

Mechanisms of polysaccharide degradation of particle-
associated microbial communities

I n a u g u r a l d i s s e r t a t i o n

zur

Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Universität Greifswald

vorgelegt von

Doreen Schultz

geboren am 28.06.1991

in Siegen-Weidenau

Schillsdorf, den 28.01.2022

Dekan: Prof. Dr. Gerald Kerth

1. Gutachter: Prof. Dr. Katharina Riedel

2. Gutachter: Prof. Dr. Andreas Tholey

Tag der Promotion: 27.01.2022

„Ohne Fleiß, kein Preis.“
Für den besten Opa der Welt.
Ich vermisse dich!



Table of contents

Abbreviations.....	VI
1. Abstract	- 1 -
2. Zusammenfassung	- 2 -
3. Introduction	- 3 -
3.1. Marine carbon cycle	- 3 -
3.1.1. Microbial loop	- 3 -
3.1.2. Biological carbon pump.....	- 4 -
3.1.3. Phytoplankton blooms.....	- 5 -
3.1.4. Marine microorganisms and climate change.....	- 7 -
3.2. Marine particles and associated microbial communities	- 8 -
3.2.1. Particle-associated microbial communities.....	- 9 -
3.3. Marine polysaccharides and their degradation	- 10 -
3.3.1. Heterotrophic bacteria involved in polysaccharide degradation and their polysaccharide utilization loci (PULs).....	- 12 -
3.3.2. Carbohydrate-active enzymes (CAZymes).....	- 16 -
3.3.3. TonB-dependent transporter (TBDT).....	- 17 -
3.4. Metaproteomics	- 18 -
3.4.1. Typical workflow for sample preparation and quantification.....	- 19 -
3.4.2. Challenges and bottle necks for metaproteomic analysis of marine samples-	20
-	-
4. Scope of the thesis	- 21 -
5. Material and methods	- 23 -
5.1. Material	- 23 -
5.1.1. Software	- 23 -
5.1.2. Databases	- 24 -
5.1.3. Laboratory equipment	- 24 -
5.1.4. Chemicals, kits and enzymes	- 25 -
5.1.5. Consumables.....	- 26 -
5.2. Methods	- 27 -
5.2.1. Metaproteomic analysis of filters.....	- 27 -
5.2.2. Proteomic analysis of a selected marine particle-associated bacterial isolate grown on different polysaccharides	- 39 -
6. Results	- 45 -
6.1. Establishment of a metaproteomic pipeline for PA microbial communities (Schultz <i>et al.</i> 2020).....	- 45 -
6.1.1. Protein extraction	- 45 -

Table of contents

6.1.2.	MS sample preparation	- 49 -
6.1.3.	Optimization of databases.....	- 50 -
6.2.	Comparative metaproteome analysis of FL and PA bacterioplankton harvested during 2009 spring bloom	- 52 -
6.2.1.	Taxonomic differences between FL and PA fractions	- 53 -
6.2.2.	PA specific functional differences between FL and PA bacterioplankton	- 56 -
6.3.	Investigation of phylogenetic and functional succession of marine particles during 2018 spring phytoplankton bloom	- 58 -
6.3.1.	PA bacterioplankton taxonomic succession during the 2018 spring phytoplankton bloom	- 61 -
6.3.2.	Eukaryotic taxonomic succession during the 2018 spring phytoplankton bloom.	- 66 -
6.3.3.	PA bacterioplankton functional succession during the 2018 spring phytoplankton bloom	- 70 -
6.3.4.	Eukaryotic functional succession of phytoplankton during the 2018 spring phytoplankton bloom.....	- 72 -
6.4.	Proteomic analysis of a selected PA bacterial isolate grown on laminarin and alginate	- 74 -
6.4.1.	<i>Growth comparison of Muricauda sp. MAR_2010_75</i>	- 75 -
6.4.2.	Proteomic analysis of <i>Muricauda sp. MAR_2010_75</i> grown on different polysaccharides and glucose.....	- 76 -
7.	Discussion.....	- 83 -
7.1.	Establishment of a metaproteomic pipeline for PA microbial communities.....	- 83 -
7.2.	Comparative metaproteome analysis of FL and PA bacterioplankton fractions harvested during 2009 spring phytoplankton bloom	- 85 -
7.3.	Investigation of phylogenetic and functional successions of marine particles during the 2018 spring phytoplankton bloom	- 90 -
7.3.1.	Phylogenetic succession of marine particles during the 2018 spring phytoplankton bloom.....	- 91 -
7.3.2.	Functional succession of marine particles during the 2018 spring phytoplankton bloom	- 99 -
7.4.	Proteomic analysis of a selected PA bacterial isolate grown on laminarin and alginate	- 113 -
8.	Conclusions and Outlook.....	- 120 -
References	- 123 -
9.	Appendix	- 144 -
9.1.	Supplemental Material.....	- 144 -
9.2.	Acknowledgements	- 152 -
9.3.	Curriculum vitae.....	- 153 -
9.4.	List of publications, posters and presentations	- 154 -

Table of contents

9.5. Eigenständigkeitserklärung	- 155 -
9.6. CD-ROM content	- 156 -

Abbreviations

∅	average
a. bidest.	Double distilled water (MS-water)
ATP	adenosine triphosphate
bp	base pairs
BCA	bicinchoninic acid
BLAST	Basic Local Alignment Search Tool
br	biological replicate
BSA	bovine serum albumin
c	concentration
CARD-FISH	catalyzed reporter deposition fluorescence in situ hybridization
CAZyme	carbohydrate-active enzyme
CBM	carbohydrate-binding module
CE	carbohydrate esterase
CO	carbon monoxide
CO ₂	carbon dioxide
COG	Cluster of Orthologous Groups of proteins
CTAB	cetyltrimethylammonium bromide
Da	dalton
DNA	deoxyribonucleic acid
DOM	dissolved organic matter
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
ENA	European Nucleotide Archive
EPS	extracellular polymeric substances
ESI	electrospray ionization
FDR	false discovery rate
FL	free-living
GeLC/MS	purification of proteins via 1D-SDS-gels with subsequent reverse phase liquid chromatography coupled with mass spectrometry
GH	glycoside hydrolase
GT	glycosyltransferase

Abbreviations

h	hour(s)
HAB	harmful algal bloom
HMM	Hidden Markov Model
HMW	high molecular weight
HPLC	high performance liquid chromatography
iBAQ	intensity based absolute quantification
kDa	kilo Dalton
LC-MS/MS	liquid chromatography-tandem mass spectrometry
LFQ	label free quantification
LMW	low molecular weight
LTQ	Linear Trap Quadrupole
M	molar
MIMAS	Microbial interactions in marine systems
MS/MS	tandem mass spectrometry
m/z	mass to charge ratio
NCBI	National Center for Biotechnology Information
NPP	net primary production
NSAF	Normalized Spectral Abundance Factor
OD _{xxxnm}	optical density at given nm
ORF	open reading frame
PA	particle-associated
PAGE	polyacrylamide gelelektrophoresis
Pg	penta gram
pH	<i>potentia hydrogenii</i>
PL	polysaccharide lyase
PMSF	phenylmethylsulfonyl fluoride
POM	particulate organic matter
ppm	parts per million
PUL	polysaccharide utilization locus
PSM	peptide-spectrum match
riBAQ	relative iBAQ
RNA	ribonucleic acid
rRNA	ribosomal RNA

Abbreviations

rpm	revolutions per minute
RT	room temperature
RuBisCo	Ribulose-1,5-bisphosphate-carboxylase/-oxygenase
SDS	sodium dodecyl sulfate
sp.	species
TBDR	TonB dependent receptor
TBDTs	TonB dependent transporter
TCA	trichloro acetic acid
TEP	transparent extracellular particles
tr	technical replicate
Tris	tris(hydroxymethyl)aminomethane
(v/v)	volume percent
(v/w)	weight per volume

1. Abstract

A significant fraction of the decaying algal biomass in marine ecosystems is expected to be mineralized by particle-associated (PA) heterotrophic bacterial communities, which are thus greatly contributing to large-scale carbon fluxes. Whilst numerous studies have investigated the succession of free-living (FL) marine bacteria, the community structure and functionality of PA bacterial communities remained largely unexplored and knowledge on specific contributions of these microorganisms to carbon cycling is still surprisingly limited. This has mostly been due to technical problems, i.e., caused by the enormous complexity of marine particles and the high abundance of eukaryotic microorganisms within these particles. This thesis presents (a) an optimized metaproteomics protocol for an in-depth characterization of marine PA bacteria, (b) an application example with FL and PA communities sampled during a spring phytoplankton bloom in 2009 in the North Sea, which confirmed the reliability of the optimized metaproteomic workflow, (c) the metaproteomic analysis of particulate communities sampled during a spring phytoplankton bloom in 2018, resulting in an as yet unprecedented number of identified protein groups of the bacterial response bloom and (d) a proteomic analysis of a PA bacterial isolate grown on the two naturally abundant marine polysaccharides laminarin and alginate. The observed succession of bacterial clades during metaproteomic analyses of the investigated blooms highlights individual niche occupations, also visible on genus level. Additionally, functional data shows evidence for the degradation of different marine polysaccharides e.g., laminarin, alginate and xylan supporting the important role of PA bacteria during the turnover of oceanic organic matter. Furthermore, most of the identified functions fit well with the current understanding of the ecology of an algal- or surface-associated microbial community, additionally highlighting the importance of phytoplankton-bacterial interactions in the oceans. More detailed insights into the metabolism of PA bacteria were gained by the proteomic characterization of a selected PA bacterial isolate grown on laminarin and alginate. Functional analyses of the identified proteins suggested that PA bacteria employ more diverse degradation systems partially different from the strategies used by FL bacteria.

2. Zusammenfassung

Ein signifikanter Teil zerfallender Algen-Biomasse im marinen Ökosystem wird von Partikel-assoziierten (PA) heterotrophen bakteriellen Gemeinschaften mineralisiert, welche dadurch einen wichtigen Beitrag zum marinen Kohlenstoff-Kreislauf leisten. Während zahlreiche Studien die Sukzession von frei-lebenden (FL) marinen Bakterien untersucht haben, ist die Struktur von PA Gemeinschaften und deren Funktionalität größtenteils unentdeckt und das Wissen über ihren spezifischen Beitrag zum Kohlenstoff-Kreislauf überraschend gering. Das liegt vor allem an technischen Problemen, wie der enormen Komplexität mariner Partikel und der hohen Abundanz eukaryotischer Mikroorganismen innerhalb dieser Partikel. Diese Dissertation beinhaltet (a) ein optimiertes Protokoll für die metaproteomische Analyse von marinen PA Gemeinschaften, (b) ein Anwendungsbeispiel mit Proben von FL und PA Gemeinschaften einer Frühjahrsalgenblüte im Jahr 2009, das die Verlässlichkeit des optimierten metaproteomischen Protokolls bestätigte, (c) die metaproteomische Analyse PA Gemeinschaften einer Frühjahrsalgenblüte im Jahr 2018, die zu einer bis jetzt beispiellos hohen Zahl identifizierter Proteingruppen führte, und (d) die proteomische Analyse eines PA Isolats, das auf zwei in der Natur abundanten Polysacchariden, Laminarin und Alginat, wuchs. Die beobachtete taxonomische Sukzession bakterieller Gruppen, ebenfalls sichtbar auf Genus-Level, hebt deren individuelle Anpassung an verschiedene Nischen im marinen Raum hervor. Zusätzlich gaben die funktionellen Daten Hinweise auf die Degradation verschiedener mariner Polysaccharide, wie Laminarin, Alginat und Xylan. Dies unterstreicht die wichtige Rolle von PA Bakterien innerhalb der Umsetzung organischer Materie. Außerdem unterstützten die meisten Funktionen der identifizierten Proteingruppen das aktuellen Verständnis der Ökologie von Algen- oder Oberflächen-assoziierten mikrobiellen Gemeinschaften. Diese Ergebnisse heben zusätzlich die Wichtigkeit von Interaktionen zwischen Algen und PA Bakterien in den Ozeanen hervor. Detailliertere Einblicke in den Metabolismus von PA Bakterien konnten durch die proteomische Analyse eines PA Isolats gewonnen werden. Die Ergebnisse dieser Analyse gaben Hinweise darauf, dass PA Bakterien diverse Degradationssysteme verwenden, die sich teilweise von denen der FL Bakterien unterscheiden.

3. Introduction

3.1. Marine carbon cycle

Key processes of the marine carbon cycle include among others the conversion of inorganic carbon (e.g. CO₂) to organic carbon during photosynthesis of phytoplankton, the subsequent release of dissolved organic matter (DOM) and particulate organic matter (POM) from phytoplankton, the consumption of phytoplankton biomass by zooplankton grazers as well as the microbial loop (3.1.1) (Buchan *et al.* 2014). Heterotrophic bacteria also contribute to the remineralization of organic nutrients to their inorganic forms, which are then available for phytoplankton species. Via the microbial carbon pump organic carbon is transformed into recalcitrant dissolved organic carbon (DOC) that resists further degradation and is sequestered in the oceans for thousands of years (Buchan *et al.* 2014) (Figure 1).

3.1.1. Microbial loop

A large part of the global net primary production (NPP) is reprocessed by bacterioplankton in the so called microbial loop (Figure 1). The microbial loop is a pathway during which dissolved organic carbon is returned to higher trophic levels via its incorporation into bacterial biomass (Azam 1998) and highlights the impact of heterotrophic marine bacteria on the global carbon cycle. The relevance of microorganisms to the ocean ecosystem can also be appreciated from their high number and biomass throughout the water column and the subsurface: their total number of cells is more than 10²⁹ (Flemming and Wuertz 2019; Azam and Malfatti 2007) and it is estimated that 90% of marine biomass is composed of microorganisms (Suttle 2007). Marine microorganisms form the basis of ocean food webs and thus global carbon and nutrient cycles by fixing nitrogen and carbon (Azam and Malfatti 2007). The sinking, deposition and burial of fixed carbon in form of particulate organic matter (POM) to marine sediments is a key, long-term mechanism to sequester carbon dioxide (CO₂) from the atmosphere (Figure 1).

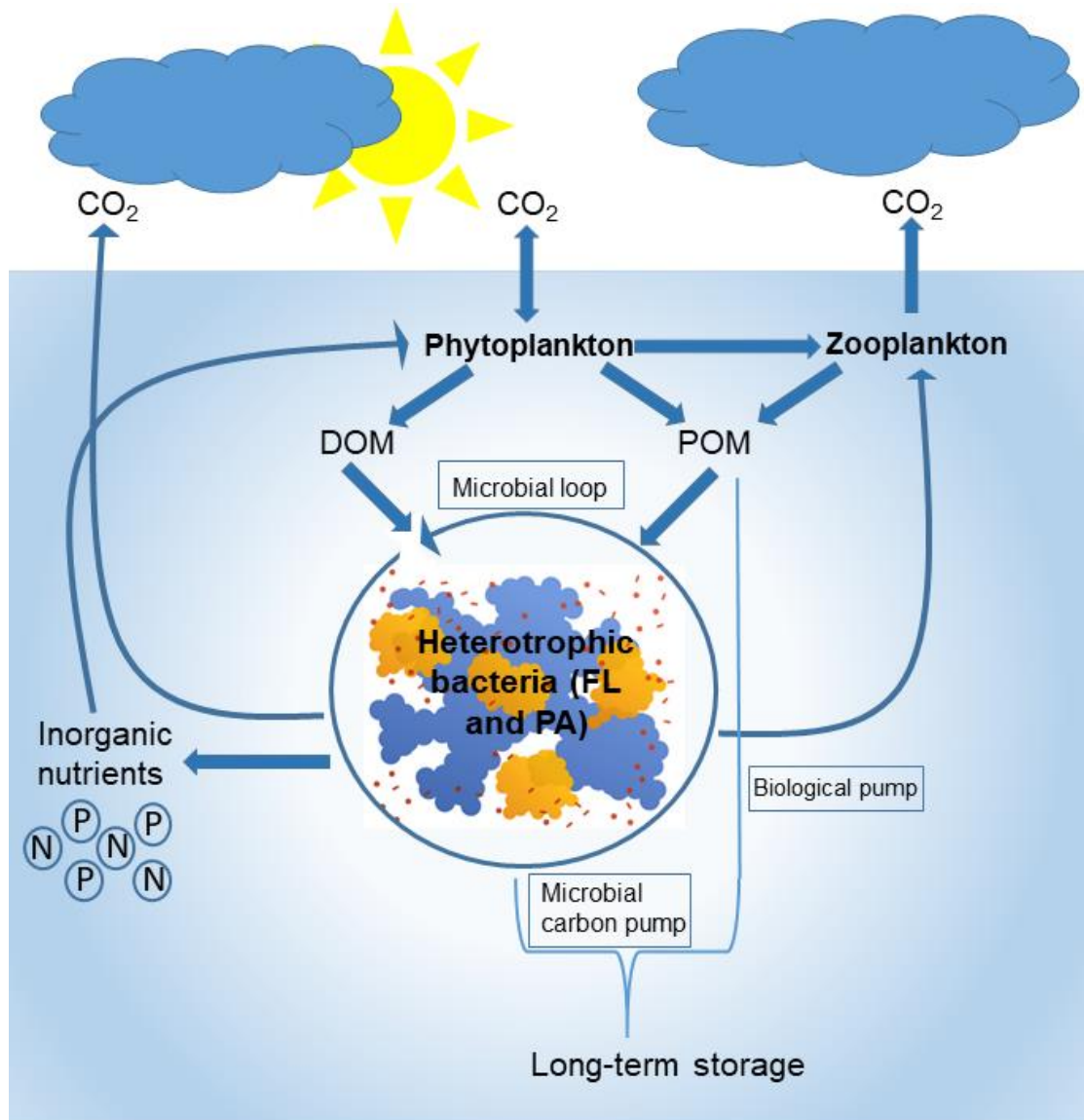


Figure 1. Marine carbon cycle. Used abbreviations: DOM (dissolved organic matter), POM (particulate organic matter), FL (free-living), PA (particle-associated). (adapted from Buchan *et al.* 2014)

3.1.2. Biological carbon pump

The biological carbon pump is the export of photosynthetically derived organic carbon from the sunlit surface ocean to deeper waters and the seafloor (Buchan *et al.* 2014). The result of this process is the depletion of nutrient concentrations in surface waters where there is net photosynthesis and elevation of nutrient concentrations at depth where there is net respiration (Sigman and Haug 2003; Sarmiento 2013). The interplay between the biological pump and ocean circulation generate important spatial gradients not only in nutrients but also in oxygen and dissolved organic carbon (Sarmiento 2013). Furthermore, this process driven carbon

storage in the deep ocean and thought to function via gravitational settling of organic particles as well as physical injection pumps and biologically mediated vertical fluxes (Boyd *et al.* 2019). However, the largest component of the biological pump is the sinking of particles (Buesseler *et al.* 2020). The interaction between bacteria and diatoms is suggested to enhance the efficiency of the biological pump through increased aggregate formation and particle sinking (Gärdes *et al.* 2011).

3.1.3. Phytoplankton blooms

Different to heterotrophic bacteria that depend upon autotrophic organisms for nutrition, photoautotrophic phytoplankton is able to convert radiant energy into biologically useful energy and synthesize metabolic compounds using only CO₂ as a source of carbon. Phytoplankton performs half of the global CO₂ fixation through photosynthesis which matches a global NPP of nearly 50 Pg C per year, and additionally performs half of the oxygen production despite the fact that they amount to only approximately 1% of the global plant biomass (Behrenfeld 2014). This can be explained by their distribution over a larger surface area, less exposure to seasonal variation and faster turnover rates compared to terrestrial plants (Behrenfeld 2014). Phytoplankton blooms recur in the oceans worldwide (Ye *et al.* 2011) and emerge when optimal growth conditions enable the exponential growth of fast-growing micro- or macroalgae (Smetacek and Zingone 2013). As the algal populations expand, the concentrations of essential nutrients decrease as the algae consume them. This depletion of essential nutrients together with viral infections and grazing by zooplankton are thought to cause the rapid decline of algal abundance. Such blooms can vary in overall phytoplankton composition, and the composition within an individual bloom can undergo multiple successive changes over its course. Some phytoplankton blooms happen periodically due to seasonal temperature and nutrient fluctuations, e.g. the spring phytoplankton blooms in the German Bight, which are often diatom-dominated (Chafee *et al.* 2018; Teeling *et al.* 2012). Diatoms are a class (*Bacillariophyceae*) of single-celled algae that have a silica-containing skeleton and are estimated to contribute nearly 20% of the global NPP (Armbrust 2009; Amin *et al.* 2012) and are thus one of the most important oceanic primary producers. Compared to other

phytoplankton groups, diatoms have relatively high sinking speeds, due to porous silica shells called frustules, which make them account for ~40% of particulate carbon export to depth (Tréguer *et al.* 2018; Boyd *et al.* 2019; Amin *et al.* 2012).

During phytoplankton blooms, heterotrophic bacterioplankton responds to successional releases of dissolved organic matter (DOM) as algal cells are lysed as a part of the microbial loop (Azam *et al.* 1983; Ducklow and Carlson 1992) (Figure 1). This process leads to rapid shifts in community structure on the order of weeks (Ward *et al.* 2017), or even days (Needham and Fuhrman 2016). Thereby individual bacterial species form response-blooms of their own, which are also characterized by a rapid growth and mortality (Needham and Fuhrman 2016; Teeling *et al.* 2012; Teeling *et al.* 2016). 16S ribosomal RNA (rRNA) gene-based studies have revealed recurring diversity patterns within these communities of free-living bacteria (Andersson *et al.* 2010; Gilbert *et al.* 2012; Fuhrman *et al.* 2015; Teeling *et al.* 2016; Chafee *et al.* 2018; Needham *et al.* 2018). Many of the bacterioplankton responding to spring bloom events are represented by fast-growing species, triggering the successions of specialized clades that are able to rapidly incorporate the algal-derived OM into their biomass by using diverse repertoires of specialized transporters and carbohydrate-active enzymes (CAZymes) (Thomas *et al.* 2011; Koropatkin *et al.* 2012; Teeling *et al.* 2012; Teeling *et al.* 2016; Buchan *et al.* 2014; Cuskin *et al.* 2015; Xing *et al.* 2015). Subsequently these organisms constitute a new pool of particulate bacterial biomass, which is available to protozoa and other higher trophic levels (Azam *et al.* 1983). Organic matter that is produced but not catabolized in turn is typically exported to sediments below. Heterotrophic members of the *Bacteroidetes*, *Gammaproteobacteria*, and the alphaproteobacterial *Roseobacter* clade are often among the dominant and recurring responders that partake in the degradation of marine phytoplankton-derived organic matter (Buchan *et al.* 2014) (see 3.3.1).

Beside algal blooms that happen periodically due to seasonal temperature and nutrient fluctuation also harmful algal blooms (HABs) can occur for example due to introduction of fertilized water into the oceans, where it unleashes the growth of algae (Smetacek and Zingone

2013; Liu *et al.* 2013). Although only 2% of all phytoplankton species are thought to produce HABs (Smayda 1997), these phenomena are occurring with increasing frequency and have a disproportionately large impact on natural ecosystems, public health and local economies (Hallegraeff 2003). Marine bacteria can both augment and buffer the influence of HABs. Algicidal bacteria for example lyse toxic phytoplankton species, leading to termination of HABs (Kodama *et al.* 2006). However, some bacterial species enhance the growth of HAB-forming species and can even increase the production of toxins (Bates *et al.* 1995). Therefore, it is important to get a better understanding of phytoplankton-bacteria interactions for example to be able to predict HAB and adopt preventive measures.

3.1.4. Marine microorganisms and climate change

Marine microbial communities are influenced by the availability of sunlight and nutrients, as well as the water temperature (Sunagawa *et al.* 2015). Rising temperatures due to climate change as one example, reduce the density of the water and consequently the circulation and stratification, which affect organismal dispersal as well as nutrient transport. The impact of climate change on ocean temperature, acidification, stratification, nutrient supply, extreme weather effects etc. affects marine microorganisms and leads to substantial environmental consequences as shifts in marine food webs, carbon export and its burial in the seabed (Hurd *et al.* 2018; Gao *et al.* 2012; Boyd 2013; Brennan and Collins 2015; Hutchins and Fu 2017; Rintoul *et al.* 2018). Additionally, marine phytoplankton, due to its distribution over a large surface area and fast turnover rates, responds rapidly on a global scale to climate variations (Cavicchioli *et al.* 2019). Therefore, it is important to learn how phytoplankton and marine microorganisms will respond to climate change (Riebesell and Gattuso 2015). Predicting the effect of climate change on phytoplankton and marine microorganisms is complicated, due to for example phytoplankton blooms, that are affected by both bottom-up (e.g. availability of nutrients, vertical mixing) and top-down control (e.g. viruses and grazers) (Behrenfeld 2014; Hutchins and Boyd 2016; Behrenfeld *et al.* 2016). Furthermore the lack of long-term data on phytoplankton production and microbial community composition (Cavicchioli *et al.* 2019) additionally hampers the prediction of climate change effects. To date only few such datasets

exist, such as the Hawaii Ocean Time-Series and the Bermuda Atlantic Time-series study (Dore *et al.* 2009; Saba *et al.* 2010; Buttigieg *et al.* 2018) but also the Helgoland Roads Long-Term Ecological Research Site in the German Bight (Wiltshire and Manly 2004). Additionally the Global Ocean Sampling Expedition (Rusch *et al.* 2007), transects of the Southern Ocean (Brown *et al.* 2012; Wilkins *et al.* 2013) and the Tara Oceans Consortium (Sunagawa *et al.* 2015; Brum *et al.* 2015; Vargas *et al.* 2015; Lima-Mendez *et al.* 2015; Guidi *et al.* 2016) provide valuable data on marine microorganisms.

3.2. Marine particles and associated microbial communities

Organic substrates remineralized by heterotrophic bacteria are not evenly distributed in the oceans. In this large oligotrophic environments marine particles form temporary nutrient-rich hot spots (Azam 1998; Simon *et al.* 2002; Stocker 2012). As a results, different heterotrophic lifestyles evolved in the oceans, namely (a) non-motile bacteria adapted to low nutrient concentrations, (b) motile bacteria sensing nutrient gradients and moving towards them, and (c) bacteria attached to or gliding on POM (Stocker 2012; Yawata *et al.* 2014). The latter two are referred to as particle-associated (PA) bacteria while the first ecotype is referred to as free-living (FL) bacteria.

Harsh environmental conditions such as high and low availability of nutrients (Stewart and Franklin 2008), as well as other environmental stressors like high cell densities (Prakash *et al.* 2003), have been shown to promote aggregate formation in phylogenetically distinct microorganisms. Marine particles consist of various kinds of organic matter, i.e. living or dead/dying zoo- or phytoplankton, bacterioplankton, as well as inorganic small particles held together by a sugary matrix. This matrix mainly consists out of transparent extracellular particles (TEP) composed of polysaccharides, which are exuded mostly by phytoplankton but also by bacteria (Alldredge *et al.* 1993). Marine particles grow while sinking and thus contribute largely to the biological pump (Figure 1), by transporting carbon to deeper waters and sediments (Volkman and Tanoue 2002). These aggregates can reach several centimetres in diameter. Additionally the exudates of living algae provide a continuous substrate source for

heterotrophic bacteria, resulting in the establishment of a phycosphere microbiome (Seymour *et al.* 2017).

3.2.1. Particle-associated microbial communities

20% of marine bacteria live attached to marine particles (Azam *et al.* 1983) and have a high impact on biogeochemical element cycles, due to their high metabolic activity (Smith *et al.* 1992; DeLong *et al.* 1993; Grossart *et al.* 2007; Ziervogel and Arnosti 2008; Ziervogel *et al.* 2010). These PA microbial communities have adapted to thrive and survive marine environments. While some bacteria are only loosely associated with algae, others colonize algal surfaces (Grossart 1999), where they form commensalistic or symbiotic communities with their host or even forage on algae (Sohn *et al.* 2005; Amin *et al.* 2012). PA bacteria are enzymatically well equipped to metabolize high molecular weight (HMW) substrates, thus providing nutrition to the attached community as well as leaving smaller carbon compounds to the surrounding water column community (Simon *et al.* 2002; Grossart 2010). Previous studies on PA bacteria have shown that these communities differ taxonomically from their FL counterparts (Bižić-Ionescu *et al.* 2014) and express specific genes to adapt to the availability of specific polysaccharides such as insoluble glycan fibers (Ganesh *et al.* 2014).

The most common technique to separate PA bacteria from their FL counterparts is sequential filtration (Bidle and Fletcher 1995; Crump *et al.* 1999; Ayo *et al.* 2001; Crespo *et al.* 2013; Bižić-Ionescu *et al.* 2014) with 0.2 µm pore-sized filters as most frequently used pore size to collect FL bacteria and 3 µm pore-sized filters as most frequently used pore size to collect PA bacteria in coastal waters (Eloe *et al.* 2011; Teeling *et al.* 2012; Crespo *et al.* 2013; D'Ambrosio *et al.* 2014; Teeling *et al.* 2016; Chafee *et al.* 2018). This separation method results in a FL fraction (3-0.2 µm) containing most unicellular cyanobacteria, members of the SAR11 and SAR86 clades and the flavobacterial marine group NS5 (Teeling *et al.* 2012; Teeling *et al.* 2016; Mestre *et al.* 2020) and a PA fraction (> 3 µm) enriched in *Rhodobacterales*, *Alteromonadales*, *Bacteroidetes*, *Planctomycetes* and *Verrucomicrobia* (Mestre *et al.* 2017). However, it is

important to notice that marine particles have a limited life span, forcing PA bacteria to spend some time FL on their way to the next particle (Bižić-Ionescu *et al.* 2014).

Our knowledge on specific physiological functions of PA microorganisms is, however, surprisingly limited due to other methodological problems, i.e., in the difficulty to retrieve proteins from these polysaccharide-rich entities.

Interactions between photoautotrophic phytoplankton and heterotrophic bacteria have been extensively studied in both natural environments and laboratory cultures (Buchan *et al.* 2014; Christie-Oleza *et al.* 2017; Amin *et al.* 2012; Azam and Malfatti 2007). Among these interactions, synergistic and antagonistic relationships have been commonly observed. It is generally thought that photoautotrophs and heterotrophs develop synergistic relationships due to nutrient cycling. Photoautotrophs provide labile OM for growth of heterotrophs, and in return benefit from the supply of essential micronutrients from heterotrophic bacteria e.g. vitamins and bioavailable trace metals (Amin *et al.* 2012; Christie-Oleza *et al.* 2015; Kazamia *et al.* 2012) as well as the removal of reactive oxygen species (ROS) (Morris *et al.* 2008; Morris *et al.* 2011). Besides mutualistic interactions, competition and antagonistic behaviors could also be observed between photoautotrophs and heterotrophs. Such interactions include for example competition for essential inorganic nutrients like phosphate and nitrate as well as the production of algicidal compounds by heterotrophic bacteria (Shi *et al.* 2013; Crenn *et al.* 2018).

3.3. Marine polysaccharides and their degradation

Marine polysaccharides are a major carbon and energy source for heterotrophic bacteria (Teeling *et al.* 2012) as they can constitute for 50% and even up to 90% of algal dry matter depending on algae species and growth stage (Kraan 2012; Myklestad 1974). They are food for macro- and microorganisms recycling the carbon and energy stored in them. Microbial degradation of these diverse and abundant polysaccharide sources is regarded as a bottleneck in the marine carbon cycle, allowing transfer of organic matter to higher trophic levels (Azam and Malfatti 2007; Gasol *et al.* 2008; Buchan *et al.* 2014). Recalcitrant polysaccharides may

also escape rapid degradation and transport carbon via vertical sinking to the deep ocean thereby removing carbon for millennia from the atmosphere.

Marine polysaccharides are natural macromolecules consisting of monosaccharide sugars that are interconnected by glycosidic bonds. Monosaccharide composition is often dominated by glucose, mannose, fucose, arabinose, xylose, rhamnose and galactose (reviewed in Gügi *et al.* 2015; Mühlenbruch *et al.* 2018). Various modifications of polysaccharides exist, leading to diverse functions as storage compounds, cell wall constituents or as intra- or extracellular matrix components. Marine algae produce many types of polysaccharides with different complexities and varying physicochemical properties that are distinct from those of land plants (Popper *et al.* 2011; Popper and Tuohy 2010; Ficko-Blean *et al.* 2015). Green macroalgae contain ulvans, red macroalgae contain agars, carrageenans and porphyrans, brown algae contain alginates, fucans and laminarin, and diatoms contain chrysolaminarin and sulfated mannans (Popper *et al.* 2011). Their properties are mainly influenced by the length, the sugar composition, the position of the glycosidic bonds and the stereochemistry at the anomeric carbon (α/β conformation). Additionally, different modifications at the hydroxyl groups like alkylation, acylation, sulfation, dehydration or additional glycoside side chains can be observed in nature and lead to the characteristics of a specific polysaccharide (Hehemann *et al.* 2014; Percival 1979). However, so far only few structures of marine polysaccharides have been resolved (Gügi *et al.* 2015; Le Costaouëc *et al.* 2017; Jesus Raposo *et al.* 2014).

Phytoplankton blooms produce huge amounts of beta-glucans such as laminarin, a soluble β -1,3-glucan with β -1,6 side chains. Laminarin is a major food resource for heterotrophic marine bacteria and diatoms alone are estimated to produce 5-15 Gt of this polysaccharide per year (Alderkamp *et al.* 2007). Besides in diatoms, glucans of the laminarin type act as storage compound also in brown algae, *Chrysophyceae* and haptophytes such as *Phaeocystis* spp. (Beattie A., Hirst E. L., Percival E. 1961; Janse *et al.* 1996; Davis *et al.* 2003). Enzymes for the degradation of laminarin are abundant in ocean surface waters, but also within deeper parts

of the water column and in sediments (Keith and Arnosti 2001; Arnosti *et al.* 2005) highlighting the major role of this polysaccharide.

Another frequent polysaccharide in the oceans is alginate, which is a cell wall polysaccharide as well as an intracellular component in brown algae and can account for up to 40% of algal dry weight (Haug *et al.* 1966; Davis *et al.* 2003). Alginate is an unbranched copolymer of 1,4-linked β -D-mannuronate and α -L-guluronate that are arranged in alternating blocks.

Other components of algal organic matter include proteins, lipids and nucleic acids (Villacorte *et al.* 2015; Haug and Myklestad 1976) but also the osmolyte dimethylsulfoniopropionate (DMSP) (Yoch 2002), which plays an important role in the oceanic sulfur cycle. Moreover, methanol, which is hypothesized to be a product of polysaccharide demethoxylation (Mincer and Aicher 2016) and glycolate, which is a waste product of photorespiration (Schada von Borzyskowski *et al.* 2019; Tolbert N. E., Zill L. P. 1956) are part of algal OM. Furthermore, a greater variety of secondary metabolites, small molecules and humic substances are part of OM. Together they represent an important pool of carbon and thus energy for the heterotrophic community even if they might be less accessible than polysaccharides and proteins, due to greater diversity of structures among secondary metabolites and humic substance and greater resistance to degradation.

3.3.1. Heterotrophic bacteria involved in polysaccharide degradation and their polysaccharide utilization loci (PULs)

Recent studies have shown that marine polysaccharides can be used by diverse specialized bacterial taxa, including members of *Gammaproteobacteria*, *Planctomyces*, *Bacteroidetes* and *Verrucomicrobia* (Chafee *et al.* 2018; Wemheuer *et al.* 2015; Hahnke *et al.* 2015; Lage and Bondoso 2014; Martinez-Garcia *et al.* 2012; Teeling *et al.* 2012). The decomposition of polysaccharides is notably different in gram-negative bacteria than in gram-positive bacteria and *Archaea*. The latter two have to decompose polysaccharides to di- and monosaccharides that can subsequently be taken up by for example ABC-type transporters. Many gram-negative polysaccharide-degrading bacteria harbor so-called “polysaccharide utilization loci” (PULs)

(Sonnenburg *et al.* 2010; Martens *et al.* 2009). The first described PUL was the starch utilization system (Sus) of the human gut symbiont *Bacteroidetes thetaiotaomicron* (Shipman *et al.* 2000). PULs are operon-like arrangements that harbor genes for numerous functions necessary for the degradation of a polysaccharide (Martens *et al.* 2009; Hehemann *et al.* 2014; Hehemann *et al.* 2010; Grondin *et al.* 2017): (a) binding of the polysaccharide to the outer membrane, e.g. by SusD-like proteins, (b) initial degradation to oligosaccharides by endo-glycoside hydrolases (GHs) or polysaccharide lyases (PLs) at the outer membrane, (c) binding and transport of the resulting oligomers into the periplasmic space by TonB-dependent-receptors (TBDRs=SusC-like) (Foley *et al.* 2016; Bjursell *et al.* 2006), (d) processing the oligomers by several carbohydrate-active enzymes (CAZymes) into dimers or monomers, (e) transport of the dimers/monomers through the cytoplasmic membrane, (f) funneling the monomers into the basic metabolism through further degradation, e.g., through cleavage of decorating sulfate or acetate esters and demethylation by auxiliary CAZymes, and (g) proteins for sensing and regulation of gene expression. The size of a PUL correlates with the complexity of its target polysaccharide as a larger set of enzymes is required to degrade polysaccharides with several modifications and variations.

Well-known polysaccharide degraders, which are not only found in the marine environment but also in the gut of humans and animals, are bacteria from the gram-negative clade *Bacteroidetes* (Grondin *et al.* 2017). Previous studies showed high abundances of *Bacteroidetes* in the bacterioplankton during blooms of phytoplankton in the coastal East Antarctica (Williams *et al.* 2013) and in the coastal North Sea (Teeling *et al.* 2012). In recent years they have been recognized for their pivotal role in the decomposition of biopolymers, including polysaccharides (Cottrell and Kirchman 2000; Bauer *et al.* 2006; Kirchman 2002; González *et al.* 2008; González *et al.* 2011; Gómez-Pereira *et al.* 2010; Fernández-Gómez *et al.* 2013; Thomas *et al.* 2011). One typical feature observed in marine and gut *Bacteroidetes* is the so-called “selfish” polysaccharide utilization that describes the direct uptake of larger oligomers, produced by a first extracellular degradation (Cuskin *et al.* 2015; Reintjes *et al.* 2017). After this import the depolymerization to monosaccharides takes place in the periplasm,

which allows a protection of the oligosaccharide from surrounding bacteria in the diffusion-open environment. Oligosaccharides can be bound at the interface of SusCD complexes. It was recently shown, that these two proteins form a “pedal bin” complex in *Bacteroidetes thetaiotaomicron*, with SusD acting as a substrate binding lid on top of the SusC-like TonB-dependent transporter (TBDT). Upon binding of an oligosaccharide, the SusD lid closes and conformational changes lead to substrate release into the periplasm (Glenwright *et al.* 2017). However only the *Bacteroidetes* are known to have evolved SusCD complexes, for which it was recently proposed, that they do not only transport polysaccharides, but also other macromolecules such as peptides (Glenwright *et al.* 2017) or nucleic acids (Schauer *et al.* 2008).

The largest family of the *Bacteroidetes* phylum is the *Flavobacteriaceae* whose members thrive a wide variety of habitats. The members of this family are non-spore forming, rod-shaped, aerobic bacteria and commonly referred to as flavobacteria (Rosenberg *et al.* 2014). They are common in terrestrial and freshwater environments, as well as numerically predominant in marine habitats (Kirchman 2002). Marine flavobacteria have been found to establish an either FL or PA lifestyle. (Bennke *et al.* 2016; DeLong *et al.* 1993). In particular members of the class *Flavobacteriia* are associated with the degradation of high molecular weight (HMW) POM in marine environments, because they were often found enriched on detritus, and colonizing surfaces of higher plants (Crump and Koch 2008) as well as macroalgae (see e.g. Barbeyron *et al.* 2001; Martin *et al.* 2015; Nedashkovskaya *et al.* 2003; Nedashkovskaya *et al.* 2004a; Nedashkovskaya *et al.* 2004b; Nedashkovskaya *et al.* 2013; Nedashkovskaya *et al.* 2005; Nedashkovskaya *et al.* 2006; Ivanova and Dedysh 2006). Furthermore, this class was also found in close association with invertebrate animals such as corals (Sweet *et al.* 2011; Frias-Lopez *et al.* 2002; Rohwer *et al.* 2002), echinoderms (Romanenko *et al.* 2007) and sponges (Yoon and Oh 2012). Many marine flavobacteria degrade HMW macromolecules like complex polysaccharides and proteins and thus contribute to the carbon turnover in marine environments (Bennke *et al.* 2016; Mann *et al.* 2013; Barbeyron *et al.* 2016b). Additionally, several studies have indicated that sequenced genomes of marine *Flavobacteriaceae* feature

high proportions of PULs (Bennke *et al.* 2016; Mann *et al.* 2013; Barbeyron *et al.* 2016b; Fernández-Gómez *et al.* 2013; Kappelmann *et al.* 2018; Xing *et al.* 2015), reinforcing their adaptations towards biopolymer degradation.

It is suggested that due to the high structural diversity of polysaccharides and their wide variety of sugar monomers and moieties, polysaccharide-degrading bacteria specialize on polysaccharide subsets and their monomers and therefore harbor only a restricted set of polysaccharide-degrading enzymes. In most bacterial genomes CAZymes account for roughly 2% of the genes and seldom exceed 5% of the genes in bacteria, that are specialized on polysaccharide degradation (Mann *et al.* 2013; Coutinho *et al.* 2003; Henrissat and Coutinho 2001; Lairson *et al.* 2008). This leads to the hypothesis, that niche-adapted specialists, degrading dedicated polysaccharides and their sugar monomers, are necessary for the coordinated break down of complex polysaccharide mixtures in nature. This type of resource and niche partitioning could already be shown in human gut *Bacteroidetes* (Hehemann *et al.* 2012; Martens *et al.* 2011).

Teeling *et al.* (2012) investigated the free-living fraction of a spring diatom bloom and could identify marine *Flavobacteria* and *Gammaproteobacteria* as key polysaccharide degraders, which showed consecutive blooms of *Formosa* (*Flavobacteria*), *Reinekea* (*Gammaproteobacteria*) and *Polaribacter* (*Flavobacteria*) as most prominent clades with a relative abundance of 15-25% of total bacterial abundance. With a combination of metagenome and metaproteome analyses they could show a succession of protein functions (especially CAZymes and transporters), which indicates that the above-mentioned clades adapted to distinct polysaccharide substrates and that this succession of the clades is driven by step-wise degradation of complex algal polysaccharide structures.

3.3.1.1. The genus *Muricauda*

The genus *Muricauda* belongs to the class of *Flavobacteriia* of the Phylum *Bacteroidetes* and was e.g. identified as one of the most abundant genera in a *Synechococcus* cultivation experiment (Zheng *et al.* 2020). The in this thesis used strain *Muricauda* sp. MAR_2010_75

was isolated from phytoplankton in the North Sea at the island Sylt (Hahnke and Harder 2013) and is therefore referred to as PA. *Muricauda sp. MAR_2010_75* has 96,3% 16S rRNA gene identity to *Muricauda flavrescens*, forms round orange colonies when it grows on agar and shows gliding motility. Its predicted degradation capacity of polysaccharide classes based on PUL-associated CAZyme annotations is starch, glycogen, β -mannan, β -xylose, rhamnose and N-acetylglucosamin (Kappelmann *et al.* 2018).

3.3.2. Carbohydrate-active enzymes (CAZymes)

For utilization of polysaccharides heterotrophic bacteria need various specific enzymes, which are called carbohydrate-active enzymes (CAZymes), i.e., enzymes that synthesize, modify or cleave glycosidic bonds. CAZymes are classified into over 300 families and are listed in the CAZy.org database (Lombard *et al.* 2014). The degradative enzymes are grouped into *endo*- or *exo*-acting enzymes, depending on the cleavage position in the polysaccharide chain. *Endo*-enzymes cleave bonds within the chain and *exo*-enzymes catalyze the depolymerization from the ends of the chain. The following enzyme classes belong to the CAZymes:

Glycoside hydrolases. The class with the highest number of families within the CAZy database are the glycoside hydrolases (GHs) which catalyze the hydrolytic cleavage of the glycosidic bonds. GHs are involved in nearly all polysaccharide degradation processes and are often very specific for a certain oligosaccharide substructure within a polysaccharide. For example, GHs active on the β -glucan laminarin do not cleave in α -glucans like starch and vice versa (Roche-Mayzaud and Mayzaud 1987).

Polysaccharide lyases. Some polysaccharides like alginate, pectin or ulvan contain uronic acids that can be cleaved off via a non-hydrolytic elimination mechanism by polysaccharide lyases (PLs) (Garron and Cygler 2014).

Carbohydrate esterases. Carbohydrate esterases (CEs) remove additional ester linkages that can be found in some carbohydrates.

Carbohydrate-binding modules. Carbohydrate-binding modules (CBMs) increase the affinity of CAZymes to the polysaccharide.

Sulfatases. Unlike in terrestrial polysaccharides, sulfation is a common modification in marine carbohydrates (Helbert 2017). To utilize the through degradation resulting monomeric sugars in the cellular metabolism, heterotrophic bacteria need to remove these sulfate groups. This is achieved by type I polysaccharide sulfatases which were found to be highly expressed during microalgal blooms and showed high abundance in marine bacteria specialized on carbohydrate degradation (Teeling *et al.* 2012; Helbert 2017; Kappelmann *et al.* 2018; Mann *et al.* 2013; Barbeyron *et al.* 2016b). Known sulfatases are classified in the SulfAtlas database (Barbeyron *et al.* 2016a)

3.3.3. TonB-dependent transporter (TBDT)

It was previously shown that TBDRs, which are parts of the TonB-dependent uptake systems, were among the most abundantly expressed proteins during the bacterial remineralization of algal biomass during a North Sea spring phytoplankton bloom (Teeling *et al.* 2012; Francis *et al.* 2021). TBDRs seem to play a pivotal role in the uptake of oligosaccharides, which was also indicated by *in situ* metaproteome studies of samples from coastal upwelling zones in the South Atlantic Ocean (Morris *et al.* 2010) and from the Antarctic Southern Ocean (Williams *et al.* 2012). They are often found in *Bacteroidetes*, but also widely distributed among other gram-negative bacteria such as *Gammaproteobacteria* (Grondin *et al.* 2017; Hehemann *et al.* 2017; Neumann *et al.* 2015) or *Cyanobacteria* (Mirus *et al.* 2009). TBDT are known to transport bulky compounds such as iron-siderophores, nickel complexes, vitamin B₁₂, and oligosaccharides (reviewed in: Noinaj *et al.* 2010; Schauer *et al.* 2008). They consist of an outer membrane TBDR, that is linked to the cytoplasmic membrane via a TonB-ExbBD complex (Schauer *et al.* 2008). The TonB protein interlinks the components of the inner and outer membrane by β -strand pairing with the TBDR N-terminus. This connection is suggested to translate the proton motif force across the cytoplasmic membrane into a movement that opens or closes a pore formed by the TBDR via its plug domain (Krewulak and Vogel 2011; Noinaj *et al.* 2010). TBDRs

are suggested to be selective for specific substrates, due to their substrate-binding function and high sequence diversity (Schauer *et al.* 2008; Tang *et al.* 2012).

3.4. Metaproteomics

Through metagenomic studies valuable knowledge about diversity and distribution of microorganisms in natural environments can be gained. However, metagenomes do not allow linking microbial diversity and functionality. Moreover, most microbes overcome environmental changes by altering their gene/protein expression profiles and to a much lesser extent by genomic rearrangements such as horizontal gene transfer. Therefore, metatranscriptomic and metaproteomic approaches were established to be able to investigate, which genes are expressed at a given time point as well as which proteins are particularly abundant in complex biological systems. Metaproteomics, due to its opportunity to study many protein functions and responses simultaneously, has proven to be an excellent tool for improving our understanding of the complex interplay between environmental parameters, microbial community architecture and composition as well as ecosystem functioning. Various studies, describing e.g., large-scale proteome analyses of acid-mine drainage biofilms (Ram *et al.* 2005), wastewater treatment plants (Wilmes *et al.* 2008; Püttker *et al.* 2015; Salerno *et al.* 2016), and fresh-water stream biofilms (Hall *et al.* 2012) have demonstrated the success of metaproteomics to unveil molecular mechanisms involved in function, physiology, and evolution of surface-associated aquatic microbial communities. Marine metaproteomics has been widely applied (Saito *et al.* 2019; Wang *et al.* 2014), especially in habitats such as ocean scale shifts (Morris *et al.* 2010), the Atlantic (Bergauer *et al.* 2018) or Antarctic oceans (Williams *et al.* 2012), e.g. to investigate the *Roseobacter* clade (Christie-Oleza and Armengaud 2015) and bacterioplankton physiology (Wöhlbrand *et al.* 2017b). Recent ocean metaproteomics studies have provided new insights into microbial nutrient transport (Sowell *et al.* 2009; Morris *et al.* 2010), co-limitation of carbon fixation processes (Saito *et al.* 2014), the composition of microbial biofilms (Leary *et al.* 2014), dynamic of carbon flux in marine ecosystems (Moore *et al.* 2012; Moore *et al.* 2014; Bridoux *et al.* 2015) and seasonal shifts in microbial metabolic diversity (Georges *et al.* 2014).

3.4.1. Typical workflow for sample preparation and quantification

A metaproteomics experiment typically comprises three basic steps: (a) the sample preparation, including protein extraction, purification and tryptic digestion into peptides, (b) peptide separation and subsequent MS analysis and (c) peptide identification based on the obtained MS/MS spectra followed by an assignment of the peptides to protein groups in an appropriate database (Keiblinger *et al.* 2012). Therefore, raw mass spectrometry files, containing the mass spectra from peptide fragmentation, are searched against a database that contains the sequences of proteins which could theoretically be present in the sample and from which theoretical mass spectra are generated *in silico*. Various search algorithms have been developed to tackle this highly complex task (Verheggen *et al.* 2020).

(Meta)proteomic techniques also enable the quantification of proteins. One possibility is the so called label-free quantification, that can be based on MS/MS spectra assigned to a protein (spectral counting), or in the intensity of peptide precursor ions of a protein as calculated by the area under the curve or the peak height for the respective peptide (Nahnsen *et al.* 2013; Ankney *et al.* 2018). In spectral counting, the total number of tandem mass spectra that match peptides to a particular protein is used to measure the abundance of proteins in a complex mixture. Zybailov *et al.* 2006 developed the normalized spectral abundance factor (NSAF) approach for using spectral counting in quantitative proteomics. This approach takes into account the sample-to-sample variation that is obtained when replicates are used and the fact that longer proteins tend to have more peptide identifications than shorter proteins. Another frequently used method for quantification is the iBAQ algorithm (Schwanhäusser *et al.* 2011), where the intensities of the precursor peptides that map to each protein are summed together and divided by the number of theoretically observable peptides, which is considered to be all tryptic peptides between 6 and 30 amino acids in length.

3.4.2. Challenges and bottle necks for metaproteomic analysis of marine samples

Besides its great potential, metaproteomics, especially when dealing with highly complex environments such as the marine, is facing numerous challenges: (a) the tremendous heterogeneity and complexity e.g. due to additional biogenic and nonbiogenic materials within the samples, (b) protein extraction and the comparably little protein yield than can be obtained compared to other environmental samples, (c) the wide range of protein abundance levels, where numerous low abundance peaks among more abundant ones remain uncharacterized due to physical limits on the number of ions entering the MS, (d) the (up to-now) limited amount of genomic sequence information of the mostly unculturable biota (Schneider and Riedel 2010) and thus the generation of a fitting database, and (e) assigning peptides to proteins, especially if proteins are highly similar, known as the protein inference problem (Nesvizhskii and Aebersold 2005; VerBerkmoes *et al.* 2009; Muth *et al.* 2013; Wang *et al.* 2014; Timmins-Schiffman *et al.* 2017; Verheggen *et al.* 2020). Additionally, especially marine metaproteomics faces low microbial densities compared to for example soil. Microbial densities in the oceans (10^5 to 10^6 cells/g) are orders of magnitudes lower than those found in sediments (10^8 cells/g) or soil (10^9 cells/g) (Savage 1977; Schloss and Handelsman 2006; Whitman *et al.* 1998).

4. Scope of the thesis

Marine ecosystems are important to human society in a variety of ways, including maintaining Earth's habitability through microbial biogeochemical cycling and economic activities such as fisheries and aquaculture. Improving the understanding of the processes shaping the marine ecosystem is important to maintain its stability, predict consequences of e.g., environmental pollution or climate change, and prevent for example harmful algal blooms.

So far, the majority of published studies focused on free-living (FL) bacterial communities. Thereby the community structure and functionality of particle-associated (PA) bacterial communities remained largely unexplored and knowledge on specific contributions of these microorganisms to carbon cycling is still surprisingly limited. Since PA bacteria are expected to mineralize a significant fraction of the decaying algal biomass in marine ecosystems it is important to also gain insights into their role in the oceans. The particulate fraction of bacterioplankton has proved challenging to comprehensive meta-omics characterization due to its high complexity, the presence of DNA/protein-binding polysaccharides, process-interfering substances and lack of (meta)genomic information on marine particles. Thus, one aim of this thesis was the establishment of a robust and reproducible metaproteomics protocol for culture-independent analyses of seasonal taxonomic and functional successions of PA microbial communities in aquatic habitats, which helps to address and fill the above-described knowledge gap.

Recent studies investigated bacterioplankton communities during phytoplankton blooms and revealed that community structure and diversity were highly affected. However, recent research focused mainly on changes of community structure as a response to phytoplankton blooms, but functional changes have been rarely studied. Understanding the dynamics of interactions between bacterial communities and phytoplankton blooms on both taxonomic and functional level is crucial to validate the ecological impact of bloom events. Therefore, the in this study established metaproteomics protocol was used for analysing multiple time points during a spring phytoplankton bloom to investigate succession of taxonomical clades together

with expressed functions of marine particles. Thereby PA community's specific contribution to polysaccharide decomposition in marine habitats should be unravelled. Furthermore, it should be tested if PA communities express specific genes to adapt to the sessile life style and to the availability of high concentration of polysaccharides as well as if evidences for algal-bacterial interactions could be found.

FL planktonic bacteria remineralize polysaccharides using the enzymes encoded in specific polysaccharide utilization loci (PULs). However, first insights into the genomes of PA bacteria showed the absence of characterized PULs for the degradation of naturally abundant polysaccharides as laminarin and alginate. First experiments showed, that the in this thesis chosen PA marine bacterial isolate is able to grow on laminarin and alginate, despite the absence of known PUL structures for these polysaccharides. Therefore, another goal of this thesis was to get insights into the degradation of laminarin and alginate by this specific PA marine bacterial isolate employing proteomic analysis of cultures grown on the mentioned polysaccharides.

5. Material and methods

5.1. Material

5.1.1. Software

All software used in this study is shown in Table 1.

