UNIVERSITY OF CALIFORNIA

Santa Barbara

Microbial Degradation of Macroalgal Polysaccharides by Members of the *Kiritimatiellota* phylum: From Lab Cultures to Marine Sediments

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Marine Science

by

Na Liu

Committee in charge:

Professor David L. Valentine, Chair

Professor Alyson E. Santoro

Professor Elizabeth G. Wilbanks

Professor Morgan R. Raven

December 2023

The dissertation of Na Liu is approved.

_____

Alyson E. Santoro

_____

Elizabeth G. Wilbanks

_____

Morgan R. Raven

_____

David L. Valentine, Committee Chair

December 2023

Microbial Degradation of Macroalgal Polysaccharides by Members of the *Kiritimatiellota* phylum: From Lab Cultures to Marine Sediments

by

Na Liu

ACKNOWLEDGMENTS

My PhD would not have been completed without the huge support, encouragement, and love from my family, friends, committee, lab mates, and many others who helped along the way.

To David Valentine, thank you for your constant mentorship and support during my entire PhD. I would not be here without you believing in me and encouraging me to move forward during my hard times. I am especially grateful for being given the freedom to explore research projects of my interest, which leads me to see a completely different but fascinating microbial world.

To Lizzy Wilbanks, I felt so lucky that you were recruited to UCSB when I started to look for committee members. Your endless passion for microbiology inspired me, and the discussions we had in your physiology lectures shaped the foundation of my PhD work. Thank you for welcoming me to your lab space and including me as part of the group. My RNA extractions and many other molecular works would not be smoothly completed without convenient access to your lab resources.

To Alyson Santoro, thank you for your interest, guidance, and advice in the experimental design and data analysis of my research projects. Your microbial diversity course is the key to the door of microbiology for me, and I would not have thought about culturing any bacteria without the agar plates we made in the lab. I am also very grateful for you checking in and encouraging me to stay in the program when the pandemic started.

To Morgan Raven, thank you for listening to my updates and brainstorming possible directions of my projects. Although I did not have the opportunity to delve deeper into the sulfur experiment, I still appreciate the interesting conversations we had.

To Veronika Kivenson, I miss the time sitting at a desk away from you in the office. You showed me many useful skills in bioinformatics and shared important experiences with me. You helped me a lot with sequencing and annotating the very first genome of the newly isolated bacterium in such a short time.

To Xuefeng Peng, thank you for your optimism and all the small talks guiding me through many problems appeared. Your insight and knowledge in polysaccharide degradation provided important information to be included in this dissertation.

To Zhisong Cui, thank you for your guidance in the lab during your visiting year. I would not have been able to culture and isolate a new bacterium without the hands-on experience with you. Your expertise and helpful tips are invaluable to me.

To the Valentine lab (and dogs and cats with whom we shared great time together), thank you for all the laughs and tears during my time here. Your daily company and support, either on land or at sea, mean a lot to me! Frank, thank you for helping me find everything needed for my experiments, no matter big or small. Ellie and Jon, thank you for answering my basic bioinformatic questions and sharing your codes with me. Dani, I shared most office time with you and I cherish the stories we shared over these years.

To Qianhui Qin, you are not only my close lab mate but also a great friend to share life with. I was fortunate to have you in the lab both as international students. You helped me in every way from academic studies to daily life needs. I hope to meet you, Susan, and your family again in the near future!

To Braulio and Taruna, thank you for being my Wilbanks lab mates and friends. I enjoyed talking with you a lot and your presence made my lab work much easier.

To Judith, I am very fortunate to share house with you for the past five years. You made Santa Barbara my second home in the US. It was a great memory having Thanksgiving and Christmas with your big family when I could not go back home in China. Ethan, thanks for taking me to skydiving. It was truly an adventure! And of course, I would never forget the lovely (sometimes silly) moments of Speckles, Obama and Matilda.

To Yijing (Olivia), you have been my best friend since middle school, and it was my fortune to reunite with you in California. You saved many of my breakdowns and are always there for me when needed. I was lucky to witness many big moments in your life and I wish you happiness with Xan and Noodle!

To Ruixia, thank you for showing up to our first ski trip so that I was able to meet and know you. There were many great memories with you and our friends in my last two years here - the trails we hiked, the food we cooked, the veggies and flowers we grew, the festivals and birthdays we celebrated, and especially, the cheerful smiles on your face. My PhD would not be finished without these happy times with you.

To my Family, thank you Mom and Dad, for your endless love and encouragement throughout my life. Grandma, thank you for staying strong into your 90s and you are always my model, though yourself probably did not realize how deeply your wisdom and perseverance influenced and empowered me to face everything in life. Weibin, thank you for firmly staying and standing by my side even we were apart for a long distance for many years. I am no longer alone in this world with your love and support.

VITA OF NA LIU
December 2023

EDUCATION

Bachelor of Science in Geology, Minor in Biological Science, Peking University, June 2016
Doctor of Philosophy in Marine Science, University of California, Santa Barbara, December 2023

PROFESSIONAL EMPLOYMENT

2016-2020: Teaching Assistant, Department of Earth Science, University of California, Santa Barbara
Summer 2020: Teaching Associate, Department of Earth Science, University of California, Santa Barbara
2020-2023: Graduate Student Researcher, Marine Science Institute, University of California, Santa Barbara

PUBLICATIONS

Liu, N., Kivenson, V., Peng, X., Cui, Z., Lankiewicz, T.S., Gosselin, K.M., English, C.J., Blair, E.M., O'Malley, M.A. and Valentine, D.L., 2024. *Pontiella agarivorans* sp. nov., a novel marine anaerobic bacterium capable of degrading macroalgal polysaccharides and fixing nitrogen. *Applied and Environmental Microbiology* (In Press).

Yousavich, D. J., Robinson, D., Peng, X., Krause, S. J. E., Wenzhoefer, F., Janßen, F., Liu, N., Tarn, J., Kinnaman, F., Valentine, D. L., and Treude, T., 2023. Marine anoxia initiates giant sulfur-bacteria mat proliferation and associated changes in benthic nitrogen, sulfur, and iron cycling in the Santa Barbara Basin, California Borderland, *EGUsphere* [preprint], pp.2023-1198.

Qin, Q., Kinnaman, F.S., Gosselin, K.M., Liu, N., Treude, T. and Valentine, D.L., 2022. Seasonality of water column methane oxidation and deoxygenation in a dynamic marine environment. *Geochimica et Cosmochimica Acta*, *336*, pp.219-230.

Cui, Z., Kivenson, V., Liu, N., Xu, A., Luan, X., Gao, W., Paul, B., and Valentine, D.L., 2019. Complete Genome Sequence of Cycloclasticus sp. Strain PY97N, Which Includes Two Heavy Metal Resistance Genomic Islands. *Microbiology Resource Announcements* 8, no. 40: 10-1128.

AWARDS

Alumni Graduate Award for Research Excellence (2018)
Wendell Phillips Woodring Memorial Graduate Fellowship (2019)
Graduate Student Opportunity Award (2019)
Research Accelerator Award, Graduate Division (2021)
Eleanor and Richard Migues Fund for Grad Field Studies Award (2022)

ABSTRACT

Microbial Degradation of Macroalgal Polysaccharides by Members of the *Kiritimatiellota* phylum: From Lab Cultures to Marine Sediments

by

Na Liu

Macroalgae are important primary producers in the ocean and form expansive underwater forests to support diverse organisms in the ecosystem. They also contribute to a large amount of carbon export below the euphotic zone and carbon sequestration in the deep ocean in the form of detritus. As a result, there is an increasing number of studies suggesting long-term storage of carbon in macroalgae as a carbon dioxide removal strategy to mitigate climate change. However, in contrast to the growing attention of macroalgae as one of the blue carbon ecosystems, the fate of macroalgal-derived carbon after entering the deep ocean remains largely unexplored.

Polysaccharides are a major component of both living and detrital macroalgae, and the degradation of these complex polymers is an important process for the turnover of carbon in the ocean. Macroalgal polysaccharides have diverse structures varying significantly in their monomeric saccharides, glycosidic linkages, branching sites, and modifications by chemical groups like sulfate. Therefore, the degradation of macroalgal polysaccharides requires an extensive number of enzymes produced by microorganisms specialized in this process.

Carbohydrate-active enzymes (CAZymes) and sulfatases are the key enzymes typically organized in polysaccharide utilization loci (PULs) that mediate a cascade of pathways to degrade polymers to oligosaccharides and eventually to monosaccharides. However, we still have knowledge gaps in the details of the degradation processes and the various enzymes involved.

In this dissertation, approaches like cultivation, genomic, and transcriptomic sequencing were applied to study the microbial degradation of macroalgal polysaccharides from cultures grown in the laboratory to marine sediments collected in the deep Santa Barbara Basin.

A novel anaerobic bacterial strain NLcol2 was isolated from microbial mats in sediments offshore Santa Barbara, California, USA, and the name *Pontiella agarivorans* sp. nov. was proposed. It is the fifth cultivated representative in the phylum *Kiritimatiellota* with the ability to degrade macroalgal polysaccharides including agar and iota-carrageenan and is the first strain reported to be capable of nitrogen fixation. More than 10% of its genome also encodes CAZymes, sulfatases and nitrogenases that facilitate these metabolisms.

To further study the enzymes and pathways involved in macroalgae degradation, a transcriptomic study was performed on strain NLcol2 growing with macroalgal polymers of agar and dried seaweeds compared with the monosaccharide D-galactose as control condition. During the growth of strain NLcol2 on polymers, 55% of the genome was expressed differentially, among which many CAZymes and sulfatases involved in the degradation pathways were significantly up-regulated. 12 putative PULs were identified and a selfish mechanism for agar degradation was proposed. The highly expressed genes involved in flagella formation and capsule polysaccharide synthesis with growth on

seaweeds could promote motility and biofilm formation for better access and attachment to seaweed particles.

To explore the ecological role of *P. agarivorans* and its relatives in the natural marine environment and to investigate the microbial communities enriched in kelp detritus on the seafloor, marine sediments with and without kelp detritus were sampled in the deep Santa Barbara Basin. The 16S rRNA gene survey showed an enrichment of phyla (*Bacteroidota*, *Spirochaetota*, *Desulfobacterota*, *Campylobacterota*, *Verrucomicrobiota* etc.) in kelp-laden sediments with cultivated relatives specializing in polysaccharide degradation, dissimilatory sulfate reduction and nitrogen fixation. Close relatives of *P. agarivorans* are widely distributed across the basin in low abundance in general but exhibited a high relative abundance in carbon-rich sediments with kelp detritus and at shallower depths. These microbial communities together contribute to the coupled biogeochemical cycling of carbon, sulfur, and nitrogen in the Santa Barbara Basin sediments in low oxygen conditions.

TABLE OF CONTENTS

Introduction

Macroalgae, or often referred to as seaweeds, are large photosynthetic algae (up to 60 m long) encompassing a global area of 6.06 – 7.22 million km^2 (Brodie & Lewis, 2007; Duarte et al., 2022). Macroalgae can be classified based on their pigmentation and the three major types are brown seaweed (Class *Phaeophyceae*), red seaweed (Class *Rhodophyceae*), and green seaweed (Class *Chlorophyceae*) (Brodie & Lewis, 2007). They form underwater forests by mainly anchoring on the rocky shores in coastal areas, while free-living species of *Sargassum*, *S. fluitans* and *S. natans*, also dominate in the North Atlantic Ocean (Lapointe et al., 2021; Steneck et al., 2002; M. Wang et al., 2019). Latitudinally, they thrive from tropical to subpolar to polar areas, with the highest diversity in the South Atlantic Ocean (Duarte et al., 2022; Ortega et al., 2019). Macroalgal forests contribute to a high net primary productivity of ~ 1.32 Pg C/year and support food webs with diverse biota including various invertebrates, fishes, as well as parasites in the ecosystem (Duarte et al., 2022; Morton et al., 2021).

Macroalgae also serve as one of the blue carbon ecosystems which have received growing attention on their potential for climate change mitigation (Macreadie et al., 2021; National Academies of Sciences, 2022). Due to the high productivity of macroalgae, they contribute to a significant amount of organic carbon exported below the euphotic zone to the deep ocean in the form of detritus (Pedersen et al., 2020). It has been shown that 24% of macroalgae can eventually reach the seafloor at 4,000 m (Ortega et al., 2019) and the blue carbon is considered as sequestered on a geological timescale once it enters the deep ocean (Dolliver & O'Connor, 2022). It has been estimated that macroalgae could sequester about 173 Tg C/year globally, 90% of which resulted in an export to the deep sea (Krause-Jensen

& Duarte, 2016). Therefore, seaweed cultivation has been considered as a promising carbon dioxide removal (CDR) strategy (National Academies of Sciences, 2022).

Despite their importance in carbon export and sequestration, it remains unclear about the fate of macroalgae detritus on the seafloor. They can eventually be buried into marine sediments, decomposed by grazers, or be remineralized by microbial communities (Dolliver & O'Connor, 2022). It is important to know the details of its fate to inform potential impacts and the work in this dissertation provides a microbial perspective to the degradation of macroalgal-derived carbon in the ocean.

Polysaccharides are important components in the cell wall of macroalgae and comprise 15–75% of their dry biomass (Patel et al., 2023). The compositions and structures of algal polysaccharides vary significantly between different types of seaweed: the carbohydrates in brown seaweed mainly include laminarin, mannitol, alginate, fucoidan (fucan), and cellulose (Kawai & Henry, 2016; Wei et al., 2013); the polysaccharides in the red seaweed are mainly composed of agar, carrageenan, and cellulose (Usov, 2011; Wei et al., 2013); and in green seaweed, ulvan, starch, and cellulose are the major components of the carbohydrate pool (Reisky et al., 2019; Wei et al., 2013). Unlike polysaccharides typically found in terrestrial plants, marine polysaccharides are usually modified by functional groups such as sulfate, amino, methyl, or acetate groups (Arnosti et al., 2021; Helbert, 2017). The long chains formed by glycosidic bonds, and sometimes branches, and various modifications by chemical groups all add difficulties for microbes to access and degrade these polymers (Arnosti et al., 2021). However, marine sediments host highly diverse microbial communities and there are specific taxa (*Bacteroidota* phylum) that specialize in degrading complex polysaccharides (Hoshino et al., 2020; McKee et al., 2021; X.-Y. Zhu et al., 2023).

These polysaccharide degraders play an important role in the carbon cycling in the ocean and provide insights into the fate of macroalgae detritus on the seafloor.

Macroalgae is considered as a renewable energy source of methane and alcohols through anaerobic digestion (Chynoweth et al., 1981; McKennedy & Sherlock, 2015; V et al., 2021; Wei et al., 2013). In general, complex organic polymers were hydrolyzed into simple sugars, amino acids, and fatty acids and these small organic compounds can be further fermented into acetic acid, methane, and ethanol (McKennedy & Sherlock, 2015; V et al., 2021). Although studies have shown heterotrophic microbes could aerobically and anaerobically decompose macroalgae, most of them were incubated and sampled in artificial systems (Delille & Perret, 1991; Morrison et al., 2017). And the anaerobic degradation of kelp detritus that occurs naturally in deep marine environments remains unexplored. Microbial degradation of these complex macroalgal polysaccharides require an extensive number of enzymes such as carbohydrate-active enzymes (CAZymes) and sulfatases (Arnosti et al., 2021; Ficko-Blean et al., 2017; Helbert, 2017; Reisky et al., 2019; Sichert et al., 2020; Y. Zhu et al., 2016). There have been a great number of enzymes in the databases with biochemically identified structure and known substrates, however, there are more with unknown enzymatic activities (Barbeyron et al., 2016; Lombard et al., 2014).

In this dissertation, I aim to explore the anaerobic microbes that contribute to the degradation of macroalgal polysaccharides as well as the enzymes involved in the process, from a defined system of lab cultures to the natural marine habitat. The work in this dissertation mainly used a combination of traditional microbiological method of cultivation and modern molecular approaches of genomic and transcriptomic sequencing.

Chapter 1 describes the isolation and characterization of a new anaerobic bacterial species *Pontiella agarivorans* (proposed name) strain NLcol2 belonging to the *Kiritimatiellota* phylum that originated from the microbial mats in the coastal sediments offshore Santa Barbara, California, USA. A series of biochemical, physiological, and genomic analyses were performed to elucidate its ability to degrade macroalgal polysaccharides coupled with nitrogen fixation.

To better understand the degradation process of macroalgal polysaccharides, the enzymes involved, and the microbial strategy to maximize the benefits gained from hydrolysis, a transcriptomic study was carried out in **Chapter 2**. We grew strain NLcol2 on different macroalgal substrates to look for the highly expressed genes involved in macroalgae degradation.

Besides experiments on the cultures in the lab, I am also curious about the abundance and distribution of strain NLcol2 and its relatives in the natural environment, their ecological role in the marine sediments, as well as their potential to degrade kelp detritus on the seafloor.

I was fortunate to be on the Atlantis cruise in 2019 and luckily found piles of kelp detritus fallen to the deep Santa Barbara Basin surrounded and covered by microbial mats. In **Chapter 3**, we sampled and surveyed the microbial communities in kelp-laden sediments and investigated the ecological role and distribution of strain NLcol2 and its relatives in the marine sediments.

Chapter 1. *Pontiella agarivorans* sp. nov., a novel marine anaerobic bacterium capable of degrading macroalgal polysaccharides and fixing nitrogen

This chapter is being published as:

Liu, N., Kivenson, V., Peng, X., Cui, Z., Lankiewicz, T.S., Gosselin, K.M., English, C.J., Blair, E.M., O'Malley, M.A. and Valentine, D.L., 2024. *Pontiella agarivorans* sp. nov., a novel marine anaerobic bacterium capable of degrading macroalgal polysaccharides and fixing nitrogen. *Applied and Environmental Microbiology* (In Press).

