Complex algal polysaccharides as substrates

for *Maribacter* strains

Dissertation
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
- Dr. rer. nat. –

Universität Bremen

dem Fachbereich Biologie/Chemie
der Universität Bremen
vorgelegt von

Saskia Kalenborn

Bremen, August 2024

Die vorliegende Doktorarbeit wurde im Rahmen des Programms „International Max Planck Research School of Marine Microbiology (MarMic)" in der Zeit von Januar 2020 bis August 2024 am Max-Planck-Institut für Marine Mikrobiologie angefertigt.

This thesis was prepared under the framework of the International Max Planck Research School of Marine Microbiology (MarMic) at the Max Planck Institute for Marine Microbiology from January 2020 until August 2024.

Gutachter: Prof. Dr. Jens Harder

Gutachterin: Dr. Mia M. Bengtsson

Prüfer: Prof. Dr. Michael W. Friedrich

Prüfer: Prof. Dr. Jens Harder

Prüferin: Dr. Mia M. Bengtsson

Prüferin: Dr. Greta Reintjes

Datum des Promotionskolloquiums: 07.10 2024

"Sometimes the smallest things take up the most room in your heart"

- Winnie the Pooh
- A. A. Milne

Abstract

Marine bacteria play a crucial role in global nutrient cycles, yet much remains to be understood about their ecophysiological niches, metabolic dependencies, and the functions of essential proteins. The research unit FOR 2406 Proteogenomics of Marine Polysaccharide Utilization investigates the mechanisms behind bacterial polysaccharide utilization during marine phytoplankton blooms. The focus is particularly on the functional analyses of marine bacteria within the phylum *Bacteroidota*.

Flavobacteriia, a prominent class within the phylum *Bacteroidota*, are a significant component of marine bacterioplankton. Free-living members of this class play a pivotal role in the degradation of polysaccharides from lysed algae. They often harbour carbohydrate-active enzymes within polysaccharide utilization loci, operon-like genetic regions encoding proteins responsible for the hydrolysis and transport of polysaccharides. Fierce competition among free-living bacteria has led to genomic streamlining, typically focusing on a select few polysaccharides. In contrast, particle-associated bacteria face diverse substrate challenges. Some have the ability to sense and migrate toward nutrient sources, while others reside attached to particulate organic matter (POM), resulting in larger genomes compared to their free-living counterparts. Despite their ecological importance, polysaccharide utilization by particle-associated bacteria has received limited attention.

Maribacter, observed in particle-associated fractions within marine systems, represent a significant gap in physiological studies. To address this gap, we employed proteomics to investigate the complexity of polysaccharide utilization mechanisms in *Maribacter*.

In Chapter 2, we examined *Maribacter* strains that lacked annotated laminarin utilization loci in their genomes. These strains were cultivated in liquid media containing various mono- and polysaccharides, including laminarin, followed by comparative proteomics analysis. Our findings showed that the utilization of laminarin coincided with the induction of an extracellular endo-laminarinase in *Maribacter forsetii*, *Maribacter* sp. MAR_2009_72, *Maribacter* sp. Hel_I_7, and *Maribacter dokdonensis* MAR_2009_60. Interestingly, *Maribacter dokdonensis*

MAR_2009_71 lacked the large endo-laminarinase gene in its genome and exhibited no endo-laminarinase activity. These observations highlight that the *Maribacter* strains investigated in this study are capable of utilizing laminarin, without a co-localization of genes encoding the enzymatic machinery for laminarin utilization.

Chapter 3 explored the utilization of arabinogalactan by *Maribacter*. This investigation was prompted by analysing phytoplankton biomass from the North Sea, which revealed an abundance of arabinose and galactose, the primary components of arabinogalactan. We conducted growth experiments with *Maribacter* strains and employed comparative proteomics to identify expressed proteins associated with arabinogalactan degradation. The proteins identified were predominantly located within polysaccharide utilization loci, providing initial insights into the pathway for arabinogalactan utilization by particle-associated *Maribacter* sp. MAR_2009_72.

In Chapter 4, our aim was to uncover the primary substrates utilized by *Maribacter forsetii*. We found that *M. forsetii* exhibited growth on a wide array of saccharides, including arabinogalactan, chondroitin sulfate, fucoidan, carrageenans, laminarin, and numerous monosaccharides. Surprisingly, we observed growth on substrates previously deemed unusable in an *in-silico* study due to the absence of polysaccharide utilization loci, indicating a broader metabolic versatility than previously thought. Our results suggest that *M. forsetii* expressed a diverse repertoire of glycan-degrading enzymes, some encoded within polysaccharide utilization loci and others dispersed throughout the genome, utilizing various polysaccharides. This research enhances our understanding of *M. forsetii* metabolic capabilities and underscores its potential role in marine polysaccharide degradation.

This comprehensive study shed light on the metabolic capabilities and ecological roles of *Maribacter* species but also highlighted the complex mechanisms behind their polysaccharide utilization strategies, contributing to a broader understanding of marine bacterial physiology and biochemistry.

Zusammenfassung

Marine Bakterien spielen eine entscheidende Rolle in globalen Nährstoffkreisläufen, doch ihre ökologischen Nischen, metabolischen Abhängigkeiten und die Funktionen wesentlicher Proteine sind noch weitgehend unerforscht. Die Forschungsgruppe FOR 2406 "Proteogenomics of Marine Polysaccharide Utilization" untersucht die Mechanismen des bakteriellen Polysaccharid Abbau während mariner Phytoplanktonblüten. Der Schwerpunkt liegt insbesondere auf der funktionellen Analyse von marinen Bakterien innerhalb des Phylums *Bacteroidota*.

Flavobacteriia, eine bedeutende Klasse innerhalb des Phylums *Bacteroidota*, sind ein wesentlicher Bestandteil des marinen Bacterioplankton. Freilebende Mitglieder dieser Klasse spielen eine zentrale Rolle beim Abbau von Polysacchariden aus lysierten Algen. Sie besitzen häufig Kohlenhydrat-aktive Enzyme innerhalb von Polysaccharid Abbau Loci, Operon-artige genetische Region, die Proteine für die Hydrolyse und den Transport von Polysacchariden kodieren. Der intensive Wettbewerb unter freilebenden Bakterien hat zu einer genomischen Rationalisierung geführt, die sich typischerweise auf einige wenige Polysaccharide konzentriert. Im Gegensatz dazu stehen partikelassoziierte Bakterien vor vielfältigen Substratherausforderungen. Einige können Nährstoffe wahrnehmen und sich darauf zubewegen, während andere an partikulärem organischem Material haften, was zu größeren Genomen im Vergleich zu ihren freilebenden Gegenstücken führt. Trotz ihrer ökologischen Bedeutung wurde der Polysaccharid Abbau durch partikelassoziierte Bakterien bisher wenig beachtet.

Maribacter, Teil der partikelassoziierten Fraktion innerhalb mariner Systeme, stellt eine bedeutende Lücke in physiologischen Studien dar. Um diese Lücke zu schließen, haben wir Proteomic eingesetzt, um die Komplexität der Polysaccharid Abbaumechanismen in *Maribacter* zu untersuchen.

In Kapitel 2 untersuchten wir *Maribacter* Stämme, denen annotierte Laminarin Abbau Loci in ihren Genomen fehlten. Diese Stämme wurden in flüssigen Medien mit verschiedenen Mono- und Polysacchariden, einschließlich Laminarin, kultiviert und anschließend einer vergleichenden Proteom Analyse unterzogen. Unsere

Ergebnisse zeigten, dass die Nutzung von Laminarin mit der Induktion einer extrazellulären Endo-Laminarinase in *Maribacter forsetii*, *Maribacter* sp. MAR_2009_72, *Maribacter* sp. Hel_I_7 und *Maribacter dokdonensis* MAR_2009_60 einherging. Interessanterweise fehlte *Maribacter dokdonensis* MAR_2009_71 das Endo-Laminarinase-Gen in seinem Genom und zeigte keine Endo-Laminarinase-Aktivität. Diese Beobachtungen zeigen, dass die in dieser Studie untersuchten *Maribacter* Stämme Laminarin nutzen können, ohne dass die Gene für die enzymatische Maschinerie zum Laminarin Abbau an einem Ort co-lokalisiert sind.

Kapitel 3 untersuchte den Abbau von Arabinogalactan durch *Maribacter*. Diese Untersuchung wurde durch unsere Analyse von Phytoplankton-Biomasse aus der Nordsee angeregt, die reichlich Arabinose und Galactose, die Hauptkomponenten von Arabinogalactan, enthielt. Wir führten Wachstumsexperimente mit *Maribacter* Stämmen durch und verwendeten vergleichende Proteomik, um exprimierte Proteine zu identifizieren, die mit dem Abbau von Arabinogalactan assoziiert sind. Die identifizierten Proteine waren überwiegend in Polysaccharid Abbau Loci lokalisiert, was erste Einblicke in den Arabinogalactan Abbauweg durch den partikelassoziierten Stamm *Maribacter* sp. MAR_2009_72 liefert.

In Kapitel 4 war es unser Ziel, die primären Substrate von *Maribacter forsetii* zu identifizieren. Wir fanden heraus, dass *M. forsetii* auf einer Vielzahl von Polysacchariden, einschließlich Arabinogalactan, Chondroitin Sulfat, Fucoidan, Carrageenen, Laminarin und zahlreichen Monosacchariden, wachsen konnte. Überraschenderweise beobachteten wir Wachstum auf Substraten, die in einer *in-silico* Studie aufgrund des Fehlens von Polysaccharid Abbau Locus als nicht nutzbar galten, was auf eine breitere metabolische Vielseitigkeit hinweist als bisher angenommen. Unsere Ergebnisse deuten darauf hin, dass *M. forsetii* ein vielfältiges Repertoire an glykanspaltenden Enzymen exprimiert, von denen einige in Polysaccharid Abbau Loci kodiert sind, während andere über das Genom verstreut sind, wodurch es ihm ermöglicht wird, verschiedene Polysaccharide abzubauen. Diese Forschung erweitert unser Verständnis der metabolischen Fähigkeiten von

M. forsetii und unterstreicht seine potenzielle Rolle im marinen Polysaccharid Abbau.

Diese umfassende Studie beleuchtet nicht nur die metabolischen Fähigkeiten und ökologischen Rollen von *Maribacter* Stämmen, sondern hebt auch die komplexen Mechanismen hinter ihren Polysaccharidabbaustrategien hervor und trägt zu einem besseren Verständnis der marinen Bakterienphysiologie und -biochemie bei.

Abbreviations

2D gel	two-dimensional polyacrylamide gel
ANI	Average nucleotide identity
ANTS	8-amino-1,3,6-naphthalenetrisulfonate
CAZyme	carbohydrate active enzymes
CAA	casamino acids
CBM	carbohydrate-binding modules
CDD	conserved domain database
CE	carbohydrate esterases
CFU	colony forming units
DAPI	4,6-diamidino-2-phenylindole
DOC	dissolved organic carbon
DOM	dissolved organic matter
EPS	extracellular polymeric substances
FACE	fluorophore-assisted carbohydrate electrophoresis
GH	glycoside hydrolase
GT	glycosyltransferases
Gt	gigatons
FL	free living
LC	liquid chromatography
LFQ	label free quantification
MS	mass spectrometry
OD	optical density
OM	organic matter
ORF	open reading frame
PA	particle-associated
PL	polysaccharide lyase
POC	particulate organic carbon
POM	particulate organic matter
POMPU	Proteogenomics of Marine Polysaccharide Utilization
PUL	polysaccharide utilization locus
PULDB	polysaccharide utilization locus database
Rpm	rounds per minute
Sus	starch utilization system

Content

1. General Introduction

1.1 Marine carbon cycle

Earth is covered to 70% by water, constituting the marine environment a substantial habitat. Numerous biogeochemical cycles including nitrogen, phosphorus, oxygen, sulphur, iron and carbon sustain life in the ocean, all mediated by microorganisms. Primary production by phytoplankton creates the biggest fluxes of carbon and is estimated to deliver roughly 50 Gt of carbon per year (Field *et al.*, 1998; Becker *et al.*, 2020; Kulk *et al.*, 2021). The amount of photosynthetic carbon fixation is controlled by biotic and abiotic factors that can fluctuate from year to year (Mann, 2006; Miller, 2012; Hunter-Cevera *et al.*, 2016). Atmospheric inorganic carbon, mostly CO_2, is fixed by photolithoautotrophic phytoplankton to organic matter (OM), in the photic zone (Figure 1) (Field *et al.*, 1998; Azam and Malfatti, 2007). Besides phytoplankton, other photosynthetic organisms such as macroalgae (seaweeds) and seagrasses also contribute to carbon fixation in the ocean and are important carbon sinks in coastal and nearshore environments (Falkowski *et al.*, 1998; Field *et al.*, 1998; Raven *et al.*, 2011). Dissolved organic matter (DOM) and particulate organic matter (POM) are released as substrate for heterotrophic bacteria by viral lysis, through the so-called viral shunt (Figure 1). DOM includes organic matter that is smaller than 0.45 µm, ranging from simple sugars, amino acids, fatty acids to essential nutrients, such as phosphate and ammonium (Teira *et al.*, 2001; Azam and Malfatti, 2007; Thornton, 2014; Moran *et al.*, 2016). Roughly 3-50% of the phytoplankton primary production is discharged as dissolved organic carbon (DOC) (Moran *et al.*, 2022). Organic matter > 0.45 µm is characterized as POM including planktonic organisms (phytoplankton, zooplankton), decaying organic matter, fecal pellets, and other suspended particles (Azam *et al.*, 1983; Meinhard *et al.*, 2002; Thornton, 2014). The biological pump adds to the carbon sequestration by exporting POM to the bottom of the ocean. The sinking particles are buried in the ocean floor, where they are removed from the active carbon cycle for long periods of time. This sequestration of carbon in deep-sea sediments represents a significant long-term sink for atmospheric CO_2 (Turner, 2015; Boyd *et al.*, 2019; Buesseler *et al.*, 2020). Overall, $1.3×10^{24}$ Gt of carbon are sequestered via the biological pump

(Figure 1) (Nowicki *et al.*, 2022). However, more than half of the DOM is rapidly remineralized by heterotrophic bacteria in the surface water column in the so-called microbial loop (Figure 1) (Azam *et al.*, 1983; Azam, 1998; Buchan *et al.*, 2014; Moran *et al.*, 2016). The bacteria utilize DOM and in turn release inorganic carbon and other nutrients that can be taken up by phytoplankton and other primary producers. Sustaining of the photosynthetic community by the microbial loop supports that zooplankton can graze and bring nutrients and carbon to higher trophic levels, meaning that the microbial loop is a substantial process at which marine life is regulated (Azam and Malfatti, 2007; Buchan *et al.*, 2014; Worden *et al.*, 2015; Moran *et al.*, 2016).
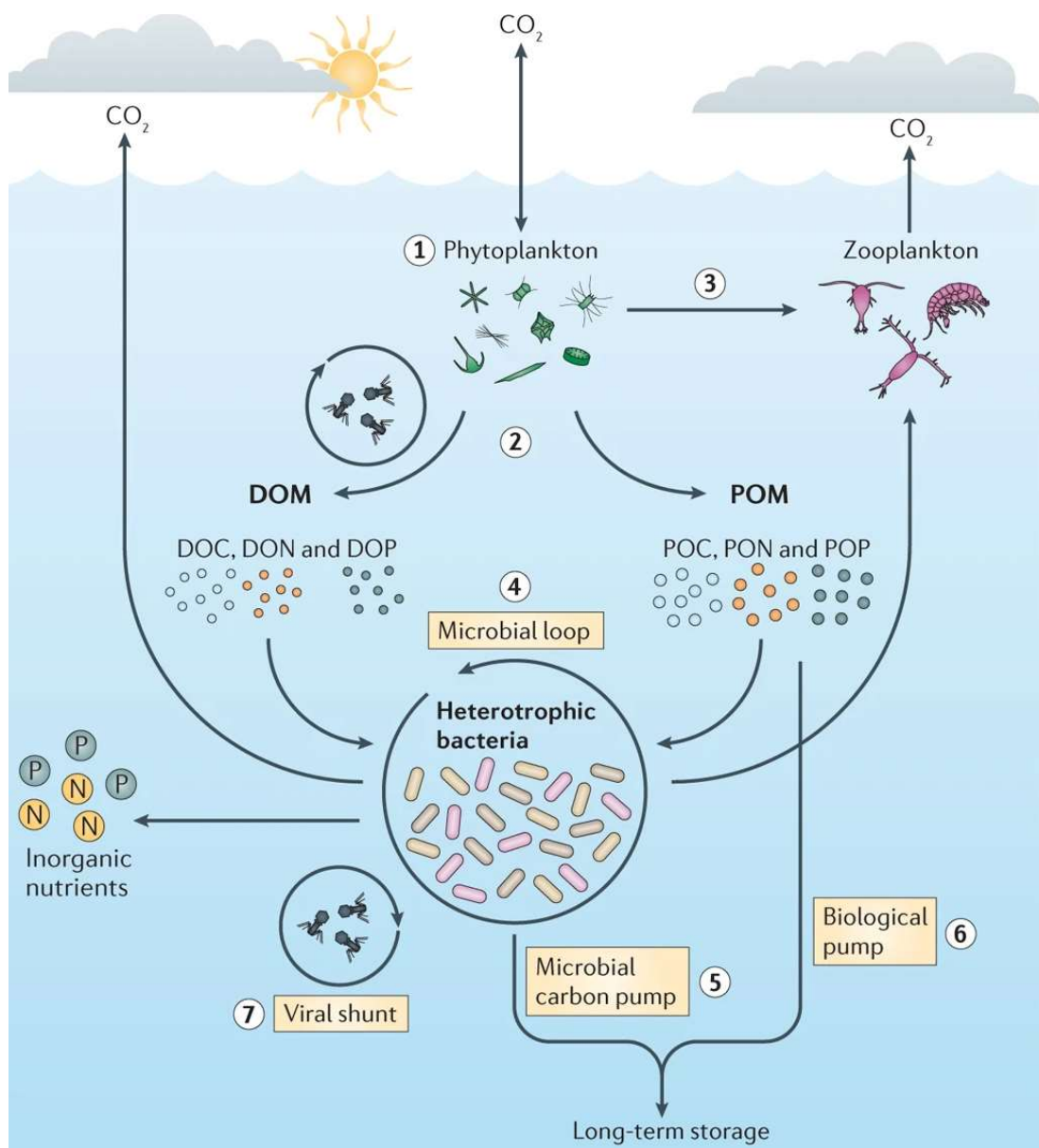
Figure 1: The marine carbon cycle 1: Conversion of inorganic carbon (such as CO_2) to organic carbon by photosynthetic phytoplankton species; 2: Release of dissolved organic matter (DOM) and particulate organic matter (POM) from phytoplankton; 3: Consumption of phytoplankton biomass by zooplankton grazers; 4: Mineralization and recycling of organic matter by diverse heterotrophic bacteria; 5: Microbial carbon pump refers to the transformation of organic carbon into recalcitrant DOC that resists further degradation and is sequestered in the ocean for thousands of years; 6: Biological pump refers to the export of phytoplankton-derived POM from the surface oceans to deeper depths via sinking 7: Viral shunt describes the contributions of viral-mediated cell lysis to the release of dissolved and particulate matter from both the phytoplankton and bacterial pools (Adapted from Buchan *et al.* (2014)). Reproduced with permission from Springer Nature.

1.2 Spring bloom

The German Bright in the North Sea is characterized by its shallow coastal area and significant tidal dynamics, and is renowned as the most significant tidal system globally (Port *et al.*, 2011). This coastal region is considered a eutrophic shelf sea, due to the abundance of essential nutrients derived from land input and bottom sediment release (Ishizaka *et al.*, 1983; Van Dongen-Vogels *et al.*, 2012). Heterotrophic bacteria play a pivotal role in this area, with diverse microbial communities closely linked to organic matter content (Llobet-Brossa *et al.*, 1998). Surface waters typically host approximately 10^6 cells per milliliter, predominantly comprising *Proteobacteria* and *Flavobacteriia* (Whitman *et al.*, 1998). Throughout the year, the North Sea experiences seasonal changes typical of temperate oceans. However, the ongoing mixing process driven by tides suppresses a water column stratification year-round.

The spring bloom is a characteristic seasonal event in temperate North Atlantic and subpolar coastal waters (Mann, 2006; Miller, 2012). In the North Sea, this phenomenon is pronounced, as depicted in Figure 2, showcasing the scale and impact of the 2020 spring bloom. The onset and intensity of the phytoplankton bloom are influenced by various abiotic and biotic factors. Primary abiotic factors are light availability, nutrient levels, and temperature (Mann, 2006; Miller, 2012), while grazing, viral lysis, and phytoplankton physiology constitute key biotic factors (Hunter-Cevera *et al.*, 2016). Typically commencing in March with increased light availability and rising temperatures, the bloom sees a surge in phytoplankton growth, accompanied by a proportional increase in heterotrophic bacteria (Sapp *et al.*, 2007a). The spring bloom can persist for weeks, ceasing when nutrient depletion triggers starvation or predation by viruses and higher organism becomes dominant.
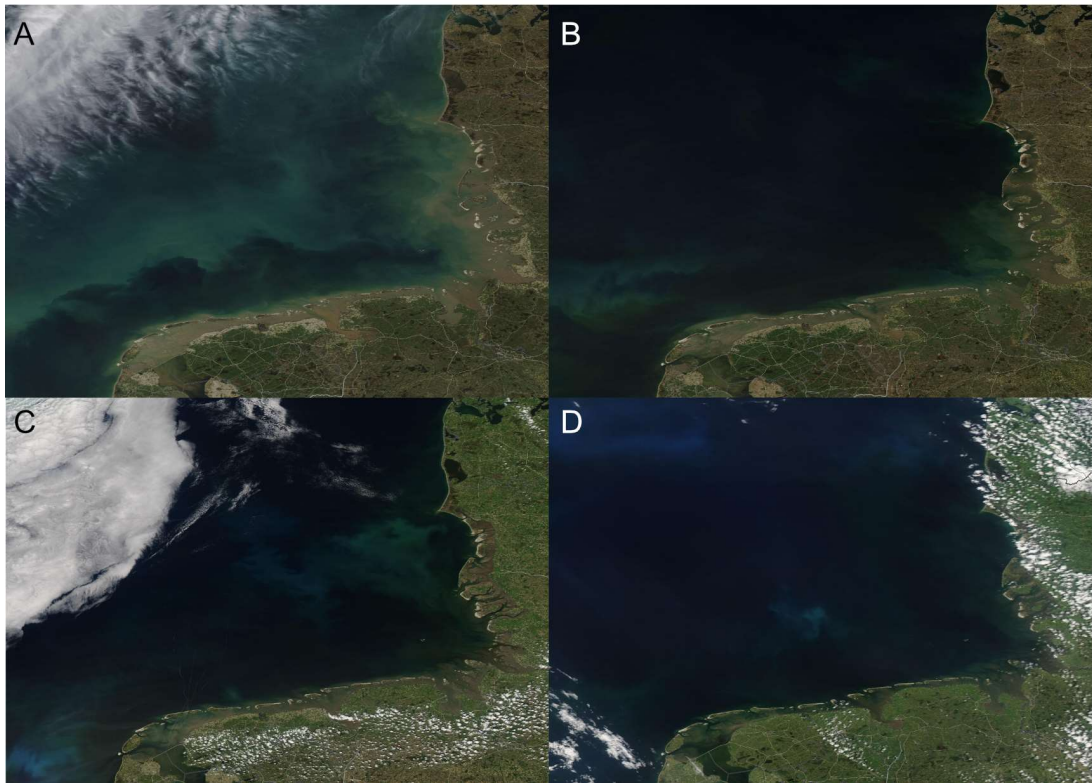
Figure 2: Satellite image of a phytoplankton bloom in the German Bight, North Sea in 2020. A: March 23 2020; B April 22 2020; C: May 29 2020; D: June 26 2020. We acknowledge the use of imagery from the Worldview Snapshots application (https://wvs.earthdata.nasa.gov), part of NASA's Earth Observing System Data and Information System (EOSDIS).

During the bloom a variety of associated organism are witnessed, not only phytoplankton, as well bacteria, viruses and protists grazers (Teeling *et al.*, 2012; Buchan *et al.*, 2014; Teeling *et al.*, 2016; Sidhu *et al.*, 2023; Wang *et al.*, 2024). A clear succession of phytoplankton was observed in all reported blooms (Teeling *et al.*, 2012; Teeling *et al.*, 2016; Sidhu *et al.*, 2023; Wang *et al.*, 2024). Naturally the blooms vary between the years in their phytoplankton and bacterial composition. Most blooms consist of two succeeding blooms. The first bloom usually dominated by diatoms (Teeling *et al.*, 2012; Teeling *et al.*, 2016; Sidhu *et al.*, 2023; Wang *et al.*, 2024). In the 2018 spring bloom the diatom bloom coincided with a evident *Chattonella* bloom (Wang *et al.*, 2024). The cell numbers of the heterotrophic bacteria start to increase with a delay to the increase of diatoms. This increase is fueled by DOM and POM of dead algae or phytoplankton cells. When the diatoms deplete, they are succeeded by *Phaeocystis* for the so called second bloom, which usually transpires in the end of April beginning of May. The number of bacteria stay consistent during this time and increase in mid-May as a response to the

Phaeocystis bloom. Even though the numbers of bacteria stay constant or increase, the community compositions changes throughout the entire bloom (Teeling *et al.*, 2012; Buchan *et al.*, 2014; Teeling *et al.*, 2016; Sidhu *et al.*, 2023; Wang *et al.*, 2024). Responding bacterial clades to diatoms blooms are *Polaribacter*, *Winogradskyella*, *Aurantivirga* (C_MB100;C_MB344), *NS4*, *NS2b*, *SAR 92* (B_MB221) (Sidhu *et al.*, 2023; Wang *et al.*, 2024). For the *Phaeocystis* common responding clades have been *Stappiaceae*, *Sulfitobacter*, *Methylophilaceae*, *SAR 86* and *Pelagibacter* (Sidhu *et al.*, 2023; Wang *et al.*, 2024). There are also clades that are present throughout the entire bloom, only their abundance changes through the bloom, *Formosa*, *Polaribacter*, *NS3a*, *Tenacibaculum*, *Ulvibacter* from Flavobacteriales. Within Gammaproteobacteria *Reinekea*, *SAR 92*, *Alteromonadaceae*, *Colwelliaceae* were present in spring blooms in the North Sea (Teeling *et al.*, 2012; Teeling *et al.*, 2016). Each is constituting 15–25% of the total picoplankton community at specific time points (Teeling *et al.*, 2012).

1.3 Marine polysaccharides

Polysaccharides form the largest fraction in marine organic matter, primarily derived from marine plants, microalgae, macroalgae and some bacteria (Bäumgen *et al.*, 2021). They represent 15-50% of the DOM (Benner *et al.*, 1992; Kraan, 2012). Multitudes of marine organism use these glycans as intracellular energy storage compounds, cell wall constituents or furthermore secreted as extracellular polysaccharides (Arnosti *et al.*, 2021).

1.3.1 Storage polysaccharides

Numerous glucans-based polysaccharides have been disclosed as energy storage in a variety of algae. Laminarin is a well-known storage molecule found in brown algae and diatoms (Rupérez *et al.*, 2002; Alderkamp *et al.*, 2007a; Kroth *et al.*, 2008; Koch *et al.*, 2018). Diatoms, in particular, are estimated to produce up to 12 gigatons of laminarin annually, making laminarin a major molecule in the marine carbon cycle

(Alderkamp *et al.*, 2007a; Becker *et al.*, 2020). Laminarin is a β-1,3 glucan typically composed of approximately 25 glucose molecules, with β-1,6 side chains varying depending on the algae or diatom species (Nelson and Lewis, 1974; Alderkamp *et al.*, 2007a; Rioux *et al.*, 2007; Kadam *et al.*, 2015; Becker and Hehemann, 2018; Becker *et al.*, 2020).

Red algae mainly produce α-polyglucans, consisting of amylose and amylopectin molecules (Shimonaga *et al.*, 2007). Both glucans are essential components of starch, the main storage polysaccharides in plants (Manners and Sturgeon, 1982). Starch as storage molecule has as well been specified for green algae and aforementioned red algae (Yu *et al.*, 2002; Busi *et al.*, 2014; Steinke *et al.*, 2022). Furthermore, glycogen is a storage molecule in animals, fungi and some bacteria (Ball and Morell, 2003). The amount in the marine environment has not been estimated yet, but utilization pathways have been investigated in marine bacteria (Kappelmann *et al.*, 2019; Sidhu *et al.*, 2023). Recently it was demonstrated that some marine *Bacteroidota* strains use excess glucose from utilizing laminarin to synthesize α-glucans as storage compounds (Beidler *et al.*, 2024).

1.3.2 Cell wall polysaccharides

The cell walls of algae are a highly combined structure of biopolymers consisting of proteins, proteoglycans, lipids, polymeric phenolics and polysaccharides (Figure 3) (Synytsya *et al.*, 2015; Domozych, 2019). A variety of cell wall polysaccharides have been reported depending on the species of algae, the anatomical part, the life cycle stage as well as the season and habitat (Haug *et al.*, 1974; Falshaw and Furneaux, 1998; Saraswathi *et al.*, 2003; Jothisaraswathi *et al.*, 2006; Guibet *et al.*, 2008; Li *et al.*, 2008; Domozych *et al.*, 2012). Green algae predominantly incorporate sulfated and carboxylic polysaccharides, such as ulvan and pectin in their cell wall (Ray and Lahaye, 1995; Lee *et al.*, 2010; Domozych *et al.*, 2012; Synytsya *et al.*, 2015). Red algae are known for cellulose, xylans, mannan fibrils and sulfated galactans (Cardoso *et al.*, 2007; Ghosh *et al.*, 2009; Synytsya *et al.*, 2015) and brown algae include cellulose fibrils, alginate and fucoidan in their cell wall (Popper *et al.*, 2011; Synytsya *et al.*, 2015; Schultz-Johansen *et al.*, 2018; Domozych, 2019).
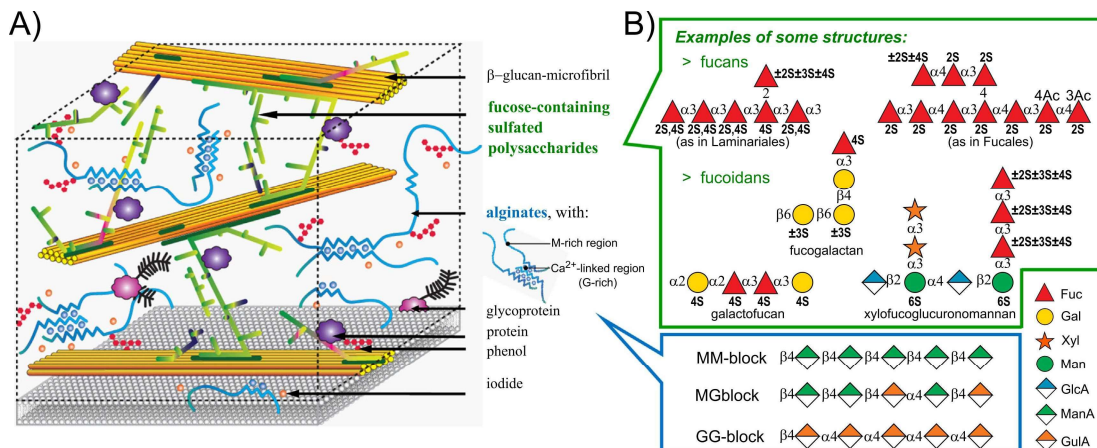
Figure 3: Schematic portrayal of the extracellular matrix (ECM) model in brown algae with key structures of FCSPs and alginates; A) ECM model for brown algae from the order *Fucales*. Alginates and FCSPs are the main constituents. Proteins and halide compounds also are present. Glycoproteins related to arabinogalactan-proteins have been reported in brown algae, but no detailed chemical characterization in these models is described so far. B) Examples of typical alginates and FCSP motifs are shown. (Mazéas *et al.*, 2023)

Arabinogalactan, more precisely arabinogalactan proteins have been confirmed in cell walls of brown algae and green seaweeds (Ciancia *et al.*, 2007; Hervé *et al.*, 2016; Fujita *et al.*, 2019; Leszczuk *et al.*, 2023; Ma and Johnson, 2023). These compounds resemble type II arabinogalactan with a complex backbone of β-1,3-linked galactans and β-1,6-linked galactans side chains with decorations such as rhamnose, fucose, mannose, xylose and glucuronic acid (Fujita *et al.*, 2019; Villa-Rivera *et al.*, 2021; Leszczuk *et al.*, 2023). In a recent study arabinogalactan was detected in the high molecular weight dissolved organic matter and particulate organic carbon (POC) fraction by monoclonal antibodies during the spring bloom in the North Sea (Vidal-Melgosa *et al.*, 2021). This matches previous studies measuring arabinose and galactose during the *Phaeocystis* bloom (Alderkamp *et al.*, 2007b; Sato *et al.*, 2018).

Fucoidan is another prominent cell wall polysaccharide found in brown algae (Ale *et al.*, 2011; Castro *et al.*, 2014; Schultz-Johansen *et al.*, 2018; Huang *et al.*, 2021; Vidal-Melgosa *et al.*, 2021; Zhang *et al.*, 2024). The structure of fucoidan is incredibly complex and varies between the species of brown algae (Figure 3b) (Berteau and Mulloy, 2003; Li *et al.*, 2008; Ale *et al.*, 2011; Castro *et al.*, 2014; Dobrinčić *et al.*, 2020; Sichert *et al.*, 2020). Fucoidan makes up roughly 3-28% of brown algae dry weight (Li *et al.*, 2021). Two types of fucan can be present in the cell walls one is homofucan, either a α-1,3 or alternating α-1,3/1,4 linked fucose

sulfated at the O-2, O-3 or O-4 position (Synytsya *et al.*, 2015; Schultz-Johansen *et al.*, 2018), or heterofucan, consisting of mannose, galactose and glucuronic acid with sulfated fucose branches (Chevolot *et al.*, 2001; Cong *et al.*, 2016).

Carrageenans are part of the cell wall of red algae (Usov, 1992; Falshaw and Furneaux, 1998; Kumar *et al.*, 2011; Yang *et al.*, 2011; Domozych, 2019). Each red algae species contains a different composition and it has been described that most carrageenans are not present in their pure type, mostly mixes of kappa and iota carrageenan (van de Velde *et al.*, 2001).

Not only algae harbor polysaccharides in their cell wall. Chondroitin sulphate, a polysaccharide present in the North Sea, is commonly found in the body wall of sea cucumbers (Vieira *et al.*, 1991; Vasconcelos and Pomin, 2017).

Each type of algae possesses specific polysaccharides in its cell wall, contributing to its unique structural features. However, many cell wall polysaccharides remain so far unidentified or poorly described due to their structural complexity (Synytsya *et al.*, 2015; Domozych, 2019).

1.3.3. Extracellular polysaccharides

The discovery of excreted products by algae dates back to 1917 (Harder, 1917). Further insights into these substances emerged in 1931 when they were described as unknown chemical nature acting as nutrient for microorganisms (Akehurst, 1931). In later years it was discovered that phytoplankton release organic substrates through their entire life cycle. Among these substrates, proteins, nucleic acids, lipids, small molecules of different kinds and polysaccharides were identified, so called extracellular polymeric substances (EPS) (Hellebust, 1974; Mague *et al.*, 1980; Fogg, 1983; Myklestad, 1995). EPS produced by microalgae and cyanobacteria often appear as mucous masses surrounding individual cells or cell groups, with extracellular proteins being the most prevalent form (Sutherland, 1982; Hoagland *et al.*, 1993; De Philippis and Vincenzini, 1998; Wotton, 2004). They can make up to 40-95% of the total EPS amount (Myklestad, 1995; González-Hourcade *et al.*, 2020). These polysaccharides are highly complex based on their monosaccharide composition. Some red algae species generate polysaccharides

with 8 to 10 different monomers, similar numbers were reported for diatoms (Abdullahi *et al.*, 2006; Rossi and De Philippis, 2016). In most cases, glucose, arabinose, galactose, mannose and fucose are the major components (Raveendran *et al.*, 2013; Rossi and De Philippis, 2016). In addition to these monosaccharides, extracellular polysaccharides may contain methylated groups, sulfate groups, and uronic acids (Allard and Casadevall, 1990; Urbani *et al.*, 2005). The production of EPS requires large amounts of the photosynthesis system and subcellular secretion machinery (Domozych and LoRicco, 2023). Recently it was stated that brown algae, here *Fucus vesiculosus*, secretes fucoidan as 0.3% of its daily biomass and takes reportedly part in the carbon sequestration of the ocean (Buck-Wiese *et al.*, 2023).

1.4 Polysaccharide utilization

A variety of enzymes are necessary for the breakdown of polysaccharides, with some located extracellularly, others within the outer or inner membrane, and some localized in the periplasm (Figure 4).
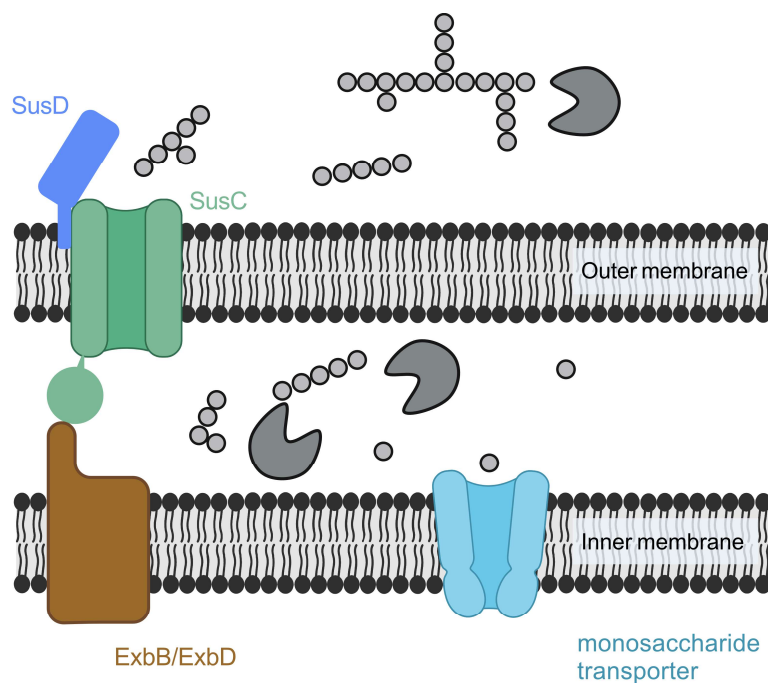
Figure 4: A model of polysaccharide uptake into the periplasm via TonB-dependent outer membrane transporters. Polysaccharides are degraded into oligosaccharides transported in the periplasm and further utilized into monosaccharides and then transported into the cytoplasm. Created with BioRender.com.

The utilization of polysaccharides by *Flavobacteriia* was initially described for the starch utilization system (Sus) in *Bacteroides thetaiotaomicron* (Shipman *et al.*, 2000). This system involves two outer membrane proteins, SusC and SusD, along with glycoside hydrolases situated in both the outer membrane and the periplasm. SusC-like and SusD-like proteins facilitate transport across the outer membrane (Reintjes *et al.*, 2017). Both SusC and SusD are responsible for binding oligosaccharides (McBride *et al.*, 2009). SusD-like proteins specifically bind to oligomeric carbohydrates, transporting them to SusC. Before SusD binds the oligomeric sugars, the polysaccharides undergo hydrolysis. SusC-like proteins, belonging to the TonB receptor-like protein family, facilitate the transportation of sugars through the outer membrane via its β-barrel channel (McBride *et al.*, 2009; Martens *et al.*, 2011). However, to accomplish this, interaction with the inner membrane protein complex consisting of ExbB, ExbD, and TonB is essential. This complex is powered by proton motive force (Figure 5) (Noinaj et al., 2010).
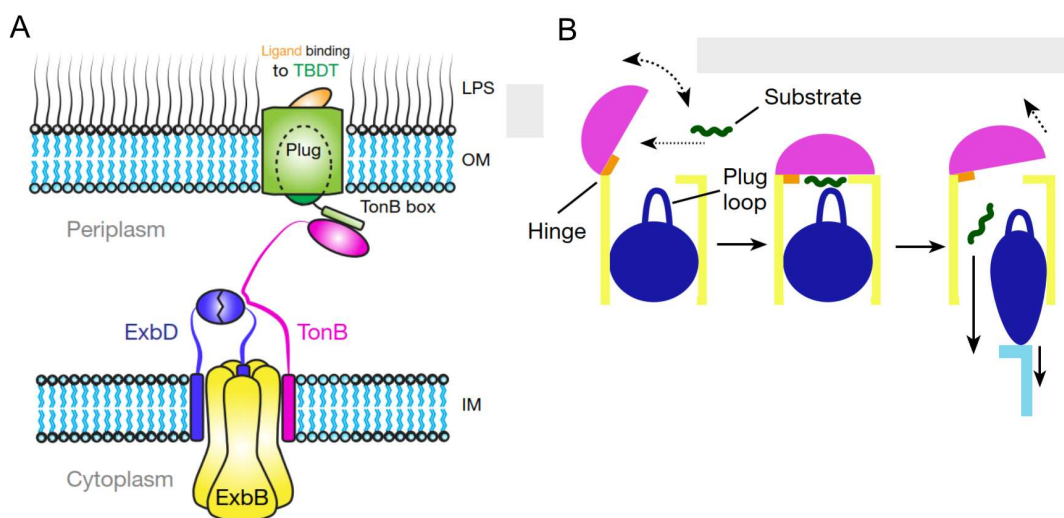
Figure 5: Structural model of the Ton complex consisting of ExbB, ExbD, and at least one TonB. LPS, lipopolysaccharide ExbB, ExbD, and TonB form the energy source for the outer membrane transporter (A); General mechanism for nutrient uptake by SusCD transport complexes. Ligand is shown as a green wavy line and the C-terminal domain of TonB that interacts with the SusC plug is cyan (B) (Celia et al., 2016; Glenwright et al., 2017). Reproduced with permission from Springer Nature.

The complex operates via a plug and loop mechanism. Upon substrate binding by the TonB-dependent transporter, rotation of ExbB and ExbD initiates, pulling the plug and enabling substrate movement into the periplasm. The TonB box plays a crucial role in this transport process, situated near the N-terminus of the plug

domain and typically concealed within the β-barrel domain in the absence of an oligosaccharide. Upon oligosaccharide binding, conformational changes expose the TonB box, facilitating interaction with other components of SusC to displace the plug (Noinaj *et al.*, 2010).

Carbohydrate-active enzymes (CAZymes) participate in the degradation of polysaccharides with high specificity (Bäumgen *et al.*, 2021; Drula *et al.*, 2022). CAZymes are classified into five main groups: glycoside hydrolases (GH), glycosyltransferases (GT), polysaccharide lyases (PL), carbohydrate esterases (CE), and carbohydrate-binding modules (CBM). GHs catalyze the hydrolysis of glycosidic bonds, while GTs facilitate the transfer of sugar moieties between molecules. PLs catalyze the cleavage of glycosidic bonds through β-elimination, whereas CEs cleave O- and N-acetyl groups from polysaccharides (Davies *et al.*, 2005). CBMs recognize specific carbohydrate structures, aiding in substrate recognition and binding for the aforementioned enzymes. The diverse functions of CAZymes are pivotal in various biological processes, including digestion, host-microbe interactions, and the degradation of plant cell walls (Cantarel *et al.*, 2009; Bäumgen *et al.*, 2021; Drula *et al.*, 2022).

In bacteria belonging to the *Bacteroidota* phylum, enzymes involved in polysaccharide utilization are often arranged within structures called polysaccharide utilization loci (PULs) (Grondin *et al.*, 2017). These PULs are operon like genetic regions encoding proteins like SusC and SusD, various CAZymes, and other enzymes necessary for breaking down specific glycans, such as sulfatases (Helbert, 2017).

1.5 Particle-associated bacteria

The distribution of organic substrates is unevenly in the ocean. Particles, such as lysed algae cells, are time sensitive nutrient rich hot spots (Azam, 1998; Meinhard *et al.*, 2002; Stocker, 2012). Bacteria have developed different life styles to adapt to their surroundings. Free-living (FL) bacteria are considered non motile (Koch, 2001). They are smaller in size, < 3 µm, as well having a smaller genome and therefore a smaller range of substrates (Turley and Stutt, 2000; Koch, 2001;

Meinhard *et al.*, 2002; Lauro *et al.*, 2009; Smith *et al.*, 2013; Heins *et al.*, 2021a; Heins *et al.*, 2021b). FL bacteria have adapted to low concentrations of nutrients and have a high expression of membrane transporter genes (Lauro *et al.*, 2009; Kappelmann *et al.*, 2019; Heins *et al.*, 2021a; Heins *et al.*, 2021b). They present the majority of the bacterial community in the environment (Turley and Stutt, 2000; Koch, 2001; Lauro *et al.*, 2009; Heins *et al.*, 2021a; Heins *et al.*, 2021b; Wang *et al.*, 2024). Furthermore it has been estimated that FL bacteria conciliate 53% of DOM fraction (Giering *et al.*, 2014). Some free-living bacteria have chemotactic sensing and use flagella to move towards particles (Stocker *et al.*, 2008; Grossart, 2010; Stocker and Seymour Justin, 2012). These specific FL bacteria have slightly larger genomes in correlation to the non-motile FL bacteria, as a result of encoding the genes for sensoring and motility (Stocker *et al.*, 2008; Grossart, 2010; Stocker and Seymour Justin, 2012; Seymour *et al.*, 2017).

Bacteria considered being particle associated (PA) are captured on filters above 3 µm (Heins *et al.*, 2021a; Heins *et al.*, 2021b; Heins and Harder, 2023; Wang *et al.*, 2024). They are not only larger in size than FL bacteria, as well their genomes are larger (Smith *et al.*, 2013). PA bacteria encode more polysaccharide-utilizing enzymes and have a higher respiration rate (Kogure *et al.*, 1981; Acinas *et al.*, 1999; Crump *et al.*, 1999; Meinhard *et al.*, 2002; Grossart *et al.*, 2003; Grossart *et al.*, 2007; Smith *et al.*, 2013; Kappelmann *et al.*, 2019). Many PA bacteria have the ability to glide. A variety of systems such as the secretion system, the generation of cell surface waves as well as "motor" proteins on the cell surface have been identified to be responsible for gliding motility (McBride, 2001; Braun *et al.*, 2005; Jarrell and McBride, 2008; Nelson Shawn *et al.*, 2008; Nan and Zusman, 2016). PA bacteria are in general low in abundance in bacterial communities as a result of infrequency particle presence (Alldredge *et al.*, 1986; Heins *et al.*, 2021b). Mostly they constitute less than 10% of the community (Alldredge *et al.*, 1986; Rieck *et al.*, 2015; Heins *et al.*, 2021b), with rare instances reaching 32% (Ghiglione *et al.*, 2007). Besides particle degradation, they also participate in particle aggregation, promoting carbon sequestration and remineralization (Azam and Malfatti, 2007; Duarte *et al.*, 2022; Reintjes *et al.*, 2023).

The phycosphere, a mucous region rich in organic matter surrounding phytoplankton cells, serves as a habitat for PA and FL motile bacteria (Sapp *et al.*,

2007b; Seymour *et al.*, 2017), akin to the rhizosphere of plants. In 1972 the phycosphere was described as a space where growth of bacteria is enhanced by extracellular products of algae (Bell and Mitchell, 1972). The main community is comprised of *Gammaproteobacteria*, *Alphaproteobacteria*, *Actinobacteria*, *Betaproteobacteria*, *Bacilli* and *Bacteroidetes* (Nicolas *et al.*, 2004; Sapp *et al.*, 2007b; Amin *et al.*, 2012; Natrah *et al.*, 2014). The interaction between bacteria and algae can be can be symbiotic or detrimental. Bacteria can secrete growth promoters that improve the growth of algae (Variem and Kizhakkedath, 2021). Excreted organic carbon of algae can as well have positive stimulatory effect on bacteria (Natrah *et al.*, 2014). Some secreted bacterial compounds can cause lysis of algae, releasing more DOM to the environment and feeding the microbial loop (Amin *et al.*, 2012; Natrah *et al.*, 2014; Variem and Kizhakkedath, 2021).

1.6 *Maribacter*

Flavobacteriia, a class within the *Bacteroidota* phylum, comprises microorganisms of the order *Flavobacteriales* and the family *Flavobacteriaceae* (Bernardet *et al.*, 1996; Bernardet *et al.*, 2002). The name *Flavobacteriia* stems from the Latin word "flavus," meaning yellow, reflecting the yellow to light orange coloration of most colonies on solid medium (Figure 6). These gram-negative rods, typically 5 µm long, are nonspore-forming chemoorganoheterotrophs, often exhibiting gliding motility.
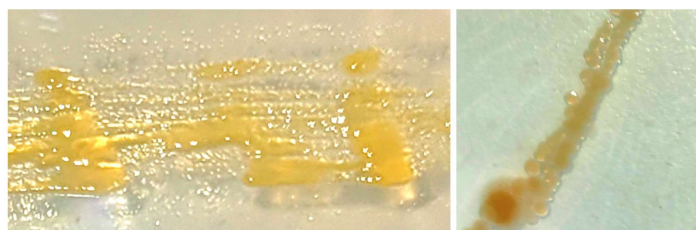
Figure 6: *Maribacter* strains grown on 2216 marine agar; left picture *Maribacter* sp. MAR_2009_72; right picture *Maribacter forsetii* DSM 18668[T]

Predominantly aerobic, *Flavobacteriia* utilize organic compounds as electron donors for energy production and biomass formation, while oxygen or nitrogen serves as an electron acceptor (Kirchman, 2002; Rosenberg *et al.*, 2014). Within

Flavobacteriia, the genus *Maribacter*, established in 2004, comprises strictly aerobic, non-motile, rod-shaped bacteria (Nedashkovskaya *et al.*, 2004). *Maribacter* strains were infrequently isolated from seawater (Heins and Harder, 2023; Lu *et al.*, 2023), and molecular community analyses revealed their sparse presence in free-living bacterial populations (Sidhu *et al.*, 2023). While surface populations of oxic sandy sediment exhibited abundances of up to 4% (Probandt *et al.*, 2018; Miksch *et al.*, 2021), higher relative abundances were noted in phytosphere populations associated with both micro- and macroalgae (Heins *et al.*, 2021a; Lu *et al.*, 2023; Wang *et al.*, 2024). Furthermore, apart from being a dominant epiphyte on various algae, *Maribacter* plays a role in *Ulva morphogenesis* through the synthesis of thallusin, underscoring the intricate interactions between particle-associated bacteria and algae within the phycosphere (Alsufyani *et al.*, 2020).

Currently 31 validly described species and more than 100 genomes have been published. Here two strains from seawater - *Maribacter forsetii* DSM 18668^T (Barbeyron *et al.*, 2008) and *Maribacter* sp. Hel_I_7 from the long-term ecological research station Kabeltonne near Helgoland – and three strains from phytoplankton catches in the Wadden Sea near Sylt - *Maribacter* sp. MAR_2009_72, *Maribacter dokdonensis* MAR_2009_60 and *Maribacter dokdonensis* MAR_2009_71 were investigated (Table 1) (Hahnke and Harder, 2013).

Table 1: Overview *Maribacter* genomes

Maribacter strain	DSMZ	Site	Type	Time	Literature	Mb	GC%	Genes
Maribacter forsetii	18668	Helgoland; Kabeltonne	surface seawater	Aug-99	Barbeyron *et al.* (2008)	4.51	35.2	3890
Maribacter dokdonensis MAR_2009_60	29383	Sylt, List	20µm phytoplankton	Oct-09	Hahnke and Harder (2013)	4.51	36.1	3914
Maribacter dokdonensis MAR_2009_71	29385	Sylt, List	20µm phytoplankton	Oct-09	Hahnke *et al.* (2015)	4.62	36.1	3955
Maribacter sp. MAR_2009_72	29384	Sylt, List	20µm phytoplankton	Oct-09	Hahnke and Harder (2013)	4.35	37.7	3696
Maribacter sp. Hel_I_7	29657	Helgoland; Kabeltonne	surface seawater	2010	Hahnke and Harder (2013)	4.78	34.8	4182

1.7 Cultivation

Cultivation is a major part of microbiology, growing microorganisms to study their behaviors and mechanism. In the 1880's the concept of pure cultures was developed by Koch and colleagues (Blevins and Bronze, 2010). Furthermore Robert Koch, Walther Hesse, Fanny Hesse and Julian Richard Petri innovated a liquid medium contain agar, so called broth, and glass made Petri dishes, filled with sterile solid medium (Petri, 1887; Hesse, 1992). Both innovations are still to this day a vital part in cultivation. The development of media went through many stages, one being that C.E. Zobell crafted a standardized medium for cultivating marine bacteria. Self-made peptone was replaced with commercial Bactopeptone, additionally he supplemented the medium with phosphate (K_2HPO_4) and iron as ferric phosphate ($Fe(III)PO_4$). This secured reproducibility and an increase in cell numbers (Zobell, 1941). He later modified the medium further by adding yeast extract (Oppenheimer and Zobell, 1952). This medium is still commercially available as marine agar and broth 2216 (Difco, BD diagnostics systems, USA). In later years Norbert Pfennig invented the vitamin solution as well as the trace element solution as an addon for media, both increased growth of the cells (Pfennig, 1965; Pfennig and Lippert, 1966). In 1974, he and colleagues reported that when using the vitamin

solution there is no need for yeast extract in the medium (Pfennig, 1974). Both solutions went through many iterations to adjust the concentration for the best growth improvement of bacteria (Widdel and Pfennig, 1977; Pfennig, 1978; Widdel and Pfennig, 1981).

The measurement of cultivation efficiency was counting cells (CFU, colony forming units) and later when microscopes were available direct cell counts (ZoBell, 1946). Sometime after the direct cell count was implemented a discrepancy between the direct cell count and CFU numbers was exposed. Direct cell count showed always much higher numbers then CFU's. Scientists of the time discovered that only 1% or less of microorganism can be cultivated on plates, an observation nowadays described as the "great plate anomaly" (Jannasch and Jones, 1959; Hobbie *et al.*, 1977; Kogure *et al.*, 1979). The majority of bacteria in the environment, roughly 99%, are difficult to cultivate on plates or not culturable at all (Amann *et al.*, 1995; Rappé and Giovannoni, 2003; Pham and Kim, 2012). The media and culture conditions cannot recreate the defined growth parameters for most bacteria (Jannasch and Jones, 1959; Hobbie *et al.*, 1977; Kogure *et al.*, 1979; Rappé and Giovannoni, 2003). The cultivable portion comprises preferentially PA bacteria (Heins and Harder, 2023).

In recent years many methods for the cultivation of so far "uncultured" microorganism have been developed. Some use nutrient poor media and other specified the media to target the growth of certain clades (Connon and Giovannoni, 2002; Rappé *et al.*, 2002; Hahn *et al.*, 2004; Watanabe *et al.*, 2009; Kasalický *et al.*, 2013). The improvement of media and cultivation conditions is still an ongoing process. Even though many steps have been taken to enhance the cultivation unculturable microorganisms, a lot of effort has to be made to explore the diversity of microbes (Schut *et al.*, 1993; Connon and Giovannoni, 2002; Song *et al.*, 2009; Carini *et al.*, 2013; Overmann *et al.*, 2017).

1.8 Proteomics

Proteomics is the study of proteins, in large scale determining the gene and cellular function. Proteins were named in 1838 by Berzelius, from the Greek word "proteius"

meaning " the first rank" (Cristea *et al.*, 2004). The main development for analytical protein chemistry happened in the 1980s and early 1990s. Traditionally proteomics was associated with showing high number of proteins on a two-dimensional polyacrylamide gel (2D gel), creating master maps of organisms to examine changes in expression patterns (Anderson and Anderson, 1996; Celis *et al.*, 1996; Wilkins *et al.*, 1996b; Pandey and Mann, 2000). 2D gels were innovated in the 1970s, but they lacked sensitivity and the determination of proteins was difficult (O'Farrell, 1975; Pandey and Mann, 2000; Cox and Mann, 2007). Proteomics changed when electrospray ionization and matrix assisted laser desorption/ionization techniques were discovered for mass spectrometry (MS) (Karas and Hillenkamp, 1988; Fenn *et al.*, 1989; Cody, 1997; Pandey and Mann, 2000; Aebersold and Goodlett, 2001; Aebersold and Mann, 2003; Cox and Mann, 2007). John B. Fenn received the Nobel prize in Chemistry of 2002 for the discovery of electrospray ionization in MS. Measurement of molecules with a molecular weight of > 10,000 mass units were immediately possible (Karas and Hillenkamp, 1988). During this time, in 1994, the word proteome came to live, which was intended to describe the overall protein content of a cell characterized in regards to its location, interaction and function at a certain time point (Swinbanks, 1995; Wilkins *et al.*, 1996b; Wilkins *et al.*, 1996a). Over the years a lot of progress has been made to use smaller volumes of biological samples, to immediately identify peptides by matching their MS fragmentation spectrum to a protein sequence database and as well the direct analysis of complex protein mixes (Aebersold and Goodlett, 2001; Aebersold and Mann, 2003; Yates *et al.*, 2005; Cox and Mann, 2011). Combining techniques such as liquid chromatography with MS/MS even increased the ability to measure complex peptide mixes (Appella *et al.*, 1995; Yates *et al.*, 1997; Chen *et al.*, 2007). All these new developments helped to better proteome coverage as well as enhance purification, characterization, quantification, sequence, structural and bioinformatical analysis of proteins for all types of organisms (Cho, 2007; Cox and Mann, 2011; Aslam *et al.*, 2017).

1.9 Aims

The primary focus of this thesis was to investigate the physiology and distinct polysaccharide utilization pathways of the particle-associated genus *Maribacter*. Specifically, the research aimed to evaluate the growth of *Maribacter* species in the presence of various glycans. This was followed by comparative proteomics to identify key players involved in polysaccharide utilization, and their respective locations and functions within the cells.

One of the glycans investigated in this study was laminarin. Previous research had indicated that no PUL for laminarin utilization was identified in the included *Maribacter* strains (Kappelmann *et al.*, 2019). Despite this, growth experiments revealed that all five *Maribacter* strains exhibited growth in the presence of laminarin. This observation prompted a deeper investigation into how these strains utilize laminarin using comparative proteomics. The primary objective was to pinpoint key enzymes involved in laminarin utilization by analyzing expression data (Chapter 2).

A major task of this thesis involved analyzing plankton biomass from the North Sea to identify substrates and discover novel PULs in *Maribacter* strains. One notable finding from this analysis was the identification of arabinogalactan as a substrate, as its degradation had not been previously studied in marine bacteria. This led to an investigation of the growth behavior of *Maribacter* strains in the presence of arabinogalactan, and a proteomic exploration of the strains that demonstrated significant growth (Chapter 3).

Bacteria belonging to the phylum *Bacteroidota* possess multiple PULs in their genomes, with many key substrates yet to be defined. This chapter aimed to investigate the substrate range of *Maribacter forsetii* to gain a better understanding of its role in the environment. To achieve this, *M. forsetii* was grown under 17 different conditions, and comparative proteomics was performed to identify enzymes expressed in correlation to specific substrates and PULs (Chapter 4).

This comprehensive approach was expected to shed light on the metabolic capabilities and ecological roles of *Maribacter* species but also highlight the intricate mechanisms behind their polysaccharide utilization strategies, contributing to a broader understanding of marine bacterial physiology and biochemistry.

2 References

Abdullahi, A.S., Underwood, G.J.C., and Gretz, M.R. (2006) Extracellular matrix assembly in diatoms (*Bacillariophyceae*): Environmental effects on polysaccharides synthesis in the model diatom *Phaeodactylum tricornutum*. *Journal of Phycology* **42**: 363-378.

Acinas, S., G., Antón, J., and Rodríguez-Valera, F. (1999) Diversity of free-living and attached bacteria in offshore Western Mediterranean waters as depicted by analysis of genes encoding 16S rRNA. *Applied and Environmental Microbiology* **65**: 514-522.