Table 1. Software used in this study

Software		Software producers
bbduk	version 35.14	http://bbtools.jgi.doe.gov
CD-HIT	program for clustering and comparing protein or nucleotide sequences; version 4.8.1	http://weizhongli-lab.org/cd-hit/download.php
dbCAN meta server	CDD version 3.18-55426 PSSMs	http://bcbl.unl.edu/dbCAN2/index.php
FastQC	version 0.11.2	http://www.bioinformatics.babraham.ac.uk/projects/fastqc
Mascot	version 2.6.0 (samples from 2009) and 2.7.0 (samples from 2018)	Matrix Science
MaxQuant	version 1.6.10.43	Max-Planck Institute for Biochemistry, Martinsried
MEGAHIT	version 1.1.3	https://github.com/voutcn/megahit
metaSPAdes	version 3.10.1	http://cab.spbu.ru/software/spades
MSConvert	64-bit, Proteowizard 3	http://proteowizard.sourceforge.net/download.html
Paver		DECODON GmbH
prodigal	version 2.6.3	http://compbio.ornl.gov/prodigal/
Prokka	version 1.11	http://vicbioinformatics.com
ProPHAnE	version 3.1.1 (samples from 2009) and 4.1 (samples from 2018)	http://prophane.de
Scaffold	versions Scaffold_4.8.7 (samples from 2009) and Scaffold_4.10.0 and Scaffold_4.11.0 (samples from 2018)	Proteome Software Inc.
SIMCA®	Version 16.0.1	Sartorius Stedim biotech
SulfAtlas	Version 1.3 (35,637 entries)	http://application.sb-roscoff.fr/blast/blast_sulfatlas/

5.1.2. Databases

All protein sequence databases used in this study are shown in Table 2.

Table 2. Protein databases used in this study

Database	Description	Supplier
NCBInr	Non redundant NCBI database; NCBIprot_20171030; 136,216,794 entries	National Center of Biotechnological information
PABD	Protein sequences of abundant bacteria and diatoms based on the data from Teeling <i>et al.</i> (2012); Uniprot_DoS_complete_20170829; 2,683,314 entries	Uniprot KB
MIMAS	Protein sequences of the FL fraction from Teeling <i>et al.</i> (2012); MIMAS_forward_reverse_all_contaminants; 1,579,724 entries	Metagenomes of the 0.2 μ m fraction from 2009
0.2 + 3 μ m 2009	02_plus_3_POMPU_nr97_fw_cont_20181015; 1,463,571 entries	Metagenomes of the 0.2 μ m and 3 μ m fractions from the 04/14/2009
3 + 10 μ m 2018	3_plus_10um_all_cont_nr_2018; 14,764,755 entries	Metagenomes of the 3 μ m and 10 μ m fractions from 03/19/2018, 04/12/2018, 04/17/2018, 04/26/2018, 05/08/2018, 05/11/2018, 05/22/2018 and 05/29/2018
Muricauda NCBI	Protein sequences of <i>Muricauda sp. MAR_2010_75</i> ; Muricauda_NCBI_fw_rev_cont_DS; 7,198 entries	National Center of Biotechnological information

5.1.3. Laboratory equipment

All laboratory equipment that was used in this study is shown in Table 3.

Table 3. Laboratory equipment used in this study.

Device	Description/type	Manufacturers
Centrifuges	Heraeus Megafuge 8R Micro Star 17R	Thermo Fisher Scientific VWR
Easy-nLC 1000	Reversed phase chromatography	Thermo Fisher Scientific
Easy-nLC 1200	Reversed phase chromatography	Thermo Fisher Scientific
Electrophoretic power supply	PowerPac™ HC	BioRad
FastPrep-24™ Homogenisator	TeenPrep and QuickPrep rotor	MP Biomedicals
LTQ Orbitrap Velos mass spectrometer		Thermo Fisher Scientific

Material and methods

Magnetic mixer	VMS-A	VWR
Micropipettes	P10 P20 P100 P200 P1000	Gilson
Orbital shaker	LOOPSTER digital	IKA®
pH Meter	FiveEasy pH/mV	Mettler Toledo
Q-Exactive™ HF Hybrid-Quadrupol Orbitrap mass spectrometer		Thermo Fisher Scientific
Qubit 4 Fluorometer		Thermo Fisher Scientific
Scanner (gel documentation)	ViewPic 700	Biostep
SDS-PAGE chambers	Criterion™	BioRad
Shaker	Polymax 1040	Heidolph
Spectrophotometer	Ultrospec 2100 pro	biochrom
Thermomixer	ThermoMixer C	Eppendorf
Ultrasonic bath	Ultrasonic cleaner	VWR
Ultrasonic probe	Sonopuls HD2200 with microtip MS 73	Bandelin electronic
Vacuum centrifuge	Vacufuge Concentrator Plus	Eppendorf
Vortexer	Vortex Genie 2	VWR
Water Bath	AQUA line AL5	LAUDA
Weighing scales	Adventurer™ Entris	OHAUS® Satorius

5.1.4. Chemicals, kits and enzymes

All chemicals, comerial kits and enzmyes that were used in this study are shown in Table 4.

Table 4. Chemicals, comerial kits and enzymes used in this study.

Chemicals/comerial kits/enzymes	Manufacturers
β-mercaptoethanol	Carl Roth
Acetic acid	Carl Roth
Acetone	Carl Roth
Acetonitrile LC-MS Grade	VWR
Acetonitrile + 0.1% acetic acid	Carl Roth
Ammoniumacetate	Carl Roth
Ammoniumbicarbonate	Sigma-Aldrich
Ammoniumsulfate	Carl Roth
Bovine serum albumin	Thermo Fisher Scientific
Bromphenol Blue	Sigma-Aldrich
Chloroform	Carl Roth
Coomassie Brilliant Blue G-250 dye	Sigma Aldrich

Compat-Able™ Protein Assay Preparation Reagent Kit	Thermo Fisher Scientific
CTAB	Carl Roth
Disodium phosphate	Carl Roth
Dithiotreitol (DTT)	GE Healthcare
EDTA	Carl Roth
Ethanol	Carl Roth
Formic acid	Carl Roth
Glycerol	Carl Roth
Glycin	Carl Roth
Hydrochloric acid	Carl Roth
Iodoacetamide	Sigma-Aldrich
Isoamylalcohol	Carl Roth
Isopropyl	Carl Roth
Laminarin from <i>Laminaria digitata</i>	Sigma Aldrich
Magnesium chloride	Carl Roth
Methanol für HPLC	VWR
Orthophosphoric Acid	Carl Roth
Page Ruler™ Prestained Protein Ladder	Thermo Fisher Scientific
Phenol	Sigma Aldrich
PMSF	Carl Roth
Pierce™ BCA Protein Assay Kit	Thermo Fisher Scientific
Proteinase K	Sigma Aldrich
Qubit™ dsDNA HS Assay kit	Thermo Fisher Scientific
SDS	AppliChem
Sodium chloride	Carl Roth
Sodium dihydrogen phosphate	Carl Roth
Sodium hydroxide	Carl Roth
Sodium phosphate	Carl Roth
Thiourea	Sigma-Aldrich
Trichloroacetic acid	Carl Roth
TRI-Reagent™	Sigma Aldrich
Tris	Carl Roth
Trypsin, sequencing grade	Promega
Urea	Merck
Water with 0.1% acetic acid	Carl Roth

5.1.5. Consumables

All consumables that were used in this study are shown in Table 5.

Table 5. Consumables used in this study.

Consumables	Description/type	Manufacturers
C18 analytical columns	100 µm x 20 cm	self-packed
C18 Millipore® ZipTip pipette tips		Merck
Conical test tubes	15 mL, sterile	Sarstedt
	50 mL, sterile	Sarstedt
Cuevettes	polystyrole, 10 mm, 1.5 ml	Sarstedt
Eppendorf tubes	1.5 ml	Eppendorf
	2 ml	
Gelloader tips	200 µl	Biozym

Glas beads	0.1-0.11 mm Ø	Scientific Industries, Inc.
Low protein binding tubes	1.5 ml	Carl Roth
MS inserts	100 µl	VWR
MS sample vials		VWR
Pipette tips	10 µl	Sarstedt
	200 µl	
	1000 µl	
Polycarbonate filter	10 µm, 142 mm Ø, TCTP	Millipore
	3 µm, 142 mm Ø, TSTP	
Polyethersulfone filter	0.2 µm, 142 mm Ø, GPWP	Millipore
Qubit™ Assay GefäÙe	polypropylene, 0.5 ml	Thermo Fisher Scientific
Reaction tubes	2 ml	Sarstedt
Scalpel	steril	Swann-Morton
TGX precast separation gel	4-20%	Biorad

5.2. Methods

5.2.1. Metaproteomic analysis of filters

Bacterial biomass sampling. Sampling of bacterioplankton was performed during spring of the years 2009 and 2018 as previously described in Teeling *et al.* 2012. Summarized, 500 L of subsurface seawater (1 m depth) were sampled weekly in 2009 and twice a week in 2018 at the station “Kabeltonne” (50° 11.3’ N, 7° 54.0’ E) between the main island Helgoland and the minor island “Düne” about 40 km offshore in the south-eastern North Sea in the German Bight. Bacterial biomass for protein extraction was sequentially filtered with peristaltic pumps onto 10 µm, 3 µm, and 0.2 µm pore-sized filters (142 mm diameter) to separate PA and FL bacteria. In detail, the samples were first filtered through 10 µm pore-size polycarbonate filters to collect large particles and eukaryotic plankton. Then, the water samples were filtered onto 3 µm pore-size polycarbonate filters to collect bacteria associated with smaller particles and algae and finally the water samples were filtered onto 0.2 µm pore-size polyethersulfone filters to collect FL bacteria. All filters were stored at -80 °C until further analysis. Usually the entire filtration process was finished within 6-8 hours.

Testing of protein extraction protocols. To test six different existing protein extraction protocols for their applicability on PA bacteria, filters from several sampling events (2-4, Table S1) containing varying amounts of biomass were chosen. Sample preparation for the metaproteomic analysis included cutting the filters into quarters and subsequently into small

pieces (1-2 mm in diameter). Pieces from one quarter filter were transferred into 15 ml falcon tubes and treated according to the respective protocol.

Protocol 1 - Phenol. Filter pieces were incubated shaking in 2.4 ml of a 0.1 M NaOH solution for 30 min at room temperature and subsequently sonicated three times for 30 s with 10% pulse and 20% amplitude. The supernatant was transferred into a new tube and protein extraction using phenol was performed according to the protocol published by Kuhn and colleagues (Kuhn *et al.* 2011) with slight modifications. In summary 5.6 ml phenol and 2.4 ml a. bidest were added to the supernatant and the samples were mixed for 1 h at RT. To separate the aqueous phase from phenol phase the sample was centrifuged for 30 min at 12,500 x g and 20 °C. The phenol phase was transferred into a new tube and washed by mixing it with an equal volume of a. bidest for 30 min. Subsequently, the sample was centrifuged as described above, the phenol phase was transferred into a new tube and the 5-fold volume of a 0.1 M ammonium acetate in methanol -solution was added for precipitation of proteins over night at – 20 °C. The next day the sample was centrifuged for 20 min at 12,500 x g and 4 °C and the resulting protein pellet was washed two times with ammonium acetate in methanol, two times with 80% ice-cold acetone_(aq) and once 70% ice-cold ethanol_(aq). Every washing step was carried out for 15 min at – 20 °C and followed by centrifugation as described above.

Protocol 2 - SDS-TCA. Filter pieces were mixed with 5 ml extraction buffer and vigorously shaken for 2 min at RT. The cell disruption by sonication, boiling and shaking was performed with following modifications according to the protein extraction protocol published by Schneider and colleagues (Schneider *et al.* 2012). The sample was sonicated three times for 40 s with one min break on ice between the cycles (10% pulse, 70% amplitude). This was followed by boiling for 20 min and shaking for 1 h at 4 °C. To remove cell debris the sample was centrifuged for 15 min at 12,500 x g and 4 °C and the supernatant was transferred into a new tube. Subsequently, proteins were precipitated with 100% TCA over night at 4 °C. The precipitated proteins were centrifuged as described above and the pellet was washed twice in ice-cold acetone.

- Extractionbuffer SDS-TCA 50 mM Tris
 1% SDS
 pH 7.0 (HCl)

Protocol 3 - TRI-Reagent®. The TRI-Reagent® (Sigma-Aldrich, Rio *et al.* 2010) is used for the simultaneous isolation of RNA, DNA, and proteins. Filter pieces were transferred into 4 ml TRI-Reagent and shaken vigorously for 5 min. Subsequently, proteins were extracted according to the manufacturer's guidelines. In summary, 800 µl chloroform were added and the suspension was again shaken vigorously for 15 s. After an incubation of 15 min at RT the sample was centrifuged for 15 min at 12,000 x g and 4 °C to separate the suspension into three phases: a red organic phase (containing proteins), and interphase (containing DNA), and a colorless upper aqueous phase (containing RNA). The aqueous phase was discarded. To precipitate DNA 1.2 ml 100% ethanol was added and the suspension was mixed by inversion, incubated for 3 min at RT and centrifuged for 5 min at 2,000 x g and 4 °C. Afterwards the supernatant was transferred into a fresh tube and used for protein extraction. The proteins were precipitated with 6 ml 2-propanol for 15 min at RT. A centrifugation step for 10 min at 12,000 x g and 4 °C followed. The supernatant was discarded and the pellet was washed three times with 8 ml of a 0.3 M guanidine hydrochloride/ 95% ethanol solution. Each washing step included an incubation for 20 min at RT and a centrifugation step for 5 min at 7,500 x g and 4 °C. After the third washing step, 2 ml of 100% ethanol were added to the protein pellet and the sample was shaken vigorously. The sample was incubated and centrifuged as described above.

Protocol 4 - Freeze and Thaw. Protein extraction was carried out using a combination of the extraction protocols of Chourey *et al.* (2010) and Thompson *et al.* (2008). To this end, filter pieces were mixed with 4 ml lysis buffer and vigorously shaken for 3 min. Subsequently, the samples were boiled for 10 min, followed by two freezing and thawing cycles with liquid nitrogen. After another boiling procedure and cooling at 4 °C, the samples were vigorously shaken for 3 min. To remove cell debris, samples were centrifuged for 20 min at 12,500 x g at 4 °C. The proteins in the supernatant were precipitated with 25% TCA over night at 4 °C. Precipitated proteins were centrifuged as described above and the resulting protein pellet was washed with ice-cold acetone.

- Lysisbuffer 50 mM Tris
 5% SDS
 0.1 mM EDTA
 0.15 M NaCl
 1 mM MgCl₂
 50 mM DTT
 pH 8.5 (HCl)

Protocol 5 - SDS-Acetone. Filter pieces were mixed with 5 ml extraction buffer and vortexed vigorously for 2 min. Proteins were extracted by sonication, boiling and shaking as described by Hall and colleagues (Hall *et al.* 2012). In summary, cells were sonicated five times for 1 min with 1 min break on ice (10% pulse, 70% amplitude). Sonication was followed by boiling for 15 min and shaking for 1 h at 4 °C. Afterwards the sample was centrifuged for 5 min at 12,5000 x g and 4 °C. The whole procedure was repeated with the pellet after addition of 3 ml extraction buffer and the supernatants were pooled. Subsequently, proteins were precipitated with five volumes acetone over night at -20 °C. The precipitated proteins were centrifuged for 20 min at 12,500 x g at -20 °C and the pellet washed twice in ice-cold acetone.

During the course of the establishment of an appropriate protein extraction protocol for PA bacterial communities following modification of the above-described protocol was tested: boiling was skipped to prevent the loss of membrane proteins.

- Extractionbuffer 50 mM Tris
 1% SDS
 pH 7.5 (HCl)

Protocol 6 – Bead Beating. Protein extraction was carried out according to the extraction protocol of Moog (2012), which is based on the protocol of Teeling and colleagues (Teeling *et al.* 2012), with slightly modified cell disruption parameters. To this end, filter pieces were covered with 4 ml lysis buffer and 2 ml glass beads were added. The cells on the filter pieces were subsequently disrupted four times for 30 s with 6.5 m/s via bead beating with a Fast Prep™-24 with a 2 min break on ice between the cycles. To remove cell debris and glass beads, samples were centrifuged for 20 min at 12,500 x g at 4 °C and the supernatant was transferred into new tubes. This washing step was repeated 2 to 4 times until the beads were colourless. The glass beads were washed with 3 ml lysis buffer and vigorously shaken.

Proteins enriched in the pooled supernatants were precipitated with 1:4 acetone at -20 °C over night. Precipitated proteins were centrifuged for 20 min at 12,500 x g at 4 °C and the resulting protein pellet washed with ice-cold acetone.

During the course of the establishment of an appropriate protein extraction protocol for PA bacterial communities following modification of the above-described protocol were tested: only one washing step, 2 times 40 s with 6.0 m/s and 3 times 30 s with 6.5 m/s.

- Lysisbuffer 50 mM Tris
 5% SDS
 0.1 M DTT
 0.01 M EDTA
 10% (v/v) glycerol
 1.7 mM PMSF
 pH 6.8 (HCl)

All resulting protein pellets were air-dried and resolved in 50-200 µl 8 M urea / 2 M thiourea buffer according to the size of the resulting protein pellet. After addition of the buffer samples were frozen at -20 °C and the following day shaken for at least 2 h at 4 °C.

- Urea/Thiourea Buffer 8 M urea
 2 M thiourea

Determination of protein concentrations. Protein concentrations were determined using the Pierce™ BCA Protein Assay Kit. Protein extracts were prepared with the Compat-Able™ Protein Assay Preparation Reagent Kit to remove interfering thiol concentrations. The preparation of the samples was performed according to the manufacturer's instructions for the micro centrifuge-tube procedure. Afterwards the concentration of protein solutions was determined according to the test-tube procedure of the manufacturer's guidelines. The Pierce™ BCA Protein Assay Kit uses the purple coloured product of the cuprous cation (Cu²⁺) and bicinchoninic acid reaction (Smith *et al.* 1985), which allows photometric measurement at 562 nm. All measurements of protein solutions were conducted in duplicates and samples were diluted appropriately to fit into the range of the standard curve. The measurements of standard samples were conducted in triplicates.

SDS-PAGE protein separation. To separate proteins according to their molecular weight the sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) (Laemmli 1970) was used. To this end 30 µg protein (method development of protein digestion and MS/MS sample preparation and application example) or 30 µl protein extract (comparison of the protein extraction protocols) was mixed with 4x SDS sample buffer and loaded on TGX precast 4-20% gels. Between the different samples one gel lane each was left empty, because the protein lanes got wider during electrophoretic separation. 3 µl of PageRuler Prestained Protein Ladder were used as a molecular weight marker. Samples were separated by electrophoresis at 150 V for 45 min. After fixation for 30 min the gels were stained with Brilliant Blue G250 Coomassie overnight. The next day, the gel was washed several times until protein bands were clearly distinguishable from the background and subsequently imaged and stored at 4 °C.

- 4 x SDS Sample buffer
 - 100 mM Tris
 - 10% SDS
 - 20% glycerol
 - 5% β-mercaptoethanol
 - 0.8% bromphenol blue
 - pH 6.8 (HCl)
- 10 x SDS-Running buffer
 - 250 mM Tris
 - 1.92 M glycine
 - 1% SDS
 - pH 8.3
- Colloidal Coomassie staining solution
 - 10% (w/v) (NH₄)₂SO₄
 - 0,12% (w/v) Coomassie® Brilliant Blue G-250
 - 10% (v/v) H₃PO₄
 - 20% (v/v) methanol
- Fixing solution
 - 50% a. bidest
 - 10% acetic acid
 - 40% ethanol

Protein digestion and MS sample preparation. Three different protocols were tested on proteins extracted from 3 µm and 10 µm filters in two technical replicates each.

Protocol 1 - 10 gel pieces. Protein lanes were cut into 10 equal-sized pieces, placed into low protein binding tubes and washed 3 times with 700 µl washing solution at 37 °C for 15 min with shaking (1200 rpm) to remove Coomassie stain and SDS residue. Prior to tryptic digestion, gel pieces were dried in a vacuum concentrator (30 °C) and re-swollen with 100 µl of a 2 ng/µl trypsin solution. Any excess of trypsin was removed followed by an overnight digestion at

37 °C, placing the reaction tubes upside down. After digestion, the gel pieces were covered with water and after a short centrifugation step at 10,000 x g for 30 s, peptides were eluted from the gel in an ultrasonic bath for 15 min. Tubes were centrifuged as described above and the supernatant containing peptides transferred into new tubes.

Protocol 2 - 20 gel pieces. The protein lanes were cut into 20 equal-sized pieces and treated as described above.

Protocol 3 - 20 gel pieces with reduction and alkylation (red. and alk.). Protein lanes were cut into 20 equal-sized pieces, placed into low protein binding tubes and washed 3 times with washing solution at 37 °C for 15 min with shaking (1200 rpm) to remove Coomassie stain and SDS residue. Gel pieces were dehydrated in 100% acetonitrile and reswollen in a 50 mM NH₄HCO₃ solution. After another dehydration step in acetonitrile, proteins were reduced in reduction solution for 30 min at 60 °C, followed by another dehydration step and an alkylation in alkylation solution for 60 min in the dark at room temperature. Prior to tryptic digestion, the gel pieces were dehydrated and washed as described above, dried in a vacuum concentrator (30 °C), re-swollen with 300 µl of a 0,1 µg/µl trypsin solution and incubated at 37 °C over night. Peptides were eluted from the gel pieces by an eight-step procedure (200 µl of solution each), using acetonitrile for 5 min, 1% (v/v) acetic acid in water solution for 10 min, acetonitrile for 5 min, 5% (v/v) acetic acid, elution of peptides in ultrasonic bath for 10 min and two times acetonitrile with an additional elution step in an ultrasonic bath in between. Peptide-containing supernatants were pooled and completely dried in a vacuum concentrator (30 °C). Samples were subsequently resolved in 20 µl washing solution.

- Washing solution 200 mM NH₄HCO₃
 30% (v/v) acetonitrile
- Washing solution (red and alk) 100 mM NH₄HCO₃
 50% (v/v) methanol
- Reduction solution 50 mM NH₄HCO₃
 10 mM DTT
- Alkylation solution 50 mM NH₄HCO₃
 50 mM IAA

Zip Tip Purification. The eluted peptides from each protocol were desalted with C18 Millipore® ZipTip columns according to the manufacturer's guidelines. In summary the Zip Tip pipette tip

was prewetted twice with 10 µl wetting solution and subsequently equilibrated for binding by washing it twice with 10 µl equilibration solution. Peptides were bound to the Zip Tip column by aspirating and dispensing the peptide solution 10 times. Peptides were then washed twice with 10 µl washing solution followed by the elution of the peptides into 10 µl elution solution by aspirating and dispersing it 3 times. The elution solution containing the peptides was transferred into MS sample vials with 100 µl inserts. To assure the equal volume of samples, 15 µl a. bidest was added and filled vials were vacuum centrifuged to dryness. 11 µl a. bidest was added to each sample, vials were vortexed and incubated for 1 h at 4 °C. The vials were stored at – 20 °C until LC-MS/MS measurement.

- Wetting solution 70% (v/v) acetonitrile
30% (v/v) a. bidest
- Equilibration solution 3% (v/v) acetonitrile
0.1% (v/v) acetic acid
96.9% (v/v) a. bidest
- Washing solution 0.1 (v/v) acetic acid
99.9% (v/v) a. bidest
- Elution solution 60% (v/v) acetonitrile
0.1% (v/v) acetic acid
39.9% (v/v) a. bidest

Constructions of a protein sequence database from marine metagenomes.

Extraction of environmental DNA. Environmental DNA was extracted by the Max Planck Institute for Marine Microbiology in Bremen, according to a modified standard protocol of Zhou *et al.* (1996). In detail, one polycarbonate filter was cut into 4 pieces and mixed with 13.5 ml extraction buffer. Subsequently, 100 µl 10 mg/ml Proteinase K was added and the sample was incubated shaking at 37 °C for 30 min. 1.5 ml 20% SDS was added and the sample was incubated shaking at 65 °C for 2 h. The sample was centrifuged at 6,000 x g for 10 min at RT and the supernatant was transferred to a fresh tube. Subsequently, an equal volume of chloroform/isoamylalcohol (24:1)-solution was added and the sample was mixed carefully by shaking for 20 min and was subsequently centrifuged at 10,000 x g for 10 min at RT. Afterwards the aqueous upper phase was transferred into a new tube and the DNA was precipitated by addition of 0.6 volumes isopropanol. The sample was moderately shaken over night at 4°C. After centrifugation at 50,000 x g for 20 min at RT, the pellet was washed with 10

ml 80% (v/v) ethanol and dried. The pellet was resuspended in 200 µl TE buffer and stored at -20 °C until sequencing. DNA-concentration was determined with the Qubit™ dsDNA HS Assay Kit according to the manufacturer's guidelines.

- Extraction buffer
 - 1 M Tris
 - 0.5 M EDTA
 - 1 M Na-phosphate buffer
 - 5 M NaCl
 - 10% CTAB
 - pH 8.0
- Na-phosphate buffer
 - 1 M Na₂HPO₄
 - 1 M NaH₂PO₄
- TE buffer
 - 10 mM Tris
 - 1 mM EDTA

DNA-sequencing, metagenome assembly and gene annotation. DNA was sequenced at the Max Planck Sequencing Centre (Cologne, Germany), using the Illumina HiSeq 2500 platform and 2 x 250 bp chemistry. The further processing of the metagenome sequences was done by Dr. Thomas Ben Francis at the Max Planck Institute for Marine Microbiology in Bremen. In summary, sequences were trimmed using bbduk v35.14 (<http://bbtools.jgi.doe.gov>) with the following parameters: ktrim = r k = 28 mink = 12 hdist = 1 tbo = t tpe = t qtrim = rl trimq = 20 minlength = 100. Read quality for each sample was then confirmed using FastQC v0.11.2 (Andrews 2010). Trimmed and filtered reads from the three size fractions were assembled individually. The 0.2 µm pore-sized filter sample was assembled with metaSPAdes v3.10.1 (Nurk *et al.* 2017) with kmers of length 21, 33, 55, 77, 99, and 127, and error correction mode switched on. Assembly of the larger size fraction was done with MEGAHIT v1.1.3 (Li *et al.* 2015) with kmers 21, 33, 55, 77, 99, 127, 155, 183, and 211. Assembled contigs longer than 1,500 base pairs were kept for gene predictions. Genes were predicted and annotated using Prokka v1.11 (Seemann 2014), which implements prodigal v2.6.3 (Hyatt *et al.* 2010) for ORF prediction. Raw read sequences and assembled contig sequences have been deposited in the European nucleotide archive (ENA) under the project accession numbers PRJEB2888 (2009 metagenomes and PRJEB38290 (2018 metagenomes).

Construction of a protein sequence database. A list of common contaminants was added to downloaded protein sequences or translated ORF sequences, which were found through the

metagenomic analyses of the different size fractions. Redundant sequences were eliminated (100% redundancy for one single size fraction, 97% redundancy for combined sequences of different size fractions; elimination of shorter sequences) using CD-Hit, a program for clustering and comparing protein or nucleotide sequences.

LC-MS/MS data acquisition and data analysis.

LC-MS/MS measurements. LC-MS/MS measurements were done by Dr. Claudia Hirschfeld (samples from 2009) and Dr. Pierre Mücke (samples from 2018) in the group of Prof. Dr. Dörte Becher at the Institute of Microbiology at the University of Greifswald. Peptides were separated by reversed-phase chromatography on an Easy-nLC 1000 with self-packed C18 analytical columns (100 μ m \times 20 cm) and coupled to an LTQ Orbitrap Velos mass spectrometer using a non-linear binary gradient of 80 minutes from 5 % solvent A to 99 % solvent B and a flow rate of 300 nL/min. Survey scans at a resolution of 30,000 were recorded in the Orbitrap analyser (m/z 300 - 1700) and the 20 most intense precursor ions were selected for CID fragmentation in the LTQ. Dynamic exclusion of precursor ions was enabled; single-charged ions and ions with unknown charge state were excluded from fragmentation. Internal lock mass calibration was enabled (lock mass 445.120025).

- Solvent A 0.1% (v/v) acetic acid
 99.9% (v/v) a. bidest
- Solvent B 0.1% (v/v) acetic acid
 99.9% (v/v) acetonitrile

Database search. The mass spectrometry raw data was converted into mgf files using MSConvert and subsequently subjected to database searching via Mascot. Four different protein sequence databases were used for peptide to spectrum matching: I) the non-redundant NCBI database (NCBI nr - NCBIprot_20171030 database (136,216,794 entries)), II) a database containing protein sequences of abundant bacteria and diatoms (PABD) based on the study of Teeling *et al.* (2012), and retrieved from Uniprot KB (Uniprot_DoS_complete_20170829 database (2,638,314 entries)), III) a database containing protein sequences of the free-living fraction from Teeling *et al.* (2012) (0.2 μ m 2009 (MIMAS) - MIMAS_forward_reverse_all_contaminants database (1,579,724 entries)), and IV) a

database based on translated metagenomes from 0.2 μm and 3 μm filters (see 5.2.1: Constructions of a protein sequence database from marine metagenomes for details) (0.2 + 3 μm 2009 - 02_plus_3_POMPU_nr97_fw_cont_20181015 database (1,463,571 entries)). Data from samples from 2018 were searched against a database based on translated metagenomes from 3 and 10 μm filters of eight different sampling time-points (3 + 10 μm 2018 - 3_plus_10um_all_cont_nr_2018 (14,764,755 entries)). Mascot search was performed with a fragment ion mass tolerance of 0.80 Da and a parent ion tolerance of 10.0 ppm. Oxidation of methionine was specified as a variable modification; trypsin was set as digestion enzyme and a maximum of two missed cleavages was allowed.

Data analysis. Scaffold was used to validate MS/MS-based peptide and protein group identifications. Peptide identifications were accepted if they could be established at greater than 95.0% probability by the Peptide Prophet algorithm (Keller *et al.* 2002) with Scaffold delta-mass correction. Protein group identifications were accepted if they could be established at greater than 99.0% probability and contained at least one identified peptide. Protein group probabilities were assigned by the Protein Prophet algorithm (Nesvizhskii *et al.* 2003). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Peptides that were only found in one of the replicates were excluded from subsequent data analysis. For data obtained from samples from 2018 the X! Tandem search engine (The Global Proteome Machine Organization), which is included in Scaffold, was used. X! Tandem automatically searches for missed cleavages, semi-tryptic peptides, post-translational modifications and point mutations and with this improves the number and confidence of protein group identifications by combining its results with the Scaffold search engine results. Mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (Perez-Riverol *et al.* 2019) with the data set identifiers PXD12699 (data from 2009) and several of the Institute of Microbiology of the University Greifswald (data from 2018).

For further data analysis, the software *ProPHAnE* (Proteomics result Pruning and Homology group Annotation Engine) (Schneider *et al.* 2011; Schiebenhoefer *et al.* 2020) was used. For

the taxonomical classification of the identified protein groups the NCBI NR database (version 2018-08-02; e-value 0.01, query cover 0.9, max-target-seqs 1) and the diamond blastp algorithm (version 0.8.22) were used. For functional classification of the identified protein groups the eggnoG database (version 4.5.1, downloaded at 2018-07-31) and the algorithm e-mapper were used for the data from 2009. For the data from 2018 following databases and parameters were used. For the taxonomical classification of the identified protein groups the NCBI NR database (version 2019-09-30; e-value 0.001, diamond blastp search algorithm) and the UniprotKB database (version 2019-09-30; e-value 0.001; diamond blastp search algorithm) were used. For functional classification of the identified protein groups the eggnoG database (version 4.5.1; e-value 0.001; emapper search algorithm), the PFAMs database (version 32; e-value 0.001; hmmscan search algorithm), the TIGRFAMs database (version 15.0; e-value 0.001; hmmscan search algorithm), the FOAM database (version rel1a; e-value 0.001; hmmscan search algorithm), the CAzY/dbCAN database (version 8; e-value 0.001; hmmscan and hmmsearch search algorithm) and the ResFAMs(core) database (version 1.2; e-value 0.001; hmmscan search algorithm) were used. Functional classification of the identified protein groups used for the data interpretation was carried out manually. As the PFAMs database led to the highest number of both annotated bacterial and eukaryotic protein groups, this database was used as main annotation source. Results of the other databases were compared to the PFAM database results and differences were controlled via BLAST search and possibly changed to the verified annotation. Functional clustering of protein groups used for data interpretation and creation of figures was based on TIGRFAMs descriptions.

The NSAF (normalized spectral abundance factor)-value was used for the examination of the abundance of protein groups (Zybailov *et al.* 2006). For the calculation of this value the sum of spectral counts, the number of aa of the longest protein sequence (samples from 2009) and number of aa of the mean length of protein sequences (samples from 2018) of the respective protein groups and the number of overall identified protein groups is taken into account. In the method of spectral counting (Zhu *et al.* 2010b) it is assumed that the number of peptides of a distinct protein that contributes to the peptide mixture, correspond to the concentration of this

protein group in the sample and that the number of peptides mirror the number of recorded MS-spectra. The NSAF value of a protein group can have a value between 0 and 1. A value nearer to 1 indicates a higher amount of the respective protein group in the sample.

PCA (principal component analysis) and OPLS-DA (orthogonal partial least squares-discriminant analysis) plots were used to visualize differences and similarities within the complex metaproteomics dataset of the spring phytoplankton bloom 2018. The plots were further used to investigate, if the differences between the samples were based on pore-size of the used filters or on sampling time point during the bloom. To generate the plots the SIMCA® software was used. The NSAF-values of all identified protein groups assigned to the same taxonomical class or family were summed up as variables. To handle the inhomogeneity within the absolute values of the variables and secure that each variable has the same influence on the statistical model, the variables were centred and scaled to “Unit Variance” (UV).

5.2.2. Proteomic analysis of a selected marine particle-associated bacterial isolate grown on different polysaccharides

Isolation and Cultivation of *Muricauda* sp. MAR_2010_75

Isolation. The strain *Muricauda* sp. MAR_2010_75 was isolated from the colleagues Hahnke and Harder (2013) from phytoplankton near List on the German island Sylt in the North Sea. The strain belongs to the genus *Muricauda* within the class of *Flavobacteriia* of the Phylum *Bacteroidetes*. The strain has 96,3% 16S rRNA gene identity to *Muricauda flavrescens*. *Muricauda* sp. MAR_2010_75 forms round orange colonies when it grows on agar and shows gliding motility.

Cultivation. The cultivation of the strain was performed at the Max Plank Institute for Marine Microbiology in Bremen. The strain was cultivated in HaHa100V medium according to Hahnke *et al.* (2015). Before inoculation 2 g/l of a specific carbon source were added to the medium. In this study glucose and the polysaccharides laminarin and alginate were used. For the inoculation of the liquid culture 0.4% (v/v) of a respective preculture was used. Precultures

Material and methods

were inoculated by picking a colony of the bacterium from an agar plate (marine agar 2216; 55.1 g/l). Liquid cultures were incubated at 25 °C till they reached an OD_{600nm} of roughly 0,25 (reached after 45,5-47,5 h).

Strain	NCBI Accession number
Muricauda sp. MAR_2010_75	PRJNA24852

- Haha100V medium
 - 26.37 g/l NaCl
 - 5.67 g/l MgCl₂ x 6 H₂O
 - 6.8 g/l MgSO₄ x 7 H₂O
 - 1.47 g/l CaCl₂ x 2 H₂O
 - 0.72 g/l KCl
 - 0.1 g/l KBr
 - 0.02 g/l H₃BO₃
 - 0.02 g/l SrCl₂
 - 0.003 g/l NaF
- Supplementation Haha100V
 - 2 ml/l trace elements solution (Pfennig and Trüper 1981)
 - 1 ml/l SeW solution (Widdel and Bak 1992)
 - 5 mg/l KH₂PO₄
 - 0.8 g/l NH₄Cl
 - 0.1 g/l yeast extract
 - 0.1 g/l casaminoacids
 - 0.1 g/l trypton peptone
 - 1 ml 7 vitamins solution (Winkelmann and Harder 2009)
 - 1 ml vitamin B12 solution (Widdel and Bak 1992)
 - 1 ml riboflavin solution (Winkelmann and Harder 2009)
 - 1 thiamin solution (Winkelmann and Harder 2009)
 - pH 7.5 (1 M NaHCO₃)
- Trace elements solution
 - 2.1 g/l FeSO₄ x 7 H₂O
 - 5.2 g/l Na₂-EDTA
 - 30 mg/l H₃BO₃
 - 100 mg/l MnCl₂ x 4 H₂O
 - 190 mg/l CoCl₂ x 6 H₂O
 - 24 mg/l NiCl₂ x 6 H₂O
 - 10 mg/l CuCl₂ x 2 H₂O
 - 144 mg/l ZnSO₄ x 7 H₂O
 - 36 mg/l Na₂MoO₄ x 2 H₂O
 - pH 6.0 (5 M NaOH)
- SeW solution
 - 18 mg/l Na₂SeO₃ x 5 H₂O
 - 18 mg/l Na₂WO₄ x 2 H₂O
 - 200 mg/l NaOH
- 7 vitamins solution
 - 4 mg 4-aminobenzoic acid
 - 1 mg D (+) - biotin
 - 10 mg nicotinic acid
 - 5 mg D (+) – panthothenic acid, Ca-salt
 - 15 mg pyridoxine dichloride
 - 4 mg folic acid

Material and methods

- Vitamin B12 solution
1.5 mg lipoic acid
in 100 ml 10 mM NaH₂PO₄ + Na₂HPO₄, pH 7.1
50 mg/l cyanocobalamin
- Riboflavin solution
5 mg riboflavin
in 100 ml 25 mM NaH₂PO₄
pH 3.4
- Thiamin solution
5 mg thiamin
in 100 ml 25 mM NaH₂PO₄
pH 3.4

Harvest. Cultures were transferred into 50 ml test tubes and centrifuged for 20 min at 5,000 rpm and 4 °C. The resulting cell pellet was resuspended in 1 ml medium and transferred into a 1.5 ml tube. After another centrifugation step for 15 min at 130,000 rpm and 4 °C the pellet was frozen in liquid nitrogen and stored at – 20 °C till further analysis. The identity of the strain was confirmed by use of partial 16S rRNA sequencing.

Protein extraction. For protein extraction, cell pellets were resuspended in 500 µl lysis buffer and transferred into a reaction tube containing 250 µl glassbeads. The cells were subsequently disrupted four times for 30 s with 6.5 m/s via bead beating with a Fast Prep™-24 with 2 min break on ice between the cycles. To remove cell debris and glass beads, samples were centrifuged for 10 min at 12,500 x g at 4 °C and the supernatant was transferred into new 2 ml Eppendorf tubes. The glass beads were washed with 500 µl lysis buffer and vigorously shaken. Proteins enriched in the pooled supernatants were precipitated with 1:4 acetone at -20 °C overnight. Precipitated proteins were centrifuged for 20 min at 12,500 x g at 4 °C and the resulting protein pellet washed with ice-cold acetone. The resulting protein pellet was air-dried and resolved in 200 µl 8 M urea / 2 M thiourea buffer. After addition of the buffer samples were frozen at -20 °C and the following day shaken for at least 2 h at 4 °C.

- Lysis buffer
50 mM Tris
5% SDS
0.1 M DTT
0.01 M EDTA
10% (v/v) Glycerol
1.7 mM PMSF
pH 6.8 (HCl)
- Urea/thiourea buffer
8 M urea
2 M thiourea

Determination of protein concentrations. Protein concentrations were determined using the Pierce™ BCA Protein Assay Kit. Protein extracts were prepared with the Compat-Able™

Protein Assay Preparation Reagent Kit to remove interfering thiol concentrations. The preparation of the samples was performed according to the manufacturer's instructions for the micro centrifuge-tube procedure. Afterwards the concentration of protein solutions was determined according to the test-tube procedure of the manufacturer's guidelines. All measurements of protein solutions were conducted in duplicates and samples were diluted appropriately to fit into the range of the standard curve. The measurements of standard samples were conducted in triplicates.

SDS-PAGE protein separation. To separate proteins 30 µg protein was mixed with 4x SDS sample buffer and loaded on TGX precast 4-20% gels. Between the different samples one gel lane each was left empty, because the protein lanes got wider during electrophoretic separation. 3 µl of PageRuler Prestained Protein Ladder were used as a molecular weight marker. Samples were separated by electrophoresis at 150 V for 45 min. After fixation for 30 min the gels were stained with Brilliant Blue G250 Coomassie overnight. The next day, the gel was washed several times until protein bands were clearly distinguishable from the background and subsequently imaged and stored at 4 °C.

Protein digestion and MS sample preparation.

Protein in-gel digest. Protein lanes were cut into 10 equal-sized pieces, placed into low protein binding tubes and washed 3 times with 700 µl washing solution at 37 °C for 15 min with shaking (1200 rpm) to remove Coomassie stain and SDS residue. Prior to tryptic digestion, gel pieces were dried in a vacuum concentrator (30 °C) and re-swollen with 100 µl of a 2 ng/µl trypsin solution. Any excess of trypsin was removed followed by an overnight digestion at 37 °C, placing the reaction tubes upside down. After digestion, the gel pieces were covered with water and after a short centrifugation step at 10,000 x g for 30 s, peptides were eluted from the gel in an ultrasonic bath for 15 min. Tubes were centrifuged as described above and the supernatant containing peptides transferred into new tubes.

Zip Tip Purification. The eluted peptides from each protocol were desalted with C18 Millipore® ZipTip columns according to the manufacturer's guidelines. In summary the Zip Tip pipette tip

was prewetted twice with 10 µl wetting solution and subsequently equilibrated for binding by washing it twice with 10 µl equilibration solution. Peptides were bound to the Zip Tip column by aspirating and dispensing the peptide solution 10 times. Peptides were then washed twice with 10 µl washing solution followed by the elution of the peptides into 10 µl elution solution by aspirating and dispersing it 3 times. The elution solution containing the peptides was transferred into MS sample vials with 100 µl inserts. To assure the equal volume of samples, 15 µl a. bidest was added and filled vials were vacuum centrifuged to dryness. 11 µl a. bidest was added to each sample, vials were vortexed and incubated for 1 h at 4 °C. The vials were stored at – 20 °C until LC-MS/MS measurement.

LC-MS/MS data acquisition and data analysis.

LC-MS/MS measurements. LC-MS/MS measurements were done by Dr. Daniela Zühlke in the group of Prof. Dr. Katharina Riedel at the Institute of Microbiology at the University of Greifswald. Peptides were separated by reversed-phase chromatography on an Easy-nLC 1200 with self-packed C18 analytical columns (100 µm × 20 cm) and coupled to an Q Exactive™ HF Hybrid-Quadrupole-Orbitrap mass spectrometer using a non-linear binary gradient of 80 minutes from 5 % solvent A to 99 % solvent B and a flow rate of 300 nL/min. Survey scans at a resolution of 60,000 were recorded in the Orbitrap analyser (m/z 333 - 1650) and the 15 most intense precursor ions were selected for HCD fragmentation. Dynamic exclusion of precursor ions was enabled; single-charged ions and ions with unknown charge state were excluded from fragmentation. Internal lock mass calibration was enabled (lock mass 445.12003).

Database search. The mass spectrometry raw data was subjected to database searching via MaxQuant software (Cox and Mann 2008). A database containing all available protein sequences of *Muricauda sp. MAR_2010_75* (Muricauda_NCBI_fw_rev_cont_DS; 7,198 entries) was used for peptide to spectrum matching: Peptide search was performed with the Andromeda search algorithms (Cox *et al.* 2011). For positive protein identification the following criteria had to be met: a minimum peptide length of 6 aa, a maximum of 2 missed cleavages,

oxidation of methionine was specified as a variable modification and trypsin was set as digestion enzyme. The false discovery rate (FDR) was estimated and protein identifications with FDR < 1% were considered acceptable. Furthermore, a protein has to be identified in at least 2 out of 3 biological replicates and a minimum of two unique peptides per protein was required for relative quantification using the iBAQ (intensity based absolute quantification) algorithm provided by MaxQuant.

Mass spectrometry proteomics data is available on the servers of the Institute of Microbiology of the University Greifswald.

Data analysis. Provided iBAQ values, whereby the peak area is divided by the sum of all theoretical peptides, were used to manually calculate riBAQ (relative iBAQ) values for semi quantitative comparison of protein abundances. riBAQ values give the relative protein abundance in percent of all proteins in the same sample by normalizing each protein's iBAQ value to the sum of all iBAQ values (Shin *et al.* 2013).

For further data analysis and annotation of identified protein the web CD-Search Tool (Lu *et al.* 2020) was used which is a NCBI interface to search the Conserved Domain Database and uses RPS-BLAST to quickly scan a set of precalculated position-specific scoring matrices (PSSMs). For the search the CDD database was used a superset which includes NCBI-curated domains and data imported from Pfam, SMART, COG, PRK and TIGRFAMs. For the search the expected e-value was set to 0.01. Furthermore the dbCAN (automated Carbohydrate-active enzyme Annotation) meta server (Zhang *et al.* 2018b) was used for the identification of CAZymes among the identified proteins. For this annotation the bio sequence analysis tool HMMER, which uses profile hidden Markov models, was used for the annotation of CAZyme domain boundaries according to the dbCAN CAZyme domain HMM database and the protein alignment tool DIAMOND was used to obtain blast hits in the CAZyme database. The proteome data was combined with the locus tags from NCBI. PULs predicted in the study from Kappelmann *et al.* 2018 were translated into the locus tags from NCBI via start- and stop codons.

6. Results

6.1. Establishment of a metaproteomic pipeline for PA microbial communities (Schultz *et al.* 2020)

Metaproteomic analyses of PA microbial communities are hindered by their high complexity, the high amount of eukaryotic proteins, the sugary particle-matrix as well as the lack of (meta)genomic information on PA-specific pro- and eukaryotes (Wöhlbrand *et al.* 2017a; Saito *et al.* 2019). Metaproteomics analyses of FL bacterioplankton (sampled on 0.2 µm filters) during spring phytoplankton blooms from 2009 to 2012 off the German island Helgoland (station “Kabeltonne”, 54°11′03″N, 7°54′00″E) resulted in the identification of several thousand protein groups (Teeling *et al.* 2012, Kappelmann *et al.* 2018). However, the PA microbial communities retained on 3 and 10 µm pore-sized filters were substantially more difficult to analyze by the integrated metagenomic/metaproteomic approach employed at that time (Teeling *et al.* 2012). For sample preparation, 500 L raw seawater were sequentially filtrated through 10 µm, 3 µm (PA fractions) and 0.2 µm (FL fraction) pore-sized filters. Everything retained on 3 µm and larger pore-sized filters is regarded as marine particles in the following.

Among others this thesis aimed to develop a suitable and effective metaproteomics pipeline for in depth analyses of taxonomic composition and functionality of marine PA microbial communities. Therefore, six different already well established protein extraction protocols for various particulate samples were tested using biomass from 3 and 10 µm pore-sized filters of the MIMAS bio archive (www.mimas-project.de). These samples covered different stages of the spring phytoplankton bloom, which had been sampled in 2009 during the sampling campaign by Teeling *et al.* 2012. Moreover, various MS sample preparation protocols and different protein sequence databases were analyzed.

6.1.1. Protein extraction

Efficient protein extraction is an essential part in the successful metaproteomics analyses of microbial communities. Therefore first of all, five different protein extraction methods were

tested, which employed different strategies and were already successfully applied for metaproteome analyses of microbial communities from different environments. These protocols were applied for sewage sludge (phenol extraction; Kuhn *et al.* 2011), leaf litter (SDS-TCA; Schneider *et al.* 2012), stream hyporheic biofilms (SDS-acetone; Hall *et al.* 2012), hypersaline microbial mats (bead beating; Moog 2012), and soil (freezing and thawing; Chourey *et al.* 2010; Thompson *et al.* 2008). Additionally, the commercially available TRI-Reagent[®] (Sigma Aldrich) for simultaneous isolation of RNA, DNA and proteins was tested (Table 6 Appendix; Figure 2A). Filter samples used for protocol evaluation were sampled at different sampling events in the 2009 spring bloom sampling campaign (9th of February, 7th of April, 21st of April, and 16th of June 2009). Filters were cut into pieces and underwent the respective extraction protocol (Figure 2A; 2-4 sampling events; Table S 1 Appendix). Total protein amounts extracted from the filters by each of the tested methods were quite variable (Table 6, Table S 1 Appendix, Figure 2B). Highest protein yield as determined by the Pierce[™] BCA Protein Assay and 1D SDS-PAGE was obtained using the SDS-acetone or bead-beating approach for both analyzed fractions (Table 6, Figure 2). Therefore, these two methods were used for the optimization of the downstream MS sample preparation procedure.

Results

Table 6. Comparison of the six tested protein extraction protocols

Protocol.	1 - Phenol	2 - SDS-TCA	3 - TRI-reagent®	4 - Freeze and thaw	5 - SDS-acetone	6 – Bead-beating
Reference	Kuhn <i>et al.</i> (2011)	Schneider <i>et al.</i> (2012)	Sigma Aldrich	Chourey <i>et al.</i> (2010) Thompson <i>et al.</i> (2008)	Hall <i>et al.</i> (2012)	Moog (2012)
Originally used for	sewage sludge from biomembrane reactors	leaf litter	simultaneous extraction of RNA, DNA, and proteins	soil	stream hyporheic biofilms	hypersaline microbial mats
Composition of the protein extraction buffer	0.1 M NaOH	1% (w/v) SDS, 50 mM Tris/HCl, pH 7	TRI-Reagent® (guanidine thiocyanate and phenol monophasic solution)	5% (w/v) SDS, 50 mM Tris/HCl, 0.1 mM EDTA, 0.15 M NaCl, 1 mM MgCl ₂ , 50mM DTT, pH 8.5	1% (w/v) SDS, 50 mM Tris/HCl, pH 6.8	5% (w/v) SDS, 0.05 mM Tris/HCl, 0.1 M DTT, 0.01 M EDTA, 10% (v/v) glycerol, 1.7 mM PMSF, pH 6.8
Cell disruption methodology	sonication 3 x 30 s (20% power output)	sonication 3 x 40 s (20% power output)	TRI-Reagent®	2 freeze and thaw cycles (liquid nitrogen, rt), 10 min boiling	sonication 5 x 1 min (20% power output), 15 min boiling, procedure repeated on the pellet	FastPrep® 6.5 m/s, 4 x 30 s
Additional protein purification	phenol extraction (2x)	/	chloroform extraction, ethanol extraction	/	/	/
Protein precipitation	0.1 M ammonium-acetate in methanol (1:5)	10% TCA	2-propanol (1:1.5)	25% TCA	acetone (1:5)	acetone (1:4)
Mean total protein amount [µg] 3 - 10 µm fraction	8.9	22.7	16.1	25.2	38.6	27.3
Mean total protein amount [µg] ≥ 10 µm fraction	8.8	12.8	38.2	24.3	102.1	114.2

Results

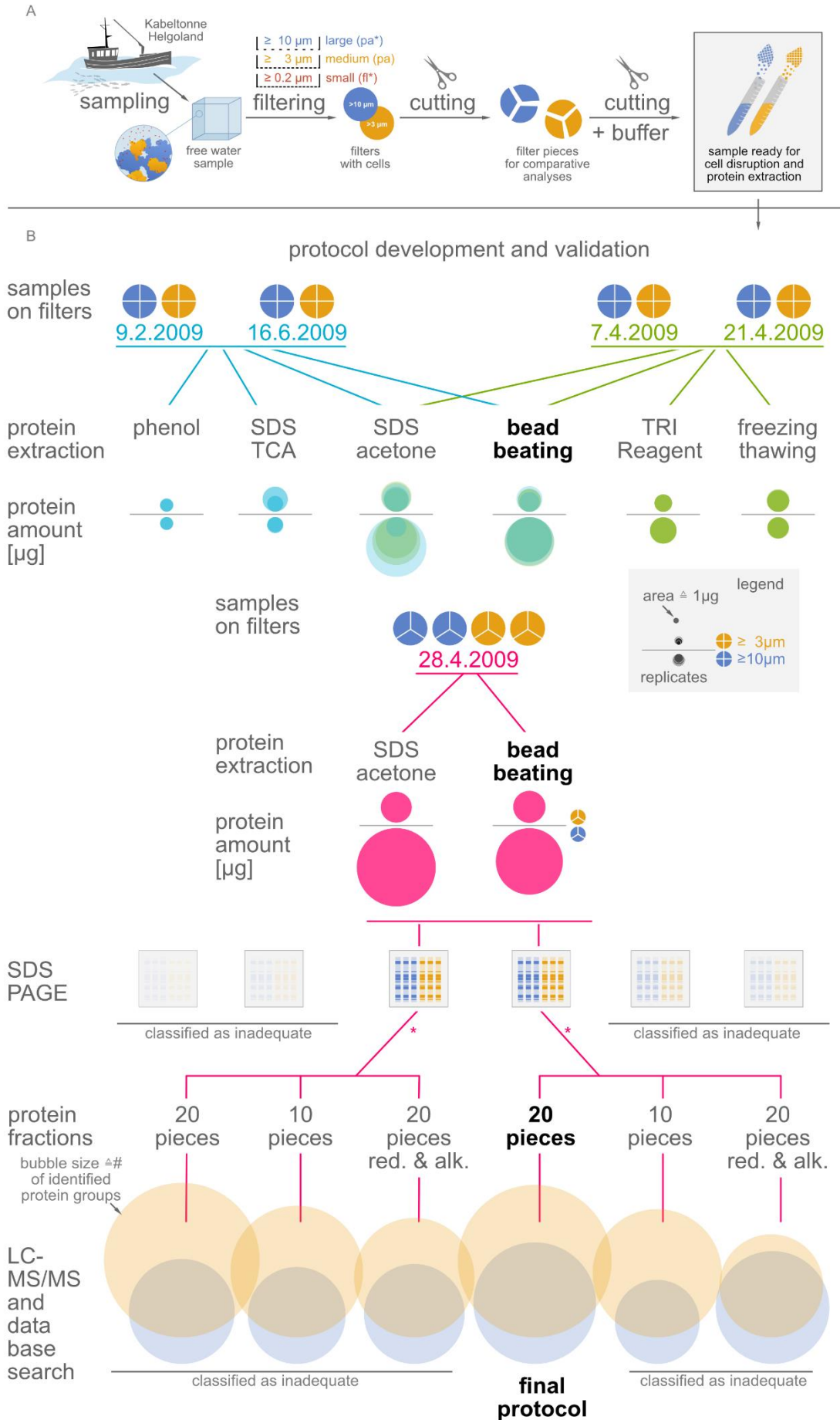


Figure 2. (A) Sampling strategy and (B) evaluation of protein extraction and MS sample preparation protocols. (A) Water samples collected at “Kabeltonne” Helgoland during the spring bloom 2009 were sequentially filtered to obtain the 0.2 - 3 μm (FL-free-living) and two PA (particle-associated, 3 - 10 μm = medium, $\geq 10 \mu\text{m}$ = large) fractions as described in Teeling *et al.* (2012). Filters were initially cut into three or four pieces, which were subsequently shredded and mixed with the respective extraction buffer. **(B)** Filters (medium particles = yellow; large particles = blue) from different sampling time points (turquoise, green and red) were processed according to the six different protocols described in the Experimental Procedure section. With regard to the extracted protein amount, the bead beating and SDS-acetone approaches obviously outcompeted the four other protocols. However, the SDS-acetone protocol was less reproducible than the bead beating protocol. Considering bead beating and SDS-acetone as best performing protocols, they were employed to test different MS sample preparation approaches, i.e., different number of SDS gel fractions for tryptic digestion together with protein reduction (red.) and alkylation (alk.) prior to tryptic digestion. The subsequent LC-MS/MS analyses revealed best results for the bead beating protocol followed by GeLC-MS/MS from 20 fractions without protein reduction and alkylation as shown in the bottom line of the figure. Bubble sizes for the large (blue) and medium (yellow) particles correspond to the number of identified protein groups (see also **Figure 3**). (Schultz *et al.* 2020)

6.1.2. MS sample preparation

Total protein amount was extracted from filters collected on the 28th of April 2009 employing the SDS-acetone and bead-beating method and separated by 1D SDS-PAGE (Figure 2B). To investigate if an increase in the total number of individual gel sub-fractions will lead to more identified protein groups, gel lanes (two technical replicates for each protocol) were cut in either 10 or 20 equal-sized fractions. Subsequently, proteins were *in-gel* trypsin digested and the resultant peptides were subjected to LC-MS/MS analysis. Additionally, it was investigated if reduction and alkylation of the proteins prior to tryptic digestion increased protein group identification rates (Figure 2B). The acquired spectra were searched in the so far available 0.2 μm 2009 (MIMAS) database (Teeling *et al.* 2012). The increase of individual gel sub-fractions led to nearly doubled amount of measured LC-MS/MS spectra. A maximum number of identified protein groups (1230 identified protein groups) was achieved with the SDS-acetone method combined with 20 gel fractions without reduction and alkylation for the 3-10 μm fractions. For the $\geq 10 \mu\text{m}$ fraction highest number of identified protein groups (760 identified protein groups) was obtained with the bead-beating method combined with 20 gel fractions and without reduction and alkylation (Figure 2B, Figure 3).