Abstract

Marine macroalgae produce abundant and diverse polysaccharides which contribute substantially to the organic matter exported to the deep ocean. Microbial degradation of these polysaccharides plays an important role in the turnover of macroalgal biomass. Various members of the *Planctomycetes-Verrucomicrobia-Chlamydia* (PVC) superphylum are degraders of polysaccharides in widespread anoxic environments. In this study, we isolated a novel anaerobic bacterial strain NLcol2^T from microbial mats on the surface of marine sediments offshore Santa Barbara, California, USA. Based on 16S rRNA gene and phylogenomic analyses, strain NLcol2^T represents a novel species within the *Pontiella* genus in the *Kiritimatiellota* phylum (within the PVC superphylum). Strain NLcol2^T is able to utilize various monosaccharides, disaccharides, and macroalgal polysaccharides such as agar and iota-carrageenan. A near-complete genome also revealed an extensive metabolic capacity for anaerobic degradation of sulfated polysaccharides, as evidenced by 202 carbohydrate-active enzymes (CAZymes) and 165 sulfatases. Additionally, its ability of

nitrogen fixation was confirmed by nitrogenase activity detected during growth on nitrogen-free medium, and the presence of nitrogenases (*nifDKH*) encoded in the genome. Based on the physiological and genomic analyses, this strain represents a new species of bacteria that may play an important role in the degradation of macroalgal polysaccharides and with relevance to the biogeochemical cycling of carbon, sulfur, and nitrogen in marine environments. Strain NLcol2^T (= DSM 113125^T = MCCC 1K08672^T) is proposed to be the type strain of a novel species in *Pontiella* genus, and the name *Pontiella agarivorans* sp. nov. is proposed.

Importance

Growth and intentional burial of marine macroalgae is being considered as a carbon dioxide reduction strategy but elicits concerns as to the fate and impacts of this macroalgal carbon in the ocean. Diverse heterotrophic microbial communities in the ocean specialize on these complex polymers such as carrageenan and fucoidan, for example, members of the *Kiritimatiellota* phylum. However, only four type strains within the phylum have been cultivated and characterized to date, and there is limited knowledge about the metabolic capabilities and functional roles of related organisms in the environment. The new isolate strain NLcol2^T expands the known substrate range of this phylum and further reveals the ability to fix nitrogen during anaerobic growth on macroalgal polysaccharides, thereby informing the issue of macroalgal carbon disposal.

1.1 Introduction

Marine macroalgae are important primary producers in coastal ecosystems. They sequester about 173 TgC yr^{-1} into their biomass and are considered as part of the “blue

carbon" in the ocean (Krause-Jensen & Duarte, 2016). Seaweed cultivation has been considered as one of the promising strategies to mitigate the increasing amount of anthropogenic CO_2 and climate change (National Academies of Sciences, 2022). A recent study shows that 24% of macroalgae will eventually reach the seafloor and thus export the fixed carbon to the deep ocean (Ortega et al., 2019). Polysaccharides are important components among the fixed carbon, which includes agar, carrageenan, and fucoidan etc. (Popper et al., 2011; Wei et al., 2013). In contrast to terrestrial plants, marine polysaccharides are usually decorated by sulfate and other functional groups, which require specialized enzymes for removal, and thereby limit the range of microbes that can access and degrade these compounds (Helbert, 2017).

Members of the PVC superphylum (named for *Plantomycetes, Verrucomicrobia, Chlamydiae)* include degraders of recalcitrant glycopolymers, though much of their true functional diversity has been obscured by the lack of cultivated representatives (Cardman et al., 2014; Glöckner et al., 2003; Kim et al., 2016; Martinez-Garcia et al., 2012). The PVC superphylum also consists of phyla *Kiritimatiellota* and *Lentisphaerae* as well as uncultured candidate phyla from environmental samples (Rivas-Marín & Devos, 2018). The *Kiritimatiellota* phylum was established in 2016 (previously named as *Kiritimatiellaeota*), and was recognized as the Subdivision 5 of *Verrucomicrobia* in the PVC superphylum (Oren & Garrity, 2021; Spring et al., 2016). The geographic distribution of 16S rRNA gene sequences reveals that bacteria in phylum *Kiritimatiellota* are common to anoxic environments ranging from the intestine of animals to hypersaline sediments and wastewater (Spring et al., 2016). However, there are only four cultivated strains reported to date, and we know little about their metabolic capabilities and functional role in the environment. The

first cultivated strain, *Kiritimatiella glycovorans* L21-Fru-AB^T, is a halophilic saccharolytic bacterium isolated from an anoxic cyanobacterial mat from a hypersaline lake on the Kiritimati Atoll (Spring et al., 2015). *Pontiella desulfatans* F1^T and *Pontiella sulfatireligans* F21^T were isolated from Black Sea sediments and are capable of degrading sulfated polysaccharides like iota-carrageenan and fucoidan (van Vliet et al., 2019, 2020). *Tichowtungia aerotolerans* S-5007^T was isolated from surface marine sediment and can grow under microaerobic conditions (Mu et al., 2020).

In this study, we enriched and isolated a novel anaerobic bacterial strain NLcol2^T from the marine sediments offshore Santa Barbara, California, USA, which belongs to the *Kiritimatiellota* phylum. We fed the strain with agar, iota-carrageenan, and fucoidan as carbon substrate to test whether it is able to degrade these polysaccharides or not. Among other isolates of *Kiritimatiellota*, ammonium has been identified as the nitrogen source, but nitrogen fixation has not been observed. However, macroalgal polysaccharides are depleted in nutrients including nitrogen, therefore, we used nitrogen gas as the sole nitrogen source to test its ability of nitrogen fixation. Strain NLcol2^T is characterized by phylogenomic, morphological, chemotaxonomic, and physiological traits. We further investigated its metabolic potential by analyzing CAZymes, sulfatases, and nitrogenases in the genome in detail.

1.2 Materials and Methods

1.2.1 Inoculum source, enrichment, and isolation of strain NLcol2^T

Strain Nlcol2^T was enriched and isolated from microbial mats found on the surface of marine sediments at Shane Seep (34.40616 N, 119.89047 W) within the Coal Oil Point seep field offshore Santa Barbara, California, USA. Microbial mat samples were collected at 20

m depth with an in-situ temperature of 15 °C in October 2017. The seep area is characterized by a large amount of hydrocarbon gas emissions, microbial mat coverage, and high sulfide and alkalinity in sediment porewater (Ding & Valentine, 2008; Eichhubl et al., 2000; Washburn et al., 2005). The samples used for inoculum contained both microbial mats and partially decomposed macroalgae (**Figure 1.1a**). The microbial mats were scraped off their attached surface as the inoculum source. The cultures were enriched anaerobically in semi-solid agar (0.25% w/v, BD Difco Agar, Granulated) in the top layer of the sulfide gradient media (**Figure 1.1b**) modified from Kamp et al., 2006. Cultures were maintained at room temperature in the dark and were transferred into fresh media every two to three weeks for a year.

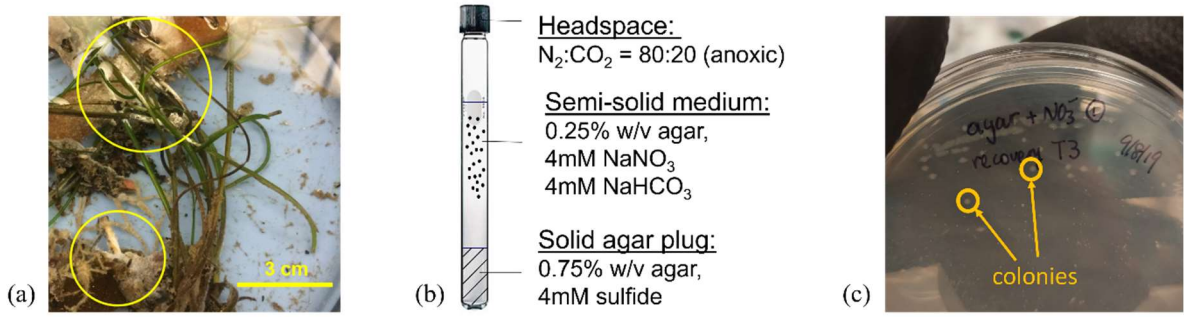


Figure 1.1. (a) Microbial mats collected on the surface of marine sediments were used as inoculum for the enrichment and isolation of bacterial cultures. **(b)** Anoxic enrichment culture setup with agar gradient seawater media. **(c)** Isolation of culturable strains on agar plates. Colonies were highlighted in yellow circles.

Further isolation of strain Nlcol2^T was performed by streaking on agar plates in an anaerobic chamber (Coy Laboratory Products) (**Figure 1.1c**). The medium is the same as the top agar medium in enrichment cultures, except that 1.5 % w/v agar (BD Difco Agar, Noble) was added as both gelling agent and substrate and 2 mM sulfide added as reducing agent. The Petri dishes were kept in the anaerobic chamber at room temperature (22 °C). Single colonies formed after three weeks and were picked from agar plates. Streak plating was

repeated for three more rounds to ensure the purity of the culture. Pure culture was subsequently maintained in liquid media with D-galactose (1g/L) as substrate at 22 °C and was transferred every other week. A full modified medium contained: 28.0 g NaCl, 10.0 g $MgCl_2 \cdot 6\ H_2O$, 3.8 g $MgSO_4 \cdot 7\ H_2O$, 0.6 g $CaCl_2 \cdot 2\ H_2O$, 1.0 g KCl, 37 mg K_2HPO_4, 4 mg Na_2MoO_4, 50 mg $Na_2S_2O_5$, 2 mg $FeCl_3 \cdot 6\ H_2O$, 10.0 mL modified Wolin's Mineral Solution (see DSMZ medium 141), 0.5 mL Na-resazurin solution (0.1% w/v), 1.0 g D-galactose, 1.0 g NH_4Cl (optional), 0.75 g Na_2CO_3, 0.5 g $Na_2S \cdot 9\ H_2O$, 10.0 mL Wolin's Vitamin Solution (see DSMZ medium 141), in 1000 mL distilled water. All ingredients except carbonate, sulfide and vitamins were dissolved under N_2/CO_2 (80:20) atmosphere in Hungate tubes or serum bottles and autoclaved. Carbonate was added from a sterile anoxic stock solution prepared under N_2/CO_2 (80:20) atmosphere. Sulfide and vitamins were added from sterile anoxic stock solutions prepared under 100% N_2 gas. Purity of the isolate was checked by full-length 16S rRNA gene sequencing and observation of morphology under the microscope.

1.2.2 Phylogenetic reconstruction by 16S rRNA gene

Full-length 16S rRNA gene of strain Nlcol2^T was sequenced by GENEWIZ (Azenta Life Sciences), from colonies grown on agar plates. 16S rRNA gene sequence was searched using the website tool BLASTn (Camacho et al., 2009) against the 16S rRNA database and compared to the sequence identity to the other four isolated strains in the *Kiritimatiellota* phylum.

To construct a phylogenetic tree based on the 16S rRNA gene, 106 sequences over 1200 bp from the *Kiritimatiellales* order in SILVA Ref NR SSU r138.1 database (released August 2020, accessed November 2021) (Quast et al., 2012) were selected for alignment. The full-

length 16S rRNA genes of strain Nlcol2^T, *Tichowtungia aerotolerans* strain S-5007^T, and two *Verrucomicrobia* (ABEA03000104, AF075271 as outgroups) were also added to the alignment using SINA Aligner v1.2.11 (Pruesse et al., 2012). The alignment was trimmed using the "gappyout" method in TrimAl v1.4 (Capella-Gutiérrez et al., 2009) to remove ambiguous ends and columns with >95% gaps. All trimmed nucleotide sequences represent >50% of the 1568 alignment columns. A maximum-likelihood tree was constructed using RaxML v.8.2.9 (Stamatakis, 2014) with GTRGAMMA model of evolution. Rapid bootstrap search was stopped after 1000 replicates with MRE-based criterion. The best-scoring ML tree with support values was visualized in the iTOL server (Letunic & Bork, 2019).

1.2.3 Genome sequencing and analyses

Genomic DNA was extracted from the isolate cultures using FastDNA Spin Kit for Soil (MP Biomedicals, OH). Genomic DNA library preparation and sequencing were performed at the University of California Davis Genome Center on Illumina HiSeq 4000 platform with 150-base pair (bp) paired-end reads. Trimmomatic v.0.36 (Bolger et al., 2014) and Sickle v.1.33 (Joshi & Fass, 2011/2011) were used to remove adapter and low quality or short reads. Trimmed reads were assembled into contigs using MEGAHIT v.1.1.1 (D. Li et al., 2015). Contigs longer than 2500 bp were kept and the trimmed reads were mapped back to those contigs using Bowtie2 v.2.3.4.1 (Langmead & Salzberg, 2012) and Samtools v.1.7 (H. Li et al., 2009). Contigs were visualized using Anvi'o v.3 interactive interface (Eren et al., 2015) and manual binning was performed based on coverage, GC content, and tetranucleotide frequency signatures. Completion and redundancy for the reconstructed genome was determined using CheckM v.1.0.7 (Parks et al., 2015).

Open reading frame (ORF) features and protein-coding gene sequences were predicted using Prodigal v.2.6.3 (Hyatt et al., 2010). Annotation was assigned to proteins using HMMER v.3.1b2 (Eddy, 2011) hmmscan searching against the Pfam v.32.0 (El-Gebali et al., 2019) and TIGRFAMs v.15.0 (Haft, 2001) databases with a maximum e-value of 1×10^{-7}, corresponding to a bit score of > 30 to balance the trade-offs between false positives and missed matches. Information on protein family, domain and conserved site were confirmed using InterProScan5 (Mitchell et al., 2019). The amino acid sequences of protein-coding genes were further searched against NCBI's Conserved Domain Database (CDD) (Marchler-Bauer et al., 2017) using the RPS-BLAST program v.2.7.1. The cdd2cog script (Leimbach, 2016) was used to assign COG (Cluster of Orthologous Groups) categories (Tatusov et al., 2003) to each protein-coding gene. Protein sequences were also submitted to the BlastKOALA server (Kanehisa et al., 2016) for KEGG Orthology (KO) ID assignments. Ribosomal RNA genes were determined by RNAmmer v.1.2 (Lagesen et al., 2007). tRNA genes were predicted by tRNAscan-SE 2.0 server (Lowe & Chan, 2016). Metabolic pathways were reconstructed using KEGG Mapper (Kanehisa et al., 2012) and MetaCyc database (Caspi et al., 2020).

For phylogenomic analyses, high-quality genomes in the *Kiritimatiellales* order from NCBI's GenBank database and the Genome Taxonomy Database (GTDB) r95 were selected (accessed on Feb 1, 2021). *Opitutus terrae* PB90-1 from the *Verrucomicrobia* phylum was selected as the outgroup. All genomes meet the GTDB quality criterion based on completeness and redundancy from CheckM: completeness – 5×redundancy > 50. 120 single-copy genes were searched and aligned using GTDB-Tk v1.4.0 (Chaumeil et al., 2020). The concatenated alignment was further trimmed using TrimAl v1.4 (Capella-

Gutiérrez et al., 2009) with "gappyout" parameter, which results in a final alignment with 4488 amino acid columns. Maximum-likelihood phylogenetic tree was calculated using RaxML v.8.2.9 (Stamatakis, 2014) with PROTGAMMALG model of evolution. Rapid bootstrap search was stopped after 350 replicates with MRE-based criterion. The best-scoring ML tree with support values was visualized in the iTOL server (Letunic & Bork, 2019). The average nucleotide identity (ANI) and average amino acid identity (AAI) between genomes were calculated using the ANI/AAI calculator (Rodriguez-R & Konstantinidis, 2016).

Carbohydrate-active enzymes were predicted using dbCAN2 meta server (H. Zhang et al., 2018). In brief, uploaded protein sequences were searched against the dbCAN CAZyme domain HMM database v.7, CAZy database (www.cazy.org) and PPR library using HMMER, DIAMOND, and Hotpep programs respectively (H. Zhang et al., 2018). Only genes predicted by no less than two programs were defined as CAZymes for further analysis. CAZyme gene clusters were predicted by the CGC-Finder on dbCAN2 server with at least one CAZyme and one transporter detected within a maximum distance of two genes (H. Zhang et al., 2018). To classify sulfatases into families and subfamilies, gene sequences with an annotated sulfatase domain (PF00884) were searched and classified by the SulfAtlas database v.1.1 (Barbeyron et al., 2016) using the BLASTp program (Camacho et al., 2009). Additionally, SignalP v.5 (Almagro Armenteros et al., 2019) was used to predict signal peptides for translocation of sulfatases into the periplasmic space and outside of the cells.

To better understand the evolution of nitrogen fixation in the *Kiritimatiellota* phylum, reannotation and phylogenetic analysis of the *nifH* gene were performed for all 52 genomes in this phylum from NCBI's GenBank database (accessed on Mar 3, 2020). The same

annotation pipeline described above was used to keep consistency and allow better comparison. *nifH* gene sequences were aligned with 879 full-length *nifH* genes from the genomes of cultivated diazotrophs (https://wwwzehr.pmc.ucsc.edu/Genome879/) using MUSCLE v.3.8 (Edgar, 2004). Two light-independent protochlorophyllide reductases were included as outgroups: ChlL from *Trichormus variabilis* ATCC 29413 (WP_011320185.1) and BchL from *Chlorobium limicola* DSM 245 (WP_012467085.1). The alignment was trimmed in Jalview v.2.10.5 (Waterhouse et al., 2009) to remove ambiguous ends and the columns with >95% gaps. All trimmed amino acid sequences represent >81% of the alignment columns. A maximum-likelihood tree was constructed using RaxML v.8.2.9 (Stamatakis, 2014) with LG substitution model plus GAMMA model of rate heterogeneity. Rapid bootstrap search was stopped after 350 replicates with MRE-based criterion. The best-scoring ML tree with support values was visualized in the iTOL server (Letunic & Bork, 2019).

1.2.4 Microscopy

To obtain high-resolution images, cell morphology was examined under the transmission electron microscope (TEM). For TEM imaging, cells grown on the agar plates were fixed with modified Karnovsky's fixative (2% paraformaldehyde and 2.5% glutaraldehyde in 0.1 M sodium phosphate buffer) and spun down into a cell pellet. Cells were rinsed in 0.1 M sodium phosphate buffer and fixed again with 1% osmium tetroxide in the same buffer. After another rinse, they were dehydrated in 50% EtOH, 75% EtOH, 95% EtOH, 100% EtOH and propylene oxide twice. Cells were pre-infiltrated in 1:1 propylene oxide:resin (Epon/Alradite mixture) overnight, infiltrated in 100% resin and embedded in fresh resin at 60 °C overnight. Ultrathin sections were cut using a Diatome diamond knife. Sections were

picked up on copper grids and imaged in a FEI Talos 120C transmission electron microscope at the Biological Electron Microscopy Facility, University of California Davis.

1.2.5 Chemotaxonomic analysis

The cellular fatty acid composition of strain Nlcol2^T was determined from cells grown at 22 °C to late-log phase in liquid medium with 1.0 g/L D-galactose as carbon source and nitrogen gas as nitrogen source. Cells were centrifuged down at 10,000 × g for 10 mins and were frozen in -80 °C.