Aebersold, R., and Goodlett, D.R. (2001) Mass spectrometry in proteomics. *Chemical Reviews* **101**: 269-296.

Aebersold, R., and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* **422**: 198-207.

Akehurst, S.C. (1931) XII.—Observations on pond life, with special reference to the possible causation of swarming phytoplankton. *Journal of the Royal Microscopical Society* **51**: 237-265.

Alderkamp, A.-C., Van Rijssel, M., and Bolhuis, H. (2007a) Characterization of marine bacteria and the activity of their enzyme systems involved in degradation of the algal storage glucan laminarin. *FEMS Microbiology Ecology* **59**: 108-117.

Alderkamp, A.-C., Buma, A.G.J., and van Rijssel, M. (2007b) The carbohydrates of *Phaeocystis* and their degradation in the microbial food web. *Biogeochemistry* **83**: 99-118.

Ale, M.T., Maruyama, H., Tamauchi, H., Mikkelsen, J.D., and Meyer, A.S. (2011) Fucose-containing sulfated polysaccharides from brown seaweeds inhibit proliferation of melanoma cells and induce apoptosis by activation of caspase-3 in vitro. *Marine Drugs* **9**: 2605-2621.

Allard, B., and Casadevall, E. (1990) Carbohydrate composition and characterization of sugars from the green microalga *Botryococcus braunii*. *Phytochemistry* **29**: 1875-1878.

Alldredge, A.L., Cole, J.J., and Caron, D.A. (1986) Production of heterotrophic bacteria inhabiting macroscopic organic aggregates (marine snow) from surface waters. *Limnology and Oceanography* **31**: 68-78.

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* **71**: 3340-3349.

Amann, R.I., Ludwig, W., and Schleifer, K.H. (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological Reviews* **59**: 143-169.

Amin, S.A., Parker, M.S., and Armbrust, E.V. (2012) Interactions between diatoms and bacteria. *Microbiology and Molecular Biology Reviews* **76**: 667-684.

Anderson, N.G., and Anderson, N.L. (1996) Twenty years of two-dimensional electrophoresis: Past, present and future. *Electrophoresis* **17**: 443-453.

Appella, E., Padlan, E.A., and Hunt, D.F. (1995) Analysis of the structure of naturally processed peptides bound by class I and class II major histocompatibility complex molecules. *EXS* **73**: 105-119.

Arnosti, C., Wietz, M., Brinkhoff, T., Hehemann, J.H., Probandt, D., Zeugner, L., and Amann, R. (2021) The biogeochemistry of marine polysaccharides: Sources, inventories, and bacterial drivers of the carbohydrate cycle. *Annual Review of Marine Science* **13**: 81-108.

Aslam, B., Basit, M., Nisar, M.A., Khurshid, M., and Rasool, M.H. (2017) Proteomics: Technologies and their applications. *Journal of Chromatographic Science* **55**: 182-196.

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

Azam, F., and Malfatti, F. (2007) Microbial structuring of marine ecosystems. *Nature Reviews Microbiology* **5**: 782-791.

Azam, F., Fenchel, T., Field, J.G., Gray, J.S., Meyer-Reil, L.A., and Thingstad, F. (1983) The ecological role of water-column microbes in the sea. In *Classic Papers with Commentaries*. Miller, T.E., and Travis, J. (eds): University of Chicago Press, pp. 384-390.

Ball, S.G., and Morell, M.K. (2003) From bacterial glycogen to starch: Understanding the biogenesis of the plant starch granule. *Annu Rev Plant Biol* **54**: 207-33.

Barbeyron, T., Carpentier, F., Haridon, S., Schüler, M., Michel, G., and Amann, R. (2008) Description of *Maribacter forsetii* sp. nov., a marine *Flavobacteriaceae* isolated from North Sea water, and emended description of the genus *Maribacter*. *International Journal of Systematic and Evolutionary Microbiology* **58**: 790-797.

Bäumgen, M., Dutschei, T., and Bornscheuer, U.T. (2021) Marine polysaccharides: Occurrence, enzymatic degradation and utilization. *ChemBioChem* **22**: 2247-2256.

Becker, S., and Hehemann, J.-H. (2018) Laminarin quantification in microalgae with enzymes from marine microbes. *Bio-protocol* **8**: e2666.

Becker, S., Tebben, J., Coffinet, S., Wiltshire, K., Iversen, M.H., Harder, T. *et al.* (2020) Laminarin is a major molecule in the marine carbon cycle. *Proceedings of the National Academy of Sciences* **117**: 6599-6607.

Beidler, I., Steinke, N., Schulze, T., Sidhu, C., Bartosik, D., Krull, J. *et al.* (2024) Alpha-glucans from bacterial necromass indicate an intra-population loop within the marine carbon cycle. *Nature Communications* **15**: 4048.

Bell, W., and Mitchell, R. (1972) Chemotactic and growth responses of marine bacteria to algal extracellular products. *The Biological Bulletin* **143**: 265-277.

Benner, R., Pakulski, J.D., McCarthy, M., Hedges, J.I., and Hatcher, P.G. (1992) Bulk chemical characteristics of dissolved organic matter in the ocean. *Science* **255**: 1561-1564.

Bernardet, J.-F., Nakagawa, Y., and Holmes, B. (2002) Proposed minimal standards for describing new taxa of the family *Flavobacteriaceae* and emended description of the family. *International Journal of Systematic and Evolutionary Microbiology* **52**: 1049-1070.

Bernardet, J.-F., Segers, P., Vancanneyt, M., Berthe, F., Kersters, K., and Vandamme, P. (1996) Cutting a gordian knot: Emended classification and description of the genus *Flavobacterium*, emended description of the family *Flavobacteriaceae*, and proposal of *Flavobacterium hydatis* nom. nov. (Basonym, *Cytophaga aquatilis* Strohl and Tait 1978). *International Journal of Systematic and Evolutionary Microbiology* **46**: 128-148.

Berteau, O., and Mulloy, B. (2003) Sulfated fucans, fresh perspectives: Structures, functions, and biological properties of sulfated fucans and an overview of enzymes active toward this class of polysaccharide. *Glycobiology* **13**: 29R-40R.

Blevins, S.M., and Bronze, M.S. (2010) Robert Koch and the 'golden age' of bacteriology. *International Journal of Infectious Diseases* **14**: e744-e751.

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* **568**: 327-335.

Braun, T.F., Khubbar, M.K., Saffarini, D.A., and McBride, M.J. (2005) *Flavobacterium johnsoniae* gliding motility genes identified by mariner mutagenesis. *Journal of Bacteriology* **187**: 6943-6952.

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* **12**: 686.

Buck-Wiese, H., Andskog, M.A., Nguyen, N.P., Bligh, M., Asmala, E., Vidal-Melgosa, S. *et al.* (2023) Fucoid brown algae inject fucoidan carbon into the ocean. *Proceedings of the National Academy of Sciences* **120**: e2210561119.

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* **117**: 9679-9687.

Busi, M.V., Barchiesi, J., Martín, M., and Gomez-Casati, D.F. (2014) Starch metabolism in green algae. *Starch - Stärke* **66**: 28-40.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The carbohydrate-active enzymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Research* **37**: 233-238.

Cardoso, M.A., Noseda, M.D., Fujii, M.T., Zibetti, R.G.M., and Duarte, M.E.R. (2007) Sulfated xylomannans isolated from red seaweeds *Chondrophycus papillosus* and *C. flagelliferus* (*Ceramiales*) from Brazil. *Carbohydrate Research* **342**: 2766-2775.

Carini, P., Steindler, L., Beszteri, S., and Giovannoni, S.J. (2013) Nutrient requirements for growth of the extreme oligotroph 'Candidatus *Pelagibacter ubique*' HTCC1062 on a defined medium. *The ISME journal* **7**: 592-602.

Castro, L.S.E.P.W., Pinheiro, T.S., Castro, A.J.G., Dore, C.M.P.G., da Silva, N.B., Faustino Alves, M.G.d.C. *et al.* (2014) Fucose-containing sulfated polysaccharides from brown macroalgae *Lobophora variegata* with antioxidant, anti-inflammatory, and antitumoral effects. *Journal of Applied Phycology* **26**: 1783-1790.

Celia, H., Noinaj, N., Zakharov, S.D., Bordignon, E., Botos, I., Santamaria, M. *et al.* (2016) Structural insight into the role of the Ton complex in energy transduction. *Nature* **538**: 60-65.

Celis, J.E., Gromov, P., Østergaard, M., Madsen, P., Honoré, B., Dejgaard, K. *et al.* (1996) Human 2-D PAGE databases for proteome analysis in health and disease:. *FEBS Letters* **398**: 129-134.

Chen, G., Pramanik, B.N., Liu, Y.-H., and Mirza, U.A. (2007) Applications of LC/MS in structure identifications of small molecules and proteins in drug discovery. *Journal of Mass Spectrometry* **42**: 279-287.

Chevolot, L., Mulloy, B., Ratiskol, J., Foucault, A., and Colliec-Jouault, S. (2001) A disaccharide repeat unit is the major structure in fucoidans from two species of brown algae. *Carbohydrate Research* **330**: 529-535.

Cho, W.C.S. (2007) Proteomics technologies and challenges. *Genomics, Proteomics & Bioinformatics* **5**: 77-85.

Ciancia, M., Quintana, I., Vizcargüénaga, M.I., Kasulin, L., de Dios, A., Estevez, J.M., and Cerezo, A.S. (2007) Polysaccharides from the green seaweeds *Codium fragile* and *C. vermilara* with controversial effects on hemostasis. *International Journal of Biological Macromolecules* **41**: 641-649.

Cody, R.B. (1997) Electrospray ionization mass spectrometry. In *Fundamentals, Instrumentation and Applications*. New York, USA: Wiley.

Cong, Q., Chen, H., Liao, W., Xiao, F., Wang, P., Qin, Y. *et al.* (2016) Structural characterization and effect on anti-angiogenic activity of a fucoidan from *Sargassum fusiforme*. *Carbohydrate Polymers* **136**: 899-907.

Connon, S., A., and Giovannoni, S., J. (2002) High-throughput methods for culturing microorganisms in very-low-nutrient media yield diverse new marine isolates. *Applied and Environmental Microbiology* **68**: 3878-3885.

Cox, J., and Mann, M. (2007) Is Proteomics the new genomics? *Cell* **130**: 395-398.

Cox, J., and Mann, M. (2011) Quantitative, high-resolution proteomics for data-driven systems biology. *Annual Review of Biochemistry* **80**: 273-299.

Cristea, I.M., Gaskell, S.J., and Whetton, A.D. (2004) Proteomics techniques and their application to hematology. *Blood* **103**: 3624-3634.

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* **65**: 3192-3204.

Davies, G.J., Gloster, T.M., and Henrissat, B. (2005) Recent structural insights into the expanding world of carbohydrate-active enzymes. *Current Opinion in Structural Biology* **15**: 637-645.

De Philippis, R., and Vincenzini, M. (1998) Exocellular polysaccharides from *Cyanobacteria* and their possible applications. *FEMS Microbiology Reviews* **22**: 151-175.

Dobrinčić, A., Balbino, S., Zorić, Z., Pedisić, S., Bursać Kovačević, D., Elez Garofulić, I., and Dragović-Uzelac, V. (2020) Advanced technologies for the extraction of marine brown algal polysaccharides. *Marine Drugs* **18**: 168.

Domozych, D. (2019) Algal cell walls. In *Encyclopedia of Life Sciences*, pp. 1-11.

Domozych, D., Ciancia, M., Fangel, J.U., Mikkelsen, M.D., Ulvskov, P., and Willats, W.G. (2012) The cell walls of green algae: A journey through evolution and diversity. *Frontiers in Plant Science* **3**: 82.

Domozych, D.S., and LoRicco, J.G. (2023) The extracellular matrix of green algae. *Plant Physiology* **194**: 15-32.

Drula, E., Garron, M.-L., Dogan, S., Lombard, V., Henrissat, B., and Terrapon, N. (2022) The carbohydrate-active enzyme database: Functions and literature. *Nucleic Acids Research* **50**: 571-577.

Duarte, C.M., Gattuso, J.-P., Hancke, K., Gundersen, H., Filbee-Dexter, K., Pedersen, M.F. *et al.* (2022) Global estimates of the extent and production of macroalgal forests. *Global Ecology and Biogeography* **31**: 1422-1439.

Falkowski, P.G., Barber, R.T., and Smetacek, V. (1998) Biogeochemical controls and feedbacks on ocean primary production. *Science* **281**: 200-206.

Falshaw, R., and Furneaux, R.H. (1998) Structural analysis of carrageenans from the tetrasporic stages of the red algae, *Gigartina lanceata* and *Gigartina chapmanii* (*Gigartinaceae*, *Rhodophyta*). *Carbohydrate Research* **307**: 325-331.

Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., and Whitehouse, C.M. (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**: 64-71.

Field, C.B., Behrenfeld, M.J., Randerson, J.T., and Falkowski, P. (1998) Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science* **281**: 237-240.

Fogg, G.E. (1983) The ecological significance of extracellular products of phytoplankton photosynthesis. **26**: 3-14.

Fujita, K., Sakamoto, A., Kaneko, S., Kotake, T., Tsumuraya, Y., and Kitahara, K. (2019) Degradative enzymes for type II arabinogalactan side chains in *Bifidobacterium longum* subsp. *longum*. *Applied Microbiology and Biotechnology* **103**: 1299-1310.

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* **54**: 217-231.

Ghosh, T., Pujol, C.A., Damonte, E.B., Sinha, S., and Ray, B. (2009) Sulfated xylomannans from the red seaweed *Sebdenia Polydactyla*: Structural features, chemical modification and antiviral activity. *Antiviral Chemistry and Chemotherapy* **19**: 235-242.

Giering, S.L.C., Sanders, R., Lampitt, R.S., Anderson, T.R., Tamburini, C., Boutrif, M. *et al.* (2014) Reconciliation of the carbon budget in the ocean's twilight zone. *Nature* **507**: 480-483.

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* **541**: 407-411.

González-Hourcade, M., del Campo, E.M., Braga, M.R., Salgado, A., and Casano, L.M. (2020) Disentangling the role of extracellular polysaccharides in desiccation tolerance in lichen-forming microalgae. First evidence of sulfated polysaccharides and ancient sulfotransferase genes. *Environmental Microbiology* **22**: 3096-3111.

Grondin, J.M., Tamura, K., Déjean, G., Abbott, D.W., and Brumer, H. (2017) Polysaccharide utilization loci: Fueling microbial communities. *Journal of Bacteriology* **199**: e00860-16.

Grossart, H.-P. (2010) Ecological consequences of bacterioplankton lifestyles: Changes in concepts are needed. *Environmental Microbiology Reports* **2**: 706-714.

Grossart, H.-P., Kiørboe, T., Tang, K., and Ploug, H. (2003) Bacterial colonization of particles: Growth and interactions. *Applied and Environmental Microbiology* **69**: 3500-3509.

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* **266**: 194-200.

Guibet, M., Boulenguer, P., Mazoyer, J., Kervarec, N., Antonopoulos, A., Lafosse, M., and Helbert, W. (2008) Composition and distribution of carrabiose moieties in hybrid κ-/ι-carrageenans using carrageenases. *Biomacromolecules* **9**: 408-415.

Hahn, M.W., Stadler, P., Wu, Q.L., and Pöckl, M. (2004) The filtration–acclimatization method for isolation of an important fraction of the not readily cultivable bacteria. *Journal of Microbiological Methods* **57**: 379-390.

Hahnke, R.L., and Harder, J. (2013) Phylogenetic diversity of *Flavobacteria* isolated from the North Sea on solid media. *Systematic and Applied Microbiology* **36**: 497-504.

Hahnke, R.L., Bennke, C.M., Fuchs, B.M., Mann, A.J., Rhiel, E., Teeling, H. *et al.* (2015) Dilution cultivation of marine heterotrophic bacteria abundant after a spring phytoplankton bloom in the North Sea. *Environmental Microbiology* **17**: 3515-3526.

Harder, R. (1917) *Ernährungsphysiologische Untersuchungen an Cyanophyceen: hauptsächlich dem endophytischen Nostoc punctiforme*: Fischer.

Haug, A., Larsen, B., and Smidsrød, O. (1974) Uronic acid sequence in alginate from different sources. *Carbohydrate Research* **32**: 217-225.

Heins, A., and Harder, J. (2023) Particle-associated bacteria in seawater dominate the colony-forming microbiome on ZoBell marine agar. *FEMS Microbiology Ecology* **99**: fiac151.

Heins, A., Amann, R.I., and Harder, J. (2021a) Cultivation of particle-associated heterotrophic bacteria during a spring phytoplankton bloom in the North Sea. *Systematic and Applied Microbiology* **44**: 126232.

Heins, A., Reintjes, G., Amann, R.I., and Harder, J. (2021b) Particle collection in Imhoff sedimentation cones enriches both motile chemotactic and particle-attached bacteria. *Frontiers in Microbiology* **12**: 643730.

Helbert, W. (2017) Marine polysaccharide sulfatases. *Frontiers in Marine Science* **4**: Article 6.

Hellebust, J. (1974) Extracellular products. In *Algal Physiology and Biochemistry*. ND, S. (ed). Berkeley, CA: University of California Press, pp. 838-63.

Hervé, C., Siméon, A., Jam, M., Cassin, A., Johnson, K.L., Salmeán, A.A. *et al.* (2016) Arabinogalactan proteins have deep roots in eukaryotes: Identification of genes and epitopes in brown algae and their role in *Fucus serratus* embryo development. *New Phytologist* **209**: 1428-1441.

Hesse, W. (1992) Walther and Angelina Hesse-Early contributors to bacteriology. *ASM News* **58**: 425-428.

Hoagland, K.D., Rosowski, J.R., Gretz, M.R., and Roemer, S.C. (1993) Diatom extracellular polymeric substances: Function, fine structure, chemistry and physiology. *Journal of Phycology* **29**: 537-566.

Hobbie, J.E., Daley, R.J., and Jasper, S. (1977) Use of nuclepore filters for counting bacteria by fluorescence microscopy. *Applied and Environmental Microbiology* **33**: 1225-1228.

Huang, G., Vidal-Melgosa, S., Sichert, A., Becker, S., Fang, Y., Niggemann, J. *et al.* (2021) Secretion of sulfated fucans by diatoms may contribute to marine aggregate formation. *Limnology and Oceanography* **66**: 3768-3782.

Hunter-Cevera, K.R., Neubert, M.G., Olson, R.J., Solow, A.R., Shalapyonok, A., and Sosik, H.M. (2016) Physiological and ecological drivers of early spring blooms of a coastal phytoplankter. *Science* **354**: 326-329.

Ishizaka, J., Takahashi, M., and Ichimura, S. (1983) Evaluation of coastal upwelling effects on phytoplankton growth by simulated culture experiments. *Marine Biology* **76**: 271-278.

Jannasch, H.W., and Jones, G.E. (1959) Bacterial populations in sea water as determined by different methods of enumeration. *Limnology and Oceanography* **4**: 128-139.

Jarrell, K.F., and McBride, M.J. (2008) The surprisingly diverse ways that prokaryotes move. *Nature Reviews Microbiology* **6**: 466-476.

Jothisaraswathi, S., Babu, B., and Rengasamy, R. (2006) Seasonal studies on alginate and its composition II: *Turbinaria conoides* (J.Ag.) Kütz. (*Fucales*, *Phaeophyceae*). *Journal of Applied Phycology* **18**: 161-166.

Kadam, S.U., Tiwari, B.K., and O'Donnell, C.P. (2015) Extraction, structure and biofunctional activities of laminarin from brown algae. *International Journal of Food Science & Technology* **50**: 24-31.

Kappelmann, L., Krüger, K., Hehemann, J.-H., Harder, J., Markert, S., Unfried, F. *et al.* (2019) Polysaccharide utilization loci of North Sea *Flavobacteriia* as basis for

using SusC/D-protein expression for predicting major phytoplankton glycans. *The ISME Journal* **13**: 76-91.

Karas, M., and Hillenkamp, F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical chemistry* **60**: 2299-2301.

Kasalický, V., Jezbera, J., Hahn, M.W., and Šimek, K. (2013) The diversity of the limnohabitans genus, an important group of freshwater bacterioplankton, by characterization of 35 isolated strains. *PLOS ONE* **8**: e58209.

Kirchman, D.L. (2002) The ecology of *Cytophaga–Flavobacteria* in aquatic environments. *FEMS Microbiology Ecology* **39**: 91-100.

Koch, A.L. (2001) Oligotrophs versus copiotrophs. *BioEssays: news and reviews in molecular, cellular and developmental biology* **23**: 657-661.

Koch, H., Dürwald, A., Schweder, T., Noriega-Ortega, B., Vidal-Melgosa, S., Hehemann, J.-H. *et al.* (2018) Biphasic cellular adaptations and ecological implications of *Alteromonas macleodii* degrading a mixture of algal polysaccharides. *The ISME Journal* **13**: 92-103.

Kogure, K., Simidu, U., and Taga, N. (1979) A tentative direct microscopic method for counting living marine bacteria. *Canadian Journal of Microbiology* **25**: 415-420.

Kogure, K., Simidu, U., and Taga, N. (1981) Bacterial attachment to phytoplankton in sea water. *Journal of Experimental Marine Biology and Ecology* **56**: 197-204.

Kraan, S. (2012) Algal polysaccharides, novel applications and outlook. In *Carbohydrates-comprehensive studies on glycobiology and glycotechnology*. London: InTech.

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* **3**: e1426.

Kulk, G., Platt, T., Dingle, J., Jackson, T., Jönsson, B., Bouman, H. *et al.* (2021) Correction: Kulk et al. Primary production, an index of climate change in the ocean: Satellite-based estimates over two decades. *Remote Sensing* **13**: 3462.

Kumar, S., Mehta, G.K., Prasad, K., Meena, R., and Siddhanta, A.K. (2011) Chemical investigation of carrageenan from the red alga *Sarconema filiforme* (*Gigartinales*, *Rhodophyta*) of Indian waters. *Natural Product Communications* **6**: 1934578X1100600928.

Lauro, F.M., McDougald, D., Thomas, T., Williams, T.J., Egan, S., Rice, S. *et al.* (2009) The genomic basis of trophic strategy in marine bacteria. *Proceedings of the National Academy of Sciences* **106**: 15527-15533.

Lee, J.-B., Koizumi, S., Hayashi, K., and Hayashi, T. (2010) Structure of rhamnan sulfate from the green alga *Monostroma nitidum* and its anti-herpetic effect. *Carbohydrate Polymers* **81**: 572-577.

Leszczuk, A., Kalaitzis, P., Kulik, J., and Zdunek, A. (2023) Review: Structure and modifications of arabinogalactan proteins (AGPs). *BMC Plant Biology* **23**: 45.

Li, B., Lu, F., Wei, X., and Zhao, R. (2008) Fucoidan: Structure and bioactivity. *Molecules* **13**: 1671-1695.

Li, Y., Zheng, Y., Zhang, Y., Yang, Y., Wang, P., Imre, B. *et al.* (2021) Brown algae carbohydrates: Structures, pharmaceutical properties, and research challenges. *Marine Drugs* **19**: 620.

Llobet-Brossa, E., Rosselló-Mora, R., and Amann, R. (1998) Microbial community composition of Wadden Sea sediments as revealed by fluorescence in situ hybridization. *Applied and Environmental Microbiology* **64**: 2691-2696.

Lu, D.-C., Wang, F.-Q., Amann, R.I., Teeling, H., and Du, Z.-J. (2023) Epiphytic common core bacteria in the microbiomes of co-located green (*Ulva*), brown (*Saccharina*) and red (*Grateloupia*, *Gelidium*) macroalgae. *Microbiome* **11**: 126.

Ma, Y., and Johnson, K. (2023) Arabinogalactan proteins – Multifunctional glycoproteins of the plant cell wall. *The Cell Surface* **9**: 100102.

Mague, T.H., Friberg, E., Hughes, D.J., and Morris, I. (1980) Extracellular release of carbon by marine phytoplankton; a physiological approach. *Limnology and Oceanography* **25**: 262-279.

Mann, K.H. (2006) *Dynamics of marine ecosystems: Biological-physical interactions in the oceans*. Malden, Mass Oxford: Malden, Mass Oxford: Blackwell Pub.

Manners, D.J., and Sturgeon, R.J. (1982) Reserve carbohydrates of algae, fungi, and lichens. In *Plant Carbohydrates I: Intracellular Carbohydrates*. Loewus, F.A., and Tanner, W. (eds). Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 472-514.

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* **9**: e1001221.

Mazéas, L., Yonamine, R., Barbeyron, T., Henrissat, B., Drula, E., Terrapon, N. *et al.* (2023) Assembly and synthesis of the extracellular matrix in brown algae. *Seminars in Cell & Developmental Biology* **134**: 112-124.

McBride, M., J., Xie, G., Martens Eric, C., Lapidus, A., Henrissat, B., Rhodes Ryan, G. *et al.* (2009) Novel features of the polysaccharide-digesting gliding bacterium *Flavobacterium johnsoniae* as revealed by genome sequence analysis. *Applied and Environmental Microbiology* **75**: 6864-6875.

McBride, M.J. (2001) Bacterial gliding motility: Multiple mechanisms for cell movement over surfaces. *Annual Review of Microbiology* **55**: 49-75.

Meinhard, S., Hans-Peter, G., Bernd, S., and Helle, P. (2002) Microbial ecology of organic aggregates in aquatic ecosystems. *Aquatic Microbial Ecology* **28**: 175-211.

Miksch, S., Meiners, M., Meyerdierks, A., Probandt, D., Wegener, G., Titschack, J. *et al.* (2021) Bacterial communities in temperate and polar coastal sands are seasonally stable. *ISME Communications* **1**: 29.

Miller, C.B. (2012) *Biological oceanography*. Chichester, Great Britain: Wiley-Blackwell.

Moran, M.A., Ferrer-González, F.X., Fu, H., Nowinski, B., Olofsson, M., Powers, M.A. *et al.* (2022) The Ocean's labile DOC supply chain. *Limnology and Oceanography* **67**: 1007-1021.

Moran, M.A., Kujawinski, E.B., Stubbins, A., Fatland, R., Aluwihare, L.I., Buchan, A. *et al.* (2016) Deciphering ocean carbon in a changing world. *Proceedings of the National Academy of Sciences* **113**: 3143-3151.

Myklestad, S.M. (1995) Release of extracellular products by phytoplankton with special emphasis on polysaccharides. *Science of The Total Environment* **165**: 155-164.

Nan, B., and Zusman, D.R. (2016) Novel mechanisms power bacterial gliding motility. *Molecular Microbiology* **101**: 186-193.

Natrah, F.M.I., Bossier, P., Sorgeloos, P., Yusoff, F.M., and Defoirdt, T. (2014) Significance of microalgal–bacterial interactions for aquaculture. *Reviews in Aquaculture* **6**: 48-61.

Nedashkovskaya, O.I., Kim, S.B., Han, S.K., Lysenko, A.M., Rohde, M., Rhee, M.-S. *et al.* (2004) *Maribacter* gen. nov., a new member of the family *Flavobacteriaceae*, isolated from marine habitats, containing the species *Maribacter sedimenticola* sp. nov., *Maribacter aquivivus* sp. nov., *Maribacter orientalis* sp. nov. and *Maribacter ulvicola* sp. nov. *International Journal of Systematic and Evolutionary Microbiology* **54**: 1017-1023.

Nelson Shawn, S., Bollampalli, S., and McBride Mark, J. (2008) SprB Is a cell surface component of the *Flavobacterium johnsoniae* gliding motility machinery. *Journal of Bacteriology* **190**: 2851-2857.

Nelson, T.E., and Lewis, B.A. (1974) Separation and characterization of the soluble and insoluble components of insoluble laminaran. *Carbohydrate Research* **33**: 63-74.

Nicolas, J.L., Corre, S., and Cochard, J.C. (2004) Bacterial population association with phytoplankton cultured in a bivalve hatchery. *Microbial Ecology* **48**: 400-413.

Noinaj, N., Guillier, M., Barnard, T.J., and Buchanan, S.K. (2010) TonB-dependent transporters: Regulation, structure, and function. *Annual Review of Microbiology* **64**: 43-60.

Nowicki, M., DeVries, T., and Siegel, D.A. (2022) Quantifying the carbon export and sequestration pathways of the ocean's biological carbon pump. *Global Biogeochemical Cycles* **36**: e2021GB007083.

O'Farrell, P.H. (1975) High resolution two-dimensional electrophoresis of proteins. *Journal of Biological Chemistry* **250**: 4007-4021.

Oppenheimer, C.H., and Zobell, C.E. (1952) The growth and viability of sixty-three species of marine bacteria as influenced by hydrostatic pressure. *Journal of Marine Research* **11**: 10-18.

Overmann, J., Abt, B., and Sikorski, J. (2017) Present and future of culturing bacteria. *Annual Review of Microbiology* **71**: 711-730.

Pandey, A., and Mann, M. (2000) Proteomics to study genes and genomes. *Nature* **405**: 837-846.

Petri, R.J. (1887) Eine kleine Modification des Koch'schen Plattenverfahrens. *Cntralbl Bakteriol Pasitenkunde* **1**: 279-280.

Pfennig, N. (1965) Anreicherungskulturen für rote und grüne Schwefelbakterien. *Zb Bakt, 1, Abt Org Suppl* **1**: 503-504.

Pfennig, N. (1974) *Rhodopseudomonas globiformis*, sp. n., a new species of the *Rhodospirillaceae*. *Archives of Microbiology* **100**: 197-206.

Pfennig, N. (1978) *Rhodocyclus purpureus* gen. nov. and sp. nov., a ring-shaped, Vitamin B12-requiring member of the family *Rhodospirillaceae*. *International Journal of Systematic and Evolutionary Microbiology* **28**: 283-288.

Pfennig, N., and Lippert, K.-D. (1966) Über das Vitamin B12-Bedürfnis phototropher Schwefelbakterien. *Archiv für Mikrobiologie* **55**: 245-256.

Pham, V.H.T., and Kim, J. (2012) Cultivation of unculturable soil bacteria. *Trends in Biotechnology* **30**: 475-484.

Popper, Z.A., Michel, G., Hervé, C., Domozych, D.S., Willats, W.G.T., Tuohy, M.G. *et al.* (2011) Evolution and diversity of plant cell walls: From algae to flowering plants. *Annual Review of Plant Biology* **62**: 567-590.

Port, A., Gurgel, K.-W., Staneva, J., Schulz-Stellenfleth, J., and Stanev, E.V. (2011) Tidal and wind-driven surface currents in the German Bight: HFR observations versus model simulations. *Ocean Dynamics* **61**: 1567-1585.

Probandt, D., Eickhorst, T., Ellrott, A., Amann, R., and Knittel, K. (2018) Microbial life on a sand grain: From bulk sediment to single grains. *The ISME Journal* **12**: 623-633.

Rappé, M.S., and Giovannoni, S.J. (2003) The uncultured microbial majority. *Annual Review of Microbiology* **57**: 369-394.

Rappé, M.S., Connon, S.A., Vergin, K.L., and Giovannoni, S.J. (2002) Cultivation of the ubiquitous *SAR11* marine bacterioplankton clade. *Nature* **418**: 630-633.

Raveendran, S., Yoshida, Y., Maekawa, T., and Kumar, D.S. (2013) Pharmaceutically versatile sulfated polysaccharide based bionano platforms. *Nanomedicine: Nanotechnology, Biology and Medicine* **9**: 605-626.

Raven, J.A., Giordano, M., Beardall, J., and Maberly, S.C. (2011) Algal and aquatic plant carbon concentrating mechanisms in relation to environmental change. *Photosynthesis Research* **109**: 281-296.

Ray, B., and Lahaye, M. (1995) Cell-wall polysaccharides from the marine green alga *Ulva "rigida"* (*ulvales*, *chlorophyta*). Extraction and chemical composition. *Carbohydrate Research* **274**: 251-261.

Reintjes, G., Arnosti, C., Fuchs, B.M., and Amann, R. (2017) An alternative polysaccharide uptake mechanism of marine bacteria. *The ISME Journal* **11**: 1640-1650.

Reintjes, G., Heins, A., Wang, C., and Amann, R. (2023) Abundance and composition of particles and their attached microbiomes along an Atlantic Meridional Transect. *Frontiers in Marine Science* **10**: 1051510.

Rieck, A., Herlemann, D.P.R., Jürgens, K., and Grossart, H.-P. (2015) Particle-associated differ from free-living bacteria in surface waters of the Baltic Sea. *Frontiers in Microbiology* **6**: 1297.

Rioux, L.E., Turgeon, S.L., and Beaulieu, M. (2007) Characterization of polysaccharides extracted from brown seaweeds. *Carbohydrate Polymers* **69**: 530-537.

Rosenberg, E., Long, E.F., Lory, S., Stackebrandt, E., and Thompson, F. (2014) *The Prokaryotes: Other major lineages of bacteria and the archaea*. Berlin, Germany: Springer

Rossi, F., and De Philippis, R. (2016) Exocellular polysaccharides in microalgae and *Cyanobacteria*: Chemical features, role and enzymes and genes involved in their biosynthesis. In *The Physiology of Microalgae*. Borowitzka, M.A., Beardall, J., and Raven, J.A. (eds). Cham: Springer International Publishing, pp. 565-590.

Rupérez, P., Ahrazem, O., and Leal, J.A. (2002) Potential antioxidant capacity of sulfated polysaccharides from the edible marine brown seaweed *Fucus vesiculosus*. *Journal of Agricultural and Food Chemistry* **50**: 840-845.

Sapp, M., Wichels, A., Wiltshire, K.H., and Gerdts, G. (2007a) Bacterial community dynamics during the winter–spring transition in the North Sea. *FEMS Microbiology Ecology* **59**: 622-637.

Sapp, M., Schwaderer, A.S., Wiltshire, K.H., Hoppe, H.-G., Gerdts, G., and Wichels, A. (2007b) Species-specific bacterial communities in the phycosphere of microalgae? *Microbial Ecology* **53**: 683-699.

Saraswathi, S.J., Babu, B., and Rengasamy, R. (2003) Seasonal studies on the alginate and its biochemical composition I: *Sargassum polycystum* (Fucales), *Phaeophyceae*. *Phycological Research* **51**: 240-243.

Sato, K., Hara, K., Yoshimi, Y., Kitazawa, K., Ito, H., Tsumuraya, Y., and Kotake, T. (2018) Yariv reactivity of type II arabinogalactan from larch wood. *Carbohydrate Research* **467**: 8-13.

Schultz-Johansen, M., Cueff, M., Hardouin, K., Jam, M., Larocque, R., Glaring, M.A. *et al.* (2018) Discovery and screening of novel metagenome-derived GH107 enzymes targeting sulfated fucans from brown algae. *The FEBS Journal* **285**: 4281-4295.

Schut, F., de Vries Egbert, J., Gottschal Jan, C., Robertson Betsy, R., Harder, W., Prins Rudolf, A., and Button Don, K. (1993) Isolation of typical marine bacteria by dilution culture: Growth, maintenance, and characteristics of isolates under laboratory conditions. *Applied and Environmental Microbiology* **59**: 2150-2160.

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* **2**: 17065.

Shimonaga, T., Fujiwara, S., Kaneko, M., Izumo, A., Nihei, S., Francisco, P.B. *et al.* (2007) Variation in storage α-polyglucans of red algae: Amylose and semi-amylopectin types in *porphyridium* and glycogen type in *cyanidium*. *Marine Biotechnology* **9**: 192-202.

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* **182**: 5365-5372.

Sichert, A., Corzett, C.H., Schechter, M.S., Unfried, F., Markert, S., Becher, D. *et al.* (2020) *Verrucomicrobia* use hundreds of enzymes to digest the algal polysaccharide fucoidan. *Nature Microbiology* **5**: 1026-1039.

Sidhu, C., Kirstein, I.V., Meunier, C.L., Rick, J., Fofonova, V., Wiltshire, K.H. *et al.* (2023) Dissolved storage glycans shaped the community composition of abundant bacterioplankton clades during a North Sea spring phytoplankton bloom. *Microbiome* **11**: 77.

Smith, M.W., Zeigler Allen, L., Allen, A.E., Herfort, L., and Simon, H.M. (2013) Contrasting genomic properties of free-living and particle-attached microbial assemblages within a coastal ecosystem. *Frontiers in Microbiology* **4**: 4:120.

Song, J., Oh, H.-M., and Cho, J.-C. (2009) Improved culturability of *SAR11* strains in dilution-to-extinction culturing from the East Sea, West Pacific Ocean. *FEMS Microbiology Letters* **295**: 141-147.

Steinke, N., Vidal-Melgosa, S., Schultz-Johansen, M., and Hehemann, J.-H. (2022) Biocatalytic quantification of α-glucan in marine particulate organic matter. *MicrobiologyOpen* **11**: e1289.

Stocker, R. (2012) Marine microbes see a sea of gradients. *Science* **338**: 628-633.

Stocker, R., and Seymour Justin, R. (2012) Ecology and physics of bacterial chemotaxis in the Ocean. *Microbiology and Molecular Biology Reviews* **76**: 792-812.

Stocker, R., Seymour, J.R., Samadani, A., Hunt, D.E., and Polz, M.F. (2008) Rapid chemotactic response enables marine bacteria to exploit ephemeral microscale nutrient patches. *Proceedings of the National Academy of Sciences* **105**: 4209-4214.

Sutherland, I.W. (1982) Biosynthesis of microbial exopolysaccharides. In *Advances in Microbial Physiology*. Rose, A.H., and Morris, J.G. (eds): Academic Press, pp. 79-150.

Swinbanks, D. (1995) Australia backs innovation, shuns telescope. *Nature* **378**: 653-653.

Synytsya, A., Čopíková, J., Kim, W.J., and Park, Y.I. (2015) Cell wall polysaccharides of marine algae. In *Springer Handbook of Marine Biotechnology*. Kim, S.-K. (ed). Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 543-590.

Teeling, H., Fuchs, B.M., Bennke, C.M., Krüger, K., Chafee, M., Kappelmann, L. *et al.* (2016) Recurring patterns in bacterioplankton dynamics during coastal spring algae blooms. *eLife* **5**: e11888.

Teeling, H., Fuchs, B.M., Becher, D., Klockow, C., Gardebrecht, A., Bennke, C.M. *et al.* (2012) Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* **336**: 608-611.

Teira, E., José Pazó, M., Serret, P., and Fernández, E. (2001) Dissolved organic carbon production by microbial populations in the Atlantic Ocean. *Limnology and Oceanography* **46**: 1370-1377.

Thornton, D.C.O. (2014) Dissolved organic matter (DOM) release by phytoplankton in the contemporary and future ocean. *European Journal of Phycology* **49**: 20-46.

Turley, C.M., and Stutt, E.D. (2000) Depth-related cell-specific bacterial leucine incorporation rates on particles and its biogeochemical significance in the Northwest Mediterranean. *Limnology and Oceanography* **45**: 419-425.

Turner, J.T. (2015) Zooplankton fecal pellets, marine snow, phytodetritus and the ocean's biological pump. *Progress in Oceanography* **130**: 205-248.

Urbani, R., Magaletti, E., Sist, P., and Cicero, A.M. (2005) Extracellular carbohydrates released by the marine diatoms *Cylindrotheca closterium*, *Thalassiosira pseudonana* and *Skeletonema costatum*: Effect of P-depletion and growth status. *Science of The Total Environment* **353**: 300-306.

Usov, A.I. (1992) Sulfated polysaccharides of the red seaweeds. *Food Hydrocolloids* **6**: 9-23.

van de Velde, F., Peppelman, H.A., Rollema, H.S., and Tromp, R.H. (2001) On the structure of κ/ι-hybrid carrageenans. *Carbohydrate Research* **331**: 271-283.

Van Dongen-Vogels, V., Seymour, J.R., Middleton, J.F., Mitchell, J.G., and Seuront, L. (2012) Shifts in picophytoplankton community structure influenced by changing upwelling conditions. *Estuarine, Coastal and Shelf Science* **109**: 81-90.

Variem, S.S., and Kizhakkedath, V.K. (2021) Phycosphere associated bacteria; a prospective source of bioactive compounds. *Biologia* **76**: 1095-1098.

Vasconcelos, A.A., and Pomin, V.H. (2017) The sea as a rich source of structurally unique glycosaminoglycans and mimetics. *Microorganisms* **5**: 51.

Vidal-Melgosa, S., Sichert, A., Francis, T.B., Bartosik, D., Niggemann, J., Wichels, A. *et al.* (2021) Diatom fucan polysaccharide precipitates carbon during algal blooms. *Nature Communications* **12**: 1150.

Vieira, R.P., Mulloy, B., and Mourão, P.A. (1991) Structure of a fucose-branched chondroitin sulfate from sea cucumber. Evidence for the presence of 3-O-sulfo-beta-D-glucuronosyl residues. *Journal of Biological Chemistry* **266**: 13530-13536.

Villa-Rivera, M.G., Cano-Camacho, H., López-Romero, E., and Zavala-Páramo, M.G. (2021) The role of arabinogalactan type II degradation in plant-microbe interactions. *Frontiers in Microbiology* **12**: 730543.

Wang, F.-Q., Bartosik, D., Sidhu, C., Siebers, R., Lu, D.-C., Trautwein-Schult, A. *et al.* (2024) Particle-attached bacteria act as gatekeepers in the decomposition of complex phytoplankton polysaccharides. *Microbiome* **12**: 32.

Watanabe, K., Komatsu, N., Ishii, Y., and Negishi, M. (2009) Effective isolation of bacterioplankton genus *Polynucleobacter* from freshwater environments grown on photochemically degraded dissolved organic matter. *FEMS Microbiology Ecology* **67**: 57-68.

Whitman, W.B., Coleman, D.C., and Wiebe, W.J. (1998) Prokaryotes: The unseen majority. *Proceedings of the National Academy of Sciences* **95**: 6578-6583.

Widdel, F., and Pfennig, N. (1977) A new anaerobic, sporing, acetate-oxidizing, sulfate-reducing bacterium, *Desulfotomaculum* (emend.) *acetoxidans*. *Archives of Microbiology* **112**: 119-122.

Widdel, F., and Pfennig, N. (1981) Studies on dissimilatory sulfate-reducing bacteria that decompose fatty acids. *Archives of Microbiology* **129**: 395-400.

Wilkins, M.R., Sanchez, J.C., Gooley, A.A., Appel, R.D., Humphery-Smith, I., Hochstrasser, D.F., and Williams, K.L. (1996a) Progress with proteome projects: Why all proteins expressed by a genome should be identified and how to do it. *Biotechnology and genetic engineering reviews* **13**: 19-50.

Wilkins, M.R., Pasquali, C., Appel, R.D., Ou, K., Golaz, O., Sanchez, J.-C. *et al.* (1996b) From proteins to proteomes: Large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Bio/Technology* **14**: 61-65.

Worden, A.Z., Follows, M.J., Giovannoni, S.J., Wilken, S., Zimmerman, A.E., and Keeling, P.J. (2015) Rethinking the marine carbon cycle: Factoring in the multifarious lifestyles of microbes. *Science* **347**: 1257594.

Wotton, R.S. (2004) The utiquity and many roles of exopolymers (EPS) in aquatic systems. *Scientia Marina* **68**: 13-21.

Yang, B., Yu, G., Zhao, X., Ren, W., Jiao, G., Fang, L. *et al.* (2011) Structural characterisation and bioactivities of hybrid carrageenan-like sulphated galactan from red alga *Furcellaria lumbricalis*. *Food Chemistry* **124**: 50-57.

Yates, J.R., Gilchrist, A., Howell, K.E., and Bergeron, J.J.M. (2005) Proteomics of organelles and large cellular structures. *Nature Reviews Molecular Cell Biology* **6**: 702-714.

Yates, J.R., McCormack, A.L., Schieltz, D., Carmack, E., and Link, A. (1997) Direct analysis of protein mixtures by tandem mass spectrometry. *Journal of Protein Chemistry* **16**: 495-497.

Yu, S., Blennow, A., Bojko, M., Madsen, F., Olsen, C.E., and Engelsen, S.B. (2002) Physico-chemical characterization of floridean starch of red algae. *Starch - Stärke* **54**: 66-74.

Zhang, Y.-S., Zhang, Y.-Q., Zhao, X.-M., Liu, X.-L., Qin, Q.-L., Liu, N.-H. *et al.* (2024) Metagenomic insights into the dynamic degradation of brown algal

polysaccharides by kelp-associated microbiota. *Applied and Environmental Microbiology* **90**: e02025-23.

Zobell, C., E. (1941) Studies on marine bacteria. I. The cultural requirements of heterotrophic aerobes. *J Mar Res* **4**: 42-75.

ZoBell, C.E. (1946) *Marine microbiology, a monograph on hydrobacteriology*. Waltham, Mass: Chronica Botanica Company.

Chapter 2

Genes for laminarin degradation are dispersed in the genomes of particle-associated *Maribacter* species

Saskia Kalenborn[1], Daniela Zühlke[2], Greta Reintjes[3], Katharina Riedel[2], Rudolf I. Amann[1], Jens Harder[1*]

[1]Department of Molecular Ecology, Max Planck Institute for Marine Microbiology, Bremen, Germany; [2]Department for Microbial Physiology and Molecular Biology, University of Greifswald, Germany; [3] Microbial Carbohydrate Interaction Group, Department of Biology and Chemistry, University of Bremen, Germany

Contribution of the candidate in % of the total work

Experimental concept and design – 60%

Experimental work/acquisition of experimental data – 90%

Preparation of figures and tables – 100%

Drafting of the manuscript – 80%

Chapter 2

2.1 Abstract

Laminarin is a cytosolic storage polysaccharide of phytoplankton and macroalgae and accounts for over 10% of the world´s annually fixed carbon dioxide. Algal disruption, e.g., by viral lysis releases laminarin. The soluble sugar is rapidly utilized by free-living planktonic bacteria, in which sugar transporters and the degrading enzymes are frequently encoded in polysaccharide utilization loci. The annotation of flavobacterial genomes failed to identify canonical laminarin utilization loci in several particle-associated bacteria, in particular in strains of *Maribacter*. In this study, we report in vivo utilization of laminarin by *Maribacter forsetii* accompanied by additional cell growth and proliferation. Laminarin utilization coincided with the induction of an extracellular endo-laminarinase, SusC/D outer membrane oligosaccharide transporters and a periplasmic glycosyl hydrolase family 3 protein. An ABC transport system and sugar kinases were expressed. Endo-laminarinase activity was also observed in *Maribacter* sp. MAR_2009_72, *Maribacte*r sp. Hel_I_7 and *Maribacter dokdonensis* MAR_2009_60. *Maribacter dokdonensis* MAR_2009_71 lacked the large endo-laminarinase gene in the genome and had no endo-laminarinase activity. In all genomes, genes of induced proteins were scattered across the genome rather than clustered in a laminarin utilization locus. These observations revealed that the *Maribacter* strains investigated in this study participate in laminarin utilization, but in contrast to many free-living bacteria, there is no co localisation of genes encoding the enzymatic machinery for laminarin utilization.

2.2 Introduction

Laminarin is a major carbon storage polysaccharide in algae. Annual net biosynthesis rates have been estimated to be in the range of 6 to 18 Gt carbon, one eighth of the annual biomass production on earth (Alderkamp *et al.*, 2007; Becker *et al.*, 2020). The water-soluble compound is a linear glucose polymer with decorations. Around 25 glucose molecules are linked by β-(1,3)-glycosidic bonds and modified in algal species with β-(1,6)-linked glucose side chains (Nelson and Lewis, 1974; Alderkamp *et al.*, 2007; Rioux *et al.*, 2007; Kadam *et al.*, 2015; Becker

et al., 2020). It becomes available as an important carbon source for microorganisms when algae are lysed by zooplankton predation or viral infection. Then, as part of the dissolved organic carbon fraction (DOC), it sustains bacterioplankton blooms following algal blooms in regions with a dynamic annual climate, i. e. temperate and polar regions (Teeling *et al.*, 2012; Teeling *et al.*, 2016; Sidhu *et al.*, 2023). DOC favours growth of free-living bacteria, the abundance of particle attached bacteria remains low (Heins *et al.*, 2021b). In bacteria of the phylum *Bacteroidota*, polysaccharide utilization genes are often colocalized in genomic loci or islands, termed polysaccharide utilization loci (PULs) (Grondin *et al.*, 2017).

The starch utilization system (Sus) was the first described PUL (Shipman *et al.*, 2000). Efficient polysaccharide degradation pathways include an extracellular hydrolysis, preferentially at the cell surface including a forwarding of lysis products to a transporter. The oligosaccharides are actively translocated into the periplasm where they are hydrolysed and the monomers are transported into the cytoplasm. The outer membrane transport system SusC/D receives its energy from a proton gradient via an ExbBD-TonB system in the cytoplasmic membrane and a periplasmic domain to open the transport channel on the periplasmic side (Noinaj *et al.*, 2010). The depolymerisation of marine polysaccharides is catalysed by glycoside hydrolases, glycosyltransferases, polysaccharide lyases and carbohydrate esterases which digest the polysaccharides with a high specificity, often assisted by carbohydrate-binding modules in multidomain proteins. These protein groups are summarized as carbohydrate-active enzymes (CAZymes) (Bäumgen *et al.*, 2021; Drula *et al.*, 2022). Anionic polymers require additional enzymes including sulfatases for utilizing i.e., ulvan, carrageenan, porphyran and fucose-containing sulfated polysaccharides. For laminarin, up to three PULs have been found in a single genome in free-living bacteria, i.e., in *Formosa* clade B (Unfried et al., 2018). An *in-silico* study on 53 flavobacterial genomes identified PULs for laminarin utilization in free-living flavobacterial strains, but not in 10 out of 12 strains in a clade of genera of particle-associated strains, mainly isolated from a 20 μm particle net catch (Kappelmann *et al.*, 2019).

Particle-associated bacteria include chemotactic motile free-living bacteria and particle attached bacteria (Heins *et al.*, 2021a; Heins *et al.*, 2021b). From the genera

identified by size filtration and diatom settlement experiments as strongly particle-associated, we selected *Maribacter* to study laminarin degradation. *Maribacter* was first described in 2004 as a new genus in the family *Flavobacteriaceae* (Nedashkovskaya *et al.*, 2004). In nine samples taken over a spring algal bloom, 56 *Maribacter* strains were among 3572 isolated strains isolated from bulk seawater (Alejandre-Colomo *et al.*, 2020). In large comparative studies, *Maribacter* strains were predominantly isolated from particle fractions (> 3 µm filter retentate) and rarely from bulk seawater which was not size-fractionated and contained the natural population of particles (Heins and Harder, 2023; Lu *et al.*, 2023). Molecular community analyses showed that *Maribacter* are rare in free-living bacterial populations as defined by bacterioplankton fraction of 0.2 - 3 µm size in sequential filtration (Sidhu *et al.*, 2023) (and references therein). Abundances of up to 4% were detected in surface populations of oxic sandy sediment (Probandt *et al.*, 2018; Miksch *et al.*, 2021). Higher relative abundances were reported for phycosphere populations on micro- and macroalgae (Heins *et al.*, 2021b; Lu *et al.*, 2023; Wang *et al.*, 2024). In addition, being a dominant epiphyte on many algae, *Maribacter* synthesize with thallusin a controlling factor of *Ulva morphogenesis*. This demonstrates the interplay of particle-associated bacteria with algae in the phycosphere (Alsufyani *et al.*, 2020).

Currently 32 validly described species and more than 180 genomes have been published. In this study, two strains from seawater - *Maribacter forsetii* DSM 18668^T (Barbeyron *et al.*, 2008) and *Maribacter* sp. Hel_I_7 from the long-term ecological research station Kabeltonne near Helgoland – were investigated together with three strains from phytoplankton catches in the Wadden Sea near Sylt - *Maribacter* sp. MAR_2009_72, *Maribacter dokdonensis* MAR_2009_60 and *Maribacter dokdonensis* MAR_2009_71 (Hahnke and Harder, 2013). For all these strains representing four species (Supplement Table S1), the genomes have been published, which opened the opportunity to use proteomics for the identification and location of genes and proteins involved in laminarin degradation. We detected that genes for laminarin degradation are dispersed in the genomes of four particle-associated *Maribacter* species.

2.3 Material and Methods

2.3.1 Growth experiments

Maribacter forsetii DSM 18668[T] (Barbeyron *et al.*, 2008), *Maribacter* sp. Hel_I_7 (DSM 29657), *Maribacter* sp. MAR_2009_72 (DSM 29384), *Maribacter dokdonensis* MAR_2009_60 (DSM 29383) and *Maribacter dokdonensis* MAR_2009_71 (DSM 29385) (Supplement Table S1) were reactivated from glycerol stocks maintained in the laboratory since the isolation of the strains (Hahnke *et al.*, 2013; Hahnke *et al.*, 2015). All strains were grown in liquid HaHa_100V medium with 0.3 g/L casamino acids as sole carbon source, limiting the growth to an optical density (OD) at 600 nm below 0.2 (Hahnke *et al.*, 2015). The addition of 2 g/L of a carbohydrate source enabled growth beyond OD of 0.3. Laminarin from *Laminaria digitata* and glucose (Sigma Aldrich/Merck KGaA, Darmstadt, Germany) were used in growth experiments. In 250 ml photometer-sidearm Erlenmeyer flasks, 50 ml cultures were inoculated with 0.4% v/v of a pre-grown culture in the same medium and incubated with 110 rpm at room temperature. In addition to three cultures for proteomic analyses, a fourth culture was used to monitor bacterial growth by measuring optical and cell density. Cells were harvested at an OD of 0.15 for casamino acid cultures serving as control cultures and at an OD of 0.25 for sugar containing cultures. Cells were pelleted by centrifugation in 50 ml tubes with $3080x\,g$ for 30 min at 4°C. Pellets were resuspended in 1 ml medium and centrifuged in 1.5 ml tubes at 15870x g for 15 min at 4°C. The wet biomass was weighed and stored at -20°C.

To determine cell density, the cells were filtered onto a 47 mm (polycarbonate filter (0.2 μm pore size, Millipore, Darmstadt, Germany), applying a vacuum of 200 mbar. For microscopic cell counting, filter pieces were stained with DAPI (4,6-diamidino-2-phenylindole; Sigma-Aldrich, Steinheim, Germany) at a final concentration of 1 μg/ml. Subsequently, the filters were embedded in an antifading mounting solution of CitiFluor and Vectashield (3:1 v/v) (CitiFluor Ltd., London, United Kingdom; Vector Laboratories, Inc., Burlingame, CA, USA). The cell abundance was determined by counting the number of DAPI signals per counting grid with a defined area. A minimum of 30 grids were counted per filter piece (>280 cells), and

the total cell abundance was calculated by scaling the grid to the total filter area and divided by the sample volume.

2.3.2 Saccharide uptake visualization

Uptake studies with fluorescently labelled laminarin were performed with *Maribacter* strains, *Christiangramia* (formerly known as *Gramella*) *forsetii* DSM 17595 as positive control and *Escherichia coli* DSM 498 as negative control. *Flavobacteriia* were grown on carbon-limited (HaHa_100V medium with 0.3 g/L casamino acids as sole carbon source (CAA medium)) or including 20 mg/L laminarin to obtain induced cells. *E. coli* was grown in 1/10 of the marine medium (10% CAA medium in water) containing 0.5 g/L tryptone and 0.25 g/L yeast extract. Uptake experiments with 40 µM labelled laminarin were inoculated with OD (600nm) = 0.01. Non-induced and induced cells were also incubated without the labelled laminarin to determine cellular autofluorescence. 1 ml samples were taken after 1 min, 1 h, 2 h, 4 h and 24 h. After 1 h of fixation with 1% v/v formaldehyde, samples were diluted with 5 ml medium and cells were filtered onto 0.2 µm filter (diameter 47 mm) and stained with DAPI (1 ng/µl) for automatic microscopy (Reintjes *et al.*, 2017).

Automated imaging of cells used a Zeiss AxioImager.Z2 microscope stand (Carl Zeiss, Oberkochen, Germany) as described in Bennke *et al.* (2016). A 63 × magnification and 1.4 numerical aperture oil emersion plan apochromatic objective (Carl Zeiss)) was used with fixed exposure times of 50, 100, and 150 ms. Images were not optimized for Min/Max display normalization, but acquired in linear range 0-255 (based on 8-bit grey scale) to facilitate cross-comparison of signal intensities on images. Cell detection was performed using the ACMEtool software for automated cell enumeration (https://www.mpi-bremen.de/en/automated-microscopy.html#section19794) (Bennke *et al.*, 2016). Cells were identified by their DAPI signal and a signal at the same location on the filter in the laminarin images was considered as labelled cell. Positive cells had a minimum overlap of 30% area and a minimum area of 20 pixels (0.3 µm^2) (Reintjes *et al.*, 2017).

2.3.3 Protein preparation and mass spectrometry (MS)

Proteins were extracted from cells using a bead-beating method (Schultz *et al.*, 2020). A pellet of 20 to 200 mg wet weight was disintegrated with 0.25 ml glass beads (0.1 mm) in 500 µl lysis buffer. Protein was quantified using the Roti-Nanoquant assay (Carl Roth, Karlsruhe, Germany). For protein separation on denaturing polyacrylamide gels (SDS-PAGE), 50 µg of protein was mixed with 10 µl of 4x SDS buffer (20% glycerol, 100 mM Tris/HCl, 10% (w/v) SDS, 5% β-mercaptoethanol, 0.8% bromphenol blue, pH 6.8) and loaded on Tris-glycine-extended (TGX) precast 4-20% gels (Biorad, Neuried, Germany). Electrophoresis was performed for 8 min at 150 V. The gel was fixed in 10% v/v acetic acid and 40% v/v ethanol for 30 min, stained with Brilliant Blue G250 Coomassie and the protein band was excised. The proteins were dissected in one gel piece and the pieces were washed with 50 mM ammonium bicarbonate in 30% v/v acetonitrile. Evaporation in a SpeedVac (Eppendorf, Hamburg, Germany) yielded dry gel pieces which were reswollen with 2 ng/µl trypsin (sequencing grade trypsin, Promega, USA). After 15 min incubation at room temperature the excess liquid was removed. Samples were digested overnight at 37°C. The gel pieces were then covered with MS grade water and peptides were eluted with ultrasonication. The peptides were desalted using Pierce™ C18 Spin Tips (Thermofisher, Schwerte, Germany) according to the manufacturer's guidelines. The eluates were dried in the SpeedVac and then stored at -20°C. For the MS analysis the samples were thawed and taken up in 10 µl of Buffer A (99.9% acetonitrile + 0.1% acetic acid).

Tryptic peptides of *Maribacter forsetii* and *Maribacter* sp. MAR_2009_72 samples were analyzed on an EASYnLC 1200 coupled to a Q Exactive HF mass spectrometer (Thermo Fisher Scientific, Waltham, USA). Peptides were loaded onto a self-packed analytical column (3 µm C18 particles; Dr. Maisch GmbH, Ammerbuch, Germany) using buffer A (0.1% acetic acid) with a flow rate of 2 µL/min within 5 min and separated using an 85 min binary gradient from 4 to 50% buffer B (0.1% acetic acid in acetonitrile) and a flow rate of 300 nL/min. Samples were measured in parallel mode; survey scans in the Orbitrap were recorded with a resolution of 60,000 with m/z of 333-1650. The 15 most intense peaks per scan were selected for fragmentation. Precursor ions were dynamically excluded from

fragmentation for 30 sec. Single-charged ions as well as ions with unknown charge state were rejected. Internal lock mass calibration was applied (lock mass 445.12003 Da).

LC-MS/MS analyses of *Maribacter* sp. Hel_I_7, *Maribacter dokdonensis* MAR_2009_60 and *Maribacter dokdonensis* MAR_2009_71 samples were carried out on an EASY-nLC II coupled to a LTQ XL Orbitrap mass spectrometer (Thermo Fisher Scientific, Waltham, USA). Peptides were loaded onto a self-packed analytical column (3 μm C18 particles; Dr. Maisch GmbH, Ammerbuch, Germany) using buffer A (0.1% acetic acid) with a flow rate of 500 nL/min and separated using a 56 min binary gradient from 5 to 40% buffer B (0.1% acetic acid in acetonitrile) and a flow rate of 300 nL/min. Samples were measured in parallel mode; survey scans in the Orbitrap were recorded with a resolution of 30,000 with m/z of 300-2000. The 5 most intense peaks per scan were selected for fragmentation. Precursor ions were dynamically excluded from fragmentation for 30 sec. Single-charged ions as well as ions with unknown charge state were rejected. Internal lock mass calibration was applied (lock mass 445.12003 Da).