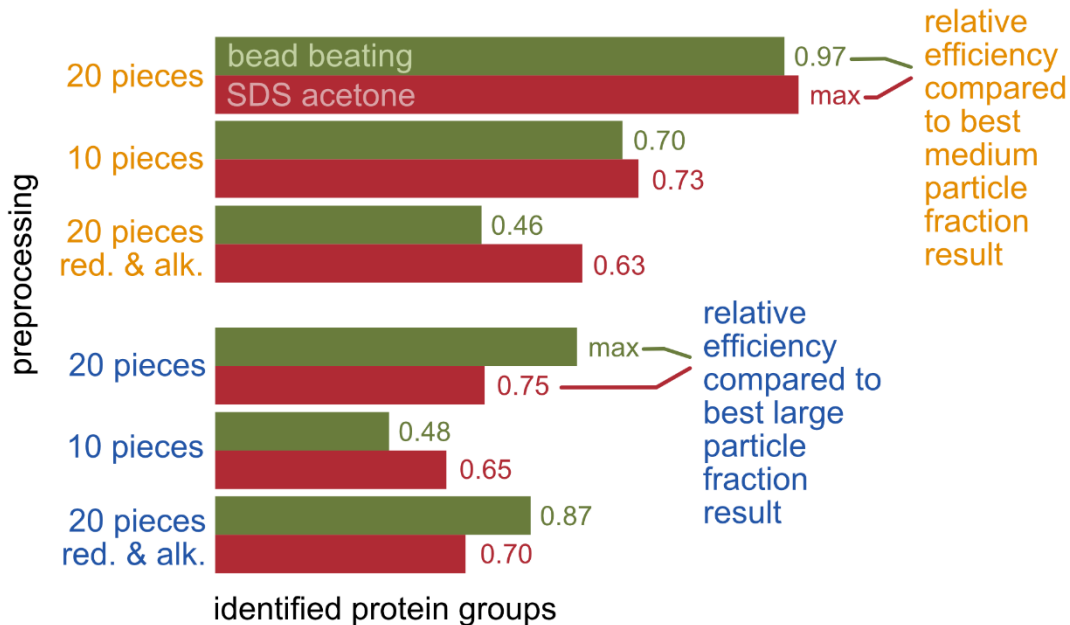


Figure 3. Protein group identifications obtained by different extraction and protein pre-fractionation protocols. For the medium particle size fraction (3 - 10 μm , yellow), 20 gel fractions after standard treatment, i.e. without protein reduction (red.) and alkylation (alk.), resulted in the highest number of identified protein groups, no matter which protein extraction protocol (SDS-acetone (red) or bead beating (green)) was applied. For the large particle fraction ($\geq 10 \mu\text{m}$, blue) the general trend was similar. However, the bead beating protocol performed better compared to the SDS-acetone protocol. (Schultz *et al.* 2020)

6.1.3. Optimization of databases

Simultaneously to the optimization of the metaproteomics protocol, metagenomic sequencing, assembly and annotation of FL (0.2 μm pore-sized filters) and PA (3 and 10 μm pore-sized filters) fractions of water samples collected during the Helgoland spring bloom 2009 was performed. The metagenomic database used for database searches included sequences of FL bacteria (0.2 μm pore-sized filters) and microbial communities present in the medium particulate fraction (3 μm pore-sized filters).

The LC-MS/MS spectra received by the application of the bead-beating protocol were searched against four different databases to identify the database resulting in the highest number of reliably identified protein groups (Figure 4): (a) the non-redundant NCBI database (NCBI_{nr}), (b) a database with Uniprot sequences of abundant bacteria and diatoms identified by Teeling *et al.* 2012 (PABD), (c) the database by Teeling *et al.* 2012 containing protein sequences based on translated metagenomes of FL bacteria (0.2 μm pore-sized filters from different sampling time points) of the spring bloom 2009 (MIMAS) and (d) a database based in

the metagenomes of the 0.2 and 3 μm pore-sized filters from samples of the 14th of April 2009 (0.2 + 3 μm 2009). For both analyzed fractions highest number of identified protein groups were sustained with the 0.2 + 3 μm database (2120 identified protein groups for the 3-10 μm fraction, 1870 identified protein groups for the $\geq 10 \mu\text{m}$ fraction).

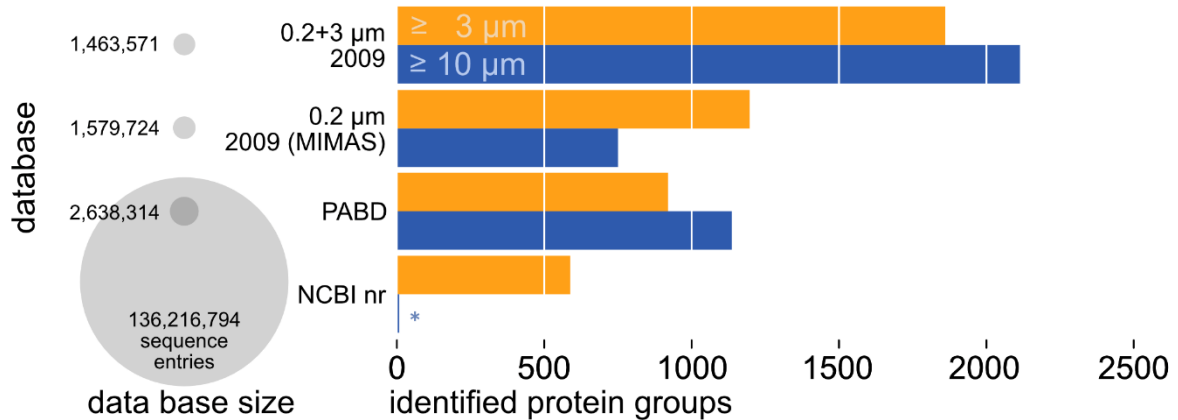


Figure 4. Number of identified protein groups obtained with different databases. (a) the non-redundant NCBI database (NCBI nr, 136,216,794 entries), **(b)** a database with Uniprot sequences of known abundant bacteria and diatoms identified by the study of Teeling *et al.* (2012) (PABD, 2,638,314 entries), **(c)** a metagenome-based database employed for the FL bacterial fraction within the study of Teeling *et al.* (2012) (MIMAS, 1,579,724 entries) and **(d)** a database based on translated metagenomes of the FL fraction on the 0.2 μm filters and particles on the 3 μm filters sampled on the 14th of April 2009 (0.2 + 3 μm 2009, 1,463,572 entries). (Schultz *et al.* 2020)

The bead-beating method was used for protein extraction in all subsequent analyses as it resulted in more reproducible protein yields compared to the SDS-acetone extraction protocol (Figure 2B), was less time-intensive and led to the identification of the highest number of unique protein groups. Summarized, the optimized metaproteomics pipeline for marine particles (Figure 5) included protein extraction by bead-beating, protein fractionation by 1D SDS-PAGE (20 fractions), followed by *in-gel* tryptic digestion, LC-MS/MS measurements and database search against a matching metagenome database.

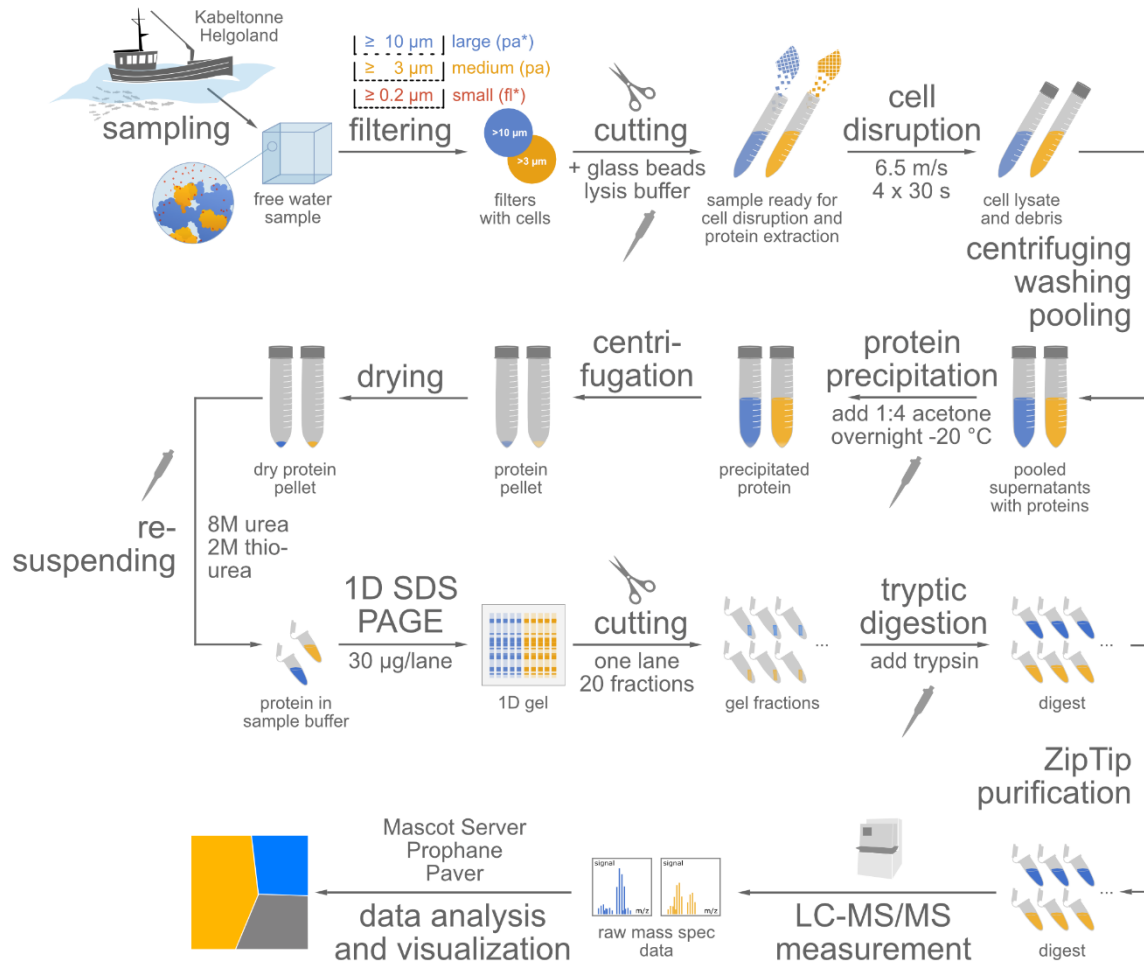


Figure 5. Final metaproteomics protocol for analysis of PA communities. Protein extraction from filters was conducted using 5% (w/v) SDS-containing lysis buffer, cell disruption by FastPrep-mediated bead-beating, separation of proteins by 1D-SDS-PAGE, tryptic *in-gel* digestion, LC-MS/MS analyses on an Orbitrap Velos™ mass spectrometer, MASCOT database search against the metagenome-based database (0.2 + 3 μm 2009) and data-processing and visualization with the *in-house*-developed bioinformatics tools *ProPhane 3.1* and *Paver*. (Schultz *et al.* 2020)

6.2. Comparative metaproteome analysis of FL and PA bacterioplankton harvested during 2009 spring bloom

To evaluate if the optimized protocol is appropriate for a comparative metaproteomics analysis of FL and PA microbial communities, it was applied to various fractions of a microbial community sampled at the 14th of April in 2009 (0.2 – 3 μm = FL; 3 – 10 μm and $\geq 10 \mu\text{m}$ = PA). Five technical replicates (tr) of each sample underwent the final optimized workflow (Figure 5) and resultant MS/MS-data was searched against the matching metagenome-based database (0.2 + 3 μm 2009). Using the optimized pipeline 360,000 to 460,000 spectra per tr and about 20,000 spectra per gel fraction were identified, which subsequently resulted in the

identification of 9,354 protein groups (19.4% of spectral IDs; 89,240 out of 460,000), 2,263 protein groups (10.2% of spectral IDs; 36,720 out of 360,000), and 2,771 protein groups (10.7% of spectral IDs; 47,080 out of 440,000) for the 0.2 – 3 μm (Table S1 CD-ROM), 3 – 10 μm (Table S2 CD-ROM), and $\geq 10 \mu\text{m}$ (Table S3 CD-ROM) fractions, respectively.

1,956 of the identified protein groups of the two PA fractions could also be identified in the FL fraction whereas 276 protein groups were exclusively identified in the PA fractions (Figure 6).

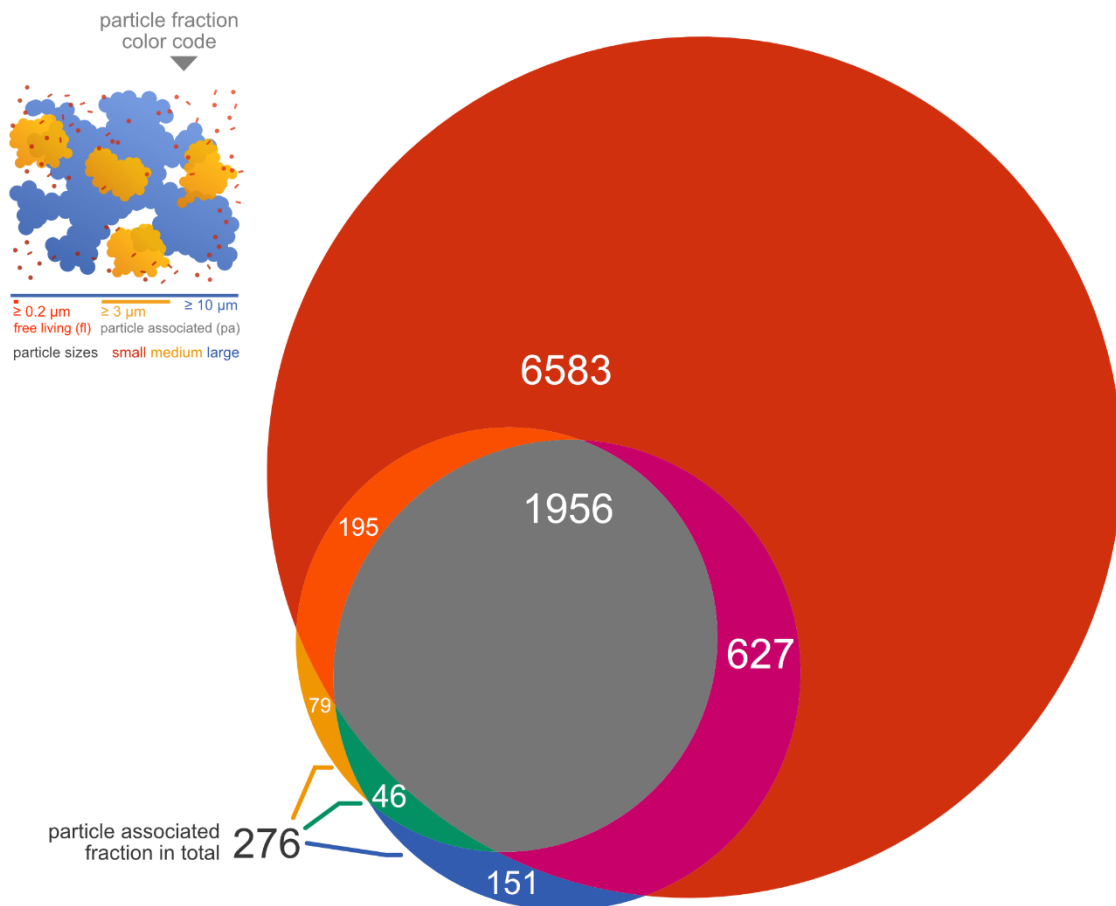


Figure 6. Venn diagram showing the number of fraction-specific and shared protein groups of the planktonic and particulate fractions. The FL (0.2 μm) fraction is depicted in red, the medium PA fraction (0.2 – 3 μm) in yellow and the large PA fraction ($\geq 10 \mu\text{m}$) in blue. (Schultz *et al.* 2020)

6.2.1. Taxonomic differences between FL and PA fractions

Despite the similarity of the FL and PA metaproteomic datasets, phylogenetic assignment of the identified protein groups suggested some remarkable taxonomic differences between the FL and PA fractions (Figure 7; Table S4 CD-ROM). PA fractions comprised substantially more

Results

eukaryotic protein groups than the FL fraction (Figure 7A). These protein groups made up 43% (3 – 10 μm fraction) and 54% ($\geq 10 \mu\text{m}$ fraction) of the protein groups identified from particles, but only 11% of the protein groups identified in the planktonic fraction. Many of these protein groups could be assigned to known phytoplankton taxa (Figure 8), including numerous microalgal groups, e.g., diatoms, *Pelagophytes*, *Raphidiophytes*, *Cryptophytes*, *Dinoflagellates* and *Haptophytes*, but also fungi and various protozoa (Table S1 CD-ROM, Table S2 CD-ROM). For instance, abundant protein groups could be assigned to *Oomycetes* (water molds, e.g., *Peronosporales*, *Saprolegniales*) and fungi (e.g., *Cryptomycota*). Noteworthy, the number of viral protein groups was almost three times higher in the two particulate fractions compared to their planktonic counterparts (Figure 7B).

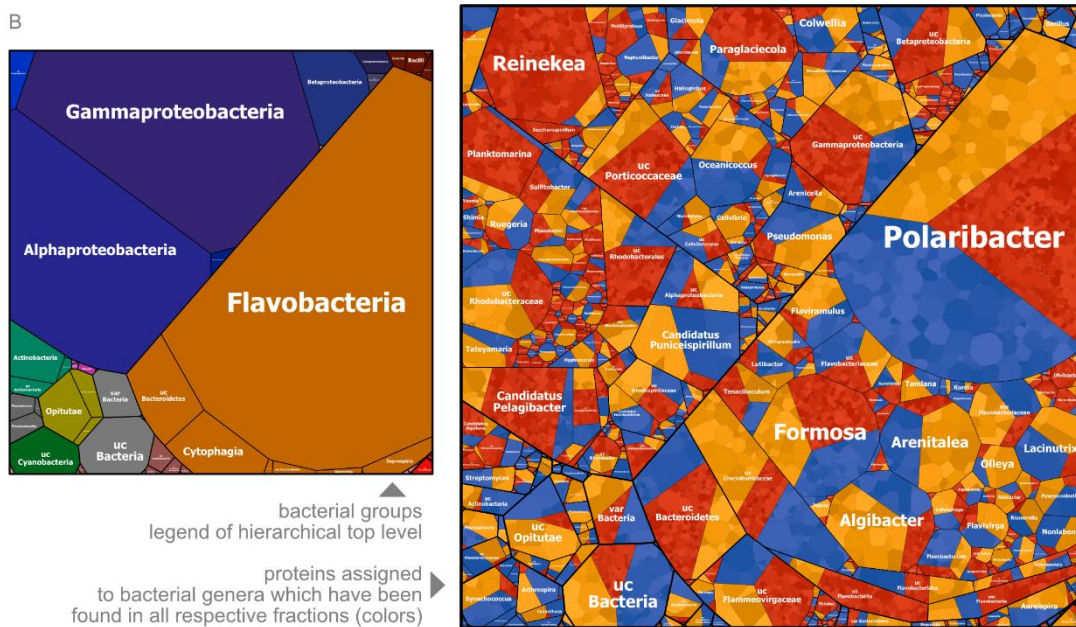
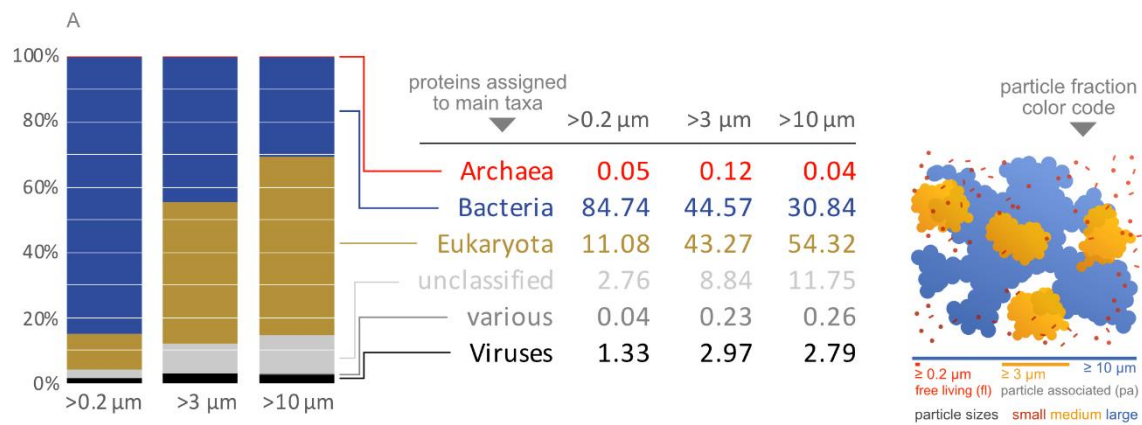


Figure 7. Taxonomic affiliation of protein groups of FL and PA metaproteomes during the spring bloom on 14th of April 2009 at “Kabeltonne” Helgoland. (A) Distribution of pro- and eukaryotes in the FL (0.2 - 3 μm) and PA (3 - 10 μm , \geq 10 μm) fractions based on the relative abundance of protein groups assigned to the different phylogenetic groups. **(B)** Voronoi tree maps visualizing the phylogenetic assignment of bacterial protein groups identified in FL (red) and PA (yellow and blue) fractions. Cell size corresponds to the relative abundance of the respective bacterial genus on protein level. Protein groups of *Reinekea* for example are most abundant in the FL fraction and are therefore encoded by a large red tree map cell. In the PA fractions they can be detected only in traces, resulting in very small cell sizes (coloured in yellow and blue). *Algibacter* protein group abundance, on the other hand, was notably higher in the PA fractions, compared to the FL fraction (Schultz *et al.* 2020).

The most abundant bacterial phyla within both fractions, were *Proteobacteria* (FL 55%; PA 41% and 39%, 3 μm and 10 μm pore-sized filters) and *Bacteroidetes* (FL 40%, PA 48% and 47%, 3 μm and 10 μm pore-sized filters). Protein groups assigned to *Alpha-*, *Beta-* and *Gammaproteobacteria* were in general more abundant in the FL bacteria, whereas protein groups assigned to *Cyanobacteria* (e.g. *Synechococcus*, *Arthrospira*), *Opituae*, *Flavobacteriia* (e.g. *Arenitalea*, *Olleya*, *Algibacter*, *Lacinutrix*), and some proteobacterial genera (e.g. *Oceanicoccus*, *Canidatus* *Puniceispirillum*, *Neptuniibacter*, *Halioglobus*, *Ramlibacter*, *Pseudomonas*) were more abundant in the PA fraction (Figure 7).

Results

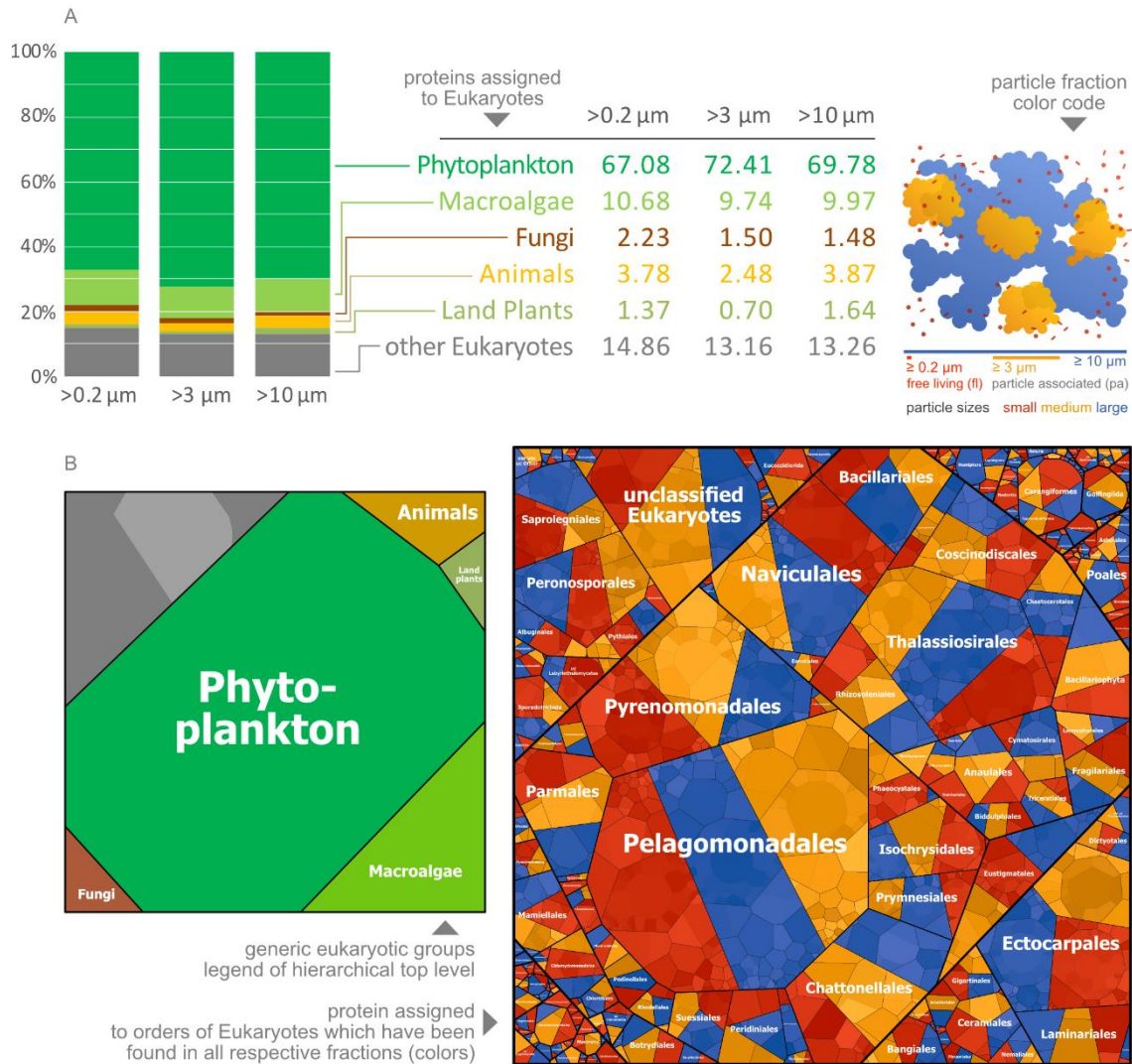


Figure 8. Phylogenetic assignment of eukaryotic protein groups present in the FL and PA fractions during the spring bloom on 14th of April 2009 at “Kabeltonne” Helgoland. (A) Distribution of different eukaryotes in the FL (0.2 μm) and PA (0.2 - 3 μm and $\geq 10 \mu\text{m}$) fractions as shown by relative protein abundances assigned to the different eukaryotic phylogenetic groups. **(B)** Voronoi tree map visualizing the relative abundance of eukaryotic taxa based on the abundance of assigned protein groups extracted from the FL (red) and PA (yellow and blue) fractions. Cell size corresponds to the relative abundance of the respective genus. In this preliminary analysis, protein group identification is based on metagenomic (DNA-based) information from the filtered fractions, which suffers limitations for eukaryotic protein group identification, probably resulting in incomplete functional and taxonomic profiles. (Schultz *et al.* 2020)

6.2.2. PA specific functional differences between FL and PA bacterioplankton

The SusC/D utilization system, specific GHs, i.e. GH family 1, 13, and 16 (including beta-glucosidases, alpha-1,4-amylases, and exo- and endo-1,3-beta-glycanases), GTs and TBDT showed higher overall expression levels in the PA fractions than in the FL fraction (Figure 9B). Additionally, sulfatases were mainly expressed by *Flavobacteriia*, especially *Formosa sp.* (Figure 9A and B).

Results

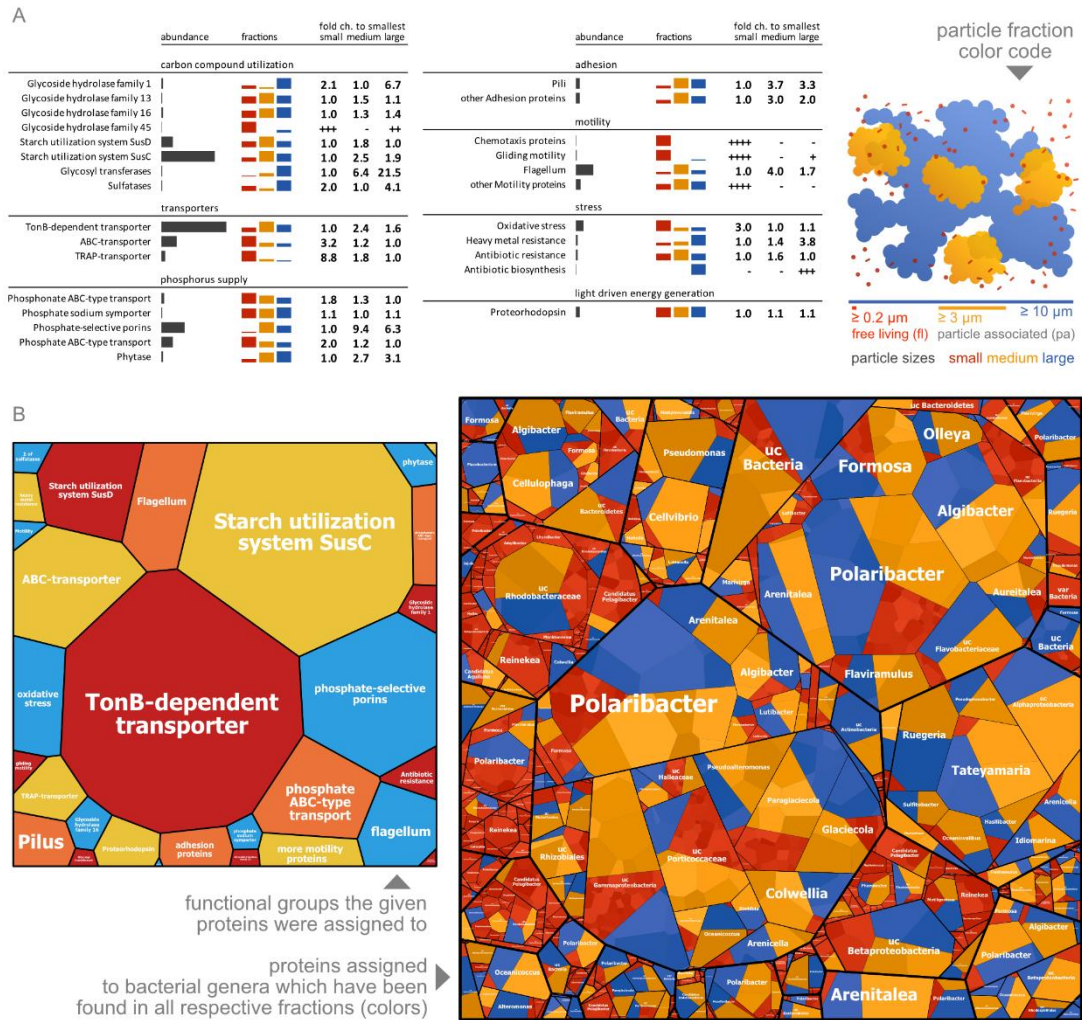


Figure 9. Functional assignment of protein groups in FL and PA metaproteomes during the spring bloom on 14th of April 2009 at “Kabeltonne” Helgoland. (A) Total (black bars) and relative (red for FL, yellow and blue for PA) abundance of selected protein groups with assigned functions in the FL (0.2 – 3 μm = small) and PA (3 - 10 μm = medium and $\geq 10 \mu\text{m}$ = large) fractions. **(B)** Voronoi tree maps showing the phylogenetic assignment of selected functional protein groups identified in FL (red) and PA (yellow and blue) fractions. Cell size corresponds to the relative abundance of the respective genus within specific functional categories. (Schultz *et al.* 2020)

In FL bacteria phosphate and phosphonate ABC-type transporter were highly expressed, whereas PA bacteria seem to express mainly phytases and phosphate-selective porins. Moreover, several protein groups involved in oxidative stress defence seem to be less abundant in the PA fractions, whereas protein groups for heavy-metal and antibiotic resistance were strongly expressed in the $\geq 10 \mu\text{m}$ fraction (Figure 9A and B). Protein groups involved in adhesion and motility, i.e. flagella and type IV pili, as well as proteorhodopsin were more abundant in the particulate fractions.

6.3. Investigation of phylogenetic and functional succession of marine particles during 2018 spring phytoplankton bloom

To investigate succession of taxonomical clades and expressed functions of marine particles during the course of a spring phytoplankton bloom, the PA fractions (3 – 10 µm and ≥ 10 µm) of three sampling time points (three tr of each sample) of a bloom in 2018 ranging from pre-bloom conditions to the peak of the bacterial response bloom were analyzed, using the above optimized metaproteomics pipeline.

The spring phytoplankton bloom 2018 at Helgoland roads started at the end of April with chlorophyll a (chl a) amounts rising from under 5 µg/L up to 22 µg/L during a first peak in the beginning of May and after a decline to around 6 µg/L a second bloom peak established in the mid of May with chl a amounts of around 34 µg/L. The bacterial response bloom started in the beginning of May with total cell counts (tcc) ranging from around 7×10^5 /ml before the bloom to 3.2×10^6 /ml during its peak at the end of May. For the metaproteomics analysis three sampling time points based on the ongoing of the bacterial response bloom were chosen to investigate the taxonomic and functional succession of marine particles: pre- bloom (17th of April; chl a 4.6 µg/L; tcc 7.22×10^5 /ml), bloom rise (8th of May; chl a 7.3 µg/L; tcc 1.44×10^6 /ml) and bloom peak (24th of May; chl a 4.7 µg/L; tcc 3.15×10^6 /ml) (spring-bloom sampling campaign 2018 at Helgoland Roads, chl a and tcc measurements by MPI Bremen, person in authority Dr. Bernhard Fuchs) (Figure 10).

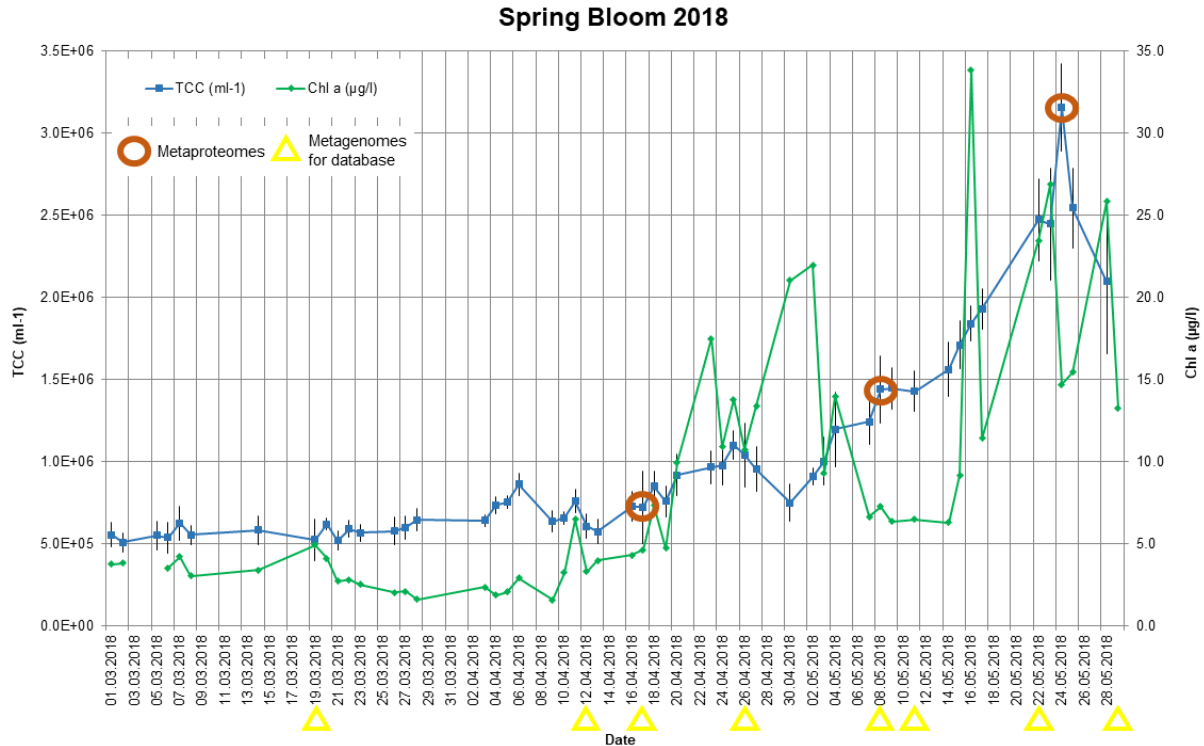


Figure 10. Spring bloom 2018 at “Kabeltonne” Helgoland. Measured chlorophyll a (Chl A) amounts in $\mu\text{g/L}$ reflecting the ongoing of the phytoplankton bloom from beginning of March to end of May 2018 are depicted in green. Total cell counts (TCC) in ml^{-1} reflecting the ongoing of the bacterial response bloom are depicted in blue. Orange circles highlight the time points of metaproteomics analysis of the spring phytoplankton bloom (17th of April = pre-bloom, 8th of May = bloom rise, 24th of May = bloom peak according to the bacterial response bloom). The yellow triangles highlight the time points of metagenomic analysis used for the construction of a protein sequence reference database for metaproteomic analysis. Data provided by Dr. Bernhard Fuchs at the Max-Planck Institute of Marine Microbiology in Bremen.

Protein extraction and MS-sample preparation was performed according to the established metaproteomics pipeline described in 6.1. The resulting MS/MS-data was searched against a matching metagenome-based protein sequence database (3 + 10 μm filter 2018; based on eight sampling time points from mid-March to end of May during the investigated bloom, Figure 10). 339,980 to 390,869 spectra per tr were recorded, which subsequently led to the identification of 1,614 to 3,084 (14.7% - 18% of spectral IDs) (Table 7Table S5 CD-ROM).

Results

Table 7. Metaproteomic investigation of a metaproteomic analysis of marine particles during a spring phytoplankton bloom in 2018. Used abbreviation: FDR (false discovery rate).

sample	number of identified protein groups	number of measured spectra [Ø]	number of identified spectra in %	FDR in %
pre-bloom 3 – 10 µm	3,084	386,463	15.0	0.6
pre-bloom ≥ 10 µm	2,523	390,869	15.4	0.7
bloom rise 3 – 10 µm	3,710	376,626	14.7	0.3
bloom rise ≥ 10 µm	3,431	388,107	17.4	0.8
bloom peak 3 – 10 µm	2,504	347,049	17.4	0.6
bloom peak ≥ 10 µm	1,614	339,980	18.0	0.7

Pre-bacterial bloom samples contained more identified protein groups assigned to eukaryotes than the bloom rising and bloom peak samples (Figure 11). These protein groups comprised 80% (3 – 10 µm fraction) and 86% (≥ 10 µm fraction) of the protein groups identified in pre-bloom samples, but contributed only 55% (3 – 10 µm fraction) and 62% (≥ 10 µm fraction) of the protein groups identified in bloom peak samples.

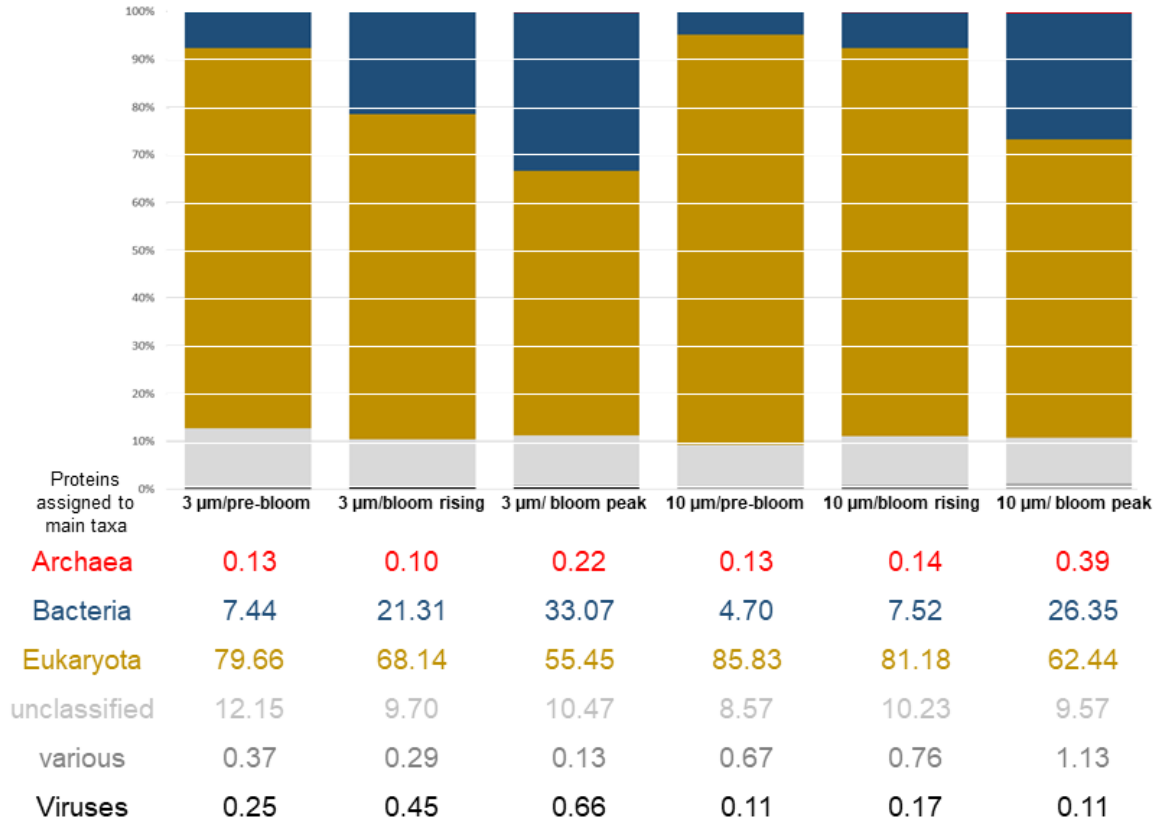


Figure 11. Taxonomic affiliation of protein groups of PA metaproteomes during the spring phytoplankton bloom 2018 at “Kabeltonne” Helgoland. Distribution of pro- and eukaryotes in the 3 – 10 µm (3 µm) and ≥ 10 µm (10 µm) fractions in pre-bloom (17th of April), bloom rising (5th of May) and bloom peak (24th of May) samples based on the relative abundance of protein groups assigned to the different phylogenetic groups.

6.3.1. PA bacterioplankton taxonomic succession during the 2018 spring phytoplankton bloom

To get a first overview of the data PCA and OPLS-DA plots were generated (Figure 12). The principal component analysis (PCA) is a qualitative method of analysis during which a coordinate transformation is performed. The principal components are organized according to their importance and depict an orthogonal coordinate system. Thus, it is possible to project a high-dimensional dataset into a low-dimensional space and to show the coherence in a graphic way. With the orthogonal partial least squares discriminant analysis (OPLS-DA) it is possible to identify which variables are driving the separation between different groups. For the above described analyses the NSAF-values of all identified protein groups assigned to the same taxonomical family were summed up as variables. To handle the inhomogeneity within the

Results

absolute values of the variables and secure that each variable has the same influence on the statistical model, the variables were centred and scaled to “Unit Variance” (UV). For the OPLS-DA plots the two different filter sizes (3 and 10 μm) as well as the three different sampling time points (pre-bloom, bloom rise and bloom peak) were defined as groups. Within the different plots samples that are located in proximity to each other and cluster are similar and show only a slight variance, while samples with the greatest distance show the highest variance. The PCA plot shows, that both size fractions of the pre-bloom sampling time point and the $\geq 10 \mu\text{m}$ fraction of the sampling time point during the rise of the bacterial bloom cluster together (Figure 12A). The results of the OPLS-DA plots show, that the two analysed size fractions of the individual sampling time points cluster together in the plot based on sampling time point (Figure 12B). In the OPLS-DA plot based on filter size the 3 – 10 μm and $\geq 10 \mu\text{m}$ fractions show different trends: the variance between the different sampling time points is bigger in the 3 – 10 μm fractions and within the $\geq 10 \mu\text{m}$ fraction the bloom rise and bloom peak sampling time points cluster together (Figure 12C).

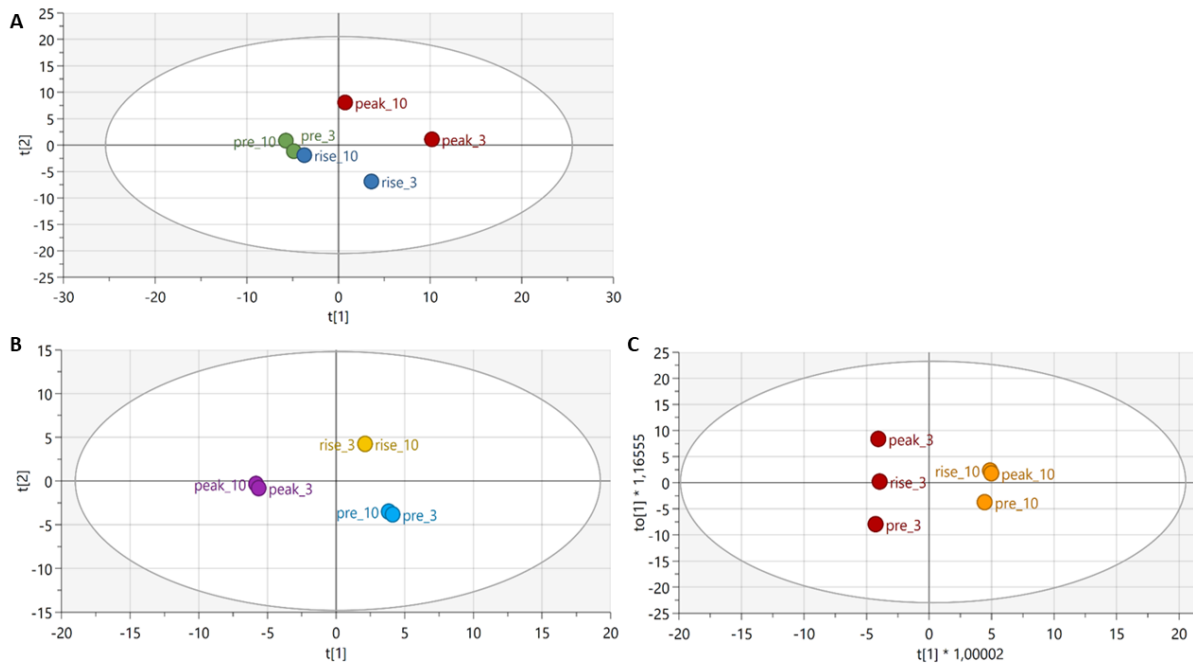


Figure 12. Multivariate analysis of bacterial taxonomical families identified during metaproteomic analysis of marine particles during a spring phytoplankton bloom in 2018. (A) PCA plot of bacterial families. (B) OPLS-DA plot of bacterial families based on sampling time points during the spring phytoplankton bloom 2018 at “Kabeltonne” Helgoland. (C) OPLS-DA plot of bacterial families based on pore-size of the used filters (3 and 10 μm).

The most abundant bacterial phyla within the analyzed PA fractions, were *Proteobacteria* (3 μm pre-bloom 61.7%, bloom rise 64.3%, bloom peak 62.3%; 10 μm pre-bloom 62.3%, bloom rise 64.1%; bloom peak 75.5%) and *Bacteroidetes* (3 μm pre-bloom 11.8%, bloom rise 13.8%, bloom peak 16.8%; 10 μm pre-bloom 12.9%; bloom rise 12.2%, bloom peak 10.0%) (Figure 13; Table S6 CD-ROM). Protein groups expressed by *Cyanobacteria*, *Planctomycetes* and *Verrucomicrobia* were generally more abundant in the 3 – 10 μm fraction, whereas protein groups expressed by *Actinobacteria* and *Proteobacteria*, especially by the classes *Beta*- and *Deltaproteobacteria*, were higher abundant in the ≥ 10 μm fraction. Protein groups assigned to the alphaproteobacterial genera *Canidatus Pelagibacter*, *Planktomarina* and *Tateyamaria* as well as to the flavobacterial genera *Ulvibacter* and *Formosa*, were generally more dominant in the 3 – 10 μm fraction, whilst protein groups assigned to the alphaproteobacterial genus *Roseobacter* and to the flavobacterial genus *Tenacibaculum* were more abundant in the ≥ 10 μm fraction (Figure 13). Besides the differences in taxonomic distributions between the two particulate fractions, the phylogenetic assignment of the identified protein groups also indicated some notable taxonomic differences between the different stages of the investigated bacterial bloom in response to the spring phytoplankton bloom 2018. Protein groups of bacterial taxa generally more abundant in the pre-bloom stage were assigned to *Cyanobacteria* and the gammaproteobacterial genus *Pseudomonas*. Protein groups of bacterial taxa generally more abundant in the rise of the bacterial bloom were assigned to the verrucomicrobial genus *Verrucomicrobiae*, the planctomycetal genus *Phycisphaera* in the 3 – 10 μm fraction, as well as the flavobacterial genus *Polaribacter* in the ≥ 10 μm fraction. During the peak of the bacterial bloom protein groups assigned to the alphaproteobacterial genera *Sulfitobacter*, *Tateyamaria*, *Roseovarius* and *Roseobacter*, as well as to the verrucomicrobial genus *Opituae* were generally more abundant (Figure 13). Additionally, protein groups assigned to the alphaproteobacterial genus *Planktomarina*, and the flavobacterial genus *Tenacibaculum* were more abundant during the peak of the bloom in the 3 – 10 μm fraction and protein groups assigned to the actinobacterial genus *Rhodococcus*, the alphaproteobacterial genera *Hyphomonas* and *Roseobacter*, the gammaproteobacterial

genera *Oceanicoccus* and *Psychrospira* as well as to the flavobacterial genera *Muricauda* and *Leeuwenhoekiella* were more abundant in the $\geq 10 \mu\text{m}$ fraction.

Results

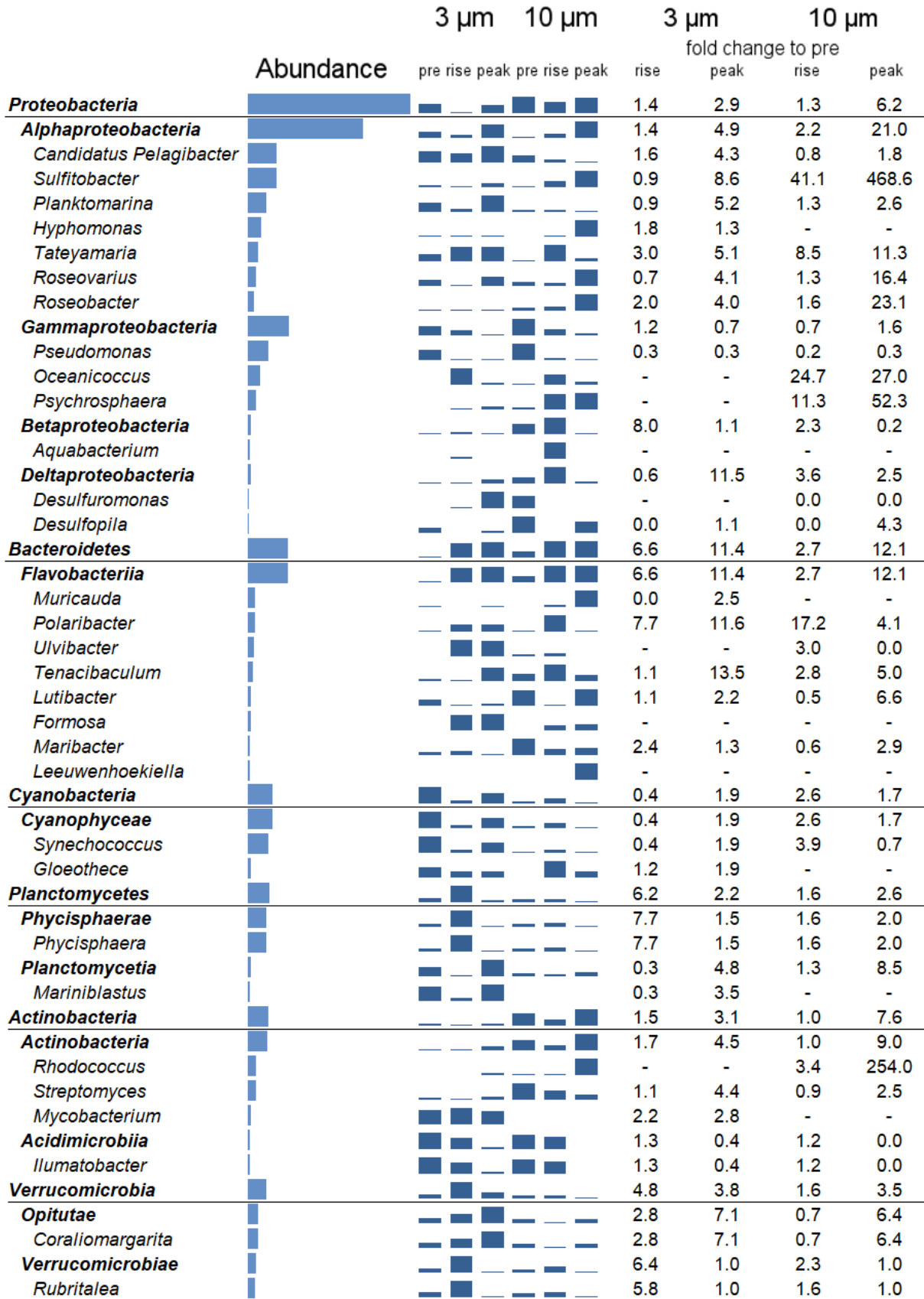


Figure 13. Taxonomic affiliation of prokaryotic protein groups in metaproteomes of pre-bloom, bloom rise and bloom peak samples during the spring bloom 2018 at "Kabeltonne" Helgoland. Total (horizontal bars) and relative (vertical bars) abundance of selected protein groups with assigned taxa. Additionally, fold changes compared to pre-bloom metaproteome calculated based on NSAF values are depicted.