Cellular fatty acids were extracted twice using a modified Folch method (Folch et al., 1957) with a chloroform: methanol mixture (2:1) and tridecanoic acid as an internal standard. The samples were partitioned and the organic phase containing the total lipid extract (TLE) was retained. Transesterification of the TLE was performed by adding toluene and 1% sulfuric acid in methanol to the TLE after it was brought to complete dryness under N_2. The acidic methanol/toluene TLE was heated at 90 °C for 90 minutes to produce fatty acid methyl esters (FAME). The FAMEs were extracted from the acidic methanol by adding hexane and water, vortexing, centrifuging, and removing the top (hexane) fraction to a new vial twice. The combined transesterified hexane extracts were dried under N_2 to a final volume of 300 µL. Each extract was spiked with methyl heptadecanoate to calculate the recovery of the internal standard and analyzed by gas chromatography with flame ionization detection (GC-FID).

Concentration analysis was performed with an HP 5890 Series II GC-FID. Chromatography was performed with a 30 m × 0.25 mm internal diameter (ID), 0.25 µm pore size, fused silica Omegawax capillary column with a splitless 1-µL injection. Initial oven temperature was set at 50 °C and held for 2 min, followed by a 10 °C min^{-1} ramp to

150 °C, then a 5 °C min^{-1} ramp to the final temperature of 265 °C. A certified reference material (FAME 37, Supelco) was run to calculate retention times and identify peaks. Peak identification was further confirmed from their mass spectra.

Analyses of catalase, oxidase, and API ZYM assay for semi-quantitation of enzymatic activities (e.g. beta-galactosidase) were carried out by DSMZ Services, Leibniz-Institut DSMZ – Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, Germany.

1.2.6 Physiology

Bacterial growth of strain Nlcol2^T was monitored by measuring optical density (OD) of liquid cultures at 600 nm wavelength. Growth at different temperature (4, 10, 14, 22, 26, 31, 37, 55 °C), salinity (0%, 1%, 2%, 2.5%, 3%, 4%, 5%, 6% NaCl) and pH (4.0, 5.0, 5.5, 6.0, 6.5, 7.0, 8.0, 9.0) conditions were determined in triplicates when growing on D-galactose with ammonium supplied. Growth was tested on various substrates (1 g/L) in triplicates at optimum temperature, salinity and pH conditions with ammonium supplied: D-glucose, D-galactose, D-fructose, L-fucose, L-rhamnose, D-mannose, D-mannitol, meso-inositol, D-arabinose, D-xylose, D-cellobiose, lactose, sucrose, maltose, xylan from corn core (TCI), starch (Sigma-Aldrich), cellulose (Sigma-Aldrich), alginic acid (Acros Organics), agarose (Sigma-Aldrich), agar (BD Difco Agar, Noble), iota-carrageenan (TCI), fucoidan from *Macrocystis pyrifera* (Sigma-Aldrich), commercially bought dried red algae (*Porphyra* spp.), commercially bought dried brown algae (*Saccharina japonica*), and the giant kelp (*Macrocystis pyrifera*) harvested from offshore Santa Barbara, California, USA.

To test the utilization of several nitrogen sources by strain Nlcol2^T, we cultured them with sodium nitrate (1 g/L), ammonium chloride (1 g/L), and without any nitrogen species

supplemented in the liquid media. Two sets of tubes with headspace gases of nitrogen gas or helium gas were made as experimental and control group respectively. Triplicate cultures were supplied with 1 g/L D-galactose as substrate and incubated at room temperature (22 °C) for 14 days. Growth was monitored by OD (600 nm) measurements.

1.2.7 Metabolite analysis from galactose fermentation

To quantify the metabolic products of strain Nlcol2^T from galactose fermentation, cultures were grown in triplicates at room temperature (22 °C) with D-galactose as carbon source for 10 days. No ammonium was added in the media and N_2 gas was served as the sole nitrogen source. Growth was monitored by measuring optical density (OD) at 600 nm wavelength. 2mL of culture was subsampled each day (twice a day during exponential phase) for quantification of metabolites.

The chromatography protocols used in this study are similar to those previously described (Gilmore et al., 2019; Peng et al., 2021). Galactose, acetate, succinate, and malate concentrations were measured on an Agilent Infinity 1260 (Agilent Technologies, Santa Clara, CA, USA) high-performance liquid chromatograph (HPLC) using an Aminex HPX-87H analytical column (part no. 1250140, Bio-Rad, Hercules, CA, USA) protected by, first, a 0.22 μm physical filter, followed by a Coregel USP L-17 guard cartridge (Concise Separations, San Jose, CA, USA). Separations were performed at 60 °C with a flow rate of 0.6 mL/min and a 5 mM sulfuric acid (H_2SO_4) mobile phase. Acetate, succinate, and malate were measured using a variable wavelength detector set to 210 nm, while galactose was measured using a refractive index detector set to 35 °C. Samples and standards for HPLC were acidified to a concentration of 5 mM H_2SO_4, incubated for 5 min at room temperature, and spun at maximum speed in a tabletop centrifuge for 5 min to pellet bacterial cells. The

samples were removed from above the cell pellet, and 0.22 μm filtered through a polyethersulfone (PES) membrane into HPLC vials with 300 μL polypropylene inserts. Standard curves for each compound of interest were constructed using triplicate standards of 0.1, 0.5, and 1.0 g/L. Peaks were integrated using OpenLab CDS analysis software (version 2.6, Agilent Technologies).

Hydrogen gas production was measured on a Fisher Scientific TRACE 1300 Gas Chromatograph (Thermo Fisher Scientific, Waltham, MA) using a TRACE TR- 5 GC Column (part no. 260E113P, Thermo Fisher Scientific) at 30 °C, with an Instant Connect Pulsed Discharge Detector (PDD) (part no. 19070014, Thermo Fisher Scientific) at 150 °C, and ultra-high purity He as a carrier gas. All injections of samples and standards were 100 μL in volume. Supplier-mixed standards of 50 ppm, 500 ppm, and 1% hydrogen were run before and after injecting samples, and hydrogen peaks were integrated using Chromeleon Chromatography Data System (CDS) Software (version 6.8, Thermo Fisher Scientific). CO_2 was not considered due to the carbonate-buffered medium and N_2/CO_2 atmosphere.

A tentative fermentation balance was formulated based on the concentrations of galactose, succinate, acetate, malate, and hydrogen measured above. The changes of concentrations in mmol/L were taken as coefficients for these compounds. The biomass was formulated with a C:N molar ratio of 106:16 following the canonical Redfield ratio and the coefficient was determined by the balance of carbon. Nitrogen and ammonium were also included for electron balance and as part of the biomass.

1.2.8 Metabolite analysis from agar and i-carrageenan degradation

Cultures were grown at 33 °C with 1 g/L agar (BD Difco Agar, Noble) and 1 g/L iota-carrageenan (TCI) as carbon sources and ammonium was supplied in the media. 10 mL of

culture was sampled and filtered through 0.22 um polyethersulfone (PES) membrane (Millipore Millex) both on Day 0 immediately after inoculation, as well as on Day 9 and Day 7 for agar and carrageenan incubations respectively. Growth was confirmed by OD (600 nm) measurements.

Agar and carrageenan concentrations were quantified as polymeric galactose, the main sugar component of the two polymers. Polymeric galactose was quantified as the difference between total galactose and free galactose. To measure total galactose, 5ml of 0.22um-filtered media was acid hydrolyzed to cleave glycosidic linkages and release galactose. Samples were hydrolyzed in 1M HCl at 100°C for 20 hours. Following hydrolysis samples were neutralized by N_2 evaporation and diluted 1:1000 with ultrapure water. Galactose was quantified using high performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) on a DIONEX ICS5000+ equipped with a CarboPac PA10 column using an isocratic elution of 18mM NaOH for 20 minutes (Engel & Händel, 2011). Free galactose was measured by HPAEC-PAD before acid hydrolysis. Incubation media was 1:25 diluted with ultrapure water to reduce the salt concentration and quantified using the same gradient program described above.

Acetate and succinate concentrations were measured on the Agilent Infinity 1260 HPLC using a similar protocol described above in section 1.2.7, except using a refractive index detector.

1.2.9 Acetylene reduction assay

To test the nitrogenase activity of strain Nlcol2^T when growing with nitrogen gas as the sole nitrogen source, acetylene reduction assay was performed following Hardy et al., 1968. In short, acetylene (C_2H_2) can be reduced to ethylene (C_2H_4) when nitrogenases actively fix

nitrogen gas at the same time. Cultures were grown on D-galactose in triplicates at 22 °C and triplicate media bottles without inoculation were used as controls. 1.2 mL of acetylene was injected to all culture and control bottles, which contained 80 mL of liquid and 80 mL of headspace pressurized at 150 kPa at the beginning. Gas concentrations and OD_{600} were measured at 6 time points during the 18-day incubation. Acetylene and ethylene concentrations were resolved on a Shimadzu 8A Gas Chromatograph with a flame ionization detector (GC-FID). 1.5 mL samples and standards were injected, then carried by N_2 at a flow rate of 20 mL/min through an n-octane on Res-Sil C packed column (Restek, Centre County, PA, USA) set at 25 °C. 0.5% and 1.0% GASCO calibration gas mixtures of acetylene and ethylene (Cal Gas Direct Incorporated, Huntington Beach, CA, USA) were used for the standard curves.

1.3 Results and Discussions

1.3.1 Phylogenetic analyses

Phylogenetic placement of strain NLcol2^T was determined by comparing full-length 16S rRNA gene, single-copy genes, and whole-genome similarity metrics including average nucleotide identity (ANI) and average amino acid identity (AAI).

Strain Nlcol2^T was classified within the R76-B128 clade (*Pontiellaceae* family in GTDB database) of the *Verrucomicrobia* phylum under the current SILVA taxonomy (SILVA Ref NR SSU r138.1). Full-length 16S rRNA gene of the isolate shares 84.1%, 88.9%, 92.9%, and 94.5% identity with the four reported cultivated strains in the *Kiritimatiellota* phylum: *Kiritimatiella glycovorans* strain L21-Fru-AB^T, *Tichowtungia aerotolerans* strain S-5007^T, *Pontiella sulfatireligans* strain F21^T, and *Pontiella desulfatans* strain F1^T, respectively. Strain Nlcol2^T is more closely related with *P. desulfatans* and *P. sulfatireligans* than *K.*

glycovorans and *T. aerotolerans*. The 16S rRNA gene identities compared to *P. desulfatans* and *P. sulfatireligans* were absolutely higher than the 86.5% threshold for family level, but fall on the edge of the threshold for a new genus as 94.5% (Yarza et al., 2014). A maximum-likelihood tree of 16S rRNA gene sequences from the *Kiritimatiellota* phylum was reconstructed by RaxML (**Figure 1.2**). The R76-B128 clade (*Pontiellaceae* family) formed a monophyletic group with MSBL3 clade (*Tichowtungiaceae* family) as the sister group, both of which are in a different cluster from the *Kiritimatiellaceae* family. It is clear that strain Nlcol2^{T} is not affiliated with *K. glycovorans* within the *Kiritimatiellaceae* family nor with *T. aerotolerans* within the *Tichowtungiaceae* family (MSBL3 clade), but belongs to the *Pontiellaceae* family (R76-B128 clade) within the *Kiritimatiellales* order as do *P. desulfatans* and *P. sulfatireligans* (van Vliet et al., 2019).

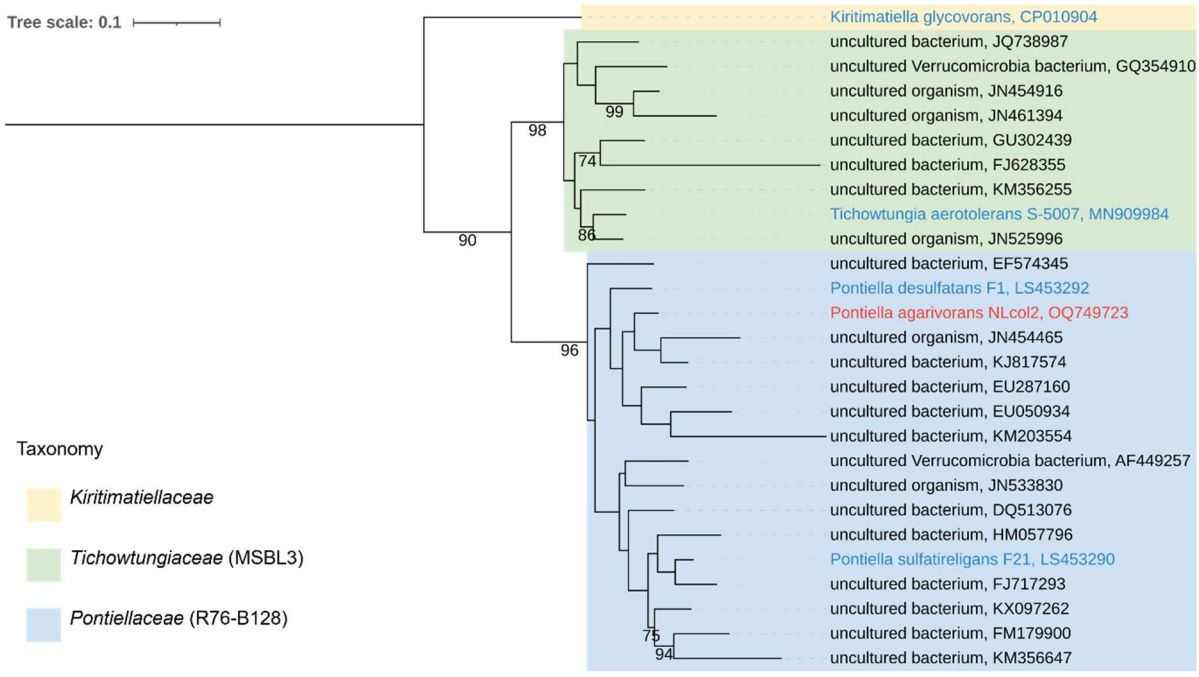


Figure 1.2. Maximum-likelihood phylogenomic tree of 16S rRNA genes of the Kiritimatiellales order. Strain Nlcol2 is labeled in red and other cultivated strains are labeled in blue. The tree was rooted with two Verrucomicrobia sequences (not shown) and pruned with 27 selected representative sequences. Bootstrap values over 70 are shown on the nodes.

To resolve the phylogeny of strain Nlcol2^T in detail, we further performed genome-level phylogenetic analyses using the Genome Taxonomy DataBase toolkit (Chaumeil et al., 2020). A concatenated phylogenomic tree was reconstructed from 120 bacterial single-copy genes of genomes in the *Kiritimatiellales* order (**Figure 1.3**). Here, strain Nlcol2^T falls within the *Pontiellaceae* family with a bootstrap value of 100. Additionally, the average amino acid identity (AAI) values of the genomes between strain Nlcol2^T and *P. desulfatans* and *P. sulfatireligans* are 69.94% and 68.51%, which are slightly above the threshold of 65% for same genus (Konstantinidis et al., 2017). However, within *Pontiella* genus, it represents a different group from *P. desulfatans* and *P. sulfatireligans*. Moreover, the average nucleotide identity (ANI) values of the genomes between strain Nlcol2^T and *P. sulfatireligans* and *P. desulfatans* are 72.73% and 73.71% respectively, which was much lower than the 95% ANI criterion for the same species (Jain et al., 2018; Konstantinidis & Tiedje, 2005). Therefore, we propose that strain Nlcol2^T represents a novel species within the *Pontiella* genus according to the phylogenetic analyses above.

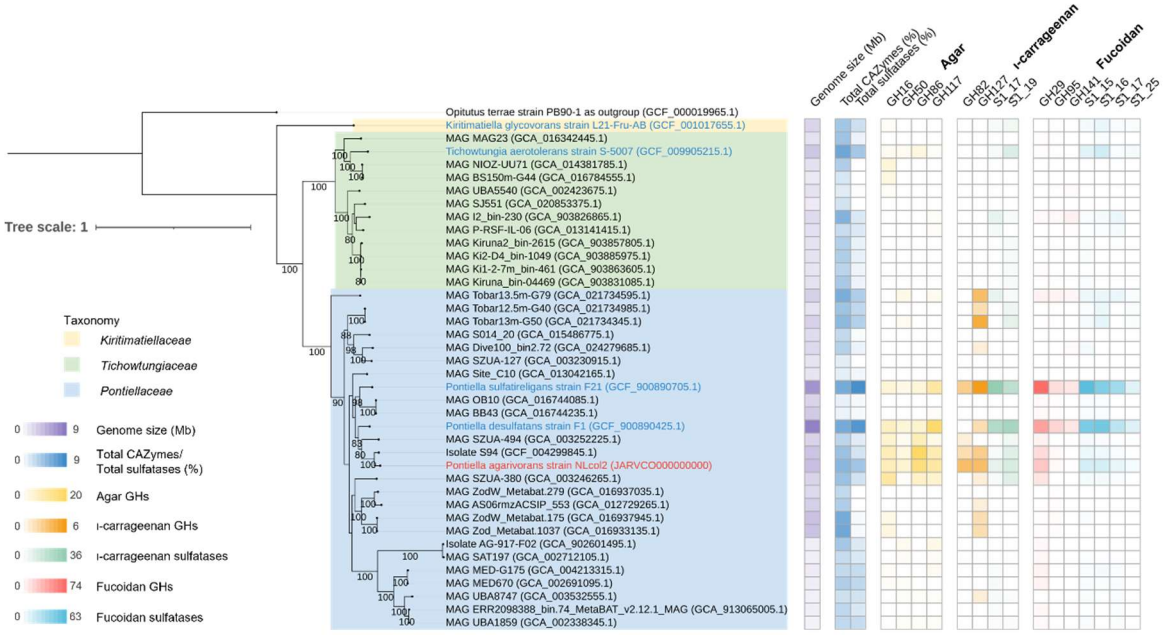


Figure 1.3. Concatenated maximum-likelihood phylogenomic tree of 120 bacterial single-copy genes from 39 genomes in the Kiritimatiellales order. Strain Nlcol2 is labeled in red and other cultivated strains are labeled in blue. A Verrucomicrobia genome was selected as outgroup and the tree was rooted there. Bootstrap values over 80 are shown on the nodes. Genome size (Mb), total CAZymes (%) and total sulfatases (%) as a percentage of all protein-coding genes in each genome are presented as reference. The number of glycoside hydrolase (GH) and sulfatase homologs involved in the degradation pathways of agar, iota-carrageenan, and fucoidan are presented in the heatmap.