The mass spectrometry files were analyzed in MaxQuant version 2.2.0.0 in the standard settings - see supplement information for details - against the strain specific protein database downloaded from NCBI (see below) (Tyanova *et al.*, 2016; Sayers *et al.*, 2022) and common laboratory contaminants. Statistical analysis was performed in Perseus version 2.0.7.0 (Tyanova and Cox, 2018). Proteins were recognized as being expressed when they had label free quantification (LFQ) intensities in one out of three biological replicates. No proteins were retrieved from the control cultures of MAR_2009_72. Analysis was performed with a biomass grown on galactose. Proteins were considered being upregulated when the p-value and the difference were positive in the statistical t-test analysis (Supplement Tables S8-S12).

Reference genomes were downloaded from NCBI: *Maribacter forsetii* NZ_JQLH01000001.1 (download 2022/10/06), *Maribacter* sp. MAR_2009_72 NZ_VISB01000001.1 (2022/09/27), *Maribacter* sp. Hel_I_7 NZ_JHZW01000001.1, NZ_JHZW01000002.1 and NZ_JHZW01000003.1 (2022/10/06), *Maribacter dokdonensis* MAR_2009_60 NZ_LT629754.1 (2022/10/06) and *Maribacter dokdonensis* MAR_2009_71 NZ_FNTB00000000.1 (2022/07/18).

2.3.4 Bioinformatic analyses

Protein annotation was refined using several of databases. CAZymes were considered to be identified if two out of three search algorithms in dbCAN3 were positive in the web interface search (Zheng *et al.*, 2023). The conserved domain database (CDD) (Lu *et al.*, 2019), the SulfAtlas web interface (Stam *et al.*, 2022), InterPro (Paysan-Lafosse *et al.*, 2022), PULDB (Terrapon *et al.*, 2018), deepTMHMM (Hallgren *et al.*, 2022), SignalP (Teufel *et al.*, 2022), Blastkoala (Kanehisa *et al.*, 2016) and UniProt (The UniProt, 2023) provided additional information.

Alignment for the SusC and GH16_3 protein tree were performed using MAFFT online in automode (see supplement information for details) (Katoh *et al.*, 2019). SusC proteins encoded in laminarin utilizing loci were retrieved from publications (Krüger *et al.*, 2019). The GH16_3 sequence were taken from CAZy (Cantarel *et al.*, 2009). The trees were calculated using the maximum likelihood method in Mega 11 with the default settings (Tamura *et al.*, 2021a).

Proteins homolog to the laminarinase of *M. forsetii* (WP_051941695.1) were identified by BlastP using the NCBI web service (Altschul *et al.*, 1990). Genomes were downloaded from NCBI and the annotations of seven open reading frames (ORFs) upstream and of seven ORFs downstream of the homologous GH16_3 genes were analyzed manually for the presence of polysaccharide utilization loci as defined by a SusC/D pair and at least one CAZymes within eight ORFs.

For the visualization of the data the following programs and packages were used: R version 4.3.2 (R Core Team, 2023), ggplot2 (Wickham, 2016), gggenes (Wilkins, 2023) and Proksee (Grant *et al.*, 2023).

2.3.5 Enzyme activity tests and sugar quantifications

Cell pellets stored at -20°C were lyzed in 1 mL 50 mM 3-(N-morpholino) propanesulfonic acid (MOPS) by sonification on ice using a Sonoplus HD70 Bandelin MS73 (BANDELIN, Berlin, Deutschland) with a titanium sonotrode MS73 (BANDELIN) for 4 min at 50% power and 50% cycle (0.5 sec on, 0.5 sec off). The soluble fraction was separated by centrifugation for 15 min at 16.000x *g* at 21°C.

Enzyme were inactivated by pasteurization (80°C for 1 h). Protein concentration was determined by the Bradford protein assay using bovine serum albumin as standard (Bradford, 1976). Laminarinase assays were performed with 10 to 25 µg protein - of cell lysate, soluble fractions or resuspended membrane fractions - and 250 µg laminarin in 100 µl MOPS for 20 h at 21°C. After vacuum drying at 45°C for 1 hour (Eppendorf Concentrator plus, Eppendorf, Hamburg, Germany), the sugars were labelled with 2 µl of 0.15 M 8-amino-1,3,6-naphthalenetrisulfonate (ANTS) in the presence of cyanoborohydride and analyzed on 25% v/v acrylamide gels following a protocol for fluorophore-assisted carbohydrate electrophoresis (FACE) (Becker *et al.*, 2017). Separation was performed at 100 V for 30 minutes, followed by 200 V for 60 minutes. Gels were documented using a Bio-Rad GelDoc EZ Gel Imaging System (Cambridge Scientifix, Watertown, US).

The activity of β-glucosidase was detected with 1 mM 4-nitrophenyl-glucopyranoside in 50 mM potassium phosphate, pH 7.0, and the quantification of nitrophenol at 405 nm using an extinction coefficient of 7500 M^{-1} cm^{-1}.

Glucose concentrations were quantified with a glucose oxidase-peroxidase assay in 0.2 ml microtiter wells (Elabscience, Wuhan, China) using a spectrophotometer (SPECTROstar® Nano, BMG LABTECH, Ortenberg, Germany). Total carbohydrate concentration was determined by the phenol–sulfuric acid method with glucose as standard, adapted for a microtiter plate (DuBois *et al.*, 1956). 25 µl of sample were mixed with 15 µl of 5% phenol and then acidified with 100 µl of concentrated sulfuric acid. The microtiter plate was incubated for 20 min at 30 °C and the absorbance was measured at 490 nm using a plate spectrophotometer.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol *et al.*, 2022) partner repository with the dataset identifier: *Maribacter forsetii* PXD049038 and 10.6019/PXD049038; *Maribacter* sp. MAR_2009_72 PXD049039 and 10.6019/PXD049039; *Maribacter dokdonensis* MAR_2009_60 PXD049040 and 10.6019/PXD049040; *Maribacter dokdonensis* MAR_2009_71 PXD049041 and 10.6019/PXD049041 and *Maribacter* sp. Hel_I_7 PXD049042 and 10.6019/PXD049042.

2. 4 Results

2.4.1 Growth of *Maribacter forsetii* on laminarin

Maribacter forsetii KT02ds 18-6^T (DSM 18668 ^{T}) grew with a maximal growth rate of $\mu = 0.09$ h^{-1} at room temperature in a liquid mineral medium containing 0.3 g/L carbohydrate-free casamino acids as limiting carbon source. The cell density reached an OD of 0.196. The addition of 2 g/L glucose or laminarin enabled growth rates of 0.10 h^{-1} and 0.09 h^{-1} and maximal OD of 0.578 and 0.501, respectively. The proliferation of cells was confirmed by counting DAPI-stained cells (Figure 1). Cell concentrations correlated with the optical density. Quantification of reactive aldehyde groups by the phenol-sulfuric acid method revealed a consumption of 42% for glucose and 62% for laminarin. Small amounts of reactive aldehydes (0.38 mM) were transiently detected during growth on casamino acids. These observations established an *in vivo* metabolism of laminarin by *Maribacter forsetii*.
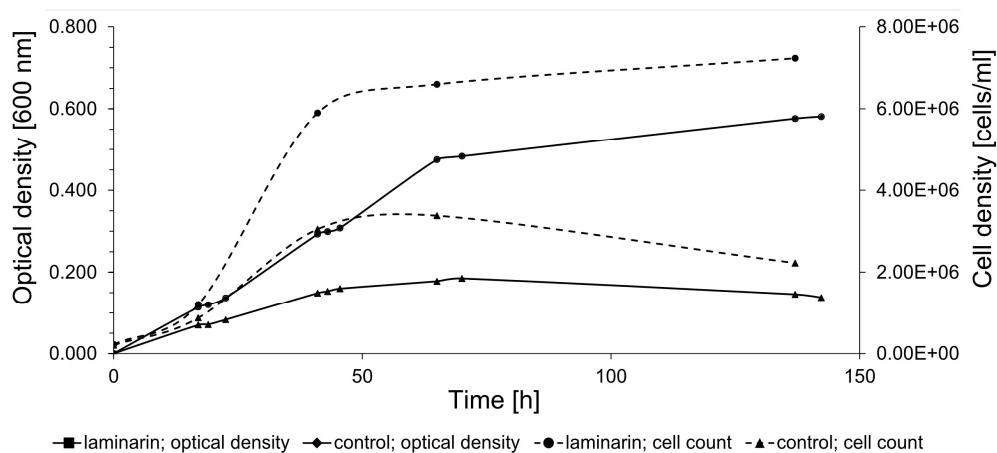
Figure 1: Growth curve of *Maribacter forsetii* in modified HaHa_100V medium with 2 g/L laminarin at room temperature. The control culture has a limited amount of casamino acids. The optical density was measured at 600 nm and the cell density was determined by cell counts using DAPI staining.

2.4.2 *Maribacter forsetii* utilizes specific proteins for laminarin degradation

The proteomes of the differently grown cultures showed significant variations. Overall, 1971 proteins were detected in laminarin grown biomass in at least one of the three biological replicates. We observed 1974 proteins in glucose grown cells and 2018 proteins in cells from the control experiment grown with a limited amount of casamino acids. 1748 proteins were detected at least once for each substrate. The analysis showed that 59 proteins were identified only in glucose grown cells, 53 only in laminarin grown cells and 125 in the control culture biomass (Figure 2A, Supplement Table S3). A principal component analysis of the proteomes revealed a greater difference between the control and the sugar-derived proteomes and a smaller difference between the proteomes of glucose and of laminarin grown cells (Figure 2B).
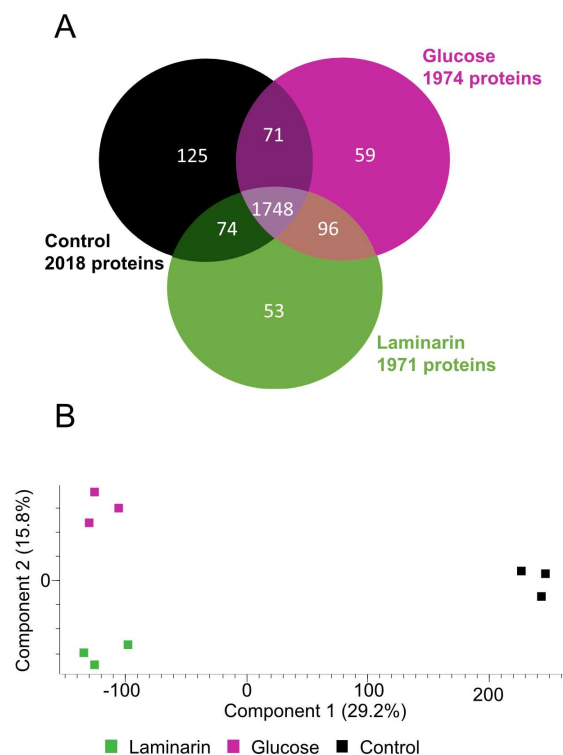
Figure 2: Comparison of three *Maribacter forsetii* proteomes of glucose, laminarin and casamino acid grown cells. A: Venn diagram showing the overlap of detected proteins in at least one of three biological replicates. B: Principal component analysis showing the differences between the detected proteomes.

Proteins of the basic cellular metabolism with an upregulation in the presence of laminarin included the enzymes involved in glycolysis, citric acid cycle and oxidative

phosphorylation as well as the pentose phosphate pathway. In addition, several proteins with homologies to characterized enzymes for the degradation of laminarin were upregulated (Figure 3; Table 1 and Supplement Table S8).

Table 1: Results of two sample t-test calculated for volcano plot visualization in Perseus (version 2.0.7.0) for enzymes highlighted in Figure 3, possible key role players in laminarin degradation in *Maribacter forsetii*. Difference indicates the difference between the label free quantification values in log2 of the laminarin proteome versus the control proteome (mean values of the three biological replicates); p-value in -log10; significance: + sign shows if the specific lies above or below the significance (p-value) threshold. Missing observations were imputed in Perseus from NaN (Not a Number) to constant value 0. False discovery rate (FDR): 0.05, Artificial within groups variance (default: 0). It controls the relative importance of t-test p-value and difference between means. At s0=0 only the p-value matters, while at nonzero s0 also the difference of means plays a role (s0: 0.15) (Tusher *et al.*, 2001)

Product	Difference	p-value	Significance	Protein id	Locus Tag (NCBI)
GH16_3	3.20	3.20	+	WP_051941695.1	P177_RS20320
SusC	27.92	9.83	+	WP_245233002.1	P177_RS04595
SusD	27.91	7.65	+	WP_036152296.1	P177_RS04600
GH3	0.46	0.60		WP_157486485.1	P177_RS07335
ABC substrate binding	26.00	7.55	+	WP_036150709.1	P177_RS00465
ABC permease	0.72	1.07		WP_036156362.1	P177_RS15985
ABC ATP-binding	0.26	0.38		WP_036150993.1	P177_RS01135
ExbD	1.05	1.25		WP_036151815.1	P177_RS03320
ExbB	0.49	0.67		WP_036151807.1	P177_RS03305
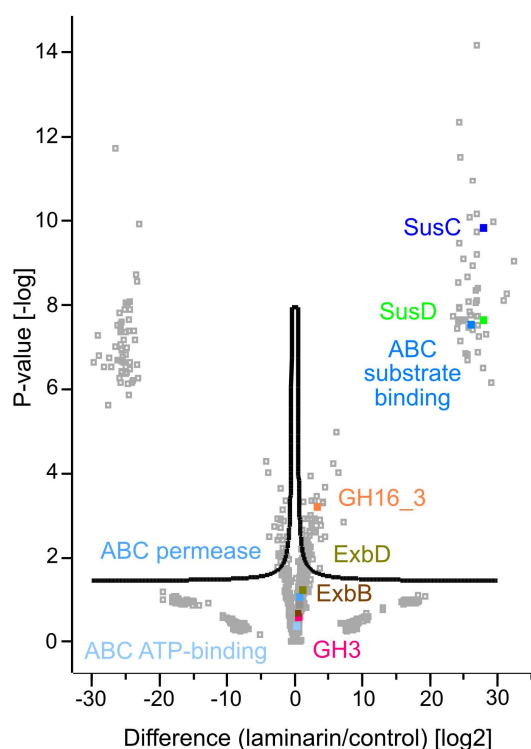
Figure 3: Volcano plot showing the statistical analysis of up and downregulated proteins of laminarin and casamino acid grown *Maribacter forsetii* cells. Label free quantification (LFQ) intensities were transformed to log2. The average mean value of the three biological replicates of either laminarin or control was used for the calculations. Missing observations were imputed in Perseus from NaN to constant value 0. FDR: 0.05, s0:0.15.

A glycosylhydrolase family GH16_3, assigned to laminarinases as a β-glucan degrading endohydrolase (P177_RS20320), was upregulated in comparison to the control sample. The protein was predicted to be extracellular, as expected for the degradation of the polysaccharide. It was not detected in the proteome of glucose grown cells. To support the GH16_3 characterization as an endo-β-1,3-laminarinase of *Maribacter forsetii*, we calculated a phylogenetic tree to examine the association with other GH16_3 proteins that have been characterized experimentally as either endo-β-1,3-laminarinase or endo-β-1,4-galactosidase (Supplement Figure 1). The GH16_3 of *M. forsetii* was clustered together with GH16_3 from the other investigated *Maribacter* strains and with experimentally verified laminarinases. Separate branches included GH16_3 galactosidases as well as GH16_12 and GH16_16 of *Maribacter dokdonensis* MAR_2009_60 and MAR_2009_71.

The expected product of GH16_3 activity is a range of laminarin oligosaccharides which are actively transported into the periplasm by a TonB-dependent SusC/D transporter system (Labourel *et al.*, 2015). Several SusC and SusD proteins were detected in the proteomes. The energy-delivering protein complex in the cytoplasmic membrane was expressed with two copies of ExbB (P177_RS03305, P177_RS16115) and three copies of ExbD (P177_RS03320, P177_RS03315, P177_RS16110) in the proteome. ExbBD proteins were upregulated in laminarin grown cells (Supplement Table S3). The SusC protein P177_RS04595 had the highest upregulation, followed by three other SusC proteins. The first one has a specific hit in the reverse position specific iterated BLAST of NCBI-CDD, whereas the other three SusC proteins have only one superfamily hit, indicating that the annotation of the transporters is not clearly established. SusD (P177_RS04600) of the respective *susCD* gene pair was also highly upregulated. In addition, four more SusD proteins were upregulated. The aforementioned SusC/D pair was detected in control cells and expressed in glucose grown cells. These observations suggest the use of a SusC/D transport system in the laminarin metabolism.

The conversion of oligomeric laminarin to glucose in the periplasm can be catalysed by a number of glycosyl hydrolases. Genome analysis revealed one GH1, two GH3 and two GH5 proteins, however, only one GH3 (P1177_RS07335) was expressed and upregulated in laminarin compared to glucose or control.

A candidate for the transport into the cell is an ABC transport system. The substrate binding protein P177_RS00465 was highly expressed in laminarin and in glucose grown cells. The permease P177_RS15985 and the ATP-binding protein P177_RS01135 may complete the ABC transporter. No major facilitator superfamily transporters were expressed in cells grown on laminarin, glucose or casamino acids.

The phosphorylation of glucose to glucose-6-phosphate is a kinase reaction. We found several expressed and upregulated kinases in the genome. The most upregulated kinase was classified as a carbohydrate kinase (P177_RS13360). The pathway tool Kegg BlastKoala suggested the polyphosphate glucose phosphotransferase P177_RS04875. This enzyme is only upregulated in comparison to the control.

These findings suggested a degradation pathway of laminarin by *Maribacter forsetii* similar to the pathways in free-living *Flavobacteriia* (Figure 4).
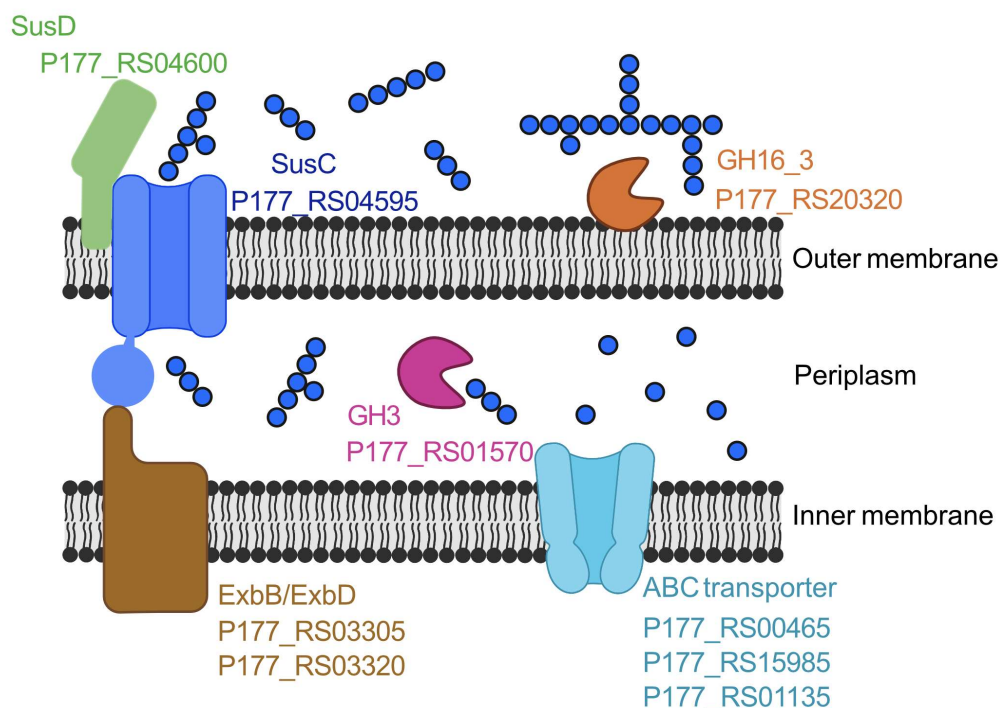
Figure 4: Proposed laminarin degradation pathway in *Maribacter forsetii*. GH16_3: laminarinase, P177_RS20320; SusD: nutrient uptake outer membrane protein (P177_RS04600); SusC: TonB-linked outer membrane protein (P177_RS04595); GH3: β-glucosidase (P177_RS01570); ExbB: proton channel family protein (P177_RS03305); ExbD: biopolymer transporter (P177_RS03320); ABC Transporter: ABC transporter substrate binding (P177_RS00465), ABC transporter permease (P177RS_15985), ABC transporter ATP-binding protein (P177_RS01135). Created with BioRender.com.

The genes for these proteins are spread over the genome and not localized in a polysaccharide utilisation locus (Figure 5).
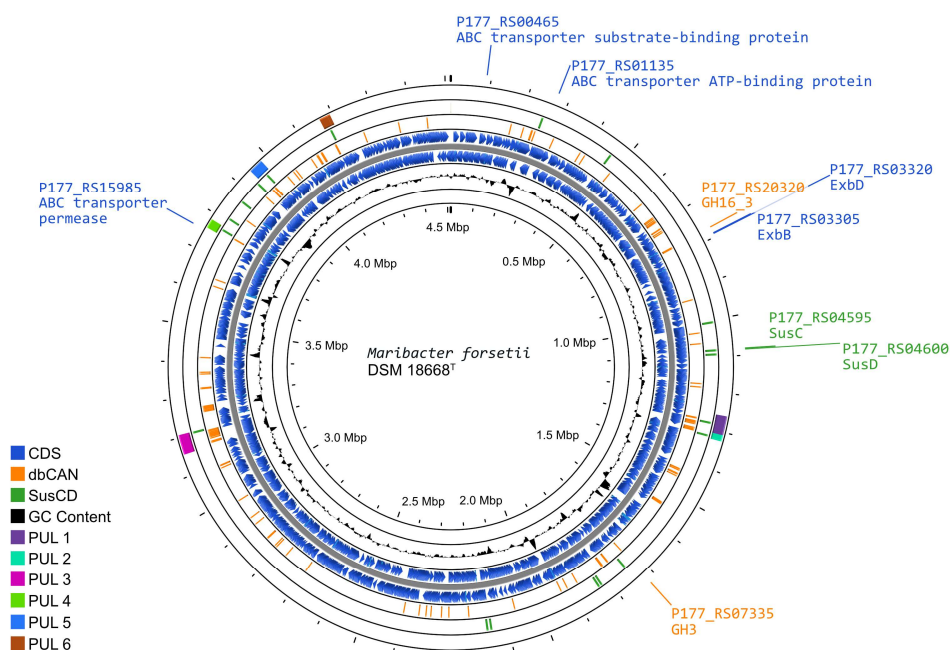
Figure 5: Representation of the *Maribacter forsetii* genome, displaying the overall genome structure, as well as the enzymes playing a possible key role in laminarin degradation. The rings (outside to inside), present (1) location and length of polysaccharide utilization loci (PUL), defined by PULDB and Kappelmann et al. (2019) (purple); (2) location of SusC and SusD genes, with SusCs having the motif Tigr04056 (green); (3) genes for CAZymes annotated by dbCAN3, annotated in at least two out of three search algorithms (orange); (4/5) coding sequences in forward and reverse direction from NCBI genome annotation of *Maribacter forsetii* (blue); (6) G/C content

The neighbouring genes of the GH16_3 gene (P177_RS20320, Figure 6) were not expressed in laminarin grown cells. None of the SusC and SusD genes has a CAZyme in the neighborhood in the same reading direction. Other neighboring genes were expressed, i.e., for the SusC/D pair with the highest expression values in laminarin-grown cells (P177_RS04595 and P177_RS04600), an alkaline phosphatase was expressed in laminarin and glucose, a FAD-dependent oxidoreductase in two out of three glucose samples and a phosphonatase-like hydrolase in two out three replicates in laminarin and glucose. The GH3 gene (P177_RS07335) has a neighbouring anhydro-N-acetylmuramic acid kinase gene and more distantly a GH10 gene in the same reading direction. In coincidence with these findings, the genes for the ExbBD and ABC transporter proteins are widely distributed across the genome.
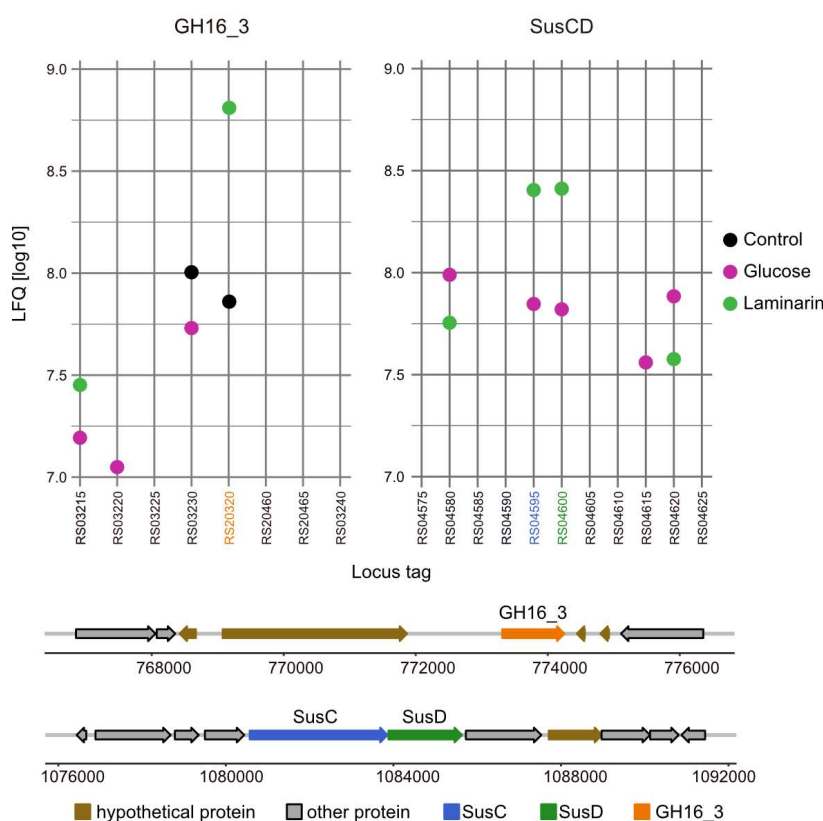
Figure 6: Gene organization and expression of enzymes proposed to degrade laminarin in *Maribacter forsetii*. Expression data are shown in LFQ values (log10) of the genetic region containing genes of GH16_3 and of the highest upregulated SusC/D pair of *Maribacter forsetii*.

2.4.3 Enzyme activities

We investigated the presence of endo-laminarinase activity in cell extracts using the labelling of sugars with a fluorophore and size separation by acrylamide gel electrophoresis. Extracts of *Maribacter forsetii* and *Maribacter* sp. Hel_I_7, *Maribacter dokdonensis* MAR_2009_60 and *Maribacter* sp. MAR_2009_72 catalyzed the formation of a size range of oligosaccharides (Figure 7A), but the enzymes of *Maribacter dokdonensis* MAR_2009_71 did not show this laminarin hydrolysis (Figure 7B). The oligosaccharide formation is the characteristic pattern expected for an endo-laminarinase (Labourel *et al.*, 2014; Labourel *et al.*, 2015). The activity was heat-labile, increased with increasing protein concentration and was membrane-associated. It partly dissolved in MOPS-buffer that had less ionic strength than seawater. Efficient removal of the enzyme activity from the membrane

fraction was obtained by washing the membrane particles with detergents, including n-dodecyl-β-D-maltopyranoside, Tween 20, Tween 80 and Triton X-100 in buffers as described in Orwick-Rydmark *et al.* (2016). The endo-laminarinase activity of four strains confirmed the expression of the GH16_3 encoded in their respective genomes.

Extracts of *Maribacter dokdonensis* MAR_2009_71 showed a laminarin-dependent formation of a molecule with the retention time of glucose in FACE gels, catalyzed by the soluble fraction of the crude extract (Figure 7B). The formation of glucose by heat-sensitive enzymes was quantified by a glucose oxidase peroxidase assay. Hydrolysis of 4-nitrophenyl-ß-D-glucopyranosid revealed a heat-labile β-glucosidase activity of 18 nmol s^{-1} (μg protein)$^{-1}$ in soluble extracts. These observations indicated a different degradation pathway for laminarin in *Maribacter dokdonensis* MAR_2009_71. In addition, the absence of a GH16_3 gene in the genome coincided with the absence of an oligosaccharide ladder in the enzyme assay.
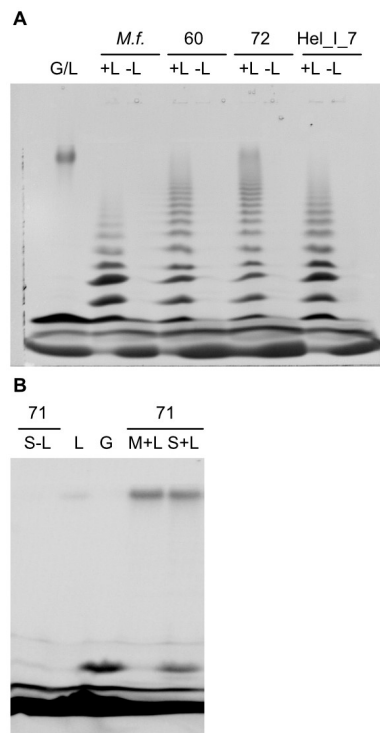
Figure 7: Hydrolysis of laminarin by soluble extracts of *Maribacter* visualized as fluorescently labelled sugars separated by polyacrylamide gel electrophoresis. A) strains with endo-laminarinase activity: G/L: standard of laminarin and glucose, +L/-L: assay with or without laminarin, *M.f.*: *Maribacter forsetii*, 60: MAR_2009_60, 72: MAR_2009_72, Hel_I_7, B) 71: MAR_2009_71 has no endolaminarinase activity: S-L: assay of soluble fraction without laminarin, L: laminarin standard, G: glucose standard, M+L: assay of membrane fraction with laminarin, S+L: assay of soluble fraction with laminarin

2.4.4 Genomes and proteomes of other *Maribacter* strains

Homologues of the laminarin-degrading proteins of *Maribacter forsetii* were also detected in the other investigated *Maribacter* strains (Table 2 and Supplement Table S4-S7). The endo-laminarinase activity coincided with the GH16_3 genes in the four genomes. The proteomic analyses revealed higher LFQ values for the corresponding proteins only for *Maribacter forsetii* and MAR_2009_72 grown on laminarin. It was also detected in the proteomes of Hel_I_7, but not in those of MAR_2009_60. This showed the sensitivity limitations of proteomics in this study in comparison to enzyme assays. No other genes for an endo-laminarinase were detected in the genomes. Other glycosyl hydrolase family GH16 genes were GH16_12 and GH16_16 in MAR_2009_60 and MAR_2009_71, annotated in the respective genomes as β-porphyranase and β-agarase (Viborg *et al.*, 2019).

Table 2: Comparison of proteomic data sets of laminarin-grown cells of Maribacter forsetii, Maribacter sp. MAR_2009_72, Maribacter dokdonensis MAR_2009_60, Maribacter dokdonensis MAR_2009_71, and Maribacter sp. Hel_I_7; numbers of proteins expressed in laminarin-grown cells/annotated in the genome.

	M. forsetii	MAR_2009_72	MAR_2009_60	MAR_2009_71	Hel_I_7
GH16_3	1/1	1/1	0/1	0/0	0/1
GH16_12	0/0	0/0	0/1	0/1	0/0
GH16_16	0/0	0/0	0/1	0/1	0/0
SusC	10/19	10/22	7/23	5/24	6/16
SusC-specific hit	4/12	5/14	4/17	2/15	2/12
SusC-superfamily hit	6/7	5/8	3/6	3/9	4/4
SusD	12/20	8/22	7/23	9/24	6/21
GH1	0/1	0/3	0/1	0/1	0/1
GH3	1/4	2/6	0/4	0/5	0/5
GH5	0/2	0/2	0/1	0/2	0/2
GH17	0/0	0/0	0/0	0/0	0/0
GH30	0/0	0/0	0/0	0/0	0/0
ExbB	2/2	2/2	2/2	2/2	1/2
ExbD	3/5	3/5	3/3	3/3	2/5
ABC substrate binding	2/4	1/1	0/3	0/3	0/2
ABC permease	2/19	1/17	0/20	0/21	0/21
ABC ATP-binding	5/17	8/16	1/19	1/18	2/19

The degradation of oligomeric laminarin can be performed by enzymes affiliating to GH1, GH3 and GH5, which were encoded in all five genomes, sometimes in multiple copies. Only MAR_2009_72 had expressed GH3 proteins, in this case two GH3 proteins. ABC transporters were encoded in all genomes. A complete ABC transport system was expressed in MAR_2009_72 besides *Maribacter forsetii* and ABC ATP-binding proteins were expressed in the other three strains. All five *Maribacter* strains had an upregulated polyphosphate glucose phosphotransferase, classified as EC 2.7.1.63, most likely responsible for the conversion of glucose to glucose-6-phosphate after the transport into the cytoplasm.

The contrasting enzymatic activities of the two strains of *Maribacter dokdonensis* in our study have likely a molecular basis in the strain-specific part of the genomes. MAR_2009_60 has an endo-laminarinase GH16_3 and its degradation pathway appears to be similar to *Maribacter forsetii*. To characterize strain MAR_2009_71, we analyzed the proteomes and especially the strain-specific genes. The genome encodes 228 proteins that are absent in MAR_2009_60 and 406 proteins with less than 70% identity to the best Blast hit among the proteins in MAR_2009_60. The strain specific genes are located in one large island of 206 genes and several smaller islands of 9 to 44 genes, besides individual genes. The genome encodes 60 glycosyl hydrolases, including 10 strain-specific ones. dbCAN predicted 47 glycosyl transferases, including 10 GT4 and 4 other GT with a strain-specificity. One strain-specific GT4 (BLW30_RS12510) was expressed in laminarin grown cells. The gene is part of a strain-specific genetic island of 26 genes (BLW30_RS12470 - BLW30_RS12590) comprising 9 upregulated genes including a sugar transferase, a second GT4 and a nucleotide sugar dehydrogenase. A GT4 with 57% identity to the aforementioned enzyme was upregulated in MAR_2009_60.

In MAR_2009_71, the other upregulated GH or GT was a GH109 (BLW30_RS06230), whereas the homolog in MAR_2009_60 (BLT83_RS16160) was not upregulated on laminarin compared to the control. Proteomic analysis of MAR_2009_71 showed the upregulation of three SusC, four SusD and a sugar phosphate isomerase/epimerase (Supplement Table S6). A significant upregulation in both MAR_2009_71 and MAR_2009_60 was detected for three proteins annotated as SusC, an outer membrane protein and a sensor protein.

All strains had expressed and upregulated SusC/D proteins together with expressed ExbBD. In earlier studies, the substrate specificity of SusC proteins was predicted largely by proteomic and phylogenetic studies. We calculated a protein tree to see whether the upregulated SusCs from the investigated *Maribacter* strains were affiliated with published SusC proteins that had been identified to be involved in laminarin degradation (Figure 8, Supplement Table S2). SusC proteins encoded in laminarin specific PULs from other *Flavobacteriia* formed together with *Maribacter* SusC proteins with a specific hit for the TIGR04056 SusC motif one branch in the tree. Separated were the SusC proteins with only a superfamily hit forming a second clade close to the outgroup. *Maribacter dokdonensis* MAR_2009_60 has three upregulated SusC proteins with a specific hit, whereas the other strains have only one SusC for the transport of laminarin-derived oligosaccharides. The *Maribacter forsetii* SusC is related to one of *Maribacter dokdonensis* MAR_2009_60. The other two SusC from MAR_2009_60 affiliate with the one of MAR_2009_71, forming a species-specific branch. A third branch is formed by MAR_2009_72 and Hel_I_7 together with proteins of *Christiangramia* (ex *Gramella*).
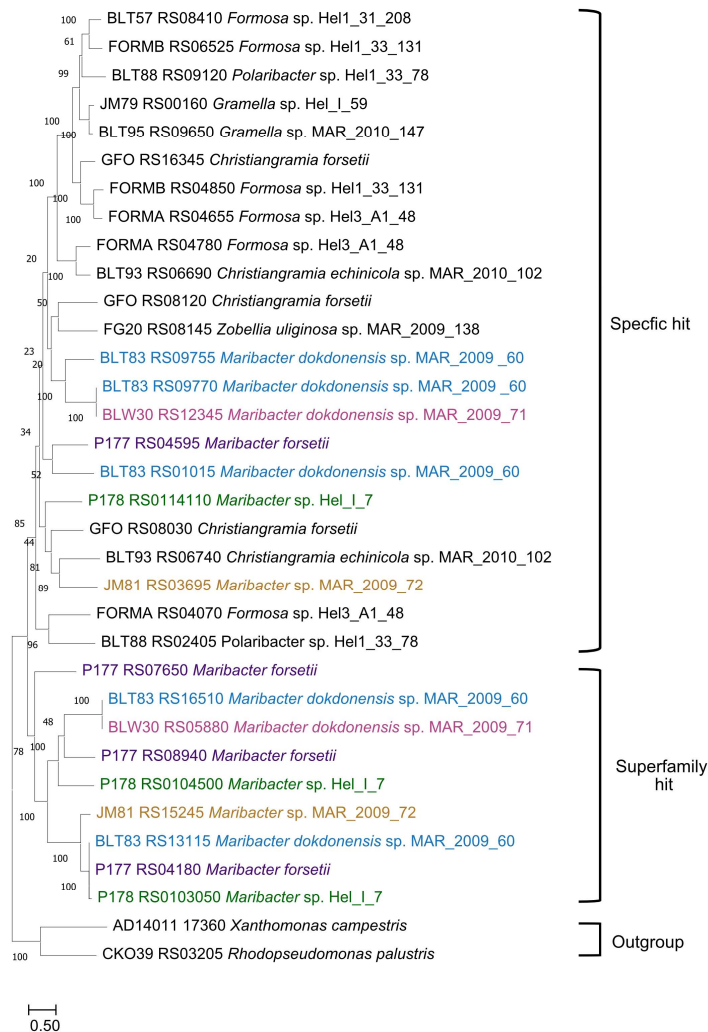
Figure 8: Phylogenetic tree showing the likeness between previous published SusCs affiliated with laminarin degradation and upregulated SusCs from the five *Maribacter* strains in this study.
The evolutionary history was inferred by using the Maximum Likelihood method and Le_Gascuel_2008 model. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the JTT model, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (4 categories (+*G*, parameter = 5.1185)). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site.
Each SusC was analyses using the CDD for identification of its specificity. The strain nomenclature was used as provided in NCBI (Feb.2024). *Maribacter forsetii* purple; *Maribacter* sp. Hel_I_7 green; *Maribacter* sp. MAR_2009_72 orange; *Maribacter dokdonensis* MAR_2009_60 blue; *Maribacter dokdonensis* MAR_2009_71 pink.

2.4.5 Uptake of fluorescently labelled laminarin

The energy driven uptake of oligosaccharides via SusC/D proteins in the outer membrane is a key factor for the ecological success of *Flavobacteriia*. This transport has been demonstrated by the accumulation of fluorescently labelled sugars in the

periplasm of free-living planktonic bacteria, referred to as selfish uptake (Reintjes *et al.*, 2017; Fischer *et al.*, 2019). None of the here investigated *Maribacter* strains showed selfish uptake following the standard protocol (Reintjes *et al.*, 2017). To introduce a quantitative measure for this negative result, we switched from a threshold-integration method to a linear data acquisition and used the mean grey value (MGV) of the cells in the detection of the fluorescent signal (Figure 9). No specific uptake of laminarin was detectable for *Maribacter* strains and *E. coli DSM 498* which served as negative control. In both cases the automatic cell detection system identified a limited number of cells, around 10%, as positive. These cells had a MGV 10 units above the MGV of the background of the individual slide, in the overall range from 0 (black) to 255 (white). In contrast, *Christiangramia forsetii* cells showed a clear fluorescence signal for the selfish uptake by (i) an inducible process, (ii) a high cell number (~ 90% of DAPI cells) and (iii) a clear intensity 40 units above the background.
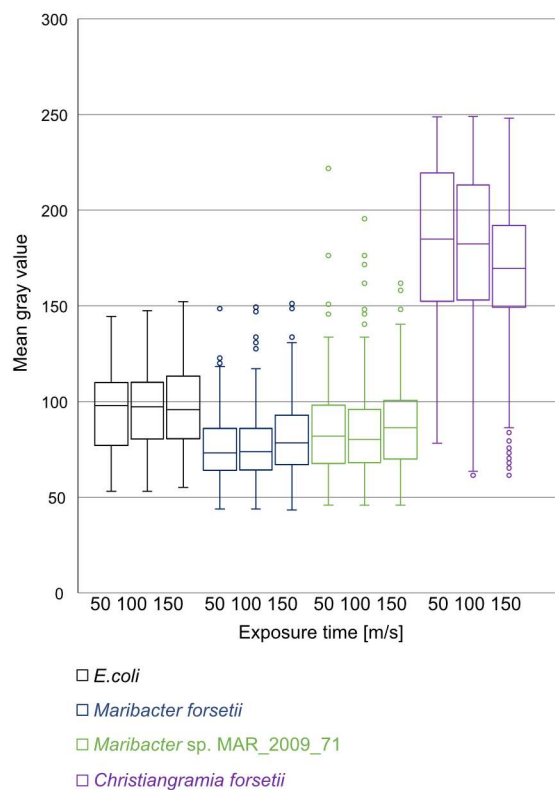
Figure 9: Mean grey value distribution of cells automatically detected as positive for uptake of labelled laminarin at 50, 100 and 150 ms exposure time. Images of cells samples after 1 min, 1, 2, 4 and 24 h were processed and automatically counted. Cell numbers were 68, 105 and 104 for *E.coli*, 261, 365 and 343 for *Maribacter forsetii*, 217, 308 and 332 for *Maribacter dokdonensis* MAR_2009_71 and 775, 2459 and 1635 for *Christiangramia forsetii* and represent areas of ten filters identical in size.

2.4.6 *M. forsetii* laminarinase homologues

The non-redundant protein dataset of NCBI included 57 homologous proteins that contained all domains of the *M. forsetii* laminarinase. Length varied between 687 and 705 AA. Lowest protein identity was 40%. *Maribacter* genomes encode 40 different homologues (Supplement Table S13). The genes were not engulfed within PULs in all *Maribacter* genomes analysed, including *M. algarum*, *M. aquivivus* DSM 16478, *M. arcticus* DSM 23546, *M. confluentis*, *M. huludaoensis*, *M. hydrothermalis*, *M. litoralis*, *M. sedimenticola* DSM 19840, *M. spongiicola*, *M. stanieri* DSM 19891, *M. zhoushanensis*, *M.* sp. 6B07, *M.* sp. 1_MG-2023, *M.* sp. BPC-D8, *M.* sp. MJ134 and *M.* sp. 4U21. Outside of *Maribacter*, homologues in *Aurantibacter* sp., *Eudoraea adriatica*, *E. chungangensis*, *Zobellia uliginosa*, *Z. barbeyronii*, *Croceitalea vernalis* and two *Flavobacteriaceae* (ASW18X and TMED265) were not part of a PUL. The large gene was encoded within PULs in *Croceitalea dokdonensis*, *C.* sp P059, *Croceivirga radicis* and *Flavobacterium* ASW18X (86.3% ANI to *C. radicis*). Besides a SusC/D pair, these PULs had either a catalytic GH16 domain protein or GH2, GH17 and several carbohydrate-binding module (CBM) proteins. *Croceitalea* sp. MZPC5 had a CAZyme cluster including a GH16 catalytic domain protein, GH17, two CBM proteins and an MFS transporter.

2.5 Discussion

Our study provided insight into the laminarin utilization in particle-associated *Maribacter* strains. They are capable of growing on laminarin as major carbon source. A previous *in-silico* study did not detect laminarin PULs in *Maribacter* genomes (Kappelmann *et al.*, 2019). Analysing the proteomes, we detected the expression of enzymes known to be involved in laminarin degradation. Their genes have, however, a genetic neighbourhood that did not fulfil definitions for canonical PULs: physically linked CAZyme genes located around a *susCD* pair whereby different search windows were applied (Grondin *et al.*, 2017; Unfried *et al.*, 2018; Krüger *et al.*, 2019; Francis *et al.*, 2021; Sidhu *et al.*, 2023). Our study demonstrated that the absence of genetic loci should not be used to predict physiological traits. In

our understanding, laminarin is not a major substrate for particle-associated *Maribacter* and therefore there was no selection for tight coregulation in one genetic locus.

The expressed GH16_3 is annotated as laminarinase subfamily, a large and diverse subfamily, which can be found in all four kingdoms (Viborg *et al.*, 2019). The GH16-laminarinase-like domain is the active site of hydrolysis. It has a size of 230 AA and is highly conserved, however, proteins with this domain are present in proteins up to over 1700 AA in nearly 50 protein superfamilies, usually together with specific carbohydrate binding domains or as yet unannotated regions, according to protein family models in CDD of NCBI. All these proteins are usually annotated in genomes as family 16 glycosylhydrolase or glycoside hydrolase family 16 protein on the basis of the presence of the GH16-laminarinase-like domain.

The GH16_3 enzyme of *Maribacter forsetii* and of other *Maribacter* strains is a large protein of ~ 700 AA and has homologues in some flavobacterial genomes. Our analysis resulted in the following annotation for the *Maribacter forsetii* GH16_3: a flexible N-terminal signal peptide for the sec secretion system (AA 1-20) with a lipid attachment site (AA 1-18), a polycistronic kidney disease (PKD) domain at AA 48-109, a laminarin-binding domain and the catalytic site GH16-laminarin-like domain that is interrupted by a domain of unknown function. The seven-strand β-sandwich domain annotated as PKD domain can bind the protein onto the outer membrane because the characteristic feature of β-sandwiches is a hydrophobic site (Bycroft *et al.*, 1999; Cheng *et al.*, 2013). Laminarin binding by the whole N-terminus including the second domain (AA 6-294) is based on homology to the laminarin-binding surface glucan binding protein B of *Bacteroides fluxus* BfSGBP-B (Tamura *et al.*, 2021b). In the *Bacteroides fluxus* PUL, the gene for BfSGBP-B is followed by a GH16_3 gene with a BACON domain (*Bacteroides*-associated carbohydrate-bind (putative) domain) preceding the GH16-laminarin-like domain (WP_034523369.1). In the *Maribacter* enzyme, the active hydrolase site is only preceded by the laminarin-binding domain. But the GH16 domain is interrupted after the β-sheet 5, following the nomenclature of *Zobellia* endo-laminarinase. LamA from *Zobellia galactivorans* has a small additional loop between the β-sheet 5 and the β-sheet 6, after AA 213 of LamA or AA 81 of LamC (Labourel *et al.*, 2015). The inserted domain of unknown function in the *Maribacter* enzymes was not assigned by

NCBI/CDD or InterPro to a described domain. It was with 29% identity over 290 AA related to the aforementioned laminarin-binding surface glucan binding protein B of *Bacteroides fluxus* BfSGBP-B. The GH16-laminarin-like domain has the expected active site consisting of 67 amino acids and the WPA motif (Viborg *et al.*, 2019). These homologies suggest that the large extracellular GH16_3 enzymes of *Maribacter* and other *Flavobacteriia* have two laminarin-binding domains besides the catalytic site with an endo-β-(1,3)-laminarinase activity. A model of the GH16_3 of *Maribacter* sp. MAR_2009_72 published by the European Bioinformatics Institute confirmed the domains (https://alphafold.ebi.ac.uk/entry/A0A559NTH3).

Recent studies have enlarged the number of enzyme activities associated with the GH16-laminarinase-like domain, but also provided more insight into the substrate specificity of GH16_3 enzymes (Matard-Mann *et al.*, 2017; Crouch *et al.*, 2020). Determinants are the amino acids in the six fingers. A finger is a succession of β-strand + tight turn or loop or helix + β-strand extending from the protein "bowl" (Matard-Mann *et al.*, 2017). A detailed comparison of GH16 glucanases, laminarinases, porphyranases, carrageenases, xyloglucanases and other activities was the basis for a deeper bioinformatics analysis of *Maribacter* GH16_3 laminarinases (Crouch *et al.*, 2020). Galactosidases have in finger one a conserved motif (WKLCTYN**N**A**W**SQ) with Asp and Trp blocking an equatorial C4 hydroxyl group of glucose (Crouch *et al.*, 2020). This feature was not conserved in *Maribacter* GH16_3 laminarinases indicating sufficient space for glucose is available. This difference is as well reflected in the phylogenetic GH16_3 analysis.

A feature conserved in β-(1,3)-glucanases, but not in GH16 enzymes with other activities, are two Trp´s before and after finger three. *Maribacter* GH16_3 laminarinases have these Trp´s within the motif GGTWPAL**W**ALGANFDEVG**W**P. Together with the for endo-laminarinases typical oligosaccharide degradation pattern – in contrast to strain MAR_2009_71, which has no GH16_3, uses an unidentified degradation system and may be considered as negative control for the presence of endo-laminarinases – and the presence of laminarin-binding domains in the GH16_3 proteins, we conclude that the identified GH16_3 proteins are endo-laminarinases. Definitely, ultimative proof will be a future biochemical study on the importance of two laminarin-binding domains in the action of this laminarinase.

Uptake of laminarin oligo- and polysaccharides into the periplasm requires an outer membrane transport system. SusC/D transport systems for laminarin oligosaccharides were originally identified as expressed proteins in laminarin grown biomass, a direct biochemical assay does not exist. All five *Maribacter* strains expressed at least one SusC protein that has a specific similarity to the TIGR04056 motif for sugar transporting SusC proteins. The motif includes the N-terminal extension of TonB-dependent transporter (PFAM13715) characteristic for SusC (Pollet *et al.*, 2021). Our phylogenetic analyses affiliated the expressed SusC proteins with laminarin specific SusCs of other *Flavobacteriia*. The gene *susC* of *Maribacter forsetii* (P177_RS4595) did not encode an N-terminal signal peptide domain. This detail was also observed in the genomes of *Formosa* strain A and B and *Christiangramia forsetii* (FORMA_RS04655, FORMB_RS04850 and GFO_RS08120).

For oligosaccharide hydrolysis in the periplasm, β-exo-glucanases of the GH3 family were expressed. The *Maribacter forsetii* gene P177_RS07335 is predicted to have the periplasmic β-D-glucosidase domain (COG1472) and a C-terminal uncharacterized 140 AA region. A constraint alignment with the larger C-terminal domain of barley GH3 isozyme ExoI indicated conserved amino acids, suggesting a glucan binding for the domain (Varghese *et al.*, 1999). Transport of glucose into the cytosol is presumably performed via an ABC transport system. Two complete systems consisting of substrate binding, permease and ATP binding protein were expressed in *Maribacter forsetii* when grown on laminarin. The pathway is expected to import glucose. The imported glucose can be further hydrolysed in the glycolysis and then replenish the citric acid cycle.

Comparison of the laminarin utilization of *Maribacter forsetii* with the other four *Maribacter* strains of this study brought similarities, but also differences to light. *Maribacter* sp. MAR_2009_72 uses the same degradation pathway for laminarin as *Maribacter forsetii*. All key enzymes were expressed in its proteome: GH16_3, SusC/D, ExbBD, GH3 and ABC transport system. The presence of GH16_3 in the genomes of MAR_2009_60 and Hel_I_7 and the endo-laminarinase activity of the cellular extracts suggested a similar pathway; however, our proteomic analysis did not detect the corresponding GH16_3 proteins. For *Maribacter dokdonensis* MAR_2009_71 a GH16_3 was not annotated in the genome and we observed the

lack of endo-laminarinase activity. These findings suggested that growth of MAR_2009_71 on laminarin is based on a so-far-uncharacterized set of enzymes.

Free-living and particle-associated bacteria differ in their genome size and numbers of genes for polysaccharides degradation. For example, the genus *Winogradskyella* includes bacteria of both lifestyles and the genome size correlates with the lifestyle, being that particle-associated bacteria have larger genomes (Alejandre-Colomo *et al.*, 2021). These comprise more degradative functions and motility, either swimming or gliding. Within the particle-associated bacteria, many particle-attached bacteria have the capacity to glide on surfaces (Seymour *et al.*, 2017). All five *Maribacter* strains have the machinery for gliding annotated in their genome. We found most gliding associated enzymes expressed in the proteomes and gliding was observed on plates. The investigated *Maribacter* strains have a range of PULs, including ones for alginate and alpha-glucans (Kappelmann *et al.*, 2019). The separation into free-living and particle-associated bacteria does not coincide with the genus border. *Winogradskyella* and *Polaribacter* include strains with both lifestyles and PUL for laminarin (Xing *et al.*, 2015; Avcı *et al.*, 2020; Alejandre-Colomo *et al.*, 2021). *Christiangramia* (ex *Gramella*) is particle-associated and has a PUL for laminarin (Kabisch *et al.*, 2014).

This leaves the question open why has a laminarin utilization loci not evolved in *Maribacter*? Settlement experiments showed a diatom-associated lifestyle of *Maribacter* (Heins *et al.*, 2021b). We argue that laminarin is rarely available for *Maribacter* in relation to algal surface polysaccharides and algal exudates. Laminarin is an intracellular carbon storage for algae and concealed for bacteria. Only upon algal lysis, an ephemeral event, laminarin becomes available, but diffuses rapidly away, with a diffusion coefficient of one third of that of glucose (Elyakova *et al.*, 1994). Thus, laminarin appears to have little effect on how the genomes of particle-associated *Maribacter* strains continues to evolve. We interpret this absence of a laminarin PUL as evidence for a specialization on less soluble algal polysaccharides and algal exudates.

Bacterioplankton ecotypes are often defined by experimental methods, i.e., the separation into free-living and particle-associated bacteria by sequential filtration. Alternative separation method, i.e., settlement, gave access to free-living motile bacteria of the phycosphere and separate large free-living bacteria from bacteria

attached to settling particles. Living alga with a constant production of exudates differ from marine snow particle that represent a single-fed batch substrate. According to our substrate flux hypothesis, live on laminarin-rich dead organic matter may support the evolution of laminarin PULs in particle-associated bacteria specialized on dead particles. Homologues of the laminarinase of *M. forsetii* are present in laminarin PULs together with GH17, GH2 and other CAZymes in *Croceivirga radicis*, isolated from a rotten mangrove root, *Croceitalea dokdonensis*, isolated from the rhizosphere of the brown alga *Ecklonia kurome*, and two *Croceitalea* strains from sea surfaces. The presence of the large laminarinase gene outside and within PULs reflects the huge diversity of niches for microbial live-in nature.

2.6 Acknowledgement

2.7 Supplement

2.7.1 Supplement Methods

Parameter for MaxQuant analysis as provided by the software

Parameter   Value
Version   2.2.0.0
Include contaminants   True
PSM FDR   0.01
PSM FDR Crosslink   0.01

Protein FDR 0.01
Site FDR   0.01
Use Normalized Ratios for Occupancy   True
Min. peptide Length 7
Min. score for unmodified peptides   0
Min. score for modified peptides 40
Min. delta score for unmodified peptides   0
Min. delta score for modified peptides   6
Min. unique peptides   0
Min. razor peptides 1
Min. peptides   1
Use only unmodified peptides and   True
Modifications included in protein quantification Oxidation (M);Acetyl (Protein N-term)
Peptides used for protein quantification Razor
Discard unmodified counterpart peptides   True
Label min. ratio count   2
Use delta score   False
iBAQ True
iBAQ log fit   True
Match between runs   False
Find dependent peptides   False
Decoy mode revert
Include contaminants   True
Advanced ratios   True
Fixed andromeda index folder
Combined folder location
Second peptides   True
Stabilize large LFQ ratios True
Separate LFQ in parameter groups   False
Require MS/MS for LFQ comparisons   True
Calculate peak properties False
Main search max. combinations 200
Advanced site intensities True
Write msScans table   False
Write msmsScans table   True
Write ms3Scans table   True
Write allPeptides table   True
Write mzRange table   True
Write DIA fragments table False
Write DIA fragments quant table False
Write pasefMsmsScans table   True
Write accumulatedMsmsScans table   True
Max. peptide mass [Da]   4600
Min. peptide length for unspecific search   8
Max. peptide length for unspecific search   25
Razor protein FDR True
Disable MD5 False
Max mods in site table   3

Match unidentified features   False
Epsilon score for mutations
Evaluate variant peptides separately   True
Variation mode   None
MS/MS tol. (FTMS) 20 ppm
Top MS/MS peaks per Da interval. (FTMS)   12
Da interval. (FTMS) 100
MS/MS deisotoping (FTMS)   True
MS/MS deisotoping tolerance (FTMS)   7
MS/MS deisotoping tolerance unit (FTMS)   ppm
MS/MS higher charges (FTMS)   True
MS/MS water loss (FTMS) True
MS/MS water loss (FTMS for cross link) False
MS/MS ammonia loss (FTMS)   True
MS/MS ammonia loss (FTMS for cross link)   False
MS/MS dependent losses (FTMS)   True
MS/MS recalibration (FTMS)   False
MS/MS tol. (ITMS) 0.5 Da
Top MS/MS peaks per Da interval. (ITMS)   8
Da interval. (ITMS) 100
MS/MS deisotoping (ITMS)   False
MS/MS deisotoping tolerance (ITMS)   0.15
MS/MS deisotoping tolerance unit (ITMS)   Da
MS/MS higher charges (ITMS)   True
MS/MS water loss (ITMS) True
MS/MS water loss (ITMS for cross link) False
MS/MS ammonia loss (ITMS)   True
MS/MS ammonia loss (ITMS for cross link)   False
MS/MS dependent losses (ITMS) True
MS/MS recalibration (ITMS)   False
MS/MS tol. (TOF)   25 ppm
Top MS/MS peaks per Da interval. (TOF)   16
Da interval. (TOF)   100
MS/MS deisotoping (TOF) True
MS/MS deisotoping tolerance (TOF)   0.01
MS/MS deisotoping tolerance unit (TOF)   Da
MS/MS higher charges (TOF)   True
MS/MS water loss (TOF)   True
MS/MS water loss (TOF for cross link)   False
MS/MS ammonia loss (TOF)   True
MS/MS ammonia loss (TOF for cross link)   False
MS/MS dependent losses (TOF) True
MS/MS recalibration (TOF)   False
MS/MS tol. (Unknown)   20 ppm
Top MS/MS peaks per Da interval. (Unknown) 12
Da interval. (Unknown)   100
MS/MS deisotoping (Unknown)   True
MS/MS deisotoping tolerance (Unknown)   7
MS/MS deisotoping tolerance unit (Unknown) ppm

MS/MS higher charges (Unknown)   True
MS/MS water loss (Unknown)   True
MS/MS water loss (Unknown for cross link)   False
MS/MS ammonia loss (Unknown)   True
MS/MS ammonia loss (Unknown for cross link)   False
MS/MS dependent losses (Unknown)   True
MS/MS recalibration (Unknown) False
Site tables   Oxidation (M)Sites.txt

Parameters in Perseus 2.0.7.0

Volcanoplot
t-test
Side: both
Number of randomizations: 250
Preserver grouping in randomizations: None
FDR: 0.05
s0: 0.15

Principle component analysis
Number of components: 5
Cutoff method: Benjamini-Hochberg
Benjamini-Hochberg FDR: 0.05
Relative enrichment: None

Parameters for the phylogenetic trees

Mafft alignment parameters in automode:
Scoring matrix for amino acid sequences: BLOSUM62
Gap opening penaltiy:1.53
Strategy: Auto (FFT-NS-1,FFT-NS-2, FFT-NS-i or L-INS-I; depends on data size)

Mega 11:
The evolutionary history was inferred by using the Maximum Likelihood method and Le_Gascuel_2008 model. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the JTT model, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (4 categories). The trees are drawn to scale, with branch lengths measured in the number of substitutions per site.

