. L. (2009): Novel features of the polysaccharide digesting gliding bacterium *Flavobacterium johnsoniae* as revealed by genome sequence analysis. *Appl. Environ. Microb.* **75**, 6864-6875.

MCBRIDE, M. J. & ZHU, Y. (2012): Gliding motility and Por secretion system genes are widespread among members of the phylum *Bacteroidetes*. *J. Bacteriol.* ***195,*** 270-278.

MOLLER, I., SORENSEN, I., BERNAL, A.J., BLAUKOPF, C., LEE, K., OBRO, J., PETTOLINO, F., ROBERTS, A., MIKKELSEN, J.D., KNOX, J.P., BACIC, A., WILLATS, W.G.T. (2007): High-throughput mapping of cell-wall polymers within and between plants using novel microarrays. *Plant J.* **50**, 1118-1128.

MORRIS, R. M., RAPPÉ, M. S., CONNON, S. A., VERGIN, K. L., SIEBOLD, W. A., CARLSON, C. A. & GIOVANNONI, S. J. (2002): SAR11 clade dominates ocean surface bacterioplankton communities. *Nature* **420**, 806-810.

MUNOZ, R., YARZA, P., LUDWIG, W., EUZÉBY, J., AMANN, R., SCHLEIFER, K. H., GLÖCKNER, F.O., & ROSSELLÓ-MÓRA, R. (2011): Release LTPs104 of the All-Species Living Tree. *Syst. Appl. Microbiol.* **34**, 169-170.

MYKLESTAD, S., HAUG, A., LARSEN, B. (1972): Production of carbohydrates by the marine diatom *Chaetoceros affinis* var. willei (gran) Hustedt. II preliminary investigation of the extracellular polysaccharide. *J. Exp. Mar. Biol. Ecol.* **9**, 137-144.

NEDASHKOVSKAYA, O. I., KIM, S. B., HAN, S. K., RHEE, M. S., LYSENKO, A. M., FALSEN, E., FROLOVA, G. M., MIKHAILOV, V. V., BAE, K. S. (2004): *Ulvibacter litoralis* gen. nov., sp nov., a novel member of the family *Flavobacteriaceae* isolated from the green alga *Ulva fenestrata*. *Int. J. Syst. Evol. Microbiol.* **54**, 119-123.

NEDASHKOVSKAYA, O. I., KIM, S. B., VANCANNEYT, M., SNAUWAERT, C., LYSENKO, A. M., ROHDE, M., FROLOVA, G. M., ZHUKOVA, N. V., MIKHAILOV, V. V., BAE, K. S., OH, H. W. & SWINGS, J. (2006): *Formosa agariphila* sp. nov., a budding bacterium of the family *Flavobacteriaceae* isolated from marine environments, and emended description of the genus *Formosa*. *Int. J. Syst. Evol. Microbiol.* **56**, 161-167.

NEWTON, R. J., JONES, S. E., EILER, A., MCMAHON, K. D. & BERTILSSON, S. (2011): A guide to the natural history of freshwater lake bacteria. *Microbiol. Mol. Biol. Rev.* **75**, 14-49.

NICOLAS, J.-L., CORRE, C. & COCHARD, J.-C. (2004): Bacterial population association with phytoplankton cultured in a bivalve hatchery. *Microb. Ecol.* **48**, 400-413.

OLIVER, J. L., BARBER, R. T., SMITH, W. O. & DUCKLOW, H. W. (2004): The heterotrophic bacterial response during the Southern Ocean Iron Experiment (SOFeX). *Limnol. Oceanogr.* **49**, 2129-2140.

PEDROTTI, M. L., BEAUVAIS, S., KERROS, M. E., IVERSEN, K. & PETERS, F. (2009): Bacterial colonization of transparent exopolymeric particles in mesocosms under different turbulence intensities and nutrient conditions. *Aquat. Microb. Ecol.* **55**, 301-312.

PERNTHALER, J. & AMANN, R. (2005): Fate of heterotrophic microbes in pelagic habitats: focus on populations. *Microbiol. Mol. Biol. Rev.* **69**, 440-461.

PERNTHALER, J. (2005): Predation on prokaryotes in the water column and its ecological implications. *Nat. Rev. Microbiol.* **3**, 537-546.

PINHASSI, J., SALA, M. M., HAVSKUM, H., PETERS, F., GUADAYOL, O., MALITS, A. & MARRASÉ, C. (2004): Changes in bacterioplankton composition under different phytoplankton regimens. *Appl. Environ. Microb.* **70**, 6753-6766.

PINHASSI, J., SIMÓ, R., GONZÁLEZ, J. M., VILA, M., ALONSO-SÁEZ, L., KEINE, R. P., MORAN, M. A. & PEDRÓS-ALIÓ, C. (2005): Dimethylsulfoniopropionate turnover is linked to the composition and dynamics of the bacterioplankton assemblage during a microcosm phytoplankton bloom. *Appl. Environ. Microbiol.* **71**, 7650-7660.

POMEROY, L., WILLIAMS, P., AZAM, F. & HOBBIE, F. (2007): 'The Microbial Loop'. *Oceanogr.* 28-33.

POMMIER, T., PINHASSI, J. & HAGSTROM, A. (2005): Biogeographic analysis of ribosomal RNA clusters from marine bacterioplankton. *Aquat. Microb. Ecol.* **41**, 79-89.

POPPER, Z.A., MICHEL, G., HERVE, C., DOMOZYCH, D.S., WILLATS, W.G.T., TUOHY, M.G., KLOAREG, B., STENGEL, D.B. (2011): Evolution and diversity of plant cell walls: from algae to flowering plants. *Annu. Rev. Plant Biol.* **62**, 567-590.

PORTER, F. G. & FEIG, Y. S. (1980): The use of DAPI for identifying and counting aquatic microflora. *Limnol. Oceanogr.* **25**, 943-948.

RAABE, T., & WILTSHIRE, K. H. (2009): Quality control and analyses of the long-term nutrient data from Helgoland Roads, North Sea. *J. Sea Res.* **61**, 3-16.