1.3.2 General features of the genome

The draft genome of strain Nlcol2^T is 95% complete with 4% redundancy. The genome consists of 12 contigs (N50 is 1,265,434 bp) with a total length of 4,436,865 bp and the mean coverage is 593x. DNA G+C content is 52.4 mol%. 5S, 16S and 23S rRNA genes and 50 tRNA genes were found in the genome.

3,611 ORFs were predicted by Prodigal, among which 2,757 proteins in the genome were assigned with COG (Cluster of Orthologous Groups) functional category codes. The number of genes in each functional category is shown in **Figure 1.4**. More genes are involved in carbohydrate (260) and amino acid (188) transport and metabolism than those of nucleotides (65) and lipids (61), which is similar to that in *Kiritimatiella glycovorans* (12). A further detailed analysis of genes involved in macroalgal polysaccharide degradation and nitrogen fixation is presented in sections 1.3.4 and 1.3.5.

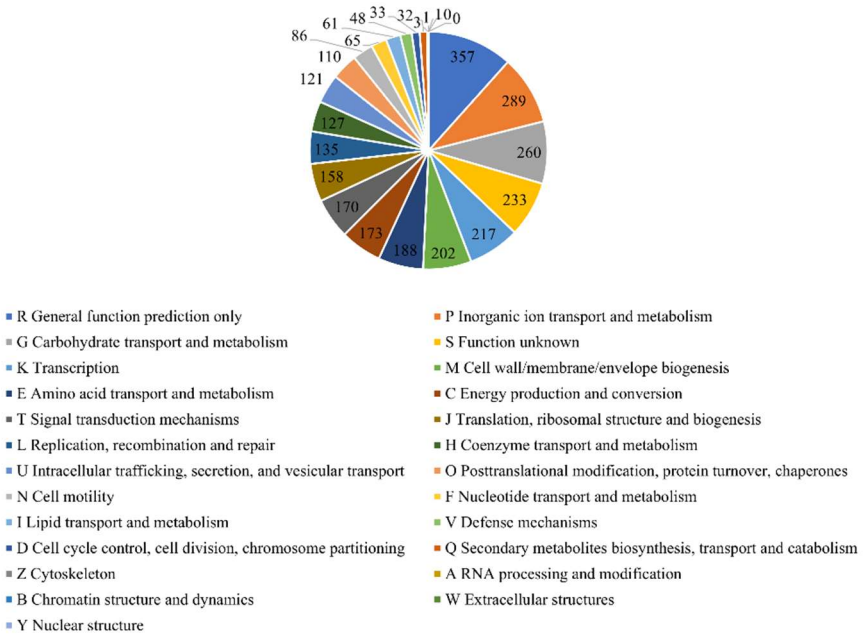
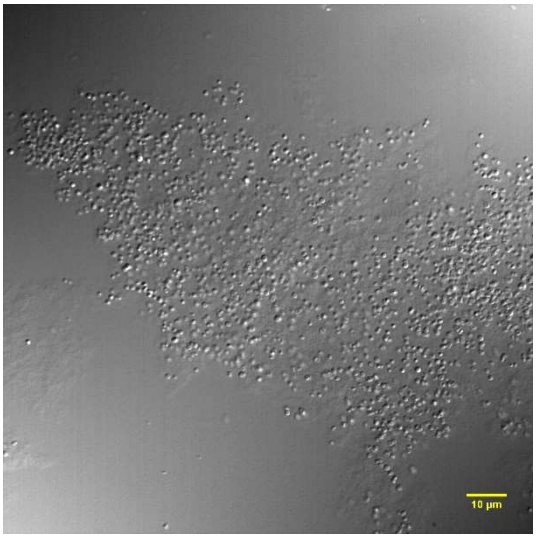


Figure 1.4. Pie chart showing number of genes assigned in COG functional categories.

1.3.3 Morphologic and chemotaxonomic characterization of strain Nlcol2^T

Single colonies on agar plates were white or ivory, circular, and smooth after growing anaerobically for 2 weeks at 22 °C. Bacterial cells of strain Nlcol2^T have a round to ovoid shape with a size of 1 µm in diameter observed under microscope (**Figure 1.5a**). Cells divide by binary fission and genes of bacterial cell division complex including FtsZ family were present. No motility or flagella were observed, although a full set of genes coding for flagellar assembly was present in the genome. No spore formation was observed. A Gram-negative cell wall structure of outer membrane, periplasmic space and cytoplasmic membrane was shown by electron microscopy (**Figure 1.5b**). There are also genes coding for proteins involved in lipopolysaccharide export and peptidoglycan synthesis in the genome. Some bacteria in the PVC superphylum exhibit compartments inside the cells (Santarella-Mellwig et al., 2010), but like other strains in the *Kiritimatiellota* phylum, no compartmentalization of the cytoplasm was observed in strain Nlcol2^T. There were unknown

inclusions or granules present inside the cells, and genes involved in the synthesis and utilization of polyphosphate and glycogen were found in the genome, which may serve as phosphate and energy storage materials, respectively.

(a) 

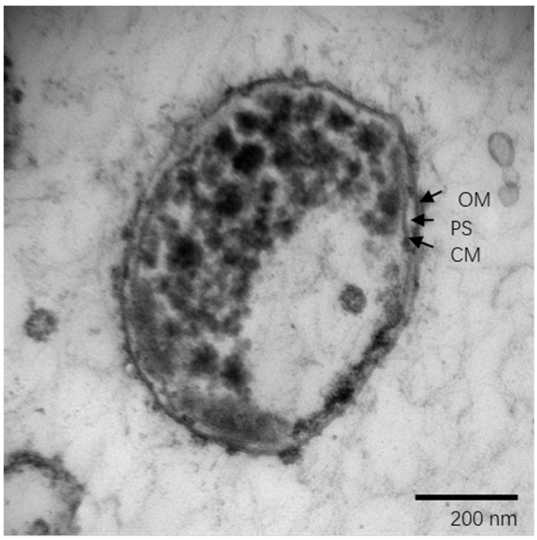
(b) 

Figure 1.5. Microscopic images of strain Nlcol2^T. (a) Bright field microscopy image – round dots are bacterial cells, ~1 um in diameter; (b) Transmission electron microscopy (TEM) image shows a Gram-negative cell wall structure. OM, outer membrane; PS, periplasmic space; CM, cell membrane.

Major cellular fatty acids (>10% of total) of strain Nlcol2^T include C18:0, *i*-C12:0, *i*-C18:0 and *i*-C14:0, in order of abundance. The major cellular fatty acid profile is quite different from *K. glycovorans* and *T. aerotolerans*, but almost the same as that in *P. desulfatans* and *P. sulfatireligans*, except that *P. sulfatireligans* also has *i*-C16:0 as one of the major components (**Table 1.1**). Again, this agrees with the phylogenetic placement of strain Nlcol2^T in the *Pontiella* genus, being more closely related with *P. desulfatans* than *P. sulfatireligans*. However, strain Nlcol2^T can be further distinguished by a relatively higher composition of *i*-C18:0 than *i*-C14:0, while *P. desulfatans* has more *i*-C14:0 than *i*-C18:0 (**Table S1.1**). Other cellular fatty acids detected in strain Nlcol2^T include C16:0, *i*-C16:0, C20:0, and *i*-C20:0 (**Table S1.1**).

Strain Nlcol2^T tested negative for both catalase and oxidase, which is common in strict anaerobes (**Table 1.1**). Beta-galactosidase was tested positive with ~ 5 nanomoles of substrate hydrolyzed in the API Zym assay.

Table 1.1. Comparison of phenotypic characteristics between strain Nlcol2^T and four other isolated strains in the *Kiritimatiellota* phylum. Notations: NA, data not available. +, positive; -, negative; +/-, unstable, ceasing growth upon the second transfer. Data for strains other than Nlcol2^T were referenced from literatures: a) van Vliet et al., 2020 (17); b) Spring et al., 2016 (14); c) Mu et al., 2020 (18). * Substrates were D-galactose for strain Nlcol2^T and D-glucose for other strains. ** For strain S-5007^T, acetate production was predicted from genomic data.

Strains	*P. agarivorans* Nlcol2^T	*P. desulfatans* F1$^{T\ a)}$	*P. sulfatireligans* F21$^{T\ a)}$	*K. glycovorans* L21-Fru-AB$^{T\ b)}$	*T. aerotolerans* S-5007$^{T\ c)}$
Isolation source	Microbial mat on marine sediment	Anoxic marine sediment	Anoxic marine sediment	Hypersaline microbial mat	Marine sediment
Cell shape	Spherical	Spherical	Spherical	Spherical	Spherical
Cell size (μm)	1.0	0.5-1.2	0.5-1.0	1.0-2.0	0.5-0.8
Motility	-	-	-	-	-
Genome size (Mbp)	4.44	8.66	7.40	2.95	3.88
DNA G+C content (mol%)	52.4	56.0	54.6	63.3	53.1
Major cellular fatty acids (>10% of total)	C18:0, *i*-C12:0, *i*-C18:0	C18:0, *i*-C12:0, *i*-C14:0	C18:0, *i*-C12:0, *i*-C18:0	*i*-C14:0, C18:0	C18:0, *i*-C12:0, *i*-C18:0, *i*-C16:0
Catalase activity	-	-	-	-	weak
Oxidase activity	-	-	+	-	-
Growth Temperature (°C)					
Range	10-37	10-30	0-25	20-40	15-45
Optimum	31	25	25	28	33-35
Growth Salinity (g/L NaCl)					
Range	10-60	10-31	10-50	20-180	5-80
Optimum	25-30	23	23	60-70	30-40
Growth pH					
Range	6.0-9.0	6.5-8.5	6.0-8.5	6.5-8.0	6.0-8.5
Optimum	8.0	7.5	7.5	7.5	7.0-7.5
Substrate utilization					
Glucose	+	+	+	+	+
Galactose	+	+	+	+/-	+
Fructose	+	+	+	-	+
Fucose	-	+	+	-	NA
Rhamnose	-	+	+	+/-	+
Mannose	+	-	+	+	-
Mannitol	+	-	+	-	-

Arabinose	-	+	-	-	+
Xylose	+	+	+	+	-
Lactose	+	+	+	-	NA
Cellobiose	+	+	+	-	+
Sucrose	+	+	+	-	-
Maltose	+	+	+	-	-
Fucoidan	+	+	+	+/-	NA
iota-Carrageenan	+	-	+	+/-	NA
Xylan	+	-	-	NA	NA
Agar	+	-	-	-	-
Major non-gaseous fermentation products *	Succinate, acetate, malate	Acetate, ethanol, lactate	Acetate, ethanol, lactate	Ethanol, acetate	Acid (maybe acetate **)
Nitrogen sources	N_2, NH_4^+	NH_4^+	NH_4^+	NH_4^+	NH_4^+

1.3.4 Physiology of growth

Strain Nlcol2^T exhibited consistent growth between 10-37 °C (optimum 31 °C), with NaCl concentration between 10-60 g/L (optimum 25-30 g/L), and with pH 6.0-9.0 (optimum pH 8.0) when D-galactose was utilized as the substrate. It was determined as a mesophilic and neutrophilic bacterium, which is similar to the other four isolated strains from the *Kiritimatiellota* phylum (**Table 1.1**). Growth with ammonium supplied in the medium was faster than when dependent on nitrogen fixation. The doubling times are 15 h and 65 h when growing with and without ammonium respectively, at room temperature (22 °C). Strain Nlcol2^T was considered as obligately anaerobic, being unable to grow with the presence of oxygen and even in non-reduced medium lacking sulfide as the reducing agent.

Strain Nlcol2^T was able to grow on various carbohydrate substrates under optimal conditions with ammonium supplied, which includes D-glucose, D-galactose, D-fructose, D-mannose, D-mannitol, D-xylose, D-cellobiose, lactose, sucrose, maltose, xylan, agarose, agar, and iota-carrageenan (**Figure 1.6**). No growth was observed when supplied with L-

fucose, L-rhamnose, D-arabinose, meso-inositol, starch, cellulose, alginic acid, or with fucoidan from *Macrocystis pyrifera*.

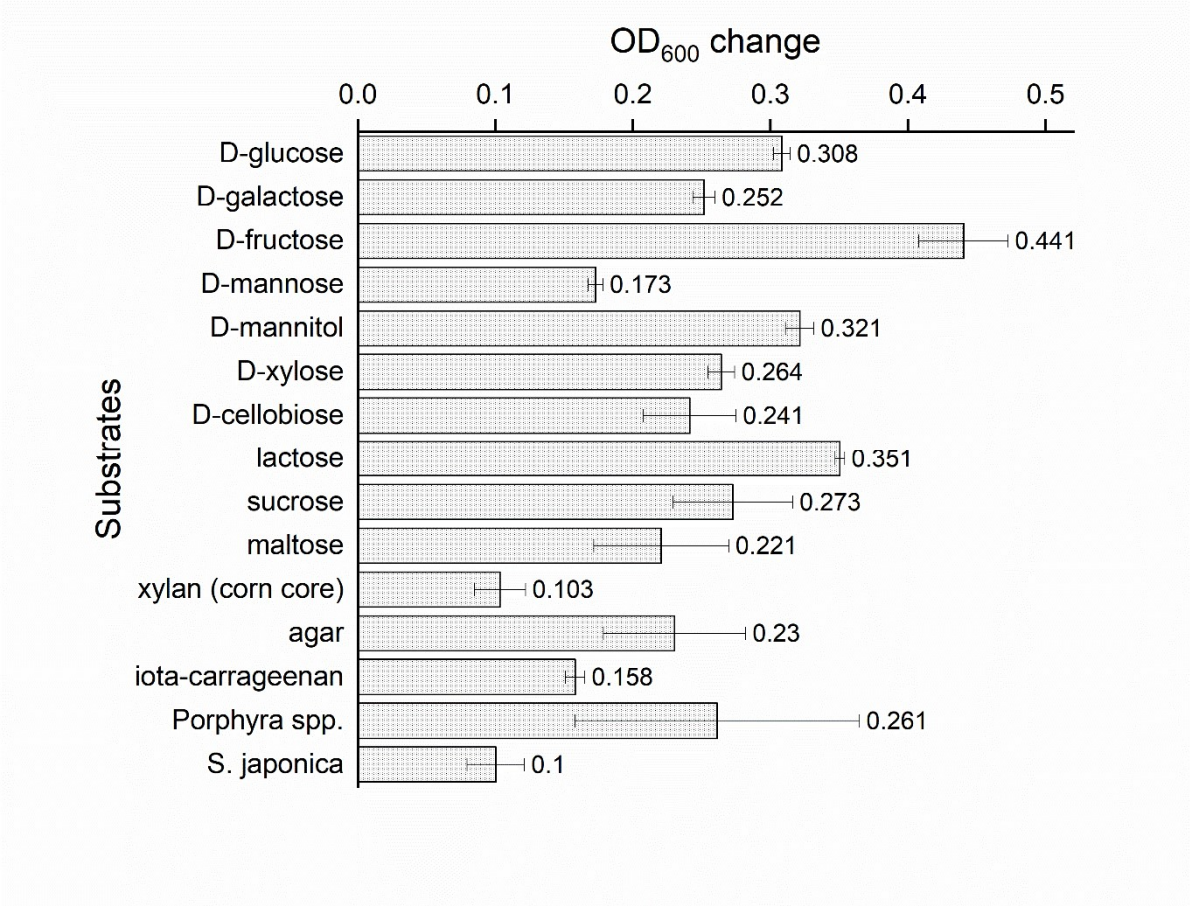


Figure 1.6. Bacterial growth observed by the change of optical density at 600nm (OD_{600}) between the start and end timepoints of the incubations when supplied with various substrates. The list of substrates only includes those that can support the growth of strain Nlcol2.

When growing on D-galactose, major fermentation products formed were succinate and acetate, with small amounts of malate and hydrogen gas also detected during the incubation (**Figure 1.7)**. Initially, the culture was supplied with 4.71 ± 0.12 mM D-galactose, and only 0.43 ± 0.06 mM D-galactose remained after the 10-day incubation period. Taking all fermentation products into consideration, the fractional electron recovery for galactose fermentation by strain Nlcol2^T was about 75%. The remaining electrons could be shunted to

and utilized by nitrogen fixation and biomass formation. A tentative fermentation balance was formulated as below, including measured fermentation products:

$$Galactose \rightarrow Succinate + Acetate + Malate + Hydrogen + \{biomass\}$$

$$4.3\ C_6H_{12}O_6 \rightarrow 3.2\ C_4H_6O_4 + 2.4\ C_2H_4O_2 + 0.7\ C_4H_6O_5 + 0.12\ H_2 + \{biomass\}$$

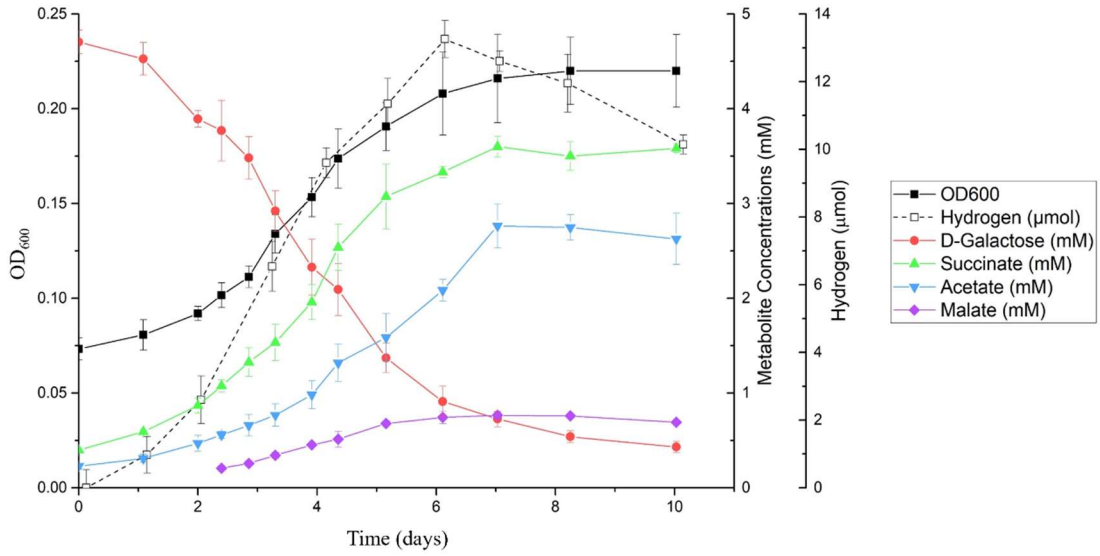


Figure 1.7. Time series of major metabolites from D-galactose fermentation by strain Nlcol2^T. Growth curve was measured by optical density at 600nm (OD_{600}). D-galactose decreased while fermentation products of succinate, acetate, malate, and hydrogen gas were produced during anaerobic bacterial growth on D-galactose.