6.3.2. Eukaryotic taxonomic succession during the 2018 spring phytoplankton bloom

To get a first overview of the data, again, PCA and OPLS-DA plots were generated. For the analyses the NSAF-values of all identified protein groups assigned to the same taxonomical class were summed up as variables. To handle the inhomogeneity within the absolute values of the variables and secure that each variable has the same influence on the statistical model, the variables were centred and scaled to “Unit Variance” (UV). For the OPLS-DA plots the two different filter sizes as well as the three different sampling time points were defined as groups. Within the different plots samples that are located in proximity to each other and cluster are similar and show only a slight variance, while samples with the greatest distance show the highest variance. The PCA plot shows that the two analysed fractions differ from each other and that the samples from pre-bloom and peak of the bloom cluster together for the 3 – 10 μm fraction (Figure 14A). The results of the OPLS-DA plots show that the two analysed size fractions of the individual sampling time points cluster together in the plot based on sampling time point (Figure 14B). In the OPLS-DA plot based on filter size the 3 – 10 μm and $\geq 10 \mu\text{m}$ fractions show different trends: the variance between the different sampling time points is bigger in the $\geq 10 \mu\text{m}$ fractions. Within the 3 – 10 μm the bloom rise and bloom peak sampling time points cluster while in the $\geq 10 \mu\text{m}$ fraction the pre-bloom and bloom peak sampling time points cluster together (Figure 14C).

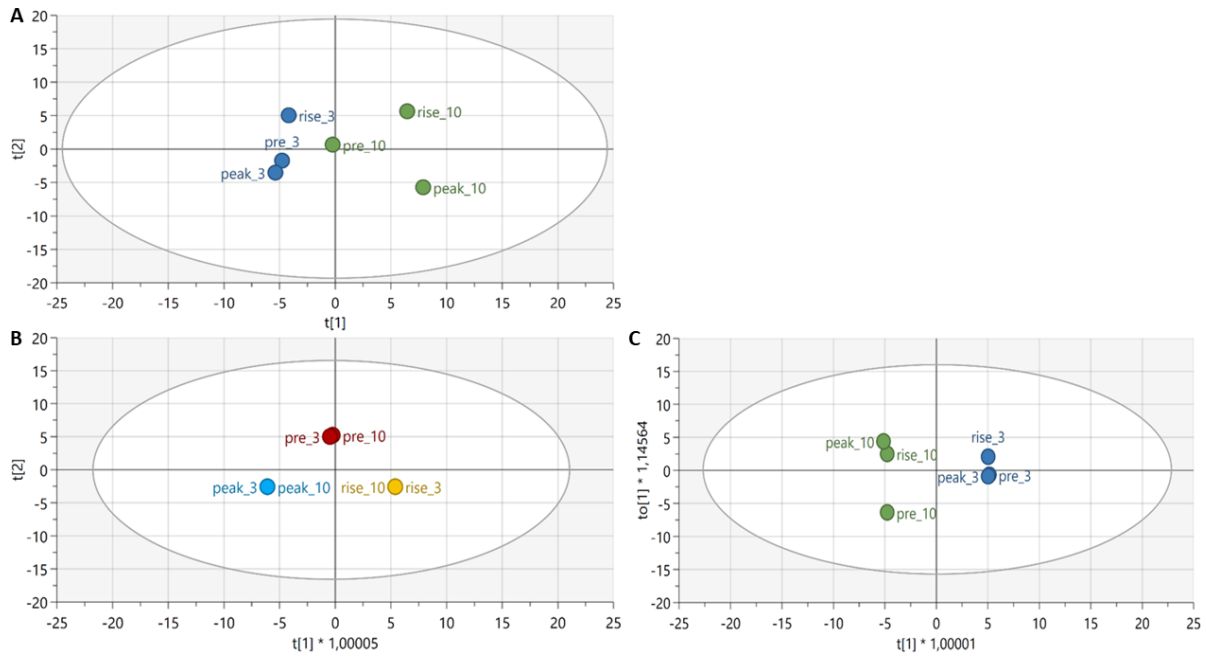


Figure 14. Multivariate analysis of eukaryotic taxonomical classes identified during metaproteomic analysis of marine particles during a spring phytoplankton bloom in 2018. (A) PCA plot of eukaryotic classes. (B) OPLS-DA plot of eukaryotic classes based on sampling time points during the spring phytoplankton bloom. (C) OPLS-DA plot of eukaryotic classes based on pore-size of the used filters (3 and 10 μm).

Many of the identified eukaryotic protein groups were assigned to known phytoplankton taxa: 89.7% and 78.8% of identified eukaryotic protein groups in the pre-bloom samples for 3 – 10 μm and $\geq 10 \mu\text{m}$ fractions respectively, 86.7% and 46.4% in the bloom rise samples and 86.7% and 50.7% in the bloom peak samples (Figure 15; Table S7 CD-ROM). These include numerous microalgal groups, e.g., diatoms (*Bacillariophyta*), *Cryptophytes* and *Chlorophytes* but also various *Arthropoda*, protozoa and fungi such as *Ascomycota* and *Basidiomycota* were identified. Within the *Bacillariophyta* the order *Thalassiosirales* and genus *Thalassiosira* were the most abundant. Protein groups expressed by protozoa and fungal *Chytridiomycota* were generally more dominant in the 3 – 10 μm fraction, whilst protein groups assigned to *Arthropoda* and the fungal phylum *Basidiomycota* were more abundant in the $\geq 10 \mu\text{m}$ fractions. Besides the differences in taxonomic distributions between the two particulate fractions, the phylogenetic assignment of the identified protein groups also indicated some notable taxonomic differences between the different stages of the investigated bacterial bloom in response to the spring phytoplankton bloom 2018 (Figure 15). Protein

groups of eukaryotic taxa generally more abundant in the pre-bloom stage were assigned to the microalgal phyla *Bacillariophyta* and *Cryptophyta* and to protozoa in the 3 – 10 μm fraction. Protein groups assigned to the microalgal phyla *Chlorophyta*, *Ochrophyta* and *Streptophyta* were generally more abundant during the rise of the bacterial bloom as well as protozoa and fungal phylum *Ascomycota* in the ≥ 10 μm fractions. Protein groups assigned to the fungal phyla *Chytridiomycota* and *Basidiomycota* were generally abundant in the pre-bloom and bloom rise samples in the 3 – 10 μm and ≥ 10 μm fractions, respectively. During the peak of the bacterial bloom protein groups assigned to the fungal phylum *Chytridiomycota* and therein the order *Chytridiales* were generally dominant, as well as macroalgal phylum *Rhodophyta* and therein mainly the order *Ceramiales* in the ≥ 10 μm fractions (Figure 15).

Results

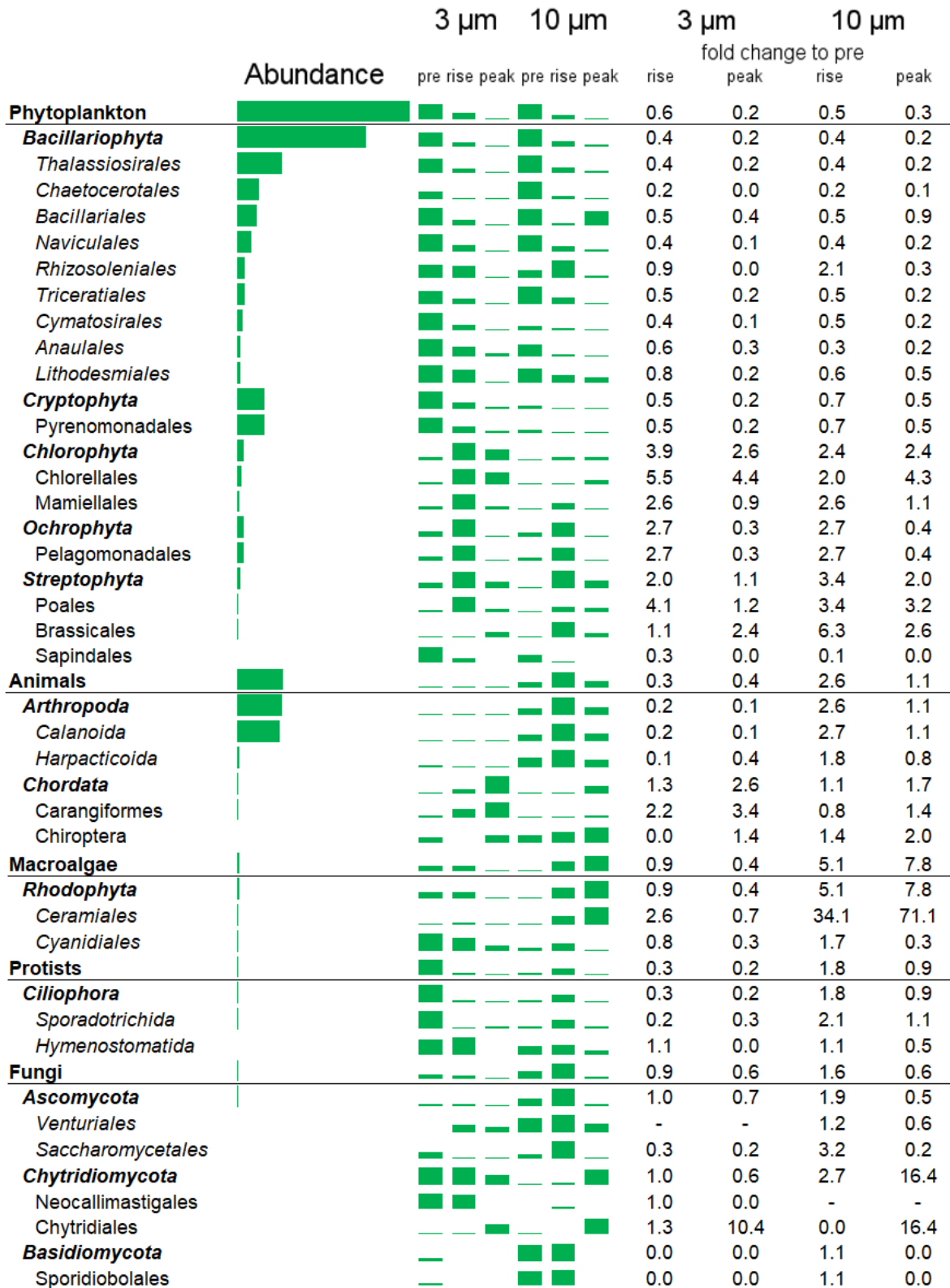


Figure 15 Taxonomic affiliation of eukaryotic protein groups in metaproteomes of pre-bloom, bloom rise and bloom peak samples during the spring bloom 2018 at "Kabeltonne" Helgoland. Total (horizontal bars) and relative (vertical bars) abundance of selected protein groups with assigned taxa. Additionally, fold changes compared to pre-bloom metaproteome calculated based on NSAF values are depicted.

6.3.3. PA bacterioplankton functional succession during the 2018 spring phytoplankton bloom

In addition to ribosomal subunits and chaperonins, most prevalent identified protein groups assigned to bacteria belonged to transport and binding proteins (Figure 16; Table S8 CD-ROM). Therein the most abundant group were ABC transporter proteins, which were generally more abundant in the 3 – 10 μm fraction and during the peak of the bacterial bloom. Phosphate- and amino acid transporter, oligosaccharide- and monosaccharide transporter belonged to the dominant ABC transporter classes. The second most abundant transporter group were porins, where gram-negative porins, phosphate-selective porins O and P as well as OmpA were the dominant groups. Additionally, protein groups belonging to the SusC/D utilization system and TBDT were identified as abundant transport and binding proteins. Thereby TBDT were more abundant during the rise and peak of the bloom while e.g., SusC was more abundant in the pre-bloom and bloom rise samples. The expression of the group transport and binding proteins was dominated by *Alpha-* and *Gammaproteobacteria*, bacteroidetal *Flavobacteriia* and *Verrucomicrobia*. Besides transporters also protein groups belonging to Carbohydrate metabolism could be identified, namely specific GHs, i.e., GH family 2, 17 (β -1,3-glucanases), 27, 28, 38, GTs, i.e. family 1, 2 and 39 as well as PLs, i.e., family 4 and 17 (including alginate lyases) and CEs. GHs were generally more abundant in the pre-bloom stage in the 3 – 10 μm fraction and in the bloom rise phase in the $\geq 10 \mu\text{m}$ fractions. Moreover, protein groups belonging to maltose and xylose metabolism could be identified.

Furthermore, protein groups belonging to the group transcription and protein synthesis were found with high overall expression, especially in the bloom rise and peak samples.

Results

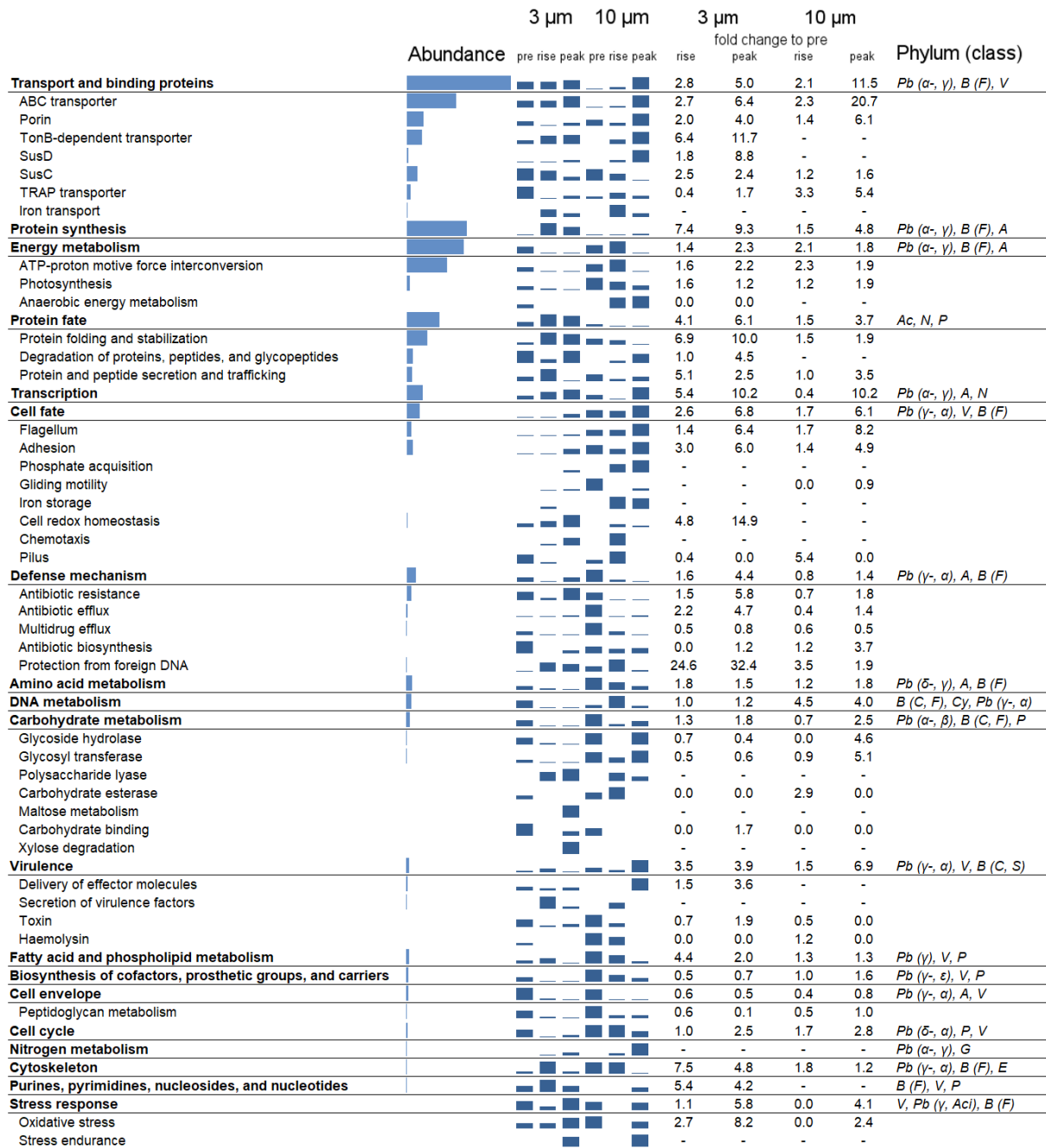


Figure 16. Functional assignment of prokaryotic protein groups in metaproteomes of pre-bloom, bloom rise and bloom peak samples during the spring bloom 2018 at "Kabeltonne" Helgoland. Total (horizontal bars) and relative (vertical bars) of selected protein groups with assigned functions. Additionally fold changes compared to pre-bloom metaproteome calculated based on NSAF values. Used abbreviations: Pb-Proteobacteria, B-Bacteroidetes, F-Flavobacteriia, V-Verrucomicrobiae, A-Actinobacteria, Ac-Acidobacteria, N-Nitrospirae, P-Planctomycetes, Cy-Cyanobacteria, C-Cytophagia, S-Saprospira, G-Gemmatimonadetes, E-Elusimicrobia and Aci-Acidithiobacillia.

Protein groups involved in motility, i.e., flagella- and gliding-mediated motility were more abundant in the $\geq 10 \mu\text{m}$ fractions, as well as protein groups involved in adhesion, phosphate acquisition, iron storage and toxin production. During the pre-bloom stage protein groups

involved in the biosynthesis of cofactors, prosthetic groups and carriers, in gliding motility (in $\geq 10 \mu\text{m}$ fractions), pilus mediated motility (in the 3 – 10 μm fractions), antibiotic resistance as well as antibiotic- and multidrug efflux (in $\geq 10 \mu\text{m}$ fractions) were generally more expressed. In the stage of the bloom rise protein groups involved in pilus-mediated motility (in $\geq 10 \mu\text{m}$ fractions), adhesion (in the 3 – 10 μm fractions), toxin production (in $\geq 10 \mu\text{m}$ fractions), protection from foreign DNA and secretion of virulence factors were generally more abundant. Finally, protein groups involved in flagellum-mediated motility, adhesion (in $\geq 10 \mu\text{m}$ fractions), phosphate acquisition, cell redox homeostasis (in the 3 – 10 μm fractions), antibiotic resistance (in the 3 – 10 μm fractions), delivery of effector molecules as well as oxidative stress showed a higher overall abundance in the peak of the bacterial bloom (Figure 16; Table S8 CD-ROM). Identified protein groups were mainly assigned to the most abundant taxonomic phyla and classes, thus *Alpha-* and *Gammaproteobacteria* as well as *Flavobacteriia* together with one of the following phyla: *Verrucomicrobiae*, *Actinobacteria*, *Planctomycetes* and *Acidobacteria*. One exception is the group protein fate, within which the identified protein groups were mainly assigned to *Actinobacteria*, *Nitrospirae* and *Planctomycetes* (Figure 16). Throughout the bacterial response bloom similar patterns to the overall abundance of the phyla could be observed (Figure S 2).

6.3.4. Eukaryotic functional succession of phytoplankton during the 2018 spring phytoplankton bloom

Most of the identified protein groups assigned to phytoplankton species belonged to the group Energy metabolism and therein Photosynthesis (Figure 17; Table S9 CD-ROM). Photosynthesis-related protein groups were generally more abundant in the 3 – 10 μm fractions as well as during the peak of the bacterial bloom and mainly assigned to the classes *Bacillariophyceae*, *Coscinodiscophyceae* (both phylum *Bacillariophyta*) and to the phyla *Haptophyta* and *Cryptophyta*. Protein groups assigned to the group of transporter and binding proteins were mainly transmembrane ATPases (V-type ATPases), which were more abundant in the $\geq 10 \mu\text{m}$ fraction and in the pre-bloom stage, as well as eukaryotic porins, which were

Results

more dominant in the 3 – 10 µm fractions and in the pre-bloom stage. Moreover, also carbohydrate metabolism associated protein groups could be identified, i.e. GHs family 1, 3, 63 (glucosidases), 78 (rhamnosidases) and 81 (endo-β-1,3-glucanases), GTs family 1, 2, 24, and 28 as well as CBM family 25. They were mainly expressed in the pre-bloom stage and by the phyla *Bacillariophyta* (classes *Coscinodiscophyceae*, *Bacillariophyceae*), *Chlorophyta* and *Streptophyta*.



Figure 17. Functional assignment of phytoplankton protein groups in metaproteomes of pre-bloom, bloom rise and bloom peak samples during the spring bloom 2018 at “Kabeltonne” Helgoland. Total (horizontal bars) and relative (vertical bars) of selected protein groups with assigned functions. Additionally fold changes compared to pre-bloom metaproteome calculated based on NSAF values. Used abbreviations: B-Bacillariophyta, C-Coscinodiscophyceae, Ba-Bacillariophyceae, H-Haptophyta, Cr-Cryptophyta, M-Mediophyceae, D-Dinophyceae, Ce-Cercozoa, Ch-Chlorophyta, S-Streptophyta, O-Ochrophyta and F-Fragilariophyceae.

Furthermore, protein groups associated with protein synthesis were found with high overall expression especially in the pre-bloom phase.

Cytoskeletal protein groups, transport protein groups and protein groups participating in oxidative stress response were more abundant in the $\geq 10 \mu\text{m}$ fractions. During the pre-bloom stage protein groups involved in the biosynthesis of cofactors, prosthetic groups and carriers, i.e. chlorophyll, terpenoids and steroids, as well as protein groups involved in antibiotic resistance and –efflux as well as in the biosynthesis of antibiotics and plant resistance proteins, were generally more expressed. In the stage of the bloom rise cytoskeletal protein groups, protein groups involved in vesicle mediated transport (in the $\geq 10 \mu\text{m}$ fractions) and the biosynthesis of S-adenosylmethionine were generally more abundant. Finally, protein groups involved in the detoxification of reactive electrophilic compounds (in the $\geq 10 \mu\text{m}$ fractions) and nitrogen metabolism (in the $\geq 10 \mu\text{m}$ fractions) were higher expressed in the peak of the bacterial bloom. Protein groups involved in cell redox homeostasis (in the 3 – 10 μm fractions) and adhesion (in the $\geq 10 \mu\text{m}$ fractions) were dominant in both the bloom rise and peak samples (Figure 17, Table S9 CD-ROM). Identified protein groups were mainly assigned to the most abundant taxonomic phyla and classes, thus *Bacillariophyta* and *Haptophyta* as well as *Cryptophyta* and *Chlorophyta* (Figure 17).

6.4. Proteomic analysis of a selected PA bacterial isolate grown on laminarin and alginate

Laminarin is an algal carbon storage molecule and a major molecule in the ocean carbon cycle (Becker *et al.* 2020). FL planktonic bacteria remineralize laminarin using enzymes encoded in specific polysaccharide utilization loci (PULs). Despite the prominent role of laminarin none of the four well characterized or hypothesized PULs for the degradation of laminarin could be identified in genomes of PA bacteria (Kappelman *et al.* 2018). This is also true for the PA marine isolate *Muricauda sp.* MAR_2010_75, which was isolated from phytoplankton in the North Sea at the island Sylt (Hahnke and Harder 2013). To test if a potential laminarin and also alginate PUL can be predicted by a proteomic analysis, the degradation of the two

polysaccharides by the PA isolate *Muricauda sp.* MAR_2010_75 was investigated. For this purpose the growth of this strain in a medium without sugars, on the monosaccharide glucose as well as on the polysaccharides laminarin and alginate was compared. Subsequently proteomic analyses of the cultures grown on glucose as a control and the two polysaccharides were performed.

6.4.1. Growth comparison of Muricauda sp. MAR_2010_75

The first step was to test if *Muricauda sp.* MAR_2010_75 is able to utilize the provided sugars by growth experiments. Figure 18 shows the growth curves of the strain grown on glucose, alginate, laminarin and without a sugar addition. The cultures without the addition of sugars reached a maximal OD of 0.2 after 65.5 hours, whereas the cultures grown with glucose showed the highest maximal OD at 0.47 after nearly 90 hours. The cultures grown with polysaccharides reached their maximal ODs at 0.39 for laminarin after 65.5 hours and 0.35 for alginate after 70 hours. All of the cultures reached the stationary phase after approximately 90 hours but at different ODs (Table S 2 Appendix).

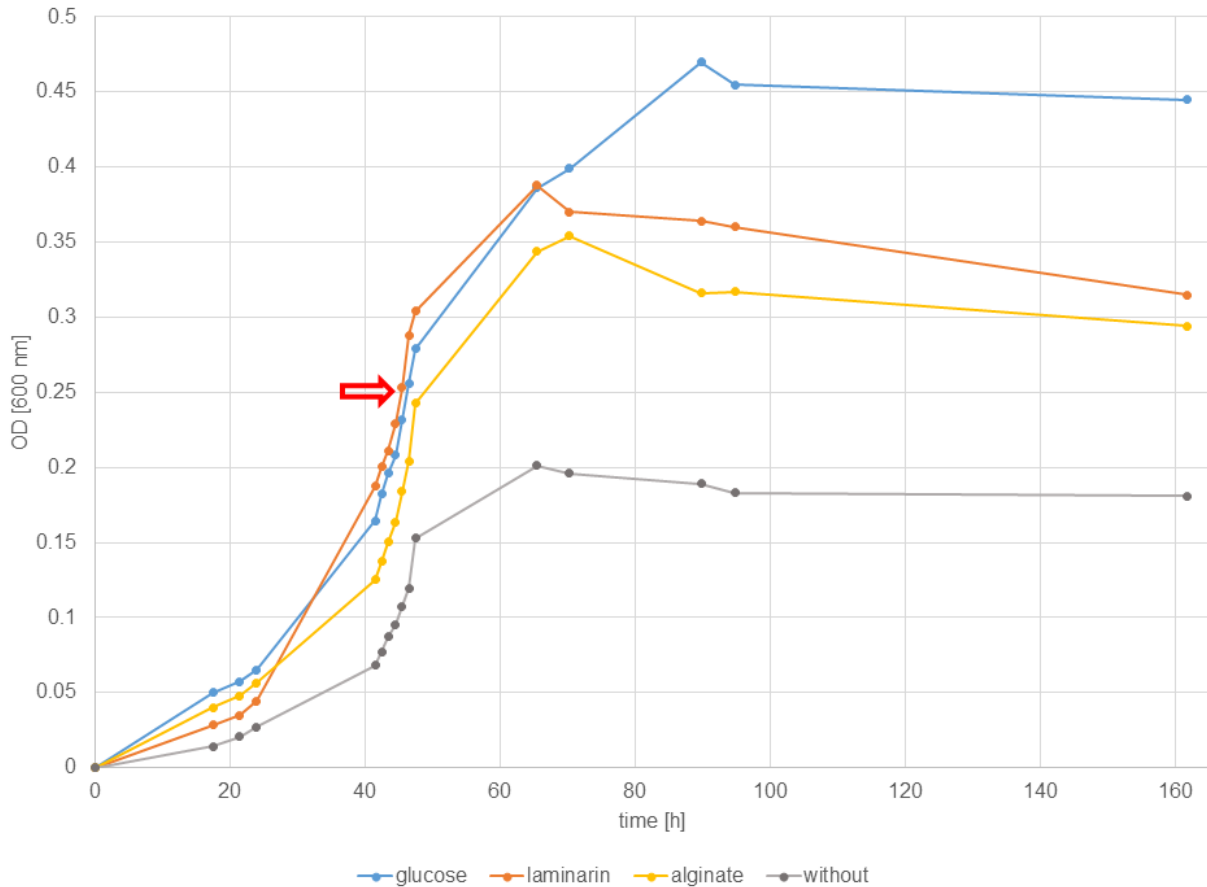


Figure 18. Growth comparison of *Muricauda sp.* MAR_2010_75 in Haha100V medium without sugar, with the monosaccharide glucose and with the polysaccharides laminarin and alginate (2 g/l). Cultures were inoculated with 0.4% (v/v) of an over night culture and grown till the harvest point or stationary phase was reached. The red arrow marks the harvest point of the cultures grown with sugars at an OD_{600nm} of 0.25.

6.4.2. Proteomic analysis of *Muricauda sp.* MAR_2010_75 grown on different polysaccharides and glucose

For the proteomic analysis proteins of three biological replicates (br) of cultures grown on glucose, laminarin and alginate were extracted via the above-described beat-beating based method and separated by 1D-SDS PAGE. For the MS sample preparation gel lanes were cut into 10 pieces and proteins were *in-gel* trypsin digested. LC-MS/MS measurement was performed with a Q Exactive Orbitrap MS and database search was performed with MaxQuant against a database containing species-specific protein sequences from NCBI. The quantification of the protein groups that were identified in at least 2 out of 3 br was based on the relative intensity-based absolute quantification (riBAQ). The riBAQ gives the relative protein abundance of all proteins in the same sample in percentage (Shin *et al.* 2013). For

cultures grown with glucose 2,173 protein groups could be identified which represent 60.4% of the *Muricauda sp. MAR_2010_75* genome. For cultures grown with laminarin 2,140 proteins could be identified (genome coverage of 59.5%) and for cultures grown on alginate 2,060 proteins could be identified (genome coverage of 48.7%) (Table S10 CD-ROM). 1677 identified protein groups were identified in cultures grown on all three substrates. Cultures grown on glucose showed the highest number of specifically identified protein groups (164), followed by laminarin (132) (Figure 19).

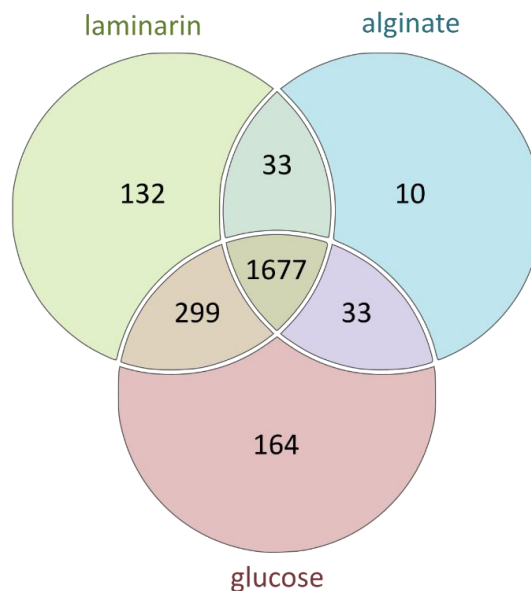


Figure 19. Venn diagram showing the number of substrate-specific and shared protein groups of *Muricauda sp. MAR_2010_75* grown on glucose, laminarin and alginate. The identified protein groups for cultures grown on glucose are depicted in green, on laminarin in blue and on alginate in yellow.

Proteome of *Muricauda sp. MAR_2010_75* grown on laminarin. PULs can be manually detected based on the presence of CAZyme clusters and co-occurring SusC/D cluster, which in most cases are also part of a PUL (Bjursell *et al.* 2006; Kappelmann *et al.* 2018). CAZymes were annotated based on the dbCAN2 database. In the genome of *Muricauda sp. MAR_2010_75* 88 GH family proteins were annotated via the dbCAN2 database of which 71 could be identified during growth on glucose and 73 during the growth on laminarin (Table 8). Known laminarin PULs contain GH families 2, 3, 5, 16, 17 and 30, whereof GH16 and GH17 cleave branched laminarin polysaccharides into oligosaccharides (Labourel *et al.* 2015; Barras and Stone 1969; Krüger *et al.* 2019). Both, GH16 and GH17, could not be identified in samples grown on laminarin. However, two GH2, two GH3 and one GH30_8 were identified. One GH2 protein (FG28_RS13650) was not induced, while the other (FG28_RS07890) showed a fold change (fc) of 5.7 compared to glucose. One of the GH3 protein (FG28_RS00590) showed a fc of 2.1 compared to glucose-grown cells, while the other (FG28_13800) showed no induction. The GH30_8 protein (FG28_RS07895) could only be identified in the samples grown on laminarin (

Table 9) and was part of a predicted glucuronoxylan PUL (Kappelmann *et al.* 2018) (Table S10 CD-ROM). However, considering the identified proteins belonging to the same operon or genomic neighborhood of the identified CAZyme from other laminarin PULs, no laminarin PUL structure could be observed (Figure S 3 Appendix). For the identification of SusC proteins

Results

TIGRFAM profile TIGR04056 (OMP_RagA_SusC) was used. With this approach 10 SusC could be identified in samples grown on laminarin (Table 8, Table S10 CD-ROM). Two showed a fc over 2 compared to glucose (Table 10). Namely FG28_RS02370 showed a fc of 13.7, is part of a predicted β -mannan PUL (Kappelmann *et al.* 2018) and the second most abundant SusC according to its riBAQ. The most abundant SusC (FG28_RS02640) showed no induction compared to glucose. The second induced SusC (FG28_RS13765) showed a fc of 11.5 (Table 10) and was part of a predicted PUL with unknown substrate (Kappelmann *et al.* 2018) (Table S10 CD-ROM). The latter two SusC proteins were also identified in the samples grown on alginate. However, considering the identified proteins belonging to the same operon or genomic neighborhood of the identified SusC protein, no known laminarin PUL structure could be observed (Figure S 3 Appendix).

Table 8. Number of annotated CAZymes via dbCAN2 database and number of SusC-proteins with the TIGRFAM profile TIGR04056. Shown are the numbers of CAZymes and SusC-proteins with the TIGRFAM profile TIGR04056 in the genome of *Muricauda sp.* 2010_75 and the identified ones in the proteomes of the species grown on glucose, laminarin and alginate. CAZymes were counted if they were annotated with only one or both of the used search algorithms (HAMMER; DIAMOND). Used abbreviations: AA (auxiliary activities), CBM (carbohydrate-binding module), CE (carbohydrate esterase), GH (glycosyl hydrolase), GT (glycosyl transferase), PL (polysaccharide lyase).

CAZyme/SusC TIGR04056	Number annotated in genome	Number annotated in proteome after growth on glucose	Number annotated in proteome after growth on laminarin	Number annotated in proteome after growth on alginate
AA	5	5	5	5
CBM	19	11	8	8
CE	19	14	16	12
GH	88	71	73	56
GT	43	30	28	21
PL	3	2	2	2
SusC	12	8	10	4

Table 9. Identified GH proteins in the proteomes of *Muricauda sp.* MAR_2010_75 grown on laminarin and alginate with their fold changes compared to growth on glucose and locus_tags. Only GH (glycosyl hydrolase) proteins with a fold change over 1.5 and GH proteins only identified in samples grown on laminarin are shown.

GH family	Locus_tag	fold change laminarin/glucose
GH2	FG28_RS07890	5.7
	FG28_RS1650	1.7
GH3	FG28_RS00590	2.1
GH10	FG28_RS07875	1.9
GH13_19	FG28_RS18610	1.6
GH13_31	FG28_RS07530	only identified during growth on laminarin

Results

GH18	FG28_RS03875	only identified during growth on laminarin
GH20	FG28_RS01070 FG28_RS00580	1.7 only identified during growth on laminarin
GH29	FG28_RS03965	3.8
GH30_8	FG28_RS07895	only identified during growth on laminarin
GH31	FG28_RS14045	6.4
GH37	FG28_RS16860	2.1
GH43_1	FG28_RS07930	1.8
GH43_28	FG28_RS13795	3.5
GH43_29	FG28_RS02610	2.0
GH51	FG28_RS02585	2.7
GH67	FG28_RS07860	3.2
GH74	FG28_RS07535	3.4
GH97	FG28_RS00805	2.4
GH101	FG28_RS00605	3.5
GH105	FG28_RS17990	1.8
GH109	FG28_RS19285 FG28_RS01285 FG28_RS16245 FG28_RS17865 FG28_RS05995 FG28_RS12240	1.5 2.4 1.7 1.8 5.0 only identified during growth on laminarin
GH144	FG28_RS13780	only identified during growth on laminarin
GH146	FG28_RS02605 FG28_RS02595	1.9 2.1

A coupling of peptide metabolism and laminarin utilization was suggested (Unfried *et al.* 2018). In the *Formosa sp.* Hel1_33_131 proteome 41 peptidases were expressed in the presence of laminarin of which 9 showed a significantly higher protein abundance compared to glucose or were exclusively expressed. This coupling was also observed in the proteome of *Muricauda sp.* MAR_2010_75, in which 82 peptidases were expressed in the presence of laminarin of which 15 were induced with a fc over 2 compared to glucose and one was exclusively expressed.

Proteome of Muricauda sp. MAR_2010_75 grown on alginate. Known alginate lyases belong to the PL families 6, 7, 12 and 17. However none of these could be identified in samples grown on alginate. For the identification of SusC proteins TIGRFAM profile TIGR04056 (OMP_RagA_SusC) was used. With this approach four SusC could be identified in samples grown on alginate (Table 8, Table S10 CD-ROM). All of this SusC proteins were also identified

Results

in samples grown on laminarin (Table 10). One showed a fc over 2 compared to glucose. Namely FG28_RS13765 showed a fc of 11.9 and is part of a predicted PUL with unknown substrate (Kappelmann *et al.* 2018) and was also induced in samples grown on laminarin. Two of the identified SusC (FG28_RS02370, FG28_RS02640) showed no induction compared to glucose. The last of the identified SusC (FG28_RS09405) showed a fc of 0.2 compared to glucose and thus was decreased during growth on alginate (Table 10, Table S10 CD-ROM). However, considering the identified proteins belonging to the same operon or genomic neighborhood of the identified SusC protein, no alginate PUL structure could be observed (Figure S 3).

Table 10. Identified SusC proteins with the TIGRFAM profile TIGR04056 in the proteomes of *Muricauda sp.* MAR_2010_75 grown on laminarin and alginate with fold changes compared to growth on glucose and locus_tags.

locus_tag	fold change laminarin/glucose	fold change alginate/glucose
FG28_RS02370	13.7	1.1
FG28_RS02640	1.53	1.14
FG28_RS07675	0.74	not identified
FG28_RS07925	1.47	not identified
FG28_RS09405	1.30	0.23
FG28_RS13665	only identified during growth on laminarin	not identified
FG28_RS13765	11.52	11.87
FG28_RS14040	only identified during growth on laminarin	not identified
FG28_RS14995	2.6	not identified
FG28_RS19660	0.98	not identified

Global responses to growth on polysaccharides. Besides DNA-binding proteins, elongation factors, chaperones, ATP synthases and RNA-polymerases also SusD-like proteins, Omp_RagA_SusC superfamily proteins and MotA_ExbB family proteins protein belonged to the 15 most abundant proteins under all growth conditions (Table 11). A PorT family protein as well as a dihydropicolinate synthase family protein showed lower abundance in the proteome of cultures grown on glucose compared to cultures grown on polysaccharides.

Table 11. Top 15 abundant proteins according to riBAQ in *Muricauda sp.* MAR_2010_75 proteome during growth on glucose, laminarin and alginate.

Glucose		Laminarin		Alginate	
Protein (locus_tag)	riBAQ	Protein (locus_tag)	riBAQ	Protein (locus_tag)	riBAQ

Results

Elongation factor TU (FG28_RS11990)	0.038	HU family DNA-binding protein (FG28_RS15560)	0.052	Elongation factor TU (FG28_RS11990)	0.05
HU family DNA-binding protein (FG28_RS15560)	0.032	SusD-like superfamily protein (FG28_RS03065)	0.042	Hypothetical protein (FG28_RS03070)	0.031
SusD-like superfamily protein (FG28_RS03065)	0.031	OmpA-like protein (FG28_RS05710)	0.034	Omp_RagA_SusC superfamily protein (FG28_RS03060)	0.028
Hypothetical protein (FG28_RS03070)	0.028	Elongation factor TU (FG28_RS11990)	0.032	OmpA-like protein (FG28_RS05710)	0.028
OmpA-like protein (FG28_RS05710)	0.026	Hypothetical protein (FG28_RS03070)	0.030	SusD-like superfamily protein (FG28_RS03065)	0.027
Omp_RagA_SusC superfamily protein (FG28_RS03060)	0.023	Omp_RagA_SusC superfamily protein (FG28_RS03060)	0.028	HU family DNA-binding protein (FG28_RS15560)	0.023
Chaperonin GroEL (FG28_RS03870)	0.014	SusD superfamily protein (FG28_RS04380)	0.023	Dihydropicolinate synthase family protein (FG28_RS02625)	0.015
Glyceraldehyde-3-phosphate dehydrogenase (FG28_RS15245)	0.012	Omp_RagA_SusC superfamily protein (FG28_RS04375)	0.017	MotA_ExbB family protein (FG28_RS09795)	0.014
ATP synthase subunit beta (FG28_RS04930)	0.011	Dihydropicolinate synthase family protein (FG28_RS02625)	0.013	PorT family protein (FG28_RS12515)	0.013
ATP synthase subunit alpha (FG28_RS06595)	0.009	Chaperonin GroEL (FG28_RS04930)	0.012	ATP synthase subunit beta (FG28_RS04930)	0.012
OmpA family protein (FG28_RS17520)	0.009	MotA_ExbB family protein (FG28_RS09795)	0.012	OmpA family protein (FG28_RS17520)	0.010
MotA_ExbB family protein (FG28_RS09795)	0.009	PorT family protein (FG28_RS12515)	0.010	RNA-polymerase subunit alpha (FG28_RS02885)	0.010
RNA-polymerase subunit alpha (FG28_RS02885)	0.008	ATP synthase subunit beta (FG28_RS04930)	0.009	Ribosomal protein (FG28_RS08545)	0.009
Catalase/peroxidase (FG28_RS05165)	0.008	Glyceraldehyde-3-phosphate dehydrogenase (FG28_RS15245)	0.009	Chaperonin GroEL (FG28_RS03870)	0.009
Elongation factor TU ((FG28_RS09880)	0.008	ATP synthase subunit alpha (FG28_RS06595)	0.008	DUF1573-containing protein (FG28_RS01295)	0.009

Muricauda sp. MAR_2010_75 seems to activate type 9 secretion systems (T9SS) during the growth on polysaccharides. 24 T9SS-related proteins could be identified in cultures grown on laminarin, whereof six showed a fc over 2 compared to glucose and two a fc under 0.5. Furthermore, 22 T9SS-related proteins could be identified in cultures grown on alginate,

Results

whereof 2 showed a fc over 2 compared to glucose und three a fc under 0.5 (Table S 3 Appendix). Among protein groups with the highest fc and thus induction in cultures grown in laminarin were Usp and Usp-like proteins as well as NosZ and aminotransferases and some of the already above described SusC proteins. In cultures grown on alginate several hypothetical proteins as well as an aminotransferase, OmpA_C-like protein and SusC and SusD-like proteins showed the highest fold changes (Table 12).

Table 12. Proteins with highest induction according to fc compared to glucose as control in *Muricauda sp.* MAR_2010_75 proteomes during growth on laminarin and alginate.

Laminarin		Alginate	
Protein (locus_tag)	Fold change to glucose	Protein (locus_tag)	Fold change to glucose
Usp universal stress protein (FG28_RS14350)	66.0	Aminotransferase class I/II (FG28_RS17780)	40.7
Usp-like protein (FG28_RS14315)	34.9	M13 family metalloproteinase (FG28_RS02645)	18.3
NosZ (FG28_RS13870)	34.2	OmpA_C-like protein (FG28_RS08785)	15.6
Hypothetical protein (FG28_RS05460)	31.6	DUF285 containing-protein (FG28_RS20320)	15.4
FM1-binding family protein (FG28_RS14250)	23.1	SusC TIGR04056 (FG28_RS13765)	11.9
Usp-like protein (FG28_RS14325)	22.0	4-hydroxyphenylpyruvate dioxygenase (FG28_RS09935)	10.9
Usp-like protein (FG28_14340)	21.9	Hypothetical protein (FG28_RS09160)	8.8
Aminotransferase class I/II (FG28_RS17780)	20.6	Hypothetical protein (FG28_RS02725)	8.8
Aminotransferase class I/II (FG28_RS00305)	20.4	Urocanate hydratase (FG28_RS09195)	8.3
Urocanate hydratase (FG28_RS09195)	18.7	Hypothetical protein (FG28_RS10485)	8.2
2-hydroxyacid dehydrogenase (FG28_RS14375)	17.8	Cupin-like superfamily protein (FG28_RS09930)	8.0
Hypothetical protein (FG28_RS10650)	14.8	Hypothetical protein (FG28_RS15830)	6.7
Omp_RagA_SusC superfamily protein (FG28_RS09515)	13.8	SusD-like-2 superfamily protein (FG28_RS03350)	6.6
SusC TIGR04056 (FG28_RS02370)	13.7	Fumarylacetoacetase (FG28_RS09925)	6.4
SusC TIGR04056 (FG28_RS13765)	11.5	Omp_RagA_SusC superfamily protein (FG28_RS09925)	6.4

7. Discussion

7.1. Establishment of a metaproteomic pipeline for PA microbial communities

Efficient protein extraction is an essential step for successful metaproteomics analysis of microbial communities. Therefore, five different protein extraction methods using different strategies and the commercially available TRI-Reagent® were tested on filter samples originating from several sampling events during the 2009 spring bloom sampling campaign. SDS-acetone and bead-beating based protocols turned out to be most efficient for protein extraction from particles, as they resulted in the highest protein yields (Figure 2B, Table 6). A possible explanation for the efficiency of the bead-beating protocol is the effective disintegration of the particle matrix by EDTA added to the extraction buffer (Passow 2002). The SDS-acetone and bead-beating protocols were used for optimizing the subsequent MS sample preparation procedure.

Therefore, total protein amount was extracted by the chosen methods from filters collected on the 28th of April 2009 and separated by 1D SDS-PAGE (Figure 2B). Even though MS sample preparation via GeLC-MS/MS is more time-consuming compared to 1D or 2D-LC approaches, it has proven valuable to purify protein extracts and remove polymeric contaminants (e.g. Lassek *et al.* 2015; Keiblinger and Riedel 2018) and yields comparable results as LC-based peptide fractionation (Hinzke *et al.* 2019). To investigate if an increase in the total number of individual gel subfractions leads to a higher number of identified protein group, gel lanes were cut in either 10 or 20 equal-sized fractions. Additionally, it was tested if reduction and alkylation of the proteins prior to tryptic digestion increased protein group identification rates (Figure 2B). Searching the resulting spectra in the so far available 0.2 µm 2009 (MIMAS) database (Teeling *et al.* 2012) showed that the best results (Figure 2B, Figure 3) were obtained by higher fractionation (20 gel pieces) without reduction and alkylation. Subsequent to the MS-sample preparation the obtained LC-MS/MS spectra were searched against different MS target databases. The creation of a MS target database is a balancing act between considering all

sequences of proteins possibly present in the analyzed sample on the one hand and reducing the database size to save computational costs and ease protein interference on the other hand. Searching metaproteomics MS/MS data against large databases, e.g., the NCBI non-redundant database with 136,216,794 entries (10/03/2017), presents some major challenges, such as an increasing number of false positive or false negative peptide to spectrum matches (PSMs) together with (sub)sequence redundancies caused by conserved regions of different proteins. Simultaneously to the optimization of the metaproteomics protocol metagenomic sequencing, assembly and annotation of FL (0.2 µm pore-sized filters) and PA (3 and 10 µm pore-sized filters) fractions of water samples collected during the Helgoland spring bloom 2009 were performed. Unfortunately, most probably due to the high amount of eukaryotic DNA, length and number of assembled sequences of the large particulate fraction (10 µm pore-sized filters) were not sufficient for a valid data interpretation. Thus, the metagenomic database used for subsequent database searches of samples from the spring bloom 2009 was only composed of sequences of FL bacteria (0.2 µm pore-sized filters) and microbial communities present in the medium particulate fraction (3 µm pore-sized filters).

As the success of metaproteome analyses depends upon the database used for protein group identification (e.g. Schneider and Riedel 2010; Teeling *et al.* 2012), the LC-MS/MS spectra obtained with the bead-beating protocol were searched against four different databases to identify the database that gives the highest number of reliably identified protein groups (Figure 4). Best results were obtained with the 0.2 + 3 µm 2009 database (Figure 4), which is not surprising as it is well accepted that metaproteomic data are most informative in combination with complementary omics approaches (e.g. Banfield *et al.* 2005; Ram *et al.* 2005). Therefore, e.g. Timmins-Schiffman *et al.* (2017) recommended to sequence corresponding site and time specific metagenomes in order to generate accurate metaproteomic databases as best practice for environmental proteomics. One possible reason for this is, that protein sequences of many of the marine bacterioplankton and phytoplankton species present in these particular samples, are not present in public available databases and databases that do not rely on consistent metagenomic or metatranscriptomics analyses, as these species likely have not

previously been sequenced. However, the success of a metaproteomics study also relies on the data analysis during which software like Prophane can help to overcome for example challenges during peptide to protein mapping. If fragment spectra of unique peptides (peptides that can be assigned to only one protein) are missing, a distinction between eligible proteins and hence a definite protein identification is impossible. This problem is also known as the protein inference problem (Nesvizhskii and Aebersold 2005). A common approach to deal with this inference problem is the grouping of identified proteins into protein groups based on shared peptide matches (Koskinen *et al.* 2011) Prophane (Schiebenhoefer *et al.* 2020; Schneider *et al.* 2011) provides a fully automatic pipeline to assist in metaproteomics data analysis and screens groups of peptide-sharing proteins for commonalities on functional or taxonomic level. Prophane uses the lowest common ancestor (LCA) approach, in which proteins are assigned to their lowest common ancestor in a hierarchical tree (Schiebenhoefer *et al.* 2020).

In the final optimized workflow (Figure 5), proteins are extracted using a combination of SDS-containing lysis buffer and cell disruption by bead-beating, separated by SDS-PAGE, *in-gel* digested and analysed by LC-MS/MS in 20 fractions, before MASCOT search against a metagenome-based database and data processing/visualization with the bioinformatics tool *Prophane*.

7.2. Comparative metaproteome analysis of FL and PA bacterioplankton fractions harvested during 2009 spring phytoplankton bloom

To investigate if the optimized protocol is suitable for a comparative metaproteomics analysis of FL and PA microbial communities, it was applied to various fractions of a microbial community sampled at the 14th of April in 2009 (0.2 – 3 µm = FL; 3 – 10 µm and ≥ 10 µm = PA). Using the optimized procedure, it was possible to record 360,000 to 460,000 spectra per technical replicate and about 20,000 spectra per gel fraction, which subsequently led to the identification of 9,354 protein groups, 2,263 protein groups, and 2,771 protein groups for the 0.2 – 3 µm (Table S1 CD-ROM), 3 – 10 µm (Table S2 CD-ROM), and ≥ 10 µm (Table S3 CD-

ROM) fractions, respectively. Comparable studies addressing metaproteomics analyses of marine sediments of the Bering Sea (Moore *et al.* 2012), the coastal North Sea, and the Pacific Ocean (Wöhlbrand *et al.* 2017a) identified less than 10% of the protein group identification numbers resulting from the here presented novel metaproteomics pipeline.

1,956 of the identified protein groups of the two PA fractions were also identified in the FL fraction whereas 276 protein groups were exclusively found in the PA fractions (Figure 6). This suggests that protein expression profiles of planktonic and particulate bacteria vary less than expected. However, this might also be due to the fact that PA bacteria are known to be tychoplanktic and, e.g. as offspring cells searching for a place to settle, may thus only temporarily be part of the planktonic community (Ghiglione *et al.* 2007; Grossart 2010; Crespo *et al.* 2013). Additionally, chemosensory and motile taxa, predominantly alphaproteobacterial *Rhodobacteraceae* and *Gammaproteobacteria* (Seymour *et al.* 2017) often stay on or close to a particle during nutrient-rich conditions, before swimming to the next particle (Ghiglione *et al.* 2007; Grossart 2010; Yawata *et al.* 2014), thus simultaneous presence as FL and PA could be a consequence of highly dynamic responses to substrate availability (Seymour *et al.* 2017). It is further important to notice that particles have a limited life span (Iversen and Ploug 2010), hence PA bacteria must spend some time FL on their way to the next particle (Bižić-Ionescu *et al.* 2014). Moreover, clogging of filter pores by particles may cause retention of FL bacteria, thus contaminating the PA fractions by planktonic bacteria. Additionally, also mechanical stress is discussed to disintegrate fragile bacteria during filtration (Ferguson *et al.* 1984) leading to accumulation of their protein groups in another fraction. Nevertheless, still many taxa are well separated in a 0.2 – 3 μm FL and over 3 μm PA fraction (Heins *et al.* 2021). In addition, the lack of $\geq 10 \mu\text{m}$ pore-sized filter metagenomic sequences hampers comprehensive protein group identifications in this PA fraction that may result in a virtually lower abundance than expected. Furthermore, this lacking information can not be found in public databases as genomes of microbial eukaryotes therein are also limited. However, also other studies comparing the community composition of FL and PA bacterioplankton in coastal

areas using for example high-throughput 16S rRNA gene sequencing methods, revealed minor differences between the fractions (Campbell and Kirchman 2013; Ortega-Retuerta *et al.* 2013).