RAPPÉ, M. S., KEMP, P. F. & GIOVANNONI, S. J. (1997): Phylogenetic diversity of marine coastal picoplankton 16S rRNA genes cloned from the continental shelf off Cape Hatteras, North Carolina. *Limnol. Oceanogr.* **42**, 811-826.

REEVES, A. R., WANG, G. R. & SALYERS, A. A. (1997): Characterization of four outer membrane proteins that play a role in utilization of starch by *Bacteroides thetaiotaomicron*. *J. Bacteriol.* **179**, 643-649.

RIEMANN, L., STEWARD, G. F. & AZAM, F. (2000): Dynamics of bacterial community composition and activity during a mesocosm diatom bloom. *Appl. Environ. Microbiol.* **66**, 578-587.

RINTA-KANTO, J. M., SUN, S., SHARMA, S., KIENE, R. P. & MORAN, M. A. (2012): Bacterial community transcription patterns during a marine phytoplankton bloom. *Environ. Microbiol.* **14**, 228-239.

ROSSBY, T. (1996): The North Atlantic Current and surrounding waters: At the crossroads. *Rev. Geophys.* **34**, 463-481.

RUIZ-PONTE, C., CILIA, V., LAMBERT, C. & NICOLAS, J. L. (1998): *Roseobacter gallaeciensis* sp. nov., a new marine bacterium isolated from rearings and collectors of the scallop *Pecten maximus*. *Int. J. Syst. Bacteriol.* **48**, 537-542.

SALCHER, M. M., PERNTHALER, J. & POSCH, T. (2011): Seasonal bloom dynamics and ecophysiology of the freshwater sister clade of SAR11 bacteria 'that rule the waves' (LD12). *ISME J.* **5**, 1242-1252.

SAPP, M., SCHWADERER, A.S., WILTSHIRE, K.H., HOPPE, H.-G., GERDTS, G., WICHELS, A. (2007) Species-specific bacterial communities in the phycosphere of microalgae? *Microb. Ecol.* **53**, 683-699.

SCHÄFER, H., ABBAS, B., WITTE, H. & MUYZER, G. (2002): Genetic diversity of 'satellite' bacteria present in cultures of marine diatoms. *FEMS Microbiol. Ecol.* **42**, 25-35.

SCHATTENHOFER, M., FUCHS, B. M., AMANN, R., ZUBKOV, M. V., TARRAN, G. A. & PERNTHALER, J. (2009): Latitudinal distribution of prokaryotic picoplankton populations in the Atlantic Ocean. *Environ. Microbiol.* **11**, 2078-2093.

SCHAUER, K., RODIONOV, D. A. & DE REUSE, H. (2008): New substrates for TonB-dependent transport: do we only see the 'tip of the iceberg'? *Trends Biochem. Sci.* **33**, 330-338.

SCHAUER, R., BIENHOLD, C., RAMETTE, A. & HARDER, J. (2010): Bacterial diversity and biogeography in deep-sea surface sediments of the South Atlantic Ocean. *ISME J.* **4**, 159-170.

SCHMITZ, W. J. & MCCARTNEY, M. S. (1993): On the North Atlantic circulation. *Rev. Geophys.* **31**, 29-49.

SHIPMAN, J. A., BERLEMAN, J. E. & SALYERS, A. A. (2000): Characterization of four outer membrane proteins involved in binding starch to the cell surface of *Bacteroides thetaiotaomicron*. *J. Bacteriol.* **182**, 5365-5372.

SIEVERT, S. M., KUEVER, J. & MUYZER, G. (2000): Identification of 16S ribosomal DNA-defined bacterial populations at a shallow submarine hydrothermal vent near Milos Island (Greece). *Appl. Environ. Microbiol.* **66**, 3102-3109.

SIMON, M., GLÖCKNER, F. O. & AMANN, R. (1999). Different community structure and temperature optima of heterotrophic picoplankton in various regions of the Southern Ocean. *Aquat. Microb. Ecol.* **18**, 275-284.

SMETACEK, V. (1999): Diatoms and the ocean carbon cycle. *Protist* **150**, 25-32.

SMITH, C., ROCHA, E. & PASTER, B. (2006): The medically important *Bacteroidetes* spp. in health disease. In: M. DWORKIN, FALKOW, S., ROSENBERG, E., SCHLEIFER, K. AND STACKEBRANDT, E. (EDS.). *The Prokaryotes*, Vol. 7, Springer-Verlag, New York, pp. 381-427.

STALEY, J. T. & GOSINK, J. J. (1999): Poles apart: Biodiversity and biogeography of sea ice bacteria. *Ann. Rev. Microbiol.* **53**, 189-215.

STORM, S. L., BRAINARD, M. A., HOLMES, J. L. & OLSON, M. B. (2001): Phytoplankton blooms are strongly impacted by microzooplankton grazing in coastal North Pacific Waters. *Mar. Biol.* 138, 355-368.

TADA, Y., TANIGUCHI, A., NAGAO, I., MIKI, T., UEMATSU, M., TSUDA, A. & HAMASAKI, K. (2011): Differing growth responses of major phylogenetic groups

of marine bacteria to natural phytoplankton blooms in the western North Pacific Ocean. *Appl. Environ. Microbiol.* **77**, 4055-4065.

TEELING, H., FUCHS, B. M., BECHER, D., KLOCKOW, C., GARDEBRECHT, A., BENNKE, C. M., KASSABGY, M., HUANG, S. X., MANN, A.J., WALDMANN, J., WEBER, M., KLINDWORTH, A., OTTO, A., LANGE, J., BERNHARDT, J., REINSCH, C., HECKER, M., PEPLIES, J., BOCKELMANN, F. D., CALLIES, U., GERDTS, G., WICHELS, A., WILTSHIRE, K. H., GLOCKNER, F. O., SCHWEDER, T., AMANN, R. (2012): Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* **336**, 608-611.

THIELE, S., FUCHS, B., AMANN, R. (2011): Identification of microorganisms using the ribosomal RNA approach and fluorescence *in situ* hybridization. In: WILDERER P. (ED.) *Treatise on Water Science*, Academic Press, Oxford, pp. 171-189.