1.3.5 Anaerobic degradation of macroalgal polysaccharides

1.3.5.1 CAZyme analyses

Microbial degradation of macroalgal polysaccharides involves complex metabolic pathways and requires a large number of enzymes during the process (Ficko-Blean et al., 2017; Hehemann, Correc, et al., 2012; Reisky et al., 2019; Sichert et al., 2020). CAZymes, especially glycoside hydrolases (GHs) and polysaccharide lyases, can break down polysaccharides into oligosaccharides (Drula et al., 2022). In the genome of strain Nlcol2^T,

202 genes (5.6% of predicted ORFs) were predicted to be CAZymes and associated carbohydrate-binding modules (CBM) by dbCAN2 meta server (H. Zhang et al., 2018). Among these, 164 genes were annotated to be in the glycoside hydrolase (GH) families. GH2, GH29, GH86 and GH117 are the most abundant families mainly represented by β-galactosidase, α-L-fucosidase, β-agarase and α-1,3-L-neoagarooligosaccharide hydrolase. 100 GHs were predicted with signal peptide sequences indicating 61% of GHs target to the cell membrane or can be exported outside of the cell. Extracellular and membrane associated GHs may hydrolyze large extracellular polymers that cannot otherwise enter the cell. Four porins and nine sugar transporters of the major facilitator superfamily were also present in the genome which may help with the acquisition of carbohydrate molecules by the cell.

1.3.5.2 Sulfatase analyses

As most marine polysaccharides are sulfated, another group of enzymes called sulfatases are needed in the degradation pathway, which can cleave sulfate ester groups off the carbohydrate backbone (Barbeyron et al., 2016). It has been shown that *Kiritimatiellota* as well as PVC superphylum have large numbers of copies of sulfatase genes in their genomes (van Vliet et al., 2019), and it is the same case in strain Nlcol2^{T}. We found 165 sulfatase genes (PF00884), comprising 4.6% of predicted ORFs in the genome.

Sulfatases are activated via post-translational modification by other enzymes before functioning. The most common one is formylglycine-generating enzyme (FGE) which transforms a cysteine or serine residue into a catalytic formylglycine (Appel & Bertozzi, 2015). These fGly-sulfatases are classified as type I sulfatases (family S1) which contain all carbohydrate sulfatases and is the largest sulfatase family (Helbert, 2017). Sulfatases were classified into 22 subfamilies in the SulfAtlas database (Barbeyron et al., 2016), all of which

belongs to family S1 fGly-sulfatases. The most abundant subfamilies (>5% of total sulfatases) are S1_16, S1_7, S1_15, S1_24, S1_8, S1_19, S1_17 and S1_20. Homologous sulfatases with known enzymatic activities within these subfamilies include D-galactose-6-sulfate 6-O-sulfohydrolase, endo-/exo-xylose-2-sulfate-2-O-sulfohydrolase, endo-/exo-galactose-4-sulfate-4-O-sulfohydrolase, endo-3,6-anhydro-D-galactose-2-sulfate-2-O-sulfohydrolase, exo-fucose-2-sulfate-2-O-sulfohydrolase etc., and the known substrates of these sulfatases include alpha-/iota-/kappa-carrageenan, fucan, ulvan etc.. These results imply that strain Nlcol2^T has the potential to target a vast variety of sulfated polysaccharides, similar to isolates *K. glycovorans* (Spring et al., 2016), *P. desulfatans*, and *P. sulfatireligans* (van Vliet et al., 2019, 2020). However, due to limited number of characterized fGly-sulfatases, there are still many unknowns about the specific substrates and/or reactions catalyzed by sulfatases in each subfamily (Barbeyron et al., 2016). 128 sulfatases have the best match genes from organisms in the PVC superphylum and 32 from *Bacteroidota*. 96% of sulfatases (158) were predicted to have a signal peptide sequence, indicating most sulfatases could be membrane-anchored or exported outside of the cell.

Although less well studied, the anaerobic sulfatase-maturing enzyme can mature either cysteine or serine sulfatases under anaerobic conditions (Berteau et al., 2006; Wagner & Horn, 2006). There are also five genes encoding formylglycine-generating enzyme and one encoding anaerobic sulfatase-maturing enzyme, which are essential for the activation of sulfatase by post-translational modification (Appel & Bertozzi, 2015; Berteau et al., 2006).

1.3.5.3 Growth on macroalgal polysaccharides

We further confirmed the ability of strain Nlcol2^T to grow on different macroalgal polysaccharides in live cultures. Bacterial growth was observed in anaerobic cultures with

agarose, agar, and iota-carrageenan, but not fucoidan. Many commercially bought algal polysaacharides are contaminated with co-extracted impurities, so we took direct measurements of polysaccharides to confirm the degradation of agar and iota-carrageenan by strain Nlcol2^T. Agar and iota-carrageenan concentrations were quantified as polymeric galactose after acid hydrolysis. Polymeric galactose of agar and iota-carrageenan decreased by 88% and 91% respectively, while the fermentation products of succinate and acetate increased by 88% ~ 93% along with bacterial growth (**Figure 1.8**). The carbon recovery rates are 78% and 87% for agar and iota-carrageenan degradation, respectively. This indicates that strain Nlcol2^T is able to degrade agar and iota-carrageenan and their growth were mainly fueled by these polysaccharides but not the impurities. This is the first strain reported with the ability of utilizing agar as substrate in the *Kiritimatiellota* phylum. We further tested their growth on seaweeds and cells also exhibit consistent growth on dried red algae (*Porphyra* spp.) and dried brown algae (*Saccharina japonica*), but not on the giant kelp (*Macrocystis pyrifera*). Since agar, porphyran, and carrageenan are all sulfated polysaccharides extracted mainly from red algae with a similar structure (Wei et al., 2013; Q. Zhang et al., 2004), it is not surprising that cells can grow on *Porphyra* spp. directly.

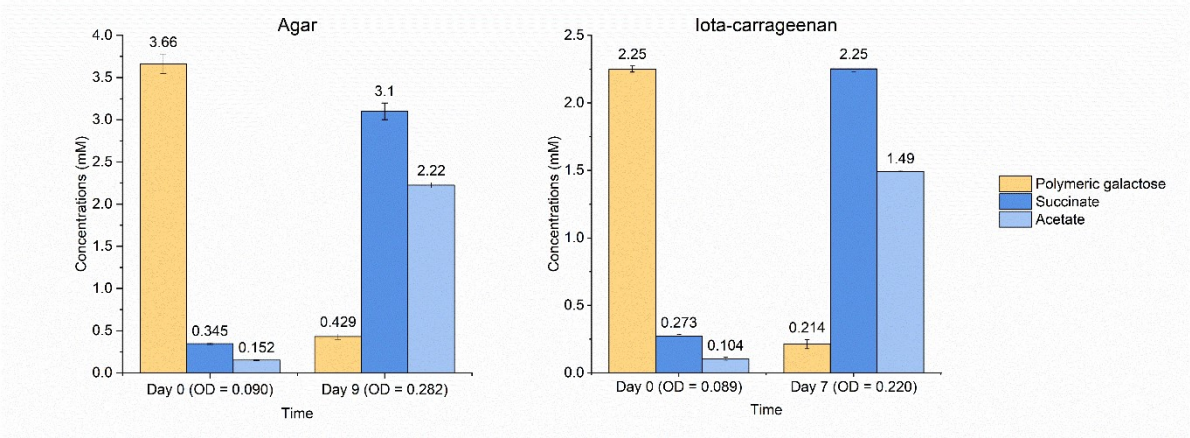


Figure 1.8. Agar and iota-carrageenan degradation by strain Nlcol2^T. Agar and iota-carrageenan concentrations were quantified as polymeric galactose after acid hydrolysis. Metabolites of succinate and acetate were measured as well. Bacterial growth was measured by optical density at 600nm (OD_{600}). Agar and iota-carrageenan concentrations decreased while fermentation products of succinate and acetate were produced during the growth of cultures.

Agar is a mixture of agarose and agaropectin which is commonly used as solidifying agent for culture media. Agarose is composed of alternating α-1,3 linked D-galactose and β-1,4 linked 3,6-anhydro-α-L-galactose with little sulfate modification, while agaropectin is heavily modified with sulfate (Anderson et al., 1965; Araki, 1956; Duckworth & Yaphe, 1971). Carrageenan is structurally related to agarose, except the β-linked unit is D-galactose-6-sulfate (Anderson et al., 1965). Fucoidan is also a sulfated polysaccharide composed mainly of L-fucose units adorned with sulfate esters, while minor xylose, galactose, mannose, glucuronic acid can be present too (Cong et al., 2016). Algal polysaccharide degradation has been well studied in *Zobellia galactanivorans* Dsij^T, the marine *Bacteroidota* model for the discovery of agarases, porphyranases, and carrageenases (Ficko-Blean et al., 2017; Hehemann, Correc, et al., 2012; Hehemann, Kelly, et al., 2012). We found potential genes not only involved in the degradation pathways of agar and iota-carrageenan, but fucoidan as well (**Figure 1.9; Table S1.2a, b, c**). Homologous genes encoding for potential β-agarases, ι-carrageenanases, and associated sulfatases were found in the genome of strain NLcol2^T and could be involved in degrading agar and iota-carrageenan into galactose and anhydrogalactose, which then can be directed to the central metabolism for energy. Potential fucoidanases were also found in the genome, but this contrasts with the experimental observation that cells did not grow with fucoidan from *M. pyrifera* as sole carbon source. However, bacterial growth was not supported by L-fucose either (see section 3.4), indicating that strain Nlcol2^T may house potential fucoidanases only to remove fucose

from fucoidan, but cannot further metabolize fucose and cannot gain energy from fucoidan degradation to support its growth. Alternatively, these genes may not encode fucoidanases to degrade fucoidan, but may encode enzymes for other purposes, for example, removing the fucose cap from mucin-like molecules (Glover et al., 2022).

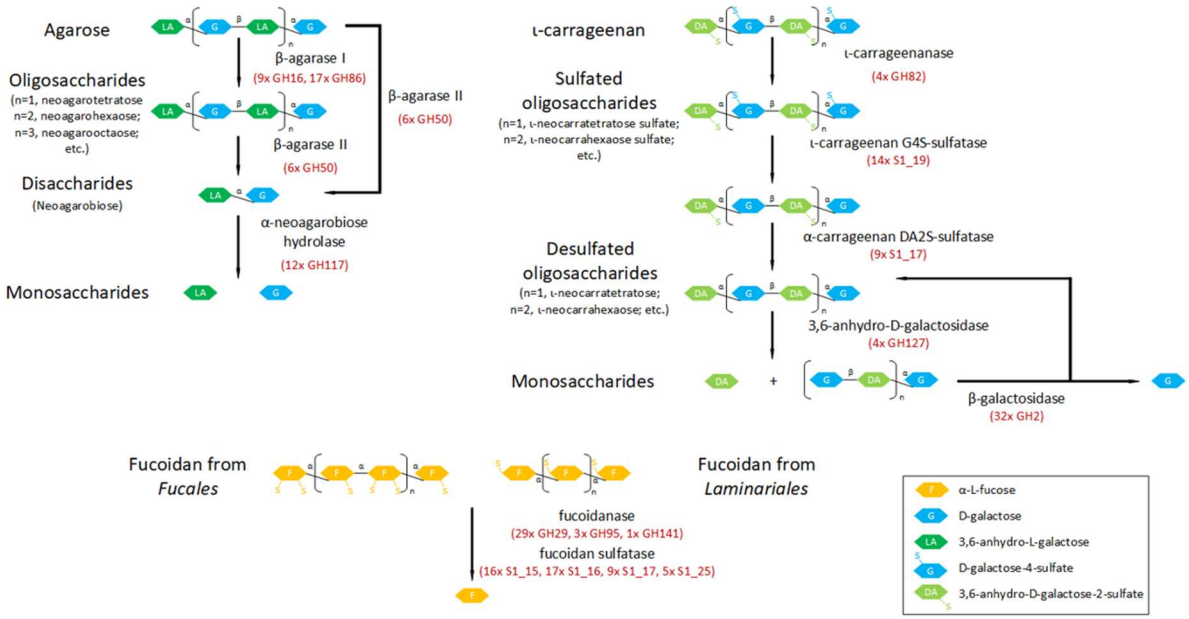


Figure 1.9. Homologs of key enzymes involved in the metabolic pathways of agar, iota-carrageenan, and fucoidan degradation were found in the genome. The number of gene copies encoding homologs of enzymes in the pathway are also listed. GH, glycoside hydrolase; S1_#, sulfatase classified in S1 family_subfamily. The GHs and sulfatases listed are potential agarases, carrageenanses, and fucoidanases.

A neighborhood analysis of the genome shows that GHs and sulfatases are often located nearby (within the distance of five genes), suggesting certain sulfatases and glycoside hydrolases could be regulated together to degrade sulfated polysaccharide (Grondin et al., 2017). In some cases, histidine kinase (PF07730, PF02518), response regulator (PF00072), and TonB-dependent transporters (PF00593, PF03544) are in the neighborhood too. The histidine kinase and response regulator together form a two-component signal transduction system that may help bacteria sense available substrates and respond to the changing environments (Skerker et al., 2005). There are cases when sulfatases themselves cluster

together, for example 4 or 6 copies in a row. A complete pathway for assimilatory sulfate reduction is also present in the genome and the cells may utilize the cleaved sulfate group for biosynthesis of reduced sulfur compounds.

A comparative study of GHs and sulfatases in selected genomes of the *Kiritimatiellales* order revealed that not all genomes harbor enzymes involved in degradation pathways of agar, iota-carrageenan, and fucoidan, and some bacteria don't have any GHs or sulfatases at all (**Figure 1.3**). However, certain genomes in the *Pontiella* genus show a relatively larger component of GHs and sulfatases. This indicates that these bacteria may adopt the lifestyle of utilizing macroalgal polysaccharides like agar, carrageenan, and fucoidan as carbon and energy sources, while other clades in the *Kiritimatiellales* order may specialize on other substrates available in their living environments. Some genomes in the *Pontiellaceae* family do not have high number of GHs or sulfatases either. This may indicate that these carbohydrate-related genes could be laterally transferred into the *Pontiella* genus but some were lost during evolution living in the environments where other available substrates were preferred. For example, such phenomenon was reported that the lateral gene transfer of porphyranases was from the marine *Bacteroidota*, *Zobellia galactanivorans* to the human gut bacterium *Bacteroides plebeius* (Hehemann et al., 2010). Another explanation would be that these MAGs were incomplete, and the GHs or sulfatases investigated were not easy to be captured.

1.3.6 Nitrogen fixation

We further tested nitrogen-fixing ability in live cultures of strain Nlcol2^{T}. The strain was able to grow on nitrogen gas as the sole N source in a nitrogen-free medium with D-galactose as carbon source. No growth was observed when nitrogen was replaced by helium

in the headspace. Bacterial growth was also supported by ammonium but not nitrate (**Figure 1.10**), and neither assimilatory nor dissimilatory nitrate reductase was present in the genome. Nitrogenase activity was detected by acetylene reduction assay. The production of ethylene from acetylene during bacterial growth on nitrogen gas as the sole nitrogen source showed that the cultures expressed active nitrogenases and could fix nitrogen gas into bioavailable forms to support their growth (**Figure 1.11**). This nitrogen-fixing ability may give them the advantage to survive in nitrogen-limiting environments.

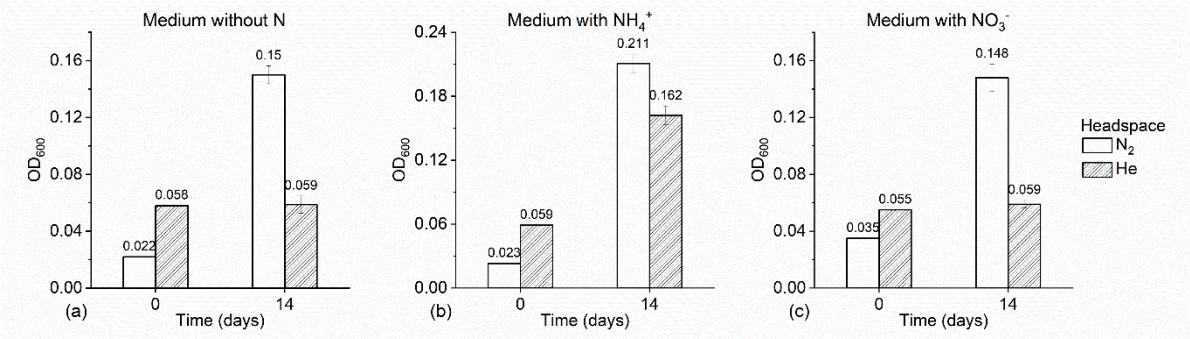


Figure 1.10. Comparisons of cultures growing in media with (a) no N, (b) ammonium or (c) nitrate supplemented. Headspace gases are nitrogen gas (white bar) or helium (grey bar). OD_{600} was measured at the beginning (day 0) and at the end (day 14). Bacterial growth was supported by nitrogen gas, ammonium, but not nitrate.

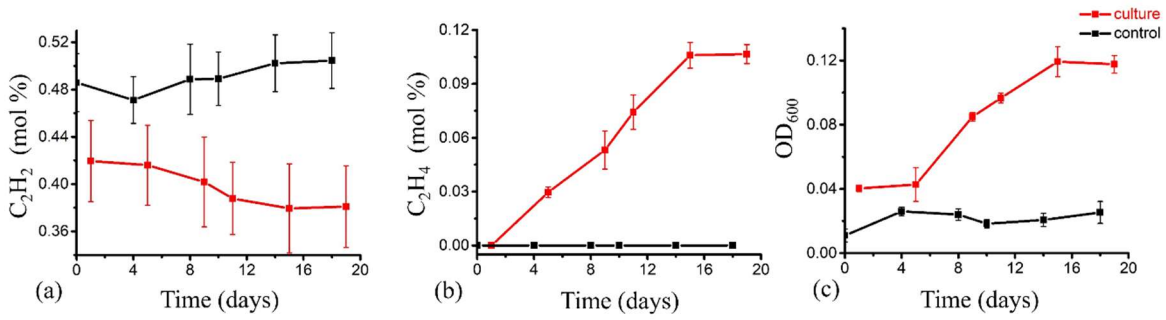


Figure 1.11. Nitrogenase activity of strain Nlcol2^T detected by acetylene reduction assay. Acetylene (a), ethylene (b) percentage and OD_{600} (c) change with time in liquid culture with D-galactose as substrate and nitrogen gas as the sole N source. Cultures with bacteria are labeled in red, and media without inoculum (control) are labeled in black. The decrease of acetylene and the increase of ethylene indicated active nitrogen fixation during bacterial growth.