2.7.2 Supplement Tables

Supplement Tables are available at Edmond: https://doi.org/10.17617/3.HCEGVG

2.7.3 Supplement Figures
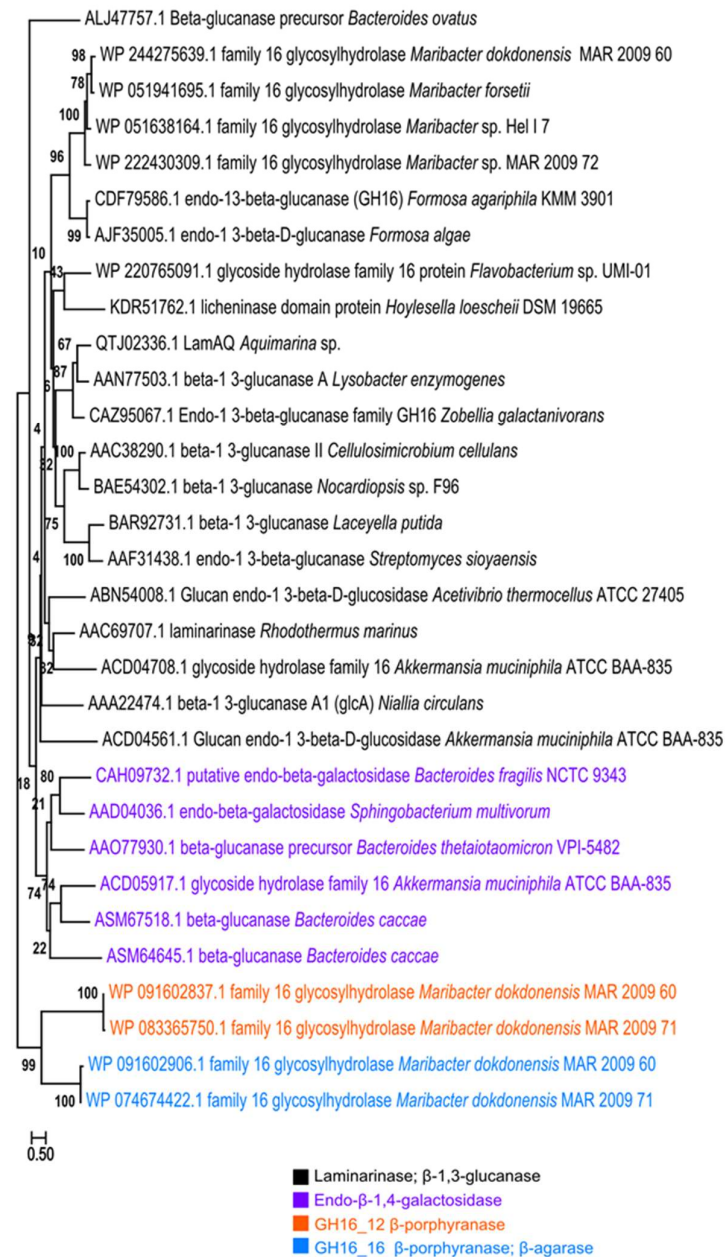
Supplement Figure 1: Phylogenetic tree shows the likeness between enzymes of the GH16_3 family, characterized in the CAZy database and the ones from the five *Maribacter* strains in this study. The evolutionary history was inferred by using the Maximum Likelihood method and Le_Gascuel_2008 model. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the JTT model, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (4 categories (+G, parameter = 3.7828)). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site.

2.8 References

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* **59**: 108-117.

Alejandre-Colomo, C., Harder, J., Fuchs, B.M., Rosselló-Móra, R., and Amann, R. (2020) High-throughput cultivation of heterotrophic bacteria during a spring phytoplankton bloom in the North Sea. *Systematic and Applied Microbiology* **43**: 126066.

Alejandre-Colomo, C., Francis, B., Viver, T., Harder, J., Fuchs, B.M., Rossello-Mora, R., and Amann, R. (2021) Cultivable *Winogradskyella* species are genomically distinct from the sympatric abundant candidate species. *ISME Communications* **1**: 51.

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* **71**: 3340-3349.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology* **215**: 403-410.

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* **14**: 1369-1383.

Barbeyron, T., Carpentier, F., Haridon, S., Schüler, M., Michel, G., and Amann, R. (2008) Description of *Maribacter forsetii* sp. nov., a marine *Flavobacteriaceae* isolated from North Sea water, and emended description of the genus *Maribacter*. *International Journal of Systematic and Evolutionary Microbiology* **58**: 790-797.

Bäumgen, M., Dutschei, T., and Bornscheuer, U.T. (2021) Marine polysaccharides: Occurrence, enzymatic degradation and utilization. *ChemBioChem* **22**: 2247-2256.

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* **83**: e03389-16.

Becker, S., Tebben, J., Coffinet, S., Wiltshire, K., Iversen, M.H., Harder, T. *et al.* (2020) Laminarin is a major molecule in the marine carbon cycle. *Proceedings of the National Academy of Sciences* **117**: 6599-6607.

Bennke, C.M., Reintjes, G., Schattenhofer, M., Ellrott, A., Wulf, J., Zeder, M., and Fuchs, B.M. (2016) Modification of a high-throughput automatic microbial cell enumeration system for shipboard analyses. *Applied and Environmental Microbiology* **82**: 3289-3296.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**: 248-254.

Bycroft, M., Bateman, A., Clarke, J., Hamill, S.J., Sandford, R., Thomas, R.L., and Chothia, C. (1999) The structure of a PKD domain from polycystin-1: Implications for polycystic kidney disease. *The EMBO Journal* **18**: 297-305.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The carbohydrate-active enzymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Research* **37**: 233-238.

Cheng, P.-N., Pham, J.D., and Nowick, J.S. (2013) The supramolecular chemistry of β-sheets. *Journal of the American Chemical Society* **135**: 5477-5492.

Crouch, L.I., Liberato, M.V., Urbanowicz, P.A., Baslé, A., Lamb, C.A., Stewart, C.J. *et al.* (2020) Prominent members of the human gut microbiota express endo-acting O-glycanases to initiate mucin breakdown. *Nature Communications* **11**: 4017.

Drula, E., Garron, M.-L., Dogan, S., Lombard, V., Henrissat, B., and Terrapon, N. (2022) The carbohydrate-active enzyme database: Functions and literature. *Nucleic Acids Research* **50**: 571-577.

DuBois, M., Gilles, K.A., Hamilton, J.K., Rebers, P.A., and Smith, F. (1956) Colorimetric method for determination of sugars and related substances. *Analytical Chemistry* **28**: 350-356.

Elyakova, L.A., Pavlov, G.M., Isakov, V.V., Zaitseva, I.I., and Stepenchekova, T.A. (1994) Molecular characteristics of subfractions of laminarin. *Chemistry of Natural Compounds* **30**: 273-274.

Fischer, T., Schorb, M., Reintjes, G., Kolovou, A., Santarella-Mellwig, R., Markert, S. *et al.* (2019) Biopearling of interconnected outer membrane vesicle chains by a marine flavobacterium. *Applied and Environmental Microbiology* **85**: e00829-19.

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* **15**: 2336-2350.

Grant, J.R., Enns, E., Marinier, E., Mandal, A., Herman, E.K., Chen, C.-y. *et al.* (2023) Proksee: In-depth characterization and visualization of bacterial genomes. *Nucleic Acids Research* **51**: W484-W492.

Grondin, J.M., Tamura, K., Déjean, G., Abbott, D.W., and Brumer, H. (2017) Polysaccharide utilization loci: Fueling microbial communities. *Journal of Bacteriology* **199**: e00860-16.

Hahnke, R.L., and Harder, J. (2013) Phylogenetic diversity of *Flavobacteria* isolated from the North Sea on solid media. *Systematic and Applied Microbiology* **36**: 497-504.

Hahnke, R.L., Probian, C., Fuchs, B.M., and Harder, J. (2013) Variations in pelagic bacterial communities in the North Atlantic Ocean coincide with water bodies. *Aquatic Microbial Ecology* **71**: 131-140.

Hahnke, R.L., Bennke, C.M., Fuchs, B.M., Mann, A.J., Rhiel, E., Teeling, H. *et al.* (2015) Dilution cultivation of marine heterotrophic bacteria abundant after a spring

phytoplankton bloom in the North Sea. *Environmental Microbiology* **17**: 3515-3526.

Hallgren, J., Tsirigos, K.D., Pedersen, M.D., Almagro Armenteros, J.J., Marcatili, P., Nielsen, H. *et al.* (2022) Preprint DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. *bioRxiv* 2022.04.08.487609.

Heins, A., and Harder, J. (2023) Particle-associated bacteria in seawater dominate the colony-forming microbiome on ZoBell marine agar. *FEMS Microbiology Ecology* **99**: fiac151.

Heins, A., Amann, R.I., and Harder, J. (2021a) Cultivation of particle-associated heterotrophic bacteria during a spring phytoplankton bloom in the North Sea. *Systematic and Applied Microbiology* **44**: 126232.

Heins, A., Reintjes, G., Amann, R.I., and Harder, J. (2021b) Particle collection in Imhoff sedimentation cones enriches both motile chemotactic and particle-attached bacteria. *Frontiers in Microbiology* **12**: 643730.

Kabisch, A., Otto, A., König, S., Becher, D., Albrecht, D., Schueler, M. *et al.* (2014) Functional characterization of polysaccharide utilization loci in the marine *Bacteroidetes* 'Gramella forsetii' KT0803. *The ISME Journal* **8**: 1492–1502.

Kadam, S.U., Tiwari, B.K., and O'Donnell, C.P. (2015) Extraction, structure and biofunctional activities of laminarin from brown algae. *International Journal of Food Science & Technology* **50**: 24-31.

Kanehisa, M., Sato, Y., and Morishima, K. (2016) BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *Journal of Molecular Biology* **428**: 726-731.

Kappelmann, L., Krüger, K., Hehemann, J.-H., Harder, J., Markert, S., Unfried, F. *et al.* (2019) Polysaccharide utilization loci of North Sea *Flavobacteriia* as basis for using SusC/D-protein expression for predicting major phytoplankton glycans. *The ISME Journal* **13**: 76-91.

Katoh, K., Rozewicki, J., and Yamada, K.D. (2019) MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics* **20**: 1160-1166.

Krüger, K., Chafee, M., Ben Francis, T., Glavina del Rio, T., Becher, D., Schweder, T. *et al.* (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* **13**: 2800-2816.

Labourel, A., Jam, M., Jeudy, A., Hehemann, J.-H., Czjzek, M., and Michel, G. (2014) The β-Glucanase ZgLamA from *Zobellia galactanivorans* evolved a bent active site adapted for efficient degradation of algal laminarin. *Journal of Biological Chemistry* **289**: 2027-2042.

Labourel, A., Jam, M., Legentil, L., Sylla, B., Hehemann, J.-H., Ferrières, V. *et al.* (2015) Structural and biochemical characterization of the laminarinase ZgLamCGH16 from *Zobellia galactanivorans* suggests preferred recognition of branched laminarin. *Acta crystallographica Section D : Structural biology* **71**: 173-184.

Lu, D.-C., Wang, F.-Q., Amann, R.I., Teeling, H., and Du, Z.-J. (2023) Epiphytic common core bacteria in the microbiomes of co-located green (*Ulva*), brown (*Saccharina*) and red (*Grateloupia*, *Gelidium*) macroalgae. *Microbiome* **11**: 126.

Lu, S., Wang, J., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R. *et al.* (2019) CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Research* **48**: 265-268.

Matard-Mann, M., Bernard, T., Leroux, C., Barbeyron, T., Larocque, R., Préchoux, A. *et al.* (2017) Structural insights into marine carbohydrate degradation by family GH16 κ-carrageenases. *Journal of Biological Chemistry* **292**: 19919-19934.

Miksch, S., Meiners, M., Meyerdierks, A., Probandt, D., Wegener, G., Titschack, J. *et al.* (2021) Bacterial communities in temperate and polar coastal sands are seasonally stable. *ISME Communications* **1**: 29.

Nedashkovskaya, O.I., Kim, S.B., Han, S.K., Lysenko, A.M., Rohde, M., Rhee, M.-S. *et al.* (2004) *Maribacter* gen. nov., a new member of the family *Flavobacteriaceae*, isolated from marine habitats, containing the species *Maribacter sedimenticola* sp. nov., *Maribacter aquivivus* sp. nov., *Maribacter orientalis* sp. nov. and *Maribacter ulvicola* sp. nov. *International Journal of Systematic and Evolutionary Microbiology* **54**: 1017-1023.

Nelson, T.E., and Lewis, B.A. (1974) Separation and characterization of the soluble and insoluble components of insoluble laminaran. *Carbohydrate Research* **33**: 63-74.

Noinaj, N., Guillier, M., Barnard, T.J., and Buchanan, S.K. (2010) TonB-dependent transporters: Regulation, structure, and function. *Annual Review of Microbiology* **64**: 43-60.

Orwick-Rydmark, M., Arnold, T., and Linke, D. (2016) The use of detergents to purify membrane proteins. *Current Protocols in Protein Science* **84**: 4.8.1-4.8.35.

Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B.L., Salazar, Gustavo A. *et al.* (2022) InterPro in 2022. *Nucleic Acids Research* **51**: 418-427.

Perez-Riverol, Y., Bai, J., Bandla, C., García-Seisdedos, D., Hewapathirana, S., Kamatchinathan, S. *et al.* (2022) The PRIDE database resources in 2022: A hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Research* **50**: D543-D552.

Pollet, R.M., Martin, L.M., and Koropatkin, N.M. (2021) TonB-dependent transporters in the *Bacteroidetes*: Unique domain structures and potential functions. *Molecular Microbiology* **115**: 490-501.

Probandt, D., Eickhorst, T., Ellrott, A., Amann, R., and Knittel, K. (2018) Microbial life on a sand grain: From bulk sediment to single grains. *The ISME Journal* **12**: 623-633.

R Core Team (2023) R: A language and environment for statistical computing. R Foundation for Statistical Computing. In. Vienna, Austria.

Reintjes, G., Arnosti, C., Fuchs, B.M., and Amann, R. (2017) An alternative polysaccharide uptake mechanism of marine bacteria. *The ISME Journal* **11**: 1640-1650.

Rioux, L.E., Turgeon, S.L., and Beaulieu, M. (2007) Characterization of polysaccharides extracted from brown seaweeds. *Carbohydrate Polymers* **69**: 530-537.

Sayers, E.W., Bolton, E.E., Brister, J.R., Canese, K., Chan, J., Comeau, Donald C. *et al.* (2022) Database resources of the national center for biotechnology information. *Nucleic Acids Research* **50**: 20-26.

Schultz, D., Zühlke, D., Bernhardt, J., Francis, T.B., Albrecht, D., Hirschfeld, C. *et al.* (2020) An optimized metaproteomics protocol for a holistic taxonomic and functional characterization of microbial communities from marine particles. *Environmental Microbiology Reports* **12**: 367-376.

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* **2**: 17065.

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* **182**: 5365-5372.

Sidhu, C., Kirstein, I.V., Meunier, C.L., Rick, J., Fofonova, V., Wiltshire, K.H. *et al.* (2023) Dissolved storage glycans shaped the community composition of abundant bacterioplankton clades during a North Sea spring phytoplankton bloom. *Microbiome* **11**: 77.

Stam, M., Lelièvre, P., Hoebeke, M., Corre, E., Barbeyron, T., and Michel, G. (2022) SulfAtlas, the sulfatase database: State of the art and new developments. *Nucleic Acids Research* **51**: 647-653.

Tamura, K., Stecher, G., and Kumar, S. (2021a) MEGA11: Molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution* **38**: 3022-3027.

Tamura, K., Dejean, G., Van Petegem, F., and Brumer, H. (2021b) Distinct protein architectures mediate species-specific beta-glucan binding and metabolism in the human gut microbiota. *Journal of Biological Chemistry* **296**: 100415.

Teeling, H., Fuchs, B.M., Bennke, C.M., Krüger, K., Chafee, M., Kappelmann, L. *et al.* (2016) Recurring patterns in bacterioplankton dynamics during coastal spring algae blooms. *eLife* **5**: e11888.

Teeling, H., Fuchs, B.M., Becher, D., Klockow, C., Gardebrecht, A., Bennke, C.M. *et al.* (2012) Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* **336**: 608-611.

Terrapon, N., Lombard, V., Drula, É., Lapébie, P., Al-Masaudi, S., Gilbert, H.J., and Henrissat, B. (2018) PULDB: The expanded database of polysaccharide utilization loci. *Nucleic Acids Research* **46**: D677-D683.

Teufel, F., Almagro Armenteros, J.J., Johansen, A.R., Gíslason, M.H., Pihl, S.I., Tsirigos, K.D. *et al.* (2022) SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nature Biotechnology* **40**: 1023-1025.

The UniProt, C. (2023) UniProt: The universal protein knowledgebase in 2023. *Nucleic Acids Research* **51**: D523-D531.

Tusher, V.G., Tibshirani, R., and Chu, G. (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences*

 98: 5116-5121.

Tyanova, S., and Cox, J. (2018) Perseus: A bioinformatics platform for integrative analysis of proteomics data in cancer research. In *Cancer Systems Biology: Methods and Protocols*. von Stechow, L. (ed). New York, NY: Springer New York, pp. 133-148.

Tyanova, S., Temu, T., and Cox, J. (2016) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols* **11**: 2301-2319.

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* **12**: 2894–2906.

Varghese, J.N., Hrmova, M., and Fincher, G.B. (1999) Three-dimensional structure of a barley β-D-glucan exohydrolase, a family 3 glycosyl hydrolase. *Structure* **7**: 179-190.

Viborg, A.H., Terrapon, N., Lombard, V., Michel, G., Czjzek, M., Henrissat, B., and Brumer, H. (2019) A subfamily roadmap of the evolutionarily diverse glycoside hydrolase family 16 (GH16). *J Biol Chem* **294**: 15973-15986.

Wang, F.-Q., Bartosik, D., Sidhu, C., Siebers, R., Lu, D.-C., Trautwein-Schult, A. *et al.* (2024) Particle-attached bacteria act as gatekeepers in the decomposition of complex phytoplankton polysaccharides. *Microbiome* **12**: 32.

Wickham, H. (2016) *ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag.

Wilkins, D. (2023) gggenes: Draw gene arrow maps in 'ggplot2'. In. https://wilkox.org/gggenes/.

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* **9**: 1410-1422.

Zheng, J., Ge, Q., Yan, Y., Zhang, X., Huang, L., and Yin, Y. (2023) dbCAN3: Automated carbohydrate-active enzyme and substrate annotation. *Nucleic Acids Research* **51**: W115-W121.

Chapter 3

Proteomic insight into arabinogalactan utilization by particle-associated *Maribacter* sp. MAR_2009_72

Saskia Kalenborn[1], Daniela Zühlke[2], Katharina Riedel[2], Rudolf I. Amann[1], Jens Harder[1]

[1]Department of Molecular Ecology, Max Planck Institute for Marine Microbiology, Bremen, Germany; [2]Department for Microbial Physiology and Molecular Biology, University of Greifswald, Germany

Contribution of the candidate in % of the total work

Experimental concept and design – 60%

Experimental work/acquisition of experimental data – 90%

Preparation of figures and tables – 100%

Drafting of the manuscript – 80%

Chapter 3

3.1 Abstract

Arabinose and galactose are major, rapidly metabolized components of marine particulate and dissolved organic matter. In this study, we observed for the first time large microbiomes for the degradation of arabinogalactan and report a detailed investigation of arabinogalactan utilization by the flavobacterium *Maribacter* sp. MAR_2009_72. Cellular extracts hydrolysed arabinogalactan in vitro. Comparative proteomic analyses of cells grown on arabinogalactan, arabinose, galactose, and glucose revealed the expression of specific proteins in the presence of arabinogalactan, mainly glycoside hydrolases (GH). Extracellular glycan hydrolysis involved five alpha-L-arabinofuranosidases affiliating with glycoside hydrolase families 43 and 51, four unsaturated rhamnogalacturonylhydrolases (GH105) and a protein with a glycoside hydrolase family-like domain. We detected expression of three induced TonB-dependent SusC/D transporter systems, one SusC, and nine glycoside hydrolases with a predicted periplasmatic location. These are affiliated with the families GH3, GH10, GH29, GH31, GH67, GH78, and GH115. The genes are located outside of and within canonical polysaccharide utilization loci classified as specific for arabinogalactan, for galactose-containing glycans, and for arabinose-containing glycans. The breadth of enzymatic functions expressed in *Maribacter* sp. MAR_2009_72 as response to arabinogalactan from the terrestrial plant larch suggests that *Flavobacteriia* are main catalysts of the rapid turnover of arabinogalactans in the marine environment.

3.2 Introduction

Marine environments contain many different polysaccharides as dissolved organic matter (DOM) or in particulate organic matter (POM). These are a vital carbon source for microorganisms, released from algae as exudates or during lysis by zooplankton predation or viral infection. Monosaccharide analysis of planktonic biomass from the North Sea revealed already in 1982, a dominance of glucose followed by arabinose, galactose, and mannose (Ittekkot *et al.*, 1982; Urbani *et al.*, 2005; Alderkamp *et al.*, 2007; Scholz and Liebezeit, 2013; Huang *et al.*, 2021). These monomers are the building blocks of algal polysaccharides: the abundant

β-homoglycans laminarin, cellulose, and xylan are often complemented with species-specific glycans such as agar, alginate, carrageenan, fucoidan, mannan, pectin, porphyran, and ulvan. The degradation of these glycans has been studied intensively in marine systems, however, details for arabinogalactan are missing (Bäumgen *et al.*, 2021). Recently arabinogalactan was detected in the high molecular weight dissolved organic matter (HMWDOM) and POM fraction using monoclonal antibodies during the algal spring bloom in the North Sea (Vidal-Melgosa *et al.*, 2021). This coincides with the high arabinose and galactose content of *Phaeocystis* spp., a haptophyte blooming in the North Sea (Alderkamp *et al.*, 2007; Sato *et al.*, 2018). The antibody-based quantification also showed a decrease in arabinogalactan content towards the end of the spring bloom, suggesting a fast turnover of the compound— contrasting with the accumulation of fucose-containing sulfated polysaccharides (Vidal-Melgosa *et al.*, 2021). The major source of arabinose and galactose in algae are likely arabinogalactan proteins, which anchor polysaccharide cell walls in the outer membrane of plants and algae (Silva *et al.*, 2020; Leszczuk *et al.*, 2023). The model compound for arabinogalactan type II is arabinogalactan from larch wood. It contains D-galactose and L-arabinose in a 6:1 molar ratio as well as traces of rhamnose, fucose, mannose, xylose, and D-glucuronic acid (Fujita *et al.*, 2019; Villa-Rivera *et al.*, 2021; Leszczuk *et al.*, 2023). Type II arabinogalactans have a complex backbone structure consisting of β-1,3-linked galactan backbone with β-1,6-linked galactan side chains (Kelly, 1999; Wang and LaPointe, 2020). Type I has a β-1,4-linked galactan backbone, whereby C3 can be linked with L-arabinofuranose (Hinz *et al.*, 2005). Plant arabinogalactan is degraded by aerobic bacteria and fungi as well as by anaerobic fermenting bacteria in gut systems, including *Bifidobacterium* and *Bacteroidetes* (Shulami *et al.*, 2011; Ndeh *et al.*, 2017; Cartmell *et al.*, 2018; Luis *et al.*, 2018; Wang and LaPointe, 2020; Sasaki *et al.*, 2021). The latter phylum encompasses also aerobic *Flavobacteriia* that have been identified as specialists for polysaccharide degradation in marine systems (Sidhu *et al.*, 2023). For this first study on the degradation of arabinogalactan by marine microorganisms, we selected a flavobacterial strain with a published genome and a particle-associated lifestyle, *Maribacter* sp. MAR_2009_72 (Kappelmann *et al.*, 2019; Heins *et al.*, 2021a). Strains of the genus *Maribacter* are rarely isolated from sea water, but they are

more abundant in particle fractions (Nedashkovskaya *et al.*, 2004; Heins and Harder, 2023; Lu *et al.*, 2023; Sidhu *et al.*, 2023). Abundances of up to 4% were detected in the oxic surface layer of sandy sediments (Probandt *et al.*, 2018; Miksch *et al.*, 2021). Even higher abundances were observed in micro- and macroalgae phycosphere populations (Heins *et al.*, 2021b; Lu *et al.*, 2023). This makes *Maribacter* strains ideal candidates for studying the degradation of algal cell wall polysaccharides. The uptake and degradation of polysaccharides in Bacteroidetes is often encoded in polysaccharide utilization loci (PULs). The first PUL was described for *Bacteroides thetaiotaomicron* for starch utilization (Shipman *et al.*, 2000). Polysaccharide utilization starts with the extracellular hydrolysis of polysaccharides into oligosaccharides on the surface of the cell. The oligosaccharides are transported into the periplasm via the SusC/D transport system, which is energized by a proton gradient via an ExbB/D-TonB system in the cytoplasmic membrane and by a domain in the periplasm to open the β-barrel channel of SusC for the transport (Noinaj *et al.*, 2010). The hydrolysis of polysaccharides is achieved by glycoside hydrolases (GH), glycoside transferases, polysaccharide lyases, and carbohydrate esterases with a high specificity, sometimes assisted by carbohydrate binding modules. These five groups of proteins are classified as carbohydrate active enzymes (CAZymes) (Bäumgen *et al.*, 2021; Drula *et al.*, 2022). For the degradation of arabinogalactan from larch wood, PULs were so far characterized for gut bacteria including *Bifidobacterium longum* ssp. *longum* NCC2705, *Bacteroides caccae* ATCC 43185, and *Bacteroides thetaiotaomicron* (Ndeh *et al.*, 2017; Cartmell *et al.*, 2018; Luis *et al.*, 2018; Wang and LaPointe, 2020). Here, we analyzed *Maribacter* sp. MAR_2009_72 proteomes using cells grown on arabinogalactan, arabinose, galactose, and glucose. Those proteomes were compared to identify the proteins induced by arabinogalactan. This study expands a recent *in-silico* study that did not report on arabinogalactan specific PULs (Kappelmann *et al.*, 2019) and provides experimental observations for a better interpretation of marine metagenomes.

3.3 Material and Methods

3.3.1 Growth experiments

Maribacter sp. MAR_2009_72 (DSM 29384), originally isolated from a phytoplankton catch in the Wadden Sea near the island Sylt, Germany, was revived from glycerol stocks that had been preserved in the laboratory since the initial isolation (Hahnke and Harder, 2013). The strain was grown in the liquid medium HaHa_100V with 0.3 g/L of casamino acids as the sole carbon source (Hahnke *et al.*, 2015). This limited growth to an optical density (OD) at 600 nm below 0.2. Growth beyond an OD of 0.3 was achieved by adding 2 g/L of a carbohydrate source, here arabinose, galactose, glucose (Sigma Aldrich/Merck KGaA, Darmstadt, Germany), and larch arabinogalactan (The Dairy School, Auchincruive, Scotland). The supplier of arabinogalactan had specified the monosaccharide composition as 81% galactose, 14% arabinose, and 5% other, whereby the other fraction was not defined. For proteomics, three cultures of 50 ml were inoculated with 0.4% v/v of a pregrown culture in the same medium and incubated at room temperature at 110 rpm. A fourth culture per substrate was maintained to monitor bacterial growth by measuring OD at 600 nm beyond the harvest point. Cells were harvested at an OD of 0.25. Cells were pelleted by centrifugation in 50 ml tubes with 3080x g for 30 min at 4°C. Pellets were resuspended in 1 ml medium and centrifuged in 1.5 ml tubes at 15870x g for 15 min at 4°C. The wet biomass was weighed and stored at −20°C. For microbiome size determinations, colony-forming units (CFU) were determined with 4 g/L larch wood arabinogalactan as organic carbon source on marine plates (Hahnke and Harder, 2013), using 4 g/L glucose or ZoBell's 2216 marine agar plates as reference. Inoculation of serial diluted sea or sediment pore water was performed with a 96 pin-holder. Inoculations were at room temperature. Partial 16S rRNA gene sequences of strains were obtained by colony PCR and Sanger sequencing (Hahnke and Harder, 2013). Partial 16S rRNA gene sequences have been deposited at GenBank under the accession numbers PP600029 to PP600099.

3.3.2 Protein preparation and mass spectrometry

Proteins were extracted from cells using a bead-beating method following the protocol by Schultz *et al.* (2020). A pellet of wet weight ranging from 20 to 200 mg was disrupted using 0.25 ml glass beads in 500 µl of lysis buffer. The protein content was quantified using the Roti Nanoquant assay (Carl Roth, Karlsruhe, Germany). For protein purification on denaturing polyacrylamide gels (SDS-PAGE), 50 µg of protein was combined with 10 µl of 4x SDS buffer [composed of 20% glycerol, 100 mM Tris/HCl, 10% (w/v) SDS, 5% β-mercaptoethanol, 0.8% bromophenol blue, pH 6.8] and loaded onto Tris-glycine-extended precast 4%–20% gels (Bio-Rad, Neuried, Germany). Electrophoresis was conducted at 150 V for 8 min. Subsequently, the gel was fixed in a solution of 10% v/v acetic acid and 40% v/v ethanol for 30 min, stained with Brilliant Blue G250 Coomassie, and the desired protein band was excised. The proteins were extracted from the gel in one piece and then washed with a solution of 50 mM ammonium bicarbonate in 30% v/v acetonitrile. The gel pieces were dried using a SpeedVac (Eppendorf, Hamburg, Germany), and then rehydrated with 2 ng/µl trypsin (sequencing grade trypsin, Promega, USA). After a 15 min incubation at room temperature, excess liquid was removed, and the samples were digested overnight at 37°C. Following digestion, the gel pieces were covered with water suitable for mass spectrometry (MS), and peptides were eluted using ultrasonication. The peptides were subsequently desalted using Pierce™ C18 Spin Tips (Thermo Fisher, Schwerte, Germany) in accordance with the manufacturer's guidelines. The eluted peptides were dried using a SpeedVac and stored at −20°C. For MS analysis, the samples were thawed and reconstituted in 10 µl of Buffer A (99.9% acetonitrile + 0.1% acetic acid). Tryptic peptides of *Maribacter* sp. MAR_2009_72 were analyzed using an EASYnLC 1200 system coupled to a Q Exactive HF mass spectrometer (Thermo Fisher Scientific, located in Waltham, USA). Peptides were loaded onto a custom-packed analytical column containing 3 µm C18 particles (Dr. Maisch GmbH, Ammerbuch, Germany). The loading was performed using buffer A (0.1% acetic acid) at a flow rate of 2 µl/min. Peptide separation was achieved through an 85-min binary gradient, transitioning from 4% to 50% buffer B, composed of 0.1% acetic acid in acetonitrile, at a flow rate of 300 nl/min. Samples were measured in parallel mode; survey scans

in the Orbitrap were recorded with a resolution of 60 000 with a m/z range of 333 to 1650. The 15 most intense peaks per scan were selected for fragmentation. Precursor ions were dynamically excluded from fragmentation for 30 s. Single-charged ions as well as ions with unknown charge state were rejected. Internal lock mass calibration was applied (lock mass 445.12003 Da). The MS files were analyzed in MaxQuant version 2.2.0.0 in the standard settings against the strain specific protein database downloaded from NCBI: *Maribacter* sp. MAR_2009_72 NZ_VISB01000001.1 (27 September 2022) and common laboratory contaminants (Tyanova *et al.*, 2016; Sayers *et al.*, 2022). Statistical analysis was performed in Perseus version 2.0.7.0 (Tyanova and Cox, 2018). Proteins were recognized as being expressed when they had label free quantification intensities (LFQ) intensities in one out of three biological replicates.

3.3.3 Bioinformatic analyses

Protein annotation was refined using several databases. CAZymes were considered to be identified, if two out of three search algorithms in dbCAN3 were positive (Zheng *et al.*, 2023). The conserved domain database (CDD) (Lu *et al.*, 2019), the SulfAtlas web interface (Stam *et al.*, 2022),InterPro (Paysan-Lafosse *et al.*, 2022), PULDB (Terrapon *et al.*, 2018), deepTMHMM (Hallgren *et al.*, 2022), SignalP (Teufel *et al.*, 2022), Blastkoala (Kanehisa *et al.*, 2016), and UniProt (The UniProt, 2023) provided additional information.

The search pattern used in PULDB were the following for PUL 1: GH10+GH43_1+GH67; for PUL 7: GH43_18+GH78 +GH105+GH105+GH115; and for PUL 8: GH43_19+GH43_34 +GH51+GH105+GH105.

For the visualization of the data the following programs and packages were used: R version 4.3.2 (R Core Team, 2023), ggplot2 (Wickham, 2016), gggenes (Wilkins, 2023), and Proksee (Grant *et al.*, 2023). The MS proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol *et al.*, 2022) partner repository with the dataset identifier PXD049074 and 10.6019/PXD049074.

3.4 Results

3.4.1 Growth on arabinogalactan

Maribacter sp. MAR_2009_72 grew in presence of larch wood arabinogalactan to a maximum OD of 0.338 and at a maximum growth rate µ = 0.06 h−1 (Figure 1). When 2 g/l of galactose or arabinose were provided in the medium a maximum OD of 0.419 and 0.446 was measured with respective growth rates of 0.07 h−1 and 0.06 h−1. Glucose supported the largest biomass formation, with an OD of 0.526 and µ = 0.05 h−1. The arabinogalactan cultures required more time to enter the exponential growth phase than the cultures with monosaccharides as substrates. The physiological reaction was confirmed in vitro. Soluble cell extracts hydrolysed larch wood arabinogalactan to an abundant molecule with the size of a hexose (galactose) and a minor molecule with the size of a pentose (arabinose) in fluorophore-assisted carbohydrate electrophoresis gels (Supplement Figure S1).
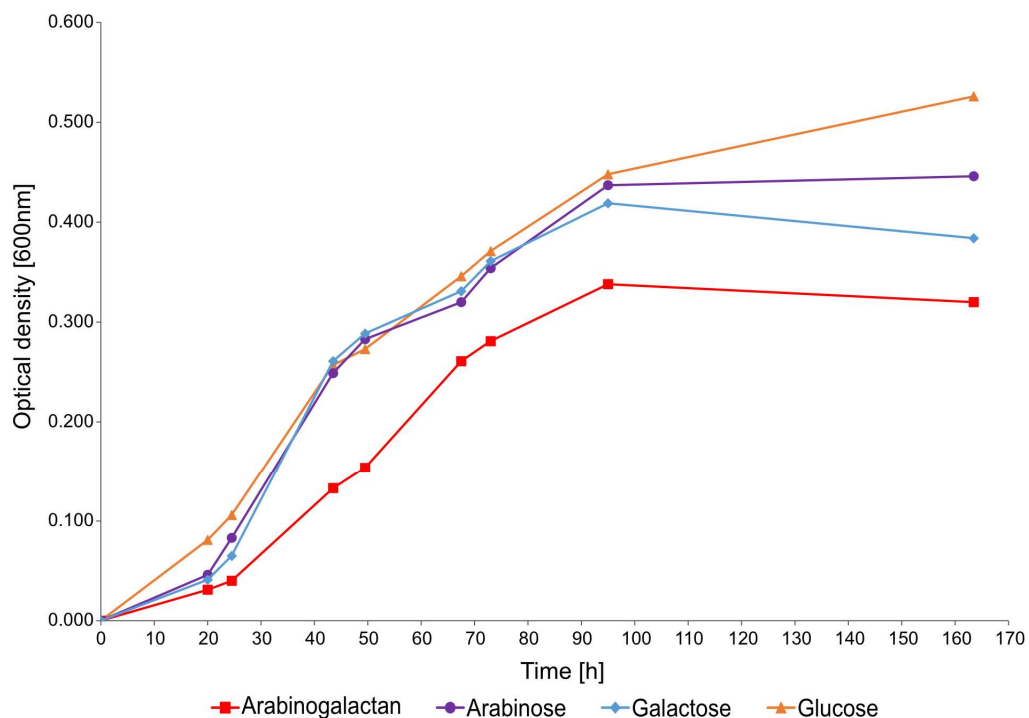
Figure 1: Growth curve of *Maribacter* sp. MAR_2009_72 in presence of four different carbon sources; arabinogalactan, arabinose, galactose, and glucose. MAR_2009_72 was grown in 50 ml of modified HaHa_100V with 2 g/L of the respective carbon source at room temperature at 110 rpm. The OD was measured at 600 nm.

3.4.2 Protein expression in *Maribacter* sp. MAR_2009_72

The comparative proteomic analysis was based on glucose as reference against arabinose, galactose, and arabinogalactan. We identified 1874 proteins in the arabinogalactan proteome (Figure 2A). Overall, these four conditions shared 1636 proteins. Only a small number of proteins were found to be unique to a particular growth condition. The glucose proteome had 36 unique proteins, the arabinose proteome 17 proteins, and the galactose proteome 19 proteins. Arabinogalactan had 52 unique proteins. We used the expression data, here label free quantification intensities (LFQ), to visualize the difference between the four conditions in a principal component analysis (PCA) (Figure 2B).

A

B
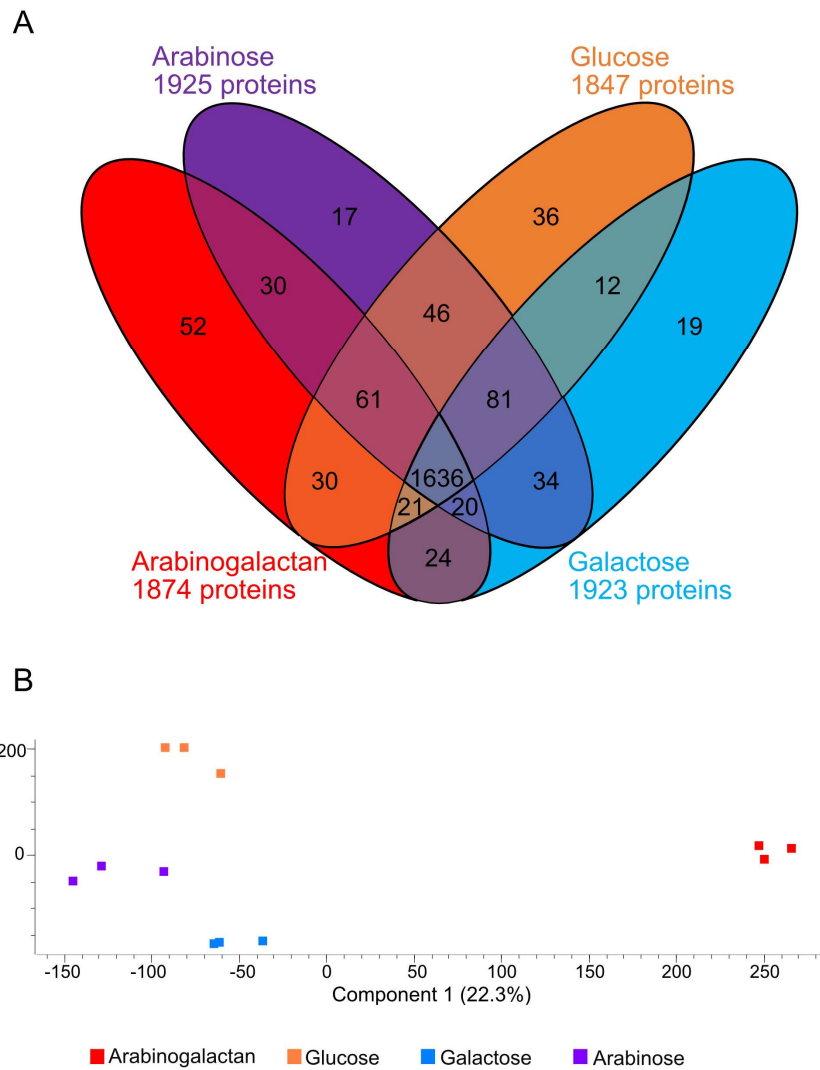
Figure 2: Comparison of the number of detected proteins in arabinogalactan, arabinose, galactose, and glucose. (A) Venn diagram showing the overlap of detected proteins in at least one of three biological replicates. (B) Principal component analysis shows the differences between the expression intensities of the four proteomes of MAR_2009_72.

The PCA plot indicated that the arabinogalactan proteome had the most contrasting expression pattern. The PCA analysis documented that the differences between monosaccharide proteomes were less pronounced than to the arabinogalactan proteome. *Maribacter* sp. MAR_2009_72 has a genome of 4.35 Mb encoding 3635 proteins (Figure 3).
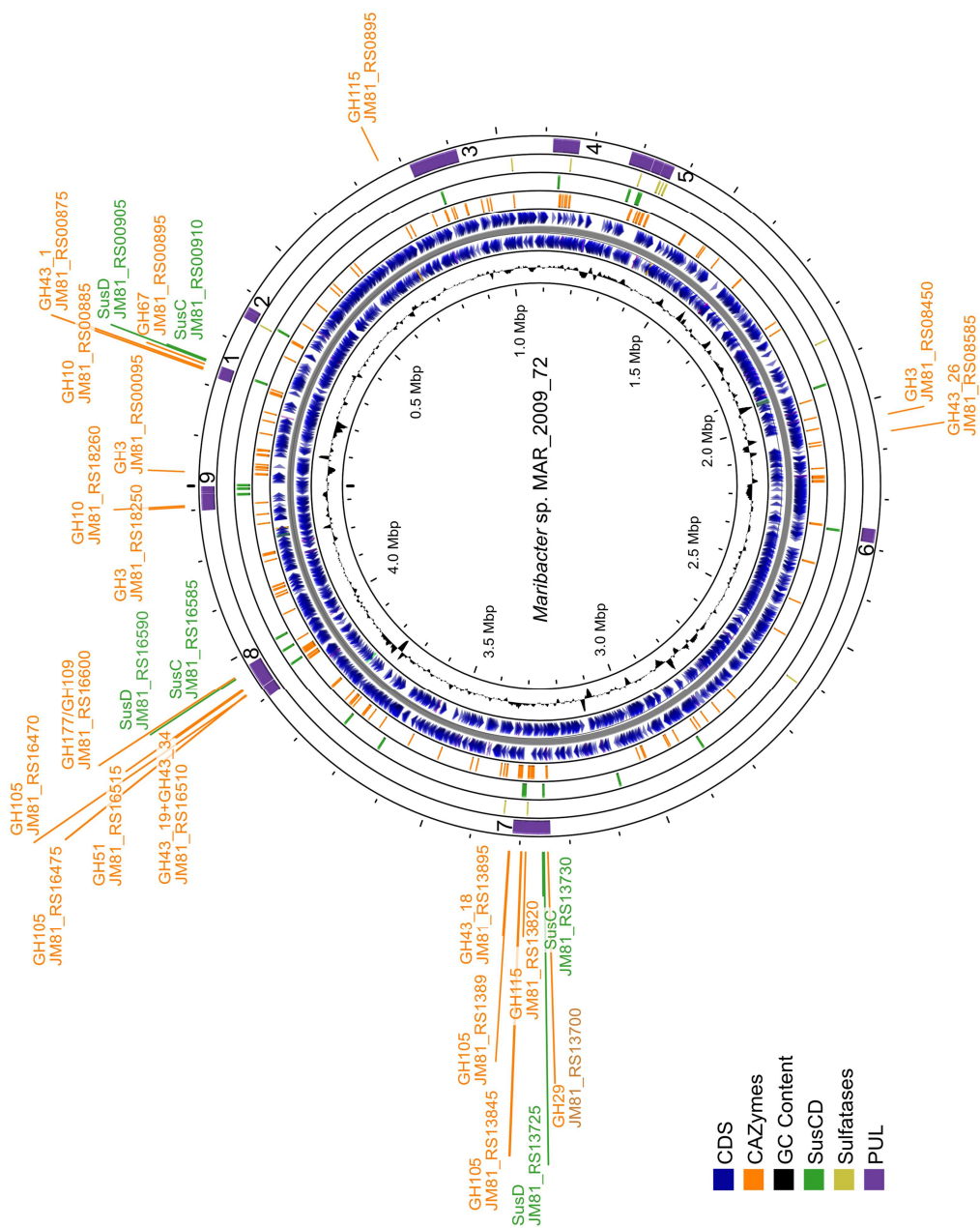
Figure 3: Full genome overview of *Maribacter* sp. MAR_2009_72 showcasing the GC content (ring one (most inner ring)), all annotated coding genes (CDS, ring two and three) in forward and reverse direction, CAZymes identified by dbCAN3 (ring four), SusC/D (ring five), sulfatases (ring six), and PULs (ring seven). Furthermore, we highlighted CAZymes and SusC/Ds that might be important for arabinogalactan utilization.

Nine PULs contain one or several SusC/D transporters and neighboring CAZymes. We labelled the PULs based on the arrangement in the genome, with PUL 1 being closest to the origin of replication (Supplement Table S1). The expression values revealed a proteomic response to arabinogalactan in PUL 1, 7, and 8 and outside

of PULs. PUL 1 encodes 13 proteins of which three out of four CAZymes and one SusC/D pair were expressed in arabinogalactan grown cells (Figure 4).

PUL 1

Figure 4: Gene organization and expression of PUL 1 of *Maribacter* sp. MAR_2009_72 grown in the presence of arabinogalactan, arabinose, galactose, and glucose. Expression intensities in the plot are the mean values of three biological replicates of each condition shown in LFQ values [log10].

The SusC/D pair (JM81_RS00910 and JM81_RS00905) was only expressed in the arabinogalactan proteome. Four other proteins were clearly induced by arabinogalactan, arabinose, and galactose. The α-L-arabinofuranosidase GH43_1 (JM81_RS00875) was 10-fold induced relative to the glucose proteome. A GH10,

an endo-β-1,4-xylanase, showed a similar expression pattern with a 5-fold difference to glucose. The third enzyme was a GH67, an α-glucuronidase, which had the strongest induction in arabinose and galactose proteomes. The fourth induced protein of the operon with an expression in the arabinogalactan proteome affiliated to the superfamily of protein or cofactor modifying RimK-type glutamate ligases with an ATP-grasp binding domain (JM81_RS00865). PUL 7 contains a single SusC/D pair and a tandem of SusC/D pair in one genetic region. It encodes 42 enzymes, 13 being classified as CAZymes, three SusC/D pairs, and one sulfatase (Figure 5). One SusC/D pair and 6 CAZymes were expressed in arabinogalactan grown cells. SusC (JM81_RS13730) and SusD (JM81_RS13725) were expressed in the galactose and arabinogalactan proteomes 100-fold and 10-fold stronger than in the arabinose proteome, respectively, and not in the glucose proteome. The tandem SusC/D pairs were not detected in any of the proteomes. An α-L-fucosidase of the GH29 family (JM81_RS13700) had the highest expression among the CAZymes in this PUL. The GH29 was expressed in similar intensities in all four conditions, suggesting a constitutive expression of this periplasmic enzyme. Less intense, but also expressed in all proteomes was a xylan-α-1,2-glucuronidase belonging to the GH115 family (JM81_RS13820), with the strongest expression on galactose. Two GH105 unsaturated rhamnogalacturonyl hydrolases (EC 3.2.1.172) (JM81_RS13845 and JM81_RS13890) were expressed in all four growth conditions, with the exception of JM81_RS13890, which was not detected in the arabinose proteome. A GH43_18 (JM81_RS13895) was expressed in all four proteomes with similar expression intensities. An α-L-rhamnosidase GH78 (JM81_RS13900) was expressed under all growth conditions. During our analysis a hypothetical protein (JM81_RS13825) with a six-hairpin GH like family domain sparked our interest. It was expressed in all four proteomes, with higher intensities in arabinogalactan, arabinose, and galactose proteomes.
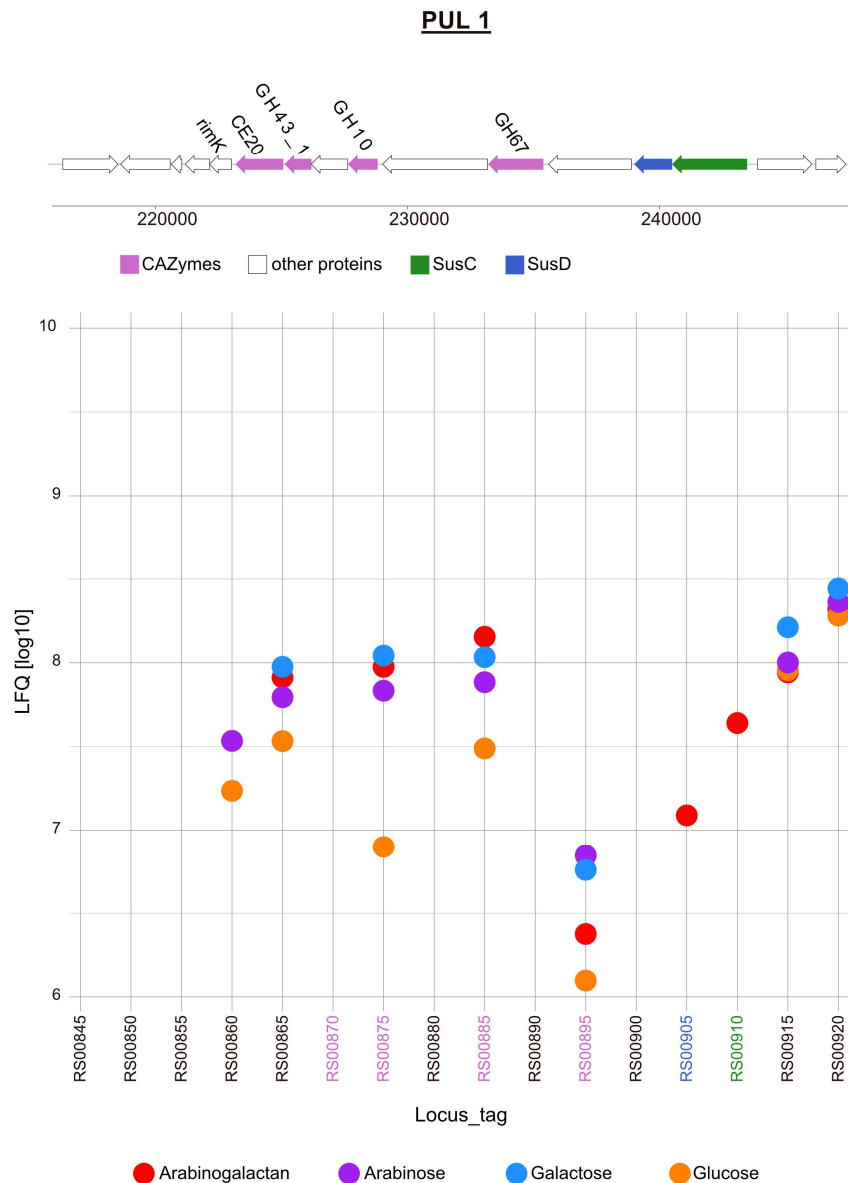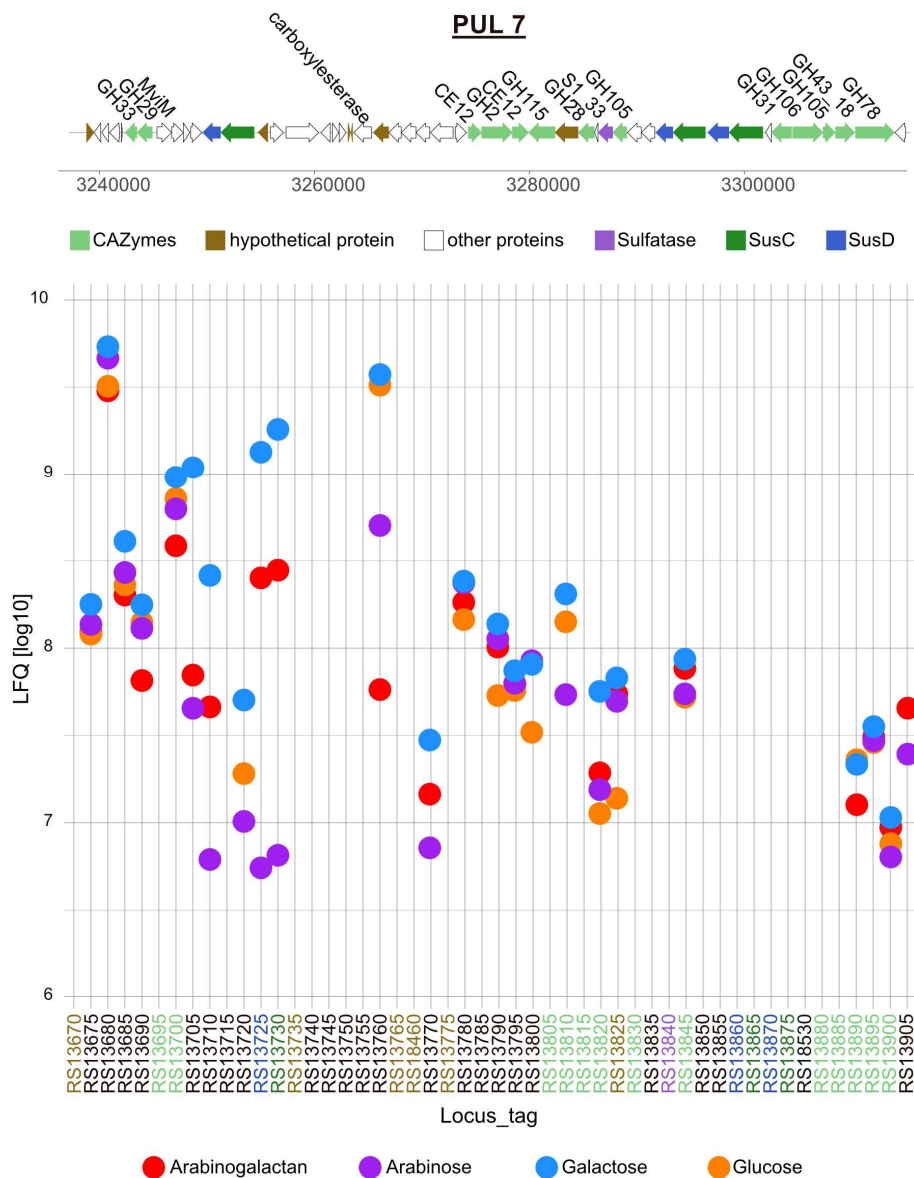
Figure 5: Gene organization and expression of PUL 7 of *Maribacter* sp. MAR_2009_72 grown in the presence of arabinogalactan, arabinose, galactose, and glucose. Expression intensities in the plot are the mean values of three biological replicates of each condition shown in LFQ values [log10].

PUL 8 encodes a total of 58 proteins, including 11 CAZymes and two SusC/D pairs (Figure 6). A total of five CAZymes, two SusCs but only one SusD were expressed in arabinogalactan grown cells. SusC (JM81_RS16585) and SusD (JM81_RS16590) were expressed in the arabinose and arabinogalactan proteome. Another SusC (JM81_RS16455) showed expression, slightly lower than the other SusC, in the arabinose proteome and slightly less for arabinogalactan. Two GH105 proteins (JM81_RS16470 and JM81_RS16475) annotated as unsaturated

rhamnogalacturonyl hydrolases were expressed similar in all proteomes. JM81_RS16510 includes two domains, GH43_19 and GH43_34. It was expressed in the arabinose, arabinogalactan, and galactose proteome, whereby the highest intensities were measured for arabinose. Another α-L-arabinofuranosidase, a GH51 (JM81_RS16515), was expressed in a similar pattern to the GH43_19+GH43_34 protein. These two genes are followed by genes of the arabinose metabolism to the pentose phosphate pathway - ribulokinase, L-ribulose-5- phosphate 4-epimerase, and L-arabinose isomerase - and a gene for a galactose mutarotase. All proteins in this operon were expressed in the arabinose, arabinogalactan, and galactose proteome, with highest intensities in arabinose proteomes. Unknown is the function of a GH109, a member of the Gfo/ldh/MocA superfamily of NAD(P) dependent oxidoreductases, that had the highest expression in the arabinogalactan proteome. The expression of a mannonate dehydratase (JM81_RS16615) hinted at a sugar acid metabolism. Interestingly, PUL 8 is preceded by an operon with sugar acid metabolizing enzymes. The following enzymes were induced in the arabinogalactan proteome in comparison to glucose: 5-dehydro-4-deoxy-d-glucuronate isomerase, gluconate-5-dehydrogenase, a sugar kinase, 2-dehydro-3-deoxyphosphogluconate aldolase, and tagaturonate reductase.
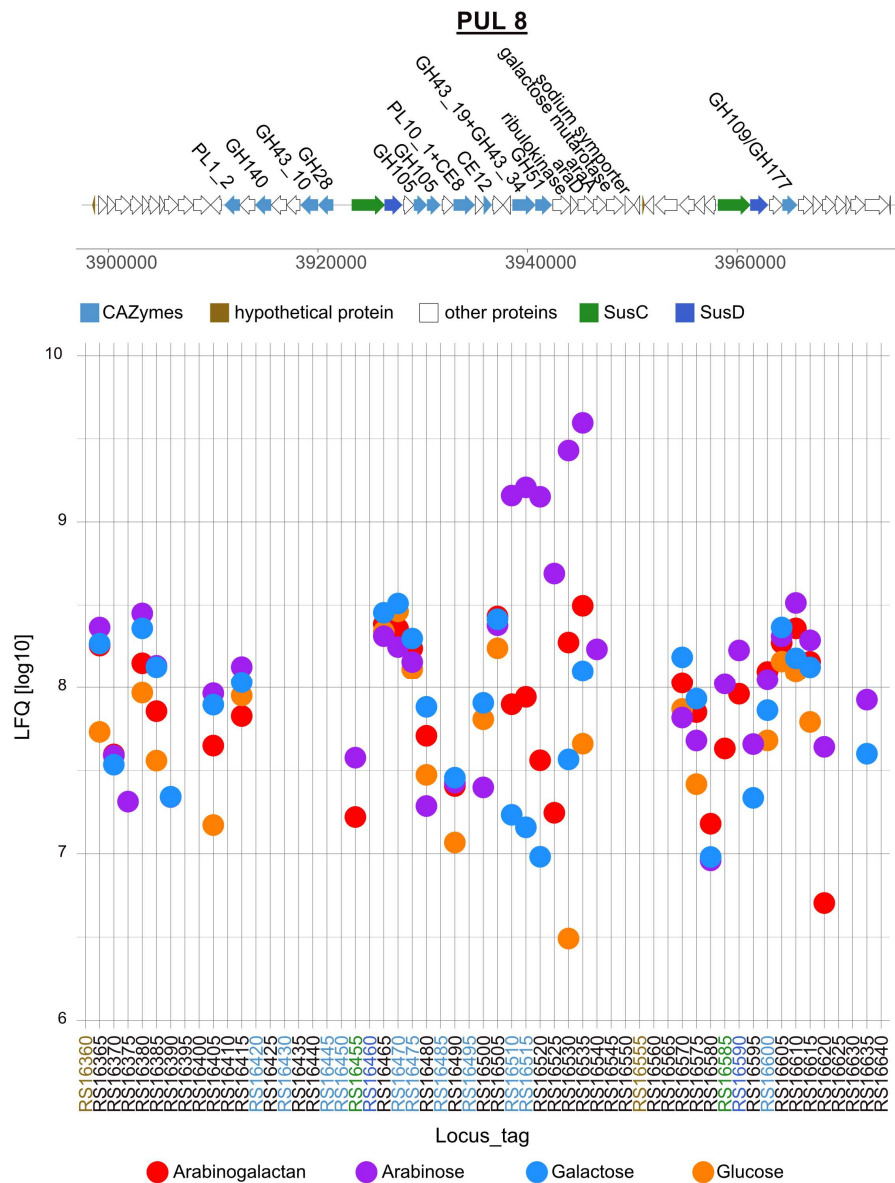
Figure 6: Gene organization and expression of PUL 8 of *Maribacter* sp. MAR_2009_72 grown in the presence of arabinogalactan, arabinose, galactose, and glucose. Expression intensities in the plot are the mean values of three biological replicates of each condition shown in LFQ values [log10].