THIELE, S., FUCHS, B. M., RAMAIAH, N. & AMANN, R. (2012): Microbial community response during the iron fertilization experiment LOHAFEX. *Appl. Environ. Microbiol.* **78**, 8803-8812.

THIELE, S. (2013): The role of specific microbial communities in the biological carbon pump. *Doctoral Thesis*, University Bremen, pp. 180.

THOMAS, F., HEHEMANN, J.-H., REBUFFET, E., CZJZEK, M. & MICHEL, G. (2011): Environmental and gut *Bacteroidetes*: the food connection. *Front. Microbiol.* **2**, 93.

TIAN, T., SU, J., FLÖSER, G., WILTSHIRE, K. & WIRTZ, K. (2011): Factors controlling the onset of spring blooms in the German Bight 2002-2005: Light, wind and stratification. *Cont. Shelf Res.* **31**, 1140-1148.

TOMCZAK, M. & GODFREY, J. (1994): Regional Oceanography: an introduction. First edition. Pergamon.

VISSER, M., BATTEN, S., BECKER, G., BOT, P., COLIJN, F. & DAMM, P. (1996): Time series analysis of monthly mean data of temperature, salinity, nutrients, suspended matter, phyto- and zooplankton at eight locations on the Northwest European shelf. *Ocean Dyn.* **48**, 299-323.

WARREN, R. A. J. (1996): Microbial hydrolysis of polysaccharides. *Ann. Rev. Microb.* **50**, 183-212.

WHITMAN, W. B., COLEMAN, D. C. & WIEBE, W. J. (1998): Prokaryotes: The unseen majority. *PNAS* **95**, 6578-6583.

WILLIAMS, T. J., WILKINS, D., LONG, E., EVANS, F., DEMAERE, M. Z., RAFTERY, M. J. & CAVICCHIOLI, R. (2013): The role of planktonic *Flavobacteria* in processing algal organic matter in coastal East Antarctica revealed using metagenomics and metaproteomics. *Environ. Microbiol.* 15, 1302-1317.

WILTSHIRE, K. H. & MANLY, B. F. J. (2004): The warming trend at Helgoland Roads, North Sea: Phytoplankton response. *Helgol. Mar. Res.* **58**, 269-273.

WILTSHIRE, K.H., KRABERG, A., BARTSCH, I., BOERSMA, M., FANKE, H.-D., FREUND, J., GEBÜHR, C., GERDTS, G., STOCKMANN, K. & WICHELS A. (2010): Helgoland Roads, North Sea: 45 Years of Change. *Estuaries and Coasts* **33**, 295-310.

WILTSHIRE, K.H., MALZAHN, A.M., WIRTZ, K., GREVE, W., JANISCH, S., MANGELSDORF, P., MANLY, B.F.J. & BOERSMA, M. (2008): Resilience of North Sea phytoplankton spring bloom dynamics: An analysis of long-term data at Helgoland Roads. *Limnol. Oceanogr.* **53**, 1294-1302.

WOEBKEN, D., FUCHS, B. M., KUYPERS, M. A. A. & AMANN, R. (2007): Potential interactions of particle-associated anammox bacteria with bacterial and archaeal partners in the Namibian upwelling system. *Appl. Environ. Microbiol.* **73**, 4648-4657.

WOESE, C. R. (1987): Bacterial Evolution. *Microbiol. Rev.* **51**, 221-271.

WOYKE, T., XIE, G., COPELAND, A., GONZÁLEZ, J. M., HAN, C., KISS, H., SAW, J. H., SENIN, P., YANG, C., CHATTERJI, S., CHENG, J.F., EISEN, J. A., SIERACKI, M. E. & STEPANAUSKAS, R. (2009): Assembling the marine metagenome, one cell at a time. *PLoS One* **4**, e5299.

XU, J., BJURSELL, M. K., HIMROD, J., DENG, S., CARMICHAEL, L. K., CHIANG, H. C., HOOPER, L. V., & GORDON, J. I. (2003): A genomic view of the human *Bacteroides thetaiotaomicron* symbiosis. *Science* **299**, 2074-2076.

YARZA, P., RICHTER, M., PEPLIES, J., EUZEBY, J., AMANN, R., SCHLEIFER, K.-H., LUDWIG, W., GLÖCKNER, F. O. & ROSSELLÓ-MÓRA, R. (2008): The all-species living tree project: a 16S rRNA-based phylogenetic tree of all sequenced type strains. *Syst. Appl. Microbiol.* **31**, 241-250.

ZEDER, M., & PERNTHALER, J. (2009): Multi-spot live-image autofocusing for high-throughput microscopy of fluorescently stained bacteria. *Cytom. Part A.* **75**, 781-788.

ZEDER, M., PETER, S., SHABAROVA, T. & PERNTHALER, J. (2009): A small population of planktonic *Flavobacteria* with disproportionally high growth during the springphytoplankton bloom in a prealpine lake. *Environ. Microbiol.* **11**, 2676-86.

ZEDER, M., KOHLER, E. & PERNTHALER, J. (2010): Automated quality assessment of autonomously acquired microscopic images of fluorescently stained bacteria. *Cytom. Part A.* **77**, 76-85.

ZUBKOV, M. V., SLEIGH, M. A., TARRAN, G. A., BURKIL, P. H. & LEAKEY, R. J. G. (1998): Picoplanktonic community structure on an Atlantic transect from 50° N to 50° S. *Deep Sea Res. I* **45**, 1339-1355.

ZWART, G., CRUMP, B. C., KAMST-VAN AGTERVELD, M. P., HAGEN, F. & HAN, S.-K. (2002): Typical freshwater bacteria: an analysis of available 16S rRNA gene sequences from plankton of lakes and rivers. *Aquat. Microb. Ecol.* **28**, 141-155.

ZWART, G., HIORNS, W. D., METHÉ, B. A., VAN AGTERVELD, M. P., HUISMANS, R., NOLD, S. C., ZEHR, J. P. & LAANBROEK, H. J. (1998): Nearly identical 16S rRNA sequences recovered from lakes in North America and Europe indicate the existence of clades of globally distributed freshwater bacteria. *Syst. Appl. Microbiol.* **21**, 546-556.