Phylogenetic assignment of FL and PA metaproteomes. As expected, PA fractions contained considerably more eukaryotic protein groups than the FL fraction (Figure 7A). Many of these protein groups were assigned to known phytoplankton taxa (e.g., diatoms, *Pelgaophytes*, *Raphidiophytes*, *Cryptophytes*, *Dinoflagellates* and *Haptophytes*) (Figure 8), reflecting the ongoing phytoplankton bloom. Moreover, abundant protein groups were also assigned to *Oomycetes* (water molds, e.g. *Peronosporales*, *Saprolegniales*) and Fungi (e.g. *Cryptomycota*), indicating that saprotrophic or parasitic eukaryotes (Nigrelli and Thines 2013; Jones *et al.* 2011), may contribute to phytoplankton biomass degradation during algal blooms. The most abundant bacterial phyla within both, the FL and PA fractions, were *Proteobacteria* and *Bacteroidetes*. In the study of Teeling *et al.* 2012 protein groups derived from planktonic samples were mostly assigned to *Bacteroidetes*, *Alpha-* and *Gammaproteobacteria*. The most prominent clades within the *Bacteroidetes* were *Flavobacteriia* and therein the genera *Ulvibacter*, *Formosa* and *Polaribacter*. *Gammaproteobacteria* were dominated by the clades *Reinekea* and *Roseobacter*. These findings are in good accordance with the results obtained for the planktonic fraction in the here presented study. Protein groups expressed by *Alpha-*, *Beta-* and *Gammaproteobacteria* were generally more dominant in the FL bacteria, whilst protein groups assigned to *Cyanobacteria* (e.g., *Synechococcus*, *Arthrospira*), *Opituae*, *Flavobacteriia* (e.g. *Arenitalea*, *Olleya*, *Algibacter*, *Lacinutrix*), and some proteobacterial genera (e.g. *Oceanicoccus*, *Canidatus Puniceispirillum*, *Neptuniibacter*, *Halioglobus*, *Ramlibacter*) were more abundant in the PA fraction (Figure 7). This is in good accordance to other studies which reported *Bacteroidetes* in both, FL and PA, bacterioplankton (DeLong *et al.* 1993; Eilers *et al.* 2001; Abell and Bowman 2005; Alonso *et al.* 2007). Moreover, *Flavobacteriia* have been found highly abundant during phytoplankton blooms indicating that they play an important role as consumers of algal-derived organic matter (Simon *et al.* 1999; Riemann *et al.* 2000; Pinhassi *et al.* 2004; Grossart *et al.* 2005; Teeling *et al.* 2016; Chafee *et al.* 2018). In previous studies *Flavobacteriia* and *Alphaproteobacteria* were mainly identified as particle-preferring,

because they are often enriched in the PA fraction in contrast to e.g. *Gammaproteobacteria* (Crump *et al.* 1999; Bižić-Ionescu *et al.* 2014). However, Heins *et al.* 2021 found specialization within these classes even within families. In accordance with this thesis within the *Flavobacteriaceae* the genera *Algibacter* and *Arenitalea* were found enriched in the PA fraction and within the *Alphaproteobacteria* the genus *Planktomarina* was found enriched in the FL fraction. These findings highlight the individual lifestyles and niche occupations already shown within clades e.g. within the SAR11 clade (*Canidatus Pelagibacter ubique*) (Dadon-Pilosof *et al.* 2017; Mestre *et al.* 2020) and within genera e.g. *Vibrio* (Thompson *et al.* 2005; Hunt *et al.* 2008).

Functional assignment of FL and PA metaproteomes. Differences in the protein group profiles between FL and PA bacteria seemed more precise on the functional level (Figure 9). E.g., the SusC/D utilization system, specific GHs, GTs and TBDT were found with higher overall expression levels in the PA fractions compared to the FL fraction. This is in good accordance with the high substrate availability (Caron *et al.* 1982; Grossart *et al.* 2003; Fernández-Gómez *et al.* 2013), especially the presence of highly abundant microalgae storage polysaccharides, i.e. alpha- and beta-glucans (Kroth *et al.* 2008), in the particles. Sulfatases, capable of cleaving sulphate sugar ester bonds, are contributing to the degradation of specific sulphated algal polysaccharides such as mannans and fucans (Gómez-Pereira *et al.* 2012). This is well supported by our finding that sulfatases are strongly expressed by PA bacteria (Figure 9). However, sulfatases were strongly expressed by *Formosa sp.* bacteria which are considered as FL. One explanation for this could be that these species were frequently found attached to aggregates as single cells (Benneke *et al.* 2013). Furthermore, FL and PA bacteria seem to employ different phosphate acquisition strategies: whilst in FL bacteria phosphate and phosphonate ABC-type transporter appeared highly expressed, PA bacteria rather seem to employ phytases and phosphate-selective porins. As expected, various protein groups involved in stress response were differentially expressed. Interestingly, functions involved in oxidative stress defence appeared to be less abundant in the PA fractions (maybe due to shading, reducing solar irradiation stress in the particles), whilst protein groups for heavy metal

and antibiotic resistance were strongly expressed in the $\geq 10 \mu\text{m}$ fraction (Figure 9A and B), which also contained the highest proportion of eukaryotic protein groups (Figure 7A). This might be due to the fact that some algae take up and store heavy metals (Gaudry *et al.* 2007) and are capable of producing antibiotics (Grossart 1999). This indicates that close eukaryotic bacterial interactions in particles require such defense strategies of the associated bacteria. As it was further expected, adhesion protein group as well as protein groups involved in motility were more abundant on the particles, emphasizing their importance for biofilm/aggregate formation (O'Toole and Kolter 1998; Lemon *et al.* 2007; Houry *et al.* 2010; Burke *et al.* 2011a). However, protein groups involved in chemotaxis, gliding motility and other forms of motility could be identified as more abundant in the FL fraction. This could be explained by the above-described fraction of motile bacteria which sense nutrient gradients and swim towards their food sources. They often represent $\sim 10\%$ of cells in the coastal seas (Stocker and Seymour 2012) with large variations of up to 80% (Mitchell *et al.* 1995; Grossart *et al.* 2001). Interestingly, proteorhodopsin, an inner membrane protein involved in light dependent energy generation, which has been proposed to enable FL bacteria such as *Polaribacter* (Fernández-Gómez *et al.* 2013) and *Pelagibacter* (Giovannoni *et al.* 2005) to survive under low nutrient conditions, was abundantly identified in PA bacteria such as *Polaribacter*, *Paraglaciecola* and *Marinosulfomonas* in our analyses.

The obtained metaproteomics data demonstrate a high abundance of eukaryotes in the particles (Figure 8), which include numerous microalgal groups, but also fungi and various protozoa. This clearly sets particles apart from the FL fraction and highlights the importance of direct eukaryote bacterial interactions in particles. Previous work on FL bacteria showed that bacterial succession was largely independent of phytoplankton composition, and instead determined by broad substrate availability (Teeling *et al.* 2016). PA bacterial composition is more likely to be directly controlled by algal composition due to the intimate nature of their interactions (Grossart *et al.* 2005), although functional redundancy may be substantial (Burke *et al.* 2011a). Moreover, eukaryotes may also contribute to polysaccharide degradation in concert with bacteria. For example, fungal taxa can be abundant in marine particles and have

been shown to utilize algal polysaccharides such as laminarin (Bochdansky *et al.* 2017; Cunliffe *et al.* 2017).

7.3. Investigation of phylogenetic and functional successions of marine particles during the 2018 spring phytoplankton bloom

To investigate succession of taxonomical clades and expressed functions of marine particles during a spring phytoplankton bloom, the PA fractions (3 – 10 μm and $\geq 10 \mu\text{m}$) of three sampling time points during different stages of a bloom in 2018 at Helgoland roads in the German Bight were analyzed using the above optimized metaproteomics pipeline. The North-Sea is a semi-enclosed continental shelf sea and especially its southern region, the German Bight, is highly influenced by the run-off from the rivers Elbe and Weser and thereby constantly supplied with nutrients, making it a very productive area. The mixing of fresh- and marine water typically leads to high spatial variability with respect to environmental parameters, such as temperature, salinity, pH and organic loads. Around the island Helgoland in the German Bight, the biota and current patterns are continuously monitored for more than 5 decades known as Helgoland Roads time series (Wiltshire *et al.* 2008). This comprehensive long-term data set makes Helgoland Roads an optimal site to investigate the phylogenetic and functional successions of marine particles during a spring phytoplankton bloom in coastal oceanic environments. The spring phytoplankton bloom 2018 was accompanied by an increase of total bacterial cell counts from 0.6×10^6 cells mL^{-1} to up to 2.47×10^6 cells mL^{-1} . Sampling time points were chosen based on the bacterioplankton response bloom and labeled pre-bloom (also pre-phytoplankton bloom), bloom rise (after peak of diatom bloom) and bloom peak (post diatom bloom) (Figure 10).

Employing the optimized metaproteomic pipeline, it was possible to identify 5,607 protein groups for pre-bloom samples, 7,141 protein groups for bloom rise samples and 4,118 protein groups for the bloom peak samples (Table S5 CD-ROM). With this, the number of identified protein groups could be increased compared to the above-described application example and is, at least to my knowledge, the largest number of protein groups ever identified for marine

particles. One explanation for the increase in identified protein groups could be the application of a fitting metagenome database composed of 9 different time points for both analyzed fractions (3 + 10 μ m 2018 - 14,764,755 entries). Nevertheless, the number of identified protein groups for the bloom peak was expected to be higher, as the number of eukaryotic protein groups is lower. It is well known, that phyto- and zooplankton associated to the marine particles cause problems for the assembly and subsequent bioinformatics steps like binning and annotation during metagenomic analyses due to e.g., introns and tandem repeat regions, subsequently also hamper the identification of protein groups. Furthermore, these difficulties during metagenomic and also genomic analysis of microbial eukaryotes led to a lack of sequences in public databases. An explanation for the lower number of identified protein groups for the bloom peak samples can be the high complexity of the samples with very limited concentrations of individual proteins. In the so called "Data Dependent Acquisition" (DDA) mode during MS-analysis, peptides with the highest abundance are selected for fragmentation and the fragmentation pattern is used for the identification of peptides and proteins, respectively. Consequently, only peptides selected for fragmentation can be identified. Considering the complexity of a peptide mixture, the peptide selection during the MS measurements is a rather arbitrary process, causing missing values in time line experiments and also between technical replicates. Due to the used filter criteria (a peptide has to be identified in at least 2 out of 3 technical replicates) the number of identified protein groups was lower for more complex samples (Figure S 1).

7.3.1. Phylogenetic succession of marine particles during the 2018 spring phytoplankton bloom

As expected, the number of identified protein groups assigned to bacteria increased from pre-bloom to bloom peak samples contrary to protein groups assigned to eukaryotes (Figure 11), reflecting the ongoing of the bacterial response bloom. Overall, most of the identified protein groups in all samples were assigned to eukaryotes as already observed for the PA fractions of the application example (Figure 8). PCA plot analyses of all identified bacterial families showed

that both PA fractions of the pre-bloom sampling point and the $\geq 10 \mu\text{m}$ fraction of the sampling time point during the rise of the bacterial bloom cluster together (Figure 12A), which indicates that beta-diversity of bacterial families is increasing with the progress of the bacterial bloom, already starting during the rise of the bloom in the 3 – 10 μm fraction. OPLS-DA plots of all identified bacterial families congruently indicate, that the samples differ more based on sampling time point than on filter size. Furthermore, the sampling time points of the 3 – 10 μm fraction show a higher taxonomic beta-diversity on family level than the sampling time points of the $\geq 10 \mu\text{m}$ fraction as the bloom rise and bloom peak sampling time points cluster together (Figure 12B and C). PCA plot analysis of all identified eukaryotic classes indicate that their beta-diversity is the highest during the rise of the bacterial bloom which is in accordance with the ongoing of the diatom bloom, which already started collapsing as the bacterial bloom rose (Figure 14A). OPLS-DA plots of all identified eukaryotic classes show a similar trend as for the bacterial families: the samples seem to differ more based on sampling time point than on filter size. However, in contrast to the results for the bacterial families, eukaryotic classes seem to establish less beta-diversity as the bloom rise and bloom peak samples of the $\geq 10 \mu\text{m}$ fraction as well as the pre-bloom and bloom rise samples of the 3 – 10 μm fraction cluster together in the OPLS-DA plot based in pore size of the filters (Figure 14B and C). Nevertheless, this was somewhat expected, because the sampling time points were chosen on the basis of the bacterial response bloom.

Phylogenetic succession of bacterial taxa

The most abundant bacterial phyla within the analyzed PA fractions were *Proteobacteria* and *Bacteroidetes* at all analyzed stages of the bloom (Figure 13; Table S6 CD-ROM). In the spring bloom 2009 the comparison of FL and PA fractions (6.2; Schultz *et al.* 2020) and the study from Teeling *et al.* 2012 marine *Flavobacteriia*, *Alpha-* and *Gammaproteobacteria* were identified as key players in the FL and PA fractions, which confirmed the obtained results. Within these classes consecutively blooming *Formosa* and *Polaribacter* clades were identified

in the FL fraction (Teeling *et al.* 2012), which were also identified in the PA fraction of the 2009 spring phytoplankton bloom (Schultz *et al.* 2020). The *Polaribacter* clade was again under the most abundant clades in the spring bloom 2018 (Figure 13). The gammaproteobacterial genus *Reinekea* was also identified as abundant and consecutively blooming in the spring bloom 2009 in the FL fraction (Teeling *et al.* 2012; Schultz *et al.* 2020) and its low abundance in the PA fractions could be confirmed in the results for the 2018 spring bloom, leading to the assumption, that *Reinekea* is a mainly FL genus.

Phylogenetic differences between the two analyzed PA fractions were visible. On phyla level protein groups expressed by *Cyanobacteria*, *Planctomycetes* and *Verrucomicrobia* were generally more abundant in the 3 – 10 μm fraction, whereas protein groups expressed by *Actinobacteria* and *Proteobacteria*, especially by the classes *Gamma-* and *Deltaproteobacteria*, were higher abundant in the $\geq 10 \mu\text{m}$ fraction. *Cyanobacteria* were also identified as more abundant in the PA fractions during the spring bloom 2009 (Schultz *et al.* 2020). Furthermore, *Planctomycetes* were among the most abundant groups unique to the PA fraction and significantly overrepresented compared to FL in the Mediterranean water column (López-Pérez *et al.* 2016; Mestre *et al.* 2017) as well as frequently observed to associate with marine phytoplankton blooms (Fukunaga *et al.* 2009; Morris *et al.* 2006; Bengtsson and Øvreås 2010; Burke *et al.* 2011b; Lachnit *et al.* 2011; Lage and Bondoso 2011). *Planctomycetes* can exhibit diverse lifestyle strategies including motility, unicellular stages as well as aggregate formation (Lage 2013). *Verrucomicrobia* were observed in marine snow (Rath *et al.* 1998), in association with nanoeukaryotic cells (Petroni *et al.* 2000), enriched in PA fractions (Mestre *et al.* 2017) and suggested to be efficient biopolymer degraders (Martinez-Garcia *et al.* 2012). These findings highlight the role of the phyla *Cyanobacteria*, *Planctomycetes* and *Verrucomicrobia* as PA bacteria. On genus level protein groups assigned to the alphaproteobacterial genera *Candidatus Pelagibacter*, *Planktomarina* and *Tateyamaria* as well as to the flavobacterial genera *Ulvibacter* and *Formosa*, were generally more dominant in the 3 – 10 μm fraction, whilst protein groups assigned to the alphaproteobacterial genus *Roseobacter* and to the flavobacterial genus *Tenacibaculum* were more abundant in the

$\geq 10 \mu\text{m}$ fraction (Figure 13). Abell and Bowman 2005 found that *Tenacibaculum* spp. were unusually responsive to diatom detritus in a seawater mesocosm experiment. They suggested that the detectable populations of *Tenacibaculum* can be highly transient as they grow only to significant populations as attached population or in association with detrital OM. Additionally, the genus *Tenacibaculum* was detected in high proportions in association with a natural phytoplankton bloom (González *et al.* 2000) and was suggested to be an active colonizer of particles under algal bloom conditions (Riemann *et al.* 2000). This could be confirmed by the results of this study where the genus *Tenacibaculum* was identified as more abundant during the peak of the bacterial response bloom (Figure 13). These findings highlight the role of the genus *Tenacibaculum* as PA bacteria. The genus *Roseobacter* often dominates microbial assemblages surrounding POM (Fontanez *et al.* 2015; Mestre *et al.* 2017; Rinta-Kanto *et al.* 2012) but was also identified as abundant in FL fraction (Teeling *et al.* 2012) suggesting a special adaptation to acquire phytoplankton-derived DOM (Teeling *et al.* 2012; Landa *et al.* 2017; Landa *et al.* 2019) independent of life-style strategy. Alonso-Sáez *et al.* 2007 showed that the abundance of *Roseobacter* clade is highly correlated with the concentrations of POM and Chl *a* in the subtropical Atlantic Ocean which again reflects its affinity to nutrient enriched conditions. This corresponds to the findings of this thesis as protein groups assigned to the genus *Roseobacter* are more abundant during the peak of the bacterial bloom where the phytoplankton bloom already decayed. Candidate *Pelagibacter* (former SAR11) is one of the most abundant bacteria in the ocean and generally considered FL (Alonso and Pernthaler 2006; Heins *et al.* 2021). However, this taxon is also detected in PA fractions of sequential filtrations (Mestre *et al.* 2020; Heins *et al.* 2021) and even as one of the most abundant genera in this thesis. Heins *et al.* 2021 suggested that the cells of this taxon have less hydrophobic cell surface properties than other planktonic cells, which may facilitate the attachment to algal particles while they sink and collide with bacterial cells. The adsorption to particles could provide these cells with transient access to elevated nutrient concentrations and thus contribute to their ecological success compared to other non-motile FL bacteria. Additionally, they found evidence that the particle association of this taxon is of loose nature. These less

hydrophobic cell surface properties were previously also linked to grazing avoidance (Dadon-Pilosof *et al.* 2017).

Besides the differences between the two analyzed PA fractions, the phylogenetic assignment of the individual protein groups also indicated some notable taxonomic differences between the different stages of the investigated bacterial response bloom. Protein groups of bacterial taxa generally more abundant in the pre-bloom stage were for example assigned to *Cyanobacteria*. As cyanobacterial species were shown to be growth promoting for seaweeds (Hollants *et al.* 2013) the obtained results could be a hint for a growth promoting role of *Cyanobacteria* in the beginning of a phytoplankton bloom. Nevertheless, the abundance of *Cyanobacteria* in the pre-bloom stage can also be explained by their phototrophic lifestyle as they are not dependent on DOM released by phytoplankton cells to reach high abundances. Another taxon with higher abundance in the pre-bloom samples is the gammaproteobacterial genus *Pseudomonas*. One possible explanation for this abundance could be a mutualistic interaction with algae. It has been shown that the initiation of a *Chatonella* bloom is promoted by different *Pseudomonas* species. One example for such a growth promoting species is *Pseudomonas asplenii*, which is suggested to release heat-resistant substances, such as inorganic nutrients, and thereby elevates phosphate concentration. This mechanism seems to be dependent on the presence of the algae, suggesting a mutualistic interaction (Park *et al.* 2016). Protein groups assigned to the verrucomicrobial class *Verrucomicrobiae*, the planctomycetal genus *Phycisphaera* (in the 3 – 10 µm fraction) and the flavobacterial genus *Polaribacter* (in the ≥ 10 µm fraction) were more abundant during the rise of the bacterial response bloom. *Verrucomicrobia* are suggested to be efficient biopolymer degraders (Martinez-Garcia *et al.* 2012), also indicated by their fast response to the decaying diatom bloom in this thesis. Genomes of *Polaribacter sp.* show high peptidase proportions, small genomes and limited polysaccharide utilization capacity which makes them typical first responders (Avci *et al.* 2020). A high abundance during the rise of the bacterial response bloom may indicate that these genera mediate the decomposition of easily accessible algal polysaccharides e.g., storage compounds like the beta-1,3-linked glucose polymer laminarin.

Genera mainly abundant during the peak of the bacterial response bloom were e.g., the alphaproteobacterial genera *Sulfitobacter*, *Tateyamaria*, *Roseovarius* and *Roseobacter* as well as the verrucomicrobial genus *Opituae*. A high abundance during the peak of the bacterial response bloom may indicate that these genera mediate the decomposition of refractory algal polysaccharides e.g. sulfated and branched polysaccharides. Some genera are abundant mainly during the peak of the bacterial response bloom but only in one of the analyzed fractions. In the 3 – 10 μm fraction the alphaproteobacterial genus *Planktomarina* and the flavobacterial genera *Tenacibaculum* and *Polaribacter* show a high abundance in the bloom peak and in the $\geq 10 \mu\text{m}$ fraction the actinobacterial genus *Rhodococcus*, the alphaproteobacterial genera *Hyphomonas* and *Roseobacter*, the gammaproteobacterial genera *Oceanicoccus* and *Psychrosphaera* as well as the flavobacterial genera *Muricauda* and *Leeuwenhoekiella*.

Previous studies on FL bacterioplankton responding to a spring phytoplankton bloom found *Alphaproteobacteria* dominating the early bloom phase, *Flavobacteriia* increasing in relative abundance as the bloom commences and *Gammaproteobacteria* increasing as the bloom decays (Lucas *et al.* 2015; Teeling *et al.* 2012). The results of this thesis showed a different trend for the PA fractions concerning the *Alpha*- and *Gammaproteobacterial* abundance. Protein groups assigned to *Alphaproteobacteria* were generally more abundant during the peak of the bacterial response bloom, thus post the first diatom-dominated phytoplankton bloom and in the mid of the second phytoplankton bloom, and protein groups assigned to *Gammaproteobacteria* were generally more abundant during the pre-bloom stage (Figure 13). This may be explained by the different specialization of abundant taxa within these classes.

Phylogenetic succession of eukaryotic taxa

The spring phytoplankton bloom 2018 was dominated by diatoms of the genera *Thalassiosira*, *Chaetoceros*, *Mediopyxis* and *Rhizosolenia* and the haptophyte *Phaeocystis* (Heins *et al.* 2021). In this thesis protein groups assigned to the orders *Thalassiosirales*, *Chaetocerotales* as well as *Rhizosoleniales* reached high overall abundances and also protein groups assigned

to the order *Phaeocystis* could be identified (Figure 15; Table S7 CD-ROM). The spring phytoplankton boom at Helgoland Roads 2009 was also dominated by *Thalassiosira* species (Teeling *et al.* 2012). The mainly pre-bloom and bacterial bloom rise abundance of the phytoplankton taxa reflect the phytoplankton bloom which is decaying throughout the sampling timepoints. Protein groups of eukaryotic taxa generally more abundant in the pre-bloom stage were assigned to the microalgal phyla *Bacillariophyta* and *Cryptophyta* while protein groups assigned to the microalgal phyla *Chlorophyta*, *Ochrophyta* and *Streptophyta* were generally more abundant during the rise of the bacterial bloom (Figure 15). These results visualize that a typical phytoplankton bloom is a succession of phytoplankton taxa during the course of the bloom (Buchan *et al.* 2014).

As expected, protein groups assigned to the Arthropoda were more abundant in the $\geq 10 \mu\text{m}$ fraction due to their size. Furthermore, their abundance mainly during the rise of the bacterial response bloom can be a hint on grazing activity. For example the calanoid copepods *Neocalanus cristatus* and *Neocalanus plumchrus* are known grazers in the subarctic Pacific (Frost *et al.* 1983) and belong to the order Calanoida which is the most abundant order within the phylum Arthropoda in the metaproteomes of the spring bloom 2018.

Priest *et al.* 2021 investigated the dynamics of unicellular mycoplankton during a spring phytoplankton bloom at Helgoland Roads in 2017. They found evidence for the dominance of members of *Ascomycota*, *Basidiomycota* and *Chytridiomycota* with highest taxonomically rich community composition in the pre-bloom stage of the respective phytoplankton bloom. These were also the most abundant fungal taxa identified in this thesis with highest abundance in the pre-bloom and bacterial bloom rise stages (Figure 15). The annual recurrence indicated by the results of this thesis and Priest *et al.* 2021 was also observed in a study of Cunliffe *et al.* 2017, where they were able to isolate several annually recurring mycoplankton strains from the English Channel. Additionally, they could show their ability to degrade phytoplankton derived polysaccharides indicating that they may also contribute to phytoplankton biomass degradation during algal blooms. This could be a possible explanation for their abundance during the rise

of the bacterial bloom as the phytoplankton bloom decays. Furthermore, it was shown that fungal communities, dominated by the *Chytridiomycota*, parasitize light-stressed diatoms in the coastal Arctic environment (Hassett and Gradinger 2016). Zhang *et al.* 2018a also found evidence that phytoplankton blooms drive distinct changes within the fungal community with different effects across the taxa which reflects diverse ecological niches.

Unexpectedly the highest fold changes were observed for red algae (*Rhodophyta*) (Figure 16). Red algae are macroscopic and sessile, thus found e.g., attached to stones. A BLAST analysis of the identified protein groups showed that the corresponding peptide sequences have similarity to proteins mainly also assigned to the *Rhodophyta* but also to e.g., higher plants, *Oomycota*, *Planctomycetes* or *Haptophytes*. An explanation for the somewhat unexpected abundance of *Rhodophyta* could be a misclassification in the taxonomic annotation via the NCBI nr and Uniprot databases. However, the BLAST analysis showed, that most of the protein groups indeed belonged to the *Rhodophyta*. A possible explanation would be pieces of macroalgae in the water column but also a existence of macroalgal propagules in the plankton (Amsler *et al.* 1992). Propagules of macroalgae consist of unicellular products of sexual and asexual processes formed as a result of meiosis and mitosis such as e.g. zygotes and spores (Clayton 1992). Nevertheless, to make any clear statements about this unique high fold changes it would be necessary to investigate more time points of the bloom in smaller clearances.

Surprisingly, also protein groups of the taxon *Chiroptera* (bats) were identified. A BLAST analysis of the appropriate identified protein groups showed that the corresponding peptide sequences have similarity to proteins assigned to mainly birds but also *Saccaromycetes*.

These two examples highlight the importance of the correct annotation of eukaryotic proteins as explained more in detail in the Conclusions (page - 120 -).

7.3.2. Functional succession of marine particles during the 2018 spring phytoplankton bloom

Functional succession of bacterial taxa

Besides ribosomal proteins and chaperones the most abundant identified protein groups belong to the group transport and binding proteins (Figure 16; Table S8 CD-ROM). Hagström *et al.* 2020 examined transporter genes in metagenomic and metatranscriptomics data from a time-series survey in the Baltic Sea (0.2 – 3 µm fraction) and showed that the relative abundance of transporter stayed the same throughout the year. They suggested that most major types of transporters are present at a given time point but that their distribution patterns depend on bacterial groups being dominant at a given time of the year. Other studies revealed a differentiated expression of transporters between e.g. geographic locations or bloom vs. non-bloom conditions as well as PA vs. FL (Klindworth *et al.* 2014; Satinsky *et al.* 2014). The most abundant group of transport proteins were the ABC-transporter, which were generally more abundant in the 3 – 10 µm fraction and during the peak of the bacterial bloom (Figure 16). ABC-transporter are involved in the transport of LMW OM like di-/oligopeptides and carbohydrates including glycerol-3-phosphate and LMW carbohydrates including xylose, fructose, nucleosides, ribose and maltose. Their abundance during the peak of the bacterial bloom corresponds to their substrate spectrum, as the phytoplankton cells already decayed. Within the group of ABC-transporter phosphate and amino acids as well as oligosaccharide and monosaccharide transporter were the most abundant (Table S8 CD-ROM). The second most abundant transporter group were porins and therein gram-negative porins, phosphate-selective porins O and P as well as OmpA (Figure 16, Table S8 CD-ROM). Phosphate-selective porins O and P were mainly abundant during the peak of the bacterial response bloom with a steadily increasing abundance over the course of the bloom. Together with the high abundance of phosphate ABC-transporter these results indicate a low phosphate concentration as expected during phytoplankton blooms. Algal growth leads to rapid exhaustion of phosphate but also silicate, which diatoms use for frustule formation. Nutrient exhaustions of e.g. nitrogen, phosphorous, iron and silicate are typically limiting factors, which

together with grazing by zooplankton contribute to phytoplankton bloom termination (Schoemann *et al.* 2005; Buchan *et al.* 2014). Another phosphate acquisition strategy is the usage of phytases, a phosphatase enzyme that catalyzes the hydrolysis of phytic acid and releases a usable form of inorganic phosphorous (Mullaney *et al.* 2000). Phytic acid is mainly found in plants but was also detected in marine waters and marine sediments (Lim *et al.* 2007), and is rapidly mineralized upon introduction to the marine environment (Suzumura and Kamatani 1995) possibly due to an enhanced solubility (Lim *et al.* 2007). Phytase proteins were mainly abundant during the peak of the bacterial response bloom and in the $\geq 10 \mu\text{m}$ fraction also in the rise of the bacterial bloom (Table S8 CD-ROM). These results highlight the importance of inorganic nutrients like phosphate. In addition to phosphate also iron is an important micronutrient for most microorganisms, whose acquisition in the marine environment is hampered by its scarce bioavailability (Coale *et al.* 1996; Martin and Michael Gordon 1988). Many marine bacteria alleviate iron limitation by excreting small organic molecules with exceptionally high affinity for iron, called siderophores (Vraspir and Butler 2009). In this study no siderophores could be identified but chelatases and ferritin mainly abundant during the rise of the bacterial response bloom (Table S5 CD-ROM). Ferritin is one of the major non-haem iron storage proteins in animals, plants and microorganisms. It consists of a mineral core of hydrated ferric oxide and a multi-subunit protein shell that encloses the former and assures its stability in an aqueous environment (Chiancone *et al.* 2004).

Another abundant group of transporters are TBDT (Figure 16), whose important role in transport of DOM was recently recognized (Blanvillain *et al.* 2007; Schauer *et al.* 2008; Teeling *et al.* 2012). Most of the so far investigated TBDTs are involved in the uptake of bulky iron-siderophore complexes and vitamin B12 (Schauer *et al.* 2008). The transport through the outer membrane requires a Sus-like TonB-dependent receptor (TBDR) protein. The import may be energetically driven by the electrochemical gradient of the cytoplasmic membrane, where the energy is transduced by the TonB-complex (ExbBD proteins and TonB, here grouped as TonB-dependent transporter) to the TBDR (SusC and SusD) (Jordan *et al.* 2013; Gresock *et al.* 2015; Sverzhinsky *et al.* 2015). The number of TBDR genes in gram-negative bacteria differ widely.

The majority of taxa hosts about 14 TBDR genes, but single genomes can harbor up to 120 TBDRs (Blanvillain *et al.* 2007). High numbers of TBDR genes are typically found in *Bacteroidetes*, *Alpha-* and *Gammaproteobacteria* (Blanvillain *et al.* 2007). This is also confirmed by the results of this thesis, where TBDR and TBDR proteins are mainly assigned to these three phyla (Figure 16; Table S5 CD-ROM). Teeling *et al.* 2012 found that the blooming FL clades of the 2009 spring phytoplankton bloom were characterized by distinct substrate spectra, reflected by different CAZymes and transporter gene expression profiles. The succession of distinct FL bacterial clades coincided with the succession of distinct gene function repertoires, which indicated that changes in the algal substrate composition over the course of the bloom provided series of ecological niches in which dedicated heterotrophic planktonic bacteria could bloom. During the 2018 spring phytoplankton bloom TBDR in the PA fractions were mainly expressed by the alphaproteobacterial *Hyphomonas*, the flavobacterial *Ulvibacter* and the gammaproteobacterial *Umboniibacter*, all also mainly abundant at the peak of the bacterial response bloom (Table S5 CD-ROM). However, to confirm that also PA bacteria show similar successions as observed for the FL bacteria, more time points of a spring phytoplankton bloom need to be investigated and the number of identified bacterial protein groups need to be higher.

Before polysaccharides can be transported into the cell, they have to be degraded into oligosaccharides. Besides transporter also protein groups belonging to the group Carbohydrate metabolism could be identified including different families of glycosyl hydrolases (GH). One of the identified GH families is family 27 and annotated as an α -L-fucosidase. Fucoidans are a family of sulfated homo- and heteropolysaccharides found in various species of brown algae and mainly composed of sulfates and L-fucose with less than 10% other monosaccharides (Trincone 2018). Together with the identified sulfatases (PF00884 and PF14707; TableS8 CD-ROM) this could be an evidence for fucoidan degradation. Congruously, Tang *et al.* 2017 showed that fucoidan is hydrolyzed in aggregates and not in ambient waters. Protein groups of the GH family 17 and 2 were also identified during the spring bloom 2018. GH families 16 and 17 are mainly laminarinases but also GH family 2 is found in

PULs capable for the degradation of laminarin and other β -glucans (Ziervogel and Arnosti 2008). The polysaccharide Alginate is degraded by alginate lyases, thus presence of the PL family 17 indicates degradation of alginate. Alginate lyases initiate alginate degradation outside of the cell through endo-cleavage of the polysaccharide chain into shorter oligosaccharides, which are then substrates for the import system (Thomas *et al.* 2012; Kabisch *et al.* 2014; Bolam and Koropatkin 2012; Hehemann *et al.* 2014). Alginate is produced by brown algae (*Pheophyceae*) as structural polysaccharide and by certain genera of gram-negative bacteria e.g. *Azobacter* and *Pseudomonas sp.* for example as virulence factors (Trincon 2018). The two most abundant identified *Pseudomonas sp.* in this thesis are *Pseudomonas alcaligenes* and *Pseudomonas oleovorans* (Table S6 CD-ROM). Protein groups assigned to this species are mainly abundant in pre-bloom samples (Table S6 CD-ROM; Figure 13) and both species belong to the *Pseudomonas aeruginosa* group based on sequence similarity. *Pseudomonas aeruginosa* is known to produce alginate (Rehman *et al.* 2013). Furthermore, protein groups belonging to the GH family 28 mainly annotated as pectate lyases were identified (Table S8 CD-ROM). Pectins are biopolymers whose main structural feature is a linear (1-4)-linked chain of galacturonic acid. However, no other protein groups involved in pectin degradation could be identified. The opposite could be observed for the polysaccharide Xylan. None of the known xylanases (GH5, 7, 8, 10, 11, 43) was identified during the spring bloom 2018 but an acetyl-xylan esterase (PF05448), a xylose isomerase (PF01261), a D-xylose ABC-transporter (TIGR02634) as well as five transketolases and transaldolases. The latter convert xylose-5-phosphate into fructose-6-phosphate and glyceraldehyde-3-phosphate which subsequently can enter the glycolysis cycle. All these protein groups were characterized as part of xylan degradation in the Bacteroidetes *Gramella flava* JLT2011 (Tang *et al.* 2017). Xylans are components of marine phytoplankton (Avci *et al.* 2020; Parsons *et al.* 1961) and high xylanase activities have been reported for many ocean provinces (Arnosti 2011). Taken together these results indicate the degradation of fucoidan, laminarin, alginate and xylan. This corresponds to the observations of the study of Krüger *et al.* 2019, where a restricted set of conserved

polysaccharide utilization loci (PULs) predominantly targeting laminarin, α -glucans, α -mannose-rich substrates and sulfated xylans were identified in *Bacteroidetes* species.

The presence of GTs shows the capacity of carbohydrate biosynthesis. The identified GT family 2 is probably responsible for the production of lipopolysaccharides (LPS) or extracellular polymeric substances (EPS) (Barbeyron *et al.* 2016b). Furthermore, the identified GT family 1 was identified on a fosmid in a cluster interspersed with genes encoding proteins involved in the synthesis of the capsule polysaccharide colonic acid and in LPS biosynthesis (Gómez-Pereira *et al.* 2012). EPS act as a diffusive barrier against chemicals, such as xenobiotica or biocides, protect microorganisms from physical stress like UV-radiation and dehydration as well as counteract protozoan grazing (Carvalho 2007; Flemming and Wingender 2010). Additionally, EPS ensure sorption of inorganic ions and organic compounds, promote exchange of genetic information and cell-to-cell communication and the formation of micro colonies (Flemming and Wingender 2010). *Bacteroidetes* are known to both produce and degrade EPS (Fukami *et al.* 1997; Nichols *et al.* 2005). Bennke *et al.* 2013 investigated $\geq 10 \mu\text{m}$ plankton fraction from Helgoland surface waters in April 2011. They could show that *Polaribacter* and *Ulvibacter* were able to form micro colonies within aggregates which indicates an active growth and production of EPS. In contrast *Formosa A* cells rarely formed micro colonies but were frequently found attached to aggregates as single cells. Additionally, they could show that the cells of the dominant diatom *Chaetoceros* were preferred habitats for *Polaribacter*. This diatom was also identified as one of the most abundant phytoplankton taxa in this thesis (Figure 15) as well as *Polaribacter* and *Ulvibacter* as abundant bacterial genera (Figure 13). The production of EPS has also previously been suggested for marine *Bacteroidetes* by genome annotation (Gómez-Pereira *et al.* 2012). Besides the identified GT families also exopolysaccharide biosynthesis protein YbjH (PF0602) assigned to *Verrucomicrobia* could be identified in the spring bloom 2018. Together with the high abundance of *Bacteroidetes* it can be suggested that EPS production plays a considerable role in PA fractions.

Furthermore, protein groups associated with transcription and protein synthesis were found with high overall expression especially in the bloom rise and bloom peak samples (Figure 16), reflecting higher gene expression due to high proliferation rates.

Identified protein groups were mainly assigned to the most abundant taxonomic phyla and classes, thus *Alpha-* and *Gammaproteobacteria* as well as *Flavobacteriia* together with one of the following phyla: *Verrucomicrobiae*, *Actinobacteria*, *Planctomycetes* and *Acidobacteria*. Throughout the bacterial response bloom similar patterns to the overall abundance of the phyla could be observed. The more abundant a phylum, the more protein groups assigned to this phylum were identified. For more detailed insights into the coupling of function and taxonomy more protein groups need to be identified. For example, in the group of carbohydrate metabolism a maximum number of 12 protein groups per phylum and timepoint could be identified. Also, in groups with more identified protein groups, like transport and binding proteins (maximum 209 protein groups per phylum and timepoint), only for the abundant phyla data from several timepoints is available (Figure S 2, Appendix).

Functional succession of protein groups assigned to eukaryotic taxa

As expected, most of the identified protein groups assigned to phytoplankton species belonged to the group of photosynthesis proteins (Figure 17; Table S9 CD-ROM). Besides Chlorophyll also light-harvesting protein groups specific for marine organisms could be identified: Peridinin-chlorophyll A binding proteins (PF02429) belong to a water-soluble light-harvesting complex that has a blue-green absorbing carotenoid as its main pigment and is present in most photosynthetic dinoflagellates (Hofmann *et al.* 1996). Furthermore Phycoerythrin (PF02972), a light-harvesting red protein-pigment complex present in red algae and *Cryptophytes* (Ficner and Huber 1993; van der Weij-De Wit *et al.* 2006), was identified (Table S9 CD-ROM). The high abundance of protein groups associated with protein synthesis in the pre bloom stage (Figure 17) indicates induced growth of phytoplankton cells at the onset of the phytoplankton bloom.

Protein groups assigned to motility were mainly abundant during the pre-bloom stage and in the 3 – 10 μm fraction (Figure 17). One example is the group SF-assembling/beta giardin (PF06705) which presents the major component of striated microtubule-associated fibers (SMAFs) in the flagellar basal apparatus of green flagellates (Lechtreck and Melkonian 1998).

One of the highest fold changes compared to the pre-bloom stage was found for ubiquitination protein groups in the bloom rise of the $\geq 10 \mu\text{m}$ fraction (Figure 16). However, ubiquitination affects proteins in many ways e.g. mark them for degradation via the proteasome, alter their cellular location, affect their activity as well as promote or prevent protein interactions.

Interactions between marine bacteria and phytoplankton

Diatoms and bacteria co-occurred in a common habitat throughout the oceans for more than 200 million years and fostered interactions over evolutionary time scales (Amin *et al.* 2012). Hundreds of genes in diatom genomes appear to have been acquired from bacteria. It is suggested that this acquisition plays a major role in diversity and success of diatoms (Armbrust *et al.* 2004; Bowler *et al.* 2008). One example is the diatom *Phaeodactylum tricirnutum*, for which it is estimated that 784 genes were acquired by the diatom from bacteria. These genes are involved in nitrogen and organic carbon utilization, cell wall assembly, DNA recombination as well as ornithine-urea cycle (Allen *et al.* 2011; Bowler *et al.* 2008). *Phaeodactylum* is a genus in the order *Naviculales*. During the spring bloom 2018 *Phaeodactylum tricirnutum* was the most abundant species within the *Naviculales* (Table S7 CD-ROM) and many identified protein groups for this species were involved in Carbohydrate metabolism and Nitrogen metabolism (Table S5 CD-ROM), supporting the above mentioned estimations.

The ability of diatoms to influence bacterial diversity was observed in laboratory isolates (Amin *et al.* 2012). *Alphaproteobacteria* and therein *Sulfitobacter*, and *Roseobacter*, as well as *Bacteroidetes* and to a lesser extend *Betaproteobacteria* were the most prominent isolates across examined diatoms (Schäfer *et al.* 2002). Numerous other studies demonstrated that *Proteobacteria* and *Bacteroidetes* are the main heterotrophic bacterial phyla associated with diatoms and within these phyla specific genera like *Sulfitobacter*, *Roseobacter* and

Flavobacteriia appear to be strongly associated with diatoms based on their repeated occurrence in different studies (Grossart *et al.* 2005; Kaczmarska *et al.* 2005; Sapp *et al.* 2007a; Sapp *et al.* 2007b; Sapp *et al.* 2007c). Also, the observations of this thesis correspond to these findings. For detailed discussion please see 6.3.1. Protein groups assigned to the alphaproteobacterial genus *Roseobacter* found more abundant in the $\geq 10 \mu\text{m}$ fraction (Figure 13). *Roseobacter* genera are known to cause gall formation in red algae (Egan *et al.* 2014). However the *Roseobacter* group contains over 70 genera (Simon *et al.* 2017) with highly versatile genetic repertoire (Newton *et al.* 2010; Moran *et al.* 2007) and are also shown to establish symbiotic relationships with diatoms. (Durham *et al.* 2015; Croft *et al.* 2005). They for example promote algal growth by secreting antibiotics and growth stimulants (Buchan *et al.* 2014; Amin *et al.* 2015). This corresponds to the observation that protein groups for the biosynthesis of antibiotics are most abundant in the pre-bloom stage (Figure 16).

Protein groups assigned to the flavobacterial genus *Maribacter* are more abundant in the pre-bloom samples (Figure 13). *Maribacter sp.* are known to produce algal growth- and morphogenesis-promoting factors (AGMPFs) (Alsufyani *et al.* 2020), which act similarly to auxins by promoting rhizoid initiation and cell wall formation (Spoerner *et al.* 2012).

Microbial activity against algae-derived organic matter is not limited to dead diatoms but also includes actively growing cells. Some bacteria are consistently associated with growing diatoms through specific interactions, while other bacteria colonize sinking diatom particles and decompose organic matter therein. The function of the identified protein groups fit well with the current understanding of ecology of an algal- or surface-associated microbial community and can mostly be grouped into the following categories:

- a. *Detection and movement towards host surface/particles*: Protein groups associated with chemotaxis were partially abundant (3 – 10 μm bloom rise, $\geq 10 \mu\text{m}$ bloom peak) and protein groups associated with flagellum-mediated motility (in all stages but most abundant during bloom peak) consistently abundant. These both mechanisms are important for the detection and movement towards the algal host surface during

colonization. During early growth phases, phytoplankton cells release soluble LMW molecules like amino acids, carbohydrates, sugar alcohols and organic acids (Buchan *et al.* 2014; Bjørrisen 1988) of which many are potent chemoattractants for bacteria (Seymour, JR *et al.* 2010; Miller *et al.* 2004). Many bacteria are known to use chemotaxis, a phenomenon that relies on detection of molecules in the immediate surrounding of a cell to determine swimming direction either towards or away from a chemical gradient. Chemotaxis is essential for the development and maintenance of symbiotic surface associations (e.g. in *Rhizobium* species; Munoz Aguilar *et al.* 1988). Flagellum-mediated motility is also important for biofilm formation in a range of bacteria (Hossain and Tsuyumu 2006; Houry *et al.* 2010; Lemon *et al.* 2007).

- b. *Attachment and biofilm formation*: Bacterial attachment to living cells in the marine environment has been reported in many studies including a wide range of phytoplankton (Biegala *et al.* 2002; Gärdes *et al.* 2011; Zehr 2015; Cornejo-Castillo *et al.* 2016). It is suggested that diatoms may use TEP to attract certain types of bacteria and that these bacteria recognize the presence of the diatom and initiate attachment to TEP. Several bacteria were also found to influence TEP production when they are added to bacterium-free cells of the diatom *Thalassiosira weissflogii* (same order that dominated bloom in 2018) (Gärdes *et al.* 2011). Other adhesins implicated in cell-surface and cell-cell interactions are cadherin and fascilin that contain peptide motifs conserved in all domains of life (Woyke *et al.* 2009). As expected both were also identified in this thesis (Table S5 CD-ROM). One example are identified protein groups with a bacterial cadherin-like domain (PF17803). This domain is e.g. found in a calcium-binding lectin that interacts with acidic EPS and capsular polysaccharides produced by *Rhizobium leguminosarum* (Abdian *et al.* 2013). Another example are proteins containing a LysM-domain (PF0176) which is found in a wide range of microbial extracellular proteins, where it is thought to provide anchoring to EPS. Furthermore, it is reported that the domain functions as signal for specific plant-bacteria recognition in bacterial pathogenesis (Spaink 2004) and may have a peptidoglycan binding function

(Buist *et al.* 2008). The second most abundant transporter group were porins and therein among others OmpA (Figure 16, Table S8 CD-ROM). Homologues of the OmpA protein are required for adhesion to mammalian and fish epithel cells in a range of *Proteobacteria* (Namba *et al.* 2008). Additionally, it could be shown that bacteria release EPS in response to the presence of phytoplankton cells, likely to initiate attachment (Rinta-Kanto *et al.* 2012). The presence of GT in this thesis suggest EPS production (Functional succession of bacterial taxa page - 99 -). Furthermore, protein groups containing a GGDEF domain (named after conserved central sequence pattern), here identified in 3 – 10 µm fraction mainly abundant during bloom peak (Table S5 and S8 CD-ROM) are involved in production and degradation of bis-(3'-5')-cyclic dimeric GMP (Ryjenkov *et al.* 2005; Simm *et al.* 2004). This is an important secondary messenger regulating the transition from a motile planktonic to a surface associated lifestyle by up-regulating the production of adhesion protein groups and biofilm matrix components (Kulasekara *et al.* 2005; Lee *et al.* 2007; Römling 2002; Tischler and Camilli 2004) or downregulating motility genes (Simm *et al.* 2004). Additionally, pilus protein groups were identified (3 – 10 µm pre-bloom, ≥ 10 µm bloom rise). Most common pili are involved in attachment to surfaces but some bacteria possess type IV pili, which are responsible for a type of cell movement on surfaces and is known as twitching motility. This type of motility is thought to be important for the formation of micro colonies during the initial formation of biofilms (Mattick 2002). Both flagellum and twitching motility are known to be necessary for the biofilm development of *Pseudomonas aeruginosa* (O'Toole and Kolter 1998). Another protein group typically identified in bacteria surviving in biofilms are cbb3-type cytochrome c oxidases (here ≥ 10 µm fraction mainly during bloom peak). These oxidases have a high affinity for oxygen and are associated with microaerobic metabolism in oxygen limited environments (Pitcher *et al.* 2002). Biofilms are known to be spatially heterogeneous and to contain pockets of low oxygen concentrations or even no oxygen in some areas (Beer *et al.* 1994).

c. *Response to the algal host environment.* A commonly investigated interaction between bacteria and diatoms is the bacterial production of vitamins required by different diatoms, with cobalamin (vitamin B12) as the most studied (Carlucci and Silbernagel 1969; Ryther and Guillard 1962). This is also confirmed by the results of this study, where protein groups involved in the biosynthesis of cofactors and prosthetic groups and carriers were mainly abundant in the pre bloom stage, also observed for protein groups involved in cobalamin biosynthesis (Figure 16, Table S5 CD-ROM).

Furthermore, protein groups associated with the metabolism of water-soluble polysaccharides produced by algae as e.g. xylose, enable bacteria to use these sugars as a source of carbon and energy.

In response to stress e.g. macroalgae can rapidly produce ROS and reactive nitrogen species (RNS) as defense mechanism (Kapoor *et al.* 2019; Cosse *et al.* 2007) known as oxidative burst. Marine bacteria commonly harbor antioxidant proteins, such as catalases and superoxide dismutases, capable of detoxifying reactive oxygen species (ROS) (Venisse *et al.* 2001; Dimitrieva *et al.* 2006; Munn *et al.* 2008). During the course of the spring bloom 2018 superoxide dismutase but also thioredoxin, glutaredoxins, glutathione S-transferases, sulfoxide reductases and peroxiredoxins could be identified (Table S8 CD-ROM). *Zobellia galactanivorans* for example, a well-studied PA bacterium, is well equipped to cope with oxidative burst (Barbeyron *et al.* 2016b). Superoxide is an important ROS that can cause cellular damage and can be quenched by superoxide dismutase. Thioredoxins and peroxiredoxins play a big role in the thiol-based defense and ROS damaged proteins can be repaired by methionine sulfoxide reductases (Barbeyron *et al.* 2016a). The abundance of protein groups controlling oxidative stress might represent a protective mechanism for surface communities.

d. *Lateral gene transfer.* Transposases whose functions are associated with lateral gene transfer are a source of dynamic genomic change and allow a rapid ecological adaptation (Ochman *et al.* 2000). This could provide a broad mechanism for facilitating the functional similarity of phylogenetically distinct bacteria on the surface of algae.

Transposases are the most abundant and most ubiquitous genes in nature (Aziz *et al.* 2010). They are most often transferred to other organisms within the same habitat and can be also shared by distantly related taxa (Hooper *et al.* 2009). The expression of transposases might contribute to the adaptation of the organism towards environmental stress through genome reorganization (Wemheuer *et al.* 2015; Steffen *et al.* 2014). This corresponds to the findings of this thesis, where transposases were identified as mainly abundant during the peak of the bacterial bloom in the 3 – 10 µm fraction (Figure 16; Table S8 CD-ROM).

- e. *Defense and virulence*: Many algae have the capability to produce antibiotics (Grossart 1999). Protein groups assigned to biosynthesis of antibiotics in phytoplankton were mainly abundant during the pre-bloom stage (Figure 16), as phytoplankton cells begin to increase. Examples are the lantibiotic biosynthesis dehydratase (PF14028) and tryptophan halogenase (PF04820). The latter catalyzes the chlorination of tryptophan to form 7-chlorotryptophan. This is the first step in the biosynthesis of pyrrolnitrin, an antibiotic with broad spectrum anti-fungal activity (Hammer *et al.* 1999). Lantibiotic biosynthesis dehydratases are involved in the biosynthesis of lantibiotics, a class of peptide antibiotics that contains one or more thioester bonds (Ortega *et al.* 2015). However, so far it was not possible to show that algae produce this type of antibiotics. Bacterial protein groups associated with Antibiotic- and Multidrug-efflux were mainly abundant in the ≥ 10 µm fraction and in the pre-bloom stage (Figure 13; Table S8 CD-ROM). One example are protein groups belonging to the Multi Antimicrobial Extrusion (MATE) family, which function as drug/sodium antiporters. This family mediates resistance to a wide range of e.g. fluoroquinolones, aminoglycosides and other structurally diverse antibiotics and drugs (Kuroda and Tsuchiya 2009). Besides the MATE family also antibiotic efflux pumps belonging to the Major Facilitator Superfamily (MFS) and to ABC-transporters were identified (Table S8 CD-ROM). However, also bacteria produce antibiotics. Protein groups assigned to the biosynthesis of antibiotics were as in phytoplankton cells mainly abundant during pre-bloom conditions and at the

peak of the bacterial response bloom in the $\geq 10 \mu\text{m}$ fraction (Figure 13; Table S8 CD-ROM). This suggests that the biosynthesis of antibiotics is coupled with the presence of eukaryotic cells in the pre-bloom stage and to rising number of bacterial cells in the peak of the bacterial response bloom. Namely protein groups associated with the biosynthesis of colicins were identified (PF04052; Table S8 CD-ROM). Colicins are protein antibiotics produced by strains of *Escherichia coli* and closely related bacteria. Colicins belong to the group of bacteriocins (van Nguyen *et al.* 2014) which are ribosomally synthesized antimicrobial peptides or proteins (Desriac *et al.* 2010) produced by bacteria to inhibit the growth of similar or closely related bacterial strains (Cotter *et al.* 2005). They are found in diverse bacterial species of terrestrial origins but were also found in marine animal-associated bacteria (van Nguyen *et al.* 2014; Desriac *et al.* 2010; Balcázar *et al.* 2010) as well as marine bacteria isolated from seaweeds (Prieto *et al.* 2012). Protein groups mediating antibiotic resistance in phytoplankton cells were as expected also mainly abundant during the pre-bloom stage (Figure 17). Most of the identified protein groups were assigned to chloramphenicol phosphotransferases but also protein groups involved in the resistance to tetracycline and quinolones were identified (Table S9 CD-ROM). Chloramphenicol is known as a highly potent inhibitor of bacterial protein biosynthesis due to the prevention of peptide chain elongation mediated by reversible binding to the peptidyltransferase center at the 50S ribosomal subunit of 70S ribosomes (Schlünzen *et al.* 2001). 80S ribosomes of eukaryotes are no targets of chloramphenicol. However, it has been assumed that chloramphenicol may interact with mitochondrial ribosomes whose structure is more similar to 70S ribosomes than 80S ribosomes (Schwarz *et al.* 2004). Chloramphenicol phosphotransferases mediate resistance due to inactivation of the substrate and thus mediate a resistance not limited to prokaryotes.

Protein groups related to toxin production were most abundant during the rise of the bacterial bloom in the $\geq 10 \mu\text{m}$ fraction (Figure 13; Table S8 CD-ROM). One example are protein groups belonging to the group of Tc toxin complexes like the insecticide

toxin TcB which is e.g. known to cause disease in shrimp and are found in several marine bacteria (Tang and Lightner 2014; Sheets and Aktories 2017). This is consistent with the abundance in the larger size fraction due to the size of the infected species. Tc toxin complexes bind to the cell surface, are endocytosed and subsequently perforate the host endosomal membrane by forming channels that translocate toxic enzymes into the host (Meusch *et al.* 2014).

Furthermore, protein groups associated with the protection from foreign DNA were abundant during the rise of the bacterial bloom in both fractions as well as during the peak of the bloom in the 3 – 10 μm fraction (Figure 13, Table S8 CD-ROM) suggesting an increased importance of this defense mechanism with increasing number of bacterial cells. One example are protein groups with an HU-domain that are predominantly observed in *Bacteroidetes*. Proteins with this domain are predicted to play a role in recognition and possible interception of foreign DNA such as mobile elements and thereby preventing their incorporation into the host DNA (Burroughs *et al.* 2017). Besides protein groups with HU-domain also type I restriction enzymes were identified. They belong to the restriction and modification (R-M) systems, which are found in a variety of prokaryotes and are also suggested to protect the host bacterium from the uptake of foreign DNA by endonucleolytic cleavage of DNA that lacks a site specific modification (e.g. methylation) (Piekarowicz *et al.* 2001).