Mo-dependent nitrogenase is the most common and widely studied enzyme that performs nitrogen fixation. It contains two components: an Fe protein as the reductase (*nifH*) collecting and transferring electrons, and a MoFe protein (*nifDK*) which binds dinitrogen (N_2) and converts it to ammonia (NH_3) (Seefeldt et al., 2009). Genes encoding both nitrogenase iron protein (*nifH*, PF00142) and nitrogenase molybdenum-iron protein alpha and beta subunits (*nifDK*, PF00148) are present in the genome, which together form a complete pathway of nitrogen fixation. No alternative vanadium-iron nitrogenase or iron-only nitrogenase was found. In addition to *nifHDK*, both *nifB* and *nifE* involved in the biosynthesis of nitrogenase MoFe cofactor are present in the genome. Two genes coding for nitrogen regulatory protein PII were present, which are important for the regulation of nitrogen fixation in response to nitrogen source availability (Arcondéguy et al., 2001). The rop-like protein is uncharacterized but often found in nitrogen fixation operons and may play a role in regulation (Buchko et al., 2009). There are various other *nif* genes present in other parts of the genome including *nifA, M, S, U, V* which together may help regulate the function of nitrogenase (**Table S1.3**).

Nitrogenases are highly oxygen-sensitive, but even though there are diverse anaerobes in the PVC superphylum, only a few studies demonstrated nitrogen fixation in this superphylum (Cabello-Yeves et al., 2017; Delmont et al., 2018; Khadem et al., 2010; Wertz et al., 2012) and no reports in the *Kiritimatiellota* phylum. Moreover, we have little knowledge as to where *nif* genes were acquired from by the nitrogen-fixing members in the PVC superphylum. We found 5 genomes in this phylum housing a *nifH* gene. Three were from *P. desulfatans*, *P. sulfatireligans*, and isolate S94, and two were from the marine sediments at the hydrothermal vent of South Mid-Atlantic Ridge (SZUA-380 and SZUA-

494). All *nifH* genes in this clade were classified as cluster III, which is dominated by distantly related obligate anaerobes (Zehr et al., 2003). All 6 *nifH* genes from the *Kiritimatiellota* phylum form a monophyletic clade with a bootstrap value of 89 (**Figure 1.12**). They also cluster together with sequences from *Chlorobi*, *Bacteroidota*, and *Delataproteobacteria* (mainly the *Desulfovibrio* genus), *Spirochaetes*, and some *Verrucomicrobia* to form a monophyletic clade with a bootstrap value of 85. This suggests that there could be lateral gene transfer between the *Kiritimatiellota* phylum and other phyla in this clade, but some bacteria in the *Kiritimatiellota* phylum may have lost *nif* genes during evolution. Nitrogen fixation genes in a methanotrophic Verrucomicrobial isolate *Methylacidiphilum fumariolicum* strain SolV resemble those from the *Gammaproteobacteria* which supports their acquisition of *nif* genes through lateral gene transfer (Khadem et al., 2010).

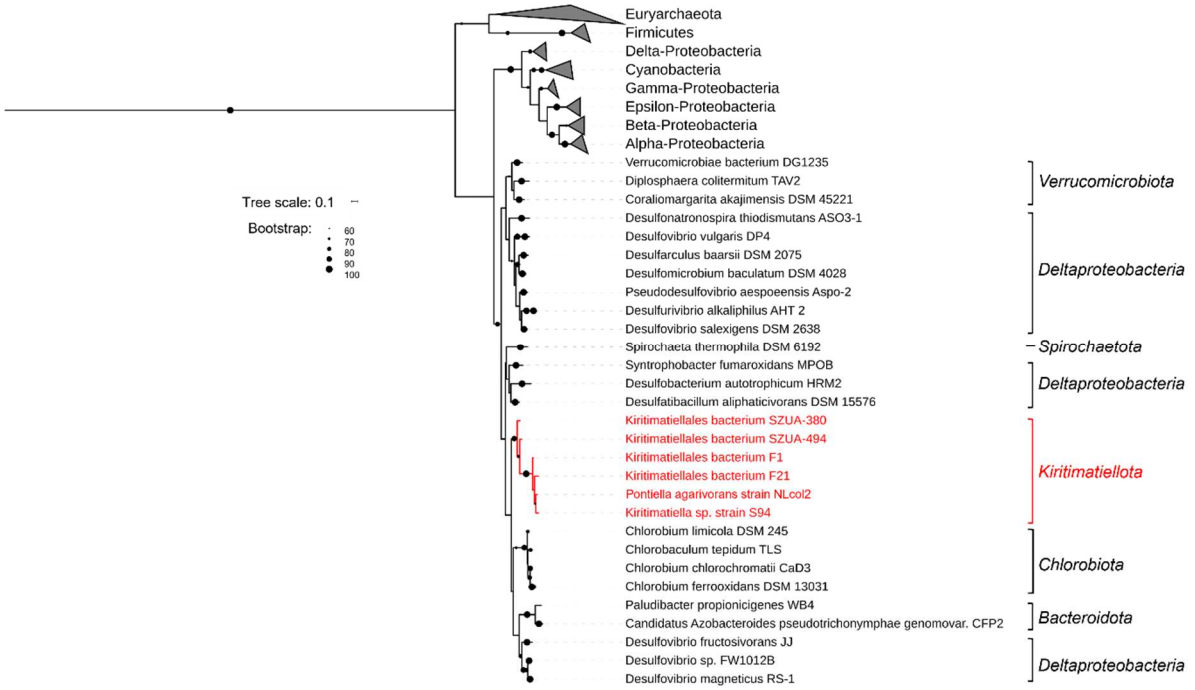


Figure 1.12. Phylogenetic placement of *nifH* genes in the Kiritimatiellota phylum among 879 *niH* genes from cultivated diazotrophs on a maximum-likelihood tree. Two light-independent protochlorophyllide reductases were included as outgroups (not shown) and the

tree was rooted there. The tree was pruned with 48 representatives of different phyla. Bootstrap values over 60 are shown in dots on the nodes.

1.4 Conclusion

In this study, we reported a novel anaerobic bacterial strain NLcol2^T isolated from microbial mats in marine sediments as the representative of a novel species in the *Pontiella* genus, which is the fifth cultivated strain in the *Kiritimatiellota* phylum. It represents the first strain to utilize agar as substrate with nitrogen-fixing ability in the *Kiritimatiellota* phylum. An extensive list of CAZymes and sulfatases shows its potential to degrade diverse macroalgae-derived sulfated polysaccharides in marine environments.

Description of *Pontiella agarivorans* sp. nov.

Pontiella agarivorans (a.ga.ri.vo'rans. N.L. neut. n. *agarum* agar, algal polysaccharide; L. pres. part. adj. *vorans* devouring, consuming; N.L. part. adj. *agarivorans* agar-devouring).

Cells are Gram-negative, anaerobic, non-motile cocci with a diameter of 1 μm. No spore formation was observed. Cells divide by binary fission. Colonies on agar plates are milky or ivory, circular, and smooth. Growth occurs at 10-37 °C (optimum 31 °C), with NaCl concentration between 10-60 g/L (optimum 25-30 g/L), and with pH 6.0-9.0 (optimum pH 8.0) when D-galactose was utilized as the substrate. The following substrates support growth: D-glucose, D-galactose, D-fructose, D-mannose, D-mannitol, D-xylose, D-cellobiose, lactose, sucrose, maltose, xylan, agarose, agar, iota-carrageenan, and fucoidan. The following compounds do not support growth under laboratory conditions: L-fucose, L-rhamnose, D-arabinose, meso-inositol, starch, cellulose, or alginate. The non-gaseous fermentation products from D-galactose are succinate, acetate, and malate (traces). Both ammonium and nitrogen gas can be utilized as nitrogen sources, but nitrate and nitrite were not utilized. Major cellular fatty acids are C18:0, *i*-C12:0, and *i*-C18:0.

The type strain NLcol2^T (= DSM 113125^T = MCCC 1K08672^T), was isolated from microbial mats on the surface of marine sediments offshore Santa Barbara, California. Genome of the type strain is 4.4 Mbp in size and DNA G+C content is 52.4 mol%. The GenBank accession number for the full-length 16S rRNA gene sequence of strain NLcol2^T is OQ749723, and the genome of strain NLcol2^T was deposited at NCBI under the accession number JARVCO000000000.

Chapter 2. Transcriptomic analysis of macroalgal polysaccharide degradation by *Pontiella agarivorans* strain NLcol2

Abstract

Marine macroalgal polysaccharides represent a substantial carbon reservoir in the world's oceans. These polymers have complex structures with variations in carbohydrate monomers, glycosidic linkages, and are often sulfated. Microbial degradation of these algal polysaccharides requires an extensive number of enzymes such as carbohydrate-active enzymes (CAZymes) and sulfatases. In this study, we investigated the transcriptomic responses of *Pontiella agarivorans* strain NLcol2 grown with carbon sources of agar, brown seaweed (*Saccharina japonica*), and red seaweed (*Porphyra haitanensis*) compared with the monosaccharide D-galactose as control condition. RNA sequencing (RNA-seq) revealed distinct gene expression profiles under different growth conditions while nitrogen-fixation genes were highly active under all conditions as nitrogen gas was the sole N source. 55% of the genome was differentially expressed with high significance (FDR < 0.05) when grown with polymers. Most relevant glycoside hydrolases (GHs) and sulfatases involved in the degradation pathways were up-regulated significantly in polymer conditions. Differential gene expression analysis also provided insight into new GHs and sulfatases with potential enzymatic activities on agar degradation. 12 putative polysaccharide utilization loci (PULs) were identified for agar degradation, containing GHs, sulfatases, transporters, and regulatory elements, with most genes showing higher expression levels for growth on agar. Finally, we found flagella-dependent cell motility and capsule polysaccharide synthesis for biofilm formation when grown with seaweed samples. Our results support a selfish mechanism in

strain NLcol2 for the depolymerization of agar to prevent the loss of oligosaccharide hydrolysates and maximize the benefits gained from agar degradation.

2.1 Introduction

Macroalgae, also commonly known as seaweeds, thrive in global oceans and are a large reservoir of carbon in the marine environment (Krause-Jensen & Duarte, 2016). The so-called blue carbon fixed in macroalgae and the export of algal detritus to the deep ocean has gained popular attention as a potential carbon dioxide removal (CDR) strategy to sequester atmospheric CO2 in the ocean (National Academies of Sciences, 2022). It has been shown that 24% of macroalgae can eventually reach the seafloor at 4,000 m (Ortega et al., 2019) and blue carbon is considered as sequestered on a geological timescale once it enters the deep ocean. However, the ultimate fate of macroalgae-derived carbon remains unclear due to a limited number of studies. And it is important to know the details of its fate – either burial into marine sediments, decomposition by grazers, or remineralization by microbial communities (Dolliver & O'Connor, 2022) – to inform potential impacts.

Marine algal polysaccharides are important cell wall components which can vary within an organism by physiological need and between different organisms based on evolutionary trends. This variability manifests with different carbohydrate monomers, glycosidic linkages, and modifications with chemical groups like sulfate (Arnosti et al., 2021). The polymer substrates used in this transcriptomic study are agar, brown seaweed (*Saccharina japonica*), and red seaweed (*Porphyra haitanensis*). Agar is a sulfated heterogenous polysaccharide extracted from red seaweeds (Class *Rhodophyceae*) (Difco™ & BBL™ Manual, 2nd Edition). It is a mixture of 70% agarose and 30% agaropectin. Agarose is composed of alternating α-1,3 linked D-galactose and β-1,4 linked 3,6-anhydro-α-L-

galactose with little sulfate modification, while agaropectin is more varied and is heavily modified with sulfate (Anderson et al., 1965; Araki, 1956; Duckworth & Yaphe, 1971; Usov, 2011). *Saccharina japonica* is a type of brown seaweed (Class *Phaeophyceae*) that is widely grown and harvested in the East Asia for its nutritional benefits. It has also attracted attention as extracted fucoidans possess antitumor activity (Vishchuk et al., 2011). Previous studies show that the sulfated polysaccharide from *S. japonica* is highly branched galactofucan, composed mainly of L-fucose, D-galactose, sulfate groups, and other minor monosaccharide residues including mannose, xylose, rhamnose, and glucuronic acid (Vishchuk et al., 2011; J. Wang et al., 2010). *Porphyra haitanensis* belongs to red algae and is also an edible and economic food with rich nutritional value. Its polysaccharide composition is mainly porphyran, a sulfated polysaccharide with a linear backbone of alternating D-galactose and 3,6-L-anhydrogalactose (Gong et al., 2018; Q. Zhang et al., 2004). It has similar disaccharide units with agarose, except for the 6-sulfate decoration on some residues in porphyran. The various biological activities of porphyran such as antioxidant and immunomodulatory activities have been the focus of research investigations in recent years (Gong et al., 2018; Wu et al., 2020; Q. Zhang et al., 2004; M. Zheng et al., 2023).

Macroalgal polysaccharides vary with their complex compositions and structures, thus their degradation by microbes involves complex metabolic pathways and requires a large number of enzymes (Ficko-Blean et al., 2017; Hehemann, Correc, et al., 2012; Reisky et al., 2019; Sichert et al., 2020). For example, hundreds of enzymes were shown to be involved in fucoidan degradation in *Verrucomicrobia* (Sichert et al., 2020). Classic genomic features for glycan degradation known as polysaccharide utilization loci (PULs) are found in

Bacteroidota bacteria, especially the starch utilization system (Sus) of *Bacteroides thetaiotaomicron* (Grondin et al., 2017; Martens et al., 2009). They contain co-regulated gene clusters of cell surface glycan-binding lipoproteins (SGBPs) for the detection of carbohydrates (typically susD homologs), TonB-dependent transporters (typically susC homologs) for the uptake of molecules into the cell, carbohydrate-active enzymes (CAZymes) for substrate-specific degradation, as well as sensors and transcriptional regulators. CAZymes, especially glycoside hydrolases (GHs) and polysaccharide lyases, can break down polysaccharides into oligosaccharides (Drula et al., 2022). Algal polysaccharide degradation has been well studied in *Zobellia galactanivorans* DsijT, the marine *Bacteroidota* model for the discovery of agarases, porphyranases, and carrageenases (Ficko-Blean et al., 2017; Hehemann, Correc, et al., 2012; Hehemann, Kelly, et al., 2012). Many non-canonical PULs have also been reported in various species without the *susCD* gene pair, providing alternative models and mechanisms for polysaccharide depolymerization (Grondin et al., 2017). As most marine polysaccharides are sulfated, sulfatases which can cleave sulfate ester groups off the carbohydrate backbone are also necessary in the polysaccharide degradation pathway in many marine bacteria (Barbeyron et al., 2016; Helbert, 2017). Therefore, PULs in marine bacteria often contain sulfatase genes and act specifically on substrates of sulfated polysaccharides (Helbert, 2017).

Different microbial strategies of polysaccharide hydrolysis and access to the resulting hydrolysate have been discovered and viewed as a cost-benefit problem. It is only beneficial to the enzyme-producing bacteria to obtain hydrolysates that can compensate the cost of biosynthesis and production of extracellular enzymes (Arnosti et al., 2021). Microorganisms with the ability to depolymerize polysaccharides can behave as either “sharing” or “selfish”

hydrolyzers. Sharing hydrolyzers release extracellular enzymes into the external environment, and thus allow "scavenging" bacteria (Reintjes et al., 2019) that have no depolymerization ability or even other competitors to take up the resulting hydrolysate. In contrast, selfish hydrolyzers only produce endo-acting enzymes for the initial hydrolysis of polysaccharides and uptake the hydrolysis products to prevent the loss of hydrolysates to scavengers. They do so by using TonB-dependent transporters to uptake oligosaccharides into the periplasmic space, followed by further depolymerization by exo-acting enzymes, and finally utilizing monosaccharides inside of the cell (Martens et al., 2009). Selfish bacteria are widespread in surface ocean waters (Reintjes et al., 2017) and this three-player model has been proposed as an important mechanism of polysaccharide hydrolysis (Arnosti et al., 2021).

Pontiella agarivorans strain NLcol2 is an anaerobic species in the *Kiritimatiellota* phylum with the ability to ferment marine macroalgal polysaccharides (Liu et al., submitted). It encodes a high number of CAZyme and sulfatase genes in its genome and is capable of degrading sulfated polysaccharides as a primary source of energy and carbon for growth. Here, we describe a transcriptomic study on strain NLcol2 to better understand its gene expression when growing on agar, dried brown seaweed *Saccharina japonica*, and dried red seaweed *Porphyra haitanensis*, compared with D-galactose as the control condition. We aim to find highly expressed CAZymes and sulfatases as well as their neighbor genes with co-regulation patterns harbored within potential PULs. These comparisons provide new insight into the enzymes and pathways for anaerobic depolymerization of macroalgal polysaccharides, as well as new perspective as to anaerobic

microbial strategies of macroalgal polysaccharide degradation and energy conservation in anoxic marine settings.

2.2 Methods

2.2.1 Growth conditions

Cultures of *Pontiella agarivorans* strain NLcol2^T were grown in anaerobic media as previously described in Chapter 1 with four different substrates (0.2% w/v) in triplicate bottles: D-(+)-galactose; agar (BD Difco Agar, Noble); dried red seaweed *Porphyra haitanensis* (commercially bought); and dried brown seaweed *Saccharina japonica* (commercially bought). Bacterial growth on the monomer D-galactose was taken as a reference condition compared to growth on the three polymers. N_2 gas was supplied as the only nitrogen source for cells, to facilitate nitrogen fixation; ammonium was not supplied. Cultures were passed on for three generations to ensure cells adapted to the relevant growth substrates. Growth was monitored by measuring optical density at 600 nm wavelength (OD_{600}). Cells were harvested in their fourth generation during mid-log phase in the anaerobic chamber and centrifuged at 5,000 × g for 10 minutes into cell pellets. Cell pellets were immediately frozen in liquid N_2 and preserved at -80 °C for at most 2 weeks until RNA extraction.