Appendix

Appendix A: Probe Evaluation

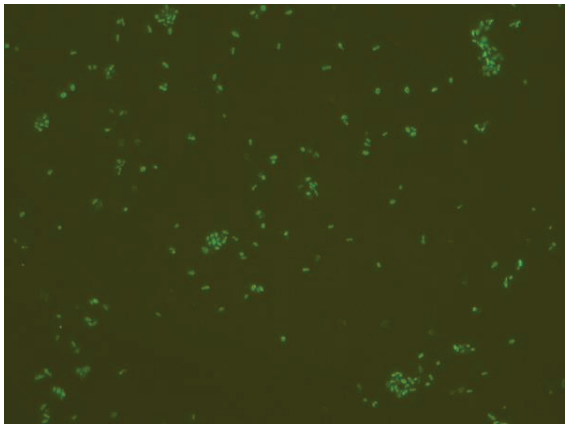
Appendix B: Genomic content of uncultured *Bacteroidetes* from contrasting oceanic provinces in the North Atlantic Ocean

Appendix C: Dilution cultivation of marine heterotrophic bacteria abundant after a spring phytoplankton bloom in the North Sea

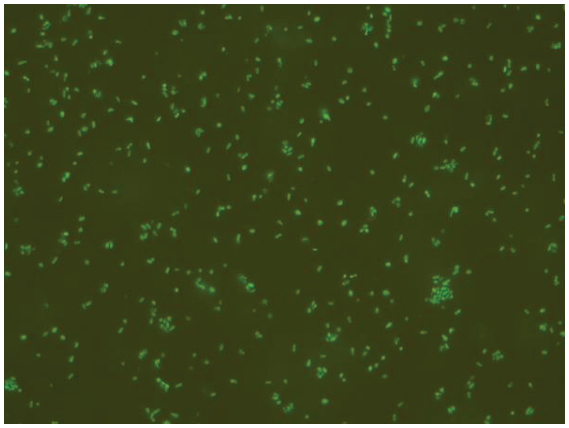
Appendix A: Probe Evaluation

Within this study new oligonucleotide probes targeting specific Bacteroidetes clades were designed. Probe design including competitors were done using the ARB probe design tool on basis of the SILVA NR 104 database release October 2010 (PRUESSE ET AL. 2007) in accordance to HUGENHOLTZ ET AL. 2001. Helpers were designed according to FUCHS ET AL. 2000. Horseradish peroxidase labeled oligonucleotide probes, unlabeled competitors and helpers were provided as lyophilized powder by Biomers (Biomers.net GmbH, Ulm, Germany). Oligonucleotides, competitors and helpers were dissolved in nuclease free water to a concentration of 8.42 pmol μl^{-1}.

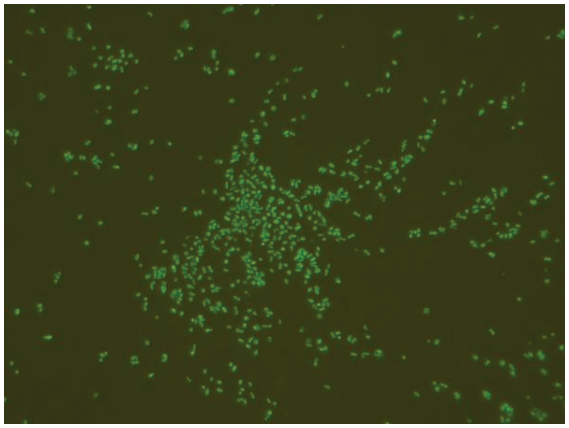
The oligonucleotide probes targeting the Formosa clade A and B were tested on pure cultures, which were obtained during the PhD of R. L. Hahnke. A formamide concentration series from 10–60% was performed and the optimal probe formamide concentration was determined by using the program *daime* 2.0 (DAIMS ET AL. 2006, Environmental Microbiology 8, 200-213).