Type IV secretion systems on the other hand are often associated with virulent host/bacterium interactions (Backert and Meyer 2006; Alvarez-Martinez and Christie 2009) but can also mediate symbiotic interactions (Sullivan *et al.* 2002). This type of secretion system mediates the intracellular transfer of effector molecules between bacteria or from bacteria to eukaryotic cells. Another secretion system associated with virulence is the type III secretion system which builds an injectosome apparatus. In pathogenic bacteria, the needle-like structure is used as a sensory probe to detect the presence of eukaryotic organisms and secrete proteins that help the bacteria to infect them (Michiels *et al.* 1991; Lountos *et al.* 2012). Protein groups annotated as one of

these secretion systems are classed into the group virulence factors and mainly abundant during the peak of the bacterial response bloom in the 3 – 10 μm fraction (Figure 16; Table S8 CD-ROM). Another identified protein group belongs to the sialidase superfamily and is assigned as neuraminidase (PF15892, PF13088), a virulence factor for many bacteria including *Bacteroidetes fragilis* and *Pseudomonas aeruginosa*. This enzyme cleaves a sialic acid residue of ganglioside-GM1 off and thus turns it into asialo-GM1 to which type IV pili preferentially bind. Ganglioside-GM1 is a modulator of cell surface and receptor activity (Gaskell *et al.* 1995). The abundance of these systems during the peak of the bacterial response bloom suggests, that they become increasingly important with higher number of bacterial cells.

7.4. Proteomic analysis of a selected PA bacterial isolate grown on laminarin and alginate

During the metaproteomic analysis in this thesis the most abundant genus in the order *Flavobacteriia* during the spring bloom 2018 was *Muricauda*, which was mainly abundant during the peak of the bacterial response bloom in the $\geq 10 \mu\text{m}$ fraction (Figure 16). Furthermore, Zheng *et al.* (2020) identified *Muricauda* as one of the most abundant genera in a *Synechococcus* cultivation experiment. In this thesis *Muricauda* sp. MAR_2010_75, which was isolated from phytoplankton in the North Sea at the island Sylt (Hahnke and Harder 2013) was chosen for a proteomic analysis of its growth on polysaccharides. The polysaccharide laminarin is an algal carbon storage molecule and a major molecule in the oceanic carbon cycle (Becker *et al.* 2020). It is estimated that diatoms produce around 5 to 15 Gt per year, an amount making laminarin a major food resource for heterotrophic bacteria (Alderkamp *et al.* 2007). FL planktonic bacteria remineralize laminarin using the enzymes encoded in specific polysaccharide utilization loci (PULs). However, first insights into the genomes of PA bacteria showed the absence of all four well characterized PULs for the degradation of laminarin (Kappelmann *et al.* 2018) which is surprising due to the high abundance and importance of this storage polysaccharide. Another important polysaccharide in the marine realm is alginate,

which can account for approximately 50% of the dry weight of brown algae (Neumann *et al.* 2015). As described above evidence for the degradation of laminarin and alginate in the metaproteomes of the spring bloom 2018 was found, supporting the importance of these two polysaccharides in the marine realm. Kappelmann *et al.* (2018) sequenced the genomes of 53 North Sea *Flavobacteriia* and manually determined 400 potential PULs based on the presence of CAZyme clusters and co-occurring *susC/D*-like gene tandems. *Muricauda sp.* MAR_2010_75 was one of the investigated PA strains. Its predicted degradation capacity of polysaccharide classes based on PUL-associated CAZyme annotations is starch, glycogen, β -mannan, β -xylose, rhamnose and N-acetylglucosamin. Thus, besides the absence of PULs for laminarin degradation also PULs for the degradation of alginate seem to be absent in this strain. PULs can be manually detected based on the presence of CAZyme clusters and co-occurring *SusC/D* cluster, which in most cases are also part of a PUL (Bjursell *et al.* 2006; Kappelmann *et al.* 2018). In this thesis CAZymes were annotated based on the dbCAN database and for the identification of *SusC* proteins the TIGRFAM profile TIGR04056 (OMP_RagA_SusC) was used.

First of all, the growth of *Muricauda sp.* MAR_2010_75 on laminarin and alginate was tested. Growth curves of this species grown without the addition of sugars, with the addition of glucose and the addition of the polysaccharides laminarin and alginate (Figure 18) showed differences in the reached maximal OD (without-0.2; glucose-0.47; laminarin-0.39; alginate-0.35). The medium without addition of sugars contains a basal set of peptone and yeast extract, explaining the growth to an OD of 0.2. This medium is carbon-limited, allowing a higher growth to higher optical densities by the addition of further carbon sources. Thus, the chosen PA strain is able to use glucose, laminarin and alginate even if no PUL for the degradation of the latter two was predicted based on its genome. The lower maximal OD for cultures grown on polysaccharides compared with the monosaccharide glucose may be explained by additional protein requirements for the utilization of polysaccharides.

For the degradation and utilization of laminarin, a soluble β -1,3-glucan with β -1,6-linked glucose side chains, the respective heterotrophic bacteria require specific enzymes and transporters. Several CAZymes are required for the depolymerization of laminarin: the main laminarin-degrading enzymes are laminarinases, which are classified into *endo*- and *exo*- β -1,3-glucanases. The first hydrolyze the glucose backbone, while the latter cleave off glucose from the ends of laminarin oligosaccharides. *Endo*-acting laminarinases are mainly grouped into GH families 16, 17, 55, 64 and 81, while GH family 3 contains *exo*-acting laminarinases (Becker *et al.* 2017; Bäumgen *et al.* 2021). Furthermore, also GH2 proteins are found in laminarin PULs (Krüger *et al.* 2019). The transport of laminarin oligomers into the periplasm is performed by SusC-like TBDT and the subsequent transport into the cytoplasm by transporters of the major-facilitator-superfamily (MFS) transporter. In the cytoplasm the oligomers are degraded to glucose, which is metabolized to pyruvate within the glycolysis/Entner-Doudoroff pathway and subsequently introduced to the citric acid cycle (Koch *et al.* 2019). In the proteomes of *Muricauda sp.* MAR_2010_75 neither GH family 16 nor GH family 17 were identified. However two GH2, two GH3 and one GH30 were identified. One of the GH2 proteins (FG28_RS07890) as well as one of the GH3 proteins (FG28_RS00590) showed an induction compared to glucose, while the other showed none. The GH30 protein (FG28_RS07895) could only be identified in the samples grown on laminarin and was part of a predicted glucuronoxytan PUL (Kappelmann *et al.* 2018) (Table S10 CD-ROM). In the highly conserved and short laminarin PUL of *Gramella forsetii* KT0803 one predicted GH3 (β -1,3-glucosidase) is framed by two predicted GH16 (β -1,3(4)-glucanase) and sometimes accompanied by a GH2 and GH30_1. It is suggested that the cell-surface associated GH16 glucanases cleave branched laminarin polysaccharides into oligosaccharides (Labourel *et al.* 2015), which are transported into the periplasm where GH3 can further cleave off glucose units (Tamura *et al.* 2017) subsequently transported into the cytoplasm. In the species *Formosa sp.* Hel1_33_131 also a GH30 enzyme was found to be involved in the degradation of laminarin, where it removes β -1,6-glucose side chains (Becker *et al.* 2017). However, a combination of GH17 enzymes depolymerizing the laminarin backbone and GH30 enzymes hydrolyzing the side

chains in necessary for an efficient laminarin depolymerization (Becker *et al.* 2017). Thus no enzyme with known laminarinases activity for the degradation of laminarin into oligosaccharides could be identified in the investigated proteomes. However, the above mentioned induced proteins may have a yet unknown laminarinase activity, which could be confirmed via their biochemical characterization. Additionally, considering the identified proteins belonging to the same operon or genomic neighborhood of the identified CAZyme from other laminarin PULs, no laminarin PUL structure could be observed (Figure S 3 Appendix). Another approach was to find potential PULs based on their SusC protein. 10 SusC proteins could be identified in samples grown on laminarin (Table S10 CD-ROM). Two showed a fc over 2 compared to glucose, namely FG28_RS02370 which is part of a predicted β -mannan PUL (Kappelmann *et al.* 2018) and the second most abundant SusC according to its riBAQ. The most abundant SusC (FG28_RS02640) showed no induction compared to glucose. The second induced SusC (FG28_RS13765) was part of a predicted PUL with unknown substrate (Kappelmann *et al.* 2018) (Table S10 CD-ROM). Again, considering the identified proteins belonging to the same operon or genomic neighborhood of the identified SusC protein, no known laminarin PUL structure could be observed (Figure S 3 Appendix).

The polysaccharide alginate is a linear copolymer of two 1-4 linked uronic acid residues, β -D-mannuronate and α -L-guluronate. For its degradation different lyases degrading the polysaccharide into oligosaccharides are necessary. Subsequently the oligosaccharides can be transported into the periplasm via TBDT. Known alginate lyases belong to the PL families 6 and 7. For the degradation of oligomeric alginate into mannuronate and guluronate in the periplasm also enzymes of the PL family 17 are necessary. These monosaccharides can then be transported into the cytoplasm via MFS transporters where they are further degraded to 2-keto-3-deoxy-6-phosphogluconate, which can be assimilated via the Entner-Doudoroff pathway to pyruvate and subsequently introduced to the citric acid cycle (Neumann *et al.* 2015; Koch *et al.* 2019). In the proteomes of *Muricauda sp.* MAR_2010_75 neither of the known alginate lyases could be identified. In the translated genome of *Muricauda sp.* MAR_2010_75 two proteins show the NCBI conserved domain motif cl19188, which is annotated as PL6

superfamily. However, only one of the two protein could be identified in the proteome of the cultures grown on alginate namely FG28_RS15365, which showed a fc of 0.46 compared to glucose and thus was downregulated. Another approach was to find potential PULs based on their SusC protein. Four SusC proteins could be identified in samples grown on alginate (Table S10 CD-ROM). One showed an induction compared to glucose (FG28_RS13765), was also induced in cultures grown on laminarin and is part of a predicted PUL with unknown substrate (Kappelmann *et al.* 2018). Two of the identified SusC proteins (FG28_RS02370, FG28_RS02640) showed no induction compared to glucose and the last of the identified SusC (FG28_RS09405) was decreased during growth on alginate (Table S10 CD-ROM). Once more considering the identified proteins belonging to the same operon or genomic neighborhood of the identified SusC protein, no alginate PUL structure could be observed (Figure S 3).

Kappelmann *et al.* (2018) observed that sometimes the sequence similarity of TBDT was too low to be considered as *susC*-like or no *susD* homolog was present or even the entire *susCD*-like gene tandem was missing (Hemsworth *et al.* 2016). It has been shown that some PULs have *susC/D* gene pairs that are separated from the corresponding CAZyme genes elsewhere in the genome (e.g. Ficko-Blean *et al.* 2017). This is a possible explanation for the missing PUL structure surrounding induced SusC proteins. However, it is important to notice that PA bacteria face a larger variety of polysaccharides in nature than their FL counterparts. It seems possible that PA bacteria established more modular degradation systems different from the strategy from FL bacteria establishing one PUL for one substrate. One hint for this could be the induction of the SusC (FG28_RS13765) in cultures grown on both investigated polysaccharides. For example, PA bacteria could harbor genes for SusC/D pairs for neutral oligosaccharides (as it is the case for laminarin) as well as for acidic oligosaccharides (as it is the case for alginate). Another explanation for the missing PUL structures could be a vivid cross-talk between different PULs. Crosstalk between different PULs was previously observed e.g. a partial co-regulation of TBDRs by both laminarin and its monomer glucose occurs, indicating that not only the glycosidic linkage type but also sugar composition controls PUL-expression in marine *Bacteroidetes*. Additionally, a crosstalk between the PULs specific for the

β -1,3-glucan laminarin and α -1,4-glucans was observed, again indicating a role of sugar composition in the co-regulation of PULs (Kabisch *et al.* 2014). For *Bacteroidetes thetaiotaomicron* VPI-5482 it has been shown that binding of a monosaccharide to a hybrid two-component system can induce a specific PUL while simultaneously other PULs are repressed. This indicates a “hierarchical PUL expression and prioritization of polysaccharide use” (Lynch and Sonnenburg 2012). As PA bacteria face a higher diversity of polysaccharides these cross-talks could be more frequent and complex and thus difficult to track. However, the absence of all known enzymes necessary for the degradation of the two investigated polysaccharides despite their visible growth on them is very surprising. Another possible explanation is that PA bacteria established multimodular enzymes. Liu *et al.* (2016) discovered the novel multifunctional enzyme Amy63 produced by the marine bacterium *Vibrio alginolyticus* 63, which possesses amylase, agarose and carrageenase activities and is a substrate promiscuous α -amylase with substrate priority orders of starch, carrageenan and agar. The use of multimodular enzymes by PA bacteria is a possible scenario, as they live under nutrient enriched conditions and face a wide variety of different polysaccharides. Another possible explanation is that PA bacteria use as yet unknown enzymes for the degradation of this abundant polysaccharides. The enrichment of genes of unknown function within polysaccharide-responsive regulons, highlight the current lack of knowledge on catabolic pathways dedicated to marine polysaccharides, compared to for example glucose utilization for which the degradation machinery has been widely studied (Thomas *et al.* 2017). However, it also has to be mentioned that several Usp and Usp-like proteins were under the most induced protein during growth on the investigated polysaccharides (Table 12). Usp is a small cytoplasmic bacterial protein whose expression is induced when the cell is exposed to stress agents. It enhances the rate of cell survival during prolonged exposure to such conditions and may provide a general “stress endurance” (Nyström and Neidhardt 1994). These stress conditions include heat-shock, nutrient starvation, as well as presence of oxidants, DNA-damaging agents or other stress agents which may arrest cell growth. Most organisms have multiple paralogs of Usp where the number of copies depends on the organism (Tkaczuk *et al.*

2013). Besides nutrient depletion during growth on polysaccharides this high abundance could also be explained by the PA lifestyle of the used strain. Polysaccharides could be an inducer of global responses to a PA lifestyle. For example algae are known to produce toxic compounds due to photorespiration (Hünken *et al.* 2008). However, further experiments are necessary to make clear statements. For example, investigation of the growth and proteome of *Muricauda sp.* MAR_2010_75 on polysaccharides with corresponding predicted PUL e.g. starch or xylan as “positive control” could be a next step. Furthermore, the biochemical characterization of the induced proteins could give hints on e.g. possible multimodularity.

Muricauda sp. MAR_2010_75 seems to induce T9SS during growth on polysaccharides (Table S 3 Appendix). T9SS is a versatile protein transport apparatus restricted to the *Bacteroidetes* phylum and as multiprotein complex, enables secretion of a wide range of effectors to the milieu or to the cell surface (Lasica *et al.* 2017; Veith *et al.* 2017; Vincent *et al.* 2018). Polysaccharides need to be depolymerized to be taken up into the cytoplasm. For this extracellular depolymerization process proteins at the outer membrane surface are needed. As T9SS is a bacteriocidal excretion system the induction during growth on polysaccharides can be explained by the need of depolymerizing enzymes at the outer membrane. In the investigated proteomes of *Muricauda sp.* MAR_2010_75 PorT and PorP, two yet poorly characterized conserved components of the T9SS and PorV, a part of the T9SS attachment complex were identified. PorT was one of the most abundant proteins identified in the cultures grown on the investigated polysaccharides (Table 11). Some members of the *Bacteroidetes* use the T9SS for the secretion of adhesins required for gliding motility or enzymes required for the supply with carbon, haem and metal sources (Sato *et al.* 2010; Shrivastava *et al.* 2013; Kita *et al.* 2016; McBride and Zhu 2013) but also for the secretion of virulence factors (Nakayama 2015).

8. Conclusions and Outlook

In this thesis, a broadly applicable metaproteomics protocol for successful extraction of proteins from marine particles is presented. Its reliability was confirmed in the subsequent comparative metaproteomics analyses of marine microbial communities sampled in 2009, which were either free-living (FL) or attached to particles (PA) that gave first insights into the expression of life style-specific functions of bacteria living on particles. Furthermore, application of the established protocol on particulate communities sampled during a spring phytoplankton bloom in 2018 resulted in an as yet unequalled number of 5,607 (pre-bloom), 7,141 (bloom rise) and 4,118 (bloom peak) identified protein groups of the bacterial response bloom. The observed succession of bacterial clades throughout the bloom highlighted individual niche occupations. Additionally, functional data supported the important role of PA bacteria during the turnover of oceanic organic matter and were in good accordance with the current understanding of the ecology of an algal- or surface-associated microbial community, additionally highlighting the importance of phytoplankton-bacterial interactions in the oceans.

Although the optimized metaproteomics workflow significantly improved the identification rate of PA protein groups, the number of protein group identifications from the particles is still considerably lower compared to FL bacterial communities (9.354 in the application example; (Teeling *et al.* 2012). It is assumed that especially the high abundance of eukaryotic protein groups poses problems in protein group identification due to the complexity and diversity of microbial eukaryote genomes. Furthermore, the presence of introns and repeats hinder metagenomic analysis necessary for the construction of a protein sequence target database and thus also hinder peptide identification (Saito *et al.* 2019). Metaproteome coverage of marine particles could be significantly improved by optimizing the used database. This could be achieved by usage of customized databases including eukaryotic metatranscriptomics (RNA-based) sequence data. It was recently shown that a combined database of transcriptomes from isolated cultures and field metatranscriptomes provided richer metaproteome results in a study from the Ross Sea of the Antarctica, where the diverse bloom

community showed high abundances in phytoplankton (Cohen *et al.* 2021). Alternatively, protein group identification could also be substantially improved by extracting already existing metatranscriptomics and metagenomic data from relevant eukaryotic taxa from public databases. Key to the latter approach is reliable information on which eukaryotic organisms make up the particles, which can be attained by 18S rRNA gene amplicon sequencing. However, the creation of MS target databases is a balancing act between considering all sequences of proteins possibly present in the analyzed sample and reducing the database size to save computational costs and ease protein inference as well as false identifications. Thus, especially in the here presented metaproteomics analysis, an improvement of assembly and annotation of eukaryotic metagenomes would be the best option to significantly improve the number of identified protein groups. During the sequencing of environmental DNA millions of small reads are produced that must be reassembled *de novo* utilizing bioinformatics tools and software. The reassembly of this reads into contigs is still a serious computational challenge and the regeneration of entire gene sequences is impossible (Riesenfeld *et al.* 2004). One possibility to overcome these challenges is the usage of the third-generation sequencing platform by Pacific Biosciences (Rhoads and Au 2015). With this system longer reads up to 60 kb can be produced and the increased depth coverage and long overlapping reads allow the reconstruction of genomes with fewer obstacles. Another challenge in the sequencing of metagenomics data is the prediction of genes especially for eukaryotic DNA in metagenomes. Metagenomes comprise of a variety of sequences from distinct microorganisms and frequently constitute not only a limited number of long contigs but also short assemblies and unassembled reads (Teeling and Glöckner 2012). Moreover, metagenomes are usually permeated with frameshifts that make gene prediction even more difficult (Hoff 2009). However, the continuous development of new annotation pipelines and algorithms e.g. GeneMark-ET (Zhu *et al.* 2010a), AUGUSTUS (Stanke *et al.* 2004; Hoff and Stanke 2013) as well as BRAKER1 and 2 (Hoff *et al.* 2016; Brůna *et al.* 2021) can help to reduce these challenges in the future.

Further experiments employing the, in this thesis presented, protocol could include the metaproteomics analysis of samples generated without filtration for example by the use of plankton nets or Imhoff sedimentation cones (Heins *et al.* 2021). This would enable to additionally analyze chemotactic PA bacteria as members of the phycosphere and overcome the limitations of filtration technique. During filtration the retentate may decrease the effective pore size during filtration (Padilla *et al.* 2015) possibly leading to the retention of otherwise FL bacteria. Imhoff sedimentation cones were developed for the quantification of settable particles in wastewater treatment plants (Imhoff 1989). The steep incline in the cones results in the concentration of particles in a small volume at the bottom which can be readily sampled. This technique was recently applied in marine microbiology (Heins *et al.* 2021).

More detailed insights into the metabolism of PA bacteria were gained by a proteomic analysis of a PA bacterial isolate grown on the two naturally abundant marine polysaccharides laminarin and alginate. Obtained results were very surprising, as no known enzyme for the degradation of the polysaccharides into oligosaccharides could be identified. Evidences were found that PA bacteria employ more diverse degradation systems partially different from the strategies used by FL bacteria. Nevertheless, further experiments are essential. For example, investigation of the growth and proteome of *Muricauda sp.* MAR_2010_75 on polysaccharides with corresponding predicted PUL e.g., starch or xylan as “positive control” could be a next step. Furthermore, comparison with proteomic analysis of a PA bacterial isolate with predicted laminarin and alginate PULs can help to interpret the obtained data. Nevertheless, the biochemical characterization of the induced proteins, including proteins with domains of unknown function (DUF) or annotated as hypothetical, would be necessary to verify e.g., possible multimodularity. Additionally, the measurement of substrate concentration in the medium throughout the cultivation process is a way to verify the mineralization of the given substrate.

References

- Abdian, P.L., Caramelo, J.J., Ausmees, N., and Zorreguieta, A. (2013) RapA2 is a calcium-binding lectin composed of two highly conserved cadherin-like domains that specifically recognize *Rhizobium leguminosarum* acidic exopolysaccharides. *The Journal of biological chemistry*, doi: 10.1074/jbc.M112.411769.
- Abell, G.C.J., and Bowman, J.P. (2005) Colonization and community dynamics of class Flavobacteria on diatom detritus in experimental mesocosms based on Southern Ocean seawater. *FEMS Microbiology Ecology*, doi: 10.1016/j.femsec.2005.01.008.
- Alderkamp, A.-C., van Rijssel, M., and Bolhuis, H. (2007) Characterization of marine bacteria and the activity of their enzyme systems involved in degradation of the algal storage glucan laminarin. *FEMS Microbiology Ecology*, doi: 10.1111/j.1574-6941.2006.00219.x.
- Allredge, A.L., Passow, U., and Logan, B.E. (1993) The abundance and significance of a class of large, transparent organic particles in the ocean. *Deep Sea Research Part I: Oceanographic Research Papers*, doi: 10.1016/0967-0637(93)90129-Q.
- Allen, A.E., Dupont, C.L., Oborník, M., Horák, A., Nunes-Nesi, A., McCrow, J.P., *et al.* (2011) Evolution and metabolic significance of the urea cycle in photosynthetic diatoms. *Nature*, doi: 10.1038/nature10074.
- Alonso, C., and Pernthaler, J. (2006) Roseobacter and SAR11 dominate microbial glucose uptake in coastal North Sea waters. *Environmental microbiology*, doi: 10.1111/j.1462-2920.2006.01082.x.
- Alonso, C., Warnecke, F., Amann, R., and Pernthaler, J. (2007) High local and global diversity of Flavobacteria in marine plankton. *Environmental microbiology*, doi: 10.1111/j.1462-2920.2007.01244.x.
- Alonso-Sáez, L., Gasol, J.M., Arístegui, J., Vilas, J.C., Vaqué, D., Duarte, C.M., and Agustí, S. (2007) Large-scale variability in surface bacterial carbon demand and growth efficiency in the subtropical northeast Atlantic Ocean. *Limnology and Oceanography*, doi: 10.4319/lo.2007.52.2.0533.
- Alsufyani, T., Califano, G., Deicke, M., Grueneberg, J., Weiss, A., Engelen, A.H., *et al.* (2020) Macroalgal-bacterial interactions: identification and role of thallusin in morphogenesis of the seaweed *Ulva* (Chlorophyta). *Journal of experimental botany*, doi: 10.1093/jxb/eraa066.
- Alvarez-Martinez, C.E., and Christie, P.J. (2009) Biological diversity of prokaryotic type IV secretion systems. *Microbiology and molecular biology reviews : MMBR*, doi: 10.1128/MMBR.00023-09.
- Amin, S.A., Hmelo, L.R., van Tol, H.M., Durham, B.P., Carlson, L.T., Heal, K.R., *et al.* (2015) Interaction and signalling between a cosmopolitan phytoplankton and associated bacteria. *Nature*, doi: 10.1038/nature14488.
- Amin, S.A., Parker, M.S., and Armbrust, E.V. (2012) Interactions between diatoms and bacteria. *Microbiology and molecular biology reviews : MMBR*, doi: 10.1128/MMBR.00007-12.
- Amsler, C.D., Reed, D.C., and Neushul, M. (1992) The microclimate inhabited by macroalgal propagules. *British Phycological Journal*, doi: 10.1080/00071619200650251.
- Andersson, A.F., Riemann, L., and Bertilsson, S. (2010) Pyrosequencing reveals contrasting seasonal dynamics of taxa within Baltic Sea bacterioplankton communities. *The ISME journal*, doi: 10.1038/ismej.2009.108.
- Andrews, S. (2010) FastQC: a quality control tool for high throughput sequence data. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>, 2010.
- Ankney, J.A., Muneer, A., and Chen, X. (2018) Relative and Absolute Quantitation in Mass Spectrometry-Based Proteomics. *Annual review of analytical chemistry (Palo Alto, Calif.)*, doi: 10.1146/annurev-anchem-061516-045357.
- Armbrust, E.V. (2009) The life of diatoms in the world's oceans. *Nature*, doi: 10.1038/nature08057.
- Armbrust, E.V., Berges, J.A., Bowler, C., Green, B.R., Martinez, D., Putnam, N.H., *et al.* (2004) The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science*, doi: 10.1126/science.1101156.
- Arnosti, C. (2011) Microbial extracellular enzymes and the marine carbon cycle. *Annual review of marine science*, doi: 10.1146/annurev-marine-120709-142731.
- Arnosti, C., Durkin, S., and Jeffrey, W.H. (2005) Patterns of extracellular enzyme activities among pelagic marine microbial communities: implications for cycling of dissolved organic carbon. *Aquatic Microbial Ecology*, doi: 10.3354/ame038135.
- Avcı, B., Krüger, K., Fuchs, B.M., Teeling, H., and Amann, R.I. (2020) Polysaccharide niche partitioning of distinct *Polaribacter* clades during North Sea spring algal blooms. *The ISME journal*, doi: 10.1038/s41396-020-0601-y.
- Ayo, B., Unanue, M., Azúa, I., Gorsky, G., Turley, C., and Iriberry, J. (2001) Kinetics of glucose and amino acid uptake by attached and free-living marine bacteria in oligotrophic waters. *Marine Biology*, doi: 10.1007/s002270000518.
- Azam, F. (1998) OCEANOGRAPHY: Microbial Control of Oceanic Carbon Flux: The Plot Thickens. *Science (New York, N.Y.)*, doi: 10.1126/science.280.5364.694.

References

- Azam, F., Fenchel, T., Field, J.G., Gray, J.S., La Meyer-Reil, and Thingstad, F. (1983) The Ecological Role of Water-Column Microbes in the Sea. *Marine Ecology Progress Series*, doi: 10.3354/meps010257.
- Azam, F., and Malfatti, F. (2007) Microbial structuring of marine ecosystems. *Nature reviews. Microbiology*, doi: 10.1038/nrmicro1747.
- Aziz, R.K., Breitbart, M., and Edwards, R.A. (2010) Transposases are the most abundant, most ubiquitous genes in nature. *Nucleic acids research*, doi: 10.1093/nar/gkq140.
- Backert, S., and Meyer, T.F. (2006) Type IV secretion systems and their effectors in bacterial pathogenesis. *Current opinion in microbiology*, doi: 10.1016/j.mib.2006.02.008.
- Balcázar, J.L., Loureiro, S., Da Silva, Y.J., Pintado, J., and Planas, M. (2010) Identification and characterization of bacteria with antibacterial activities isolated from seahorses (*Hippocampus guttulatus*). *The Journal of antibiotics*, doi: 10.1038/ja.2010.27.
- Banfield, J.F., VerBerkmoes, N.C., Hettich, R.L., and Thelen, M.P. (2005) Proteogenomic approaches for the molecular characterization of natural microbial communities. *Omics : a journal of integrative biology*, doi: 10.1089/omi.2005.9.301.
- Barbeyron, T., Brillet-Guéguen, L., Carré, W., Carrière, C., Caron, C., Czjzek, M., Hoebeke, M., and Michel, G. (2016a) Matching the Diversity of Sulfated Biomolecules: Creation of a Classification Database for Sulfatases Reflecting Their Substrate Specificity. *PloS one*, doi: 10.1371/journal.pone.0164846.
- Barbeyron, T., L'Haridon, S., Corre, E., Kloareg, B., and 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. *International Journal of Systematic and Evolutionary Microbiology*, doi: 10.1099/00207713-51-3-985.
- Barbeyron, T., Thomas, F., Barbe, V., Teeling, H., Schenowitz, C., Dossat, C., et al. (2016b) Habitat and taxon as driving forces of carbohydrate catabolism in marine heterotrophic bacteria: example of the model algae-associated bacterium *Zobellia galactanivorans* DsijT. *Environmental microbiology*, doi: 10.1111/1462-2920.13584.
- Barras, D.R., and Stone, B.A. (1969) β 1,3-Glucan hydrolases from *Euglena gracilis*. *Biochimica et Biophysica Acta (BBA) - Enzymology*, doi: 10.1016/0005-2744(69)90252-6.
- Bates, S.S., Douglas, D.J., Doucette, G.J., and Léger, C. (1995) Enhancement of domoic acid production by reintroducing bacteria to axenic cultures of the diatom *Pseudo-nitzschia multiseries*. *Natural toxins*, doi: 10.1002/nt.2620030605.
- Bauer, M., Kube, M., Teeling, H., Richter, M., Lombardot, T., Allers, E., et al. (2006) Whole genome analysis of the marine Bacteroidetes 'Gramella forsetii' reveals adaptations to degradation of polymeric organic matter. *Environmental microbiology*, doi: 10.1111/j.1462-2920.2006.01152.x.
- Bäumgen, M., Deutschei, T., and Bornscheuer, U. (2021) Marine Polysaccharides: Occurrence, Enzymatic Degradation and Utilization. *Chembiochem : a European journal of chemical biology*, doi: 10.1002/cbic.202100078.
- Beattie 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. *The Biochemical journal*, doi: 10.1042/bj0790531.
- Becker, S., Scheffel, A., Polz, M.F., and Hehemann, J.-H. (2017) Accurate Quantification of Laminarin in Marine Organic Matter with Enzymes from Marine Microbes. *Applied and Environmental Microbiology*, doi: 10.1128/AEM.03389-16.
- Becker, S., Tebben, J., Coffinet, S., Wiltshire, K., Iversen, M.H., Harder, T., Hinrichs, K.-U., and Hehemann, J.-H. (2020) Laminarin is a major molecule in the marine carbon cycle. *Proceedings of the National Academy of Sciences*, doi: 10.1073/pnas.1917001117.
- Beer, D. de, Stoodley, P., Roe, F., and Lewandowski, Z. (1994) Effects of biofilm structures on oxygen distribution and mass transport. *Biotechnology and bioengineering*, doi: 10.1002/bit.260431118.
- Behrenfeld, M.J. (2014) Climate-mediated dance of the plankton. *Nature Climate Change*, doi: 10.1038/nclimate2349.
- Behrenfeld, M.J., O'Malley, R.T., Boss, E.S., Westberry, T.K., Graff, J.R., Halsey, K.H., et al. (2016) Revaluating ocean warming impacts on global phytoplankton. *Nature Climate Change*, doi: 10.1038/nclimate2838.
- Bengtsson, M.M., and Øvreås, L. (2010) Planctomycetes dominate biofilms on surfaces of the kelp *Laminaria hyperborea*. *BMC microbiology*, doi: 10.1186/1471-2180-10-261.
- Bennke, C.M., Krüger, K., Kappelmann, L., Huang, S., Gobet, A., Schüler, M., et al. (2016) Polysaccharide utilisation loci of Bacteroidetes from two contrasting open ocean sites in the North Atlantic. *Environmental microbiology*, doi: 10.1111/1462-2920.13429.

References

- Bennke, C.M., Neu, T.R., Fuchs, B.M., and Amann, R. (2013) Mapping glycoconjugate-mediated interactions of marine Bacteroidetes with diatoms. *Systematic and applied microbiology*, doi: 10.1016/j.syapm.2013.05.002.
- Bergauer, K., Fernandez-Guerra, A., Garcia, J.A.L., Sprenger, R.R., Stepanauskas, R., Pachiadaki, M.G., Jensen, O.N., and Herndl, G.J. (2018) Organic matter processing by microbial communities throughout the Atlantic water column as revealed by metaproteomics. *Proceedings of the National Academy of Sciences of the United States of America*, doi: 10.1073/pnas.1708779115.
- Bidle, K.D., and Fletcher, M. (1995) Comparison of free-living and particle-associated bacterial communities in the Chesapeake Bay by stable low-molecular-weight RNA analysis. *Applied and Environmental Microbiology*, doi: 10.1128/AEM.61.3.944-952.1995.
- Biegala, I.C., Kennaway, G., Alverca, E., Lennon, J.-F., Vaultot, D., and Simon, N. (2002) IDENTIFICATION OF BACTERIA ASSOCIATED WITH DINOFLAGELLATES (DINOPHYCEAE) ALEXANDRIUM SPP. USING TYRAMIDE SIGNAL AMPLIFICATION-FLUORESCENT IN SITU HYBRIDIZATION AND CONFOCAL MICROSCOPY 1. *Journal of Phycology*, doi: 10.1046/j.1529-8817.2002.01045.x.
- Bižić-Ionescu, M., Zeder, M., Ionescu, D., Orlić, S., Fuchs, B.M., Grossart, H.-P., and Amann, R. (2014) Comparison of bacterial communities on limnic versus coastal marine particles reveals profound differences in colonization. *Environmental microbiology*, doi: 10.1111/1462-2920.12466.
- Björriksen, P.K. (1988) Phytoplankton exudation of organic matter: Why do healthy cells do it? 1. *Limnology and Oceanography*, doi: 10.4319/lo.1988.33.1.0151.
- Bjursell, M.K., Martens, E.C., and 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. *The Journal of biological chemistry*, doi: 10.1074/jbc.M606509200.
- Blanvillain, S., Meyer, D., Boulanger, A., Lautier, M., Guynet, C., Denancé, N., et al. (2007) Plant carbohydrate scavenging through tonB-dependent receptors: a feature shared by phytopathogenic and aquatic bacteria. *PLoS one*, doi: 10.1371/journal.pone.0000224.
- Bochdansky, A.B., Clouse, M.A., and Herndl, G.J. (2017) Eukaryotic microbes, principally fungi and labyrinthulomycetes, dominate biomass on bathypelagic marine snow. *The ISME journal*, doi: 10.1038/ismej.2016.113.
- Bolam, D.N., and Koropatkin, N.M. (2012) Glycan recognition by the Bacteroidetes Sus-like systems. *Current opinion in structural biology*, doi: 10.1016/j.sbi.2012.06.006.
- Bowler, C., Allen, A.E., Badger, J.H., Grimwood, J., Jabbari, K., Kuo, A., et al. (2008) The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature*, doi: 10.1038/nature07410.
- Boyd, P.W. (2013) Framing biological responses to a changing ocean. *Nature Climate Change*, doi: 10.1038/nclimate1881.
- Boyd, P.W., Claustre, H., Levy, M., Siegel, D.A., and Weber, T. (2019) Multi-faceted particle pumps drive carbon sequestration in the ocean. *Nature*, doi: 10.1038/s41586-019-1098-2.
- Brennan, G., and Collins, S. (2015) Growth responses of a green alga to multiple environmental drivers. *Nature Climate Change*, doi: 10.1038/nclimate2682.
- Bridoux, M.C., Neibauer, J., Ingalls, A.E., Nunn, B.L., and Keil, R.G. (2015) Suspended marine particulate proteins in coastal and oligotrophic waters. *Journal of Marine Systems*, doi: 10.1016/j.jmarsys.2014.10.014.
- Brown, M.V., Lauro, F.M., Demaere, M.Z., Muir, L., Wilkins, D., Thomas, T., et al. (2012) Global biogeography of SAR11 marine bacteria. *Molecular Systems Biology*, doi: 10.1038/msb.2012.28.
- Brum, J.R., Ignacio-Espinoza, J.C., Roux, S., Doucier, G., Acinas, S.G., Alberti, A., et al. (2015) Ocean plankton. Patterns and ecological drivers of ocean viral communities. *Science*, doi: 10.1126/science.1261498.
- Brúna, T., Hoff, K.J., Lomsadze, A., Stanke, M., and Borodovsky, M. (2021) BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. *NAR genomics and bioinformatics*, doi: 10.1093/nargab/lqaa108.
- Buchan, A., LeCleir, G.R., Gulvik, C.A., and González, J.M. (2014) Master recyclers: features and functions of bacteria associated with phytoplankton blooms. *Nature reviews. Microbiology*, doi: 10.1038/nrmicro3326.
- Buesseler, K.O., Boyd, P.W., Black, E.E., and Siegel, D.A. (2020) Metrics that matter for assessing the ocean biological carbon pump. *Proceedings of the National Academy of Sciences*, doi: 10.1073/pnas.1918114117.
- Buist, G., Steen, A., Kok, J., and Kuipers, O.P. (2008) LysM, a widely distributed protein motif for binding to (peptido)glycans. *Molecular microbiology*, doi: 10.1111/j.1365-2958.2008.06211.x.

References

- Burke, C., Steinberg, P., Rusch, D., Kjelleberg, S., and Thomas, T. (2011a) Bacterial community assembly based on functional genes rather than species. *Proceedings of the National Academy of Sciences*, doi: 10.1073/pnas.11015911108.
- Burke, C., Thomas, T., Lewis, M., Steinberg, P., and Kjelleberg, S. (2011b) Composition, uniqueness and variability of the epiphytic bacterial community of the green alga *Ulva australis*. *The ISME journal*, doi: 10.1038/ismej.2010.164.
- Burroughs, A.M., Kaur, G., Zhang, D., and Aravind, L. (2017) Novel clades of the HU/IHF superfamily point to unexpected roles in the eukaryotic centrosome, chromosome partitioning, and biologic conflicts. *Cell cycle (Georgetown, Tex.)*, doi: 10.1080/15384101.2017.1315494.
- Buttigieg, P.L., Fadeev, E., Bienhold, C., Hehemann, L., Offre, P., and Boetius, A. (2018) Marine microbes in 4D-using time series observation to assess the dynamics of the ocean microbiome and its links to ocean health. *Current opinion in microbiology*, doi: 10.1016/j.mib.2018.01.015.
- Campbell, B.J., and Kirchman, D.L. (2013) Bacterial diversity, community structure and potential growth rates along an estuarine salinity gradient. *The ISME journal*, doi: 10.1038/ismej.2012.93.
- Carlucci, A.F., and Silbernagel, S.B. (1969) EFFECT OF VITAMIN CONCENTRATIONS ON GROWTH AND DEVELOPMENT OF VITAMIN-REQUIRING ALGAE(1). *Journal of Phycology*, doi: 10.1111/j.1529-8817.1969.tb02578.x.
- Caron, D.A., Davis, P.G., Madin, L.P., and Sieburth, J.M. (1982) Heterotrophic bacteria and bacterivorous protozoa in oceanic macroaggregates. *Science (New York, N.Y.)*, doi: 10.1126/science.218.4574.795.
- Carvalho, C.C.C.R. de (2007) Biofilms: recent developments on an old battle. *Recent patents on biotechnology*, doi: 10.2174/187220807779813965.
- Cavicchioli, R., Ripple, W.J., Timmis, K.N., Azam, F., Bakken, L.R., Baylis, M., et al. (2019) Scientists' warning to humanity: microorganisms and climate change. *Nature Reviews Microbiology*, doi: 10.1038/s41579-019-0222-5.
- Chafee, M., Fernández-Guerra, A., Buttigieg, P.L., Gerds, G., Eren, A.M., Teeling, H., and Amann, R.L. (2018) Recurrent patterns of microdiversity in a temperate coastal marine environment. *The ISME journal*, doi: 10.1038/ismej.2017.165.
- Chiancone, E., Ceci, P., Ilari, A., Ribacchi, F., and Stefanini, S. (2004) Iron and proteins for iron storage and detoxification. *Biometals : an international journal on the role of metal ions in biology, biochemistry, and medicine*, doi: 10.1023/b:biom.0000027692.24395.76.
- Chourey, K., Jansson, J., VerBerkmoes, N., Shah, M., Chavarria, K.L., Tom, L.M., Brodie, E.L., and Hettich, R.L. (2010) Direct cellular lysis/protein extraction protocol for soil metaproteomics. *Journal of proteome research*, doi: 10.1021/pr100787q.
- Christie-Oleza, J.A., and Armengaud, J. (2015) Proteomics of the Roseobacter clade, a window to the marine microbiology landscape. *Proteomics*, doi: 10.1002/pmic.201500222.
- Christie-Oleza, J.A., Armengaud, J., Guerin, P., and Scanlan, D.J. (2015) Functional distinctness in the exoproteomes of marine *Synechococcus*. *Environmental microbiology*, doi: 10.1111/1462-2920.12822.
- Christie-Oleza, J.A., Sousoni, D., Lloyd, M., Armengaud, J., and Scanlan, D.J. (2017) Nutrient recycling facilitates long-term stability of marine microbial phototroph-heterotroph interactions. *Nature microbiology*, doi: 10.1038/nmicrobiol.2017.100.
- Clayton, M.N. (1992) Propagules of marine macroalgae: Structure and development. *British Phycological Journal*, doi: 10.1080/00071619200650231.
- Coale, K.H., Fitzwater, S.E., Gordon, R.M., Johnson, K.S., and Barber, R.T. (1996) Control of community growth and export production by upwelled iron in the equatorial Pacific Ocean. *Nature*, doi: 10.1038/379621a0.
- Cohen, N.R., McIlvin, M.R., Moran, D.M., Held, N.A., Saunders, J.K., Hawco, N.J., et al. (2021) Dinoflagellates alter their carbon and nutrient metabolic strategies across environmental gradients in the central Pacific Ocean. *Nature microbiology*, doi: 10.1038/s41564-020-00814-7.
- Cornejo-Castillo, F.M., Cabello, A.M., Salazar, G., Sánchez-Baracaldo, P., Lima-Mendez, G., Hingamp, P., et al. (2016) Cyanobacterial symbionts diverged in the late Cretaceous towards lineage-specific nitrogen fixation factories in single-celled phytoplankton. *Nature communications*, doi: 10.1038/ncomms11071.
- Cosse, A., Leblanc, C., and Potin, P. (2007) Dynamic Defense of Marine Macroalgae Against Pathogens: From Early Activated to Gene-Regulated Responses: Elsevier, pp. 221–266.
- Cotter, P.D., Hill, C., and Ross, R.P. (2005) Bacteriocins: developing innate immunity for food. *Nature reviews. Microbiology*, doi: 10.1038/nrmicro1273.
- Cottrell, M.T., and 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. *Applied and Environmental Microbiology*, doi: 10.1128/aem.66.4.1692-1697.2000.

References

- Coutinho, P.M., Stam, M., Blanc, E., and Henrissat, B. (2003) Why are there so many carbohydrate-active enzyme-related genes in plants? *Trends in plant science*, doi: 10.1016/j.tplants.2003.10.002.
- Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature biotechnology*, doi: 10.1038/nbt.1511.
- Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J.V., and Mann, M. (2011) Andromeda: a peptide search engine integrated into the MaxQuant environment. *Journal of proteome research*, doi: 10.1021/pr101065j.
- Crenn, K., Duffieux, D., and Jeanthon, C. (2018) Bacterial Epibiotic Communities of Ubiquitous and Abundant Marine Diatoms Are Distinct in Short- and Long-Term Associations. *Frontiers in microbiology*, doi: 10.3389/fmicb.2018.02879.
- Crespo, B.G., Pommier, T., Fernández-Gómez, B., and Pedrós-Alió, C. (2013) Taxonomic composition of the particle-attached and free-living bacterial assemblages in the Northwest Mediterranean Sea analyzed by pyrosequencing of the 16S rRNA. *MicrobiologyOpen*, doi: 10.1002/mbo3.92.
- Croft, M.T., Lawrence, A.D., Raux-Deery, E., Warren, M.J., and Smith, A.G. (2005) Algae acquire vitamin B12 through a symbiotic relationship with bacteria. *Nature*, doi: 10.1038/nature04056.
- Crump, B.C., Armbrust, E.V., and Baross, J.A. (1999) Phylogenetic analysis of particle-attached and free-living bacterial communities in the Columbia river, its estuary, and the adjacent coastal ocean. *Applied and Environmental Microbiology*, doi: 10.1128/AEM.65.7.3192-3204.1999.
- Crump, B.C., and Koch, E.W. (2008) Attached bacterial populations shared by four species of aquatic angiosperms. *Applied and Environmental Microbiology*, doi: 10.1128/AEM.00952-08.
- Cunliffe, M., Hollingsworth, A., Bain, C., Sharma, V., and Taylor, J.D. (2017) Algal polysaccharide utilisation by saprotrophic planktonic marine fungi. *Fungal Ecology*, doi: 10.1016/j.funeco.2017.08.009.
- Cuskin, F., Lowe, E.C., Temple, M.J., Zhu, Y., Cameron, E., Pudlo, N.A., *et al.* (2015) Human gut Bacteroidetes can utilize yeast mannan through a selfish mechanism. *Nature*, doi: 10.1038/nature13995.
- Dadon-Pilosof, A., Conley, K.R., Jacobi, Y., Haber, M., Lombard, F., Sutherland, K.R., *et al.* (2017) Surface properties of SAR11 bacteria facilitate grazing avoidance. *Nature microbiology*, doi: 10.1038/s41564-017-0030-5.
- D'Ambrosio, L., Zierovogel, K., MacGregor, B., Teske, A., and Arnosti, C. (2014) Composition and enzymatic function of particle-associated and free-living bacteria: a coastal/offshore comparison. *The ISME journal*, doi: 10.1038/ismej.2014.67.
- Davis, T.A., Volesky, B., and Mucci, A. (2003) A review of the biochemistry of heavy metal biosorption by brown algae. *Water Research*, doi: 10.1016/S0043-1354(03)00293-8.
- DeLong, E.F., Franks, D.G., and Alldredge, A.L. (1993) Phylogenetic diversity of aggregate-attached vs. free-living marine bacterial assemblages. *Limnology and Oceanography*, doi: 10.4319/lo.1993.38.5.0924.
- Desriac, F., Defer, D., Bourgougnon, N., Brillet, B., Le Chevalier, P., and Fleury, Y. (2010) Bacteriocin as weapons in the marine animal-associated bacteria warfare: inventory and potential applications as an aquaculture probiotic. *Marine drugs*, doi: 10.3390/md8041153.
- Dimitrieva, G.Y., Crawford, R.L., and Yüksel, G.U. (2006) The nature of plant growth-promoting effects of a pseudoalteromonad associated with the marine algae *Laminaria japonica* and linked to catalase excretion. *Journal of applied microbiology*, doi: 10.1111/j.1365-2672.2006.02831.x.
- Dore, J.E., Lukas, R., Sadler, D.W., Church, M.J., and Karl, D.M. (2009) Physical and biogeochemical modulation of ocean acidification in the central North Pacific. *Proceedings of the National Academy of Sciences*, doi: 10.1073/pnas.0906044106.
- Ducklow, H.W., and Carlson, C.A. (1992) Oceanic Bacterial Production. In *Advances in Microbial Ecology*. Marshall, K.C. (ed). Boston, MA: Springer US, pp. 113–181.
- Durham, B.P., Sharma, S., Luo, H., Smith, C.B., Amin, S.A., Bender, S.J., *et al.* (2015) Cryptic carbon and sulfur cycling between surface ocean plankton. *Proceedings of the National Academy of Sciences*, doi: 10.1073/pnas.1413137112.
- Egan, S., Fernandes, N.D., Kumar, V., Gardiner, M., and Thomas, T. (2014) Bacterial pathogens, virulence mechanism and host defence in marine macroalgae. *Environmental microbiology*, doi: 10.1111/1462-2920.12288.
- Eilers, H., Pernthaler, J., Peplies, J., Glöckner, F.O., Gerdt, G., and Amann, R. (2001) Isolation of Novel Pelagic Bacteria from the German Bight and Their Seasonal Contributions to Surface Picoplankton. *Applied and Environmental Microbiology*, doi: 10.1128/AEM.67.11.5134-5142.2001.
- Eloe, E.A., Shulse, C.N., Fadrosch, D.W., Williamson, S.J., Allen, E.E., and Bartlett, D.H. (2011) Compositional differences in particle-associated and free-living microbial assemblages from an

- extreme deep-ocean environment. *Environmental microbiology reports*, doi: 10.1111/j.1758-2229.2010.00223.x.
- Ferguson, R.L., Buckley, E.N., and Palumbo, A.V. (1984) Response of marine bacterioplankton to differential filtration and confinement. *Applied and Environmental Microbiology*, **47**: 49–55.
- Fernández-Gómez, B., Richter, M., Schüller, M., Pinhassi, J., Acinas, S.G., González, J.M., and Pedrós-Alió, C. (2013) Ecology of marine Bacteroidetes: a comparative genomics approach. *The ISME journal*, doi: 10.1038/ismej.2012.169.
- Ficko-Blean, E., Hervé, C., and Michel, G. (2015) Sweet and sour sugars from the sea: the biosynthesis and remodeling of sulfated cell wall polysaccharides from marine macroalgae. *Perspectives in Phycology*, doi: 10.1127/pip/2015/0028.
- Ficko-Blean, E., Préchoux, A., Thomas, F., Rochat, T., Larocque, R., Zhu, Y., et al. (2017) Carrageenan catabolism is encoded by a complex regulon in marine heterotrophic bacteria. *Nature communications*, doi: 10.1038/s41467-017-01832-6.
- Ficner, R., and Huber, R. (1993) Refined crystal structure of phycoerythrin from *Porphyridium cruentum* at 0.23-nm resolution and localization of the gamma subunit. *European journal of biochemistry*, doi: 10.1111/j.1432-1033.1993.tb18356.x.
- Flemming, H.-C., and Wingender, J. (2010) The biofilm matrix. *Nature Reviews Microbiology*, doi: 10.1038/nrmicro2415.
- Flemming, H.-C., and Wuertz, S. (2019) Bacteria and archaea on Earth and their abundance in biofilms. *Nature reviews. Microbiology*, doi: 10.1038/s41579-019-0158-9.
- Foley, M.H., Cockburn, D.W., and Koropatkin, N.M. (2016) The Sus operon: a model system for starch uptake by the human gut Bacteroidetes. *Cellular and molecular life sciences : CMLS*, doi: 10.1007/s00018-016-2242-x.
- Fontanez, K.M., Eppley, J.M., Samo, T.J., Karl, D.M., and DeLong, E.F. (2015) Microbial community structure and function on sinking particles in the North Pacific Subtropical Gyre. *Frontiers in microbiology*, doi: 10.3389/fmicb.2015.00469.
- Francis, T.B., Bartosik, D., Sura, T., Sichert, A., Hehemann, J.-H., Markert, S., et al. (2021) Changing expression patterns of TonB-dependent transporters suggest shifts in polysaccharide consumption over the course of a spring phytoplankton bloom. *The ISME journal*, doi: 10.1038/s41396-021-00928-8.
- Frias-Lopez, J., Zerkle, A.L., Bonheyo, G.T., and Fouke, B.W. (2002) Partitioning of bacterial communities between seawater and healthy, black band diseased, and dead coral surfaces. *Applied and Environmental Microbiology*, doi: 10.1128/aem.68.5.2214-2228.2002.
- Frost, B.W., Landry, M.R., and Hassett, R.P. (1983) Feeding behavior of large calanoid copepods *Neocalanus cristatus* and *N. plumchrus* from the subarctic Pacific Ocean. *Deep Sea Research Part A. Oceanographic Research Papers*, doi: 10.1016/0198-0149(83)90029-8.
- Fuhrman, J.A., Cram, J.A., and Needham, D.M. (2015) Marine microbial community dynamics and their ecological interpretation. *Nature reviews. Microbiology*, doi: 10.1038/nrmicro3417.
- Fukami, K., Nishijima, T., and Ishida, Y. (1997) Stimulative and inhibitory effects of bacteria on the growth of microalgae. *Hydrobiologia*, doi: 10.1023/A:1003139402315.
- Fukunaga, Y., Kurahashi, M., Sakiyama, Y., Ohuchi, M., Yokota, A., and Harayama, S. (2009) *Phycisphaera mikurensis* gen. nov., sp. nov., isolated from a marine alga, and proposal of *Phycisphaeraeaceae* fam. nov., *Phycisphaerales* ord. nov. and *Phycisphaerae* classis nov. in the phylum Planctomycetes. *The Journal of general and applied microbiology*, doi: 10.2323/jgam.55.267.
- Ganesh, S., Parris, D.J., DeLong, E.F., and Stewart, F.J. (2014) Metagenomic analysis of size-fractionated picoplankton in a marine oxygen minimum zone. *The ISME journal*, doi: 10.1038/ismej.2013.144.
- Gao, K., Xu, J., Gao, G., Li, Y., Hutchins, D.A., Huang, B., et al. (2012) Rising CO₂ and increased light exposure synergistically reduce marine primary productivity. *Nature Climate Change*, doi: 10.1038/nclimate1507.
- Gärdes, A., Iversen, M.H., Grossart, H.-P., Passow, U., and Ullrich, M.S. (2011) Diatom-associated bacteria are required for aggregation of *Thalassiosira weissflogii*. *The ISME journal*, doi: 10.1038/ismej.2010.145.
- Garron, M.-L., and Cygler, M. (2014) Uronic polysaccharide degrading enzymes. *Current opinion in structural biology*, doi: 10.1016/j.sbi.2014.07.012.
- Gaskell, A., Crennell, S., and Taylor, G. (1995) The three domains of a bacterial sialidase: a beta-propeller, an immunoglobulin module and a galactose-binding jelly-roll. *Structure (London, England : 1993)*, doi: 10.1016/s0969-2126(01)00255-6.