2.2.2 RNA extraction and sequencing

Total RNA was isolated using Qiagen RNeasy Mini kit with RNase-free DNase set (Qiagen) to remove genomic DNA. All samples were further cleaned and concentrated using Zymo RNA Clean & Concentrator with DNase kit (Zymo Research). RNA concentrations were quantified by Qubit RNA Broad Range Assay kit (Thermo Fisher Scientific), and RNA purity was checked by NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific). RNA

integrity number (RIN) was assessed by Agilent 2200 TapeStation System (Agilent Technologies). High-quality RNA samples (RINs > 8.0) were shipped to the Genome Center at University of California Davis for ribodepletion, library preparation, and RNA sequencing (RNA-seq). RNA-seq was performed with pair-ended reads at 150 bp on Illumina platform NovaSeq 6000 S4. In total, 16GB data was generated for 12 samples, and on average, 9.4 million reads were generated for each sample with a mean quality score of 35.7 ± 0.1.

2.2.3 Transcriptomic analysis

Raw reads were first trimmed by Trimmomatic v.0.39 (Bolger et al., 2014) to remove adapter sequences, short and low-quality reads using the parameters "ILLUMINACLIP:TruSeq3-PE.fa:2:30:10:2:True LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:75". Trimmed reads were examined by FastQC v.0.11.9 for quality control before mapping to the genome. Reads were aligned to the draft genome of *Pontiella agarivorans* strain NLcol2 (NCBI accession number JARVCO000000000) by Bowtie2 v.2.4.4 (Langmead & Salzberg, 2012) using default parameters. The overall alignment rate was 96.66% - 99.37% for all 12 samples using SAMtools v.1.13.0 (H. Li et al., 2009). To count the number of reads mapped to genes, featureCounts within subread v.2.0.3 (Liao et al., 2014) was used with the parameter "--fracOverlap 0.8" for counting reads with a minimum of 80% overlapping bases, and the parameter "-M --fraction" for counting fractionally for multi-mapping reads.

Read count matrices were imported to R v.4.3.1 and analyzed through the edgeR pipeline (Chen et al., 2016; Robinson et al., 2010). In short, genes with very low counts were filtered out - only genes with at least 10 counts per million (cpm) reads in at least 3 samples were kept. Normalization was performed using trimmed mean of M-values (TMM)

and accounted for the effective library sizes for each sample. Normalized reads per kilobase of transcript per million mapped reads (RPKM) were calculated for triplicates in each growth condition, taking both gene lengths and library sizes into consideration. RPKM was taken as a scale for gene expression levels within group samples. Heatmaps of log_2RPKM were made by the pheatmap package in R. The values were centered and scaled to prevent the few highly expressed genes from overshadowing genes with lower expression. The scaling was performed both in the column direction to make within-sample comparison and in the row direction to better view between-sample differences. 58 genes were selected with log_2RPKM > 2 scaled by column in all conditions.

As for between-group comparison, differential expression was determined using quasi-likelihood (QL) F-test after data was fitted by the generalized linear models in edgeR. Differentially expressed genes (DEGs) were discovered with a false discovery rate (FDR, adjusted P-value by the Benjamini-Hochberg procedure) of <0.05. Fold changes of gene expression were converted to the log2 scale. Comparisons were made between the triplicates of agar, red seaweed, and brown seaweed against D-galactose, respectively.

The number of shared DEGs were visualized in Venn diagrams for comparison between growth conditions. The counts of Venn diagrams filtered by FDR < 0.05 plus logFC > 0 for up-regulated genes and logFC < 0 for down-regulated genes. Differential gene expression of CAZymes and sulfatases under different polysaccharide conditions were presented in volcano plots.

2.2.4 Polysaccharide utilization loci (PULs) analysis

CAZymes and CAZyme gene clusters (CGCs) were predicted from the genome of *P. agarivorans* strain NLcol2T using the dbCAN3 server

(https://bcb.unl.edu/dbCAN2/index.php) (J. Zheng et al., 2023). In short, automated CAZyme annotation was performed by HMMER search for CAZyme family annotation vs. dbCAN CAZyme domain HMM database, DIAMOND search for BLAST hits in the CAZy database, and HMMER search for CAZyme subfamily annotation vs. dbCAN-sub HMM database of CAZyme subfamilies (derived from eCAMI classification of CAZyDB families). CAZyme gene clusters were predicted by the CGC-Finder on dbCAN3 server with at least one CAZyme and one transporter detected within a maximum distance of two genes.

Putative polysaccharide utilization loci (PULs) for agar degradation were manually identified based on the CGCs predicted by dbCAN3 server and homologous *susCD*-like gene pairs, as well as their differential expression profiles in agar samples compared with D-galactose. Only CGCs containing GHs were considered for potential agar PULs. SusC-like protein homologs were detected by PFAM annotations of PF00593 (TonB dependent receptor), PF07715 (TonB-dependent receptor plug domain), PF13103 (TonB C-terminal), PF03544 (Gram-negative bacterial TonB protein C-terminal), PF01618 (biopolymer transport protein ExbB), and PF02472 (biopolymer transport protein ExbD). Sus-D like protein homologs were detected by PFAM annotations of PF07980 (SusD family) and the superfamily IPR011990 (Tetratricopeptide-like helical domain superfamily) to which it belongs. For genes clustered in and around CGCs, only those with co-regulation patterns determined from log fold changes and FDR values <0.05 in comparison of agar against D-galactose were considered as part of the potential PULs. Finally, putative PULs for agar degradation were visualized in Geneious Prime 2023.1.2 (https://www.geneious.com) with both their locations and expression levels.

Additionally, subcellular localization of CAZymes and sulfatases were predicted by PSORTb v.3.0.3 (Yu et al., 2010).

2.3 Results and Discussions

2.3.1 Growth curves

Cells exhibited consistent growth in all four conditions with D-galactose, agar, brown seaweed (*Saccharina japonica*), and red seaweed (*Porphyra haitanensis*) as substrates (**Figure 2.1**). The doubling times were 73 h, 96 h, 32 h, and 17.5 h, respectively. Cultures were harvested approximately at those times for downstream RNA extraction.

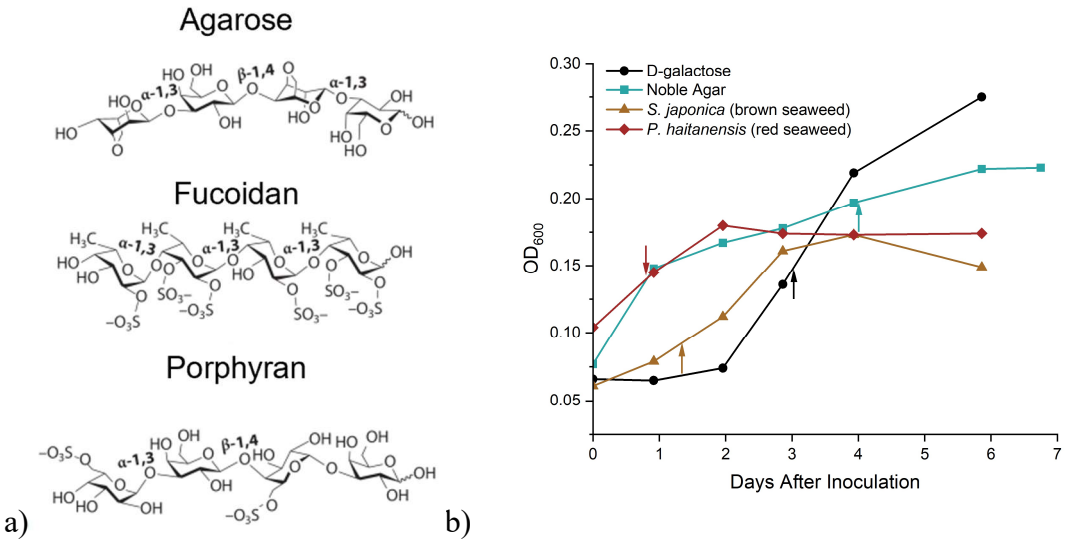


Figure 2.1. a) Structures of agarose, fucoidan, and porphyran as the major polysaccharides in agar, brown seaweed (*Saccharina japonica*), and red seaweed (*Porphyra haitanensis*), respectively. Structures are from Arnosti et al., 2021. b) Growth curves of strain NLcol2 with D-galactose, agar, brown seaweed (*S. japonica*), and red seaweed (*P. haitanensis*) as substrates. Growth was monitored with optical density OD_{600}. Arrows indicate the approximate time when subsequent cultures were harvested for RNA extraction.

2.3.2 Overview of gene expression profiles

Inter-sample relationships were examined by a multidimensional scaling (MDS), which enables visualization of the differences between the expression profiles from different

samples in two dimensions of leading fold change on a log2 scale. Triplicates from the same group of samples cluster together, while samples from different groups form separate clusters (**Figure 2.2**). This indicates that differential expression can be detected and is greater than the variance within group. Replicates of the transcriptomic samples showed high similarity and reproducibility.

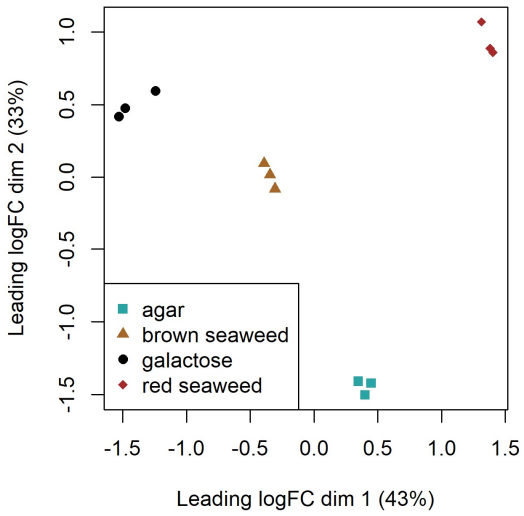


Figure 2.2. Multidimensional scaling (MDS) plot of gene expression profiles of the transcriptomic samples. The distances on two dimensions are leading log2 fold changes of pairwise comparison between the samples.

We first made within-sample comparison visualizing the top 58 genes that are highly expressed (normalized $\log_2$RPKM > 2 scaled by condition) in all four conditions (**Figure 2.3**). We identified seven *nif* genes and nitrogen fixation related regulatory genes highly expressed in all four conditions (**Table 2.1**), indicating active nitrogen fixation independent of growth substrate, on either monosaccharide or polysaccharides. There are also two genes encoding ammonium transporters and four genes encoding P_{II} proteins highly expressed in all four conditions. P_{II} proteins are transducers of the cellular nitrogen status that regulate the expression of ammonium transporters (AmtB family) by protein-protein interactions (Wirén

& Merrick, 2004). When ammonium concentration is low in the environment, P_{II} protein can sense the nitrogen limitation condition to activate ammonium transporter to uptake ammonium from the environment. This high expression of genes involved in nitrogen metabolism implies that cells dedicate substantial resources to acquiring nitrogen when grown with nitrogen-deplete saccharides.

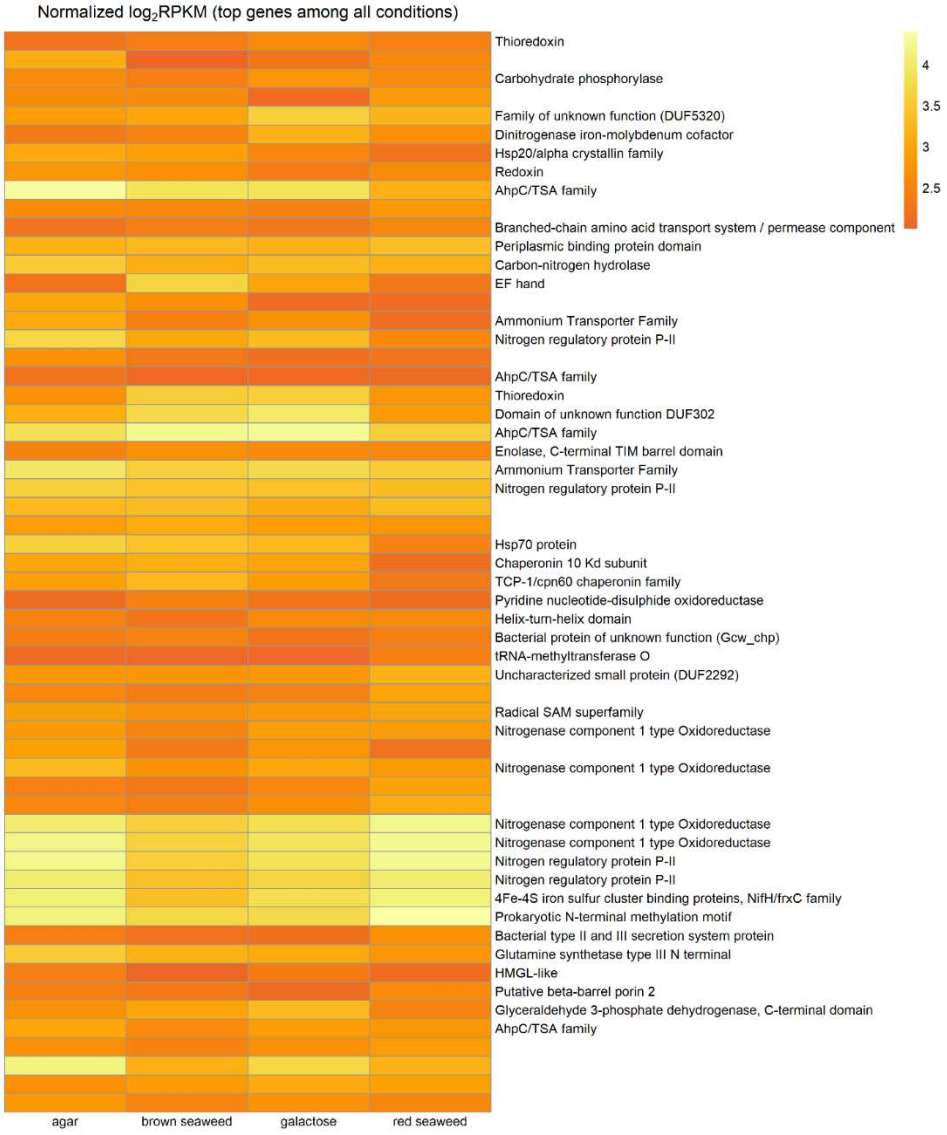


Figure 2.3. Heatmap of gene expression profiles represented by normalized log_2RPKM for top genes. The RPKM values of triplicates were calculated for all four conditions (shown in columns) and log_2RPKM values were centered and scaled in the column direction. A total of 58 highly expressed genes were selected with log_2RPKM > 2 scaled by column in all

conditions. Genes were ordered by their gene ID and their annotations were labeled on the side.

Table 2.1. List of nitrogen-related genes that are highly expressed (normalized log_2RPKM > 2 scaled by column) in all four conditions.

Gene ID	Annotation	Normalized log_2RPKM in agar	Normalized log_2RPKM in brown seaweed	Normalized log_2RPKM in galactose	Normalized log_2RPKM in red seaweed
c_000000000026_50	Dinitrogenase iron-molybdenum cofactor	2.33	2.49	3.18	2.67
c_000000000028_220	Ammonium transporter, Amt family	3.07	2.43	2.74	2.15
c_000000000028_221	Nitrogen regulatory protein P-II	3.70	3.03	3.29	2.57
c_000000000028_483	Ammonium transporter, Amt family	3.94	3.60	3.75	3.54
c_000000000028_484	Nitrogen regulatory protein P-II	3.63	3.44	3.41	3.34
c_000000000028_912	Nitrogen fixation protein NifB	2.95	2.70	2.83	3.00
c_000000000028_913	Nitrogenase component 1 type Oxidoreductase	2.84	2.51	2.89	2.86
c_000000000028_915	Nitrogenase component 1 type Oxidoreductase	3.29	2.74	3.04	2.85
c_000000000028_924	Nitrogenase component 1 type Oxidoreductase	4.03	3.61	3.82	4.20
c_000000000028_925	Nitrogenase component 1 type Oxidoreductase	4.19	3.65	3.90	4.26
c_000000000028_926	Nitrogen regulatory protein P-II	4.22	3.60	3.85	4.26
c_000000000028_927	Nitrogen regulatory protein P-II	4.03	3.40	3.68	4.03
c_000000000028_928	Nitrogenase iron protein NifH	4.08	3.37	3.80	4.11

To better view between-sample comparisons, a heatmap of normalized log_2RPKM values compared between four different conditions was made, showing clearly that different clusters of genes were relatively highly expressed (in bright yellow regions) under different conditions (**Figure 2.4**). The expression of specific gene sets induced by different substrates indicate that bacterial cells could sense, react to, and thus activate specific genes targeting different substrates available in the environment to support their growth.