An analysis with dbCAN3 identified 153 CAZymes in the genome, of which 106 were detected in the proteomes. Outside of the PULs 1, 7, and 8, several CAZymes were expressed in arabinogalactan degradation. Many expressed CAZymes had a signal peptide for export out of the cytosol (Supplement Figure S2 and Table S1). Three of the CAZymes were annotated as GH family 3 enzymes. JM81_RS00095 was expressed in all four conditions, the highest intensities were measured in the arabinogalactan proteome (Supplement Figure S2A). The second GH3

(JM81_RS08450) was expressed in all four conditions, but with a three to four times larger expression in arabinogalactan, arabinose, and galactose (Supplement Figure S2B). A third GH3 (JM81_RS18250) was as well expressed in all four conditions, but the highest intensities were measured for arabinose and galactose. It was part of an operon also including an endo1,4-β-xylanase (GH10) expressed only in arabinose and galactose grown cells (Supplement Figure S2C). All three GH3 were annotated as galactosidases. A GH43_26 (JM81_RS08585) was expressed in all four datasets, whereby the highest intensities were recorded for arabinose and nearly identical LFQs for glucose and arabinogalactan (Supplement Figure S2D). A GH115 xylan-α-1,2-glucuronidase (JM81_RS03245) was only expressed in arabinose and arabinogalactan grown cells (Supplement Figure S2E). The transport of the monosaccharides across the inner membrane may be facilitated by an ABC transport system consisting of ABC substrate-binding (JM81_RS03610), ABC permease (JM81_RS16840), and ABC ATP binding proteins (JM81_RS01625). Marine glycans are often decorated with sulfate. We identified 13 sulfatases in the genome of MAR_2009_72, of which three were expressed in arabinogalactan grown cells. JM81_RS05685, JM81_RS05692, and JM81_RS076760 were equally expressed in all four proteomes. All three were previously affiliated with the utilization of mucin, which contains to some extent galactose (Tailford *et al.*, 2015; Glover *et al.*, 2022). We also detected expressed proteins of the degradation pathway that where not classified as CAZymes. For the degradation of galactose, a galactose mutarotase (JM81_RS1635), galactokinase (JM81_RS01495), UDP-glucose-4-epimerase (JM81_RS00155), and nucleotidyltransferases (JM81_RS00330/14030/10945) were induced in the arabinogalactan proteome. Rhamnose and glucuronic acid are minor sugars in arabinogalactan from larch. Besides the aforementioned CAZymes and enzymes, three proteins related to the metabolism of sugar acids were induced in the arabinogalactan proteome. Gluconate dehydrogenases JM81_RS04310 and JM81_RS04965 yield 2-dehydro-D-gluconate. A 3-dehydro-L-gulonate 2-dehydrogenase (JM81_RS04965) was also induced.

3.4.3 Colony forming units (CFUs) microbiomes

North seawater sampled on 28 May 2022, at the beach of Sahlenburg, Cuxhaven, Germany, (N53.863607∘, E8.593004∘) yielded after 12 days of incubation a CFU microbiome of 2.57×10^4 ml on arabinogalactan and of 2.32×10^4 ml on ZoBell's 2216 marine agar. Partial 16S rRNA gene sequences of abundant CFU on arabinogalactan revealed 27 different genera within 88 colonies, with *Algoriphagus*, *Leeuwenhoekiella*, *Maribacter*, *Sulfitobacter*, *Yoonia-Loktanella*, and *Pseudoalteromonas* being present with 6–13 CFU. Only two of 63 CFU on ZoBell's 2216 marine agar, but 29 of 88 CFU on arabinogalactan plates belonged to *Flavobacteriaceae*. Additionally, we investigated autumn samples. Seawater sampled 18–21 October 2022 on Sylt, Germany (N55.031417∘, E8.441802∘), showed after 18 days of incubation 5.06 (+/−0.88) $\times 10^5$ CFU/ml for glucose and of 8.32 (+/− 2.21) $\times 10^5$ CFU/ml for arabinogalactan (n = 2). At the sandy west beach (N55.037505∘, E8.383341∘) plating of pore water yielded 1.75 (+/−1.30) $\times 10^6$ CFU/ml on for glucose and 0.90 (+/−0.64) $\times 10^6$ CFU/ml for arabinogalactan (n = 3).

3.5 Discussion

Galactose belongs to the four abundant monosaccharides in planktonic organic matter, mainly as part of polysaccharides and more complex molecules, i.e., arabinogalactan proteins. Plating sea and sediment pore water on arabinogalactan medium showed a large microbiome with the capacity to utilize arabinogalactan for growth. Together with the recent finding that particle associated bacteria dominate the readily culturable fraction of seater microbiomes (Heins and Harder, 2023) this observation indicates that arabinogalactan is a common carbon source for particle-associated bacteria. Arabinogalactan degradation pathways were so far only described for bacteria from gut and plant systems, but not for marine bacteria (Shulami *et al.*, 2011; Ndeh *et al.*, 2017; Cartmell *et al.*, 2018; Luis *et al.*, 2018; Fujita *et al.*, 2019; Wang and LaPointe, 2020; Sasaki *et al.*, 2021). These studies provided information regarding enzymes involved in arabinogalactan utilization, which includes GH families GH43, GH51, GH27, and GH28, often organized in PULs (Shulami *et al.*, 2011; Cartmell *et al.*, 2018; Luis *et al.*, 2018). Hence, we

inspected first the upregulated proteins in arabinogalactan grown cells in comparison to glucose grown cells. After a discussion of the SusC/D systems, we analyzed the uniqueness of marine PULs for arabinogalactan degradation in *Maribacter* sp. MAR_2009_72. Enzymes annotated as members of the GH family 43 are most frequently linked with arabinogalactan utilization. This protein family is one of the largest GH families and has currently 39 subfamilies, whereby only a part has been characterized enzymatically. Most common enzymatic functions are β-D-xylosidase (3.2.1.37), α-L-arabinofuranosidase (3.2.1.55), endo-α-L-arabinonase (3.2.1.99), and 1,3-β-galactosidase (3.2.1.145) (Mewis et al. 2016), with no correlation between subfamily classification and enzyme activity (Kelly, 1999; Wang and LaPointe, 2020), but as common feature the presence of three active sites in the enzymes. Still, many need a biochemical characterization (Pons *et al.*, 2004; Mewis *et al.*, 2016). The genome of *Maribacter* sp. MAR_2009_72 encodes five GH43 and four of these were shown to be expressed in this study. Subfamilies GH43_1 (JM81_RS00875) and GH43_26 (JM81_RS08585) are characterized so far solely as α-L-arabinofuranosidases. They hydrolyze via retention of configuration (Pitson *et al.*, 1996) the terminal nonreducing residues in α-L-arabinoside side chains of hemicelluloses, like arabinoxylan, arabinogalactan, and arabinan (Shallom *et al.*, 2002b; Shallom *et al.*, 2002a). JM81_RS16510 encodes a protein that consists out of two GH43 domains, GH43_19 and GH43_34. Over 70% of subfamily GH43_34 proteins contain a second GH43 domain (Mewis *et al.*, 2016). GH43_34 domains are usually located at the C-terminus and considered to be membrane-spanning (Mewis *et al.*, 2016). The function is not clear, but they may be involved in carbohydrate binding (Mewis *et al.*, 2016). Alike the GH43_1 and GH43_26 proteins, the GH43_19 domain is annotated as a catalytic α-L-arabinofuranosidase domain. The fourth GH43 protein is a member of subfamily GH43_18 (JM81_RS13895), which mostly includes α-L-arabinofuranosidases. The α-L-arabinofuranosidases function is also annotated to GH51 proteins, although the amino acid sequence differs by large from GH43 enzymes (Sevim *et al.*, 2017). The acid/base catalyst of GH51 is E176 in *Thermobacillus xylanticus* and E175 in *Geobacillus stearothermophilus* T-6 (Debeche *et al.*, 2002; Shallom *et al.*, 2002a). The GH51 from MAR_2009_72 (JM81_RS16515) has an E177, which aligns perfectly with the catalytic glutamic

acid of *T .xylanticus* and *G. stearothermophilus* T-6. The amino acid E294 in *G. stearothermophilus* T-6 was identified as catalytic nucleophile (Shallom *et al.*, 2002b), this glutamate is also conserved in GH51 of MAR_2009_72. In contrast to the inverting α-L-arabinofuranosidases of family GH43, α-L-arabinofuranosidases of GH51 hydrolyze bonds retaining the anomeric configuration (Pitson *et al.*, 1996). The family GH105 comprises unsaturated rhamnogalacturonyl hydrolases, which have been suggested to remove L-rhamnose-α-1,4-D-galacturonic acid on arabinogalactan (Munoz-Munoz *et al.*, 2021). Four GH105 proteins were expressed. The catalysis involves hydrolysis of bonds through syn-hydration of double bonds between C4 and C5 carbons of enopyranuronosyl residues of substrates (Itoh *et al.*, 2006a, b). The catalytic aspartate side chain D143 of GH105 is conserved in GH105 proteins of MAR_2009_72 (Itoh *et al.*, 2006b). The transport of the oligosaccharides involved several SusC/D pairs. PULs 1, 7, and 8 encode the three SusC/D systems that had the highest expression intensities of all SusC/Ds in the arabinogalactan proteome. On the basis of the dedicated substrate specificity of SusC/D transport systems, we propose two explanations for the induction of several SusC/D pairs: (i) the extracellular hydrolysis of larch wood arabinogalactan generates a mixture of structurally different oligosaccharides which need dedicated transport system and (ii) a signal molecule derived from larch wood arabinogalactan may induce the expression of proteins that may not be necessary for larch wood arabinogalactan, but for the degradation of marine arabinogalactans. The structural diversity of arabinogalactans in terrestrial system is well characterized (Fujita *et al.*, 2019; Villa-Rivera *et al.*, 2021; Leszczuk *et al.*, 2023), but marine arabinogalactans are understudied. In the periplasm the oligosaccharides are further hydrolyzed by a range of CAZymes. Some PULs (1 and 7) expressed enzymes that can generate monomers. Furthermore, the proteome detected CAZymes that are not encoded in PULs and are predicted to be periplasmatic. The GH10 of PUL 1 was annotated as an endo-1,4-β-xylanase, which indicates that arabinoxylans may also be a substrate for the enzymes of PUL 1. The expression of an αglucuronidase annotated to GH67, coincides with the presence of glucuronic acid in side chains of arabinogalactan. GH67 removes glucuronic acid from side chains by a single displacement mechanism using an inverting mechanism (Shulami *et al.*, 1999; Biely *et al.*, 2000; Nagy *et al.*, 2002). But it only removes glucuronic acid from

nonreducing ends of the oligo- and polysaccharides. A broader substrate range is known for GH115 proteins, which remove glucuronic acid from terminal and internal regions of oligosaccharides (Ryabova *et al.*, 2009; Aalbers *et al.*, 2015). The presence of both GH families, GH67 and two GH115, suggests that glucuronic acid is part of the decoration of arabinogalactans. The expression of the GH29 argues for fucose as a decorating sugar. Enzymes of the family GH29 are exo-α-fucosidases and cleave via a retaining mechanism (Grootaert *et al.*, 2020). Also, rhamnose as specific substrate is supported by expression of a GH78, α-L-rhamnosidase. This GH family solely includes rhamnosidases, which use an inverting mechanism to hydrolyze bonds in cooperation with their catalytic residues (Cui *et al.*, 2007). The galactan backbone hydrolysis requires a β-D-galactosidase. This enzymatic function is frequent among members of the GH family GH3. The proteome detected three expressed GH3 proteins. Final steps of the arabinogalactan pathway include the translocation through the inner membrane, likely via an ABC transport system, and cytoplasmic transformations to channel galactose, arabinose, glucuronic acid, rhamnose, and fucose into the pentose phosphate pathway and the glycolysis. We investigated the distribution of PUL 1, 7, and 8 of *Maribacter* sp. MAR_2009_72 in the PULDB database using the expressed CAZymes (Terrapon *et al.*, 2018). Homologues of PUL 1 have been characterized for human gut bacteria and Bacteroides spp. for the utilization of a range of xylan polysaccharides including arabinoxylan (Martens *et al.*, 2008; Rogowski *et al.*, 2015; Wang *et al.*, 2016). The PUL was *in-silico* detected in genomes of a large variety of Bacteroidota. In contrast, PUL 7 has so far not been studied experimentally. An *in-silico* search detected a homologous PUL structure in *Maribacter sedimenticola* DSM 19840 (Nedashkovskaya *et al.*, 2004). PUL 8 has also a homologue in *M. sedimenticola* DSM 19840 and other *Bacteroidota*. A recent metagenomic study of particle-associated bacteria detected a GH43-rich PUL in a *Maribacter* MAG, which the authors annotated as an arabinogalactan PUL (Wang *et al.*, 2024). This PUL is different to the PULs we identified for arabinogalactan in the genome of *Maribacter* sp. MAR_2009_72. Our observations revealed a substrate specificity of the three PULs. In PUL 1, arabinogalactan is the only inducer for SusC/D, and the expression of a glucuronidase and a xylanase suggests that also glycans with these sugars are substrates for the PUL (Supplement Figure S3).

This hypothesis is supported by previous studies with gut bacteria (Martens *et al.*, 2008; Rogowski *et al.*, 2015; Wang *et al.*, 2016). PULs 7 and 8 have so far not been experimentally observed. PUL 7 is characterized by a very strong induction of SusC/D by galactose and arabinogalactan (Supplement; Figure S4). Galactose is for several proteins the strongest inducer, suggesting galactans as substrate. The presence of fucosidase, glucuronidase, and rhamnosidase suggests a decoration of the marine galactans with the corresponding monosaccharides. PUL 8 is dedicated to arabinose containing glycans. The SusC/D is induced by arabinose and arabinogalactan (Supplement Figure S5). Besides GHs, the genetic region of PUL 8 includes also monosaccharide-transforming cytoplasmatic enzymes for arabinose and sugar acids. This PUL shows that the consideration of cytosolic carbohydrate-transforming enzymes in the bioinformatic analysis of PULs may improve predictions of substrate specificity. The comparative proteomic analysis of larch wood arabinogalactan degradation by *Maribacter* sp. MAR_2009_72 identified expressed proteins encoded in three PULs and outside of PULs (Figure 3). In summary, members of the GH families 43, 51, and 105 may produce a variety of oligosaccharides. At least three SusC/D systems are involved in the transport into the periplasm, where enzymes belonging to the GH families 3, 10, 29, 67, 78, and 115 produce monosaccharides. The interplay of all these enzymes allows for the utilization of arabinogalactan, which we have summarized in a graph (Figure 7).
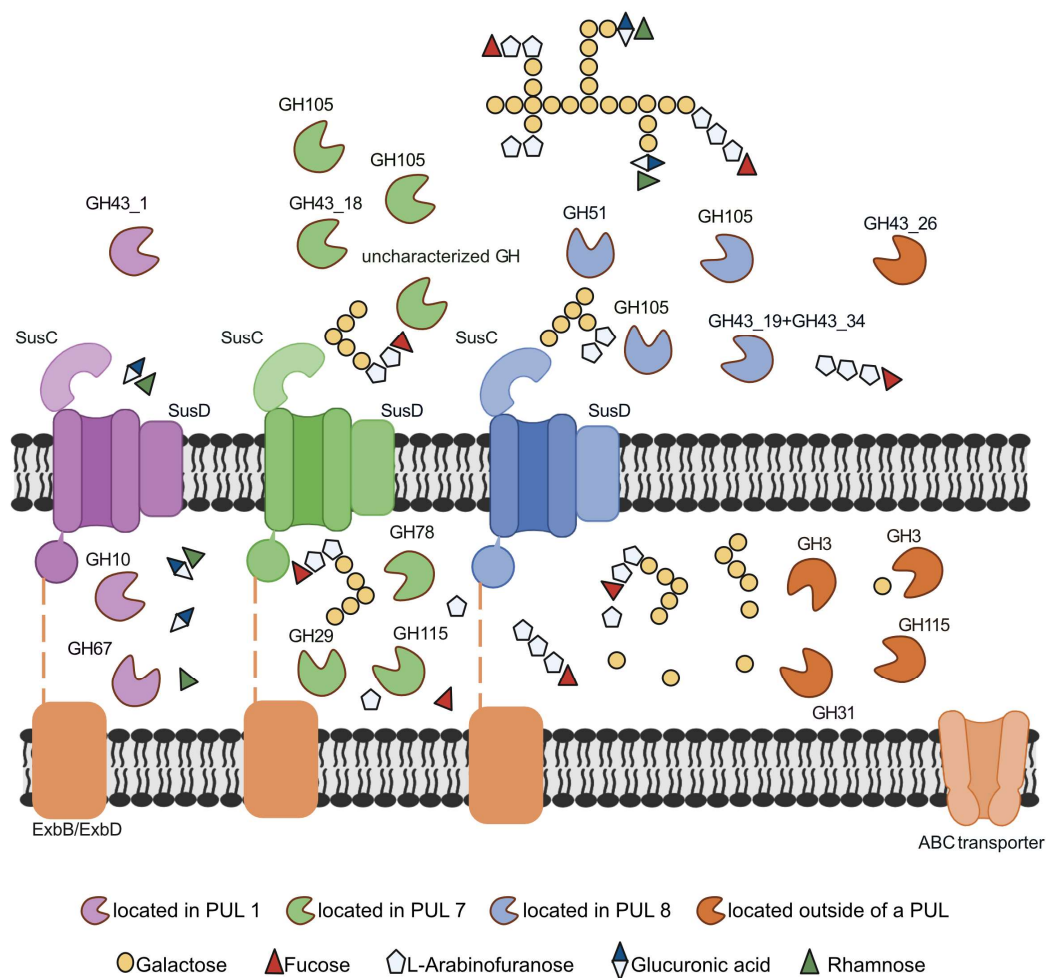
Figure 7: Proposed utilization pathway for arabinogalactan by *Maribacter* sp. MAR_2009_72. Poly- and oligosaccharide structures are not experimentally verified. PUL 1: GH43_1 (JM81_RS00875), SusC (JM81_RS00910), SusD (JM81_RS00905), GH10 (JM81_RS00885), and GH67 (JM81_RS00895). PUL 7: GH105 (JM81_RS13845), GH105 (JM81_RS13890), GH43_18 (JM81_RS13895), uncharacterized GH (JM81_RS13825), SusC (JM81_RS13730), SusD (JM81_RS13725), GH29 (JM81_RS13700), GH78 (JM81_RS13900), and GH115 (JM81_RS13820). PUL 8: GH105 (JM81_RS16475), GH105 (JM81_RS16470), GH51 (JM81_RS16515), GH43_19+GH43_34 (JM81_RS16510), SusC (JM81_RS16585), and SusD (JM81_RS16470). Outside of the PULs: GH43_26 (JM81_RS08590), GH3 (JM81_RS00095), GH3 (JM81_RS08450), GH31 (JM81_RS06015), CBM48+GH13_9 (JM81_RS06020), GH115 (JM81_RS03245), ExbB (JM81_RS07935), ExbD (JM81_RS07930), ABC substrate-binding (JM81_RS03610), ABC permease (JM81_RS16840), and ABC ATP binding (JM81_RS01625). Created with BioRender.com.

The plant polysaccharide structure is expected to be less complex than the variety of arabinogalactans present in the marine habitat (Pfeifer *et al.*, 2020). This may explain why not all CAZymes of each PUL were detected as expressed proteins. A difference between this study of a marine bacterium and previous studies on gut and plant associated bacteria was the presence of GH105 enzymes and the absence of GH27 and GH28 enzymes. Future studies might characterize marine

arabinogalactans and enzymatic studies will resolve the individual functions of the induced proteins to provide further information on the microbial utilization.

3.6 Acknowledgement

3.7 Supplement

3.7.1 Supplement Methods

Cells grown on arabinogalactan to the late logarithmic growth phase (OD 0.2 – 0.3) were harvested by centrifugation at 16000x g for 15 min at 21°C. The cell pellet was transiently stored at -20°C. Cells were lyzed in 1 mL 50 mM 3-(N-morpholino) propanesulfonic acid (MOPS), 2 mM dithiothreitol, pH 7.0 by sonication on ice using a Sonoplus HD70 Bandelin MS73 (BANDELIN, Berlin, Deutschland) with a titanium sonotrode MS73 (BANDELIN) for 4 min at 50% power and 50% cycle (0.5 sec on, 0.5 sec off). The soluble protein fraction was separated by centrifugation for 15 min at 16000x g at 21°C. The pellet was resuspended in 1 mL MOPS buffer and named membrane protein fraction. Protein concentration was determined by the Bradford protein assay using bovine serum albumin as standard (Bradford, 1976). Arabinogalactan assays were performed with 11 and 8 µg protein for the soluble and the membrane protein fraction, respectively, and 250 µg arabinogalactan in 100 µl MOPS for 20 h at 21°C. After vacuum drying at 45°C for 1 hour (Eppendorf Concentrator plus, Eppendorf, Hamburg, Germany), the sugars were labelled with 2 µl of 0.2 M 8-amino-1,3,6-naphthalenetrisulfonate (ANTS) in the presence of 5 µl 1M cyanoborohydride at 37°C for 20 h. Then the reaction was diluted with 25 µl 25% vol/vol glycerol and analyzed on 30% v/v acrylamide gels following a protocol for fluorophore-assisted carbohydrate electrophoresis (FACE) (Becker *et al.*, 2017). We introduced a concentration of 65 mM sodium chloride into

the separation gel, which allowed a separation of hexose (galactose) and pentose (arabinose). Separation was performed at 100 V for 30 minutes, followed by 200 V for 60 minutes. Gels were documented using a Bio-Rad GelDoc EZ Gel Imaging System (Cambridge Scientifix, Watertown, US).

3.7.2 Supplement Figures
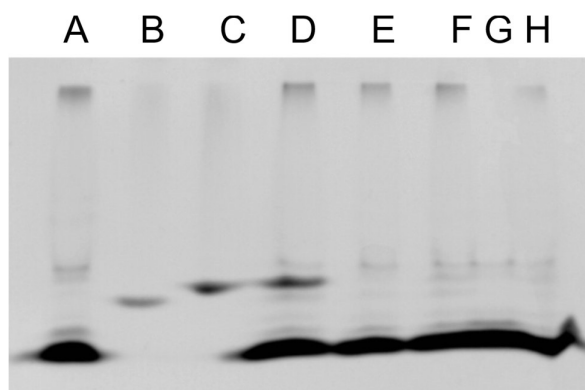
Supplement Figure 1: Fluorescently labelled carbohydrates separated by size in an acrylamide gel. A: 2 µl of 2 g/L arabinogalactan; B: 2 µl of 0.2 g/L arabinose; C: 2 µl of 0.2 g/L galactose; D: arabinogalactan degrading reaction catalyzed by soluble protein fraction; E: arabinogalactan degrading reaction catalyzed by membrane protein fraction; F: arabinogalactan degrading reaction catalyzed by heat-inactivated soluble protein fraction; G: reaction D lacking arabinogalactan; H: reaction D lacking the soluble protein fraction
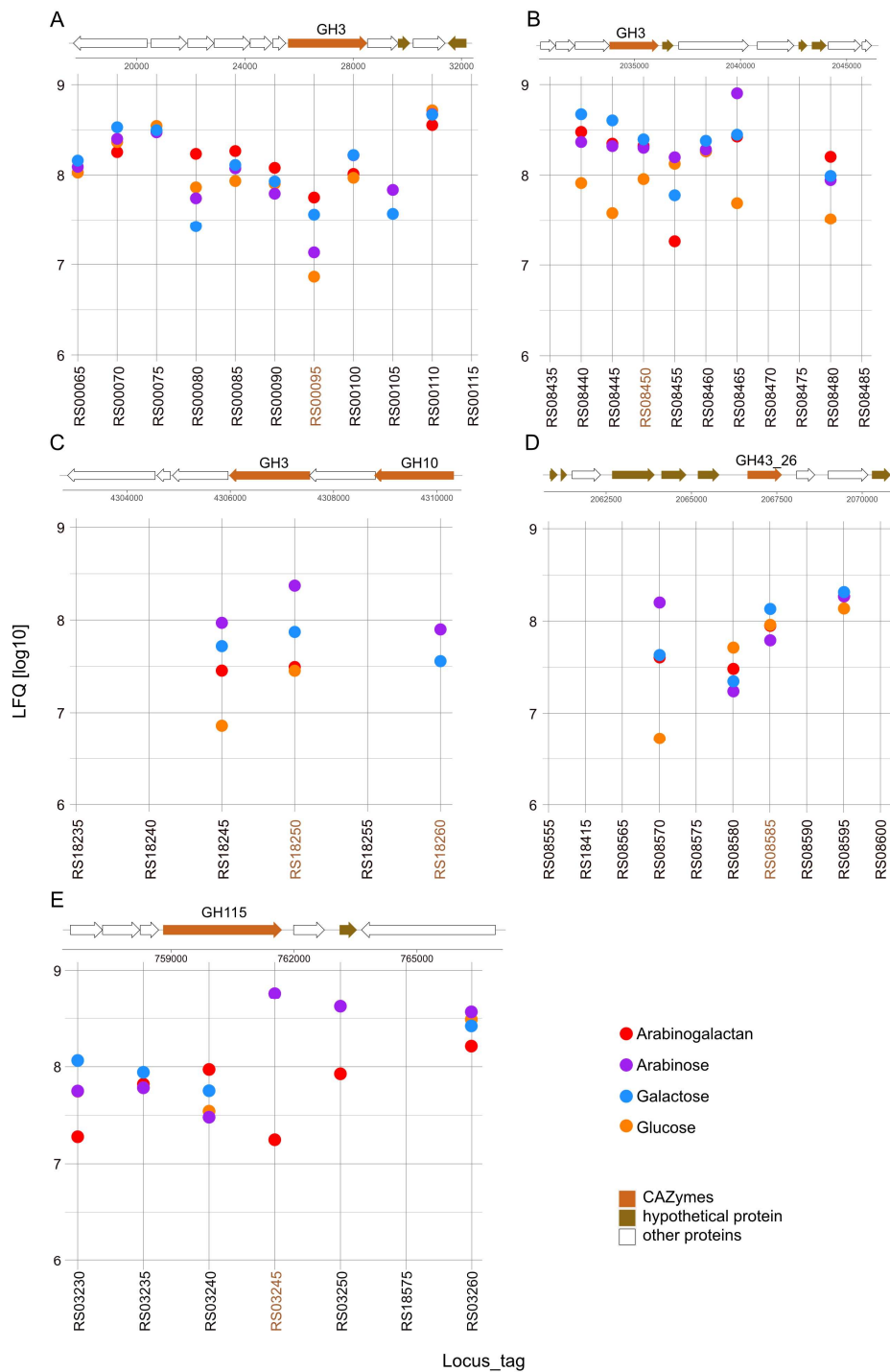
Supplement Figure 2: Gene organization and expression of CAZymes located outside of polysaccharide utilization loci of *Maribacter* sp. MAR_2009_72 grown in the presence of arabinogalactan, arabinose, galactose and glucose. Expression intensities in the plot are the mean values of three biological replicates of each condition shown in LFQ values [log10]. A: GH3 (JM81_RS00095) α-galactosidase; B: GH3 (JM81_RS08450) α-galactosidase; C: GH3 (JM81_RS18250) α-galactosidase, GH10 (JM81_RS18260) endo-1,4-β-xylanase; D: GH43_26 (JM81_RS08585) α-L-arabinofuranosidase; E: GH115 (JM81_RS03245) xylan-α-1,2-glucuronidase
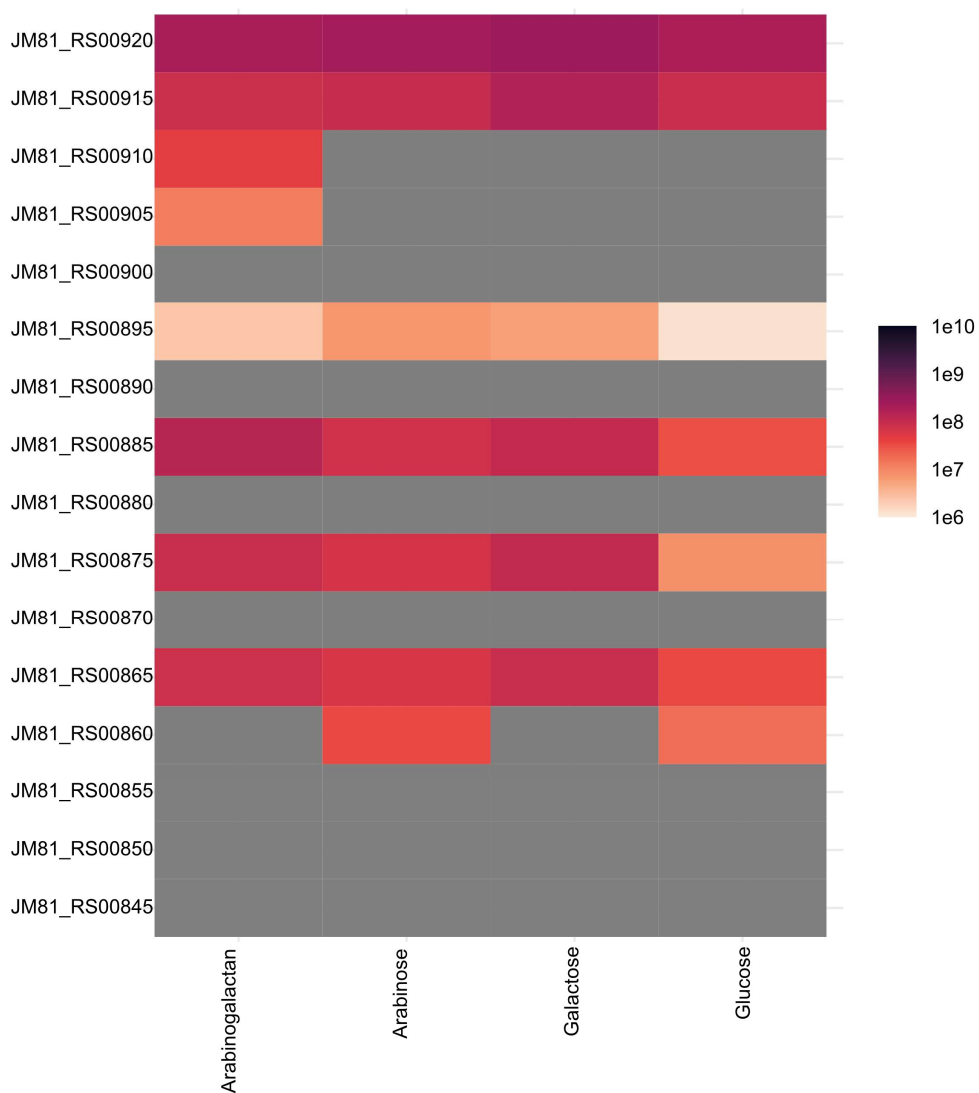
Supplement Figure 3: Heatmap of PUL 1 of *Maribacter* sp. MAR_2009_72 grown in the presence of arabinogalactan, arabinose, galactose and glucose. Expression intensities in the plot are mean values of three biological replicates of each condition shown in LFQ values [log10]. Grey means the protein was not dected.
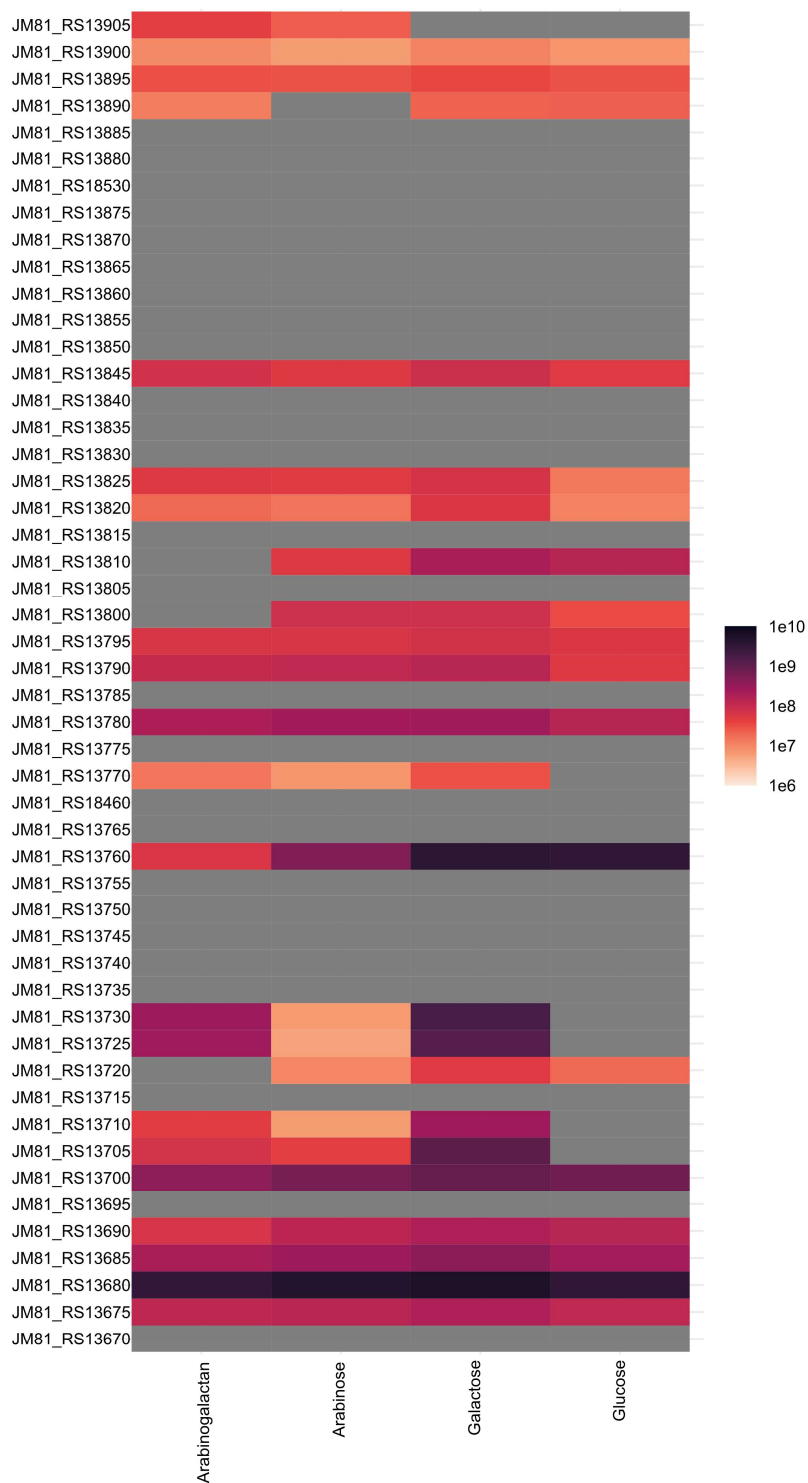
Supplement Figure 4: Heatmap of PUL 7 of *Maribacter* sp. MAR_2009_72 grown in the presence of arabinogalactan, arabinose, galactose and glucose. Expression intensities in the plot are mean values of three biological replicates of each condition shown in LFQ values [log10]. Grey means the protein was not dected.
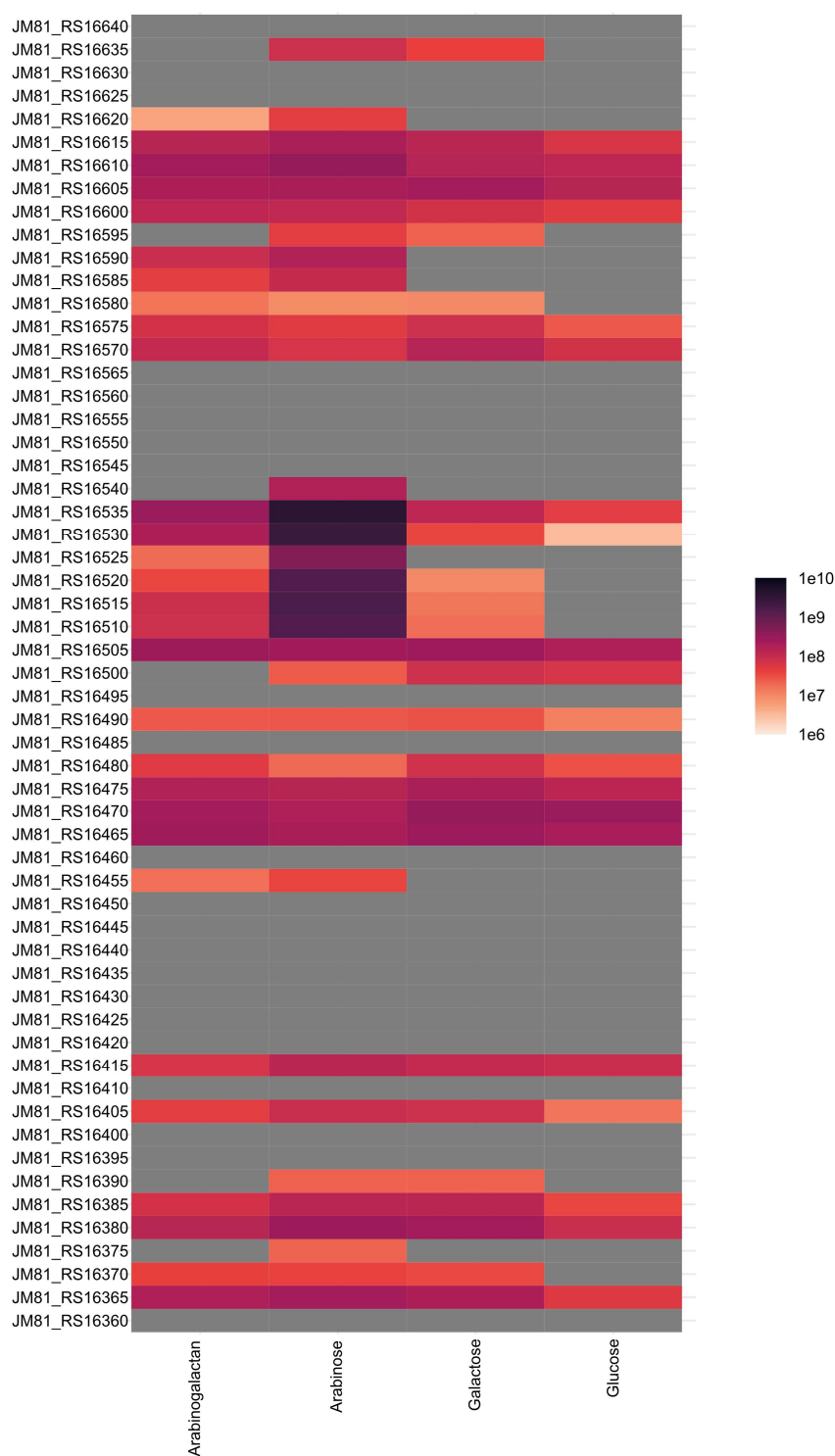
Supplement Figure 5: Heatmap of PUL 8 of *Maribacter* sp. MAR_2009_72 grown in the presence of arabinogalactan, arabinose, galactose and glucose. Expression intensities in the plot are mean values of three biological replicates of each condition shown in LFQ values [log10]. Grey means the protein was not dected.

3.8 References

Aalbers, F., Turkenburg, J.P., Davies, G.J., Dijkhuizen, L., and Lammerts van Bueren, A. (2015) Structural and functional characterization of a novel family GH115 4-O-methyl-α-glucuronidase with specificity for decorated arabinogalactans. *Journal of Molecular Biology* **427**: 3935-3946.

Alderkamp, A.-C., Buma, A.G.J., and van Rijssel, M. (2007) The carbohydrates of *Phaeocystis* and their degradation in the microbial food web. *Biogeochemistry* **83**: 99-118.

Bäumgen, M., Dutschei, T., and Bornscheuer, U.T. (2021) Marine polysaccharides: Occurrence, enzymatic degradation and utilization. *ChemBioChem* **22**: 2247-2256.

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* **83**: e03389-16.

Biely, P., de Vries, R.P., Vršanská, M., and Visser, J. (2000) Inverting character of α-glucuronidase A from *Aspergillus tubingensis*. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1474**: 360-364.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**: 248-254.

Cartmell, A., Muñoz-Muñoz, J., Briggs, J.A., Ndeh, D.A., Lowe, E.C., Baslé, A. *et al.* (2018) A surface endogalactanase in *Bacteroides thetaiotaomicron* confers keystone status for arabinogalactan degradation. *Nature Microbiology* **3**: 1314-1326.

Cui, Z., Maruyama, Y., Mikami, B., Hashimoto, W., and Murata, K. (2007) Crystal structure of glycoside hydrolase family 78 α-L-rhamnosidase from *Bacillus* sp. GL1. *Journal of Molecular Biology* **374**: 384-398.

Debeche, T., Bliard, C., Debeire, P., and O'Donohue, M.J. (2002) Probing the catalytically essential residues of the α-L-arabinofuranosidase from *Thermobacillus xylanilyticus*. *Protein Engineering, Design and Selection* **15**: 21-28.

Drula, E., Garron, M.-L., Dogan, S., Lombard, V., Henrissat, B., and Terrapon, N. (2022) The carbohydrate-active enzyme database: Functions and literature. *Nucleic Acids Research* **50**: 571-577.

Fujita, K., Sasaki, Y., and Kitahara, K. (2019) Degradation of plant arabinogalactan proteins by intestinal bacteria: Characteristics and functions of the enzymes involved. *Applied Microbiology and Biotechnology* **103**: 7451-7457.

Glover, J.S., Ticer, T.D., and Engevik, M.A. (2022) Characterizing the mucin-degrading capacity of the human gut microbiota. *Scientific Reports* **12**: 8456.

Grant, J.R., Enns, E., Marinier, E., Mandal, A., Herman, E.K., Chen, C.-y. *et al.* (2023) Proksee: In-depth characterization and visualization of bacterial genomes. *Nucleic Acids Research* **51**: W484-W492.

Grootaert, H., Van Landuyt, L., Hulpiau, P., and Callewaert, N. (2020) Functional exploration of the GH29 fucosidase family. *Glycobiology* **30**: 735-745.

Hahnke, R.L., and Harder, J. (2013) Phylogenetic diversity of *Flavobacteria* isolated from the North Sea on solid media. *Systematic and Applied Microbiology* **36**: 497-504.

Hahnke, R.L., Bennke, C.M., Fuchs, B.M., Mann, A.J., Rhiel, E., Teeling, H. *et al.* (2015) Dilution cultivation of marine heterotrophic bacteria abundant after a spring phytoplankton bloom in the North Sea. *Environmental Microbiology* **17**: 3515-3526.

Hallgren, J., Tsirigos, K.D., Pedersen, M.D., Almagro Armenteros, J.J., Marcatili, P., Nielsen, H. *et al.* (2022) Preprint DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. *bioRxiv* 2022.04.08.487609.

Heins, A., and Harder, J. (2023) Particle-associated bacteria in seawater dominate the colony-forming microbiome on ZoBell marine agar. *FEMS Microbiology Ecology* **99**: fiac151.

Heins, A., Amann, R.I., and Harder, J. (2021a) Cultivation of particle-associated heterotrophic bacteria during a spring phytoplankton bloom in the North Sea. *Systematic and Applied Microbiology* **44**: 126232.

Heins, A., Reintjes, G., Amann, R.I., and Harder, J. (2021b) Particle collection in Imhoff sedimentation cones enriches both motile chemotactic and particle-attached bacteria. *Frontiers in Microbiology* **12**: 643730.

Hinz, S.W.A., Verhoef, R., Schols, H.A., Vincken, J.-P., and Voragen, A.G.J. (2005) Type I arabinogalactan contains β-D-Galp-(1→3)-β-D-Galp structural elements. *Carbohydrate Research* **340**: 2135-2143.

Huang, G., Vidal-Melgosa, S., Sichert, A., Becker, S., Fang, Y., Niggemann, J. *et al.* (2021) Secretion of sulfated fucans by diatoms may contribute to marine aggregate formation. *Limnology and Oceanography* **66**: 3768-3782.

Itoh, T., Ochiai, A., Mikami, B., Hashimoto, W., and Murata, K. (2006a) A novel glycoside hydrolase family 105: The structure of family 105 unsaturated rhamnogalacturonyl hydrolase complexed with a disaccharide in comparison with family 88 enzyme complexed with the disaccharide. *Journal of Molecular Biology* **360**: 573-585.

Itoh, T., Ochiai, A., Mikami, B., Hashimoto, W., and Murata, K. (2006b) Structure of unsaturated rhamnogalacturonyl hydrolase complexed with substrate. *Biochemical and Biophysical Research Communications* **347**: 1021-1029.

Ittekkot, V., Degens, E.T., and Brockmann, U. (1982) Monosaccharide composition of acid-hydrolyzable carbohydrates in particulate matter during a plankton bloom. *Limnology and Oceanography* **27**: 770-776.

Kanehisa, M., Sato, Y., and Morishima, K. (2016) BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *Journal of Molecular Biology* **428**: 726-731.

Kappelmann, L., Krüger, K., Hehemann, J.-H., Harder, J., Markert, S., Unfried, F. *et al.* (2019) Polysaccharide utilization loci of North Sea *Flavobacteriia* as basis for using SusC/D-protein expression for predicting major phytoplankton glycans. *The ISME Journal* **13**: 76-91.

Kelly, G.S. (1999) Larch arabinogalactan: clinical relevance of a novel immune-enhancing polysaccharide. *Alternative medicine review : a journal of clinical therapeutic* **4**: 96-103.

Leszczuk, A., Kalaitzis, P., Kulik, J., and Zdunek, A. (2023) Review: Structure and modifications of arabinogalactan proteins (AGPs). *BMC Plant Biology* **23**: 45.

Lu, D.-C., Wang, F.-Q., Amann, R.I., Teeling, H., and Du, Z.-J. (2023) Epiphytic common core bacteria in the microbiomes of co-located green (*Ulva*), brown (*Saccharina*) and red (*Grateloupia*, *Gelidium*) macroalgae. *Microbiome* **11**: 126.

Lu, S., Wang, J., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R. *et al.* (2019) CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Research* **48**: 265-268.

Luis, A.S., Briggs, J., Zhang, X., Farnell, B., Ndeh, D., Labourel, A. *et al.* (2018) Dietary pectic glycans are degraded by coordinated enzyme pathways in human colonic *Bacteroides*. *Nature Microbiology* **3**: 210-219.

Martens, E.C., Chiang, H.C., and Gordon, J.I. (2008) Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. *Cell Host & Microbe* **4**: 447-457.

Mewis, K., Lenfant, N., Lombard, V., and Henrissat, B. (2016) Dividing the large glycoside hydrolase family 43 into subfamilies: A motivation for detailed enzyme characterization. *Applied and Environmental Microbiology* **82**: 1686-1692.

Miksch, S., Meiners, M., Meyerdierks, A., Probandt, D., Wegener, G., Titschack, J. *et al.* (2021) Bacterial communities in temperate and polar coastal sands are seasonally stable. *ISME Communications* **1**: 29.

Munoz-Munoz, J., Ndeh, D., Fernandez-Julia, P., Walton, G., Henrissat, B., and Gilbert, H.J. (2021) Sulfation of arabinogalactan proteins confers privileged nutrient status to *Bacteroides plebeius*. *mBio* **12**: 10.1128/mbio.01368-21.

Nagy, T., Emami, K., Fontes Carlos, M.G.A., Ferreira Luis, M.A., Humphry David, R., and Gilbert Harry, J. (2002) The membrane-bound α-glucuronidase from *Pseudomonas cellulosa* hydrolyzes 4-O-methyl-d-glucuronoxylooligosaccharides but not 4-O-methyl-d-glucuronoxylan. *Journal of Bacteriology* **184**: 4925-4929.

Ndeh, D., Rogowski, A., Cartmell, A., Luis, A.S., Baslé, A., Gray, J. *et al.* (2017) Complex pectin metabolism by gut bacteria reveals novel catalytic functions. *Nature* **544**: 65-70.

Nedashkovskaya, O.I., Kim, S.B., Han, S.K., Lysenko, A.M., Rohde, M., Rhee, M.-S. *et al.* (2004) *Maribacter* gen. nov., a new member of the family *Flavobacteriaceae*, isolated from marine habitats, containing the species *Maribacter sedimenticola* sp. nov., *Maribacter aquivivus* sp. nov., *Maribacter orientalis* sp. nov. and *Maribacter ulvicola* sp. nov. *International Journal of Systematic and Evolutionary Microbiology* **54**: 1017-1023.

Noinaj, N., Guillier, M., Barnard, T.J., and Buchanan, S.K. (2010) TonB-dependent transporters: Regulation, structure, and function. *Annual Review of Microbiology* **64**: 43-60.

Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B.L., Salazar, Gustavo A. *et al.* (2022) InterPro in 2022. *Nucleic Acids Research* **51**: 418-427.

Perez-Riverol, Y., Bai, J., Bandla, C., García-Seisdedos, D., Hewapathirana, S., Kamatchinathan, S. *et al.* (2022) The PRIDE database resources in 2022: A hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Research* **50**: D543-D552.

Pfeifer, L., Shafee, T., Johnson, K.L., Bacic, A., and Classen, B. (2020) Arabinogalactan-proteins of *Zostera marina L.* contain unique glycan structures and provide insight into adaption processes to saline environments. *Scientific Reports* **10**: 8232.

Pitson, S.M., Voragen, A.G.J., and Beldman, G. (1996) Stereochemical course of hydrolysis catalyzed by arabinofuranosyl hydrolases. *FEBS Letters* **398**: 7-11.

Pons, T., Naumoff, D.G., Martínez-Fleites, C., and Hernández, L. (2004) Three acidic residues are at the active site of a β-propeller architecture in glycoside hydrolase families 32, 43, 62, and 68. *Proteins: Structure, Function, and Bioinformatics* **54**: 424-432.

Probandt, D., Eickhorst, T., Ellrott, A., Amann, R., and Knittel, K. (2018) Microbial life on a sand grain: From bulk sediment to single grains. *The ISME Journal* **12**: 623-633.

R Core Team (2023) R: A language and environment for statistical computing. R Foundation for Statistical Computing. In. Vienna, Austria.

Rogowski, A., Briggs, J.A., Mortimer, J.C., Tryfona, T., Terrapon, N., Lowe, E.C. *et al.* (2015) Glycan complexity dictates microbial resource allocation in the large intestine. *Nature Communications* **6**: 7481.

Ryabova, O., Vršanská, M., Kaneko, S., van Zyl, W.H., and Biely, P. (2009) A novel family of hemicellulolytic α-glucuronidase. *FEBS Letters* **583**: 1457-1462.

Sasaki, Y., Horigome, A., Odamaki, T., Xiao, J.-Z., Ishiwata, A., Ito, Y. *et al.* (2021) Novel 3-O-α-D-galactosyl-α-L-arabinofuranosidase for the assimilation of gum arabic arabinogalactan protein in *Bifidobacterium longum* subsp. *longum*. *Applied and Environmental Microbiology* **87**: e02690-20.

Sato, K., Hara, K., Yoshimi, Y., Kitazawa, K., Ito, H., Tsumuraya, Y., and Kotake, T. (2018) Yariv reactivity of type II arabinogalactan from larch wood. *Carbohydrate Research* **467**: 8-13.

Sayers, E.W., Bolton, E.E., Brister, J.R., Canese, K., Chan, J., Comeau, Donald C. *et al.* (2022) Database resources of the national center for biotechnology information. *Nucleic Acids Research* **50**: 20-26.

Scholz, B., and Liebezeit, G. (2013) Biochemical characterisation and fatty acid profiles of 25 benthic marine diatoms isolated from the Solthörn tidal flat (southern North Sea). *Journal of Applied Phycology* **25**: 453-465.

Schultz, D., Zühlke, D., Bernhardt, J., Francis, T.B., Albrecht, D., Hirschfeld, C. *et al.* (2020) An optimized metaproteomics protocol for a holistic taxonomic and functional characterization of microbial communities from marine particles. *Environmental Microbiology Reports* **12**: 367-376.

Sevim, E., Inan Bektas, K., Sevim, A., Canakci, S., Sahin, I., and Belduz, A.O. (2017) Purification and characterization of α-L-arabinofuranosidases from *Geobacillus stearothermophilus* strain 12. *Biologia* **72**: 831-839.

Shallom, D., Belakhov, V., Solomon, D., Shoham, G., Baasov, T., and Shoham, Y. (2002a) Detailed kinetic analysis and identification of the nucleophile in α-L-arabinofuranosidase from *Geobacillus stearothermophilus* T-6, a family 51 glycoside hydrolase. *Journal of Biological Chemistry* **277**: 43667-43673.

Shallom, D., Belakhov, V., Solomon, D., Gilead-Gropper, S., Baasov, T., Shoham, G., and Shoham, Y. (2002b) The identification of the acid–base catalyst of α-arabinofuranosidase from *Geobacillus stearothermophilus* T-6, a family 51 glycoside hydrolase. *FEBS Letters* **514**: 163-167.

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* **182**: 5365-5372.

Shulami, S., Gat, O., Sonenshein Abraham, L., and Shoham, Y. (1999) The glucuronic acid utilization gene cluster from *Bacillus stearothermophilus* T-6. *Journal of Bacteriology* **181**: 3695-3704.

Shulami, S., Raz-Pasteur, A., Tabachnikov, O., Gilead-Gropper, S., Shner, I., and Shoham, Y. (2011) The L-arabinan utilization system of *Geobacillus stearothermophilus*. *Journal of Bacteriology* **193**: 2838-2850.

Sidhu, C., Kirstein, I.V., Meunier, C.L., Rick, J., Fofonova, V., Wiltshire, K.H. *et al.* (2023) Dissolved storage glycans shaped the community composition of abundant bacterioplankton clades during a North Sea spring phytoplankton bloom. *Microbiome* **11**: 77.

Silva, J., Ferraz, R., Dupree, P., Showalter, A.M., and Coimbra, S. (2020) Three decades of advances in arabinogalactan-protein biosynthesis. *Frontiers in Plant Science* **11**: 610377.

Stam, M., Lelièvre, P., Hoebeke, M., Corre, E., Barbeyron, T., and Michel, G. (2022) SulfAtlas, the sulfatase database: State of the art and new developments. *Nucleic Acids Research* **51**: 647-653.

Tailford, L.E., Crost, E.H., Kavanaugh, D., and Juge, N. (2015) Mucin glycan foraging in the human gut microbiome. *Frontiers in Genetics* **6**: 81.

Terrapon, N., Lombard, V., Drula, É., Lapébie, P., Al-Masaudi, S., Gilbert, H.J., and Henrissat, B. (2018) PULDB: The expanded database of polysaccharide utilization loci. *Nucleic Acids Research* **46**: D677-D683.

Teufel, F., Almagro Armenteros, J.J., Johansen, A.R., Gíslason, M.H., Pihl, S.I., Tsirigos, K.D. *et al.* (2022) SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nature Biotechnology* **40**: 1023-1025.

The UniProt, C. (2023) UniProt: The universal protein knowledgebase in 2023. *Nucleic Acids Research* **51**: D523-D531.

Tyanova, S., and Cox, J. (2018) Perseus: A bioinformatics platform for integrative analysis of proteomics data in cancer research. In *Cancer Systems Biology: Methods and Protocols*. von Stechow, L. (ed). New York, NY: Springer New York, pp. 133-148.

Tyanova, S., Temu, T., and Cox, J. (2016) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols* **11**: 2301-2319.

Urbani, R., Magaletti, E., Sist, P., and Cicero, A.M. (2005) Extracellular carbohydrates released by the marine diatoms *Cylindrotheca closterium*, *Thalassiosira pseudonana* and *Skeletonema costatum*: Effect of P-depletion and growth status. *Science of The Total Environment* **353**: 300-306.

Vidal-Melgosa, S., Sichert, A., Francis, T.B., Bartosik, D., Niggemann, J., Wichels, A. *et al.* (2021) Diatom fucan polysaccharide precipitates carbon during algal blooms. *Nature Communications* **12**: 1150.

Villa-Rivera, M.G., Cano-Camacho, H., López-Romero, E., and Zavala-Páramo, M.G. (2021) The role of arabinogalactan type II degradation in plant-microbe interactions. *Frontiers in Microbiology* **12**: 730543.

Wang, F.-Q., Bartosik, D., Sidhu, C., Siebers, R., Lu, D.-C., Trautwein-Schult, A. *et al.* (2024) Particle-attached bacteria act as gatekeepers in the decomposition of complex phytoplankton polysaccharides. *Microbiome* **12**: 32.

Wang, W., Yan, R., Nocek, B.P., Vuong, T.V., Di Leo, R., Xu, X. *et al.* (2016) Biochemical and structural characterization of a five-domain GH115 α-glucuronidase from the marine bacterium *Saccharophagus degradans* 2-40T*. *Journal of Biological Chemistry* **291**: 14120-14133.

Wang, Y., and LaPointe, G. (2020) Arabinogalactan utilization by *Bifidobacterium longum* subsp. *longum* NCC 2705 and *Bacteroides caccae* ATCC 43185 in monoculture and coculture. *Microorganisms* **8**: 1703.

Wickham, H. (2016) *ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag.

Wilkins, D. (2023) gggenes: Draw gene arrow maps in 'ggplot2'. In. https://wilkox.org/gggenes/.

Zheng, J., Ge, Q., Yan, Y., Zhang, X., Huang, L., and Yin, Y. (2023) dbCAN3: Automated carbohydrate-active enzyme and substrate annotation. *Nucleic Acids Research* **51**: W115-W121.

Chapter 4

Proteomic overview of glycan utilization by particle-associated *Maribacter forsetii*

Saskia Kalenborn[1], Daniela Zühlke[2], Katharina Riedel[2], Rudolf I. Amann[1], Jens Harder[1*]

[1]Department of Molecular Ecology, Max Planck Institute for Marine Microbiology, Bremen, Germany; [2]Department for Microbial Physiology and Molecular Biology, University of Greifswald, Germany

In preparation

Contribution of the candidate in % of the total work

Experimental concept and design – 60%

Experimental work/acquisition of experimental data – 95%

Preparation of figures and tables – 100%

Drafting of the manuscript – 90%

Chapter 4

4.1 Abstract

Flavobacteriia are marine specialists for the degradation of sugars. In this study, we investigated the growth and polysaccharide utilization capabilities of *Maribacter forsetii* in liquid media with casamino acids as a limiting carbon source, supplemented with 2 g/L of various glycans. *M. forsetii* demonstrated the ability to grow on a wide range of polysaccharides, including arabinogalactan, chondroitin sulfate, fucoidan, carrageenans, laminarin and many monosaccharides. Notably we observed growth on substrates previously considered to be unusable based the absence of polysaccharide loci in an in-silico study, indicating a broader metabolic versatility than anticipated.

We identified six polysaccharide utilization loci (PULs) in the genome of *M. forsetii*, containing CAZymes and SusC/D pairs. A multitude of CAZymes and SusC/D pairs were not located within PULs and their expression suggested a modular approach to polysaccharide degradation. Natural substrates for most PULs remain undisclosed, with the exception of PUL 4 and 5 showing expression of fucoidan grown cells.

Our findings suggest that *M. forsetii* employs a diverse set of glycan-degrading enzymes, some encoded within PULs and others dispersed throughout the genome, enabling the bacterium to utilize various polysaccharides. This study enhances our understanding of the metabolic capabilities of *M. forsetii* and highlights its potential role in marine polysaccharide degradation.

4.2 Introduction

The genus *Maribacter* has been first described in 2004 as part of the family *Flavobacteriaceae* (Bernardet *et al.*, 1996; Bernardet *et al.*, 2002; Nedashkovskaya *et al.*, 2004). The type strain *Maribacter forsetii* (DSM 18668), was isolated from surface sea water in August 1999 at the Kabeltonne, also known as the long-term ecological research station Helgoland Roads (Helgoland, Germany). It has been characterized using the published genome (Barbeyron *et al.*, 2008). The genus *Maribacter* has been rarely isolated from seawater (Heins and Harder, 2023; Lu *et al.*, 2023) and community analyses based on metagenomics showed very low

presence in free-living bacterial populations (Sidhu *et al.*, 2023; Wang *et al.*, 2024a). In oxic surface layers of sandy sediments abundances of up to 4% were detected (Probandt *et al.*, 2018; Miksch *et al.*, 2021) and even higher abundance in the phycosphere populations of micro and macroalgae (Heins *et al.*, 2021; Lu *et al.*, 2023). *Maribacter forsetii* was included in an *in-silico* study of 53 flavobacterial genomes, free-living and particle associated bacteria, which analysed the PULs to identify their specificity towards major phytoplankton glycans (Kappelmann *et al.*, 2019). *M. forsetii* has been predicted to only have a few PULs, such as alginate, starch, mannose-rich substrates, fucose containing sulfated polysaccharides (FCSP) and digeneaside.