References

- Gasol, J.M., Pinhassi, J., Alonso-Sáez, L., Ducklow, H., Herndl, G.J., Koblížek, M., *et al.* (2008) Towards a better understanding of microbial carbon flux in the sea*. *Aquatic Microbial Ecology*, doi: 10.3354/ame01230.
- Gaudry, A., Zeroual, S., Gaie-Levrel, F., Moskura, M., Boujrhah, F.-Z., El Moursli, R.C., *et al.* (2007) Heavy Metals Pollution of the Atlantic Marine Environment by the Moroccan Phosphate Industry, as Observed through their Bioaccumulation in *Ulva Lactuca*. *Water, Air, and Soil Pollution*, doi: 10.1007/s11270-006-9196-9.
- Georges, A.A., El-Swais, H., Craig, S.E., Li, W.K.W., and Walsh, D.A. (2014) Metaproteomic analysis of a winter to spring succession in coastal northwest Atlantic Ocean microbial plankton. *The ISME journal*, doi: 10.1038/ismej.2013.234.
- Ghiglione, J.F., Mevel, G., Pujo-Pay, M., Mousseau, L., Lebaron, P., and Goutx, M. (2007) Diel and seasonal variations in abundance, activity, and community structure of particle-attached and free-living bacteria in NW Mediterranean Sea. *Microbial ecology*, doi: 10.1007/s00248-006-9189-7.
- Gilbert, J.A., Steele, J.A., Caporaso, J.G., Steinbrück, L., Reeder, J., Temperton, B., *et al.* (2012) Defining seasonal marine microbial community dynamics. *The ISME journal*, doi: 10.1038/ismej.2011.107.
- Giovannoni, S.J., Tripp, H.J., Givan, S., Podar, M., Vergin, K.L., Baptista, D., *et al.* (2005) Genome streamlining in a cosmopolitan oceanic bacterium. *Science*, doi: 10.1126/science.1114057.
- Glenwright, A.J., Pothula, K.R., Bhamidimarri, S.P., Chorev, D.S., Baslé, A., Firbank, S.J., *et al.* (2017) Structural basis for nutrient acquisition by dominant members of the human gut microbiota. *Nature*, doi: 10.1038/nature20828.
- Gómez-Pereira, P.R., Fuchs, B.M., Alonso, C., Oliver, M.J., van Beusekom, J.E.E., and Amann, R. (2010) Distinct flavobacterial communities in contrasting water masses of the north Atlantic Ocean. *The ISME journal*, doi: 10.1038/ismej.2009.142.
- Gómez-Pereira, P.R., Schüller, M., Fuchs, B.M., Bennke, C., Teeling, H., Waldmann, J., *et al.* (2012) Genomic content of uncultured Bacteroidetes from contrasting oceanic provinces in the North Atlantic Ocean. *Environmental microbiology*, doi: 10.1111/j.1462-2920.2011.02555.x.
- González, J.M., Fernández-Gómez, B., Fernández-Guerra, A., Gómez-Consarnau, L., Sánchez, O., Coll-Lladó, M., *et al.* (2008) Genome analysis of the proteorhodopsin-containing marine bacterium *Polaribacter* sp. MED152 (Flavobacteria). *Proceedings of the National Academy of Sciences*, doi: 10.1073/pnas.0712027105.
- González, J.M., Pinhassi, J., Fernández-Gómez, B., Coll-Lladó, M., González-Velázquez, M., Puigbò, P., *et al.* (2011) Genomics of the proteorhodopsin-containing marine flavobacterium *Dokdonia* sp. strain MED134. *Applied and Environmental Microbiology*, doi: 10.1128/AEM.06152-11.
- González, J.M., Simó, R., Massana, R., Covert, J.S., Casamayor, E.O., Pedrós-Alió, C., and Moran, M.A. (2000) Bacterial community structure associated with a dimethylsulfoniopropionate-producing North Atlantic algal bloom. *Applied and Environmental Microbiology*, doi: 10.1128/AEM.66.10.4237-4246.2000.
- Gresock, M.G., Kestead, K.A., and Postle, K. (2015) From Homodimer to Heterodimer and Back: Elucidating the TonB Energy Transduction Cycle. *Journal of bacteriology*, doi: 10.1128/JB.00484-15.
- Grondin, J.M., Tamura, K., Déjean, G., Abbott, D.W., and Brumer, H. (2017) Polysaccharide Utilization Loci: Fueling Microbial Communities. *Journal of bacteriology*, doi: 10.1128/JB.00860-16.
- Grossart, H.P. (1999) Interactions between marine bacteria and axenic diatoms (*Cylindrotheca fusiformis*, *Nitzschia laevis*, and *Thalassiosira weissflogii*) incubated under various conditions in the lab. *Aquatic Microbial Ecology*, doi: 10.3354/ame019001.
- Grossart, H.P., Hietanen, S., and Ploug, H. (2003) Microbial dynamics on diatom aggregates in Øresund, Denmark. *Marine Ecology Progress Series*, doi: 10.3354/meps249069.
- Grossart, H.P., Riemann, L., and Azam, F. (2001) Bacterial motility in the sea and its ecological implications. *Aquatic Microbial Ecology*, doi: 10.3354/ame025247.
- Grossart, H.-P. (2010) Ecological consequences of bacterioplankton lifestyles: changes in concepts are needed. *Environmental microbiology reports*, doi: 10.1111/j.1758-2229.2010.00179.x.
- Grossart, H.-P., Levold, F., Allgaier, M., Simon, M., and Brinkhoff, T. (2005) Marine diatom species harbour distinct bacterial communities. *Environmental microbiology*, doi: 10.1111/j.1462-2920.2005.00759.x.
- Grossart, H.-P., Tang, K.W., Kiørboe, T., and Ploug, H. (2007) Comparison of cell-specific activity between free-living and attached bacteria using isolates and natural assemblages. *FEMS microbiology letters*, doi: 10.1111/j.1574-6968.2006.00520.x.
- Gügi, B., Le Costaouec, T., Burel, C., Lerouge, P., Helbert, W., and Bardor, M. (2015) Diatom-Specific Oligosaccharide and Polysaccharide Structures Help to Unravel Biosynthetic Capabilities in Diatoms. *Marine drugs*, doi: 10.3390/md13095993.

References

- Guidi, L., Chaffron, S., Bittner, L., Eveillard, D., Larhlimi, A., Roux, S., *et al.* (2016) Plankton networks driving carbon export in the oligotrophic ocean. *Nature*, doi: 10.1038/nature16942.
- Hagström, Å., Sundh, J., Bunse, C., Müller-Karulis, B., Osbeck, C.M.G., Jarone Pinhassi, and Zweifel, U.L. (2020) *Ecological significance of microbial transporters in the marine environment*: Unpublished.
- Hahnke, R.L., Bennke, C.M., Fuchs, B.M., Mann, A.J., Rhiel, E., Teeling, H., Amann, R., and Harder, J. (2015) Dilution cultivation of marine heterotrophic bacteria abundant after a spring phytoplankton bloom in the North Sea. *Environmental microbiology*, doi: 10.1111/1462-2920.12479.
- Hahnke, R.L., and Harder, J. (2013) Phylogenetic diversity of Flavobacteria isolated from the North Sea on solid media. *Systematic and applied microbiology*, doi: 10.1016/j.syapm.2013.06.006.
- Hall, E.K., Besemer, K., Kohl, L., Preiler, C., Riedel, K., Schneider, T., Wanek, W., and Battin, T.J. (2012) Effects of resource chemistry on the composition and function of stream hyporheic biofilms. *Frontiers in microbiology*, doi: 10.3389/fmicb.2012.00035.
- Hallegraeff GM (ed.) (2003) *Manual on harmful marine microalgae*. Paris: Unesco Publ.
- Hammer, P.E., Burd, W., Hill, D.S., Ligon, J.M., and van Pée, K. (1999) Conservation of the pyrrolnitrin biosynthetic gene cluster among six pyrrolnitrin-producing strains. *FEMS microbiology letters*, doi: 10.1111/j.1574-6968.1999.tb08775.x.
- Hassett, B.T., and Gradinger, R. (2016) Chytrids dominate arctic marine fungal communities. *Environmental microbiology*, doi: 10.1111/1462-2920.13216.
- Haug, A., Larsen, B., Smidsrød, O., Møller, J., Brunvoll, J., Bunnenberg, E., Djerassi, C., and Records, R. (1966) A Study of the Constitution of Alginic Acid by Partial Acid Hydrolysis. *Acta Chemica Scandinavica*, doi: 10.3891/acta.chem.scand.20-0183.
- Haug, A., and Myklestad, S. (1976) Polysaccharides of marine diatoms with special reference to Chaetoceros species. *Marine Biology*, doi: 10.1007/BF00388798.
- Hehemann, J.-H., Boraston, A.B., and Czjzek, M. (2014) A sweet new wave: structures and mechanisms of enzymes that digest polysaccharides from marine algae. *Current opinion in structural biology*, doi: 10.1016/j.sbi.2014.07.009.
- Hehemann, J.-H., Correc, G., Barbeyron, T., Helbert, W., Czjzek, M., and Michel, G. (2010) Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature*, doi: 10.1038/nature08937.
- Hehemann, J.-H., Kelly, A.G., Pudlo, N.A., Martens, E.C., and Boraston, A.B. (2012) Bacteria of the human gut microbiome catabolize red seaweed glycans with carbohydrate-active enzyme updates from extrinsic microbes. *Proceedings of the National Academy of Sciences of the United States of America*, doi: 10.1073/pnas.1211002109.
- Hehemann, J.-H., van Truong, Unfried, F., Welsch, N., Kabisch, J., Heiden, S.E., *et al.* (2017) Aquatic adaptation of a laterally acquired pectin degradation pathway in marine gammaproteobacteria. *Environmental microbiology*, doi: 10.1111/1462-2920.13726.
- Heins, A., Reintjes, G., Amann, R.I., and Harder, J. (2021) Particle Collection in Imhoff Sedimentation Cones Enriches Both Motile Chemotactic and Particle-Attached Bacteria. *Frontiers in microbiology*, doi: 10.3389/fmicb.2021.643730.
- Helbert, W. (2017) Marine Polysaccharide Sulfatases. *Frontiers in Marine Science*, doi: 10.3389/fmars.2017.00006.
- Hemsworth, G.R., Déjean, G., Davies, G.J., and Brumer, H. (2016) Learning from microbial strategies for polysaccharide degradation. *Biochemical Society transactions*, doi: 10.1042/BST20150180.
- Henrissat, B., and Coutinho, P.M. (2001) Classification of glycoside hydrolases and glycosyltransferases from hyperthermophiles. In *Hyperthermophilic enzymes. Part A*. Adams, M.W.W., and Kelly, R.M. (eds) . San Diego, Calif: Academic Press, pp. 183–201.
- Hinzke, T., Kouris, A., Hughes, R.-A., Strous, M., and Kleiner, M. (2019) More Is Not Always Better: Evaluation of 1D and 2D-LC-MS/MS Methods for Metaproteomics. *Frontiers in microbiology*, doi: 10.3389/fmicb.2019.00238.
- Hoff, K.J. (2009) The effect of sequencing errors on metagenomic gene prediction. *BMC genomics*, doi: 10.1186/1471-2164-10-520.
- Hoff, K.J., Lange, S., Lomsadze, A., Borodovsky, M., and Stanke, M. (2016) BRAKER1: Unsupervised RNA-Seq-Based Genome Annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics (Oxford, England)*, doi: 10.1093/bioinformatics/btv661.
- Hoff, K.J., and Stanke, M. (2013) WebAUGUSTUS--a web service for training AUGUSTUS and predicting genes in eukaryotes. *Nucleic acids research*, doi: 10.1093/nar/gkt418.
- Hofmann, E., Wrench, P.M., Sharples, F.P., Hiller, R.G., Welte, W., and Diederichs, K. (1996) Structural basis of light harvesting by carotenoids: peridinin-chlorophyll-protein from Amphidinium carterae. *Science (New York, N.Y.)*, doi: 10.1126/science.272.5269.1788.

References

- Hollants, J., Leliaert, F., Clerck, O. de, and Willems, A. (2013) What we can learn from sushi: a review on seaweed-bacterial associations. *FEMS Microbiology Ecology*, doi: 10.1111/j.1574-6941.2012.01446.x.
- Hooper, S.D., Mavromatis, K., and Kyrpides, N.C. (2009) Microbial co-habitation and lateral gene transfer: what transposases can tell us. *Genome biology*, doi: 10.1186/gb-2009-10-4-r45.
- Hossain, M.M., and Tsuyumu, S. (2006) Flagella-mediated motility is required for biofilm formation by *Erwinia carotovora* subsp. *carotovora*. *Journal of General Plant Pathology*, doi: 10.1007/s10327-005-0246-8.
- Houry, A., Briandet, R., Aymerich, S., and Gohar, M. (2010) Involvement of motility and flagella in *Bacillus cereus* biofilm formation. *Microbiology (Reading, England)*, doi: 10.1099/mic.0.034827-0.
- Hünken, M., Harder, J., and Kirst, G.O. (2008) Epiphytic bacteria on the Antarctic ice diatom *Amphiprora kufferathii* Manguin cleave hydrogen peroxide produced during algal photosynthesis. *Plant biology (Stuttgart, Germany)*, doi: 10.1111/j.1438-8677.2008.00040.x.
- Hunt, D.E., David, L.A., Gevers, D., Preheim, S.P., Alm, E.J., and Polz, M.F. (2008) Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science*, doi: 10.1126/science.1157890.
- Hurd, C.L., Lenton, A., Tilbrook, B., and Boyd, P.W. (2018) Current understanding and challenges for oceans in a higher-CO₂ world. *Nature Climate Change*, doi: 10.1038/s41558-018-0211-0.
- Hutchins, D.A., and Boyd, P.W. (2016) Marine phytoplankton and the changing ocean iron cycle. *Nature Climate Change*, doi: 10.1038/nclimate3147.
- Hutchins, D.A., and Fu, F. (2017) Microorganisms and ocean global change. *Nature microbiology*, doi: 10.1038/nmicrobiol.2017.58.
- Hyatt, D., Chen, G.-L., Locascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC bioinformatics*, doi: 10.1186/1471-2105-11-119.
- Imhoff, K. (1989) *Karl Imhoffs Handbook of urban drainage and wastewater disposal*. New York: Wiley.
- Ivanova, A.O., and Dedysh, S.N. (2006) High abundance of planctomycetes in anoxic layers of a Sphagnum peat bog. *Mikrobiologiya*, **75**: 823–827.
- Iversen, M.H., and Ploug, H. (2010) Ballast minerals and the sinking carbon flux in the ocean: carbon-specific respiration rates and sinking velocity of marine snow aggregates. *Biogeosciences*, doi: 10.5194/bg-7-2613-2010.
- Janse, I., Rijssel, M., Hall, P.-J., Gerwig, G.J., Gottschal, J.C., and Prins, R.A. (1996) The Storage Glucan of *Phaeocystis Globosa* (Prymnesiophyceae) Cells. *Journal of Phycology*, doi: 10.1111/j.0022-3646.1996.00382.x.
- Jesus Raposo, M.F. de, Morais, A.M.M.B. de, and Morais, R.M.S.C. de (2014) Bioactivity and Applications of Polysaccharides from Marine Microalgae. *Ramawat K, Mérillon J, editors. Polysaccharides, Springer, Switzerland*, doi: 10.1007/978-3-319-03751-6_47-1.
- Jones, M.D.M., Forn, I., Gadelha, C., Egan, M.J., Bass, D., Massana, R., and Richards, T.A. (2011) Discovery of novel intermediate forms redefines the fungal tree of life. *Nature*, doi: 10.1038/nature09984.
- Jordan, L.D., Zhou, Y., Smallwood, C.R., Lill, Y., Ritchie, K., Yip, W.T., Newton, S.M., and Klebba, P.E. (2013) Energy-dependent motion of TonB in the Gram-negative bacterial inner membrane. *Proceedings of the National Academy of Sciences*, doi: 10.1073/pnas.1304243110.
- Kabisch, A., Otto, A., König, S., Becher, D., Albrecht, D., Schüler, M., et al. (2014) Functional characterization of polysaccharide utilization loci in the marine Bacteroidetes 'Gramella forsetii' KT0803. *The ISME journal*, doi: 10.1038/ismej.2014.4.
- Kaczmarek, I., Ehrman, J.M., Bates, S.S., Green, D.H., Léger, C., and Harris, J. (2005) Diversity and distribution of epibiotic bacteria on *Pseudo-nitzschia multiseries* (Bacillariophyceae) in culture, and comparison with those on diatoms in native seawater. *Harmful Algae*, doi: 10.1016/j.hal.2004.10.001.
- Kapoor, D., Singh, S., Kumar, V., Romero, R., Prasad, R., and Singh, J. (2019) Antioxidant enzymes regulation in plants in reference to reactive oxygen species (ROS) and reactive nitrogen species (RNS). *Plant Gene*, doi: 10.1016/j.plgene.2019.100182.
- Kappellmann, L., Krüger, K., Hehemann, J.-H., Harder, J., Markert, S., Unfried, F., et al. (2018) Polysaccharide utilization loci of North Sea Flavobacteriia as basis for using SusC/D-protein expression for predicting major phytoplankton glycans. *The ISME journal*, doi: 10.1038/s41396-018-0242-6.
- Kazamia, E., Czesnick, H., van Nguyen, T.T., Croft, M.T., Sherwood, E., Sasso, S., et al. (2012) Mutualistic interactions between vitamin B₁₂-dependent algae and heterotrophic bacteria exhibit regulation. *Environmental microbiology*, doi: 10.1111/j.1462-2920.2012.02733.x.

- Keiblinger, K.M., and Riedel, K. (2018) Sample Preparation for Metaproteome Analyses of Soil and Leaf Litter. *Methods in molecular biology (Clifton, N.J.)*, doi: 10.1007/978-1-4939-8695-8_21.
- Keiblinger, K.M., Wilhartitz, I.C., Schneider, T., Roschitzki, B., Schmid, E., Eberl, L., Riedel, K., and Zechmeister-Boltenstern, S. (2012) Soil metaproteomics - Comparative evaluation of protein extraction protocols. *Soil biology & biochemistry*, doi: 10.1016/j.soilbio.2012.05.014.
- Keith, S.C., and Arnosti, C. (2001) Extracellular enzyme activity in a river-bay-shelf transect: variations in polysaccharide hydrolysis rates with substrate and size class. *Aquatic Microbial Ecology*, doi: 10.3354/ame024243.
- Keller, A., Nesvizhskii, A.I., Kolker, E., and Aebersold, R. (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Analytical chemistry*, doi: 10.1021/ac025747h.
- Kirchman, D.L. (2002) The ecology of Cytophaga-Flavobacteria in aquatic environments. *FEMS Microbiology Ecology*, doi: 10.1111/j.1574-6941.2002.tb00910.x.
- Kita, D., Shibata, S., Kikuchi, Y., Kokubu, E., Nakayama, K., Saito, A., and Ishihara, K. (2016) Involvement of the Type IX Secretion System in Capnocytophaga ochracea Gliding Motility and Biofilm Formation. *Applied and Environmental Microbiology*, doi: 10.1128/AEM.03452-15.
- Klindworth, A., Mann, A.J., Huang, S., Wichels, A., Quast, C., Waldmann, J., Teeling, H., and Glöckner, F.O. (2014) Diversity and activity of marine bacterioplankton during a diatom bloom in the North Sea assessed by total RNA and pyrotag sequencing. *Marine genomics*, doi: 10.1016/j.margen.2014.08.007.
- Koch, H., Dürwald, A., Schweder, T., Noriega-Ortega, B., Vidal-Melgosa, S., Hehemann, J.-H., et al. (2019) Biphasic cellular adaptations and ecological implications of *Alteromonas macleodii* degrading a mixture of algal polysaccharides. *The ISME journal*, doi: 10.1038/s41396-018-0252-4.
- Kodama, M., Doucette, G.J., and Green, D.H. (2006) Relationships Between Bacteria and Harmful Algae. In *Ecology of Harmful Algae*. Granéli, E., and Turner, J.T. (eds) . Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg, pp. 243–255.
- Koropatkin, N.M., Cameron, E.A., and Martens, E.C. (2012) How glycan metabolism shapes the human gut microbiota. *Nature Reviews Microbiology*, doi: 10.1038/nrmicro2746.
- Koskinen, V.R., Emery, P.A., Creasy, D.M., and Cottrell, J.S. (2011) Hierarchical clustering of shotgun proteomics data. *Molecular & cellular proteomics : MCP*, doi: 10.1074/mcp.M110.003822.
- Kraan, S. (2012) Algal Polysaccharides, Novel Applications and Outlook. In *Carbohydrates - Comprehensive Studies on Glycobiology and Glycotechnology*. Chang, C.-F. (ed): InTech.
- Krewulak, K.D., and Vogel, H.J. (2011) TonB or not TonB: is that the question? *Biochemistry and cell biology = Biochimie et biologie cellulaire*, doi: 10.1139/o10-141.
- Kroth, P.G., Chiovitti, A., Gruber, A., Martin-Jezequel, V., Mock, T., Parker, M.S., et al. (2008) A model for carbohydrate metabolism in the diatom *Phaeodactylum tricornutum* deduced from comparative whole genome analysis. *PLoS one*, doi: 10.1371/journal.pone.0001426.
- Krüger, K., Chafee, M., Ben Francis, T., Del Glavina Rio, T., Becher, D., Schweder, T., Amann, R.I., and Teeling, H. (2019) In marine Bacteroidetes the bulk of glycan degradation during algae blooms is mediated by few clades using a restricted set of genes. *The ISME journal*, doi: 10.1038/s41396-019-0476-y.
- Kuhn, R., Benndorf, D., Rapp, E., Reichl, U., Palese, L.L., and Pollice, A. (2011) Metaproteome analysis of sewage sludge from membrane bioreactors. *Proteomics*, doi: 10.1002/pmic.201000590.
- Kulasekara, H.D., Ventre, I., Kulasekara, B.R., Lazdunski, A., Filloux, A., and Lory, S. (2005) A novel two-component system controls the expression of *Pseudomonas aeruginosa* fimbrial cup genes. *Molecular microbiology*, doi: 10.1111/j.1365-2958.2004.04402.x.
- Kuroda, T., and Tsuchiya, T. (2009) Multidrug efflux transporters in the MATE family. *Biochimica et biophysica acta*, doi: 10.1016/j.bbapap.2008.11.012.
- Labourel, A., Jam, M., Legentil, L., Sylla, B., Hehemann, J.H., Ferrières, V., Czjzek, M., and Michel, G. (2015) Structural and biochemical characterization of the laminarinase ZgLamCGH16 from *Zobellia galactanivorans* suggests preferred recognition of branched laminarin. *Acta crystallographica. Section D, Biological crystallography*, doi: 10.1107/S139900471402450X.
- Lachnit, T., Meske, D., Wahl, M., Harder, T., and Schmitz, R. (2011) Epibacterial community patterns on marine macroalgae are host-specific but temporally variable. *Environmental microbiology*, doi: 10.1111/j.1462-2920.2010.02371.x.
- Laemmli, U.K. (1970) Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature*: 680–685.
- Lage, O.M. (2013) Characterization of a planctomycete associated with the marine dinoflagellate *Prorocentrum micans* Her. *Antonie van Leeuwenhoek*, doi: 10.1007/s10482-013-9991-4.
- Lage, O.M., and Bondoso, J. (2011) Planctomycetes diversity associated with macroalgae. *FEMS Microbiology Ecology*, doi: 10.1111/j.1574-6941.2011.01168.x.

- Lage, O.M., and Bondoso, J. (2014) Planctomycetes and macroalgae, a striking association. *Frontiers in microbiology*, doi: 10.3389/fmicb.2014.00267.
- Lairson, L.L., Henrissat, B., Davies, G.J., and Withers, S.G. (2008) Glycosyltransferases: structures, functions, and mechanisms. *Annual review of biochemistry*, doi: 10.1146/annurev.biochem.76.061005.092322.
- Landa, M., Burns, A.S., Durham, B.P., Esson, K., Nowinski, B., Sharma, S., et al. (2019) Sulfur metabolites that facilitate oceanic phytoplankton-bacteria carbon flux. *The ISME journal*, doi: 10.1038/s41396-019-0455-3.
- Landa, M., Burns, A.S., Roth, S.J., and Moran, M.A. (2017) Bacterial transcriptome remodeling during sequential co-culture with a marine dinoflagellate and diatom. *The ISME journal*, doi: 10.1038/ismej.2017.117.
- Lasica, A.M., Ksiazek, M., Madej, M., and Potempa, J. (2017) The Type IX Secretion System (T9SS): Highlights and Recent Insights into Its Structure and Function. *Frontiers in cellular and infection microbiology*, doi: 10.3389/fcimb.2017.00215.
- Lassek, C., Burghartz, M., Chaves-Moreno, D., Otto, A., Hentschker, C., Fuchs, S., et al. (2015) A metaproteomics approach to elucidate host and pathogen protein expression during catheter-associated urinary tract infections (CAUTIs). *Molecular & cellular proteomics : MCP*, doi: 10.1074/mcp.M114.043463.
- Le Costaouëc, T., Unamunzaga, C., Mantecon, L., and Helbert, W. (2017) New structural insights into the cell-wall polysaccharide of the diatom *Phaeodactylum tricornutum*. *Algal Research*, doi: 10.1016/j.algal.2017.07.021.
- Leary, D.H., Li, R.W., Hamdan, L.J., Hervey, W.J., Lebedev, N., Wang, Z., et al. (2014) Integrated metagenomic and metaproteomic analyses of marine biofilm communities. *Biofouling*, doi: 10.1080/08927014.2014.977267.
- Lehtreck, K.F., and Melkonian, M. (1998) SF-assemblin, striated fibers, and segmented coiled coil proteins. *Cell motility and the cytoskeleton*, doi: 10.1002/(SICI)1097-0169(1998)41:4<289::AID-CM2>3.0.CO;2-1.
- Lee, V.T., Matewish, J.M., Kessler, J.L., Hyodo, M., Hayakawa, Y., and Lory, S. (2007) A cyclic-di-GMP receptor required for bacterial exopolysaccharide production. *Molecular microbiology*, doi: 10.1111/j.1365-2958.2007.05879.x.
- Lemon, K.P., Higgins, D.E., and Kolter, R. (2007) Flagellar motility is critical for *Listeria monocytogenes* biofilm formation. *Journal of bacteriology*, doi: 10.1128/JB.01967-06.
- Li, D., Liu, C.-M., Luo, R., Sadakane, K., and Lam, T.-W. (2015) MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics (Oxford, England)*, doi: 10.1093/bioinformatics/btv033.
- Lim, B.L., Yeung, P., Cheng, C., and Hill, J.E. (2007) Distribution and diversity of phytate-mineralizing bacteria. *The ISME journal*, doi: 10.1038/ismej.2007.40.
- Lima-Mendez, G., Faust, K., Henry, N., Decelle, J., Colin, S., Carcillo, F., et al. (2015) Ocean plankton. Determinants of community structure in the global plankton interactome. *Science*, doi: 10.1126/science.1262073.
- Liu, D., Keesing, J.K., He, P., Wang, Z., Shi, Y., and Wang, Y. (2013) The world's largest macroalgal bloom in the Yellow Sea, China: Formation and implications. *Estuarine, Coastal and Shelf Science*, doi: 10.1016/j.ecss.2013.05.021.
- Liu, G., Wu, S., Jin, W., and Sun, C. (2016) Amy63, a novel type of marine bacterial multifunctional enzyme possessing amylase, agarase and carrageenase activities. *Scientific reports*, doi: 10.1038/srep18726.
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic acids research*, doi: 10.1093/nar/gkt1178.
- López-Pérez, M., Kimes, N.E., Haro-Moreno, J.M., and Rodriguez-Valera, F. (2016) Not All Particles Are Equal: The Selective Enrichment of Particle-Associated Bacteria from the Mediterranean Sea. *Frontiers in microbiology*, doi: 10.3389/fmicb.2016.00996.
- Lountos, G.T., Tropea, J.E., and Waugh, D.S. (2012) Structure of the cytoplasmic domain of *Yersinia pestis* YscD, an essential component of the type III secretion system. *Acta crystallographica. Section D, Biological crystallography*, doi: 10.1107/S0907444911054308.
- Lu, S., Wang, J., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R., et al. (2020) CDD/SPARCLE: the conserved domain database in 2020. *Nucleic acids research*, doi: 10.1093/nar/gkz991.
- Lucas, J., Wichels, A., Teeling, H., Chafee, M., Scharfe, M., and Gerdtts, G. (2015) Annual dynamics of North Sea bacterioplankton: seasonal variability superimposes short-term variation. *FEMS Microbiology Ecology*, doi: 10.1093/femsec/fiv099.

References

- Lynch, J.B., and Sonnenburg, J.L. (2012) Prioritization of a plant polysaccharide over a mucus carbohydrate is enforced by a *Bacteroides* hybrid two-component system. *Molecular microbiology*, doi: 10.1111/j.1365-2958.2012.08123.x.
- Mann, A.J., Hahnke, R.L., Huang, S., Werner, J., Xing, P., Barbeyron, T., *et al.* (2013) The genome of the alga-associated marine flavobacterium *Formosa agariphila* KMM 3901T reveals a broad potential for degradation of algal polysaccharides. *Applied and Environmental Microbiology*, doi: 10.1128/AEM.01937-13.
- Martens, E.C., Koropatkin, N.M., Smith, T.J., and Gordon, J.I. (2009) Complex glycan catabolism by the human gut microbiota: the *Bacteroidetes* Sus-like paradigm. *The Journal of biological chemistry*, doi: 10.1074/jbc.R109.022848.
- Martens, E.C., Lowe, E.C., Chiang, H., Pudlo, N.A., Wu, M., McNulty, N.P., *et al.* (2011) Recognition and degradation of plant cell wall polysaccharides by two human gut symbionts. *PLoS biology*, doi: 10.1371/journal.pbio.1001221.
- Martin, J.H., and Michael Gordon, R. (1988) Northeast Pacific iron distributions in relation to phytoplankton productivity. *Deep Sea Research Part A. Oceanographic Research Papers*, doi: 10.1016/0198-0149(88)90035-0.
- Martin, M., Barbeyron, T., Martin, R., Portetelle, D., Michel, G., and Vandenbol, M. (2015) The Cultivable Surface Microbiota of the Brown Alga *Ascophyllum nodosum* is Enriched in Macroalgal-Polysaccharide-Degrading Bacteria. *Frontiers in microbiology*, doi: 10.3389/fmicb.2015.01487.
- Martinez-Garcia, M., Brazel, D.M., Swan, B.K., Arnosti, C., Chain, P.S.G., Reitenga, K.G., *et al.* (2012) Capturing single cell genomes of active polysaccharide degraders: an unexpected contribution of Verrucomicrobia. *PloS one*, doi: 10.1371/journal.pone.0035314.
- Mattick, J.S. (2002) Type IV pili and twitching motility. *Annual review of microbiology*, doi: 10.1146/annurev.micro.56.012302.160938.
- McBride, M.J., and Zhu, Y. (2013) Gliding motility and Por secretion system genes are widespread among members of the phylum bacteroidetes. *Journal of bacteriology*, doi: 10.1128/JB.01962-12.
- Mestre, M., Ferrera, I., Borrull, E., Ortega-Retuerta, E., Mbedi, S., Grossart, H.-P., Gasol, J.M., and Sala, M.M. (2017) Spatial variability of marine bacterial and archaeal communities along the particulate matter continuum. *Molecular ecology*, doi: 10.1111/mec.14421.
- Mestre, M., Höfer, J., Sala, M.M., and Gasol, J.M. (2020) Seasonal Variation of Bacterial Diversity Along the Marine Particulate Matter Continuum. *Frontiers in microbiology*, doi: 10.3389/fmicb.2020.01590.
- Meusch, D., Gatsogiannis, C., Efremov, R.G., Lang, A.E., Hofnagel, O., Vetter, I.R., Aktories, K., and Raunser, S. (2014) Mechanism of Tc toxin action revealed in molecular detail. *Nature*, doi: 10.1038/nature13015.
- Michiels, T., Vanooteghem, J.C., Lambert de Rouvroit, C., China, B., Gustin, A., Boudry, P., and Cornelis, G.R. (1991) Analysis of *virC*, an operon involved in the secretion of Yop proteins by *Yersinia enterocolitica*. *Journal of bacteriology*, doi: 10.1128/jb.173.16.4994-5009.1991.
- Miller, T.R., Hnilicka, K., Dziedzic, A., Desplats, P., and Belas, R. (2004) Chemotaxis of *Silicibacter* sp. strain TM1040 toward dinoflagellate products. *Applied and Environmental Microbiology*, doi: 10.1128/AEM.70.8.4692-4701.2004.
- Mincer, T.J., and Aicher, A.C. (2016) Methanol Production by a Broad Phylogenetic Array of Marine Phytoplankton. *PloS one*, doi: 10.1371/journal.pone.0150820.
- Mirus, O., Strauss, S., Nicolaisen, K., Haeseler, A. von, and Schleiff, E. (2009) TonB-dependent transporters and their occurrence in cyanobacteria. *BMC biology*, doi: 10.1186/1741-7007-7-68.
- Mitchell, J.G., Pearson, L., Bonazinga, A., Dillon, S., Khouri, H., and Paxinos, R. (1995) Long lag times and high velocities in the motility of natural assemblages of marine bacteria. *Applied and Environmental Microbiology*, doi: 10.1128/AEM.61.3.877-882.1995.
- Moog, G. (2012) Optimierung der Probenaufarbeitung für Metaproteomanalysen mittels Massenspektrometrie. *Diplomarbeit an der Universität Greifswald*.
- Moore, E.K., Harvey, H.R., Faux, J.F., Goodlett, and Nunn, B.L. (2014) Protein recycling in Bering Sea algal incubations. *Marine Ecology Progress Series*, doi: 10.3354/meps10936.
- Moore, E.K., Nunn, B.L., Goodlett, D.R., and Harvey, H.R. (2012) Identifying and tracking proteins through the marine water column: insights into the inputs and preservation mechanisms of protein in sediments. *Geochimica et cosmochimica acta*, doi: 10.1016/j.gca.2012.01.002.
- Moran, M.A., Belas, R., Schell, M.A., González, J.M., Sun, F., Sun, S., *et al.* (2007) Ecological genomics of marine Roseobacters. *Applied and Environmental Microbiology*, doi: 10.1128/AEM.02580-06.
- Morris, J.J., Johnson, Z.I., Szul, M.J., Keller, M., and Zinser, E.R. (2011) Dependence of the cyanobacterium *Prochlorococcus* on hydrogen peroxide scavenging microbes for growth at the ocean's surface. *PloS one*, doi: 10.1371/journal.pone.0016805.

References

- Morris, J.J., Kirkegaard, R., Szul, M.J., Johnson, Z.I., and Zinser, E.R. (2008) Facilitation of robust growth of *Prochlorococcus* colonies and dilute liquid cultures by "helper" heterotrophic bacteria. *Applied and Environmental Microbiology*, doi: 10.1128/AEM.02479-07.
- Morris, R.M., Longnecker, K., and Giovannoni, S.J. (2006) *Pirellula* and OM43 are among the dominant lineages identified in an Oregon coast diatom bloom. *Environmental microbiology*, doi: 10.1111/j.1462-2920.2006.01029.x.
- Morris, R.M., Nunn, B.L., Frazar, C., Goodlett, D.R., Ting, Y.S., and Rocap, G. (2010) Comparative metaproteomics reveals ocean-scale shifts in microbial nutrient utilization and energy transduction. *The ISME journal*, doi: 10.1038/ismej.2010.4.
- Mühlenbruch, M., Grossart, H.-P., Eigemann, F., and Voss, M. (2018) Mini-review: Phytoplankton-derived polysaccharides in the marine environment and their interactions with heterotrophic bacteria. *Environmental microbiology*, doi: 10.1111/1462-2920.14302.
- Mullaney, E.J., Daly, C.B., and Ullah, A.H. (2000) Advances in phytase research. In *Advances in phytase research*. Mullaney, E.J., Daly, C.B., and Ullah, A.H. (eds) : Elsevier, pp. 157–199.
- Munn, C.B., Marchant, H.K., and Moody, A.J. (2008) Defences against oxidative stress in vibrios associated with corals. *FEMS microbiology letters*, doi: 10.1111/j.1574-6968.2008.01073.x.
- Munoz Aguilar, J.M., Ashby, A.M., Richards, A.J.M., Loake, G.J., Watson, M.D., and Shaw, C.H. (1988) Chemotaxis of *Rhizobium leguminosarum* biovar *phaseoli* towards Flavonoid Inducers of the Symbiotic Nodulation Genes. *Microbiology*, doi: 10.1099/00221287-134-10-2741.
- Muth, T., Benndorf, D., Reichl, U., Rapp, E., and Martens, L. (2013) Searching for a needle in a stack of needles: challenges in metaproteomics data analysis. *Molecular bioSystems*, doi: 10.1039/c2mb25415h.
- Myklestad, S. (1974) Production of carbohydrates by marine planktonic diatoms. I. Comparison of nine different species in culture. *Journal of Experimental Marine Biology and Ecology*, doi: 10.1016/0022-0981(74)90049-5.
- Nahnsen, S., Bielow, C., Reinert, K., and Kohlbacher, O. (2013) Tools for label-free peptide quantification. *Molecular & cellular proteomics : MCP*, doi: 10.1074/mcp.r112.025163.
- Nakayama, K. (2015) *Porphyromonas gingivalis* and related bacteria: from colonial pigmentation to the type IX secretion system and gliding motility. *Journal of periodontal research*, doi: 10.1111/jre.12255.
- Namba, A., Mano, N., Takano, H., Beppu, T., Ueda, K., and Hirose, H. (2008) OmpA is an adhesion factor of *Aeromonas veronii*, an optimistic pathogen that habituates in carp intestinal tract. *Journal of applied microbiology*, doi: 10.1111/j.1365-2672.2008.03883.x.
- Nedashkovskaya, O.I., Kim, S.B., Han, S.K., Lysenko, A.M., Rohde, M., Zhukova, N.V., et al. (2003) *Mesonia* algae gen. nov., sp. nov., a novel marine bacterium of the family Flavobacteriaceae isolated from the green alga *Acrosiphonia sonderi* (Kütz) Kornm. *International Journal of Systematic and Evolutionary Microbiology*, doi: 10.1099/ijs.0.02626-0.
- Nedashkovskaya, O.I., Kim, S.B., Han, S.K., Rhee, M.S., Lysenko, A.M., Falsen, E., et al. (2004a) *Ulvibacter litoralis* gen. nov., sp. nov., a novel member of the family Flavobacteriaceae isolated from the green alga *Ulva fenestrata*. *International Journal of Systematic and Evolutionary Microbiology*, doi: 10.1099/ijs.0.02757-0.
- Nedashkovskaya, O.I., Kim, S.B., Han, S.K., Rhee, M.-S., Lysenko, A.M., Rohde, M., et al. (2004b) *Algibacter lectus* gen. nov., sp. nov., a novel member of the family Flavobacteriaceae isolated from green algae. *International Journal of Systematic and Evolutionary Microbiology*, doi: 10.1099/ijs.0.02949-0.
- Nedashkovskaya, O.I., Kim, S.B., Lee, K.H., Bae, K.S., Frolova, G.M., Mikhailov, V.V., and Kim, I.S. (2005) *Pibocella ponti* gen. nov., sp. nov., a novel marine bacterium of the family Flavobacteriaceae isolated from the green alga *Acrosiphonia sonderi*. *International Journal of Systematic and Evolutionary Microbiology*, doi: 10.1099/ijs.0.63251-0.
- Nedashkovskaya, O.I., Kim, S.B., Vancanneyt, M., Snauwaert, C., Lysenko, A.M., Rohde, M., et al. (2006) *Formosa agariphila* sp. nov., a budding bacterium of the family Flavobacteriaceae isolated from marine environments, and emended description of the genus *Formosa*. *International Journal of Systematic and Evolutionary Microbiology*, doi: 10.1099/ijs.0.63875-0.
- Nedashkovskaya, O.I., Kukhlevskiy, A.D., and Zhukova, N.V. (2013) *Polaribacter reichenbachii* sp. nov.: a new marine bacterium associated with the green alga *Ulva fenestrata*. *Current microbiology*, doi: 10.1007/s00284-012-0200-x.
- Needham, D.M., Fichot, E.B., Wang, E., Berdjeb, L., Cram, J.A., Fichot, C.G., and Fuhrman, J.A. (2018) Dynamics and interactions of highly resolved marine plankton via automated high-frequency sampling. *The ISME journal*, doi: 10.1038/s41396-018-0169-y.
- Needham, D.M., and Fuhrman, J.A. (2016) Pronounced daily succession of phytoplankton, archaea and bacteria following a spring bloom. *Nature microbiology*, doi: 10.1038/nmicrobiol.2016.5.

References

- Nesvizhskii, A.I., and Aebersold, R. (2005) Interpretation of shotgun proteomic data: the protein inference problem. *Molecular & cellular proteomics : MCP*, doi: 10.1074/mcp.R500012-MCP200.
- Nesvizhskii, A.I., Keller, A., Kolker, E., and Aebersold, R. (2003) A statistical model for identifying proteins by tandem mass spectrometry. *Analytical chemistry*, doi: 10.1021/ac0341261.
- Neumann, A.M., Balmonte, J.P., Berger, M., Giebel, H.-A., Arnosti, C., Voget, S., *et al.* (2015) Different utilization of alginate and other algal polysaccharides by marine *Alteromonas macleodii* ecotypes. *Environmental microbiology*, doi: 10.1111/1462-2920.12862.
- Newton, R.J., Griffin, L.E., Bowles, K.M., Meile, C., Gifford, S., Givens, C.E., *et al.* (2010) Genome characteristics of a generalist marine bacterial lineage. *The ISME journal*, doi: 10.1038/ismej.2009.150.
- Nichols, C.A.M., Guezennec, J., and Bowman, J.P. (2005) Bacterial exopolysaccharides from extreme marine environments with special consideration of the southern ocean, sea ice, and deep-sea hydrothermal vents: a review. *Marine biotechnology (New York, N.Y.)*, doi: 10.1007/s10126-004-5118-2.
- Nigrelli, L., and Thines, M. (2013) Tropical oomycetes in the German Bight – Climate warming or overlooked diversity? *Fungal Ecology*, doi: 10.1016/j.funeco.2012.11.003.
- Noinaj, N., Guillier, M., Barnard, T.J., and Buchanan, S.K. (2010) TonB-dependent transporters: regulation, structure, and function. *Annual review of microbiology*, doi: 10.1146/annurev.micro.112408.134247.
- Nurk, S., Meleshko, D., Korobeynikov, A., and Pevzner, P.A. (2017) metaSPAdes: a new versatile metagenomic assembler. *Genome research*, doi: 10.1101/gr.213959.116.
- Nyström, T., and Neidhardt, F.C. (1994) Expression and role of the universal stress protein, UspA, of *Escherichia coli* during growth arrest. *Molecular microbiology*, doi: 10.1111/j.1365-2958.1994.tb00334.x.
- Ochman, H., Lawrence, J.G., and Groisman, E.A. (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature*, doi: 10.1038/35012500.
- Ortega, M.A., Hao, Y., Zhang, Q., Walker, M.C., van der Donk, W.A., and Nair, S.K. (2015) Structure and mechanism of the tRNA-dependent lantibiotic dehydratase NisB. *Nature*, doi: 10.1038/nature13888.
- Ortega-Retuerta, E., Joux, F., Jeffrey, W.H., and Ghiglione, J.F. (2013) Spatial variability of particle-attached and free-living bacterial diversity in surface waters from the Mackenzie River to the Beaufort Sea (Canadian Arctic). *Biogeosciences*, doi: 10.5194/bg-10-2747-2013.
- O'Toole, G.A., and Kolter, R. (1998) Flagellar and twitching motility are necessary for *Pseudomonas aeruginosa* biofilm development. *Molecular microbiology*, doi: 10.1046/j.1365-2958.1998.01062.x.
- Padilla, C.C., Ganesh, S., Gantt, S., Huhman, A., Parris, D.J., Sarode, N., and Stewart, F.J. (2015) Standard filtration practices may significantly distort planktonic microbial diversity estimates. *Frontiers in microbiology*, doi: 10.3389/Fmicb.2015.00547.
- Park, B.S., Joo, J.-H., Baek, K.-D., and Han, M.-S. (2016) A mutualistic interaction between the bacterium *Pseudomonas asplenii* and the harmful algal species *Chattonella marina* (Raphidophyceae). *Harmful Algae*, doi: 10.1016/j.hal.2016.04.006.
- Parsons, T.R., Stephens, K., and Strickland, J.D.H. (1961) On the Chemical Composition of Eleven Species of Marine Phytoplankters. *Journal of the Fisheries Research Board of Canada*, doi: 10.1139/f61-063.
- Passow, U. (2002) Transparent exopolymer particles (TEP) in aquatic environments. *Progress in Oceanography*, doi: 10.1016/S0079-6611(02)00138-6.
- Percival, E. (1979) The polysaccharides of green, red and brown seaweeds: Their basic structure, biosynthesis and function. *British Phycological Journal*, doi: 10.1080/00071617900650121.
- Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D.J., *et al.* (2019) The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic acids research*, doi: 10.1093/nar/gky1106.
- Petroni, G., Spring, S., Schleifer, K.H., Verni, F., and Rosati, G. (2000) Defensive extrusive ectosymbionts of *Euplotidium* (Ciliophora) that contain microtubule-like structures are bacteria related to Verrucomicrobia. *Proceedings of the National Academy of Sciences of the United States of America*, doi: 10.1073/pnas.030438197.
- Pfennig, N., and Trüper, H.G. (1981) Isolation of Members of the Families Chromatiaceae and Chlorobiaceae. In *The Prokaryotes*. Starr, M.P., Stolp, H., Trüper, H.G., Balows, A., and Schlegel, H.G. (eds) . Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 279–289.
- Piekarowicz, A., Klyz, A., Kwiatek, A., and Stein, D.C. (2001) Analysis of type I restriction modification systems in the Neisseriaceae: genetic organization and properties of the gene products. *Molecular microbiology*, doi: 10.1046/j.1365-2958.2001.02587.x.

References

- Pinhassi, J., Sala, M.M., Havskum, H., Peters, F., Guadayol, O., Malits, A., and Marrasé, C. (2004) Changes in bacterioplankton composition under different phytoplankton regimens. *Applied and Environmental Microbiology*, doi: 10.1128/AEM.70.11.6753-6766.2004.
- Pitcher, R.S., Brittain, T., and Watmough, N.J. (2002) Cytochrome cbb(3) oxidase and bacterial microaerobic metabolism. *Biochemical Society transactions*, doi: 10.1042/bst0300653.
- Popper, Z.A., Michel, G., Hervé, C., Domozych, D.S., Willats, W.G.T., Tuohy, M.G., Kloareg, B., and Stengel, D.B. (2011) Evolution and diversity of plant cell walls: from algae to flowering plants. *Annual review of plant biology*, doi: 10.1146/annurev-arplant-042110-103809.
- Popper, Z.A., and Tuohy, M.G. (2010) Beyond the green: understanding the evolutionary puzzle of plant and algal cell walls. *Plant physiology*, doi: 10.1104/pp.110.158055.
- Prakash, B., Veeregowda, B., and Krishnappa, G. (2003) Biofilms: A survival strategy of bacteria. *Current Science*: 1299–1307.
- Priest, T., Fuchs, B., Amann, R., and Reich, M. (2021) Diversity and biomass dynamics of unicellular marine fungi during a spring phytoplankton bloom. *Environmental microbiology*, doi: 10.1111/1462-2920.15331.
- Prieto, M.L., O'Sullivan, L., Tan, S.P., McLoughlin, P., Hughes, H., O'Connor, P.M., et al. (2012) Assessment of the bacteriocinogenic potential of marine bacteria reveals lichenicidin production by seaweed-derived *Bacillus* spp. *Marine drugs*, doi: 10.3390/md10102280.
- Püttker, S., Kohrs, F., Benndorf, D., Heyer, R., Rapp, E., and Reichl, U. (2015) Metaproteomics of activated sludge from a wastewater treatment plant - a pilot study. *Proteomics*, doi: 10.1002/pmic.201400559.
- Ram, R.J., VerBerkmoes, N.C., Thelen, M.P., Tyson, G.W., Baker, B.J., Blake, R.C., et al. (2005) Community proteomics of a natural microbial biofilm. *Science (New York, N.Y.)*, doi: 10.1126/science.
- Rath, J., Wu, K.Y., Herndl, G.J., and DeLong, E.F. (1998) High phylogenetic diversity in a marine-snow-associated bacterial assemblage. *Aquatic Microbial Ecology*, doi: 10.3354/ame014261.
- Rehman, Z.U., Wang, Y., Moradali, M.F., Hay, I.D., and Rehm, B.H.A. (2013) Insights into the assembly of the alginate biosynthesis machinery in *Pseudomonas aeruginosa*. *Applied and Environmental Microbiology*, doi: 10.1128/AEM.00460-13.
- Reintjes, G., Arnosti, C., Fuchs, B.M., and Amann, R. (2017) An alternative polysaccharide uptake mechanism of marine bacteria. *The ISME journal*, doi: 10.1038/ismej.2017.26.
- Rhoads, A., and Au, K.F. (2015) PacBio Sequencing and Its Applications. *Genomics, proteomics & bioinformatics*, doi: 10.1016/j.gpb.2015.08.002.
- Riebesell, U., and Gattuso, J.-P. (2015) Lessons learned from ocean acidification research. *Nature Climate Change*, doi: 10.1038/nclimate2456.
- Riemann, L., Steward, G.F., and Azam, F. (2000) Dynamics of bacterial community composition and activity during a mesocosm diatom bloom. *Applied and Environmental Microbiology*, doi: 10.1128/aem.66.2.578-587.2000.
- Riesenfeld, C.S., Schloss, P.D., and Handelsman, J. (2004) Metagenomics: genomic analysis of microbial communities. *Annual review of genetics*, doi: 10.1146/annurev.genet.38.072902.091216.
- Rinta-Kanto, J.M., Sun, S., Sharma, S., Kiene, R.P., and Moran, M.A. (2012) Bacterial community transcription patterns during a marine phytoplankton bloom. *Environmental microbiology*, doi: 10.1111/j.1462-2920.2011.02602.x.
- Rintoul, S.R., Chown, S.L., DeConto, R.M., England, M.H., Fricker, H.A., Masson-Delmotte, V., et al. (2018) Choosing the future of Antarctica. *Nature*, doi: 10.1038/s41586-018-0173-4.
- Rio, D.C., Ares, M., Hannon, G.J., and Nilsen, T.W. (2010) Purification of RNA using TRIzol (TRI reagent). *Cold Spring Harbor protocols*, doi: 10.1101/pdb.prot5439.
- Roche-Mayzaud, O., and Mayzaud, P. (1987) Purification of endo- and exolaminarin and partial characterization of the exoacting form from the copepod *Acartia Clausi*. *Comp. Biochem. Physiol.*: 105–110.
- Rohwer, F., Seguritan, V., Azam, F., and Knowlton, N. (2002) Diversity and distribution of coral-associated bacteria. *Marine Ecology Progress Series*, doi: 10.3354/MEPS243001.
- Romanenko, L.A., Uchino, M., Frolova, G.M., and Mikhailov, V.V. (2007) *Marixanthomonas ophiurae* gen. nov., sp. nov., a marine bacterium of the family Flavobacteriaceae isolated from a deep-sea brittle star. *International Journal of Systematic and Evolutionary Microbiology*, doi: 10.1099/ijs.0.64662-0.
- Römmling, U. (2002) Molecular biology of cellulose production in bacteria. *Research in Microbiology*, doi: 10.1016/S0923-2508(02)01316-5.
- Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F (eds.) (2014) *The prokaryotes. Gammaproteobacteria*. Berlin, Heidelberg: Springer.

References

- Rusch, D.B., Halpern, A.L., Sutton, G., Heidelberg, K.B., Williamson, S., Yooseph, S., *et al.* (2007) The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS biology*, doi: 10.1371/journal.pbio.0050077.
- Ryjenkov, D.A., Tarutina, M., Moskvina, O.V., and Gomelsky, M. (2005) Cyclic diguanylate is a ubiquitous signaling molecule in bacteria: insights into biochemistry of the GGDEF protein domain. *Journal of bacteriology*, doi: 10.1128/JB.187.5.1792-1798.2005.
- Ryther, J.H., and Guillard, R.R.L. (1962) Studies of marine planktonic diatoms: II. a use of *Cyclotella* Nana Hustedt for assays of Vitamin B12 in sea water. *Canadian journal of microbiology*, doi: 10.1139/m62-057.
- Saba, V.S., Friedrichs, M.A.M., Carr, M.-E., Antoine, D., Armstrong, R.A., Asanuma, I., *et al.* (2010) Challenges of modeling depth-integrated marine primary productivity over multiple decades: A case study at BATS and HOT. *Global Biogeochemical Cycles*, doi: 10.1029/2009GB003655.
- Saito, M.A., Bertrand, E.M., Duffy, M.E., Gaylord, D.A., Held, N.A., Hervey, W.J., *et al.* (2019) Progress and Challenges in Ocean Metaproteomics and Proposed Best Practices for Data Sharing. *Journal of proteome research*, doi: 10.1021/acs.jproteome.8b00761.
- Saito, M.A., McIlvin, M.R., Moran, D.M., Goepfert, T.J., DiTullio, G.R., Post, A.F., and Lamborg, C.H. (2014) Multiple nutrient stresses at intersecting Pacific Ocean biomes detected by protein biomarkers. *Science*, doi: 10.1126/science.1256450.
- Salerno, C., Benndorf, D., Kluge, S., Palese, L.L., Reichl, U., and Pollice, A. (2016) Metaproteomics Applied to Activated Sludge for Industrial Wastewater Treatment Revealed a Dominant Methylophilic Metabolism of *Hyphomicrobium zavarzinii*. *Microbial ecology*, doi: 10.1007/s00248-016-0769-x.
- Sapp, M., Schwaderer, A.S., Wiltshire, K.H., Hoppe, H.-G., Gerdt, G., and Wichels, A. (2007a) Species-specific bacterial communities in the phycosphere of microalgae? *Microbial ecology*, doi: 10.1007/s00248-006-9162-5.
- Sapp, M., Wichels, A., and Gerdt, G. (2007b) Impacts of cultivation of marine diatoms on the associated bacterial community. *Applied and Environmental Microbiology*, doi: 10.1128/AEM.02274-06.
- Sapp, M., Wichels, A., Wiltshire, K.H., and Gerdt, G. (2007c) Bacterial community dynamics during the winter-spring transition in the North Sea. *FEMS Microbiology Ecology*, doi: 10.1111/j.1574-6941.2006.00238.x.
- Sarmiento, J.L. (2013) *Ocean Biogeochemical Dynamics*. Princeton: Princeton University Press.
- Satinsky, B.M., Crump, B.C., Smith, C.B., Sharma, S., Zielinski, B.L., Doherty, M., *et al.* (2014) Microspatial gene expression patterns in the Amazon River Plume. *Proceedings of the National Academy of Sciences*, doi: 10.1073/pnas.1402782111.
- Sato, K., Naito, M., Yukitake, H., Hirakawa, H., Shoji, M., McBride, M.J., Rhodes, R.G., and Nakayama, K. (2010) A protein secretion system linked to bacteroidete gliding motility and pathogenesis. *Proceedings of the National Academy of Sciences*, doi: 10.1073/pnas.0912010107.
- Savage, D.C. (1977) Microbial ecology of the gastrointestinal tract. *Annual review of microbiology*, doi: 10.1146/annurev.mi.31.100177.000543.
- Schada von Borzyskowski, L., Severi, F., Krüger, K., Hermann, L., Gilardet, A., Sippel, F., *et al.* (2019) Marine Proteobacteria metabolize glycolate via the β -hydroxyaspartate cycle. *Nature*, doi: 10.1038/s41586-019-1748-4.
- Schäfer, H., Abbas, B., Witte, H., and Muyzer, G. (2002) Genetic diversity of 'satellite' bacteria present in cultures of marine diatoms. *FEMS Microbiology Ecology*, doi: 10.1111/j.1574-6941.2002.tb00992.x.
- Schauer, K., Rodionov, D.A., and Reuse, H. de (2008) New substrates for TonB-dependent transport: do we only see the 'tip of the iceberg'? *Trends in biochemical sciences*, doi: 10.1016/j.tibs.2008.04.012.
- Schiebenhoefer, H., Schallert, K., Renard, B.Y., Trappe, K., Schmid, E., Benndorf, D., *et al.* (2020) A complete and flexible workflow for metaproteomics data analysis based on MetaProteomeAnalyzer and Prophan. *Nature protocols*, doi: 10.1038/s41596-020-0368-7.
- Schloss, P.D., and Handelsman, J. (2006) Toward a census of bacteria in soil. *PLoS computational biology*, doi: 10.1371/journal.pcbi.0020092.
- Schlünzen, F., Zarivach, R., Harms, J., Bashan, A., Tocilj, A., Albrecht, R., Yonath, A., and Franceschi, F. (2001) Structural basis for the interaction of antibiotics with the peptidyl transferase centre in eubacteria. *Nature*, doi: 10.1038/35101544.
- Schneider, T., Keiblinger, K.M., Schmid, E., Sterflinger-Gleixner, K., Ellersdorfer, G., Roschitzki, B., *et al.* (2012) Who is who in litter decomposition? Metaproteomics reveals major microbial players and their biogeochemical functions. *The ISME journal*, doi: 10.1038/ismej.2012.11.