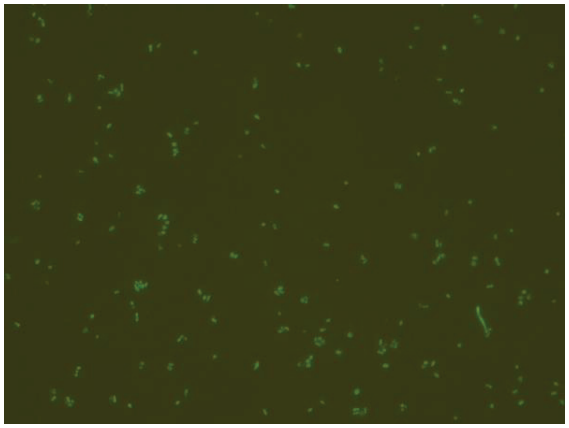
10%



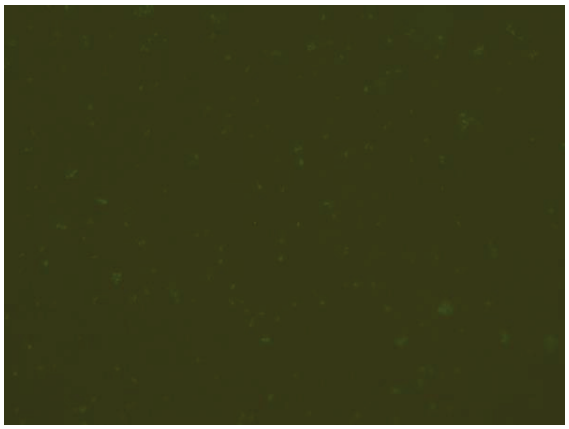
20%



30%

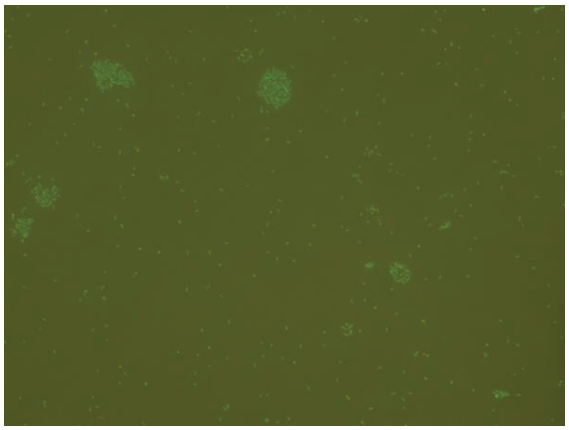


40%

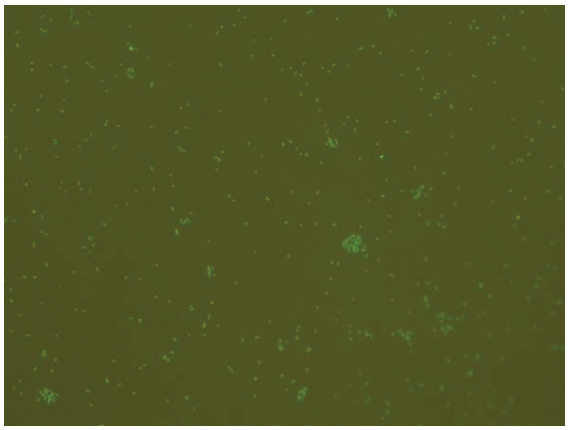


50%

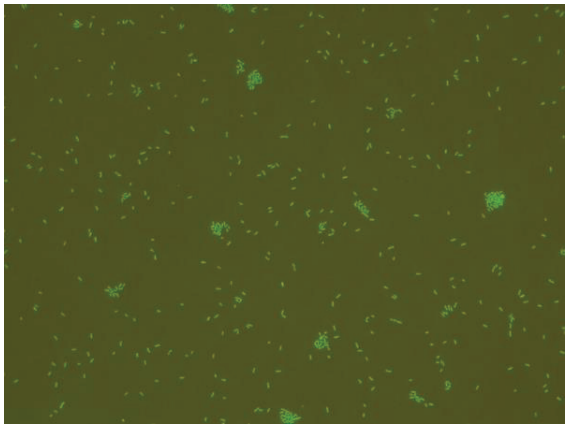
Picture gallery of epifluorescence micrographs showing the signal intensity obtained by probe FORM181A at different formamide concentrations when tested on a pure culture of Formosa spp. Hel3_A1_48 (provided by R. Hahnke).