In the marine environment a variety of polysaccharides become available for microorganisms when algae are lysed by zooplankton or viral infection, either in the dissolved organic matter (DOM) or particulate organic matter (POM) fraction. In plankton, living algae are part of the POM. Other major POM fractions are faecal pellets from zooplankton and aggregated dead algae. Whereas living algae constantly exudate low molecular weight molecules and may lose parts of their cell walls, faecal pellets and aggregates of dead algae represent a major carbon source that is only once provided. Dead particles are usually mineralized within ten days (Omand *et al.*, 2020). The niche differentiation of particle-associated bacteria towards living microalgae or dead organic matter is little explored.

Over the years the total monosaccharide composition of the North Sea, more specifically of the phytoplankton found in the North Sea, has been analysed many times. Glucose is the most abundant sugar molecule, followed by arabinose, mannose, galactose, xylose, rhamnose, galacturonic acid and fucose (Ittekkot *et al.*, 1982; Biersmith and Benner, 1998; Aluwihare and Repeta, 1999; Urbani *et al.*, 2005; Alderkamp *et al.*, 2007; Scholz and Liebezeit, 2013; Huang *et al.*, 2021; Dlugosch *et al.*, 2023). These monosaccharides are building blocks of polysaccharides, like the abundant β-glucans laminarin, cellulose and xylan as well as alginate, arabinogalactan, carrageenans, chondroitin sulphate, fucoidan and mannan. In a recent study several polysaccharides were detected in the high molecular weight dissolved organic matter (HMWDOM) and POM fraction using monoclonal antibodies (Vidal-Melgosa *et al.,* 2021). It was shown that fucose containing sulfated polysaccharides (FCSP), arabinogalactan and β-1,4 mannan

were present in the HMWDOM and POM fraction during the spring bloom in the North Sea. Laminarin (β-1,3-glucan) and α-1,4-mannan were only detected in the POM fraction for the entire spring bloom. Those polysaccharides are stored in algae as food storage compounds and structural material. Macroalgae are synthesizing laminarin, fucoidans and alginates (brown algae), xylans (green algae) and carrageenans and galactans (red algae). The carbohydrates of microalgae are less characterized (Patel *et al.*, 2023). In addition, the chemical structure and species specificity is for most microalgal polysaccharides not solved. Different species of algae have different decorations on the polysaccharides allowing utilization only by certain microorganisms, which encode the specific enzymes.

For efficient utilization of polysaccharides, PULs, operon like genetic regions, encode many of the enzymes required for the degradation and transport of polysaccharides. A PUL was first described for starch utilization in *Bacteroides thetaiotaomicron (Shipman et al., 2000)*. The hydrolysis of polysaccharides is accomplished by glycoside hydrolases (GH), glycosyl transferase (GT), polysaccharides lyases (PL) and carbohydrate esterase (CE), in a high specificity, occasionally assisted by carbohydrate binding modules (CBM). These five groups of enzymes are summarized as carbohydrate active enzymes (CAZymes) (Bäumgen *et al.*, 2021; Drula *et al.*, 2022). The transport of the extracellular hydrolysed polysaccharides through the outer membrane, is carried out via the SusCD transport system in connection with the ExbB/D-Ton-B system in the cytoplasmic membrane, which provides the energy, generated by proton gradient, to open the β-channel of the SusC, so that the oligosaccharides can be transported into the periplasm (Noinaj *et al.*, 2010).

Some degradation pathways for marine polysaccharides have been studied, but predominantly in free-living bacteria, *Formosa* B, *Pseudoalteromonas* strains and *Verrucomicrobia* (Unfried *et al.*, 2018; Hettle *et al.*, 2019; Sichert *et al.*, 2020). Others like arabinogalactan and chondroitin sulfate have only been analysed in regards to the human gut microbiota (Luis *et al.*, 2018; Ndeh *et al.*, 2018; Ndeh *et al.*, 2020; Rawat *et al.*, 2022), whereby for arabinogalactan recently a study showed a utilization pathway in *Maribacter* sp. MAR_2009_72 (Kalenborn *et al.*, 2024).

In this study, we explored the utilization of a range of glycans by *Maribacter forsetii* and identified proteins that were expressed. The genes of these proteins were expected to be placed in PULs, allowing us to assign substrates to the PULs.

4.3 Material and Methods

4.3.1 Growth experiments

Maribacter forsetii DSM 18668^{T} was reactivated from glycerol stocks maintained in the laboratory since the isolation of the strains (Barbeyron *et al.*, 2008). The strain was grown in liquid HaHa_100V medium with 0.3 g/L casamino acids as sole carbon source, limiting the growth to an optical density (OD) at 600 nm below 0.2 (Hahnke *et al.*, 2015). The addition of 2 g/L of a carbohydrate source enabled growth beyond OD of 0.3. Here we added arabinose, glucose, galactose, galacturonic acid, laminarin from *Laminaria digitata*, κ-carrageenan, λ-carrageenan, mannose, xylose (Sigma Aldrich/Merck KGaA, Darmstadt, Germany), rhamnose (Fluka/Sigma Aldrich/Merck KGaA, Darmstadt, Germany), arabinogalactan, pullulan (The Dairy School, Auchincruive, Scotland), chondroitin sulfate (BIOZOL Diagnostica Vertrieb GmbH, Eching, Germany) and fucoidan from *Lessonia negressens* (Carbosynth Limited, Compton, Berkshire, United Kingdom). As well we used algae biomass collected from the environment as carbon source. In 250 ml photometer-sidearm Erlenmeyer flasks, three 50 ml cultures were inoculated with 0.4% v/v of a pre-grown culture in the same medium and incubated with 110 rpm at room temperature. In addition to three cultures for proteomic analyses, a fourth culture was used to monitor bacterial growth by measuring the optical density (OD) over the time point of harvest. Cells were harvested at an OD of 0.15 for casamino acid cultures serving as control cultures and at an OD of 0.25 for sugar containing cultures. Cells were pelleted by centrifugation in 50 ml tubes with 3080x *g* for 30 min at 4°C. Pellets were resuspended in 1 ml medium and centrifuged in 1.5 ml tubes at 15870x *g* for 15 min at 4°C. The wet biomass was weighed and stored at -20°C. For the two cultures with algae biomass no long-term growth experiments were done. Therefore, these growth curves end at the point of harvest, which was roughly at an OD of 0.250.

4.3.2 Algae biomass preparation

The algae biomass was collected in May 2020 near List in the Wadden Sea behind around the island Sylt, Germany (55°00,683`N; 008°26,400`E). The biomass was collected using a mantra trawl, which was dragged through the surface water for roughly 8 minutes at an approximate speed of 1 m/s. The manta trawl has two nets, the outer net has a mesh size of 200 µm, providing stability, and the inner net a mesh of 80 µm. The net bucket at the end had a side window covered with a sieve gauze of 20 µm. The collected biomass was filtered through a 3 mm steel sieve and then concentrated in a small net with a mesh size of 80 µm to reduce the amount of liquid, concentrating the biomass (Supplement Figure 1). Afterwards the concentrated biomass was filled into containers and frozen at -20°C until further usage. For our experiments we defrosted some biomass in 50 ml portions and performed freeze and thaw five times. The thawing took place at 4°C. For cell disruption, a Potter-Elvehjem-Homogenizator was used to grind the biomass into smaller pieces. Roughly a 100 µm wide space was between the pestle and tube. The sample was grinded with an electronic motor at 600 rpm and the glass was moved up and down three times. The samples as well as the glass tube were kept on ice in between usage. Small samples were taken, 500 µl before and after the grinding, to see under the microscope whether the sample cells were disrupted (Supplement Figure 2). After the grinding the samples were centrifugated for 30 min at 3075x g 4°C. The supernatant was removed and a microscopic sample was taken to demonstrate the absence of algae in the supernatant. The pellet, broken cell material, was weight, resuspended in 10 ml sterile ASW (Winkelmann and Harder, 2009) and frozen until the next step. To remove soluble proteins, dialysis was performed with a 100 000 Da cut off membrane by slow stirring for 1 hour at 12 °C against artificial seawater plus freshly provided 1 mM dithiothreitol (DTT) (10 ml dialysis sample/100 ml ASW). ASW + 1 mM DTT was replaced after one hour by fresh solutions. In total, four dialyses of 1 hour were performed. This was followed by a second centrifugation step for 15 min at 3075x g 4°C. As before, the supernatant was removed, the pellet was weight and dissolved in 10 ml ASW and frozen until it was used. Before using the algae biomass for the growth experiment,

one tube was defrosted and half was transferred into a second tube, which was place into a 60°C water bath for 2 hours to inactivate proteins.

4.3.3 Protein preparation and mass spectrometry

Proteins were extracted from cells using a bead-beating method (Schultz *et al.*, 2020). A pellet of 20 to 200 mg wet weight was disintegrated with 0.25 ml glass beads in 500 µl lysis buffer. Protein was quantified using the Roti-Nanoquant assay (Carl Roth, Karlsruhe, Germany). For protein purification on denaturing polyacrylamide gels (SDS-PAGE), 50 µg of protein was mixed with 10 µl of 4x SDS buffer (20% glycerol, 100 mM Tris/HCl, 10% (w/v) SDS, 5% β-mercaptoethanol, 0.8% bromphenol blue, pH 6.8) and loaded on Tris-glycine-extended (TGX) precast 4-20% gels (Biorad, Neuried, Germany). Electrophoresis was performed for 8 min at 150 V. The gel was fixed in 10% v/v acetic acid and 40% v/v ethanol for 30 min, stained with Brilliant Blue G250 Coomassie and the protein band was excised. The proteins were dissected in one gel piece and the pieces were washed with 50 mM ammonium bicarbonate in 30% v/v acetonitrile. Evaporation in a SpeedVac (Eppendorf, Hamburg, Germany) yielded dry gel pieces which were reswollen with 2 ng/µl trypsin (sequencing grade trypsin, Promega, USA). After 15 min incubation at room temperature the excess liquid was removed. Samples were digested overnight at 37°C. The gel pieces were then covered with MS water and peptides were eluted with ultrasonication. The peptides were desalted using Pierce™ C18 Spin Tips (Thermofisher, Schwerte, Germany) according to the manufacturer's guidelines. The elutes were dried in the SpeedVac and then stored at -20°C. For the MS analysis the samples were thawed and taken up in 10 µl of Buffer A (99.9% acetonitrile + 0.1% acetic acid).

Tryptic peptides of *Maribacter forsetii* were analyzed on an EASYnLC 1200 coupled to a Q Exactive HF mass spectrometer (Thermo Fisher Scientific, Waltham, USA). Peptides were loaded onto a self-packed analytical column (3 µm C18 particles; Dr. Maisch GmbH, Ammerbuch, Germany) using buffer A (0.1% acetic acid) with a flow rate of 2 µL/min and separated using an 85 min binary gradient from 4 to 50% buffer B (0.1% acetic acid in acetonitrile) and a flow rate of 300 nL/min. Samples were measured in parallel mode; survey scans in the Orbitrap were recorded with a

resolution of 60,000 with m/z of 333-1650. The 15 most intense peaks per scan were selected for fragmentation. Precursor ions were dynamically excluded from fragmentation for 30 sec. Single-charged ions as well as ions with unknown charge state were rejected. Internal lock mass calibration was applied (lock mass 445.12003 Da).

The mass spectrometry files were analyzed in MaxQuant version 2.4.9.0 in the standard settings against the strain specific protein database downloaded from NCBI (see below) (Tyanova *et al.*, 2016; Sayers *et al.*, 2022) and common laboratory contaminants. Statistical analysis was performed in Perseus version 2.0.11.0 (Tyanova and Cox, 2018). Proteins were recognized as being expressed when they had label free quantification intensities in one out of three biological replicates.

Reference genomes were downloaded from NCBI: *Maribacter forsetii* NZ_JQLH01000001.1 (download 2023/10).

4.3.4 Bioinformatic analyses

Protein annotation was refined using several of databases. CAZymes were considered to be identified if two out of three search algorithms in dbCAN3 were positive in the web interface search (Zheng *et al.*, 2023). The conserved domain database (CDD) (Lu *et al.*, 2019), the SulfAtlas web interface (Stam *et al.*, 2022), InterPro (Paysan-Lafosse *et al.*, 2022), PULDB (Terrapon *et al.*, 2018), deepTMHMM (Hallgren *et al.*, 2022), SignalP (Teufel *et al.*, 2022), Blastkoala (Kanehisa *et al.*, 2016) and UniProt (The UniProt, 2023) provided additional information.

Alignment for the SusC protein tree was performed using MAFFT online in automode (Katoh *et al.*, 2019). SusC proteins encoded in laminarin utilizing loci were retrieved from Krüger *et al.* (2019). The tree was calculated using the maximum likelihood method in Mega 11 with the default settings (Tamura *et al.*, 2021). For the visualization of the data the following programs and packages were used: R version 4.3.2 (R Core Team, 2023), ggplot2 (Wickham, 2016), gggenes (Wilkins, 2023) and Proksee (Grant *et al.*, 2023).

4.4 Results and Discussion

We tested growth of *Maribacter forsetii* under 17 different conditions in a liquid medium containing 0.3 g/L casamino acids as limiting carbon source. To this base medium we added 2 g/L of one of the following glycans: arabinogalactan, arabinose, chondroitin sulfate, fucoidan, galactose, galacturonic acid, glucose, κ-carrageenan, λ-carrageenan, laminarin, mannose, pullulan, rhamnose, xylose and two types of algae biomasses. When grown only on the basal medium *M. forsetii* grew to a maximum optical density of 0.181 with a growth rate of $\mu = 0.05$ h^{-1} (Figure 1, Table 1). The culture with xylose as C-source had the highest overall OD 0.673 with $\mu = 0.07$ h^{-1}. Some conditions (galacturonic acid and both algae biomass) required more time to reach the exponential growth phase; these cultures did not grow to a high OD. *M. forsetii* grown in presence of galacturonic was the only condition that did not grow higher than an OD of 0.1 above the control value.

Overall *M. forsetii* seems to be a versatile bacterium that can grow in the presence of many different saccharides (Figure 1). An *in-silico* study reported that *M. forsetii* does not have a PUL for laminarin, xylose-, rhamnose- and galactose-containing polysaccharides and carrageenan utilization in its genome (Kappelmann *et al.*, 2019). Furthermore none of these glycans were reported in the type strain description (Barbeyron *et al.*, 2008). Together this raised the question whether *M. forsetii* is able to grow and utilize these glycans. Here we showed that *M. forsetii* grows on these substrates and furthermore displays expressed enzymes in the proteomic data.
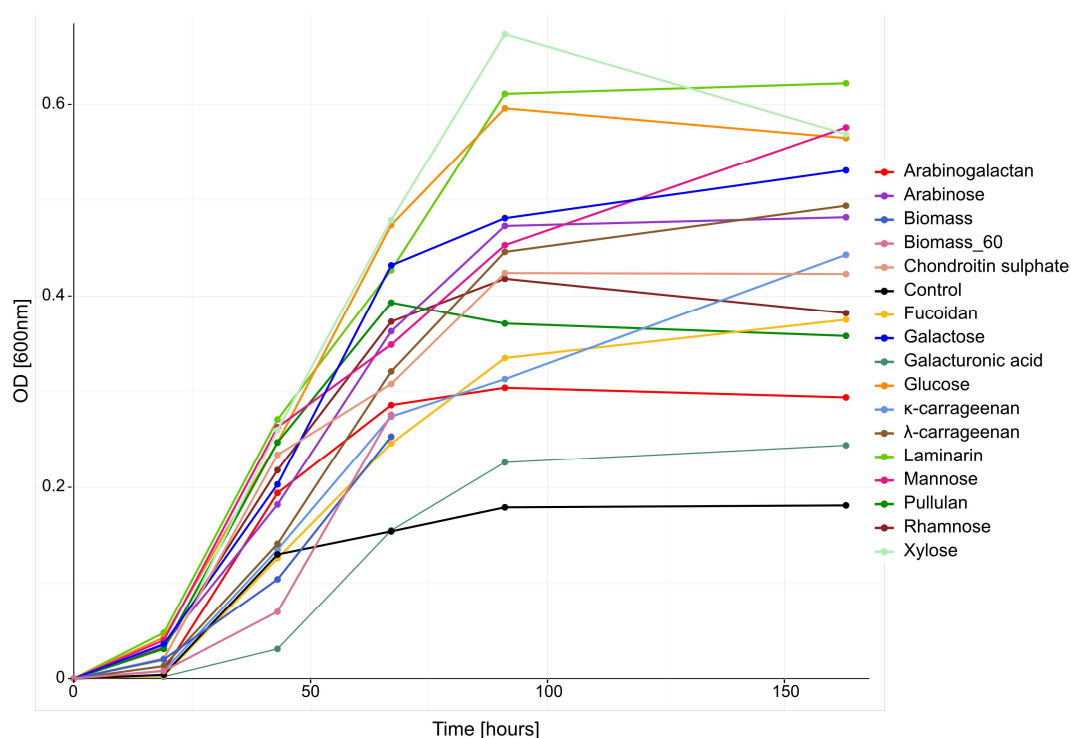
Figure 1: Growth curve of *Maribacter forsetii* in presence under 17 different conditions; arabinose, arabinogalactan, biomass from the North Sea, biomass from the North Sea heat inactivated at 60°C, chondroitin sulfate, control, fucoidan from *Lessonia negressens*, galactose, galacturonic acid, glucose, κ-carrageenan, λ-carrageenan, laminarin from *Laminaria digitata*, mannose, pullulan, rhamnose and xylose. *M. forsetii* was grown in 50 ml of modified HaHa_100V with 2 g/L of the respective carbon source at room temperature at 110 rpm. The OD was measured at 600 nm.

Biomass of all 17 growth conditions was analyzed through the proteomics pipeline and we identified between 800 to 1000 proteins for each condition, present in at least one of three biological replicates (Table 1). Only the algae biomass culture had no biological replicates and for arabinose two biological replicates were measured. The 17 conditions share 340 proteins. These proteins include many ribosomal proteins, as well as proteins involved in DNA metabolism. This group also comprises proteins involved in the glycolysis, citric acid cycle, pentose phosphate pathway and the oxidative phosphorylation (Supplement Table S1).

Table 1: Overview of substrates we grew *Maribacter forsetii* on, their respective growth rate, maximum optical density (OD) and number of identified proteins. *M. forsetii* was grown in 50 ml of modified HaHa_100V with 2 g/L of the respective carbon source at room temperature at 110 rpm. The OD was measured at 600 nm.

Substrate	Growth rate [μ]	Max OD [600nm]	No of proteins
Arabinose	0.04	0.482	982
Arabinogalactan	0.05	0.304	1048
Biomass	0.05	0.253	879
Biomass 60	0.07	0.276	883
Chondroitin sulphate	0.04	0.424	1136
Control	0.05	0.181	1081
Fucoidan	0.07	0.375	1005
Galactose	0.04	0.531	843
Galacturonic acid	0.07	0.244	1061
Glucose	0.04	0.596	1083
κ-carrageenan	0.05	0.443	947
λ-carrageenan	0.04	0.622	985
Laminarin	0.05	0.494	1066
Mannose	0.03	0.576	1099
Pullulan	0.03	0.358	1128
Rhamnose	0.05	0.418	1014
Xylose	0.07	0.673	985

To identify differences at large between the proteomic data of the various growth conditions we calculated a principal component analysis (PCA). The PCA displays differences between the proteomes based on the expression intensities (Figure 2). We observe two distinct groups. Group one includes laminarin, glucose, xylose, κ-carrageenan and λ-carrageenan. The laminarin and glucose samples are close to one another. The xylose replicates are a little further away from the rest of the group, a similar observation was made for both carrageenan's. One of the biological triplicates of λ-carrageenan is distant from the group showing that there are large differences even when grown under the same conditions. The other group consists of chondroitin sulfate, biomass, biomass_60, arabinose, arabinogalactan, fucoidan, control, rhamnose and pullulan. In this group we observed some subgroups, pullulan and chondroitin sulfate form one of them. Arabinogalactan, arabinose, rhamnose, galacturonic acid and two of the three fucoidan samples form the second subgroup. The third fucoidan samples is located far away from the other conditions in this subgroup. For the control triplicates two samples are next to each other the third is in the subgroup of pullulan and chondroitin sulfate. Two conditions form each

their own group. Galactose is situated next to the outlier of group one (λ-carrageenan) and the three biological replicates of mannose grown cells are scattered throughout the plot, showing the diversity of each of these three cultures. The two algae biomass samples are close to each other. Overall, we noticed that some proteomes like laminarin and glucose have similar expression patterns and other conditions for example galactose are very distinct from the rest.
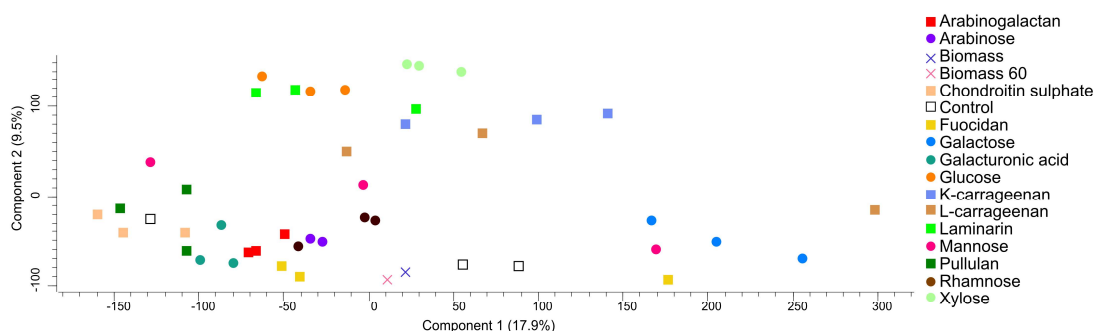
Figure 2: Principal component analysis (PCA) shows the differences between the expression intensities of all 17 proteomes *of Maribacter forsetii*.

Maribacter forsetii has a genome size of 4.51 Mb with 3890 genes in total (Figure 3). 119 proteins were identified as CAZymes by dbCAN3, with at least 2 out of three search algorithms. Furthermore *M. forsetii* has 12 sulfatases in its genome according to sulfatlas. Nineteen SusC/D pairs were identified, with all SusCs having the protein family (Pfam) ID TIGR04056. We annotated 6 PULs in the genome, each of them enclose at least one SusC/D pair as well as more than one CAZyme. In Figure 3 the distribution of the expressed proteins of each growth condition over the complete genome is displayed. It does not show the expression intensities just if the specific protein was expressed. On the first look we observed a lot of similarities between the conditions. For example, in all 17 conditions we detected only the first couple of proteins of PUL 1 were expressed. A similar remark can be made for PUL 5, here only the second half of the PUL was expressed. Reviewing the figure more closely shows distinct characteristics between conditions. One being that some proteins were only expressed under certain circumstances, for example on the left of the 2.5 Mb mark proteins only expressed in the algae biomass samples can be witnessed. In PUL 3 a gene was only expressed in fucoidan, mannose and chondroitin sulfate grown cells.

We discovered that most CAZymes and SusC/D pairs are not located in PULs in *M. forsetii*. For bacteria of the phylum *Bacteroidota* it has been mostly described that all enzymes involved in the degradation of glycans are organized in PULs (Lapébie *et al.*, 2019; McKee *et al.*, 2021). In 2021 it was mentioned that SusC/D does not always have to be located next to CAZymes (McKee *et al.*, 2021), furthermore in plant associated Bacteroidota they are considered to be part of the phosphate utilization system (Lidbury *et al.*, 2021). We noticed a low number of sulfatases (12) in comparison to the 119 CAZymes in the genome. Because sulfate is a frequent decoration of marine polysaccharides and sulfatases are as well implemented in PULs (Helbert, 2017; Arnosti *et al.*, 2021), we hypothesize, due to the moderate number of sulfatases in the genome of *M. forsetii*, that sulfated glycans are not a primary substrate.
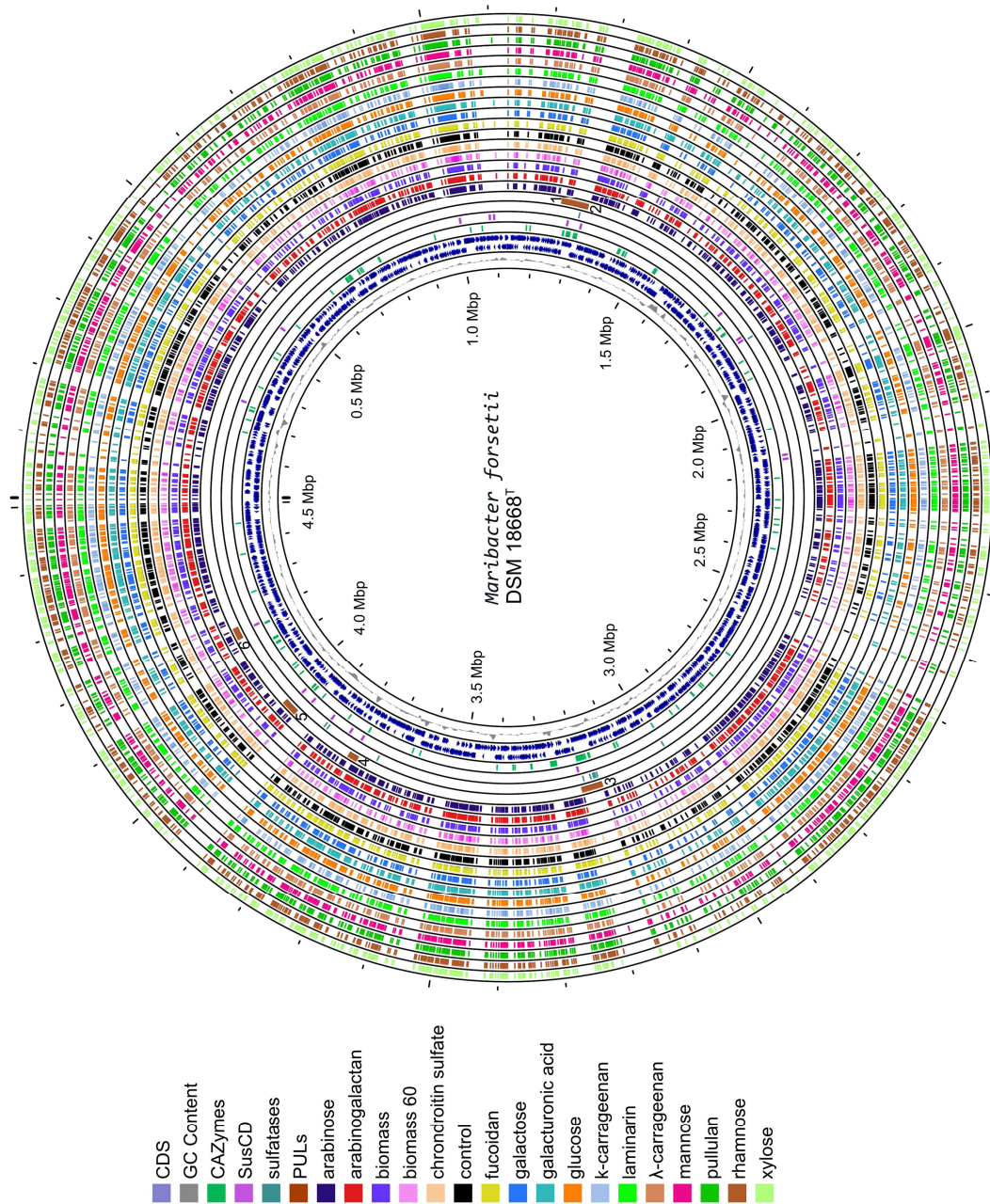
Figure 3: Full genome overview of *Maribacter forsetii* showcasing the GC content (ring one (most inner ring)), all annotated coding genes (CDS, ring 2 and 3) in forward and reverse direction, CAZymes identified by dbCAN3 (ring 4), SusC/D (ring 5), sulfatases (ring 6), and PULs (ring 7), arabinose proteome (ring 8), arabinogalactan proteome (ring 9), biomass from the North Sea proteome (ring 10), biomass from the North Sea heat inactivated at 60°C proteome (ring 11), chondroitin sulfate proteome (ring 12), control proteome (ring 13), fucoidan from *Lessonia negressens* proteome (ring 14), galactose proteome (ring 15), galacturonic acid proteome (ring 16), glucose proteome (ring 17), κ-carrageenan proteome (ring 18), laminarin from *Laminaria digitata* proteome (ring 19), λ-carrageenan proteome (ring 20), mannose proteome (ring 21), pullulan proteome (ring 22), rhamnose proteome (ring 23) and xylose proteome (ring 24).

To investigate the contrast of the conditions we looked at each of six PULs individually. Several proteins of PUL 1 were expressed in all conditions (Figure 3;

Supplement Figure 3). PUL 1 consists of 31 proteins, including one SusC/D pair and 7 CAZymes. Here no SusC/D pair and only one CAZymes, GH177 (P177_RS05435), a Gfo/Idh/MocA family oxidoreductase, was expressed in all conditions. Overall, the expressed genes do not display a difference in intensities. Alone P177_RS05395, a lipid binding SYLF-domain containing proteins, showed some change in the LFQ values. PUL 1 could potentially be specific for a mannan based polysaccharides with glucuronic acid as a possible decoration based on the CAZymes in the PUL GH92 (mannosidase), GH115 (glucuronidase) and GH88 (glucuronyl hydrolase) (Zhu *et al.*, 2010; Aalbers *et al.*, 2015; Dupoiron *et al.*, 2015; Reisky *et al.*, 2019; Zeugner *et al.*, 2021). Possible sugars could be xanthan gum or glucuronic mannan, which have mannose and glucuronic acid side chains (Williams *et al.*, 1991; Zeugner *et al.*, 2021).

The second PUL includes 9 proteins with one SusC/D pair and 3 CAZymes.

PUL 2 showed no expression data in this study (Supplement Figure 4). Both CAZymes classes present in PUL 2, GH3 (P177_RS05510) and GH144 (P177_RS05520 and P177_RS05525), are annotated in as β-glucosidases. A possible target substrate could be a sulfated β-glucan, due to the presence of a sulfatase in the PUL. Analyzing the PUL in PULDB identified one literature derived PUL for marine organisms, *Zobellia galactivorans* DsijT literature derived PUL 43 (Barbeyron *et al.*, 2016). Moreover, this PUL was found in many marine *Flavobacteriia*, it was present in all *Maribacter* strains in PULDB. Unfortunately, there is no indication of the targeted polysaccharide yet, just the assumption that it must be sulfated, due to the sulfatase S1_11 in the PUL (Lapébie *et al.*, 2019).

PUL 3 contains the most CAZymes (11) and sulfatases (6) (Supplement Figure 5). Overall, it is comprised out 31 proteins including one SusC/D pair. The SusC/D pair is only expressed as a pair in fucoidan grown cells. SusD alone was slightly expressed in the arabinogalactan, arabinose, chondroitin sulfate, galacturonic acid, κ-carrageenan, laminarin, pullulan and rhamnose proteomes. The enzyme annotated as GH29 (P177_RS), α-L-fucosidase, was only expressed in chondroitin sulfate, mannose and fucoidan, whereby the measured intensities were the highest for fucoidan grown cells. None of the sulfatases were expressed in this proteomic data set. PUL 3 has no match in the PULDB database which could have provided information on the substrate. Here we can only speculate about its specificity based

on the CAZymes and sulfatases. The only hint in the proteomic data is that the SusC/D pair was expressed in fucoidan grown cells. Fucoidan might be a possible substrate as GH29, GH95 and GH117 have been described in context with fucoidan utilization (Rebuffet *et al.*, 2011; Grootaert *et al.*, 2020; Shuoker *et al.*, 2023). The other glycosyl hydrolases GH28, GH88 and GH97 could potentially remove decorations of the fucoidan structure, such as arabinose, galactose, galacturonic acid, glucuronic acid and rhamnose (Kitamura *et al.*, 2008; Luis *et al.*, 2018; Reisky *et al.*, 2019).

The other three GH families in PUL 3, GH28, 88 and 97, seem more likely to remove decorations on side chains, such as glucuronic acid, rhamnose or galactose (Kitamura *et al.*, 2008; Luis *et al.*, 2018; Reisky *et al.*, 2019). PL40 has been characterized as an ulvan lyase (Reisky *et al.*, 2019), which might as well be the substrate for the PUL. The sulfatases in PUL 3 have been attached to fucan, mucin and ulvan utilization, whereby for S1_23 and S1_28 the substrate is unknown (Stam *et al.*, 2022). Based on all of this we hypothesize that the substrate of PUL 3 most likely is a fucoidan, that structurally differs from the fucoidan used in this study or ulvan, based on the CAZymes.

PUL 4 in the *M. forsetii* genome encodes 19 proteins, 2 being CAZymes and 1 SusC/D pair (Supplement Figure 6). Here only the SusC was highly expressed in fucoidan grown cells. P177_RS16040, a gliding motility protein was expressed in all conditions. The same is true for a type IX secretion protein (P177_RS16035). This PUL encodes ExbB and ExbD. ExbB (P177_RS16115) was only expressed in arabinogalactan and biomass grown cells, whereby for ExbD LFQ intensities were measured for all conditions, beside the two algae biomass samples. One of the two encoded PL7 (P177_RS16065) was expressed in the fucoidan proteome.

PUL 4 has multiple matches in PULDB, mostly *Maribacter* strains; *Maribacter* sp. Hel_I_7, *Maribacter dokdonensis* MAR_2009_60 and MAR_2009_71, *Maribacter aquivivus (Nedashkovskaya et al., 2004)* and others including several *Polaribacter* strains (Gosink *et al.*, 1998) and *Flagellimonas* sp. CMM7 (Shin *et al.*, 2022). Unfortunately, none has been studied and described in the literature. Based on the proteomic data and the substrate we used we assume that some type of fucoidan is the substrate, but PL7 has been mainly described for alginate degradation (Thomas *et al.*, 2013). Most likely is that the fucoidan is not as pure expected and

some alginate is present, which potentially has led to the expression of one PL7 and the SusC/D in PUL 4.

A similar conclusion can be drawn for PUL 5. PUL 5 encodes 27 proteins including 1 SusC/D pairs and 4 CAZymes (Supplement Figure 7). The SusC/D pair was expressed (P177_RS168450 and P177_RS16835), whereby only the SusC was expressed in all circumstances. The highest expression was measured for fucoidan grown cells. Three out of the six CAZymes were expressed. The PL6_1, (P177_RS16810) was expressed in arabinogalactan, arabinose, chondroitin sulfate, galactose, galacturonic acid, laminarin, mannose, pullulan, rhamnose, xylose and the highest intensities were measured for fucoidan. A similar observation was made for P177_RS16815, PL17_2, but this enzyme was overall expressed in five conditions. The enzymes annotated as PL7_5, (P177_RS16785 and P177_RS09855), had high intensities in fucoidan and low ones measured for pullulan grown cells. Besides the CAZymes, more proteins were expressed. Some showed a differentiation between the different conditions (P177_RS16765 (peroxiredoxin-like family protein), P177_RS16770 (RraA), P177_RS16795 (Gfo/Idh/MocA family oxidoreductase), P177_RS16845 (FadR/GntR family transcriptional regulator) and P177_RS16885 (response regulator)) and others were expressed under all circumstances P177_RS16855 (SDR family NAD(P)-dependent oxidoreductase), P177_RS16865 (sugar kinase) and P177_RS16875 (bifunctional 4-hydroxy-2-oxoglutarate aldolase/2-dehydro-3-deoxy-phosphogluconate aldolase)).

All four CAZymes, two PL6_1, PL7_5 and PL17_2 have been characterized as alginate lyases (Thomas *et al.*, 2013; Mathieu *et al.*, 2016; Jouanneau *et al.*, 2021). The proteomic data shows a strong expression of PUL 5 in fucoidan grown cells. Beside the CAZymes we observed the highest expression of a SDR oxidoreductase, sugar kinase and a bifunctional 4-hydroxy-2-oxoglutarate aldolase for fucoidan. According to the supplier the fucoidan of *Lessonia nigrescens* consists of approximatively 26.2% fucose, 29.1% sulfate, 13% galactose and uronic acid, leaving 19.2% of the structure unknown. We assume that *M. forsetii* is not capable of degrading the compound fucoidan, due to its lack of microcompartments for fucose degradation, which protects the cell from the toxic intermediate compound lactaldehyde (Sichert *et al.*, 2020). One theory is that *M. forsetii* can hydrolyze some

bonds of the side chains of fucoidan and degrade these further to pass into glycolysis, citric acid cycle and oxidative phosphorylation, as well as the pentose phosphate pathway, which are all expressed when grown in presence of fucoidan. Secondly, we assume that the here used fucoidan is not completely pure. We base this hypothesis on the fact that mainly alginate utilizing enzymes were expressed in fucoidan grown cells. Therefore, we rather assume that *M. forsetii* has the potential to degrade alginate.

The last PUL in the genome of *Maribacter forsetii*, encodes 1 SusC/D pair and 2 CAZymes in a range of 17 proteins (Supplement Figure 8). The SusC/D pair showed expression data for all conditions. The expressed CAZyme, a GH65 (P177_RS17740), showed the same intensity for all samples. A vanadium dependent haloperoxidase (PAP2, P177_RS17780), a β-phosphoglucomutase (pgmB, P177_RS17745) and a glycerol phosphodiester phosphoesterase (PI-PLC, P177_RS17730) were expressed in all conditions. P177_RS17770, a VCBS repeat protein was expressed in chondroitin sulfate, control, galacturonic acid, κ-carrageenan, laminarin, mannose, pullulan and xylose.

PUL 6 has some homologues in PULDB for example in *Maribacter* sp. MAR_2009_72 and Hel_I_7 as well as in *Arenibacter* sp. MAR_2009_79 and *Zobellia galactivorans* DsijT. The GH65 is characterized as a maltose phosphorylase by dbCAN. This enzyme catalyzes maltose and phosphate to glucose and β-D-glucose-1-phosphate and the latter can directly be transferred into β-D-glucose-6-phosphate by P177_RS17745 a β-phosphoglucomutase, which is a step in the pentose phosphate pathway. The aforementioned homologue in *Zobellia galactivorans* DsijT has been described as a maltose PUL (Barbeyron *et al.*, 2016). However, based on our proteomic data this PUL could potentially contribute to the degradation of various glycans.

Not all annotated CAZymes of *M. forsetii* are encoded in PULs. Most of the CAZymes in *Maribacter forsetii* are not located in PULs, out of 119 CAZymes 85 are distributed throughout the whole gnome and not in close proximity to a SusC/D pair. Of these 85 CAZymes 32 were expressed in at least one of the 17 growth conditions (Figure 4). Some of these CAZymes protrude because only in some conditions LFQ intensities were measured. PL7_5 (P177_RS09855) annotated as glucuronate lyase was expressed in 14 conditions, not for the algae biomass and

galactose. Four GT4 proteins were expressed in the data set. P177_RS02870 seems to be very specific towards glycans consisting of glucose since the only intensities measured for glucose and laminarin grown *Maribacter forsetii* cells. The other three GT4 (P177_RS07530, P177_RS02860 and P177_RS01005) have wider range of carbon sources that activate their expression. Overall, the GT family 4 has a wide range of functions, here we were not able to identify their specific target substrate . A similar perception can be gathered looking at the two expressed GT2 enzymes. One (P177_RS06030) was expressed in all conditions and the other (P177_RS06535) only in cells grown on chondroitin sulfate, laminarin, mannose and rhamnose. One of the two GH74, either a glucanase or cellobiohydrolase, (P177_RS01685) showed no expression data for the algae biomass and galactose samples. P177_RS01680 (GH74) had LFQ intensities measured for all growth conditions. A GH32, mostly described as a fructotransferase, (P177_RS17450) was expressed in arabinogalactan, arabinose, chondroitin sulfate, control, fucoidan, galacturonic acid, κ-carrageenan, pullulan and xylose grown cells. The protein annotated as GH2 (P177_RS09656) showed expression in the arabinose, biomass, biomass 60, galactose, glucose and rhamnose proteomes. The second GH2, P177_RS12820, was expressed in all conditions, confirming the wide substrate range of this GH family. The other expressed CAZymes are mostly expressed in all conditions, some not in three or four, so these all seem to have a wide range of functionality. One that protrudes a little from that group is the GH16_3. It is expressed in 15 conditions, but it shows a high expression for laminarin grown cells and only low expression intensities for the rest. Still some specificity can be seen in this case towards laminarin, since the GH16_3 is annotated as a laminarinase.
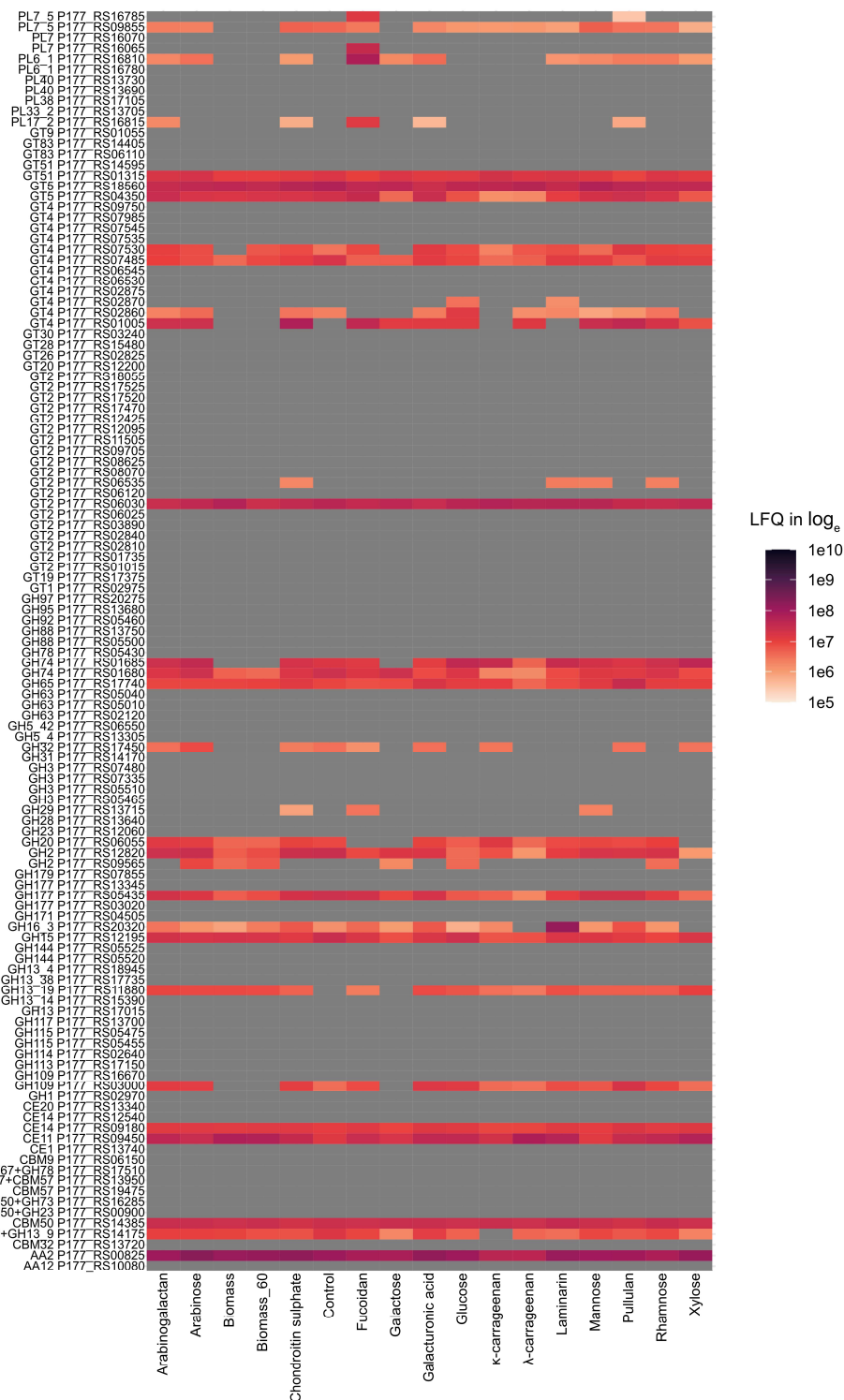
Figure 4: Heatmap of CAZymes of *Maribacter forsetii* grown in the presence of arabinogalactan, arabinose, biomass from the North Sea, biomass from the North Sea heat inactivated at 60°C, chondroitin sulfate, control, fucoidan from *Lessonia negressens*, galactose, galacturonic acid, glucose, κ-carrageenan, λ-carrageenan, laminarin from *Laminaria digitata*, mannose, pullulan, rhamnose and xylose. Expression intensities in the plot are mean values of three biological replicates of each condition shown in LFQ values [log10], for arabinose mean values of two biological replicates and both biomass proteomes do not have replicates. Grey means the protein was not dected.

Similar to the CAZymes not all SusC/D pairs are located in PULs, 11 out of 19 are not in PULs. As well important to notice is that no SusC or SusD was found in the genome without the other as a direct neighbor. Of those nine pairs not in PULs one shows a specificity towards certain C-sources (Figure 5). The SusC/D pair (P177_RS04595/04600) was expressed in glucose, xylose and laminarin grown cells, with highest intensities for laminarin. All the other do not show a specificity towards a carbon source by just looking at which conditions had expression data. For example, SusC/D pair (P177_RS01995/02000) was expressed in almost all conditions, not in λ-carrageenan and xylose. But here we noticed that for arabinogalactan and galactose the intensities were by at least a 10-fold higher than for the rest. A comparable observation was made for P177_RS08935/08940. Here the SusC was expressed in more conditions than the SusD, but still both were expressed in 10 conditions. Four of those show higher intensities glucose, κ-carrageenan, laminarin and xylose. The SusC/D pair, P177_R07645/07650, presents a similar behavior. The other three expressed pair show no specificity towards any C-source and display a mostly identical intensity between the different samples. One of the SusCs (P177_RS16435) was expressed without its neighboring SusD. It showed LFQ values for fucoidan and pullulan grown cells.
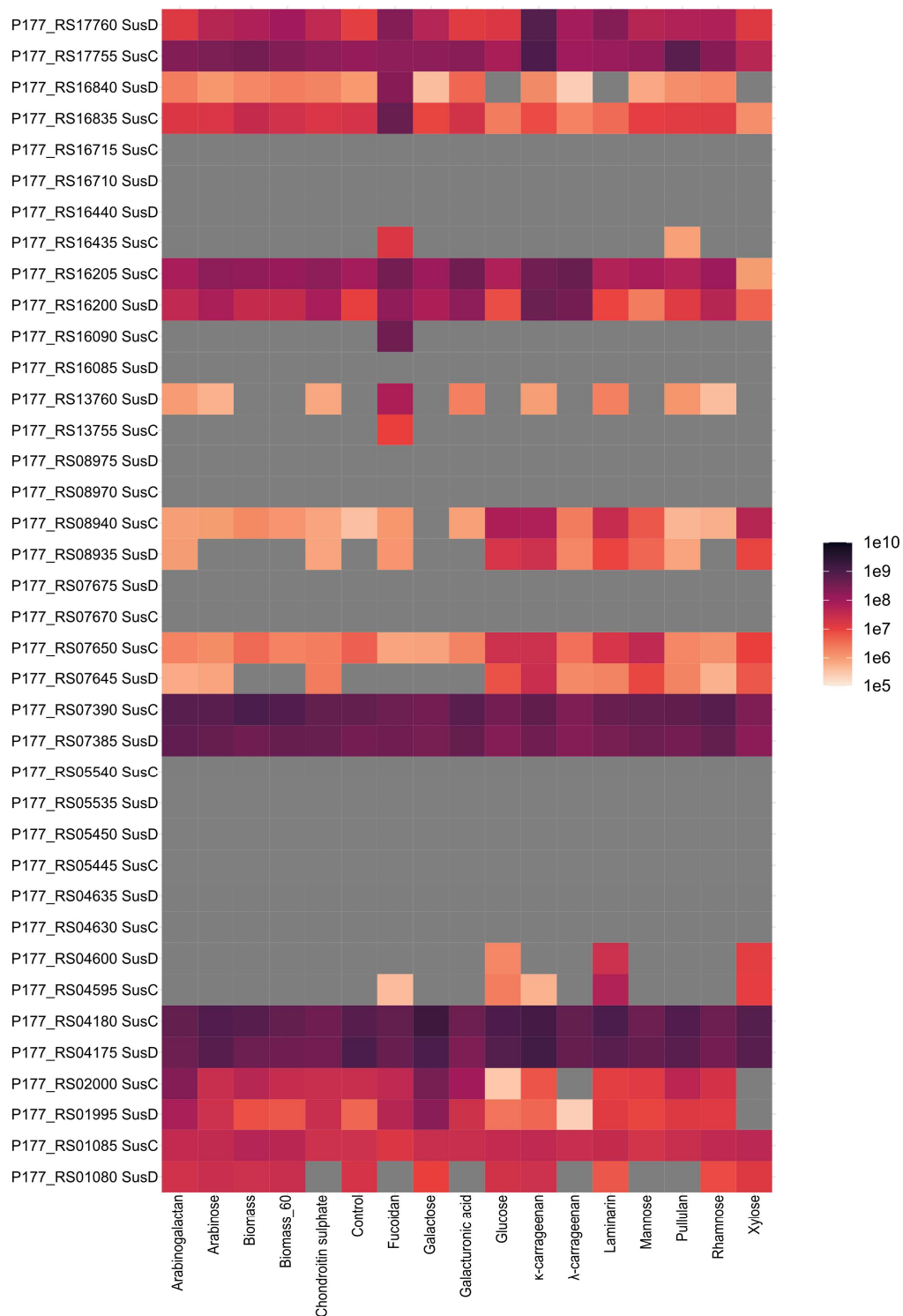
Figure 5: Heatmap of SusC/D pairs of *Maribacter forsetii* grown in the presence of arabinogalactan, arabinose, biomass from the North Sea, biomass from the North Sea heat inactivated at 60°C, chondroitin sulfate, control, fucoidan from *Lessonia negressens*, galactose, galacturonic acid, glucose, κ-carrageenan, λ-carrageenan, laminarin from *Laminaria digitata*, mannose, pullulan, rhamnose and xylose. Expression intensities in the plot are mean values of three biological replicates of each condition shown in LFQ values [log10], for arabinose mean values of two biological replicates and both biomass proteomes do not have replicates. Grey means the protein was not detected.

In previous studies SusCs were aligned in a phylogenetic tree to present how close SusCs with the same specificity towards a certain C-source are related. Here we used SusC sequences, close neighbors to the *M. forsetii* SusCs, from the Krüger *et al.* (2019) data set, which encloses SusCs from different *Bacteroidota* and included all SusCs of *Maribacter forsetii* (Figure 6). We observed that some SusCs of *M. forsetii* are closest to other *Maribacter* strains SusCs. The tree confirmed certain specificities we saw in the proteomic data. For example, P177_RS013755 (WP_051941850.1) is located on a branch that only holds FCSP specific SusCs, which coincides with our expression data. For some SusCs we can get an indication towards their specificity looking at the phylogenetic tree. P177_RS16090 was only expressed in fucoidan, in the tree it closely located to SusCs from other *Maribacter* strains that are annotated in regards to alginate degradation. This aligns with our theory that the used fucoidan was contaminated with alginate. SusC, P177_RS17755, did not show a specificity in the proteomic data, but in the tree both closest related SusCs are specific for α-glucans. For P177_RS16835 we saw expression in all conditions, its location in the tree suggest an alginate specificity. The other SusCs showed no distinct glycan affiliation in their expression data, and even with the tree no precise substrate could be defined, since their closest related SusCs were annotated for more than one glycan.
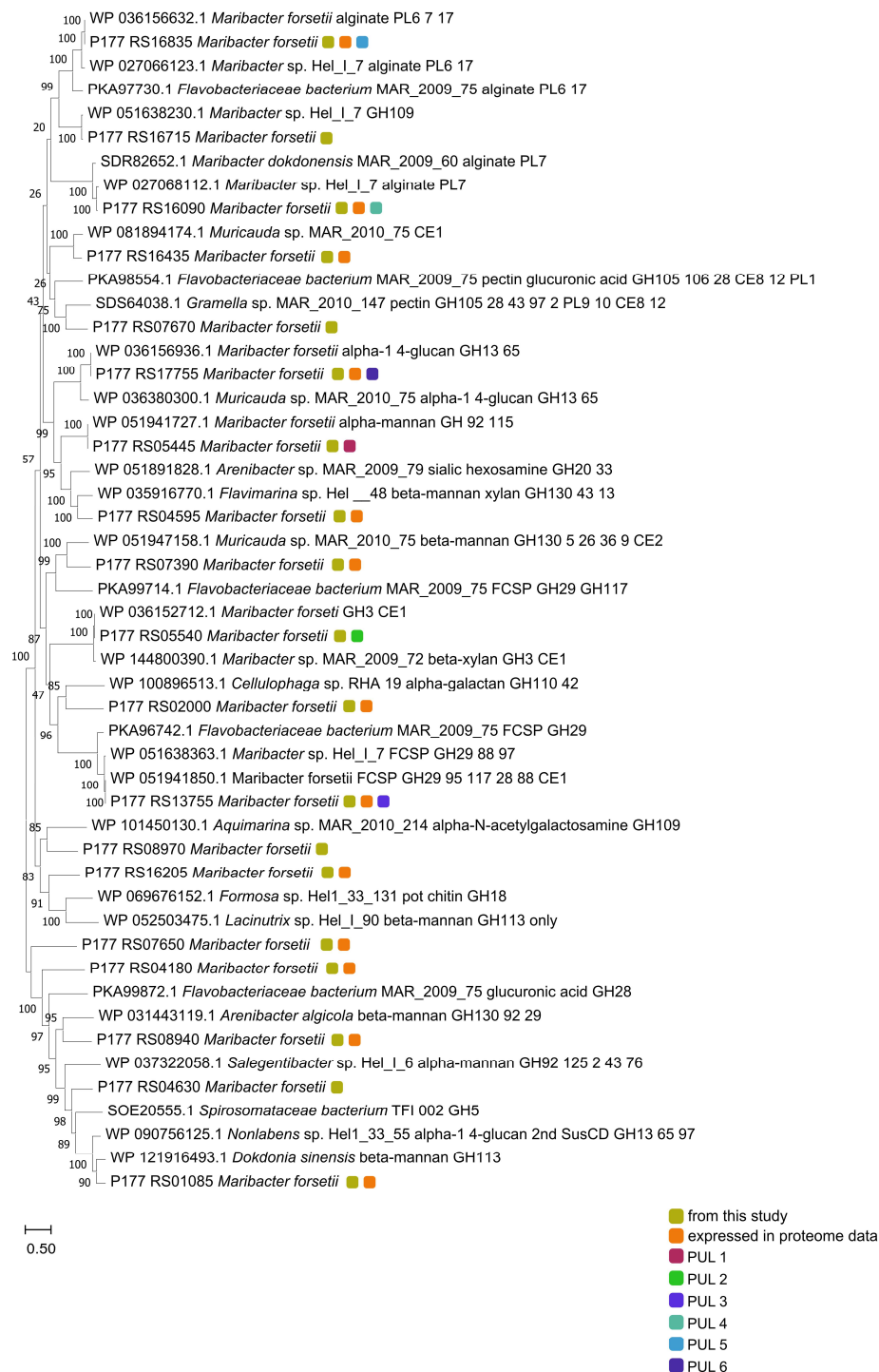
Figure 6: Phylogenetic tree showing the likeness between SusC sequences, close neighbors to the *M. forsetii* SusCs, from the Krüger *et al.* (2019) data set, which encloses SusCs from different *Bacteroidota* and all SusCs of *Maribacter forsetii*. The evolutionary history was inferred by using the Maximum Likelihood method and JTT matrix-based model. The tree with the highest log likelihood (-84700.42) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the JTT model, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The strain nomenclature was used as provided in NCBI (Feb.2024).

Looking at the overall expression data, each conditions provided similar information as concluded from the PUL analysis. *M. forsetii* seems to be able to degrade certain parts of the provided polysaccharides but cannot degrade the complete structure. The arabinogalactan utilization was recently described in a different *Maribacter* strain, *Maribacter* sp. MAR_2009_72 (Kalenborn *et al.*, 2024), but *M. forsetii* does not have any CAZymes for arabinogalactan encoded in the genome. We know that *M. forsetii* is able to grow in the presence of arabinose and galactose and moreover uses these to feed its glycolysis or pentose phosphate pathway. We hypothesize here that *M. forsetii* can cleave bonds of possible arabinogalactan side chain decorations such as glucuronic acid, rhamnose, xylose and mannose, using expressed CAZymes such as GH2, GH65 and GH74, (Fujita *et al.*, 2019; Villa-Rivera *et al.*, 2021; Leszczuk *et al.*, 2023), and utilize these.

Chondroitin sulfate degradation has been described to involve GH88, PL29, PL8 and PL33 (Ndeh *et al.*, 2018; Rawat *et al.*, 2022; Wang *et al.*, 2024b) These CAZymes are not encoded in *M. forsetii*, besides the GH88, which was not expressed during growth on chondroitin sulfate. Here as for fucoidan mostly CAZymes described for alginate, PL6_1, PL17_2, PL7 and PL7_5, have been expressed, stipulating the assumption that the used chondroitin sulfate most likely is not pure and has an alginate contamination.

In the genome of *M. forsetii* no typical CAZymes for carrageenan utilization are encoded besides the GH16, which is annotated further as GH16_3 and characterized as a laminarinase and as such described for *M. forsetii* in a recent study (Kalenborn *et al.*, unpublished/Chapter 2). In addition to the non-matching annotation, the GH16_3 was not expressed for λ-carrageenan grown cells and only low intensities were measured for κ-carrageenan. To identify a possible carrageenase we used the amino acid sequence of *Pseudoalteromonas carrageenovora* 9[T] and ran a search against the amino acid sequence of *M. forsetii*, but no homologue was recognized (Gobet *et al.*, 2018). Again, as aforementioned for arabinogalactan, chondroitin sulfate and fucoidan the theory is that *M. forsetii* can cleave certain bonds of the polysaccharide and transports these oligosaccharides into the cell, where they are then further degraded into monomers. Furthermore, we discovered that *Maribacter forsetii* is able to grow on a variety of glycans, showing how versatile the strain is. Moreover, we can see here again, as

previously described for the laminarin degradation of this bacterium (Kalenborn *et al*., unpublished/Chapter 2), that most CAZymes and SusC/D pairs are not located in PULs. As well we observed that some SusC/D pairs are substrate specific and others seem to transport various oligosaccharides across the outer membrane. We hypothesize that PUL 4 and 5 are alginate associated PULs. The four other PULs and most CAZymes were mostly not expressed in this proteomic data even though 17 different saccharides were included. Therefore, we conclude that the main substrates of *Maribacter forsetii* was not included in this study. Moreover, we hypothesize that not all marine polysaccharides have been identified, especially ones from microalgae.