References

- Schneider, T., and Riedel, K. (2010) Environmental proteomics: analysis of structure and function of microbial communities. *Proteomics*, doi: 10.1002/pmic.200900450.
- Schneider, T., Schmid, E., Castro, J.V. de, Cardinale, M., Eberl, L., Grube, M., Berg, G., and Riedel, K. (2011) Structure and function of the symbiosis partners of the lung lichen (*Lobaria pulmonaria* L. Hoffm.) analyzed by metaproteomics. *Proteomics*, doi: 10.1002/pmic.201000679.
- Schoemann, V., Becquevort, S., Stefels, J., Rousseau, V., and Lancelot, C. (2005) Phaeocystis blooms in the global ocean and their controlling mechanisms: a review. *Journal of Sea Research*, doi: 10.1016/j.seares.2004.01.008.
- Schultz, D., Zühlke, D., Bernhardt, J., Francis, T.B., Albrecht, D., Hirschfeld, C., Markert, S., and Riedel, K. (2020) An optimized metaproteomics protocol for a holistic taxonomic and functional characterization of microbial communities from marine particles. *Environmental microbiology reports*, doi: 10.1111/1758-2229.12842.
- Schwahnhauser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011) Global quantification of mammalian gene expression control. *Nature*, doi: 10.1038/nature10098.
- Schwarz, S., Kehrenberg, C., Doublet, B., and Cloeckaert, A. (2004) Molecular basis of bacterial resistance to chloramphenicol and florfenicol. *FEMS microbiology reviews*, doi: 10.1016/j.femsre.2004.04.001.
- Seemann, T. (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics (Oxford, England)*, doi: 10.1093/bioinformatics/btu153.
- Seymour, J.R., Amin, S.A., Raina, J.-B., and Stocker, R. (2017) Zooming in on the phycosphere: the ecological interface for phytoplankton-bacteria relationships. *Nature microbiology*, doi: 10.1038/nmicrobiol.2017.65.
- Seymour, JR, Ahmed, T., Durham, W.M., and Stocker, R. (2010) Chemotactic response of marine bacteria to the extracellular products of *Synechococcus* and *Prochlorococcus*. *Aquatic Microbial Ecology*, doi: 10.3354/ame01400.
- Sheets, J., and Aktories, K. (2017) Insecticidal Toxin Complexes from *Photobacterium luminescens*. *Current topics in microbiology and immunology*, doi: 10.1007/82_2016_55.
- Shi, R., Huang, H., Qi, Z., Hu, W., Tian, Z., and Dai, M. (2013) Algicidal activity against *Skeletonema costatum* by marine bacteria isolated from a high frequency harmful algal blooms area in southern Chinese coast. *World journal of microbiology & biotechnology*, doi: 10.1007/s11274-012-1168-1.
- Shin, J.-B., Krey, J.F., Hassan, A., Metlagel, Z., Tauscher, A.N., Pagana, J.M., et al. (2013) Molecular architecture of the chick vestibular hair bundle. *Nature neuroscience*, doi: 10.1038/nn.3312.
- Shipman, J.A., Berleman, J.E., and Salyers, A.A. (2000) Characterization of four outer membrane proteins involved in binding starch to the cell surface of *Bacteroides thetaiotaomicron*. *Journal of bacteriology*, doi: 10.1128/jb.182.19.5365-5372.2000.
- Shrivastava, A., Johnston, J.J., van Baaren, J.M., and McBride, M.J. (2013) *Flavobacterium johnsoniae* GldK, GldL, GldM, and SprA are required for secretion of the cell surface gliding motility adhesins SprB and RemA. *Journal of bacteriology*, doi: 10.1128/JB.00333-13.
- Sigman, D.M., and Haug, G.H. (2003) The Biological Pump in the Past. In *Treatise on Geochemistry*: Elsevier, pp. 491–528.
- Simm, R., Morr, M., Kader, A., Nimtz, M., and Römling, U. (2004) GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. *Molecular microbiology*, doi: 10.1111/j.1365-2958.2004.04206.x.
- Simon, M., Glöckner, F.O., and Amann, R. (1999) Different community structure and temperature optima of heterotrophic picoplankton in various regions of the Southern Ocean. *Aquatic Microbial Ecology*, doi: 10.3354/ame018275.
- Simon, M., Grossart, H.P., Schweitzer, B., and Ploug, H. (2002) Microbial ecology of organic aggregates in aquatic ecosystems. *Aquatic Microbial Ecology*, doi: 10.3354/ame028175.
- Simon, M., Scheuner, C., Meier-Kolthoff, J.P., Brinkhoff, T., Wagner-Döbler, I., Ulbrich, M., et al. (2017) Phylogenomics of Rhodobacteraceae reveals evolutionary adaptation to marine and non-marine habitats. *The ISME journal*, doi: 10.1038/ismej.2016.198.
- Smayda, T.J. (1997) Harmful algal blooms: Their ecophysiology and general relevance to phytoplankton blooms in the sea. *Limnology and Oceanography*: 1137–1153.
- Smetacek, V., and Zingone, A. (2013) Green and golden seaweed tides on the rise. *Nature*, doi: 10.1038/nature12860.
- Smith, D.C., Simon, M., Alldredge, A.L., and Azam, F. (1992) Intense hydrolytic enzyme activity on marine aggregates and implications for rapid particle dissolution. *Nature*, doi: 10.1038/359139a0.
- Smith, P.K., Krohn, R.I., Hermanson, G.T., Mallia, A.K., Gartner, F.H., Provenzano, M.D., et al. (1985) Measurement of Protein Using Bicinchoninic Acid. *Analytical Biochemistry*: 76–85.

References

- Sohn, J.H., Lee, J.-H., Yi, H., Chun, J., Bae, K.S., Ahn, T.-Y., and Kim, S.-J. (2005) *Kordia algicida* gen. nov., sp. nov., an algicidal bacterium isolated from red tide. *International Journal of Systematic and Evolutionary Microbiology*, doi: 10.1099/00207713-55-6-2639.
- Sonnenburg, E.D., Zheng, H., Joglekar, P., Higginbottom, S.K., Firbank, S.J., Bolam, D.N., and Sonnenburg, J.L. (2010) Specificity of polysaccharide use in intestinal bacteroides species determines diet-induced microbiota alterations. *Cell*, doi: 10.1016/j.cell.2010.05.005.
- Sowell, S.M., Wilhelm, L.J., Norbeck, A.D., Lipton, M.S., Nicora, C.D., Barofsky, D.F., et al. (2009) Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *The ISME journal*, doi: 10.1038/ismej.2008.83.
- Spaink, H.P. (2004) Specific recognition of bacteria by plant LysM domain receptor kinases. *Trends in microbiology*, doi: 10.1016/j.tim.2004.03.001.
- Spoerner, M., Wichard, T., Bachhuber, T., Stratmann, J., and Oertel, W. (2012) Growth and Thallus Morphogenesis of *Ulva mutabilis* (Chlorophyta) Depends on A Combination of Two Bacterial Species Excreting Regulatory Factors. *Journal of Phycology*, doi: 10.1111/j.1529-8817.2012.01231.x.
- Stanke, M., Steinkamp, R., Waack, S., and Morgenstern, B. (2004) AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic acids research*, doi: 10.1093/nar/gkh379.
- Steffen, M.M., Dearth, S.P., Dill, B.D., Li, Z., Larsen, K.M., Campagna, S.R., and Wilhelm, S.W. (2014) Nutrients drive transcriptional changes that maintain metabolic homeostasis but alter genome architecture in *Microcystis*. *The ISME journal*, doi: 10.1038/ismej.2014.78.
- Stewart, P.S., and Franklin, M.J. (2008) Physiological heterogeneity in biofilms. *Nature Reviews Microbiology*, doi: 10.1038/nrmicro1838.
- Stocker, R. (2012) Marine microbes see a sea of gradients. *Science*, doi: 10.1126/science.1208929.
- Stocker, R., and Seymour, J.R. (2012) Ecology and physics of bacterial chemotaxis in the ocean. *Microbiology and molecular biology reviews : MMBR*, doi: 10.1128/MMBR.00029-12.
- Sullivan, J.T., Trzebiatowski, J.R., Cruickshank, R.W., Gouzy, J., Brown, S.D., Elliot, R.M., et al. (2002) Comparative sequence analysis of the symbiosis island of *Mesorhizobium loti* strain R7A. *Journal of bacteriology*, doi: 10.1128/JB.184.11.3086-3095.2002.
- Sunagawa, S., Coelho, L.P., Chaffron, S., Kultima, J.R., Labadie, K., Salazar, G., et al. (2015) Ocean plankton. Structure and function of the global ocean microbiome. *Science (New York, N.Y.)*, doi: 10.1126/science.1261359.
- Suttle, C.A. (2007) Marine viruses--major players in the global ecosystem. *Nature Reviews Microbiology*, doi: 10.1038/nrmicro1750.
- Suzumura, M., and Kamatani, A. (1995) Origin and distribution of inositol hexaphosphate in estuarine and coastal sediments. *Limnology and Oceanography*, doi: 10.4319/lo.1995.40.7.1254.
- Sverzhinsky, A., Chung, J.W., Deme, J.C., Fabre, L., Levey, K.T., Plesa, M., et al. (2015) Membrane Protein Complex ExbB4-ExbD1-TonB1 from *Escherichia coli* Demonstrates Conformational Plasticity. *Journal of bacteriology*, doi: 10.1128/JB.00069-15.
- Sweet, M.J., Croquer, A., and Bythell, J.C. (2011) Bacterial assemblages differ between compartments within the coral holobiont. *Coral Reefs*, doi: 10.1007/s00338-010-0695-1.
- Tamura, K., Hemsworth, G.R., Déjean, G., Rogers, T.E., Pudlo, N.A., Urs, K., et al. (2017) Molecular Mechanism by which Prominent Human Gut Bacteroidetes Utilize Mixed-Linkage Beta-Glucans, Major Health-Promoting Cereal Polysaccharides. *Cell reports*, doi: 10.1016/j.celrep.2017.11.013.
- Tang, K., Jiao, N., Liu, K., Zhang, Y., and Li, S. (2012) Distribution and functions of TonB-dependent transporters in marine bacteria and environments: implications for dissolved organic matter utilization. *PloS one*, doi: 10.1371/journal.pone.0041204.
- Tang, K., Lin, Y., Han, Y., and Jiao, N. (2017) Characterization of Potential Polysaccharide Utilization Systems in the Marine Bacteroidetes *Gramella flava* JLT2011 Using a Multi-Omics Approach. *Frontiers in microbiology*, doi: 10.3389/fmicb.2017.00220.
- Tang, K.F.J., and Lightner, D.V. (2014) Homologues of insecticidal toxin complex genes within a genomic island in the marine bacterium *Vibrio parahaemolyticus*. *FEMS microbiology letters*, doi: 10.1111/1574-6968.12609.
- Teeling, H., Fuchs, B.M., Becher, D., Klockow, C., Gardebrecht, A., Bennis, C.M., et al. (2012) Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science (New York, N.Y.)*, doi: 10.1126/science.1218344.
- Teeling, H., Fuchs, B.M., Bennis, C.M., Krüger, K., Chafee, M., Kappelmann, L., et al. (2016) Recurring patterns in bacterioplankton dynamics during coastal spring algae blooms. *eLife*, doi: 10.7554/eLife.11888.
- Teeling, H., and Glöckner, F.O. (2012) Current opportunities and challenges in microbial metagenome analysis--a bioinformatic perspective. *Briefings in bioinformatics*, doi: 10.1093/bib/bbs039.

References

- Thomas, F., Barbeyron, T., Tonon, T., Génicot, S., Czjzek, M., and Michel, G. (2012) 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. *Environmental microbiology*, doi: 10.1111/j.1462-2920.2012.02751.x.
- Thomas, F., Bordron, P., Eveillard, D., and Michel, G. (2017) Gene Expression Analysis of *Zobellia galactanivorans* during the Degradation of Algal Polysaccharides Reveals both Substrate-Specific and Shared Transcriptome-Wide Responses. *Frontiers in microbiology*, doi: 10.3389/fmicb.2017.01808.
- Thomas, F., Hehemann, J.-H., Rebuffet, E., Czjzek, M., and Michel, G. (2011) Environmental and gut bacteroidetes: the food connection. *Frontiers in microbiology*, doi: 10.3389/fmicb.2011.00093.
- Thompson, J.R., Pacocha, S., Pharino, C., Klepac-Ceraj, V., Hunt, D.E., Benoit, J., et al. (2005) Genotypic diversity within a natural coastal bacterioplankton population. *Science*, doi: 10.1126/science.1106028.
- Thompson, M.R., Chourey, K., Froelich, J.M., Erickson, B.K., VerBerkmoes, N.C., and Hettich, R.L. (2008) Experimental approach for deep proteome measurements from small-scale microbial biomass samples. *Analytical chemistry*, doi: 10.1021/ac801707s.
- Timmins-Schiffman, E., May, D.H., Mikan, M., Riffle, M., Frazar, C., Harvey, H.R., Noble, W.S., and Nunn, B.L. (2017) Critical decisions in metaproteomics: achieving high confidence protein annotations in a sea of unknowns. *The ISME journal*, doi: 10.1038/ismej.2016.132.
- Tischler, A.D., and Camilli, A. (2004) Cyclic diguanylate (c-di-GMP) regulates *Vibrio cholerae* biofilm formation. *Molecular microbiology*, doi: 10.1111/j.1365-2958.2004.04155.x.
- Tkaczuk, K.L., A Shumilin, I., Chruszcz, M., Evdokimova, E., Savchenko, A., and Minor, W. (2013) Structural and functional insight into the universal stress protein family. *Evolutionary applications*, doi: 10.1111/eva.12057.
- Tolbert N. E., Zill L. P. (1956) Excretion of glycolic acid by algae during photosynthesis. *The Journal of biological chemistry*, **222**: 895–906.
- Tréguer, P., Bowler, C., Moriceau, B., Dutkiewicz, S., Gehlen, M., Aumont, O., et al. (2018) Influence of diatom diversity on the ocean biological carbon pump. *Nature Geoscience*, doi: 10.1038/s41561-017-0028-x.
- Trincon, A. (2018) Update on Marine Carbohydrate Hydrolyzing Enzymes: Biotechnological Applications. *Molecules (Basel, Switzerland)*, doi: 10.3390/molecules23040901.
- Unfried, F., Becker, S., Robb, C.S., Hehemann, J.-H., Markert, S., Heiden, S.E., et al. (2018) Adaptive mechanisms that provide competitive advantages to marine bacteroidetes during microalgal blooms. *The ISME journal*, doi: 10.1038/s41396-018-0243-5.
- van der Weij-De Wit, C.D., Doust, A.B., van Stokkum, I.H.M., Dekker, J.P., Wilk, K.E., Curmi, P.M.G., Scholes, G.D., and van Grondelle, R. (2006) How energy funnels from the phycoerythrin antenna complex to photosystem I and photosystem II in cryptophyte *Rhodomonas* CS24 cells. *The journal of physical chemistry. B*, doi: 10.1021/jp061546w.
- van Nguyen, D., Pham, T.T., Nguyen, T.H.T., Nguyen, T.T.X., and Hoj, L. (2014) Screening of marine bacteria with bacteriocin-like activities and probiotic potential for ornate spiny lobster (*Panulirus ornatus*) juveniles. *Fish & shellfish immunology*, doi: 10.1016/j.fsi.2014.06.017.
- Vargas, C. de, Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., et al. (2015) Ocean plankton. Eukaryotic plankton diversity in the sunlit ocean. *Science*, doi: 10.1126/science.1261605.
- Veith, P.D., Glew, M.D., Gorasia, D.G., and Reynolds, E.C. (2017) Type IX secretion: the generation of bacterial cell surface coatings involved in virulence, gliding motility and the degradation of complex biopolymers. *Molecular microbiology*, doi: 10.1111/mmi.13752.
- Venisse, J.S., Gullner, G., and Brisset, M.N. (2001) Evidence for the involvement of an oxidative stress in the initiation of infection of pear by *Erwinia amylovora*. *Plant physiology*, doi: 10.1104/pp.125.4.2164.
- VerBerkmoes, N.C., Deneff, V.J., Hettich, R.L., and Banfield, J.F. (2009) Systems biology: Functional analysis of natural microbial consortia using community proteomics. *Nature Reviews Microbiology*, doi: 10.1038/nrmicro2080.
- Verheggen, K., Raeder, H., Berven, F.S., Martens, L., Barsnes, H., and Vaudel, M. (2020) Anatomy and evolution of database search engines—a central component of mass spectrometry based proteomic workflows. *Mass spectrometry reviews*, doi: 10.1002/mas.21543.
- Villacorte, L.O., Ekowati, Y., Neu, T.R., Kleijn, J.M., Winters, H., Amy, G., Schippers, J.C., and Kennedy, M.D. (2015) Characterisation of algal organic matter produced by bloom-forming marine and freshwater algae. *Water Research*, doi: 10.1016/j.watres.2015.01.028.
- Vincent, M.S., Chabaliere, M., and Cascales, E. (2018) A conserved motif of *Porphyromonas* Type IX secretion effectors C-terminal secretion signal specifies interactions with the PorKLMN core complex.

References

- Volkman, J.K., and Tanoue, E. (2002) Chemical and biological studies of particulate organic matter in the ocean. *Journal of Oceanography*, doi: 10.1023/A:1015809708632.
- Vraspir, J.M., and Butler, A. (2009) Chemistry of marine ligands and siderophores. *Annual review of marine science*, doi: 10.1146/annurev.marine.010908.163712.
- Wang, D.-Z., Xie, Z.-X., and Zhang, S.-F. (2014) Marine metaproteomics: current status and future directions. *Journal of proteomics*, doi: 10.1016/j.jprot.2013.08.024.
- Ward, C.S., Yung, C.-M., Davis, K.M., Blinbry, S.K., Williams, T.C., Johnson, Z.I., and Hunt, D.E. (2017) Annual community patterns are driven by seasonal switching between closely related marine bacteria. *The ISME journal*, doi: 10.1038/ismej.2017.4.
- Wemheuer, B., Wemheuer, F., Hollensteiner, J., Meyer, F.-D., Voget, S., and Daniel, R. (2015) The green impact: bacterioplankton response toward a phytoplankton spring bloom in the southern North Sea assessed by comparative metagenomic and metatranscriptomic approaches. *Frontiers in microbiology*, doi: 10.3389/fmicb.2015.00805.
- Whitman, W.B., Coleman, D.C., and Wiebe, W.J. (1998) Prokaryotes: the unseen majority. *Proceedings of the National Academy of Sciences of the United States of America*, doi: 10.1073/pnas.95.12.6578.
- Widdel, F., and Bak, F. (1992) *Gram-Negative Mesophilic Sulfate-Reducing Bacteria*. In: Balows A., Trüper H.G., Dworkin M., Harder W., Schleifer KH. (eds) *The Prokaryotes. A Handbook on the Biology of Bacteria: Ecophysiology, Isolation, Identification, Applications*. New York, NY: Springer.
- Wilkins, D., Lauro, F.M., Williams, T.J., Demaere, M.Z., Brown, M.V., Hoffman, J.M., et al. (2013) Biogeographic partitioning of Southern Ocean microorganisms revealed by metagenomics. *Environmental microbiology*, doi: 10.1111/1462-2920.12035.
- Williams, T.J., Long, E., Evans, F., Demaere, M.Z., Lauro, F.M., Raftery, M.J., et al. (2012) A metaproteomic assessment of winter and summer bacterioplankton from Antarctic Peninsula coastal surface waters. *The ISME journal*, doi: 10.1038/ismej.2012.28.
- Williams, T.J., Wilkins, D., Long, E., Evans, F., Demaere, M.Z., Raftery, M.J., and Cavicchioli, R. (2013) The role of planktonic Flavobacteria in processing algal organic matter in coastal East Antarctica revealed using metagenomics and metaproteomics. *Environmental microbiology*, doi: 10.1111/1462-2920.12017.
- Wilmes, P., Wexler, M., and Bond, P.L. (2008) Metaproteomics provides functional insight into activated sludge wastewater treatment. *PloS one*, doi: 10.1371/journal.pone.0001778.
- Wiltshire, K.H., Malzahn, A.M., Wirtz, K., Greve, W., Janisch, S., Mangelsdorf, P., Manly, B.F.J., and Boersma, M. (2008) Resilience of North Sea phytoplankton spring bloom dynamics: An analysis of long-term data at Helgoland Roads. *Limnology and Oceanography*, doi: 10.4319/lo.2008.53.4.1294.
- Wiltshire, K.H., and Manly, B.F.J. (2004) The warming trend at Helgoland Roads, North Sea: phytoplankton response. *Helgoland Marine Research*, doi: 10.1007/s10152-004-0196-0.
- Winkelmann, N., and Harder, J. (2009) An improved isolation method for attached-living Planctomycetes of the genus Rhodopirellula. *Journal of microbiological methods*, doi: 10.1016/j.mimet.2009.03.002.
- Wöhlbrand, L., Feenders, C., Nachbaur, J., Freund, H., Engelen, B., Wilkes, H., Brumsack, H.-J., and Rabus, R. (2017a) Impact of Extraction Methods on the Detectable Protein Complement of Metaproteomic Analyses of Marine Sediments. *Proteomics*, doi: 10.1002/pmic.201700241.
- Wöhlbrand, L., Wemheuer, B., Feenders, C., Ruppertsberg, H.S., Hinrichs, C., Blasius, B., Daniel, R., and Rabus, R. (2017b) Complementary Metaproteomic Approaches to Assess the Bacterioplankton Response toward a Phytoplankton Spring Bloom in the Southern North Sea. *Frontiers in microbiology*, doi: 10.3389/fmicb.2017.00442.
- Woyke, T., Xie, G., Copeland, A., González, J.M., Han, C., Kiss, H., et al. (2009) Assembling the marine metagenome, one cell at a time. *PloS one*, doi: 10.1371/journal.pone.0005299.
- Xing, P., Hahnke, R.L., Unfried, F., Markert, S., Huang, S., Barbeyron, T., et al. (2015) Niches of two polysaccharide-degrading Polaribacter isolates from the North Sea during a spring diatom bloom. *The ISME journal*, doi: 10.1038/ismej.2014.225.
- Yawata, Y., Cordero, O.X., Menolascina, F., Hehemann, J.-H., Polz, M.F., and Stocker, R. (2014) Competition-dispersal tradeoff ecologically differentiates recently speciated marine bacterioplankton populations. *Proceedings of the National Academy of Sciences*, doi: 10.1073/pnas.1318943111.
- Ye, N., Zhang, X., Mao, Y., Liang, C., Xu, D., Zou, J., Zhuang, Z., and Wang, Q. (2011) 'Green tides' are overwhelming the coastline of our blue planet: taking the world's largest example. *Ecological Research*, doi: 10.1007/s11284-011-0821-8.

References

- Yoch, D.C. (2002) Dimethylsulfoniopropionate: its sources, role in the marine food web, and biological degradation to dimethylsulfide. *Applied and Environmental Microbiology*, doi: 10.1128/AEM.68.12.5804-5815.2002.
- Yoon, B.-J., and Oh, D.-C. (2012) *Spongiibacterium flavum* gen. nov., sp. nov., a member of the family Flavobacteriaceae isolated from the marine sponge *Halichondria oshoro*, and emended descriptions of the genera *Croceitalea* and *Flagellimonas*. *International Journal of Systematic and Evolutionary Microbiology*, doi: 10.1099/ijs.0.027243-0.
- Zehr, J.P. (2015) EVOLUTION. How single cells work together. *Science*, doi: 10.1126/science.aac9752.
- Zhang, H., Jia, J., Chen, S., Huang, T., Wang, Y., Zhao, Z., *et al.* (2018a) Dynamics of Bacterial and Fungal Communities during the Outbreak and Decline of an Algal Bloom in a Drinking Water Reservoir. *International journal of environmental research and public health*, doi: 10.3390/ijerph15020361.
- Zhang, H., Yohe, T., Le Huang, Entwistle, S., Wu, P., Yang, Z., *et al.* (2018b) dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic acids research*, doi: 10.1093/nar/gky418.
- Zheng, Q., Wang, Y., Lu, J., Lin, W., Chen, F., and Jiao, N. (2020) Metagenomic and Metaproteomic Insights into Photoautotrophic and Heterotrophic Interactions in a *Synechococcus* Culture. *mBio*, doi: 10.1128/mBio.03261-19.
- Zhou, J., Bruns, M.A., and Tiedje, J.M. (1996) DNA Recovery from Soils of Diverse Composition. *Applied and Environmental Microbiology*: 316–322.
- Zhu, W., Lomsadze, A., and Borodovsky, M. (2010a) Ab initio gene identification in metagenomic sequences. *Nucleic acids research*, doi: 10.1093/nar/gkq275.
- Zhu, W., Smith, J.W., and Huang, C.-M. (2010b) Mass spectrometry-based label-free quantitative proteomics. *Journal of biomedicine & biotechnology*, doi: 10.1155/2010/840518.
- Ziervogel, K., and Arnosti, C. (2008) Polysaccharide hydrolysis in aggregates and free enzyme activity in aggregate-free seawater from the north-eastern Gulf of Mexico. *Environmental microbiology*, doi: 10.1111/j.1462-2920.2007.01451.x.
- Ziervogel, K., Steen, A.D., and Arnosti, C. (2010) Changes in the spectrum and rates of extracellular enzyme activities in seawater following aggregate formation. *Biogeosciences*, doi: 10.5194/bg-7-1007-2010.
- Zybailov, B., Mosley, A.L., Sardi, M.E., Coleman, M.K., Florens, L., and Washburn, M.P. (2006) Statistical analysis of membrane proteome expression changes in *Saccharomyces cerevisiae*. *Journal of proteome research*, doi: 10.1021/pr060161n.

9. Appendix

9.1. Supplemental Material

Table S 1. Total protein amounts and number of sampling events of the six tested protein extraction protocols

Protocols	Sampling events/total protein amount in μg												Number of sampling events
	2/9/2009		6/16/2009		4/7/2009		4/21/2009		Mean values and standard deviations		4/28/2009		
	3 μm ($1/4$ filter)	10 μm ($1/4$ filter)	3 μm ($1/4$ filter)	10 μm ($1/4$ filter)	3 μm ($1/4$ filter)	10 μm ($1/4$ filter)	3 μm ($1/4$ filter)	10 μm ($1/4$ filter)	3 μm ($1/4$ filter)	10 μm ($1/4$ filter)	3 μm (3 filters)	10 μm (3 filters)	
Phenol	8.6	8.6	9.2	8.9	-	-	-	-	8.9 ± 0.42	8.75 ± 0.21	-	-	2
SDS-TCA	32.4	13.5	13.0	12.1	-	-	-	-	22.7 ± 13.71	12.8 ± 0.99	-	-	2
SDS-Acetone	34.7	20.4	42.6	183.8	29.4	117.5	47.7	86.7	38.6 ± 8.13	102.1 ± 67.89	-	-	4
Bead Beating	20.2	105.8	34.3	122.7	27.4	125.5	27.3	102.9	27.3 ± 5.75	114.2 ± 11.52	304.9	1126.2	5
TRI®-Reagent	-	-	-	-	16.1	38.2	15.9	39.2	16 ± 0.14	38.7 ± 0.71	-	-	2
Freezing Thawing	-	-	-	-	24.2	24.3	26.4	22.9	25.3 ± 1.56	23.6 ± 0.99	-	-	2

Appendix

Table S 2. OD at 600 nm of *Muricauda* sp. MAR_2010_75 grown in HaHa100V medium with glucose, laminarin, alginate and without addition of sugars. The red marked ODs represent the harvest time points for proteomic analysis. Mean values were used for creation of growth curves (sd= standard deviation).

Time in h	Glucose br1	Glucose br2	Glucose br3	Glucose br4	Glucose mean OD _{600nm} (sd)	Laminarin br1	Laminarin br2	Laminarin br3	Laminarin br4	Laminarin mean OD _{600nm} (sd)	Alginate br1	Alginate br2	Alginate br3	Alginate br4	Alginate mean OD _{600nm} (sd)
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17.5	0.051	0.046	0.049	0.054	0.05 (0.0034)	0.027	0.024	0.03	0.032	0.028 (0.0035)	0.041	0.042	0.039	0.038	0.040 (0.0018)
21.5	0.06	0.055	0.059	0.056	0.058 (0.0024)	0.035	0.029	0.038	0.037	0.035 (0.0040)	0.05	0.053	0.045	0.044	0.048 (0.0042)
24	0.064	0.066	0.067	0.063	0.065 (0.0018)	0.044	0.04	0.045	0.048	0.044 (0.0033)	0.06	0.059	0.054	0.053	0.057 (0.0035)
41.5	0.166	0.167	0.161	0.164	0.165 (0.0026)	0.186	0.191	0.184	0.188	0.187 (0.003)	0.129	0.123	0.122	0.126	0.125 (0.0032)
42.5	0.186	0.185	0.178	0.179	0.182 (0.0041)	0.198	0.204	0.199	0.201	0.201 (0.0026)	0.141	0.138	0.136	0.134	0.137 (0.0030)
43.5	0.199	0.195	0.195	0.192	0.196 (0.0032)	0.209	0.212	0.211	0.213	0.211 (0.0017)	0.153	0.152	0.147	0.149	0.150 (0.0028)
44.5	0.21	0.209	0.208	0.206	0.208 (0.0017)	0.229	0.232	0.225	0.231	0.229 (0.0031)	0.167	0.166	0.162	0.16	0.164 (0.0033)
45.5	0.233	0.234	0.231	0.229	0.232 (0.0022)	0.247	0.252	0.249	0.264	0.253 (0.0076)	0.188	0.185	0.181	0.181	0.184 (0.0034)
46.5	0.252	0.255	0.254	0.263	0.256 (0.0048)				0.288	0.288	0.203	0.205	0.204	0.205	0.204 (0.0010)
47.5				0.279	0.279				0.304	0.304	0.255	0.21	0.256	0.249	0.243 (0.0219)
65.5				0.386	0.386				0.388	0.388				0.344	0.344
70.25				0.399	0.399				0.37	0.370				0.354	0.354
89.75				0.47	0.47				0.364	0.364				0.316	0.316
94.75				0.455	0.455				0.36	0.360				0.317	0.317
161.75				0.445	0.445				0.315	0.315				0.294	0.294

Appendix

Time in h	without br1	without br2	without br3	without br4	without mean OD _{600nm} (sd)
0	0	0	0	0	0
17.5	0.016	0.014	0.012	0.015	0.014 (0.0017)
21.5	0.022	0.02	0.019	0.021	0.021 (0.0013)
24	0.029	0.026	0.026	0.027	0.027 (0.0014)
41.5	0.072	0.066	0.064	0.07	0.068 (0.0037)
42.5	0.081	0.074	0.078	0.076	0.077 (0.0030)
43.5	0.093	0.082	0.089	0.084	0.087 (0.0050)
44.5	0.099	0.093	0.098	0.092	0.096 (0.0035)
45.5	0.111	0.105	0.11	0.104	0.108 (0.0035)
46.5	0.123	0.118	0.121	0.117	0.120 (0.0028)
47.5	0.151	0.149	0.152	0.159	0.153 (0.0043)
65.5				0.201	0.201
70.25				0.196	0.196
89.75				0.189	0.189
94.75				0.183	0.183
161.75				0.181	0.181

Appendix

Table S 3. Type 9 secretion system (T9SS) –related proteins identified in the proteomes of *Muricauda* sp. MAR_2010_75 grown on laminarin and alginate and their fold changes (fc) compared to the control glucose.

Protein name (motif)	Locus_tag	fc Laminarin	fc Alginate
T9SS type B sorting domain-containing protein (TIGR04131)	FG28_RS00555	1.23	
	FG28_RS06240	0.69	0.55
	FG28_RS02080	1.97	1.17
	FG28_RS09995	0.49	1.35
	FG28_RS02075	1.55	1.13
	FG28_RS15815	1.42	0.71
T9SS type B sorting domain-containing protein (pfam13585)	FG28_RS12075	2.85	0.67
T9SS type A sorting domain-containing protein (pfam11721)	FG28_RS15415		
T9SS type A sorting domain-containing protein (TIGR04183)	FG28_RS04540	3.05	
	FG28_RS18335	0.54	1.06
T9SS type A sorting domain-containing protein (cl00064)	FG28_RS08530	3.84	2.98
PorT family protein (pfam13568)	FG28_RS12515	1.28	1.77
	FG28_RS12370	1.55	0.98
	FG28_RS11420	0.79	0.95
	FG28_RS09825	1.05	0.42
PorT family protein (cl21487)	FG28_RS16695	1.81	1.25
T9SS outer membrane channel protein PorV (cl21487)	FG28_RS03310	1.55	1.28
T9SS outer membrane channel protein PorP/SprF (pfam11751)	FG28_RS10355	1.02	0.99
	FG28_RS08705	2.61	1.95
	FG28_RS06235	1.93	1.40
	FG28_RS12080	1.71	1.98
	FG28_RS10315	3.00	1.81
	FG28_RS12060	0.46	0.86
T9SS outer membrane channel protein PorP/SprF (TIGR03519)	FG28_RS06245	2.19	2.34
T9SS outer membrane channel protein PorP/SprF (cl14675)	FG28_RS19175	0.89	0.48

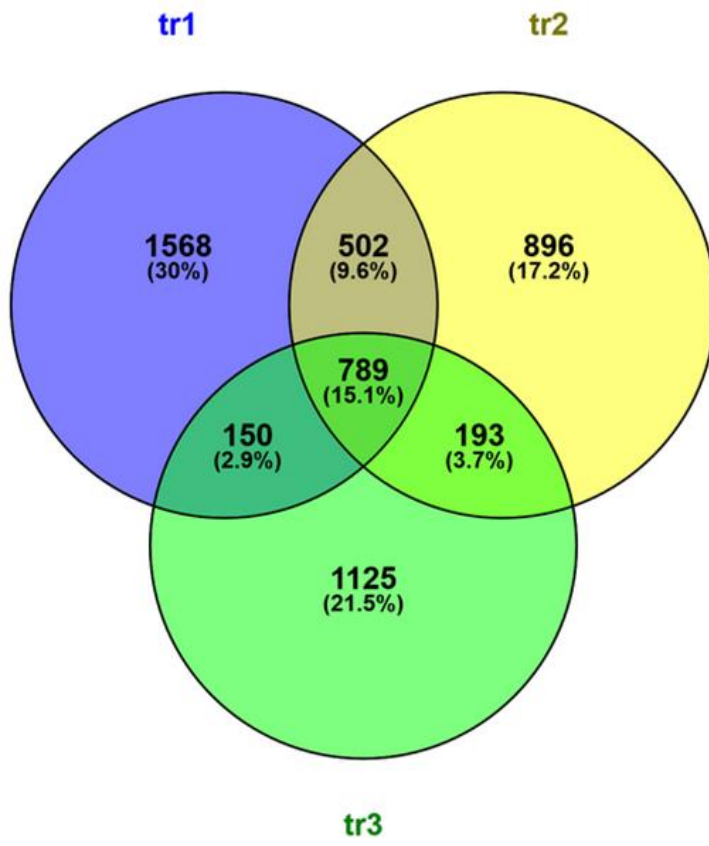


Figure S 1. Venn diagram showing the number of specific and shared identified protein groups in the three technical replicates of the metaproteomics analysis of the 24th of May during the spring bloom 2018 at “Kabeltonne” Helgoland.

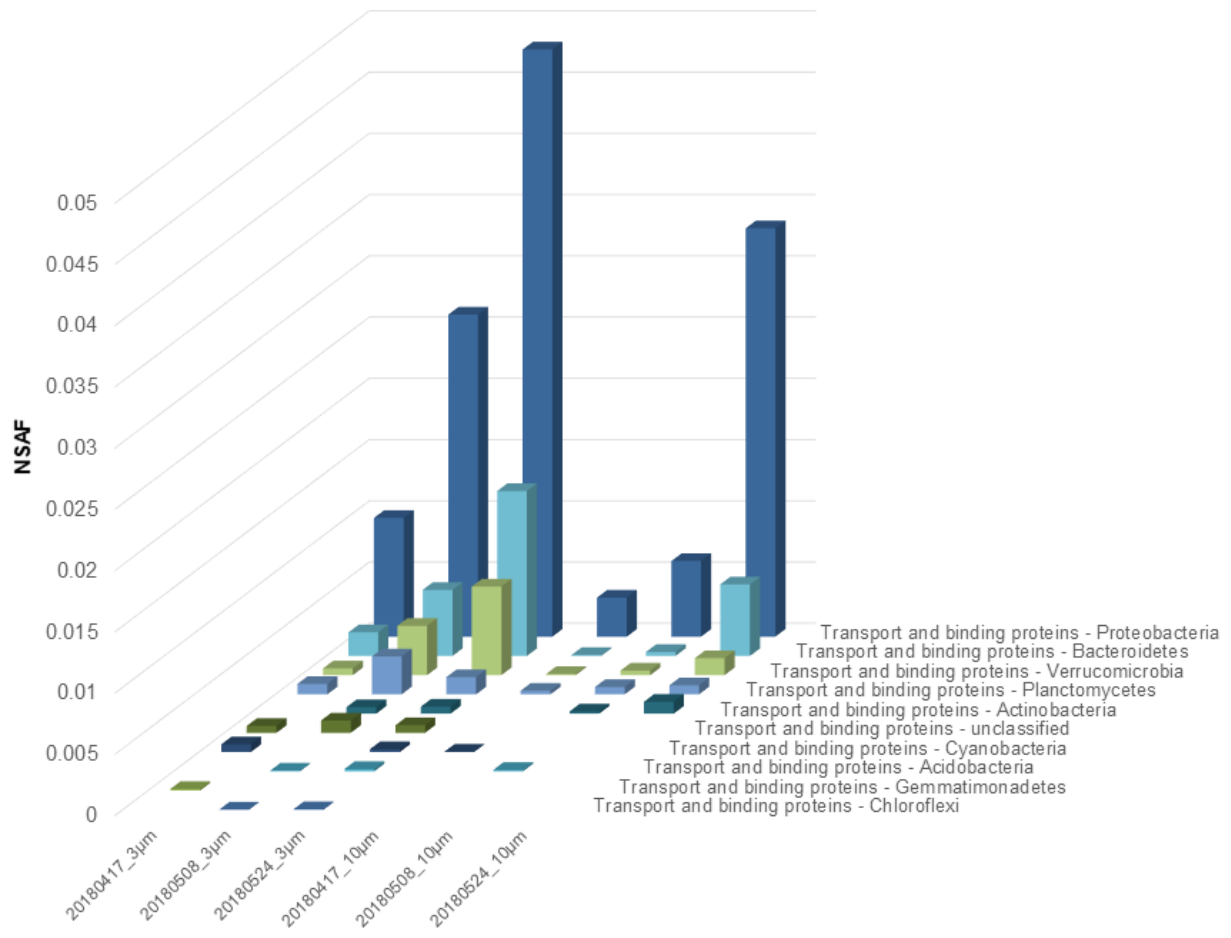
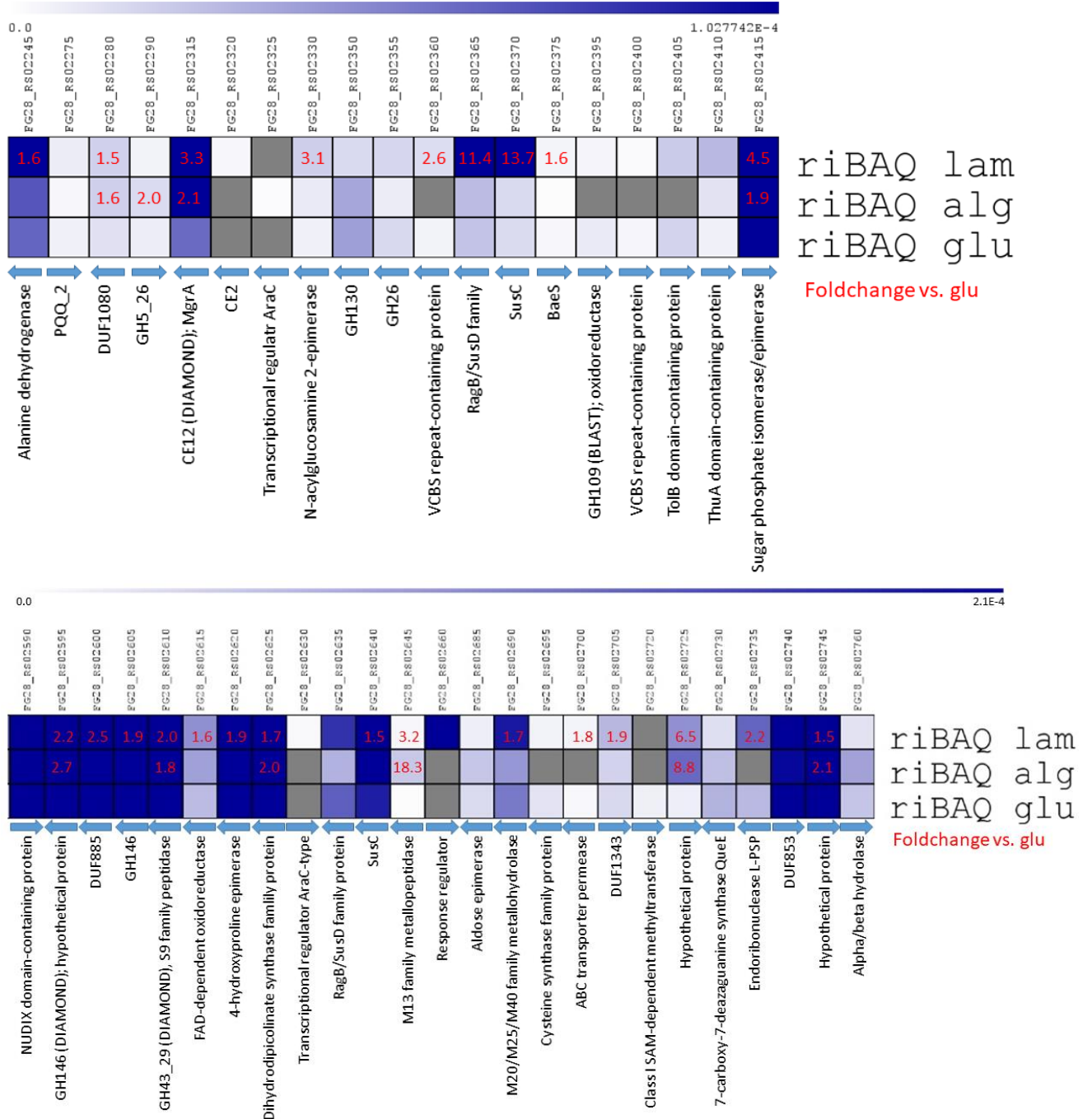


Figure S 2. Taxonomic affiliation of prokaryotic transport and binding protein associated protein groups in metaproteomes of pre-bloom, bloom rise and bloom peak samples during the spring bloom 2018 at "Kabeltonne" Helgoland. Total abundance of selected protein groups with assigned taxa.

Appendix



Appendix

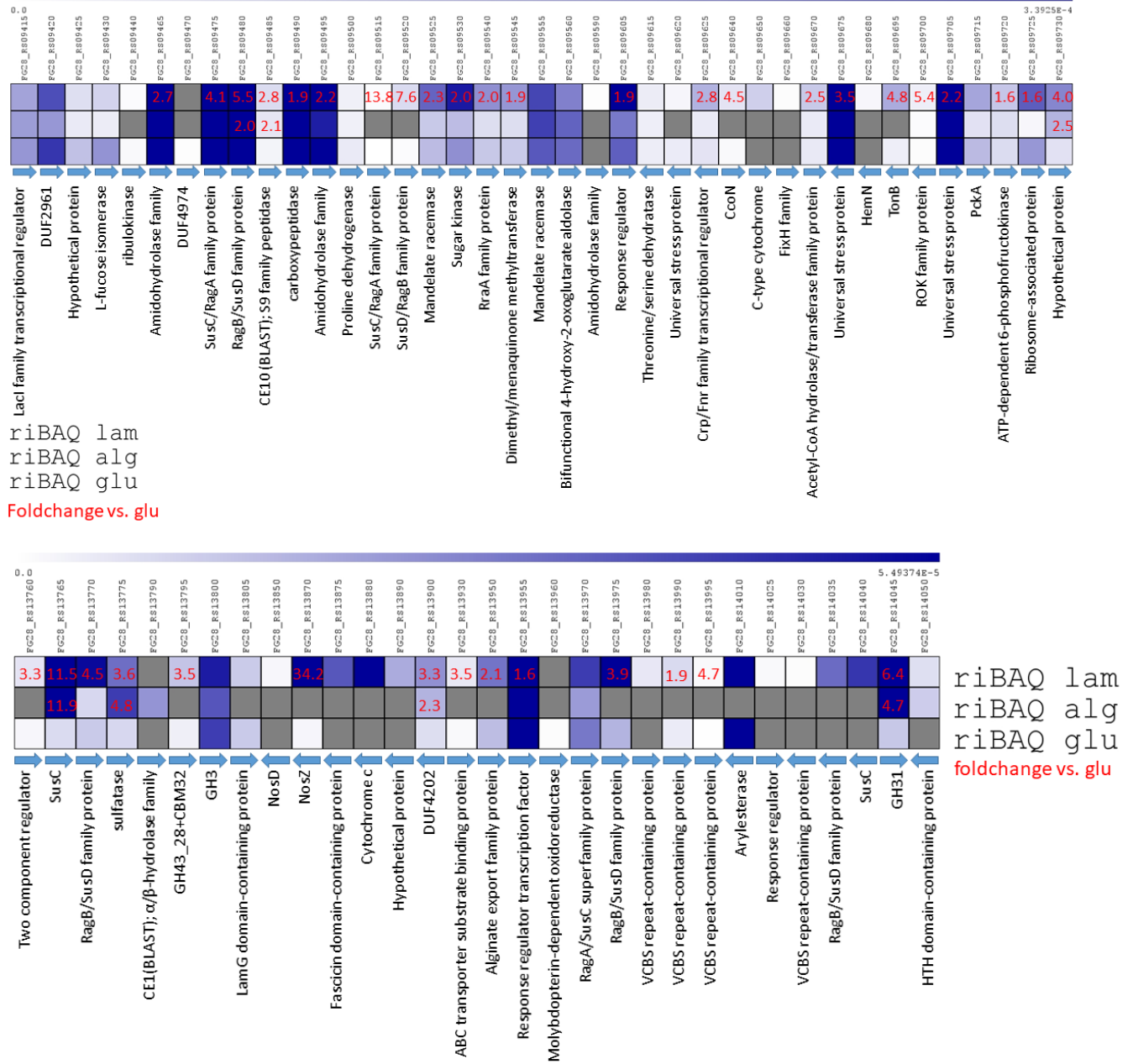


Figure S 3. Identified proteins in the proteome of *Muricauda sp. MAR_2010_75* grown on glucose as control, laminarin and alginate belonging to the same operon or genomic neighborhood of potential PUL marker proteins (CAZymes or SusC).

9.2. Acknowledgements

The work for this thesis was carried out at the Institute of Microbiology at the University of Greifswald and received funding from the German Research Foundation (DFG) in the frame of the Research Unit "POMPU" (FOR2406).

This thesis was made possible by the support and assistance of many people, who I would like to thank.

I would like to thank Prof. Dr. Katharina Riedel for giving me the opportunity to carry out this work. I appreciate your support and kind advice throughout my thesis.

I am also very grateful to my co-supervisors Dr. Daniela Zühlke and Dr. Mia Bengtsson for the engagement in this project. I appreciate the helpful discussions and constructive suggestions, which helped me a lot throughout my PhD.

Furthermore, I am thankful to the whole POMPU-consortium. It was a great experience to be a part of this research network. I am particularly grateful to Prof. Dr. Jens Harder for all his support throughout the thesis and especially during the work on the proteomic analysis of a PA isolate. The respective cultivation experiments were carried out in his lab by Saskia Kalenborn, thank you! I would also like to thank Prof. Dr. Rudolf Amann and especially Dr. Thomas Ben Francis for all the support, effort and work you put into the metagenomic analysis and Sabine Kühn for technical assistance. Furthermore, I want to thank all the kind people at the AWI Helgoland and especially Lilly and Anni for the great sampling campaigns 2017 and 2018 at Helgoland.

For all the MS measurements I am grateful to Prof. Dr. Dörte Becher, Dr. Daniela Zühlke, Dr. Claudia Hirschfeld, Dr. Pierre Mücke and Dr. Victoria Pauker from the Institute of Microbiology of the University of Greifswald.

I would like to thank the whole working group AG Riedel. Thank you for creating such a friendly and supportive atmosphere. Especially, I would like to thank my fellow PhD students for all the scientific discussions but mainly for all the support and fun. Madita, thank you for your friendship, all your support and the great time.

Mein größter Dank geht an meine Familie. Vor allem an meinen Mann Daniel und meinen Sohn Samuel. Ihr seid Alles für mich! Danke für eure Liebe, die mein Leben so lebenswert und wertvoll macht. Daniel, danke für deine unermüdliche Unterstützung, dein Verständnis und deine Ermutigung. Ich danke auch meinen Eltern für ihre Unterstützung und ihren Rückhalt. Vor Allem dir Mama, für deinen unglaublichen Einsatz in der Kinderbetreuung in den letzten Monaten dieser Arbeit. Ihr seid die Besten.

9.3. Curriculum vitae

Personal information

Name: Doreen Schultz, born Schneider

Date and place of birth: June 28th, 1991 in Siegen-Weidenau, Germany

Education

1997-2001	Gemeinschaftsgrundschule, Dreis-Tiefenbach
2001-2004	Gymnasium Stift Keppel
2004-2011	Gymnasium Altenholz, Abitur 2011

Academic studies

2011-2014	Bachelor of Science Biochemistry, University of Greifswald
2014-2016	Master of Science Molecular biology and physiology, University of Greifswald

Bachelor thesis

November 2016	Institute of Microbiology, University of Greifswald “Untersuchungen von Staphylococcus aureus Proteinen mit unbekannter Funktion“
---------------	--

Master thesis

November 2016	Institute of Microbiology, University of Greifswald “Konstruktion von Reporterstämmen zur Untersuchung physiologischer Heterogenität in Staphylococcus aureus Biofilmen“
---------------	---

PhD thesis

from December 2016	PhD student at the Institute of Microbiology, University of Greifswald “Mechanisms of polysaccharide degradation of particle-associated microbial communities”
--------------------	---

9.4. List of publications, posters and presentations

Publication

Schultz D, Zühlke D, Bernhardt J, Francis TB, Albrecht D, Hirschfeld C, Markert S, Riedel K. An optimized metaproteomics protocol for a holistic taxonomic and functional characterization of microbial communities from marine particles. *Environ Microbiol Rep.* 2020 Aug;12(4):367-376. doi: 10.1111/1758-2229.12842 . Epub 2020 Apr 23. PMID: 32281239 .

Posters

March 2017 – 5th Joint Conference of the DGHM & VAAM, Würzburg, Germany

August 2018 – 7th International Symposium on Microbial Ecology, Leipzig, Germany

August 2018 – MIMAS (Microbial Interactions in Marine Systems) II Symposium, Greifswald, Germany

Presentation

December 2018 - 3rd International Metaproteome Symposium, Leipzig, Germany

9.5. Eigenständigkeitserklärung

Hiermit erkläre ich, dass diese Arbeit bisher von mir weder an der Mathematisch-Naturwissenschaftlichen Fakultät der Universität Greifswald noch einer anderen wissenschaftlichen Einrichtung zum Zwecke der Promotion eingereicht wurde.

Ferner erkläre ich, dass ich diese Arbeit selbstständig verfasst und keine anderen als die darin angegebenen Hilfsmittel und Hilfen benutzt und keine Textabschnitte eines Dritten ohne Kennzeichnung übernommen habe.

Schillsdorf, den 27.09.2021

9.6. CD-ROM content

Table S1. Taxonomic assignment and functional annotation of all identified protein groups in the metaproteome of the FL fraction (0.2 μm) of the spring bloom 2009 and their NSAF values.

Table S2. Taxonomic assignment and functional annotation of all identified protein groups in the metaproteome of the small PA fraction (3 μm) of the spring bloom 2009 and their NSAF values.

Table S3. Taxonomic assignment and functional annotation of all identified protein groups in the metaproteome of the small PA fraction (10 μm) of the spring bloom 2009 and their NSAF values.

Table S4. Taxonomic assignment of identified protein groups and their relative abundance in the metaproteomes of the FL and PA fractions during the spring bloom 2009.

Table S5. Taxonomic assignment and functional annotation of all identified protein groups in the metaproteomes of PA fractions of the spring bloom 2018 and their NSAF values.

Table S6. Taxonomic assignment of identified bacterial protein groups and their relative abundance during the spring bloom 2018.

Table S7. Taxonomic assignment of identified eukaryotic protein groups and their relative abundance during the spring bloom 2018.

Table S8. Functional annotation of identified bacterial protein groups and their relative abundance during the spring bloom 2018.

Table S9. Functional annotation of identified phytoplankton protein groups and their relative abundance during the spring bloom 2018.

Table S10. Identified proteins in proteomes of *Muricauda* sp. MAR_2010_75 grow on glucose, laminarin and alginate including CDD annotation, dbCAN2 annotation, locus tags, riBAQ values and fold changes.