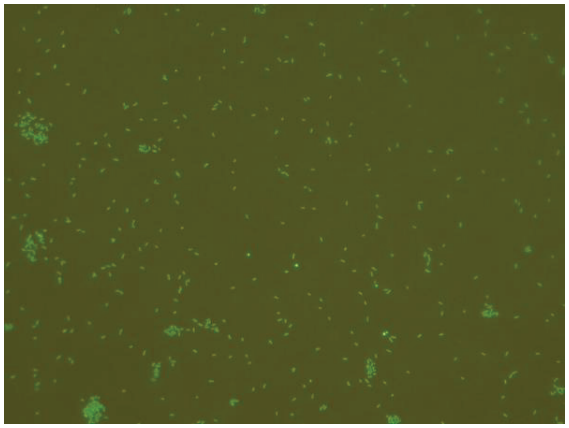
10%



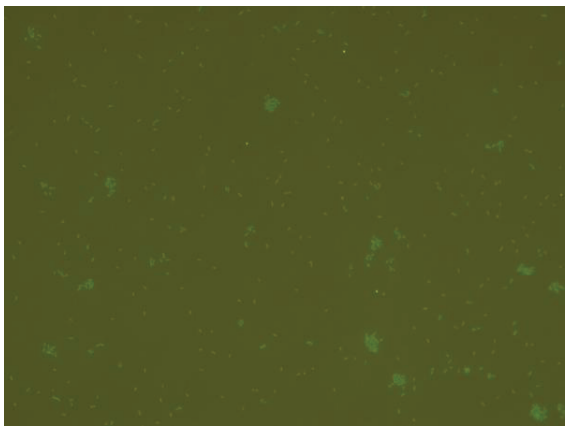
20%



30%

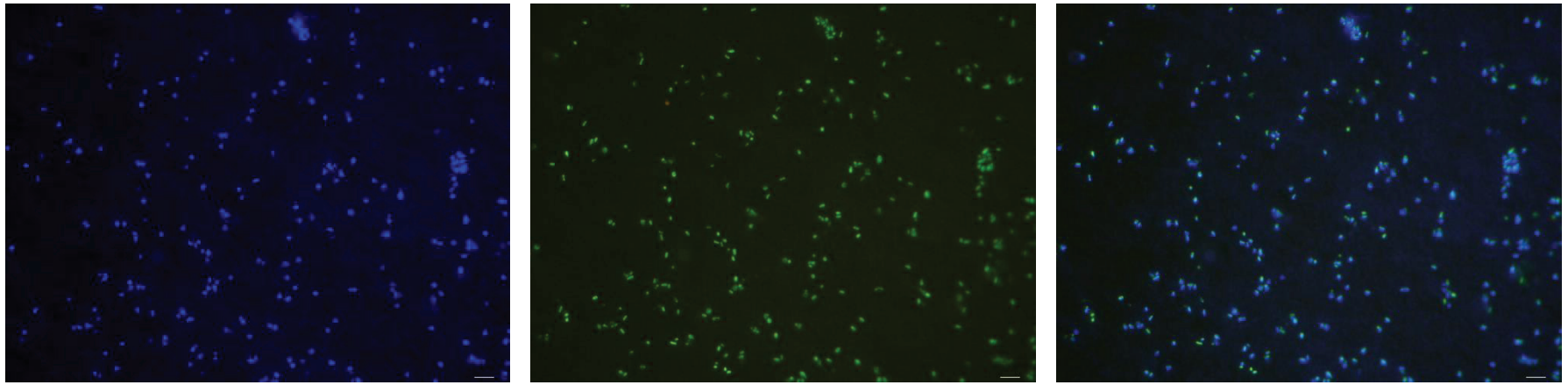


40%

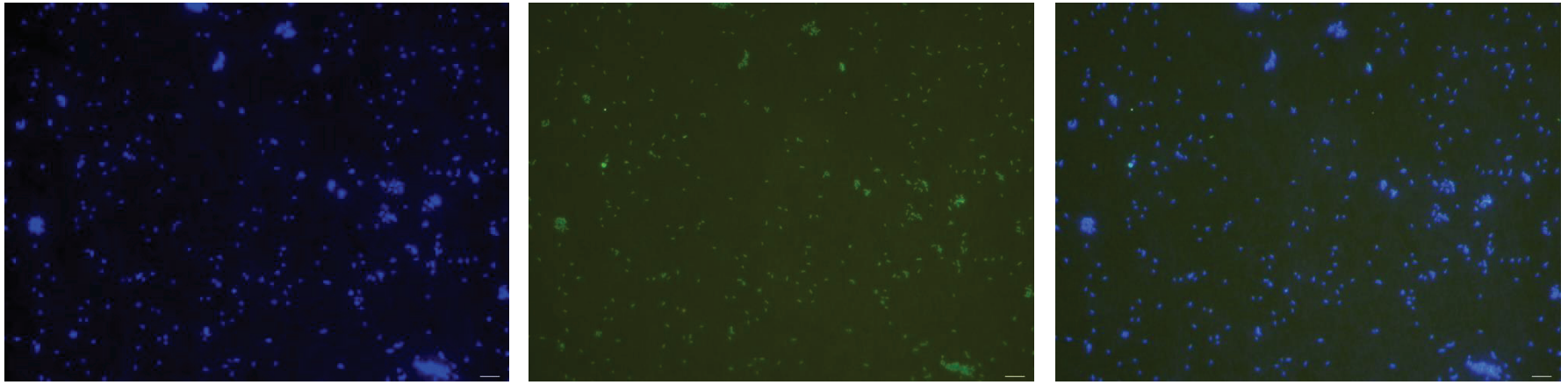


50%

Picture gallery of epifluorescence micrographs showing the signal intensity obtained by probe FORM181B at different formamide concentrations when tested on a pure culture of Formosa spp. Hel1_33_131 (provided by R. Hahnke).

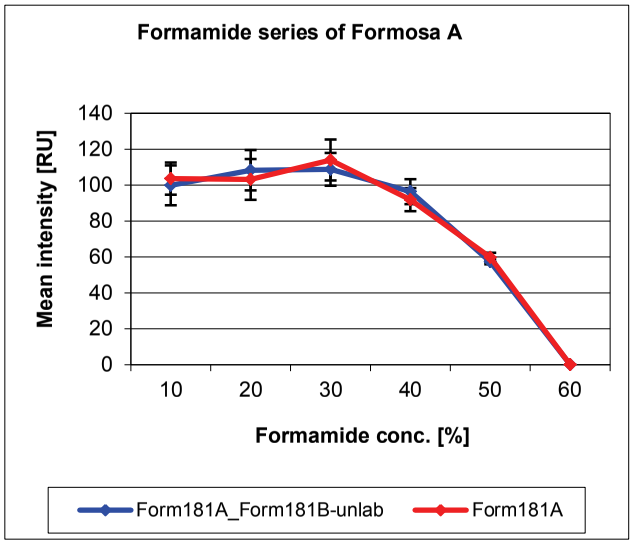


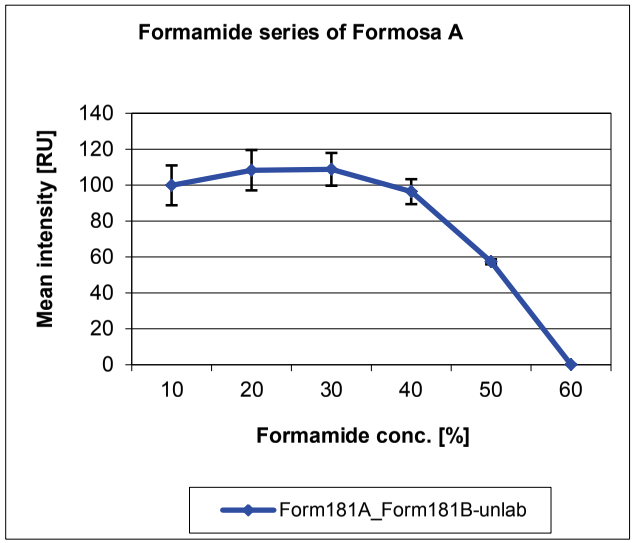
Form181A-488 on Formosa A (Hel_48) 20% FA (exposure time: DAPI = 301 ms; FISH = 150 ms)

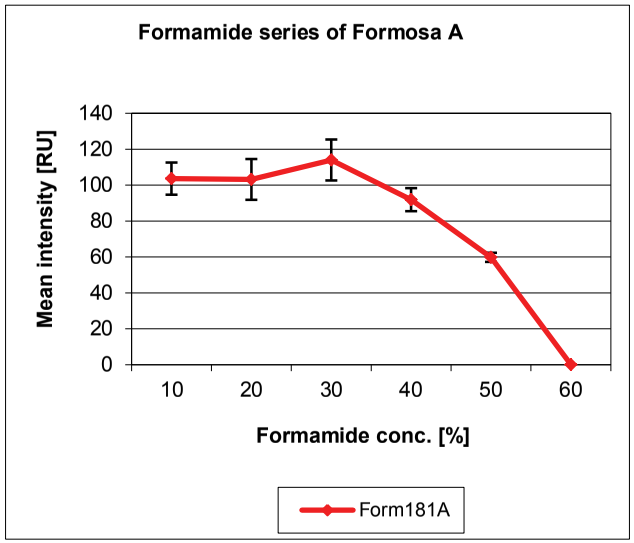


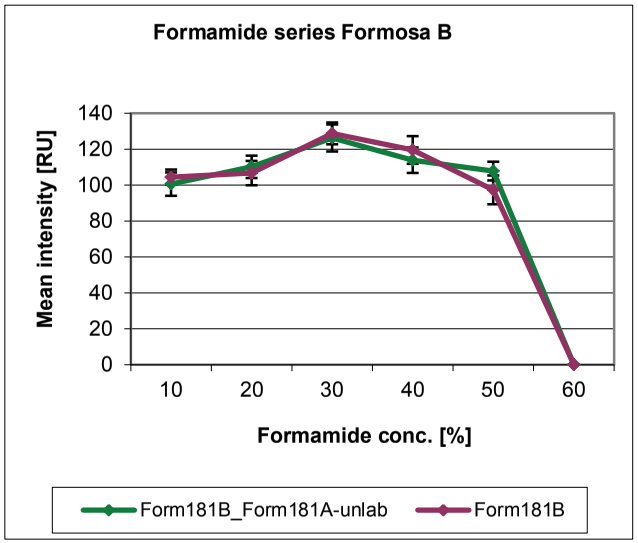
Form181B-488 on Formosa B (Hel_131) 30% FA (exposure time: DAPI = 174 ms; FISH = 174 ms)