4.5 Acknowledgement

4.6 Supplement

4.6.1 Supplement Tables

Supplement Table are available at Edmond: https://doi.org/10.17617/3.HCEGVG

4.6.2 Supplement Figures
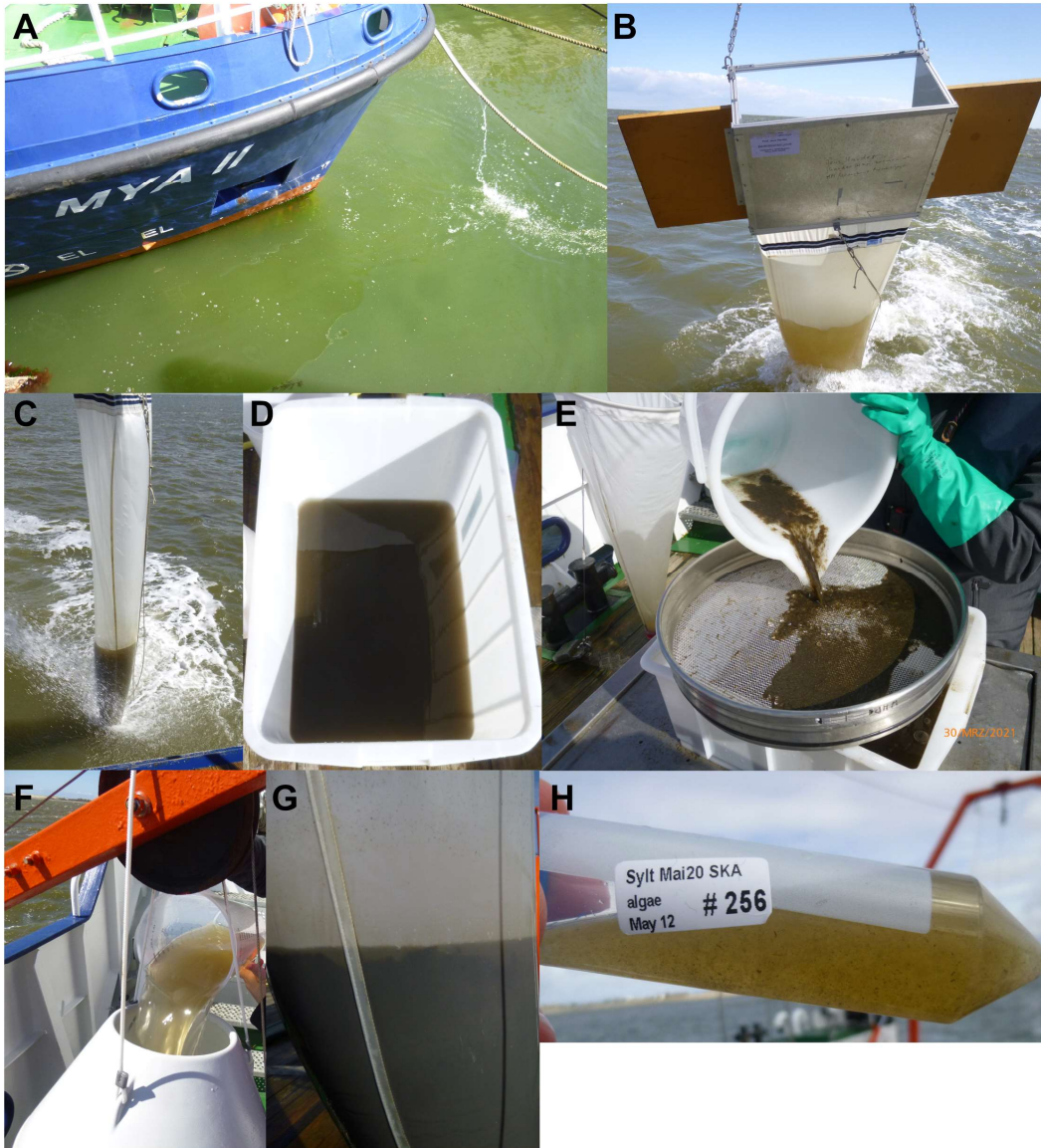
Supplement Figure 1: Pictures of taking sample of phytoplankton catches in the North Sea; A: shows the colouration of the North Sea in mid May 2020; B: Mantra trawl being hauled on to the ship; C: Mantra trawl is cleaned to not leave any phytoplankton behind in the upper parts of the net; D: sample retrieved from the mantra trawl; E: the received sample is sieved through a sieve with a mesh size of 0.2 mm to remove fish and jellyfish; F: the filtrate is poured into a second net with a mesh size of 80 µm to reduce the amount of liquid; G: shows the collect biomass in the second net, at this point already densely concentrated; H: collect biomass sample.
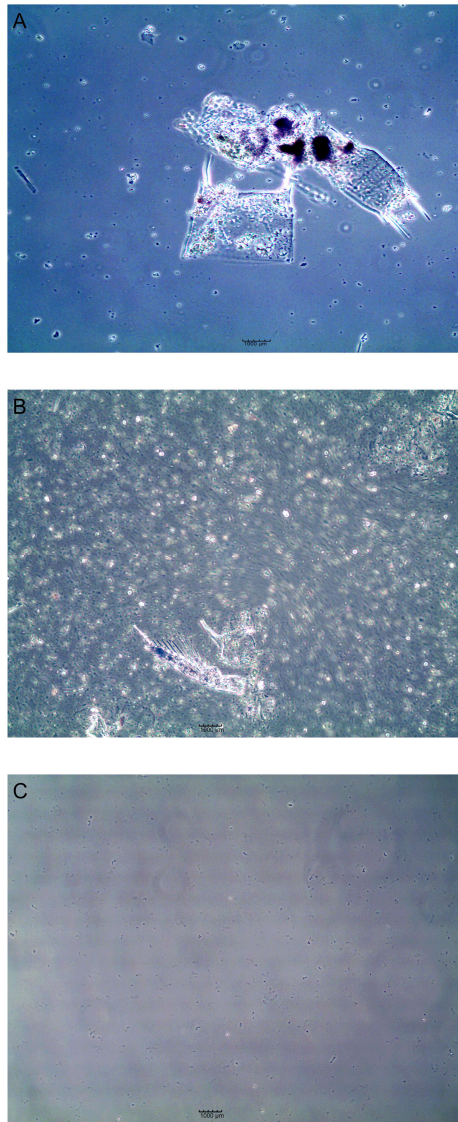
Supplement Figure 2: Microscopic pictures of collected biomass; A: untouched biomass; B: biomass after grinding; C: supernatant of biomass sample after centrifugation to demonstrate the absence of algae

PUL 1
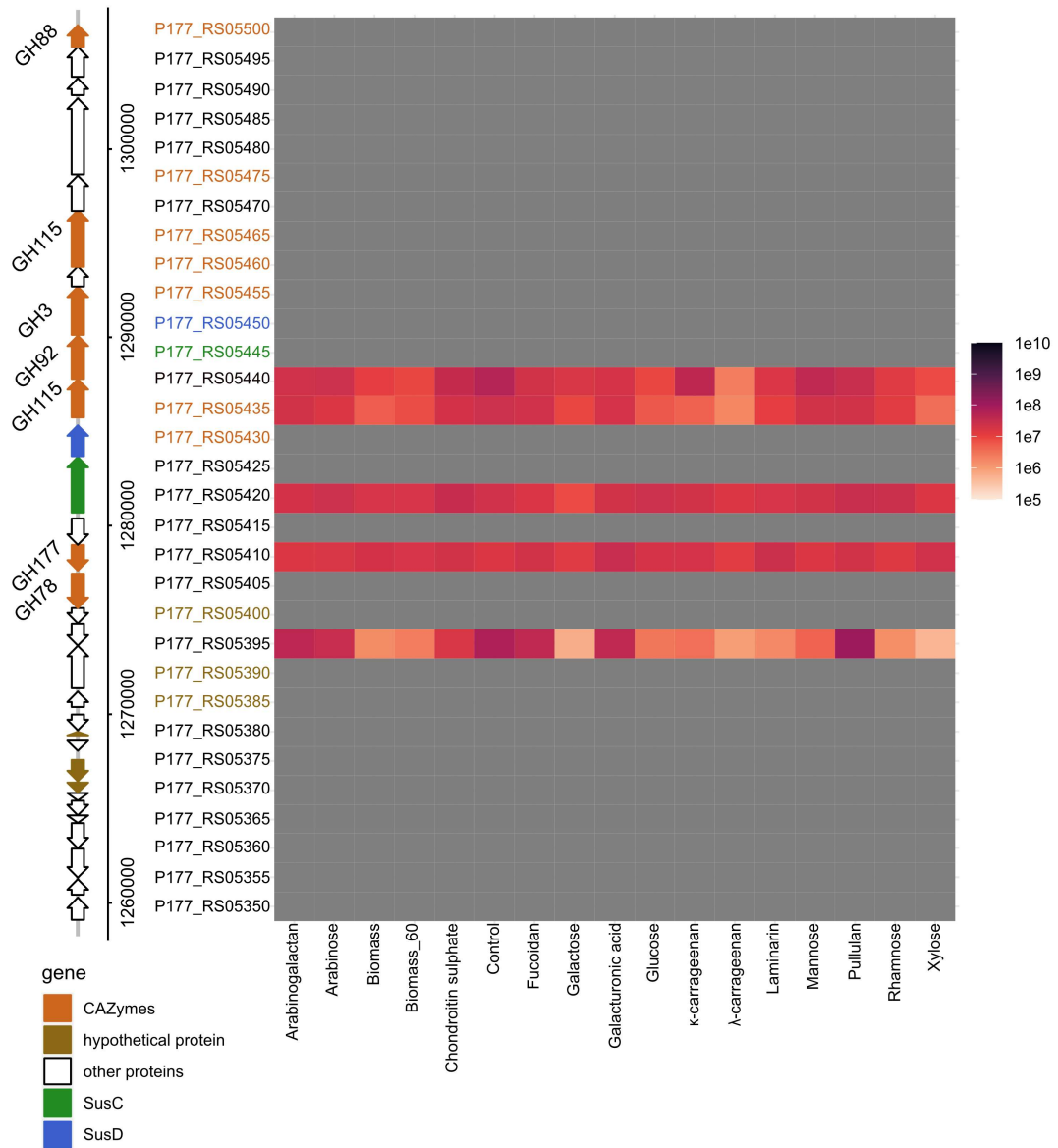
Supplement Figure 3: Gene organization and expression of PUL 1 of *Maribacter forsetii*. Expression intensities in the plot are the mean values of biological replicates of each condition shown in LFQ values [log10]. Grey means the proteins were not detected.
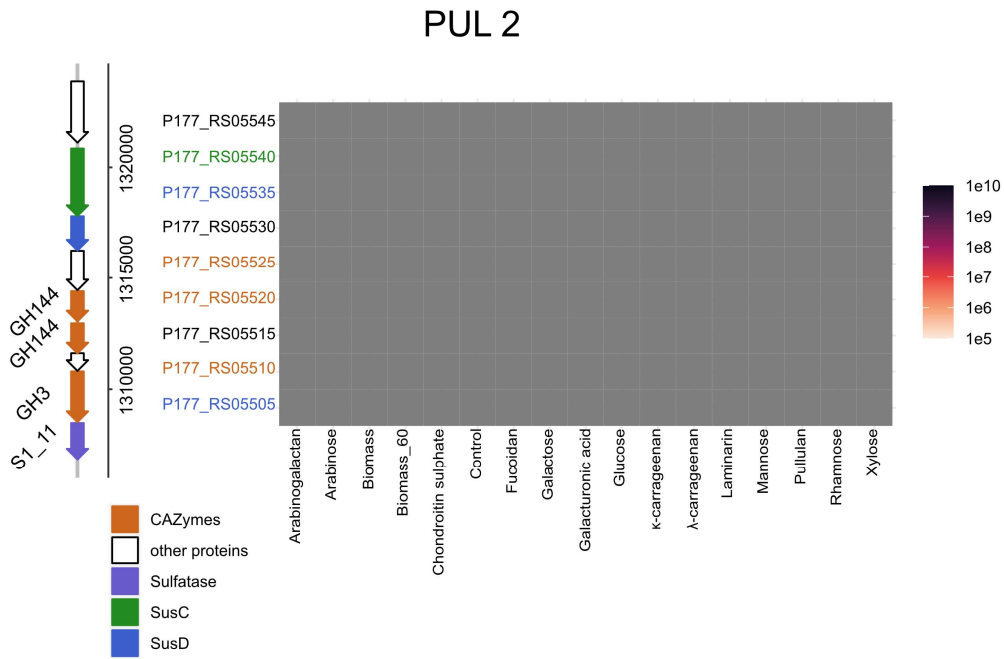
PUL 2

Supplement Figure 4: Gene organization and expression of PUL 2 of *Maribacter forsetii*. Expression intensities in the plot are the mean values of biological replicates of each condition shown in LFQ values [log10]. Grey means the proteins were not detected.

PUL 3
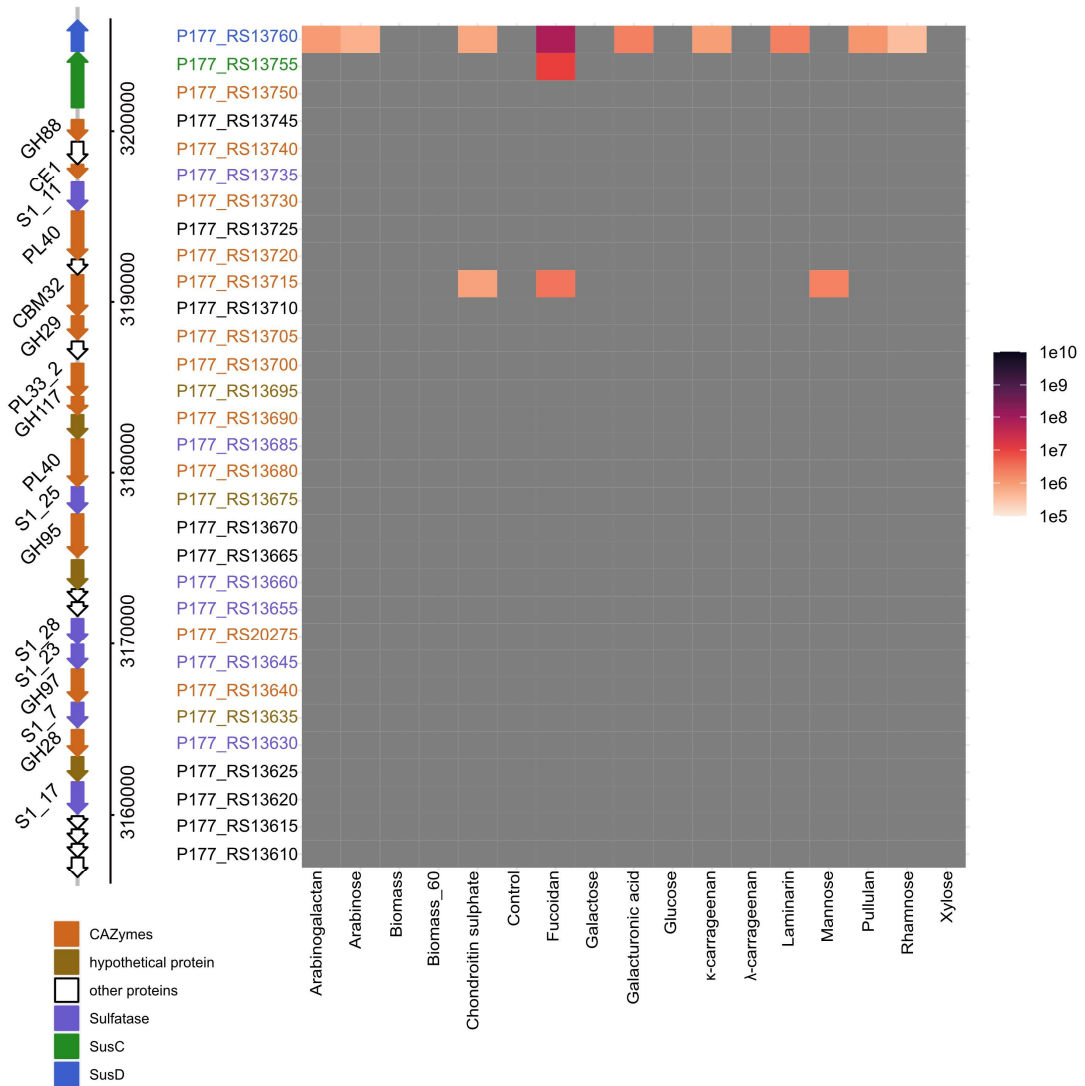
Supplement Figure 5: Gene organization and expression of PUL 3 of *Maribacter forsetii*. Expression intensities in the plot are the mean values of biological replicates of each condition shown in LFQ values [log10]. Grey means the proteins were not detected.

PUL 4
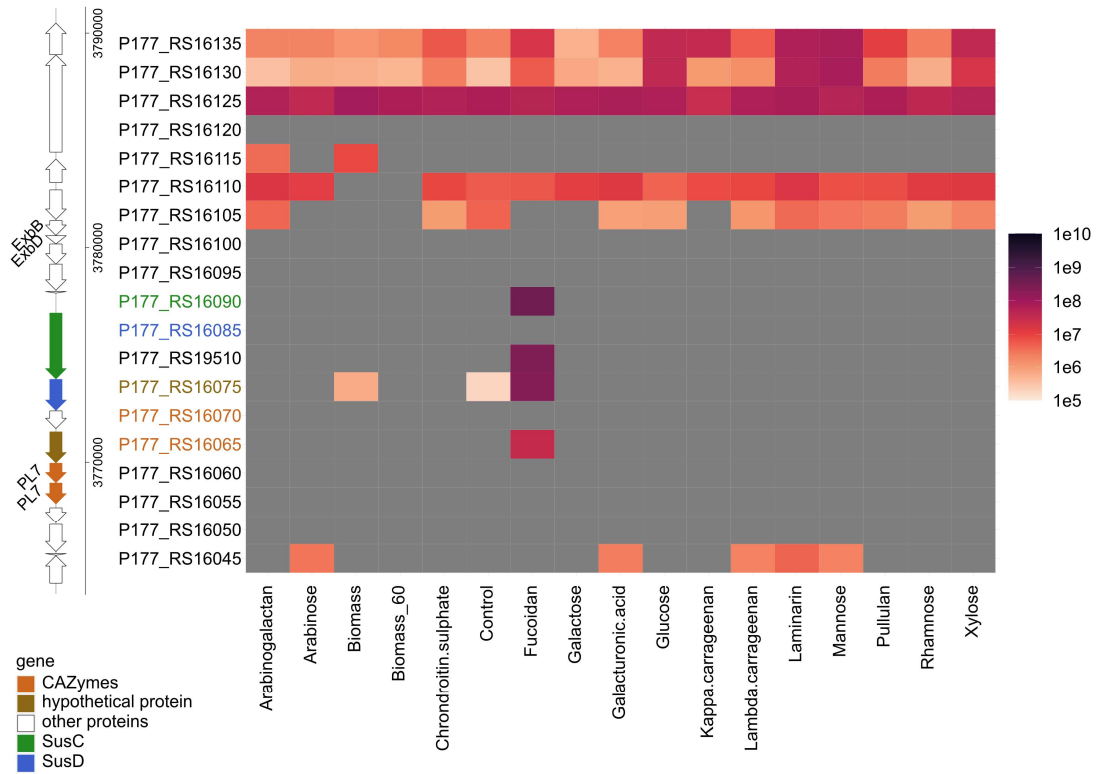
Supplement Figure 6: Gene organization and expression of PUL 4 of *Maribacter forsetii*. Expression intensities in the plot are the mean values of biological replicates of each condition shown in LFQ values [log10]. Grey means the proteins were not detected.

PUL 5
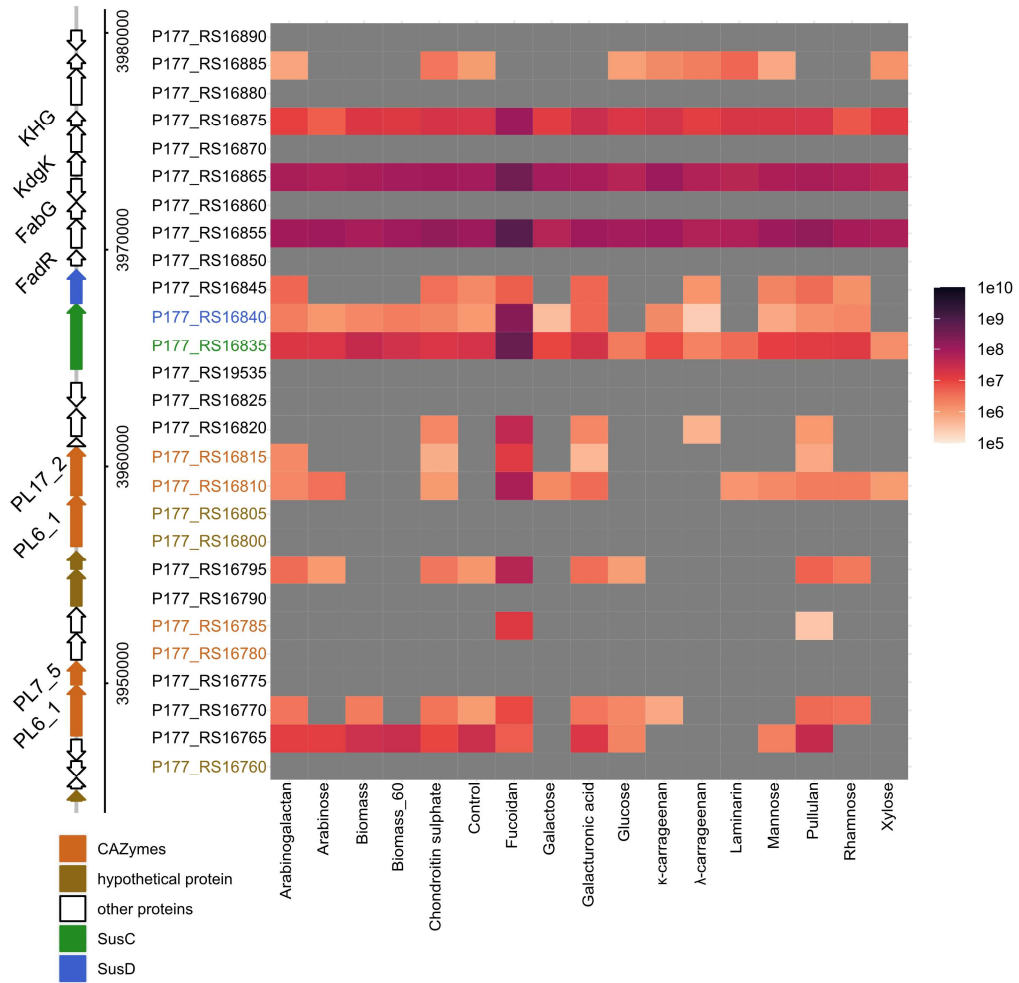
Supplement Figure 7: Gene organization and expression of PUL 5 of *Maribacter forsetii*. Expression intensities in the plot are the mean values of biological replicates of each condition shown in LFQ values [log10]. Grey means the proteins were not detected.

PUL 6
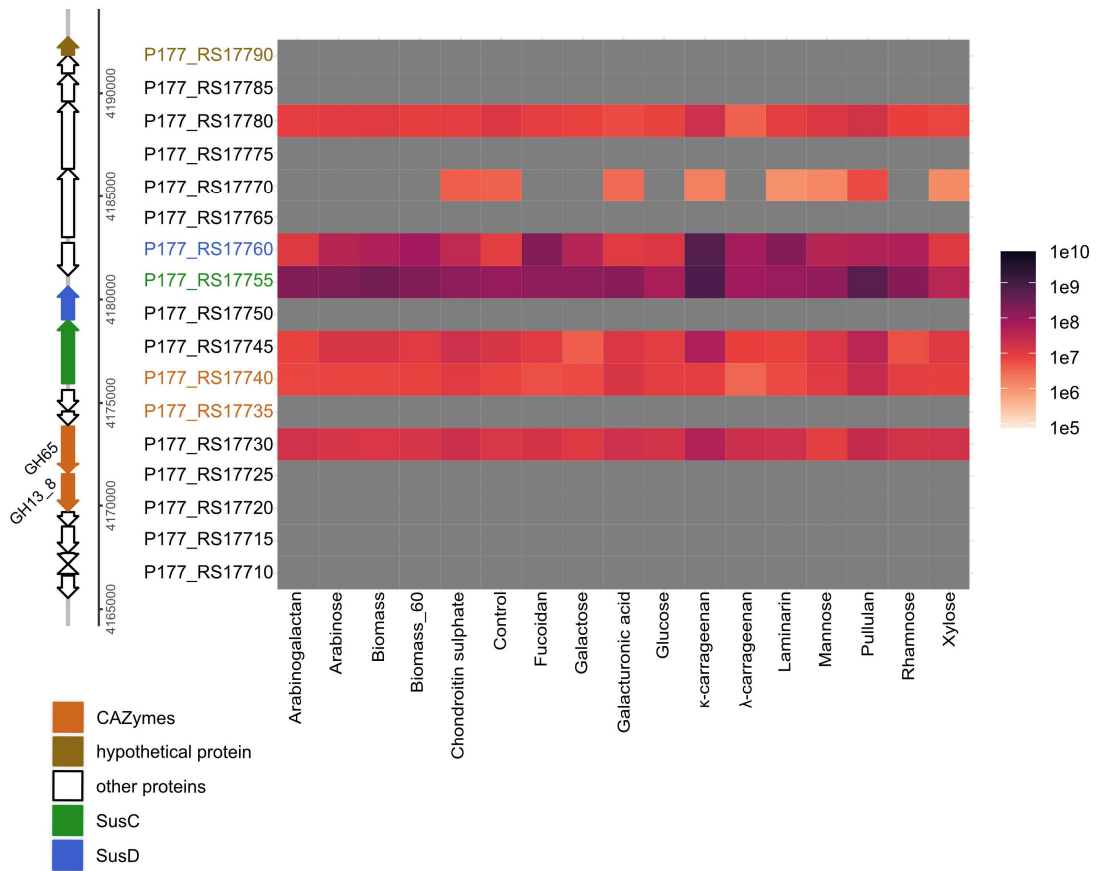
Supplement Figure 8: Gene organization and expression of PUL 6 of *Maribacter forsetii*. Expression intensities in the plot are the mean values of biological replicates of each condition shown in LFQ values [log10]. Grey means the proteins were not detected.

Chapter 4

4.7 References

Aalbers, F., Turkenburg, J.P., Davies, G.J., Dijkhuizen, L., and Lammerts van Bueren, A. (2015) Structural and functional characterization of a novel family GH115 4-O-methyl-α-glucuronidase with specificity for decorated arabinogalactans. *Journal of Molecular Biology* **427**: 3935-3946.

Alderkamp, A.-C., Buma, A.G.J., and van Rijssel, M. (2007) The carbohydrates of *Phaeocystis* and their degradation in the microbial food web. *Biogeochemistry* **83**: 99-118.

Aluwihare, L.I., and Repeta, D.J. (1999) A comparison of the chemical characteristics of oceanic DOM and extracellular DOM produced by marine algae. *Marine Ecology Progress Series* **186**: 105-117.

Arnosti, C., Wietz, M., Brinkhoff, T., Hehemann, J.H., Probandt, D., Zeugner, L., and Amann, R. (2021) The biogeochemistry of marine polysaccharides: Sources, inventories, and bacterial drivers of the carbohydrate cycle. *Annual Review of Marine Science* **13**: 81-108.

Barbeyron, T., Carpentier, F., Haridon, S., Schüler, M., Michel, G., and Amann, R. (2008) Description of *Maribacter forsetii* sp. nov., a marine *Flavobacteriaceae* isolated from North Sea water, and emended description of the genus *Maribacter*. *International Journal of Systematic and Evolutionary Microbiology* **58**: 790-797.

Barbeyron, T., Thomas, F., Barbe, V., Teeling, H., Schenowitz, C., Dossat, C. *et al.* (2016) 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* **18**: 4610-4627.

Bäumgen, M., Dutschei, T., and Bornscheuer, U.T. (2021) Marine polysaccharides: Occurrence, enzymatic degradation and utilization. *ChemBioChem* **22**: 2247-2256.

Bernardet, J.-F., Nakagawa, Y., and Holmes, B. (2002) Proposed minimal standards for describing new taxa of the family *Flavobacteriaceae* and emended

description of the family. *International Journal of Systematic and Evolutionary Microbiology* **52**: 1049-1070.

Bernardet, J.-F., Segers, P., Vancanneyt, M., Berthe, F., Kersters, K., and Vandamme, P. (1996) Cutting a gordian knot: Emended classification and description of the genus *Flavobacterium*, emended description of the family *Flavobacteriaceae*, and proposal of *Flavobacterium hydatis* nom. nov. (Basonym, *Cytophaga aquatilis* Strohl and Tait 1978). *International Journal of Systematic and Evolutionary Microbiology* **46**: 128-148.

Biersmith, A., and Benner, R. (1998) Carbohydrates in phytoplankton and freshly produced dissolved organic matter. *Marine Chemistry* **63**: 131-144.

Dlugosch, L., Bunse, C., Bunk, B., Böttcher, L., Tran, D.Q., Dittmar, T. *et al.* (2023) Naturally induced biphasic phytoplankton spring bloom reveals rapid and distinct substrate and bacterial community dynamics. *FEMS Microbiology Ecology* **99**: fiad078.

Drula, E., Garron, M.-L., Dogan, S., Lombard, V., Henrissat, B., and Terrapon, N. (2022) The carbohydrate-active enzyme database: Functions and literature. *Nucleic Acids Research* **50**: 571-577.

Dupoiron, S., Zischek, C., Ligat, L., Carbonne, J., Boulanger, A., Dugé de Bernonville, T. *et al.* (2015) The N-glycan cluster from *Xanthomonas campestris* pv. *campestris*: A toolbox for sequntial plant N-glycan processing. *Journal of Biological Chemistry* **290**: 6022-6036.

Fujita, K., Sasaki, Y., and Kitahara, K. (2019) Degradation of plant arabinogalactan proteins by intestinal bacteria: Characteristics and functions of the enzymes involved. *Applied Microbiology and Biotechnology* **103**: 7451-7457.

Gobet, A., Barbeyron, T., Matard-Mann, M., Magdelenat, G., Vallenet, D., Duchaud, E., and Michel, G. (2018) Evolutionary evidence of algal polysaccharide degradation acquisition by *Pseudoalteromonas carrageenovora* 9T to adapt to macroalgal niches. *Frontiers in Microbiology* **9**: 2740.

Gosink, J.J., Woese, C.R., and Staley, J.T. (1998) *Polaribacter* gen. nov., with three new species, *P. irgensii* sp. nov., *P. franzmannii* sp. nov. and *P. filamentus* sp. nov., gas vacuolate polar marine bacteria of the *Cytophaga-Flavobacterium-Bacteroides* group and reclassification of '*Flectobacillus glomeratus*' as *Polaribacter glomeratus* comb. nov. *International Journal of Systematic and Evolutionary Microbiology* **48**: 223-235.

Grant, J.R., Enns, E., Marinier, E., Mandal, A., Herman, E.K., Chen, C.-y. *et al.* (2023) Proksee: In-depth characterization and visualization of bacterial genomes. *Nucleic Acids Research* **51**: W484-W492.

Grootaert, H., Van Landuyt, L., Hulpiau, P., and Callewaert, N. (2020) Functional exploration of the GH29 fucosidase family. *Glycobiology* **30**: 735-745.

Hahnke, R.L., Bennke, C.M., Fuchs, B.M., Mann, A.J., Rhiel, E., Teeling, H. *et al.* (2015) Dilution cultivation of marine heterotrophic bacteria abundant after a spring phytoplankton bloom in the North Sea. *Environmental Microbiology* **17**: 3515-3526.

Hallgren, J., Tsirigos, K.D., Pedersen, M.D., Almagro Armenteros, J.J., Marcatili, P., Nielsen, H. *et al.* (2022) Preprint DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. *bioRxiv* 2022.04.08.487609.

Heins, A., and Harder, J. (2023) Particle-associated bacteria in seawater dominate the colony-forming microbiome on ZoBell marine agar. *FEMS Microbiology Ecology* **99**: fiac151.

Heins, A., Amann, R.I., and Harder, J. (2021) Cultivation of particle-associated heterotrophic bacteria during a spring phytoplankton bloom in the North Sea. *Systematic and Applied Microbiology* **44**: 126232.

Helbert, W. (2017) Marine polysaccharide sulfatases. *Frontiers in Marine Science* **4**: Article 6.

Hettle, A.G., Hobbs, J.K., Pluvinage, B., Vickers, C., Abe, K.T., Salama-Alber, O. *et al.* (2019) Insights into the κ/ι-carrageenan metabolism pathway of some marine *Pseudoalteromonas* species. *Communications Biology* **2**: 474.

Huang, G., Vidal-Melgosa, S., Sichert, A., Becker, S., Fang, Y., Niggemann, J. *et al.* (2021) Secretion of sulfated fucans by diatoms may contribute to marine aggregate formation. *Limnology and Oceanography* **66**: 3768-3782.

Ittekkot, V., Degens, E.T., and Brockmann, U. (1982) Monosaccharide composition of acid-hydrolyzable carbohydrates in particulate matter during a plankton bloom. *Limnology and Oceanography* **27**: 770-776.

Jouanneau, D., Klau, L.J., Larocque, R., Jaffrennou, A., Duval, G., Le Duff, N. *et al.* (2021) Structure–function analysis of a new PL17 oligoalginate lyase from the marine bacterium *Zobellia galactanivorans* DsijT. *Glycobiology* **31**: 1364-1377.

Kalenborn, S., Zühlke, D., Riedel, K., Amann, R.I., and Harder, J. (2024) Proteomic insight into arabinogalactan utilization by particle-associated *Maribacter* sp. MAR_2009_72. *FEMS Microbiology Ecology* **100**: fiae045.

Kanehisa, M., Sato, Y., and Morishima, K. (2016) BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *Journal of Molecular Biology* **428**: 726-731.

Kappelmann, L., Krüger, K., Hehemann, J.-H., Harder, J., Markert, S., Unfried, F. *et al.* (2019) Polysaccharide utilization loci of North Sea *Flavobacteriia* as basis for using SusC/D-protein expression for predicting major phytoplankton glycans. *The ISME Journal* **13**: 76-91.

Katoh, K., Rozewicki, J., and Yamada, K.D. (2019) MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics* **20**: 1160-1166.

Kitamura, M., Okuyama, M., Tanzawa, F., Mori, H., Kitago, Y., Watanabe, N. *et al.* (2008) Structural and functional analysis of a glycoside hydrolase family 97 enzyme from *Bacteroides thetaiotaomicron*. *Journal of Biological Chemistry* **283**: 36328-36337.

Krüger, K., Chafee, M., Ben Francis, T., Glavina del Rio, T., Becher, D., Schweder, T. *et al.* (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* **13**: 2800-2816.

Lapébie, P., Lombard, V., Drula, E., Terrapon, N., and Henrissat, B. (2019) *Bacteroidetes* use thousands of enzyme combinations to break down glycans. *Nature Communications* **10**: 2043.

Leszczuk, A., Kalaitzis, P., Kulik, J., and Zdunek, A. (2023) Review: Structure and modifications of arabinogalactan proteins (AGPs). *BMC Plant Biology* **23**: 45.

Lidbury, I.D.E.A., Borsetto, C., Murphy, A.R.J., Bottrill, A., Jones, A.M.E., Bending, G.D. *et al.* (2021) Niche-adaptation in plant-associated *Bacteroidetes* favours specialisation in organic phosphorus mineralisation. *The ISME Journal* **15**: 1040-1055.

Lu, D.-C., Wang, F.-Q., Amann, R.I., Teeling, H., and Du, Z.-J. (2023) Epiphytic common core bacteria in the microbiomes of co-located green (*Ulva*), brown (*Saccharina*) and red (*Grateloupia*, *Gelidium*) macroalgae. *Microbiome* **11**: 126.

Lu, S., Wang, J., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R. *et al.* (2019) CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Research* **48**: 265-268.

Luis, A.S., Briggs, J., Zhang, X., Farnell, B., Ndeh, D., Labourel, A. *et al.* (2018) Dietary pectic glycans are degraded by coordinated enzyme pathways in human colonic *Bacteroides*. *Nature Microbiology* **3**: 210-219.

Mathieu, S., Henrissat, B., Labre, F., Skjåk-Bræk, G., and Helbert, W. (2016) Functional exploration of the polysaccharide lyase family PL6. *PLOS ONE* **11**: e0159415.

McKee, L.S., La Rosa, S.L., Westereng, B., Eijsink, V.G., Pope, P.B., and Larsbrink, J. (2021) Polysaccharide degradation by the *Bacteroidetes*: Mechanisms and nomenclature. *Environmental Microbiology Reports* **13**: 559-581.

Miksch, S., Meiners, M., Meyerdierks, A., Probandt, D., Wegener, G., Titschack, J. *et al.* (2021) Bacterial communities in temperate and polar coastal sands are seasonally stable. *ISME Communications* **1**: 29.

Ndeh, D., Munoz Munoz, J., Cartmell, A., Bulmer, D., Wills, C., Henrissat, B., and Gray, J. (2018) The human gut microbe *Bacteroides thetaiotaomicron* encodes the founding member of a novel glycosaminoglycan-degrading polysaccharide lyase family PL29. *Journal of Biological Chemistry* **293**: 17906-17916.

Ndeh, D., Baslé, A., Strahl, H., Yates, E.A., McClurgg, U.L., Henrissat, B. *et al.* (2020) Metabolism of multiple glycosaminoglycans by *Bacteroides thetaiotaomicron* is orchestrated by a versatile core genetic locus. *Nature Communications* **11**: 646.

Nedashkovskaya, O.I., Kim, S.B., Han, S.K., Lysenko, A.M., Rohde, M., Rhee, M.-S. *et al.* (2004) *Maribacter* gen. nov., a new member of the family *Flavobacteriaceae*, isolated from marine habitats, containing the species *Maribacter sedimenticola* sp. nov., *Maribacter aquivivus* sp. nov., *Maribacter orientalis* sp. nov. and *Maribacter ulvicola* sp. nov. *International Journal of Systematic and Evolutionary Microbiology* **54**: 1017-1023.

Noinaj, N., Guillier, M., Barnard, T.J., and Buchanan, S.K. (2010) TonB-dependent transporters: Regulation, structure, and function. *Annual Review of Microbiology* **64**: 43-60.

Omand, M.M., Govindarajan, R., He, J., and Mahadevan, A. (2020) Sinking flux of particulate organic matter in the oceans: Sensitivity to particle characteristics. *Scientific Reports* **10**: 5582.

Patel, A.K., Vadrale, A.P., Singhania, R.R., Michaud, P., Pandey, A., Chen, S.-J. *et al.* (2023) Algal polysaccharides: Current status and future prospects. *Phytochemistry Reviews* **22**: 1167-1196.

Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B.L., Salazar, Gustavo A. *et al.* (2022) InterPro in 2022. *Nucleic Acids Research* **51**: 418-427.

Probandt, D., Eickhorst, T., Ellrott, A., Amann, R., and Knittel, K. (2018) Microbial life on a sand grain: From bulk sediment to single grains. *The ISME Journal* **12**: 623-633.

R Core Team (2023) R: A language and environment for statistical computing. R Foundation for Statistical Computing. In. Vienna, Austria.

Rawat, P.S., Seyed Hameed, A.S., Meng, X., and Liu, W. (2022) Utilization of glycosaminoglycans by the human gut microbiota: Participating bacteria and their enzymatic machineries. *Gut Microbes* **14**: 2068367.

Rebuffet, E., Groisillier, A., Thompson, A., Jeudy, A., Barbeyron, T., Czjzek, M., and Michel, G. (2011) Discovery and structural characterization of a novel glycosidase family of marine origin. *Environmental Microbiology* **13**: 1253-1270.

Reisky, L., Préchoux, A., Zühlke, M.-K., Bäumgen, M., Robb, C.S., Gerlach, N. *et al.* (2019) A marine bacterial enzymatic cascade degrades the algal polysaccharide ulvan. *Nature Chemical Biology* **15**: 803-812.

Sayers, E.W., Bolton, E.E., Brister, J.R., Canese, K., Chan, J., Comeau, Donald C. *et al.* (2022) Database resources of the national center for biotechnology information. *Nucleic Acids Research* **50**: 20-26.

Scholz, B., and Liebezeit, G. (2013) Biochemical characterisation and fatty acid profiles of 25 benthic marine diatoms isolated from the Solthörn tidal flat (southern North Sea). *Journal of Applied Phycology* **25**: 453-465.

Schultz, D., Zühlke, D., Bernhardt, J., Francis, T.B., Albrecht, D., Hirschfeld, C. *et al.* (2020) An optimized metaproteomics protocol for a holistic taxonomic and functional characterization of microbial communities from marine particles. *Environmental Microbiology Reports* **12**: 367-376.

Shin, J.-I., Bang, M.-S., Kim, J.Y., and Oh, C.-H. (2022) Complete Genome Sequence of *Flagellimonas* sp. Strain CMM7, Isolated from a marine green alga, *Codium minus* (Schmidt) Silva. *Microbiology Resource Announcements* **11**: e00153-22.

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* **182**: 5365-5372.

Shuoker, B., Pichler, M.J., Jin, C., Sakanaka, H., Wu, H., Gascueña, A.M. *et al.* (2023) Sialidases and fucosidases of *Akkermansia muciniphila* are crucial for growth on mucin and nutrient sharing with mucus-associated gut bacteria. *Nature Communications* **14**: 1833.

Sichert, A., Corzett, C.H., Schechter, M.S., Unfried, F., Markert, S., Becher, D. *et al.* (2020) *Verrucomicrobia* use hundreds of enzymes to digest the algal polysaccharide fucoidan. *Nature Microbiology* **5**: 1026-1039.

Sidhu, C., Kirstein, I.V., Meunier, C.L., Rick, J., Fofonova, V., Wiltshire, K.H. *et al.* (2023) Dissolved storage glycans shaped the community composition of abundant bacterioplankton clades during a North Sea spring phytoplankton bloom. *Microbiome* **11**: 77.

Stam, M., Lelièvre, P., Hoebeke, M., Corre, E., Barbeyron, T., and Michel, G. (2022) SulfAtlas, the sulfatase database: State of the art and new developments. *Nucleic Acids Research* **51**: 647-653.

Tamura, K., Stecher, G., and Kumar, S. (2021) MEGA11: Molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution* **38**: 3022-3027.

Terrapon, N., Lombard, V., Drula, É., Lapébie, P., Al-Masaudi, S., Gilbert, H.J., and Henrissat, B. (2018) PULDB: The expanded database of polysaccharide utilization loci. *Nucleic Acids Research* **46**: D677-D683.

Teufel, F., Almagro Armenteros, J.J., Johansen, A.R., Gíslason, M.H., Pihl, S.I., Tsirigos, K.D. *et al.* (2022) SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nature Biotechnology* **40**: 1023-1025.

The UniProt, C. (2023) UniProt: The universal protein knowledgebase in 2023. *Nucleic Acids Research* **51**: D523-D531.

Thomas, F., Lundqvist, L.C.E., Jam, M., Jeudy, A., Barbeyron, T., Sandström, C. *et al.* (2013) Comparative characterization of two marine alginate lyases from *Zobellia galactanivorans* reveals distinct modes of action and exquisite adaptation to their natural substrate. *Journal of Biological Chemistry* **288**: 23021-23037.

Tyanova, S., and Cox, J. (2018) Perseus: A bioinformatics platform for integrative analysis of proteomics data in cancer research. In *Cancer Systems Biology: Methods and Protocols*. von Stechow, L. (ed). New York, NY: Springer New York, pp. 133-148.

Tyanova, S., Temu, T., and Cox, J. (2016) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols* **11**: 2301-2319.

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* **12**: 2894–2906.

Urbani, R., Magaletti, E., Sist, P., and Cicero, A.M. (2005) Extracellular carbohydrates released by the marine diatoms *Cylindrotheca closterium*, *Thalassiosira pseudonana* and *Skeletonema costatum*: Effect of P-depletion and growth status. *Science of The Total Environment* **353**: 300-306.

Vidal-Melgosa, S., Sichert, A., Francis, T.B., Bartosik, D., Niggemann, J., Wichels, A. *et al.* (2021) Diatom fucan polysaccharide precipitates carbon during algal blooms. *Nature Communications* **12**: 1150.

Villa-Rivera, M.G., Cano-Camacho, H., López-Romero, E., and Zavala-Páramo, M.G. (2021) The role of arabinogalactan type II degradation in plant-microbe interactions. *Frontiers in Microbiology* **12**: 730543.

Wang, F.-Q., Bartosik, D., Sidhu, C., Siebers, R., Lu, D.-C., Trautwein-Schult, A. *et al.* (2024a) Particle-attached bacteria act as gatekeepers in the decomposition of complex phytoplankton polysaccharides. *Microbiome* **12**: 32.

Wang, Y., Ma, M., Dai, W., Shang, Q., and Yu, G. (2024b) *Bacteroides salyersiae* is a potent chondroitin sulfate-degrading species in the human gut microbiota. *Microbiome* **12**: 41.

Wickham, H. (2016) *ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag.

Wilkins, D. (2023) gggenes: Draw gene arrow maps in 'ggplot2'. In. https://wilkox.org/gggenes/.

Williams, P.A., Clegg, S.M., Day, D.H., Phillips, G.O., and Nishinari, K. (1991) Mixed gels formed with konjac mannan and xanthan gum. In *Food Polymers, Gels and Colloids*. Dickinson, E. (ed): Woodhead Publishing, pp. 339-348.

Winkelmann, N., and Harder, J. (2009) An improved isolation method for attached-living *Planctomycetes* of the genus *Rhodopirellula*. *Journal of Microbiological Methods* **77**: 276-284.

Zeugner, L.E., Krüger, K., Barrero-Canosa, J., Amann, R.I., and Fuchs, B.M. (2021) In situ visualization of glycoside hydrolase family 92 genes in marine *Flavobacteriia*. *ISME Communications* **1**: 81.

Zheng, J., Ge, Q., Yan, Y., Zhang, X., Huang, L., and Yin, Y. (2023) dbCAN3: Automated carbohydrate-active enzyme and substrate annotation. *Nucleic Acids Research* **51**: W115-W121.

Zhu, Y., Suits, M.D., Thompson, A.J., Chavan, S., Dinev, Z., Dumon, C. *et al.* (2010) Mechanistic insights into a Ca2+-dependent family of alpha-mannosidases in a human gut symbiont. *Nat Chem Biol* **6**: 125-32.

General Discussion

5. General Discussion

This thesis is part of the Proteogenomics of Marine Polysaccharide Utilization project (POMPU). The research group set out to explore functional analyses of marine bacteria of the phylum *Bacteroidota*, especially in regards to polysaccharide utilization. The subproject, this thesis belongs to, focusses on bacteria associated to phytoplankton particles.

Particle associated (PA) bacteria were studied to a lesser extent than free-living (FL) bacteria (Seymour *et al.*, 2017). In the last decade the immense impact of PA bacteria on nutrient cycling and biogeochemical processes, as well as marine biodiversity was recognized (Grossart *et al.*, 2007; Lauro *et al.*, 2009; Ziervogel *et al.*, 2010; Pedler *et al.*, 2014; Heins and Harder, 2023). They create microhabitats with marine particles, constructing diverse microbial communities and encountering many environmental conditions (Seymour *et al.*, 2017; Heins and Harder, 2023). To understand how they adapt to the environment, as well as grasp species to species interaction in these microhabitats, we need to intensify the investigation of PA bacteria. In recent years this has been implemented, demonstrating the impact of PA bacteria in the environment in various habitats (Kappelmann *et al.*, 2019; Heins *et al.*, 2021a; Heins *et al.*, 2021b; Miksch *et al.*, 2021; Giljan *et al.*, 2022; Heins and Harder, 2023; Wang *et al.*, 2024). One suggested particle associated genus is *Maribacter* (Kappelmann *et al.*, 2019). The description of particle-association of *Maribacter* has been confirmed by studies, showing that *Maribacter* strains were present and isolated from samples of the above 3 μm fraction of various marine environments (Heins *et al.*, 2021a; Heins *et al.*, 2021b; Miksch *et al.*, 2021; Heins and Harder, 2023; Lu *et al.*, 2023; Wang *et al.*, 2024).

One strain I focused on is *Maribacter forsetii* DSM 18668 (Barbeyron *et al.*, 2008). It is one of five *Maribacter* strains analyzed in the *in-silico* study of Kappelmann *et al.* (2019). With its complete genome, *M. forsetii* serves as a solid basis for studying polysaccharide utilization. As described in Chapter 4, I conducted a large proteome study with *M. forsetii* to gain insight about the bacterium's ability to degrade various saccharides. I discovered that *M. forsetii* is capable of growing in presence of a wide range of substrates, ranging from simple monosaccharides such as glucose to complex polysaccharides like fucoidan. I identified some expressed CAZymes

for certain substates, for instance a GH29, an α-L-fucosidase, for fucoidan degradation. But in this study many CAZymes and SusC/D transporters were not expressed in any of the 17 growth conditions. This let me to hypothesize that the natural substrates of *Maribacter forsetii* were not included in this study. Furthermore, this indicates as well that there are more glycans in the marine environment, that have yet to be discovered and structurally analyzed. For example, PUL 2 has many homologues in the PULDB database, but no substrate affiliation has been reported (Barbeyron *et al.*, 2016a; Terrapon *et al.*, 2018; Lapébie *et al.*, 2019). I furthermore noticed that *M. forsetii* has a low number of PULs (6), considering that 119 CAZymes and 18 SusC/D pairs are annotated in the genome. Most of its CAZymes and SusC/D pairs are not organized in PULs.

For laminarin as a substrate, I closely examined the proteomic data. In the previous mentioned *in-silico* study no PUL for laminarin was identified for all included *Maribacter* strains (Kappelmann *et al.*, 2019). In Chapter 2 I explored the laminarin degradation of *Maribacter forsetii*, after noticing its ability to grown in its presence. Not only had *M. forsetii* key enzymes such as GH16_3 annotated in the genome, moreover the enzymes were expressed when grown in the presence of laminarin. However, none of the enzymes are located close, rather dispersed over the whole genome. In addition, I discovered that *Maribacter* sp. Hel_I_7, *Maribacter* sp. MAR_2009_72 and *Maribacter dokdonensis* MAR_2009_60 (Hahnke and Harder, 2013) have like *M. forsetii* a GH16_3 annotated in their genome, but it is not located in a PUL (Chapter 2). *Maribacter dokdonensis* MAR_2009_71 (Hahnke and Harder, 2013) does not encode a GH16_3 in its genome, however it grew in presence of laminarin and activity tests showed that laminarin is degraded. It is not disclosed how the bacterium is utilizing laminarin, further experiments have to be performed to discover the novel utilization pathway in *Maribacter dokdonensis* MAR_2009_71 for laminarin. With this knowledge in mind, I explored other *Maribacter* genomes (Table 1). Most strains have at least one GH16_3 annotated, like *M. litoralis* 151 (Lee *et al.*, 2018), some have multiple copies, for example *M. aestuarii* GY20 (Lo *et al.*, 2013), *M. flavus* KCTC 42508 (Tang *et al.*, 2015), *M. halichondriae* Hal144 (Steiner *et al.*, 2024) and *M. thermophilus* HT7-2 (Hu *et al.*, 2015). Only three investigated *Maribacter* strains have no GH16_3 annotated, according to dbCAN3, them being *M. dokdonensis* MAR_2009_71, *M. aurantiacus* DSM 23546 (Khan *et*

al., 2019) and *M. caenipelagi* CECT 8455 (Jung *et al.*, 2014). For the last two mentioned strains, there is no knowledge if they are able to grow in presence of laminarin. But potentially if both are able to, they might use the same novel pathway as MAR_2009_71. Moreover, I discovered that several strains – *M. aquivivus* DSM 16478 (Nedashkovskaya *et al.*, 2004) , *M. arcticus* DSM 23546 (Cho *et al.*, 2008), *M. cobaltidurans* B1 (Fang *et al.*, 2017), *M. hydrothermalis* T28 (Lin *et al.*, 2021), *M. litopenaei* HL-LV01 (Kim *et al.*, 2023), *M. orientalis* DSM 16471 (Nedashkovskaya *et al.*, 2004), *M. sedimenticola* DSM 19840 (Nedashkovskaya *et al.*, 2004), *M. stanieri* DSM 19891 (Nedashkovskaya *et al.*, 2010) and *M. ulvicola* DSM 15366 (Nedashkovskaya *et al.*, 2004) – as analyzed by PULDB, have no PUL for laminarin, but have a GH16_3 in their genome. Possibly all of them degrade laminarin in a similar manner as *Maribacter forsetii*. Some strains have a similar laminarinase to *M. forsetii*, such as *M. aquivivus* DSM 16478, *M. arcticus* DSM 23546, *M. confluentis* CECT 8869 (Park *et al.*, 2015), *M. dokdonensis* MAR_2009_60, *M. huludaoensis* M208 (Gao *et al.*, 2023), *M. hydrothermalis* T28, *M. litoralis* HL-LV01, *M. sedimenticola* DSM 19840, *M. spongiicola* DSM 25233 (Jackson *et al.*, 2015), *M. stanieri* DSM 19891 (Nedashkovskaya *et al.*, 2010) and *M. zhoushanensis* SA7 (Gao *et al.*, 2023). As stated in Chapter 2 this is a novel finding, especially since laminarin utilization is seemingly more common in the genus of *Maribacter* even in the absence of a PUL, in contrast to previous studies in other flavobacterial genera that have always linked the utilization to PULs (Ziervogel *et al.*, 2010; Kabisch *et al.*, 2014; Barbeyron *et al.*, 2016b; Tang *et al.*, 2017; Thomas *et al.*, 2017; Chen *et al.*, 2018; Unfried *et al.*, 2018).

Table 1: Overview of 35 *Maribacter* strains, that have a published genome in NCBI, enclosing main information about the genomes, as well as specific knowledge about CAZymes and PULs, retrieved from dbCAN3 and PULDB (Terrapon *et al.*, 2018; Zheng *et al.*, 2023).

Strain	Publication	Isolation site	Source	DSMZ/ENA	Size [Mb]	GC %	genes NCBI	scaffolds	completeness	CAZymes	GH16_3	Lam PUL according to PULDB	sulfatases
M. aestuari GY20	Lo et al., 2013	Gwangyang Bay, South Korea	tidal flat sediment sample 5cm depth	JN642273.2	3.86	39.30	3538	1	chromosome	94	2	/	9
M. algarum RZ26	Zhang et al., 2020	coast of Weihai, PR China	marine red algae Gelidium amansii	MK907697.1	5.11	38.00	4436	20	scaffold	156	3	/	37
M. algicola PoM-212	Khan et al., 2019	Yellow Sea, Republic of Korea	marine alga, Porphyridium marinum	MH742311.1	4.11	41.00	3566	7	contig	130	1	/	19
M. antarcticus DSM 21422	Zhang et al., 2009	Korea Polar Research Institute, KOPRI Culture Collection of Polar Micro-organism, KCCPM strain AnM0046; Southern Ocean	culture of a green alga, Pyramimonas gelidicola	21422	4.84	37.40	4213	32	scaffold	129	1	yes	25
M. aquimaris ANRC-HE7	Zhang et al., 2023	Biological Bay adjacent to Fildes Peninsula, Antarctica	seawater	MN658399.1	4.86	40.10	4062	27	contig	235	2	/	26
M. aquivivus DSM 16478	Nedashkovskaya et al., 2004	Gulf of Peter the Great, Sea of Japan	seawater	16478	4.58	34.50	3999	11	scaffold	141	2	/	12
M. arcticus DSM 23546	Cho et al., 2008	Spitsbergen, Norway	marine sediment from Ny-Ålesund	AY771762.1	4.21	35.00	3728	22	scaffold	123	2	no	8
M. arenosus CAU 1321	Thongphrom et al., 2016	Eunwongri beach, South Korea	marine sediment at Eunwongri beach	KU719511.1	4.17	39.90	3697	24	contig	172	3	/	15
M. auraniacus KCTC 52409	Khan et al., 2019	Dengubu Island Zhejiang, China	sediment from sedimentation basin of a mariculture	KX364240.1	4.16	41.80	3600	35	contig	125	0	/	13
M. caenipelagi CECT 8455	Jung et al., 2014	Hwang-do, Yellow Sea, South Korea	tidal flat sediment	KF748920.1	4.28	35.20	3788	20	scaffold	130	0	/	2
M. cobaltidurans B1	Fang et al., 2017	South Atlantic Ocean	deep sea sediment (2819m depth)	JQ781697.2	4.64	39.70	4145	1	complete	133	1	no	4
M. confluentis CECT 8869	Park et al., 2015	Jeju Island South Korea	water sample from the place where the ocean and spring water Soesokkak meet	KR006347.1	4.7	36.70	4140	16	contig	167	2	/	6
M. dokdonensis MAR_2009_60	Hahnke et al., 2013	Sylt, List	20 µm phytoplankton catch	29383	4.51	36.10	3914	1	chromosome	144	1	/	20
M. dokdonensis MAR_2009_71	Hahnke et al., 2013	Sylt, List	20 µm phytoplankton catch	29385	4.62	36.10	3955	1	contig	146	0	/	28
M. dokdonensis DSW-8	Yoon et al., 2005	Korean Island Dokdo of the East Sea, Korea	seawater	AY960749.1	4.43	36.10	3795	2	contig	144	1	yes	21
M. flavus KCTC 42508	Tang et al., 2015	South China Sea Institute of Oceanography, Guangzhou, China	cyanobacterial culture pond of the South China Sea Institute of Oceanology	KR058351.1	4.3	41.80	3758	1	contig	130	2	/	17
M. forsetii DSM 18668	Barbeyron et al., 2008	Kabeltonne, Helgoland	seawater	18668	4.51	35.20	3890	1	contig	119	1	no	12
M. halichondriae Hal144	Steiner et al., 2024	Germany, Kiel Fjord, Schilksee	marine breadcrumb sponge Halichondria panicea	MT406525.2	4.53	41.40	4326	1	chromosome	120	2	/	25
M. huludaoensis M208	Gao et al., 2023	Zhoushan and Huludao, China	intertidal sediments	OQ165154.1	4.24	35.90	3676	15	scaffold	143	1	/	24
M. hydrothermalis T28	Lin et al., 2021	hallow-sea hydrothermal system, Kueishantao Islet, Taiwan, China	seawater	KX022625.1	4.27	34.40	3713	1	complete	116	1	no	8
M. litopenraei HL-LV01	Kim et al., 2023	marine shrimp farm in Dangjin, Republic of Korea	intestinal tract content of the Pacific white shrimp Litopenaeus vannamei	ON340574.1	3.86	39.80	4379	1	complete	88	1	no	9
M. litoralis 151	Lee et al., 2018	Sinduri beach in Taean, Republic of Korea	beach sediment	MG456900.1	4.36	35.90	3754	4	scaffold	146	1	/	21
M. luteus RZ05	Liu et al., 2020	ellow Sea, PR China	sand sample collected from the intertidal zone of the Yellow Sea	MH698963.1	4.65	38.90	3922	3	scaffold	240	3	/	30
M. orientalis DSM 16471	Nedashkovskaya et al., 2004	Gulf of Peter the Great, Sea of Japan	seawater	AY271624.1	4.16	35.10	3711	26	scaffold	132	1	no	8
M. polysaccharolyticus D37	Gao et al., 2023	Zhoushan and Huludao, China	intertidal sediments	OP132875.1	4.76	41.70	4018	21	scaffold	224	3	/	51
M. polysiphoniae DSM 23514	Nedashkovskaya et al., 2007	Gulf of Peter the Great, Sea of Japan	pacific red algae Polysiphonia japonica	23514	5.13	40.30	4419	30	scaffold	207	2	yes	18
M. sedimenticola DSM 19840	Nedashkovskaya et al., 2004	Gulf of Peter the Great, Sea of Japan	coastal sediment	19840	4.27	37.50	3684	20	scaffold	151	1	no	7
M. sp. Hel_I_7	Hahnke et al., 2013	Kabeltonne, Helgoland	seawater	26657	4.78	34.80	4182	3	contig	132	1	/	8
M. sp. MAR_2009_72	Hahnke et al., 2013	Sylt, List	20 µm phytoplankton catch	29384	4.35	37.70	3696	1	contig	151	1	/	14
M. spongiicola DSM 25233	Jackson et al., 2015	Lough Hyne Co.Cork, Ireland	marine sponge Suberites carnosus	25233	4.46	35.10	3857	18	scaffold	143	1	no	24
M. stanieri DSM 19891	Nedashkovskaya et al., 2010	Troitsa Bay, Gulf of Peter the Great, Sea of Japan	green alga Ulva fenestrata	19891	4.46	34.40	3872	9	scaffold	142	1	/	14
M. thermophilus HT7-2	Hu et al., 2015	Qingdao sea area,China	seaweed Ulva prolifera	JQ988061.1	4.05	38.90	3514	2	scaffold	115	3	/	11
M. ulvicola DSM 15366	Nedashkovskaya et al., 2004	Gulf of Peter the Great, Sea of Japan	green algae Ulva fenestrata	15366	4.51	35.20	3875	24	contig	136	1	no	16
M. vaceletii DSM 25230	Jackson et al., 2015	Lough Hyne Co.Cork, Ireland	marine sponge Leucosolenia sp.	25230	3.89	31.70	3368	15	scaffold	164	1	yes	50
M. zhoushanensis SA7	Gao et al., 2023	Zhoushan and Huludao,China	intertidal sediments	OQ165151.1	4.26	34.80	3716	18	scaffold	111	1	/	2

Within POMPU, a range of polysaccharides is characterized biochemically, but no galactan and arabinan. Realizing that galactose and arabinose are abundant sugar in the North Sea, larch wood arabinogalactan was selected for a proteomic investigation. In Chapter 3, I proposed a utilization pathway for arabinogalactan by *Maribacter* sp. MAR_2009_72, involving three of its nine PULs. I discussed that according to PULDB PUL 1 has homologues in gut associated *Bacteroidota* strains (Martens *et al.*, 2008; Rogowski *et al.*, 2015; Wang and LaPointe, 2020). For PUL 7 only a homologue in *Maribacter sedimenticola* was identified, which has not been described in the literature. A similar result was obtained for PUL 8, a homologue in *M. sedimenticola* and as well other *Bacteroidota*, but as for PUL 7 none was described in the literature. Due to the low number of homologues I decided to expand the comparison analyses by manually looking at the genomes of the *Maribacter* strains that I studied, *Maribacter forsetii*, *Maribacter* sp. Hel_I_7, *Maribacter dokdonensis* MAR_2009_60 and *Maribacter dokdonensis* MAR_2009_71. I analyzed the genomes to see whether they encode CAZymes known to be involved in arabinogalactan utilization, such as GH43 and GH51 (Ndeh *et al.*, 2017; Cartmell *et al.*, 2018; Luis *et al.*, 2018; Wang and LaPointe, 2020). All four strains either encode only one GH43 or no arabinogalactan affiliated CAZymes (Table 2). Therefore, I hypothesize these four *Maribacter* strains are not able of degrading the whole compound arabinogalactan. I rather propose that they are potentially able to hydrolyze bonds of some decorations, like rhamnose, fucose, glucuronic acid, mannose and xylose, that are attached to the main larch wood arabinogalactan structure and degrade arabinogalactan partially, using CAZymes such as GH3, GH10, GH29, GH31, GH67, GH78 and GH115 (Table 2) (Fujita *et al.*, 2019; Villa-Rivera *et al.*, 2021; Leszczuk *et al.*, 2023).

I broadened the analysis and included 30 other *Maribacter* strains. Most of them have enzymes of the family GH43 encoded in their genome, ranging from one GH43 to thirteen GH43 in a genome. However, I did not only exclusively look for GH43 in the genomes, I as well searched for GH51 and GH105, that have been identified as potential arabinogalactan degrading CAZymes (Kalenborn *et al.*, 2024). The *Maribacter* strains with the highest potential to degrade arabinogalactan, based on the presence and numbers of CAZymes in the genome are *M. aquimaris* (Zhang *et al.*, 2023), *M. confluentis*, *M. luteus* (Liu *et al.*, 2020), *M. polysiphoniae*

(Nedashkovskaya *et al.*, 2007) and *M. sedimenticola*. All of them have multiple GH43's and at least one GH51 or GH105 encoded. As aforementioned *M. sedimenticola* has 2 PULs, predicted PUL 5 and 15, that are homologues to PUL 7 and 8 in *M.* sp. MAR_2009_72. *M. polysiphoniae* has, according to PULDB, three PULs encoding a GH43, predicted PUL 2, 13 and 22. Predicted PUL 2 has several homologues in PULDB, one has been described in the literature as pectin PUL from *Flavobacterium johnsoniae* UW101 ATCC17061 (McBride *et al.*, 2009). Predicted PUL 13 has no homologues. Predicted PUL 22 has some homologues in marine and some gut associated bacteria. A homologous PUL from *Bacteroides intestinalis* DSM17393 has been published as an arabinoxylan PUL (Wang *et al.*, 2016). The other mentioned potential arabinogalactan degrading Maribacter strains, *M. aquimaris*, *M. confluentis* and *M. luteus* are not included in PULDB and therefore no comparison study could be done. Overall, it seems that arabinogalactan is not a common substrate for *Maribacter* strains, rather specific to certain strains of the genus. There is no common denominator between the potential arabinogalactan degraders, based on the information gathered in this discussion (Table 1).

Table 2: Overview of 35 *Maribacter* strains presenting specific knowledge about CAZymes and PULs in regards to arabinogalactan, annotated by dbCAN3 (Zheng *et al.*, 2023). Possible annotation of the mentioned CAZymes are the following GH43: α-L-arabinofuranosidase; GH51: α-L-arabinofuranosidase; GH105: rhamnogalacturonyl hydrolase; GH3: galactosidase; GH29: α-L-fucosidase GH31: galactosidase; GH67: α-glucuronidase; GH78: α-L-rhamnosidase and GH115: xylan-α-1,2-glucuronidase. Some of these CAZyme groups have multiple annotations; the most common ones are mentioned here.

Strain	CAZymes	sulfatases	GH43	GH51	GH105	GH3	GH10	GH29	GH31	GH67	GH78	GH115
M. aestuari GY20	94	9	0	0	0	3	0	0	1	1	0	0
M. algarum RZ26	156	37	0	0	1	5	0	4	1	0	3	0
M. algicola PoM-212	130	19	3	1	0	4	1	2	1	1	4	0
M. antarcticus DSM 21422	129	25	1	0	1	3	0	3	2	0	5	0
M. aquimaris ANRC-HE7	235	26	7	0	4	11	7	7	4	0	1	2
M. aquivivus DSM 16478	141	12	1	0	0	3	1	4	1	0	1	0
M. arcticus DSM 23546	123	8	2	1	0	4	1	2	1	1	1	0
M. arenosus CAU 1321	172	15	2	1	1	9	0	4	1	1	1	0
M. aurantiacus KCTC 52409	125	13	4	1	0	3	1	1	1	1	1	0
M. caenipelagi CECT 8455	130	2	3	1	0	3	1	1	1	1	2	0
M. cobaltidurans B1	133	4	7	1	0	3	3	1	2	2	0	1
M. confluentis CECT 8869	167	6	8	1	5	3	2	2	2	1	1	1
M. dokdonensis MAR_2009_60	144	20	1	0	0	4	1	5	1	0	2	2
M. dokdonensis MAR_2009_71	146	28	1	0	0	5	1	7	1	0	1	0
M. dokdonensis DSW-8	144	21	2	0	0	4	1	5	1	0	2	2
M. flavus KCTC 42508	130	17	4	1	0	3	1	1	1	1	0	0
M. forsetii DSM 18668	119	12	0	0	0	4	0	1	1	0	2	2
M. halichondriae Hal144	120	25	1	0	0	3	1	1	1	0	0	0
M. huludaoensis M208	143	24	1	0	1	3	1	2	1	0	4	0
M. hydrothermalis T28	116	8	1	0	0	2	0	1	2	0	1	0
M. litopenaei HL-LV01	88	9	3	0	0	3	0	1	0	1	1	0
M. litoralis 151	146	21	0	0	0	3	0	5	1	0	2	2
M. luteus RZ05	240	30	10	0	5	12	4	8	3	1	6	2
M. orientalis DSM 16471	132	8	4	1	0	4	3	1	2	1	1	0
M. polysaccharolyticus D37	224	51	6	0	0	10	0	18	2	0	1	1
M. polysiphoniae DSM 23514	207	18	13	1	2	11	3	3	5	1	1	2
M. sedimenticola DSM 19840	151	7	4	1	4	5	1	2	2	1	2	2
M. sp. Hel_I_7	132	8	1	0	0	5	1	3	2	1	2	2
M. sp. MAR_2009_72	151	14	5	1	4	5	1	2	2	1	7	2
M. spongiicola DSM 25233	143	24	0	0	1	3	0	4	1	0	5	0
M. stanieri DSM 19891	142	14	1	0	1	5	1	3	1	1	6	2
M. thermophilus HT7-2	115	11	2	0	0	3	0	0	1	0	0	0
M. ulvicola DSM 15366	136	16	0	0	0	4	0	3	3	0	2	0
M. vaceletii DSM 25230	164	50	0	0	1	3	0	17	1	0	0	0
M. zhoushanensis SA7	111	2	0	0	0	3	0	0	1	0	2	0

Having already begun the comparison of genomes, I conducted a more detailed analysis of the genomes of *M. forsetii*, *M.* sp. MAR_2009_72, *M. dokdonensis* MAR_2009_60 and MAR_2009_71 (Figure 1). They all have a similar genome size, ranging from 4.35 Mb (MAR_2009_72) to 4.62 Mb (MAR_2009_71) (Table1). A big difference between the strains is that *M. forsetii* has 25 less CAZymes, and as well a lower number in PULs, only 6. *M.* sp. MAR_2009_72 has 151 CAZymes and 9 PULs. Both *M. dokdonensis* strains have 8 PULs and MAR_2009_60 has 144 and MAR_2009_71 146 CAZymes. All mentioned PULs have at least one SusC/D pair and one CAZyme. One thing that stands out especially looking at Figure 1 is, that most CAZymes and SusC/D pairs are located outside of PULs. Due to this I suggest for future studies to expand the focus from PULs to all CAZymes in the genome to not miss any possible utilization mechanism.
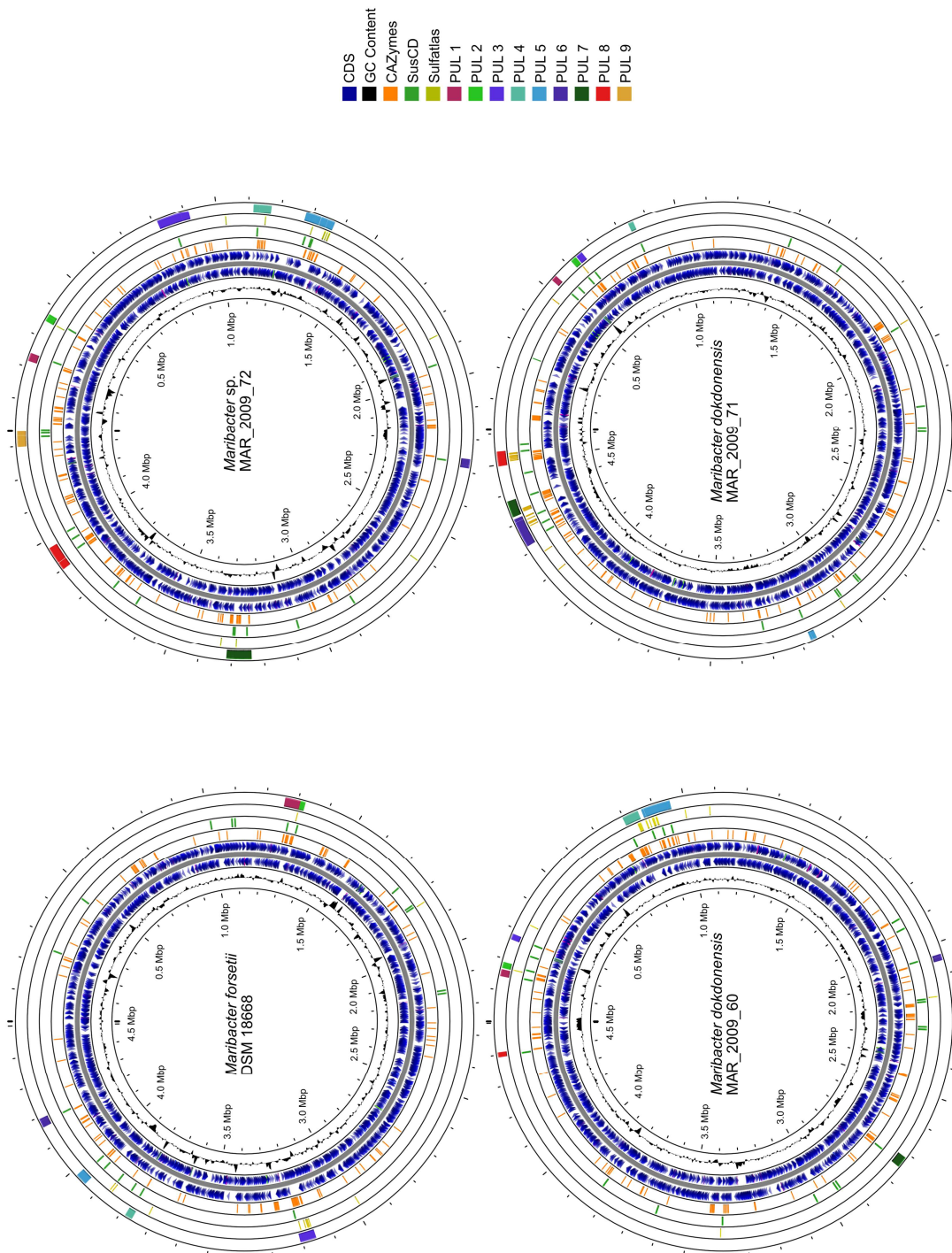
Figure 1: Full genome overview of *Maribacter forsetii*, *Maribacter* sp. MAR_2009_72, *Maribacter dokdonensis* MAR_2009_60 and MAR_2009_71 showcasing the GC content (ring one (most inner ring)), all annotated coding genes (CDS, ring 2 and 3) in forward and reverse direction, CAZymes identified by dbCAN3 (ring 4), SusC/D (ring 5), sulfatases (ring 6), and PULs (ring 7).

5.1 Maribacter diversity

In addition to comparing the genomes of these four strains, I aimed to explore the kinship of the 35 *Maribacter* strains listed in Table 1. Therefore, I calculated a phylogenetic tree based on their 16S rRNA gene sequences (Figure 2). The tree visualizes two main branches. One branch shows the *Maribacter* strains that are considered being Clade A, B and C, according to GTDB genome analyses they present separate genera (Parks *et al.*, 2022), however this finding has currently no standing in nomenclature. Clade A includes *M. arenoseus* (Thongphrom *et al.*, 2016), *M. polysiphoniae*, *M. aquimaris*, *M. luteus* and *M. polysaccharolyticus* (Gao *et al.*, 2023). In GTDB multiple unclassified *Maribacter* genomes are as well included in Clade A. Clade B has only one characterized strain/species *M. vaceletii*. The same is true for Clade C, *M. algarum* (Zhang *et al.*, 2020). A specific trait for Clade A are high numbers of CAZymes encoded in their genomes (Table 1). The usual range is between 88 to 167 CAZymes (Table 1). Strains of Clade A contain from 172 CAZymes up to 240. The strains of Clade B and C have more sulfatases than the other strains, *M. vaceletii* has 50 and *M. algarum* 37. On the other hand, *M. vaceletii* has the lowest GC percentage, 31.7%, of the analyzed *Maribacter* strains. Overall, the GC content has a 10% range in the genus *Maribacter* with 31.7% being the lowest and 41.8% being the highest GC content in *M. aurantiacus* and *M. flavus*. The second main branch carries the majority of *Maribacter* strains. There are several subbranches, such as the one with *M. arcticus* and *M. antarcticus* (Zhang *et al.*, 2009). This subbranch is specific towards polar *Maribacter* isolates. Some strains seem to be closer related to one another based on the branch length in the phylogenetic tree. For example, *M. confluentis* and *M. sedimenticola* have a short branch length. Furthermore, calculating the average nucleotide identity (ANI) for these strains shows an identity above 97% (Figure 2 and Supplement Figure 1). According to several publications strains that have an ANI above 95% should be considered the same species (Arahal, 2014; Ciufo *et al.*, 2018; Jain *et al.*, 2018). According to GTDB *M. confluentis* is a strain of the species *M. sedimenticola*, which is based on their ANI identity values true. The same statement can be made for *M.* sp. MAR_2009_72. It has an ANI of 97.83% to *M. sedimenticola* and should therefore as well be considered a strain of the *M. sedimenticola* species

(Supplement Figure 2). The strains *M. flavus* and *M. aurantiacus* have an ANI identity of 96.59%. (Supplement Figure 1). Therefore *M. aurantiacus* should be considered as a strain of the species *M. flavus*. A comparable conclusion can be drawn for *M.* sp. Hel_I_7 and *M. stanieri*, which have an ANI of 98.07%. *M.* sp. Hel_I_7 should be renamed to *M. stanieri* Hel_I_7, based on their ANI identity value (Supplement Figure 2).

The three *M. dokdonensis* strains are located on one branch in the phylogenetic tree, which is based on their species affiliation what I expected. The ANI identity values support this strain and species description, since the identity values are above 97%.

All other *Maribacter* strains have ANI values below 95% and therefore the species definition is correct (Supplement Figure 1 and 2).
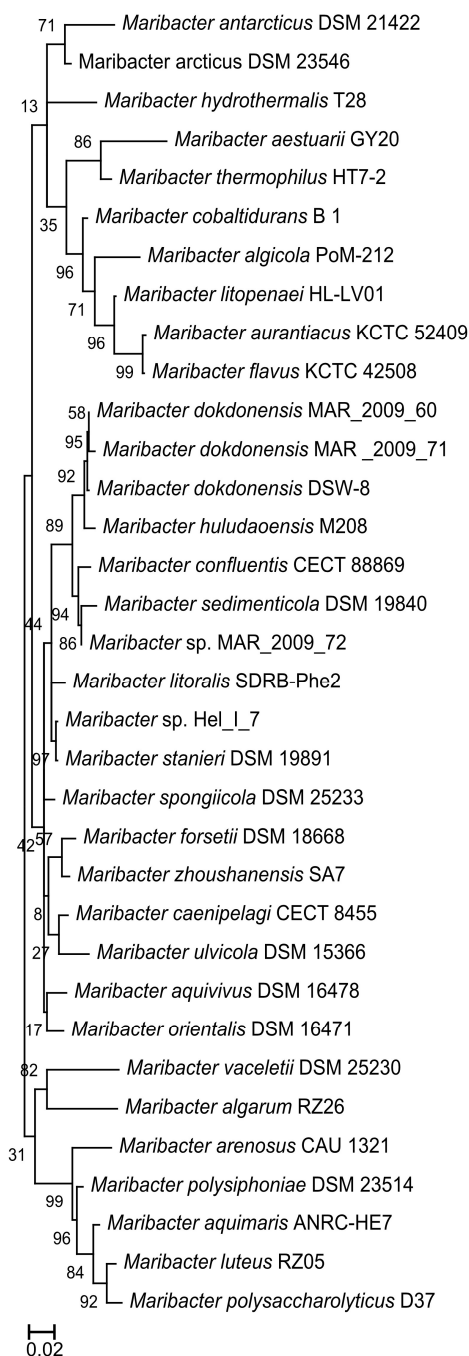
Figure 2: Phylogenetic tree showing the likeliness of the genus *Maribacter*. Only *Maribacter* strains/species included that are validly described and the four *Maribacter* strains included in this thesis. The strain nomenclature was used as provided in NCBI (July 2024). The evolutionary history was inferred by using the Maximum Likelihood method and Kimura 2-parameter model. The tree with the highest log likelihood (-5503.26) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (4 categories (+*G*, parameter = 0.1000)). The rate variation model allowed for some sites to be evolutionarily invariable ([+*I*], 43.05% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site.

5.2 Diversity of PULs

Even though I previously mentioned that we should not exclusively focus on PULs, it is still an interesting factor in strain comparisons to see how similar or different the PUL repository is among strains of the same genus. In Chapter 4 I already discussed the PULs of *M. forsetii* and in Chapter 3 the ones of *M.* sp. MAR_2009_72 affiliated with arabinogalactan. Here I included all 14 validly described strains of *Maribacter* analyzed by PULDB, as well as the four *Maribacter* strains previously mentioned in this thesis (Table 3). The majority of PUL structures are specific to one or two strains. Other PUL structures, such as 2, are present in almost all strains (Supplement Table 1). This PUL structure was described in the literature for *Zobellia galactivorans* Dsij[T] as a maltose PUL (Barbeyron *et al.*, 2016b). Number 3 has as well a couple of homologues in *Maribacter* strains and in other *Flavobacteriia*, according to PULDB, it is described in the literature as an alginate PUL in *Christiangramia* (former *Gramella*) *forsetii* (Kabisch *et al.*, 2014). A similar observation was obtained for structure 4, many homologues are encoded in the genus *Maribacter* and other *Flavobacteriia*. This structure has been described in the literature, but so far, the targeted polysaccharide remains unknown. The authors hypothesize the polysaccharide must be sulfated due to the sulfatases in the PUL (Barbeyron *et al.*, 2016b; Lapébie *et al.*, 2019). PUL structure 5 seems to be specific of strains of the *Maribacter* species *dokdonensis*. A possible substrate for this PUL could be sulfated agar, as the GH16_16 is annotated as a β-agarase (Hehemann *et al.*, 2012; Viborg *et al.*, 2019). Another PUL structures that stands out is number 10. It is present in 12 *Maribacter* strains and has been described as an arabinoxylan or plain xylan PUL in *Zobellia galactivorans* Dsij[T], *Christiangramia* (former *Gramella*) *flava* and gut associated bacteria (Martens *et al.*, 2011; Rogowski *et al.*, 2015; Barbeyron *et al.*, 2016b; Tang *et al.*, 2017). PUL structure 18 has homologs in *M. antarcticus*, *M. arcticus*, *M. cobaltidurans* and *M. ulvicola*. Moreover, it has been mentioned in publications as a potential fucose containing polysaccharide PUL (McNulty *et al.*, 2013; Barbeyron *et al.*, 2016b). *M. dokdonensis* DSW-8 (Yoon *et al.*, 2005) and *M. polysiphoniae* share PUL structure 25, which has been identified as a laminarin PUL in other marine *Bacteroidota* (Mystkowska *et al.*, 2018). Another PUL structure (No.39) identified in *M. polysiphoniae* has been

described as a pectin PUL in *Flavobacterium johnsoniae* UW101 ATCC 17061 (McBride *et al.*, 2009). *M. vaceletii* seems to encode a laminarin PUL, based on homologues in *Bacteroides thetaiotaomicron* 7330 and *Bacteroides cellulosilyticus* WH2 (McNulty *et al.*, 2013; Wu *et al.*, 2015). Furthermore *M. vaceletii* encodes the PUL structure 48, which is as well present in *Christiangramia* (former *Gramella*) *flava* JLT2011 as a described β-1,4 mannan PUL (Tang *et al.*, 2017). A third PUL, that has been described in the literature with substrate specificity, found in *M. vaceletii* is PUL structure 53, annotated for the utilization of agar (Barbeyron *et al.*, 2016b). Two more PULs encoded in *M. vaceletii* were identified in other Bacteroidota, but the substrate was not declared (Martens *et al.*, 2011; McNulty *et al.*, 2013). All of the other PULs presented in Supplement Table 1 have not been described in the literature so far. Some conclusions based on the encoded CAZymes can be drawn for some PULs, such as structure 5 as I aforementioned a potential PUL for agar utilization. But some PULs are complex, due to high numbers of CAZymes that a clear conclusion is not possible.

Table 3: PUL comparison of *Maribacter* strains analyzed in PULDB and validly described. The numbers match the numbering system of Supplement Table 1 presenting the PUL structures. Green symbolizes the presence of the PUL in the specific *Maribacter* strain. For the strains that were in depth analyzed during this thesis *M. forsetii*, MAR_2009_60/71/72 the PUL number is mentioned according to Figure 1.

Strain	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55
M.antarcticus DSM 21422																																																							
M.aquivivus DSM 16480			■							■									■	■	■																																		
M.arcticus DSM 23547		■															■	■																																				
M.cobaltidurans B8		■																				■	■	■	■																														
M.dokdonensis DSW-8		■																																												■									
M.dokdonensis MAR_2009_60			PUL 8	PUL 3	PUL 2	PUL 5	PUL 4	PUL 7	PUL 1	PUL 6																																													
M.dokdonensis MAR_2009_71			PUL 4	PUL 1	PUL 2	PUL 6	PUL 7	PUL 3		PUL 5	PUL 8																																												
M.forsetii DSM 18668			PUL 6	PUL 4	PUL 2	PUL 3	PUL 1	PUL 5																				■																										
M.hydrothermalis T28																									■			■	■																										
M.litoponaei HL-LV01																																																							
M.orientalis DSM 16471		■						■		■									■					■							■		■				■	■	■	■	■														
M.polysiphoniae DSM 23514																																																								
M.sedimenticola DSM 19840						■		■	PUL 4	PUL 1																																														
M.sp. Hel_1_7												PUL 5	PUL 7	PUL 8		PUL 9																																							
M.sp. MAR_2009_72	PUL 3	PUL 6		PUL 2																																																			
M.stanieri DSM 19891																																																							
M.ulvicola DSM 15366		■		■														■												■																									
M.vaceletii DSM 25230		■						■																											■									■	■										■

5.3 Outlook

For the future I would as aforementioned expand the focus from solely PULs to all CAZymes in a genome or MAG. Furthermore, not only the CAZymes in a PUL can provide information about the potential substrate, some PULs like PUL 8 in *M.* sp. MAR_2009_72 might encode enzymes that are part of the utilization process in the cytoplasm (Chapter 3). We all need to start looking at the whole genome of a bacterium and not pick out some parts and then hypothesize about their abilities. Many different factors come in to play for the utilization of sugars and other compounds, we need to make sure to get the whole picture and not just a small part of it.

To identify natural substrates of *Maribacter* strains and other *Bacteroidota* I would continue working with natural biomass. The methods I used needs to be improved to identify more proteins and to receive constant numbers. Potentially different equipment during the measurement in the LC-MS/MS could improve the number of detected proteins as well. In general, using proteomics seems to be a good methodology for these kinds of projects, to get an overview of the expressed proteins under certain conditions. But for future projects I would also include gene expressions experiments and more enzyme activity tests to support the proteomic data.

Overall, this thesis provided more insight in to the utilization abilities of *Maribacter* strains. It not only showed that not every saccharide is utilized by enzymes encoded in PULs, but as well that many saccharides are still to be discovered and analyzed to gain knowledge how bacteria degrade them. Marine bacteria and their involvement in the carbon cycle are getting more and more attention and we are making steps in to the right direction to understand some of the process they are involved in, but still a lot is yet to be discovered, leaving room for many more projects to investigate the utilization of polysaccharides.

5.4 Supplement

5.4.1 Supplement Methods

Information about the *Maribacter* species/strains were taken from NCBI and LPSN, as well as GTDB (Parte, 2014; Coordinators, 2018; Parks *et al.*, 2022). CAZymes were considered to be identified if two out of three search algorithms in dbCAN3 were positive in the web interface search (Zheng *et al.*, 2023). The SulfAtlas web interface (Stam *et al.*, 2022), PULDB (Terrapon *et al.*, 2018) and BLAST (Altschul *et al.*, 1990) provided additional information. The PUL structures of Supplement Table 1 were acquired from PULDB and used as search patterns to identify homologues in PULDB.

Alignment for the *Maribacter* tree was performed using MAFFT online in automode (Katoh *et al.*, 2019). The 16S rRNA gene sequences were acquired from NCBI (Coordinators, 2018). The evolutionary history was inferred by using the Maximum Likelihood method and Kimura 2-parameter model (Tamura *et al.*, 2021). The tree with the highest log likelihood (-5503.26) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (4 categories (+*G*, parameter = 0.1000)). The rate variation model allowed for some sites to be evolutionarily invariable ([+*I*], 43.05% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site.

For the genome visualization the online tool Proksee was used (Grant *et al.*, 2023). The average nucleotide identity (ANI) calculations were done using the FastANI calculator of GTDB and the *Maribacter* genome files of GTDB (Jain *et al.*, 2018).
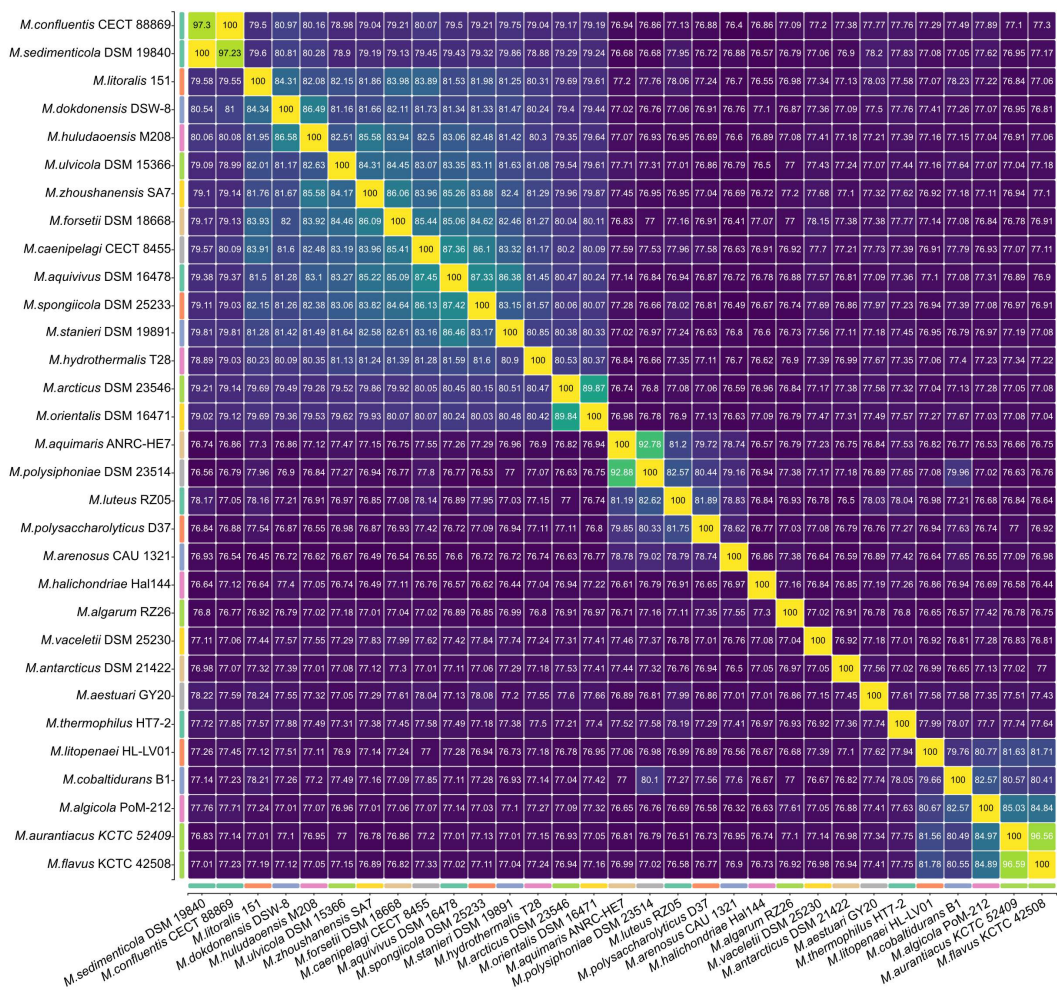
5.4.2 Supplement Tables

Supplement Table 1: PUL structures identified in the genus *Maribacter* by PULDB (Terrapon *et al.*, 2018), that have at least one SusC/D pair and one CAZyme annotated. If literature derived homologues were identified the reference is mentioned as well as the substrate of the literature derived PUL. The assigned numbers to the PUL structures match the numbering system of Table 3.

No.	PUL structure	Literature homologues	Substrate
1	GH13_9+GH28+GH92+SusC+SusD+GH23+GT2+GH15+GT20	McNulty *et al.* (2013)	Not defined
2	GH13_38+GH65+SusC+SusD	Barbeyron *et al.* (2016b)	Maltose
3	PL7+PL7+SusD+SusC	Kabisch *et al.* (2014)	Alginate
4	SusC+SusD+GH144+GH144+GH3+Sulf_1	Barbeyron *et al.* (2016b); Lapébie *et al.* (2019)	Unknown
5	GH85+GH117+SusC+SusD+GH16_16		
6	GH95+PL40+PL43+GH117+PL33_2+SusC+GH29+PL40+GH97+GH88+SusC+SusD		
7	GH78+SusC+SusD+GH92+GH29+GH3+GH115+GH88		
8	SusD+SusC+PL17_2+PL6_1+PL12+PL6_1		
9	SusD+SusC+GH20+CBM9+CBM6+GH29+GH130+CE20+GH109		
10	GH43_1+GH10+SusD+SusC	Martens *et al.* (2011); Rogowski *et al.* (2015); Barbeyron *et al.* (2016b); Tang *et al.* (2017)	Arabinoxylan, Xylan
11	Sulf_1+GH20+GH141+GH29+GH172+GH92+Sulf_1+GH20+Sulf_1+GH3+Sulf_1+Sulf_1+GH95+Sulf_1+Sulf_1+Sulf_1+GH3+GH29+Sulf_1+GH29+Sulf_1+SusD+SusC		
12	SusC+SusD+CBM67+GH5+CBM67+GH78+SusD+SusC+SusD+SusC+CBM6+GH3+GH138+GH109+GH78+GH106+GH78		
13	GH2+CE12+GH115+GH28+Sulf_1+GH105+SusD+SusC+SusD+SusC+GH31+GH106+GH105+GH43_18+GH78		
14	PL1_2+GH140+GH43_10+GH28+SusC+SusD+GH105+GH105+PL10+CE12+GH43_19+GH43_34+GH51		
15	Sulf_1+PL43+Sulf_1+Sulf_1+Sulf_1+PL40+PL33_2+GH29+PL40+GH97+GH88+SusC+SusD		
16	GH3+GH10+SusC+SusD+SusC+SusD		
17	SusC+SusD+GH16_3+GH149+GH17+GH17+GH30_1+GH17		
18	SusC+SusD+GH95	McNulty *et al.* (2013); Barbeyron *et al.* (2016b)	Fucose containing polysaccharide
19	SusC+SusD+GH31		
20	Sulf_1+GH28+Sulf_1+GH16_11+GH117+GH2+GH29+GH29+Sulf_1+GH29+Sulf_1+SusC+SusD+GH16_11+GH29		
21	GH2+SusC+SusD+GH16_16		
22	SusC+SusD+GH16_16+GH88		
23	GH2+SusC+SusD+GH43_18+GH43_26+GH43_26+GH43_33+GH5_13+GH92		
24	GH115+GH10+GH43_1+SusC+SusD		
25	SusC+SusD+GH30+GH16_3	Mystkowska *et al.* (2018)	Laminarin
26	SusC+SusD+GH117		
27	SusC+SusD+GH+GH3		
28	SusC+SusD+Sulf_1+Sulf_1+GH2		
29	SusC+SusD+GH32		

30	SusC+SusD+GH92+GH28		
31	SusC+SusD+GH146		
32	SusC+SusD+GH47+GH92+GH92+GH92+GH130_2+GH92		
33	SusC+SusD+Sulf_1+Sulf_1+GH20		
34	SusC+SusD+Sulf_1+GH9		
35	SusC+SusD+GH92+GH29+GH3+GH115+GH9+GH3		
36	SusC+SusD+GH76		
37	SusC+SusD+GH125		
38	SusC+SusD+GH127		
39	GH28+SusC+SusD+GH105+GH105+GH43_10+PL10_1	McBride *et al.* (2009)	Pectin
40	SusC+SusD+GH10+GH10+GH43_12+GH43_1+GH10		
41	SusC+SusD+GH130_4+GH30_1		
42	GH63+GH2+SusC+SusD+GH+GH117		
43	SusC+SusD+GH2+GH95+GH117+GH+Sulf_1		
44	PL7+PL7+SusC+SusD+GH95+GH20+GH172+GH36+SusC+SusD		
45	SusC+SusD+GH16_3+GH3	McNulty *et al.* (2013); Wu *et al.* (2015)	Laminarin
46	SusC+SusD+SusC+SusD+GH117+GH117+GH16_15		
47	Sulf_1+GH27+Sulf_1+GH150+GH150+Sulf_1+Sulf_1+SusC+SusD		
48	SusC+SusD+GH173+GH130_5+GH173+GH5_7	Tang *et al.* (2017)	β-1,4 mannan
49	PL29+Sulf_1+GH29+SusC+SusD+GH88+PL29+Sulf_1		
50	SusC+SusD+GH35		
51	SusC+SusD+GH36	Martens *et al.* (2011); McNulty *et al.* (2013)	
52	SusC+SusD+GH2	Martens *et al.* (2011); McNulty *et al.* (2013)	
53	GH117+GH2+SusC+SusD	Barbeyron *et al.* (2016b)	Agar
54	SusC+SusD+GH129+GH2		
55	SusC+SusD+GH97+Sulf_1+GH2+Sulf_1+GH2+GH2+GH117+GH117+GH117+GH127+GH127+Sulf_1+GH29+GH2+GH117+Sulf_1+Sulf_1+GH29+GH28+Sulf_1+GH28+Sulf_1+Sulf_1+Sulf_1+Sulf_1+GH117+SusC+SusD		

5.4.3 Supplement Figures

Supplement Figure 1: ANI values of all validly described *Maribacter* species type strains.

Supplement Figure 2: ANI values of *Maribacter* sp. MAR_2009_71, *Maribacter* sp. Hel_I_7, *Maribacter dokdonensis* MAR_2009_60 and MAR_2009_71 in comparison to all validly described *Maribacter* species type strains.

5.5 References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology* **215**: 403-410.

Arahal, D.R. (2014) Chapter 6 - Whole genome analyses: Average nucleotide identity. In *Methods in Microbiology*. Goodfellow, M., Sutcliffe, I., and Chun, J. (eds): Academic Press, pp. 103-122.

Barbeyron, T., Carpentier, F., Haridon, S., Schüler, M., Michel, G., and Amann, R. (2008) Description of *Maribacter forsetii* sp. nov., a marine *Flavobacteriaceae* isolated from North Sea water, and emended description of the genus *Maribacter*. *International Journal of Systematic and Evolutionary Microbiology* **58**: 790-797.

Barbeyron, T., Brillet-Guéguen, L., Carré, W., Carrière, C., Caron, C., Czjzek, M. *et al.* (2016a) Matching the diversity of sulfated biomolecules: Creation of a classification database for sulfatases reflecting their substrate specificity. *PLOS ONE* **11**: e0164846.

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* **18**: 4610-4627.

Cartmell, A., Muñoz-Muñoz, J., Briggs, J.A., Ndeh, D.A., Lowe, E.C., Baslé, A. *et al.* (2018) A surface endogalactanase in *Bacteroides thetaiotaomicron* confers keystone status for arabinogalactan degradation. *Nature Microbiology* **3**: 1314-1326.

Chen, J., Robb, C.S., Unfried, F., Kappelmann, L., Markert, S., Song, T. *et al.* (2018) Alpha- and beta-mannan utilization by marine *Bacteroidetes*. *Environmental Microbiology* **20**: 4127-4140.

Cho, K.H., Hong, S.G., Cho, H.H., Lee, Y.K., Chun, J., and Lee, H.K. (2008) *Maribacter arcticus* sp. nov., isolated from arctic marine sediment. *International Journal of Systematic and Evolutionary Microbiology* **58**: 1300-1303.

Ciufo, S., Kannan, S., Sharma, S., Badretdin, A., Clark, K., Turner, S. *et al.* (2018) Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI. *International Journal of Systematic and Evolutionary Microbiology* **68**: 2386-2392.

Coordinators, N.R. (2018) Database resources of the national center for biotechnology information. *Nucleic acids research* **46**: D8-D13.

Fang, C., Wu, Y.-H., Xamxidin, M., Wang, C.-S., and Xu, X.-W. (2017) *Maribacter cobaltidurans* sp. nov., a heavy-metal-tolerant bacterium isolated from deep-sea sediment. *International Journal of Systematic and Evolutionary Microbiology* **67**: 5261-5267.

Fujita, K., Sasaki, Y., and Kitahara, K. (2019) Degradation of plant arabinogalactan proteins by intestinal bacteria: Characteristics and functions of the enzymes involved. *Applied Microbiology and Biotechnology* **103**: 7451-7457.

Gao, J.-W., Ying, J.-J., Dong, H., Liu, W.-J., He, D.-Y., Xu, L., and Sun, C. (2023) Characterization of *Maribacter polysaccharolyticus* sp. nov., *Maribacter huludaoensis* sp. nov., and *Maribacter zhoushanensis* sp. nov. and illumination of the distinct adaptative strategies of the genus *Maribacter*. *Frontiers in Marine Science* **10**: 1248754

Giljan, G., Arnosti, C., Kirstein, I.V., Amann, R., and Fuchs, B.M. (2022) Strong seasonal differences of bacterial polysaccharide utilization in the North Sea over an annual cycle. *Environmental Microbiology* **24**: 2333-2347.

Grant, J.R., Enns, E., Marinier, E., Mandal, A., Herman, E.K., Chen, C.-y. *et al.* (2023) Proksee: In-depth characterization and visualization of bacterial genomes. *Nucleic Acids Research* **51**: W484-W492.

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* **266**: 194-200.

Hahnke, R.L., and Harder, J. (2013) Phylogenetic diversity of *Flavobacteria* isolated from the North Sea on solid media. *Systematic and Applied Microbiology* **36**: 497-504.

Hehemann, J.-H., Correc, G., Thomas, F., Bernard, T., Barbeyron, T., Jam, M. *et al.* (2012) Biochemical and structural characterization of the complex agarolytic enzyme system from the marine bacterium *Zobellia galactanivorans*. *Journal of Biological Chemistry* **287**: 30571-30584.

Heins, A., and Harder, J. (2023) Particle-associated bacteria in seawater dominate the colony-forming microbiome on ZoBell marine agar. *FEMS Microbiology Ecology* **99**: fiac151.

Heins, A., Amann, R.I., and Harder, J. (2021a) Cultivation of particle-associated heterotrophic bacteria during a spring phytoplankton bloom in the North Sea. *Systematic and Applied Microbiology* **44**: 126232.

Heins, A., Reintjes, G., Amann, R.I., and Harder, J. (2021b) Particle collection in Imhoff sedimentation cones enriches both motile chemotactic and particle-attached bacteria. *Frontiers in Microbiology* **12**: 643730.

Hu, J., Yang, Q.-Q., Ren, Y., Zhang, W.-W., Zheng, G., Sun, C. *et al.* (2015) *Maribacter thermophilus* sp. nov., isolated from an algal bloom in an intertidal zone, and emended description of the genus *Maribacter*. *International Journal of Systematic and Evolutionary Microbiology* **65**: 36-41.

Jackson, S.A., Kennedy, J., Morrissey, J.P., apos, Gara, F., and Dobson, A.D.W. (2015) *Maribacter spongiicola* sp. nov. and *Maribacter vaceletii* sp. nov., isolated from marine sponges, and emended description of the genus *Maribacter*. *International Journal of Systematic and Evolutionary Microbiology* **65**: 2097-2103.

Jain, C., Rodriguez-R, L.M., Phillippy, A.M., Konstantinidis, K.T., and Aluru, S. (2018) High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nature Communications* **9**: 5114.

Jung, Y.-T., Lee, J.-S., and Yoon, J.-H. (2014) *Maribacter caenipelagi* sp. nov., a member of the *Flavobacteriaceae* isolated from a tidal flat sediment of the Yellow Sea in Korea. *Antonie van Leeuwenhoek* **106**: 733-742.

Kabisch, A., Otto, A., König, S., Becher, D., Albrecht, D., Schueler, M. *et al.* (2014) Functional characterization of polysaccharide utilization loci in the marine *Bacteroidetes* 'Gramella forsetii' KT0803. *The ISME Journal* **8**: 1492–1502.

Kalenborn, S., Zühlke, D., Riedel, K., Amann, R.I., and Harder, J. (2024) Proteomic insight into arabinogalactan utilization by particle-associated *Maribacter* sp. MAR_2009_72. *FEMS Microbiology Ecology* **100**: fiae045.

Kappelmann, L., Krüger, K., Hehemann, J.-H., Harder, J., Markert, S., Unfried, F. *et al.* (2019) Polysaccharide utilization loci of North Sea *Flavobacteriia* as basis for using SusC/D-protein expression for predicting major phytoplankton glycans. *The ISME Journal* **13**: 76-91.

Katoh, K., Rozewicki, J., and Yamada, K.D. (2019) MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics* **20**: 1160-1166.

Khan, S.A., Jeong, S.E., Baek, J.H., and Jeon, C.O. (2019) *Maribacter algicola* sp. nov., isolated from a marine red alga, *Porphyridium marinum*, and transfer of *Maripseudobacter aurantiacus* Chen et al. 2017 to the genus *Maribacter* as *Maribacter aurantiacus* comb. nov. *International Journal of Systematic and Evolutionary Microbiology* **70**: 797-804.

Kim, S.Y., Choi, J.Y., Hong, Y.W., Shin, D.Y., Kim, B.J., Kang, J.K. *et al.* (2023) *Maribacter litopenaei* sp. nov., isolated from the intestinal tract of the Pacific white shrimp *Litopenaeus vannamei*. *International Journal of Systematic and Evolutionary Microbiology* **73**: 3.

Lapébie, P., Lombard, V., Drula, E., Terrapon, N., and Henrissat, B. (2019) *Bacteroidetes* use thousands of enzyme combinations to break down glycans. *Nature Communications* **10**: 2043.

Lauro, F.M., McDougald, D., Thomas, T., Williams, T.J., Egan, S., Rice, S. *et al.* (2009) The genomic basis of trophic strategy in marine bacteria. *Proceedings of the National Academy of Sciences* **106**: 15527-15533.

Lee, D.W., Lee, H., Kwon, B.-O., Khim, J.S., Yim, U.H., Park, H. *et al.* (2018) *Maribacter litoralis* sp. nov. a marine bacterium isolated from seashore. *International Journal of Systematic and Evolutionary Microbiology* **68**: 3471-3478.

Leszczuk, A., Kalaitzis, P., Kulik, J., and Zdunek, A. (2023) Review: Structure and modifications of arabinogalactan proteins (AGPs). *BMC Plant Biology* **23**: 45.

Lin, H., Lin, D., Zhang, M., Ye, J., Sun, J., and Tang, K. (2021) *Maribacter hydrothermalis* sp. nov., isolated from shallow-sea hydrothermal systems off Kueishantao Island. *Current Microbiology* **78**: 2815-2820.

Liu, A., Zhang, Y.-J., Liu, D.-K., and Li, X.-Z. (2020) *Maribacter luteus* sp. nov., a marine bacterium isolated from intertidal sand of the Yellow Sea. *International Journal of Systematic and Evolutionary Microbiology* **70**: 3497-3503.

Lo, N., Jin, H.M., and Jeon, C.O. (2013) *Maribacter aestuarii* sp. nov., isolated from tidal flat sediment, and an emended description of the genus *Maribacter*. *International Journal of Systematic and Evolutionary Microbiology* **63**: 3409-3414.

Lu, D.-C., Wang, F.-Q., Amann, R.I., Teeling, H., and Du, Z.-J. (2023) Epiphytic common core bacteria in the microbiomes of co-located green (*Ulva*), brown (*Saccharina*) and red (*Grateloupia*, *Gelidium*) macroalgae. *Microbiome* **11**: 126.

Luis, A.S., Briggs, J., Zhang, X., Farnell, B., Ndeh, D., Labourel, A. *et al.* (2018) Dietary pectic glycans are degraded by coordinated enzyme pathways in human colonic *Bacteroides*. *Nature Microbiology* **3**: 210-219.

Martens, E.C., Chiang, H.C., and Gordon, J.I. (2008) Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. *Cell Host & Microbe* **4**: 447-457.

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* **9**: e1001221.

McBride, M., J., Xie, G., Martens Eric, C., Lapidus, A., Henrissat, B., Rhodes Ryan, G. *et al.* (2009) Novel features of the polysaccharide-digesting gliding bacterium *Flavobacterium johnsoniae* as revealed by genome sequence analysis. *Applied and Environmental Microbiology* **75**: 6864-6875.

McNulty, N.P., Wu, M., Erickson, A.R., Pan, C., Erickson, B.K., Martens, E.C. *et al.* (2013) Effects of diet on resource utilization by a model human gut microbiota containing *Bacteroides cellulosilyticus* WH2, a symbiont with an extensive glycobiome. *PLOS Biology* **11**: e1001637.

Miksch, S., Meiners, M., Meyerdierks, A., Probandt, D., Wegener, G., Titschack, J. *et al.* (2021) Bacterial communities in temperate and polar coastal sands are seasonally stable. *ISME Communications* **1**: 29.

Mystkowska, A.A., Robb, C., Vidal-Melgosa, S., Vanni, C., Fernandez-Guerra, A., Höhne, M., and Hehemann, J.-H. (2018) Molecular recognition of the beta-glucans laminarin and pustulan by a SusD-like glycan-binding protein of a marine *Bacteroidetes*. *The FEBS Journal* **285**: 4465-4481.

Ndeh, D., Rogowski, A., Cartmell, A., Luis, A.S., Baslé, A., Gray, J. *et al.* (2017) Complex pectin metabolism by gut bacteria reveals novel catalytic functions. *Nature* **544**: 65-70.

Nedashkovskaya, O.I., Kim, S.B., and Mikhailov, V.V. (2010) *Maribacter stanieri* sp. nov., a marine bacterium of the family *Flavobacteriaceae*. *International Journal of Systematic and Evolutionary Microbiology* **60**: 214-218.

Nedashkovskaya, O.I., Vancanneyt, M., De Vos, P., Kim, S.B., Lee, M.S., and Mikhailov, V.V. (2007) *Maribacter polysiphoniae* sp. nov., isolated from a red alga. *International Journal of Systematic and Evolutionary Microbiology* **57**: 2840-2843.

Nedashkovskaya, O.I., Kim, S.B., Han, S.K., Lysenko, A.M., Rohde, M., Rhee, M.-S. *et al.* (2004) *Maribacter* gen. nov., a new member of the family *Flavobacteriaceae*, isolated from marine habitats, containing the species *Maribacter sedimenticola* sp. nov., *Maribacter aquivivus* sp. nov., *Maribacter orientalis* sp. nov. and *Maribacter ulvicola* sp. nov. *International Journal of Systematic and Evolutionary Microbiology* **54**: 1017-1023.

Park, S., Jung, Y.-T., Park, J.-M., Won, S.-M., and Yoon, J.-H. (2015) *Maribacter confluentis* sp. nov., isolated from the junction between the ocean and a freshwater spring. *International Journal of Systematic and Evolutionary Microbiology* **65**: 3079-3085.

Parks, D.H., Chuvochina, M., Rinke, C., Mussig, A.J., Chaumeil, P.-A., and Hugenholtz, P. (2022) GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Research* **50**: D785-D794.

Parte, A.C. (2014) LPSN--list of prokaryotic names with standing in nomenclature. *Nucleic Acids Research* **42**: 613-616.

Pedler, B.E., Aluwihare, L.I., and Azam, F. (2014) Single bacterial strain capable of significant contribution to carbon cycling in the surface ocean. *Proceedings of the National Academy of Sciences* **111**: 7202-7207.

Rogowski, A., Briggs, J.A., Mortimer, J.C., Tryfona, T., Terrapon, N., Lowe, E.C. *et al.* (2015) Glycan complexity dictates microbial resource allocation in the large intestine. *Nature Communications* **6**: 7481.

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* **2**: 17065.

Stam, M., Lelièvre, P., Hoebeke, M., Corre, E., Barbeyron, T., and Michel, G. (2022) SulfAtlas, the sulfatase database: State of the art and new developments. *Nucleic Acids Research* **51**: 647-653.

Steiner, L.X., Wiese, J., Rahn, T., Borchert, E., Slaby, B.M., and Hentschel, U. (2024) *Maribacter halichondriae* sp. nov., isolated from the marine sponge *Halichondria panicea*, displays features of a sponge-associated life style. *Antonie van Leeuwenhoek* **117**: 56.

Tamura, K., Stecher, G., and Kumar, S. (2021) MEGA11: Molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution* **38**: 3022-3027.

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* **8**: 220.

Tang, M., Wang, G., Xiang, W., Chen, C., Wu, J., Dai, S. *et al.* (2015) *Maribacter flavus* sp. nov., isolated from a cyanobacterial culture pond. *International Journal of Systematic and Evolutionary Microbiology* **65**: 3997-4002.

Terrapon, N., Lombard, V., Drula, É., Lapébie, P., Al-Masaudi, S., Gilbert, H.J., and Henrissat, B. (2018) PULDB: The expanded database of polysaccharide utilization loci. *Nucleic Acids Research* **46**: D677-D683.

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* **8**: 1808.

Thongphrom, C., Kim, J.-H., and Kim, W. (2016) *Maribacter arenosus* sp. nov., isolated from marine sediment. *International Journal of Systematic and Evolutionary Microbiology* **66**: 4826-4831.

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* **12**: 2894–2906.

Viborg, A.H., Terrapon, N., Lombard, V., Michel, G., Czjzek, M., Henrissat, B., and Brumer, H. (2019) A subfamily roadmap of the evolutionarily diverse glycoside hydrolase family 16 (GH16). *J Biol Chem* **294**: 15973-15986.

Villa-Rivera, M.G., Cano-Camacho, H., López-Romero, E., and Zavala-Páramo, M.G. (2021) The role of arabinogalactan type II degradation in plant-microbe interactions. *Frontiers in Microbiology* **12**: 730543.

Wang, F.-Q., Bartosik, D., Sidhu, C., Siebers, R., Lu, D.-C., Trautwein-Schult, A. *et al.* (2024) Particle-attached bacteria act as gatekeepers in the decomposition of complex phytoplankton polysaccharides. *Microbiome* **12**: 32.

Wang, K., Pereira, G.V., Cavalcante, J.J.V., Zhang, M., Mackie, R., and Cann, I. (2016) *Bacteroides intestinalis* DSM 17393, a member of the human colonic microbiome, upregulates multiple endoxylanases during growth on xylan. *Scientific Reports* **6**: 34360.

Wang, Y., and LaPointe, G. (2020) Arabinogalactan utilization by *Bifidobacterium longum* subsp. *longum* NCC 2705 and *Bacteroides caccae* ATCC 43185 in monoculture and coculture. *Microorganisms* **8**: 1703.

Wu, M., McNulty, N.P., Rodionov, D.A., Khoroshkin, M.S., Griffin, N.W., Cheng, J. *et al.* (2015) Genetic determinants of in vivo fitness and diet responsiveness in multiple human gut *Bacteroides*. *Science* **350**: aac5992.

Yoon, J.-H., Kang, S.-J., Lee, S.-Y., Lee, C.-H., and Oh, T.-K. (2005) *Maribacter dokdonensis* sp. nov., isolated from sea water off a Korean island, Dokdo. *International Journal of Systematic and Evolutionary Microbiology* **55**: 2051-2055.

Zhang, G.I., Hwang, C.Y., Kang, S.-H., and Cho, B.C. (2009) *Maribacter antarcticus* sp. nov., a psychrophilic bacterium isolated from a culture of the Antarctic green alga *Pyramimonas gelidicola*. *International Journal of Systematic and Evolutionary Microbiology* **59**: 1455-1459.

Zhang, J.-Y., Xia, Y., Feng, X., Mu, D.-S., and Du, Z.-J. (2020) *Maribacter algarum* sp. nov., a new member of the family *Flavobacteriaceae* isolated from the red alga *Gelidium amansii*. **70**: 3679-3685.

Zhang, Y., Zhai, Y., Mu, L., Hu, M., Fang, W., Xiao, Y., and Fang, Z. (2023) *Maribacter aquimaris* sp. nov., isolated from seawater adjacent to Fildes Peninsula, Antarctica. *Antonie van Leeuwenhoek* **116**: 753-761.

Zheng, J., Ge, Q., Yan, Y., Zhang, X., Huang, L., and Yin, Y. (2023) dbCAN3: Automated carbohydrate-active enzyme and substrate annotation. *Nucleic Acids Research* **51**: W115-W121.

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* **7**: 1007-1015.

Acknowledgement

Firstly, I would like to say "Thank you" to my supervisor and reviewer Prof. Dr. Jens Harder. I am grateful for all your support, patience, advice and honesty during my doctoral thesis. Thank you for always having an open door to discuss research as well as life.

Thank you Prof. Dr. Rudolf Amann for giving me the opportunity to do my doctoral thesis in the Molecular Ecology and as well for the valuable discussions.

A big thank you to my second reviewer Dr. Mia M. Bengtsson, who was part of all my thesis committee meetings and provided helpful guidance over the years.

A special acknowledgment goes as well to PD Dr. Bernhard Fuchs, Prof. Dr. Michael Friedrich and Ana Carolina Bercini Gusmao for being part of my examination committee.

I would like to thank my collaborators, especially Dr. Daniela Zühlke from the University of Greifswald. Thank you as well to Dr. Maria Grimm, Dr. Madita Brauer, Jürgen Bartels, Robert Knop, Silvia Dittmann, Vaikhari Kale and Dirk Albrecht for helping me in the lab and allowing me to be part of the group during my short visits.

A special thanks to Sabine Kühn for all your technical assistance over the last years.

Thank you to my office mates Almud Lonsing, Dr. Anneke Heins and Dr. Greta Reintjes for the good laughs and talks when you were present in the office.

Furthermore, as well a special thanks to Mirja Meiners, Kathrin Büttner, Anja Greiser, Cäcilia Wigand, Tina Horstmann, Bruno Stahl, Katharina Föll and Christina Probian for making work more fun. Thank you for many good conversations.

Thank you to all of my friends who supported me throughout this time, especially Judith, Caro and Andrea.

Papa, auch wenn du dies leider nicht mehr miterleben kannst, möchte ich mich bei dir bedanken. Du hast mich mein ganzes Leben unterstützt und hast es mir möglich gemacht meine Träume zu verwirklichen. Danke für alles.

Danke Mama. Ohne dich hätte ich das Project PhD nicht mehr beendet. Du hast mir mein ganzes Leben lang und besonders in den letzten anderthalb Jahren Kraft und Halt gegeben. Danke dass du immer an mich glaubst.

Universität Bremen
Fachbereich 2
Prüfungsamt Chemie
z. Hd. Frau Frauke Ernst
Leobener Straße

28359 Bremen
Deutschland

Prüfungsamt
Chemie

Frauke Ernst
Geschäftsstelle
Fachbereich 02
Leobener Str. / NW2
D-28359 Bremen

Verwaltungspavillon 06
Tel. 0421 218-62802
Fax 0421 218-9862802
frauke.ernst@uni-bremen.de
www.fb2.uni-bremen.de

Versicherung an Eides Statt

Name, Vorname	Kalenborn, Saskia
Matrikel-Nr.	
Straße	
Ort, PLZ	28357, Bremen

Ich, Saskia Kalenborn,

versichere an Eides Statt durch meine Unterschrift, dass ich die vorstehende Arbeit selbständig und ohne fremde Hilfe angefertigt und alle Stellen, die ich wörtlich dem Sinne nach aus Veröffentlichungen entnommen habe, als solche kenntlich gemacht habe, mich auch keiner anderen als der angegebenen Literatur oder sonstiger Hilfsmittel bedient habe.
Ich versichere an Eides Statt, dass ich die vorgenannten Angaben nach bestem Wissen und Gewissen gemacht habe und dass die Angaben der Wahrheit entsprechen und ich nichts verschwiegen habe.
Die Strafbarkeit einer falschen eidesstattlichen Versicherung ist mir bekannt, namentlich die Strafandrohung gemäß § 156 StGB bis zu drei Jahren Freiheitsstrafe oder Geldstrafe bei vorsätzlicher Begehung der Tat bzw. gemäß § 161 Abs. 1 StGB bis zu einem Jahr Freiheitsstrafe oder Geldstrafe bei fahrlässiger Begehung.

Ort, Datum / Unterschrift