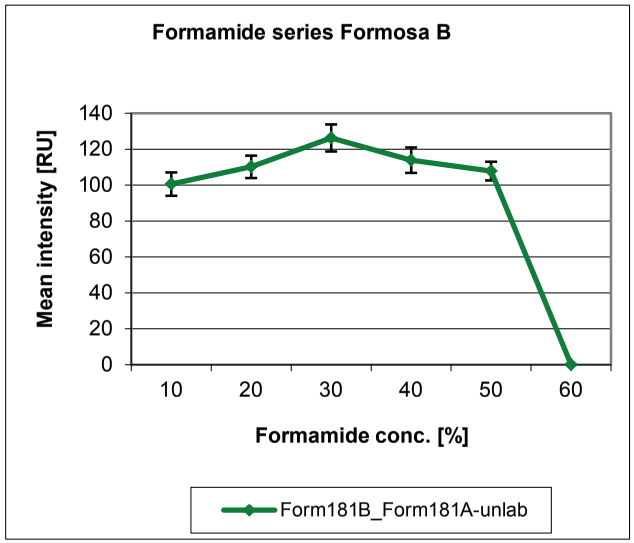
Table 1: Melting curves of probes targeting *Formosa* clade A (probe: FORM181A) and *Formosa* clade B (probe: FORM181B). Both probes could also be used as a competitor. This means FORM181B-unlabled were used as a competitor together with probe FORM181A. FORM181A-unlabled are used as competitor together with probe FORM181B. This was done to exclusively target the specific clade. Both competitors were tested if they would hinder the binding of the specific probe.

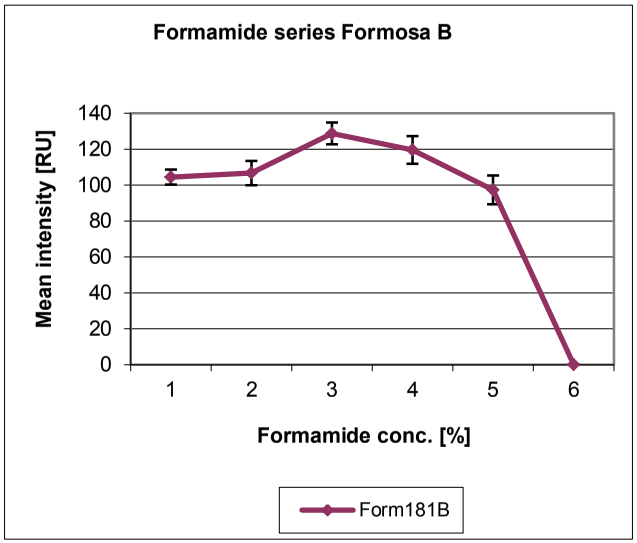












Appendix B

Genomic content of uncultured *Bacteroidetes* from contrasting oceanic provinces in the North Atlantic Ocean

Gómez-Pereira, P.R., Schüler, M., Fuchs, B.M., **Bennke, C.**, Teeling, H.,Waldmann, J., Richter, M., Barbe, V., Bataille, E., Glöckner, F.O. & Amann, R.

Published 2012 in *Environmental Microbiology* **14**, 52-66

Contribution to the manuscript: C. M. B. performed CARD-FISH and obtained in-situ abundance, preparation of figure 2

Abstract

Bacteroidetes are widespread in marine systems where they play a crucial role in organic matter degradation. Whole genome analysis of several strains has revealed a broad glycolytic and proteolytic potential. In this study, we used a targeted metagenomic approach to investigate the degradation capabilities of distinct *Bacteroidetes* clades from two contrasting regions of the North Atlantic Ocean, the Polar Biome (BPLR) and the North Atlantic Subtropical (NAST). We present here the analysis of 76 *Bacteroidetes* fosmids, of which 28 encode the 16S rRNA gene as phylogenetic marker, and their comparison to complete *Bacteroidetes* genomes. Almost all of the 16S rRNA harbouring fosmids belonged to clades that we previously identified in BPLR and NAST. The majority of sequenced fosmids could be assigned to *Bacteroidetes* affiliated with the class *Flavobacteria*. We also present novel genomic information on the classes *Cytophagia* and *Sphingobacteria*, suggesting a capability of the latter for attachment to algal surfaces. In our fosmid set we identified a larger potential for polysaccharide degradation and cell surface attachment in the phytoplan ton-rich BPLR. Particularly,

two flavobacterial fosmids, one affiliated with the genus *Polaribacter*, showed a whole armoury of enzymes that likely function in degradation of sulfated polysaccharides known to be major constituents of phytoplankton cell walls. Genes involved in protein and peptidoglycan degradation, although present in both fosmid sets, seemed to have a slight preponderance in NAST. This study provides support for the hypothesis of a distinct specialization among marine *Bacteroidetes* for the degradation of certain types of polymers.

Appendix C

Dilution cultivation of marine heterotrophic bacteria abundant after a spring phytoplankton bloom in the North Sea

Richard L. Hahnke, Christin M. Bennke, Bernhard M. Fuchs, Alexander J. Mann, Erhard Rhiel, Hanno Teeling, Rudolf Amann & Jens Harder

Manuscript in revision *Environmental Microbiology*

Contribution to the manuscript: C. M. B. probe design and testing, CARD-FISH and obtaining in-situ abundance, preparation of figure and assitance in manuscript writing

Abstract

The roles of individual bacterioplankton species in the remineralization of algal biomass are poorly understood. Evidence from molecular data had indicated that a spring diatom bloom in the German Bight of the North Sea in 2009 was followed by a rapid succession of uncultivated bacterioplankton species, including members of the genera *Ulvibacter*, *Formosa*, *Polaribacter* (class *Flavobacteria*) and *Reinekea* (class *Gammaproteobacteria)*. We isolated strains from the same site during the diatom bloom in spring 2010 using dilution cultivation in an artificial seawater medium with micromolar substrate and nutrient concentrations. Flow cytometry demonstrated growth from single cells to densities of 10^4–10^6 cells mL^{-1} and a culturability of 35%. Novel *Formosa*, *Polaribacter*, and *Reinekea* strains were isolated and had 16S rRNA gene sequence identities of >99.8% with bacterioplankton in spring or summer 2009. Selected isolates were draft sequenced and used for read recruitment of metagenomes from bacterioplankton in 2009. Metagenome reads covered 93% of a *Formosa* clade B, 91% of a *Reinekea*, and 74% of a *Formosa* clade A genome, applying a ≥94.5% nucleotide identity threshold. These novel strains represent abundant bacterioplankton species thriving on coastal phytoplankton blooms in the North Sea.

Dankesworte

Die letzten Worte dieser Arbeit möchte ich all jenen widmen, ohne deren Mithilfe und Unterstützung die vorliegende Schrift in dieser Form nicht hätte entstehen können.

Allen voran danke ich Herrn Prof. Dr. Rudolf Amann, der es mir mit der Vergabe des Arbeitsthemas und einer stets gutwissenden und versierten Betreuung mit vielen Anregungen und wertvollen Hinweisen erst ermöglichte, diese Arbeit vorzulegen.

Vielen Dank an Dr. Gunnar Gerdts für die Übernahme des Zweitgutachtens und für die Betreuung auf Helgoland. Gunnar, ohne deine Unterstützung wäre es mir nicht möglich gewesen, die zahlreichen Proben an „Land zu ziehen". An dieser Stelle möchte ich auch die Gelegenheit nutzen um Dr. Antje Wichels zu danken. Sehr gern war ich bei euch auf Helgoland. Des Weiteren möchte ich mich bei Prof. Dr. Karen Wiltshire für die Bereitstellung der physikochemischen und Phytoplankton Daten bedanken.

Ganz besonderer Dank für die fachliche und persönliche Unterstützung gilt Herrn Dr. habil. Bernhard Fuchs, der mir mit kritischen Anmerkungen und fundiertem Fachwissen stets ratsam zur Seite stand. Vielen Dank für die stehts offene Tür und die hilfreichen Diskussionen. Bernhard, ich konnte eine Menge von dir lernen. Du hast dich nie gescheut, die verschiedenen Versionen der Manuskripte zu lesen und hast mir stehts das Gefühl gegeben das meine Bemühungen nicht um sonst waren.

An dieser Stelle möchte ich mich auch bei Prof. Dr. Ulrich Fischer für die Teilnahme an meinem Prüfungskomitee sowie an meinem letzten Doktoranden Kolloquium bedanken. Weiterhin möchte ich mich bei Prof. Dr. Jens Harder für die gute Zusammenarbeit und fachliche Unterstützung bedanken.

Ebenfalls danke ich Dr. Hanno Teeling für die fachliche und persönliche Unterstützung mit der Bioinformatik. Hanno, vielen Dank, du hast mir das Feld der Bioinformatik etwas näher gebracht und mir gezeigt, das auch wir Biologen Spaß an solch kryptischen Dingen haben können.

Bedanken möchte ich mich bei allen technischen Mitarbeitern unserer Abteilung.
Jörg, vielen Dank für die Hilfe im Labor und beim Durchsehen der abertausenden automatischen Mikroskopbilder. Andreas, auch dir möchte ich danken, denn ohne deine Hilfe und Unterstüzung am konfokalem Laser Scanning Mikroskop wäre ich nie in der Lage gewesen, so tolle Aufnahmen zu gewinnen. Nicole, vielen Dank für die Unterstützung bei den Klonbanken, das wäre ohne deine Hinweise nie so reibungslos abgelaufen. Mirja, vielen Dank für die Hilfe beim mikroskopieren. Ich hatte eine tolle Zeit mit dir, wenn wir auf „Sternensuche" waren, oder mal wieder den „Mann im Mond" entdeckt haben. Sabine, dir möchte ich für die vielen Sequenzierung danken. Silke, auch wenn du zur Symbiose gehörst, deine Tipps und Tricks waren sehr hilfreich.

Ein herzliches Dankeschön richtet sich an die gesamte Abteilung „Molekulare Ökologie", denn jeder Einzelne von euch hat dazu begetragen, dass meine Zeit hier am Institut unvergesslich bleibt.

Nicht zu guter letzt danke ich all jenen die mir während meiner Zeit in Bremen, fortwährend mit Rat und Tat zur Seite standen. Vielen Dank Greta, Marion, Nicole, Mina, Stefan, Christina, Robert, Richard und sicherlich ganz vielen Weiteren, die es mir verzeihen mögen, sollte ich es vergessen haben sie persönlich zu erwähnen.

Mein ganz besonderer Dank gilt vor allem meiner Familie, meinem Freund und all jene, die mir während entscheidender und zum Teil diffiziler Phasen des Studiums und dieser finalen Arbeit hilfreich zur Seite standen!

Name: Christin M. Bennke
Anschrift: Puschkinring 56, 17491 Greifswald

Ort, Datum: Bremen, 28. Februar 2014

Erklärung

Hiermit erkläre ich, dass ich die Arbeit mit dem Titel:

Distribution and Function of marine *Bacteroidetes*

selbstständig verfasst und geschrieben habe und außer den angegebenen Quellen keine weiteren Hilfsmittel verwendet habe.

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

.....

(Unterschrift)