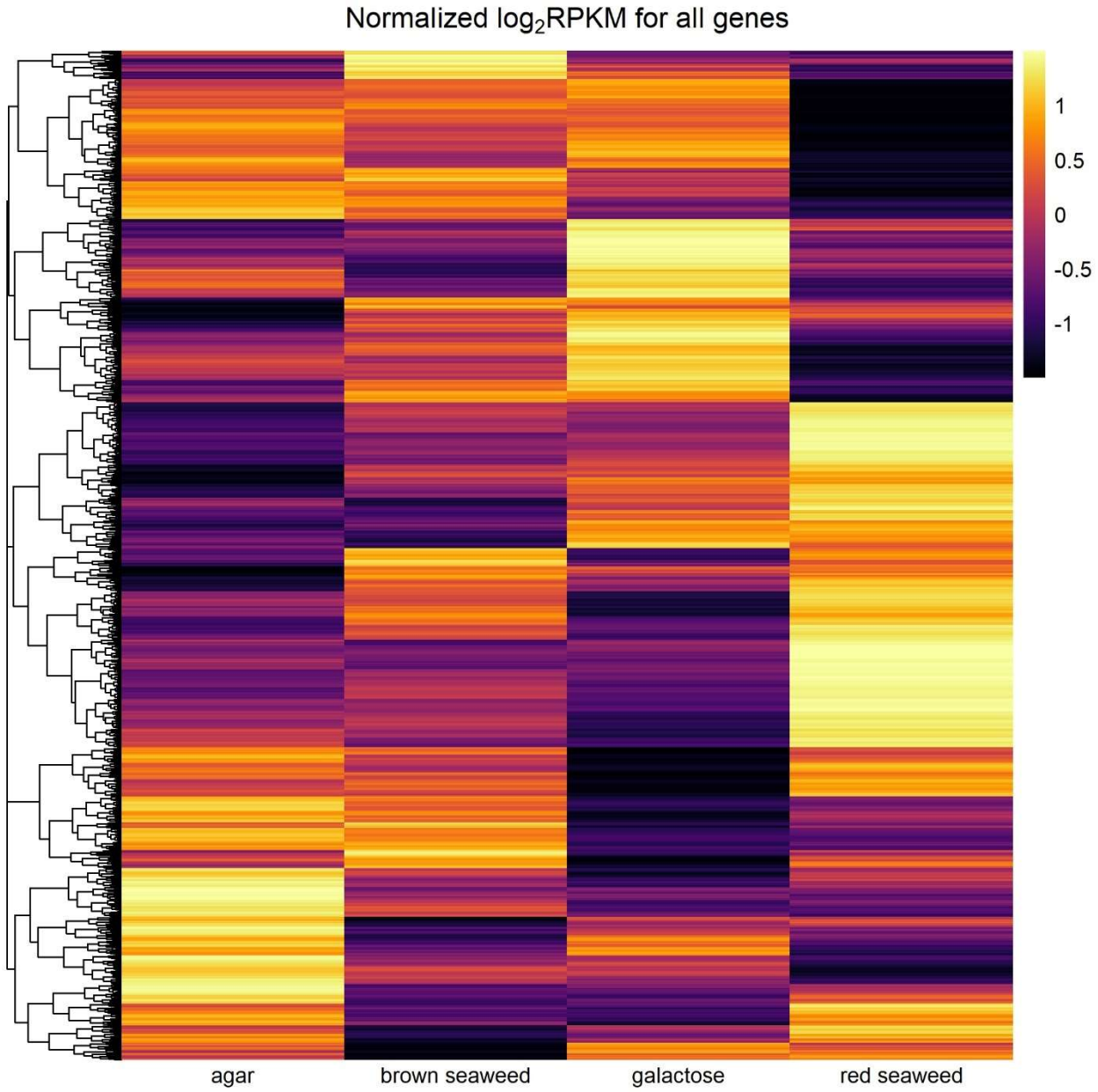


Figure 2.4. Heatmap of gene expression profiles represented by normalized log_2RPKM for all genes (shown in rows) included in the transcriptomic analysis. The RPKM values of triplicates were calculated for all four conditions (shown in columns) and log_2RPKM values were centered and scaled in the row direction. The greater value (yellow end) means relative higher abundance of transcripts expressed under the condition compared to other three conditions, and vice versa, the lower value (dark purple end) means relative lower abundance of gene expression compared to other conditions. Genes were clustered on the left side of the heatmap to highlight the regions with high levels of gene expression under different growth conditions.

To identify unique gene clusters that were relatively highly expressed under polymer conditions, we zoomed into the hot regions and searched for the top 100 genes as representatives of relatively highly expressed genes in agar, brown seaweed, and red seaweed samples respectively (**Figure 2.5**). In agar samples, we found a relatively high expression of GHs with potential agarases including GH16, GH50, GH86, GH117, as well as other CAZymes such as beta-galactosidase, beta-porphyranase, beta-L-arabinofuranosidase etc. There are also sulfatases belonging to S1_7, S1_15, S1_17, and S1_19 subfamilies at relatively high expression levels, containing potential endo-/exo-galactose-4-sulfate 4-O sulfohydrolase and galactose-6-sulfate 6-O sulfohydrolase (**Figure 2.5a**). Although the Noble Agar used in this study is recognized as extensively clean agar with high purity and contains only 0.663% of sulfate (Difco™ & BBL™ Manual, 2nd Edition), sulfatase genes were still expressed with relatively high abundance. In brown seaweed samples, there are alpha-L-fucosidases (GH29) and sulfatases (S1_16) expressed at relatively high levels. These genes may be involved in the degradation pathways of fucoidans, which are the major component of *S. japonica* (**Figure 2.5b**). In red seaweed samples, there are mostly regulatory genes and not many CAZymes or sulfatases that are relatively highly expressed among the top 100 genes (**Figure 2.5c**). However, if we expand

the search range to the top 500 genes, a number of GH16 (beta-porphyranase), GH117, GH29, and sulfatase transcripts showed up at relatively high abundance.

In general, besides CAZymes and sulfatases that are involved in the depolymerization of agar and seaweeds, transporters (e.g., L-fucose permease) and regulatory genes (e.g., histidine kinase and response regulator) were also highly expressed. These enzymes together help the cells sense substrate, activate depolymerization and transport the resulting oligosaccharides into the cell's interior.

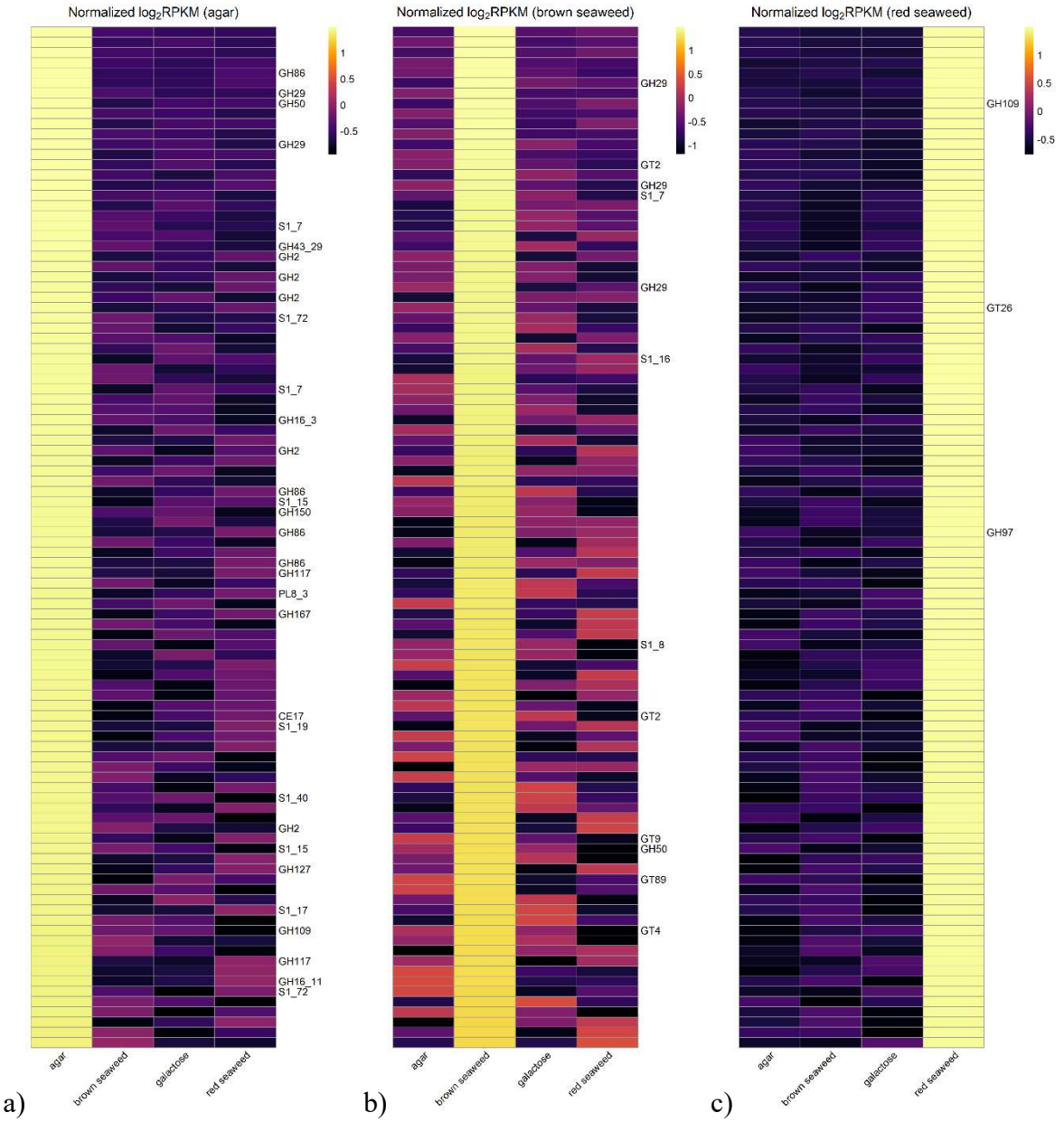


Figure 2.5. Detailed heatmaps of gene expression profiles represented by normalized log_2RPKM for the top 100 genes (shown in rows) that are relatively highly expressed in a) agar, b) brown seaweed, c) red seaweed samples respectively. The RPKM values of triplicates were calculated for all four conditions (shown in columns) and log_2RPKM values were centered and scaled in the row direction. Genes are ordered from high to low normalized log2RPKM scaled by row in each condition. Those on the top are the most distinctively expressed genes among different conditions. The greater value (yellow end) means relative higher abundance of transcripts expressed under the condition compared to other three conditions, and vice versa, the lower value (dark purple end) means relative lower abundance of gene expression compared to other conditions. Only gene annotations of CAZymes and sulfatases were labeled on the side for simplicity.

2.3.3 Differential gene expression analysis

Differentially expressed genes (DEGs) were investigated in agar, brown seaweed and red seaweed samples, each referenced to D-galactose. About 2000 genes (out of a total of 3611 ORFs in the genome) were found to be up- or down-regulated with significance (FDR < 0.05) when growing with agar and seaweeds compared to D-galactose (**Figure 2.6**). This indicates that about 55% of the genome was expressed differently when cells grew on polymers versus monosaccharide D-galactose. Among these DEGs, there are 529, 385, and 701 up-regulated genes with a substantial expression change over two fold in agar, brown seaweed, and red seaweed respectively compared with D-galactose.

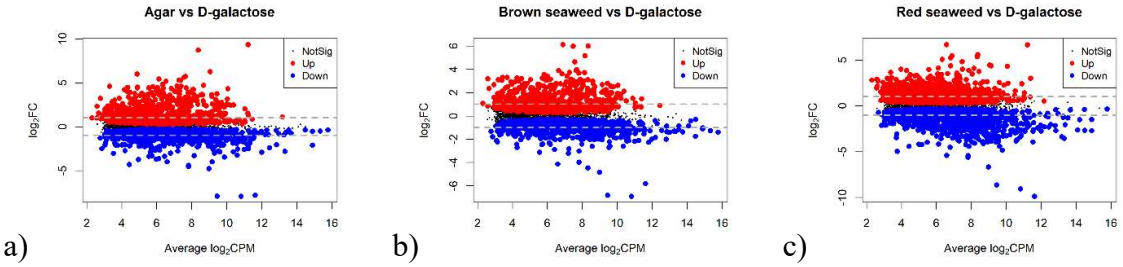
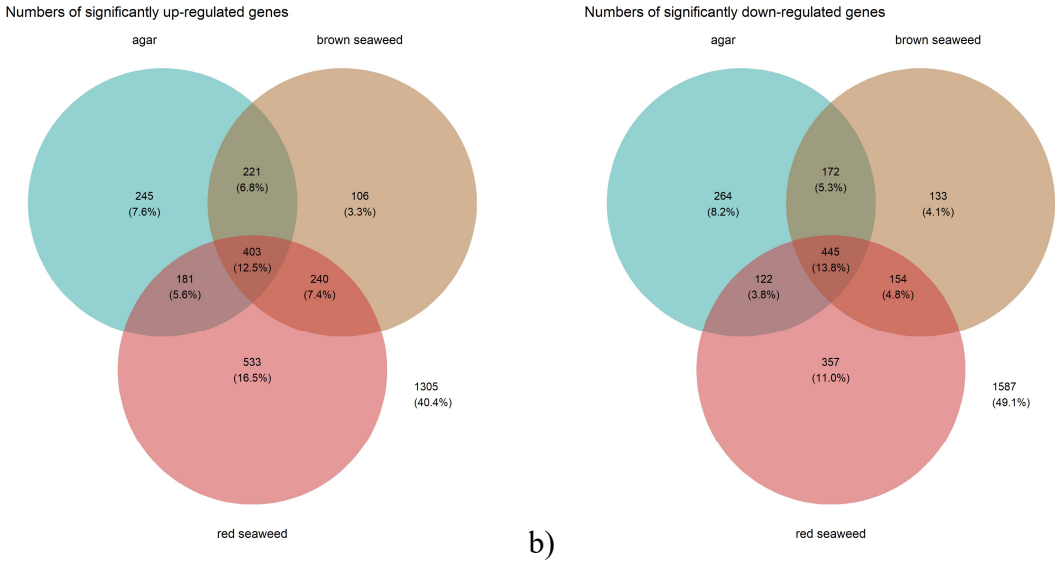


Figure 2.6. Mean difference plots of differential gene expression with a) agar, b) brown seaweed, and c) red seaweed referenced to D-galactose. Gene expression levels are on the x-axis represented by the average CPM (counts per million) values on a log2 scale of triplicates. Gene expression differences are represented by log fold changes (log_2FC) on the y-axis. Grey dashed lines indicate a log2 fold change of +/-1. Red and blue dots are up- and down-regulated genes with significance (P-value < 0.05).

To compare shared DEGs in these three polymer conditions, the numbers of significantly up- and down-regulated genes were visualized in Venn diagrams (**Figure 2.7**). Here, we found that about 13% of genes in the genome were significantly up-/down-regulated in all three polymer conditions compared with D-galactose. A subset of these genes are likely to encode polysaccharide-related enzymes and are potentially important resources for searching enzymes for polysaccharide degradation. For pairwise comparisons of DEGs shared between two conditions, about 18-20% of the genome were significantly up-/down-

regulated when strain NLcol2 was fed with either of the two polymers compared with D-galactose. There was no major difference of numbers of DEGs shared between any combination of two polymers, but the identity of DEGs differed for each comparison. As for the DEGs that were uniquely influenced by a single polymer compared with D-galactose, more DEGs (> 10%) were influenced by red seaweed alone than agar (~ 8%) than brown seaweed (~ 4%).



a) b)

Figure 2.7. Venn diagrams of significantly a) up-regulated and b) down-regulated genes shared between three conditions, relative to growth on D-galactose. Both the numbers of genes and percentages of the genome were shown in the circles. The numbers within circles are DEGs with an FDR < 0.05.

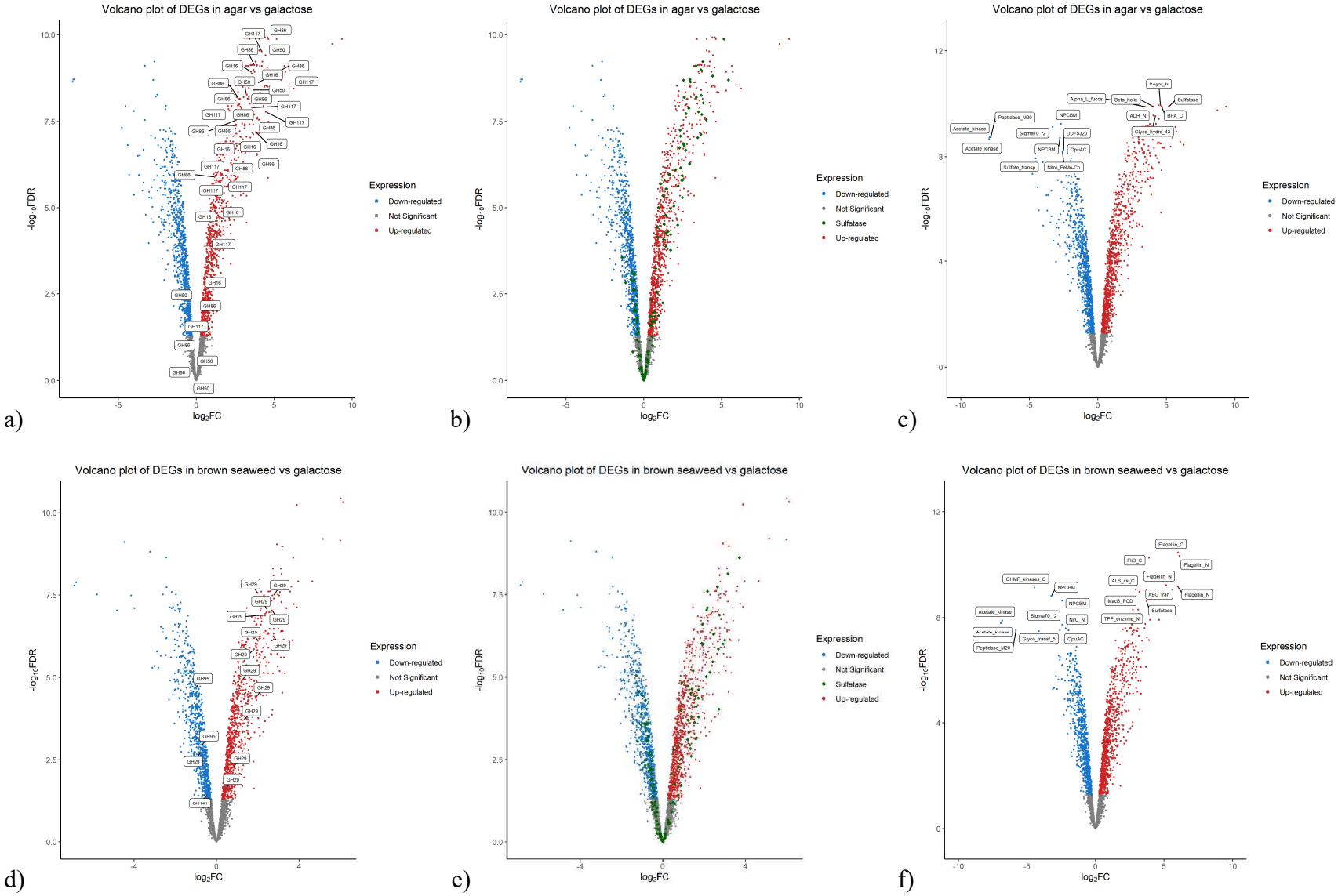
To further identify genes expressed with a large fold change that are also statistically significant, we used volcano plots to visualize significance (represented by false discovery rate, FDR) vs log fold change of the genes. The glycoside hydrolases containing potential agarases (GH16, GH50, GH86 and GH117) were labeled on the volcano plots of DEGs showing 85% of these GHs to be up-regulated with a large fold change and with high significance (**Figure 2.8a**). Brown seaweed *S. japonica* is mainly composed of fucoidan, and we found 13 out of 17 GHs containing potential fucoidanases (GH29, GH95, GH141)

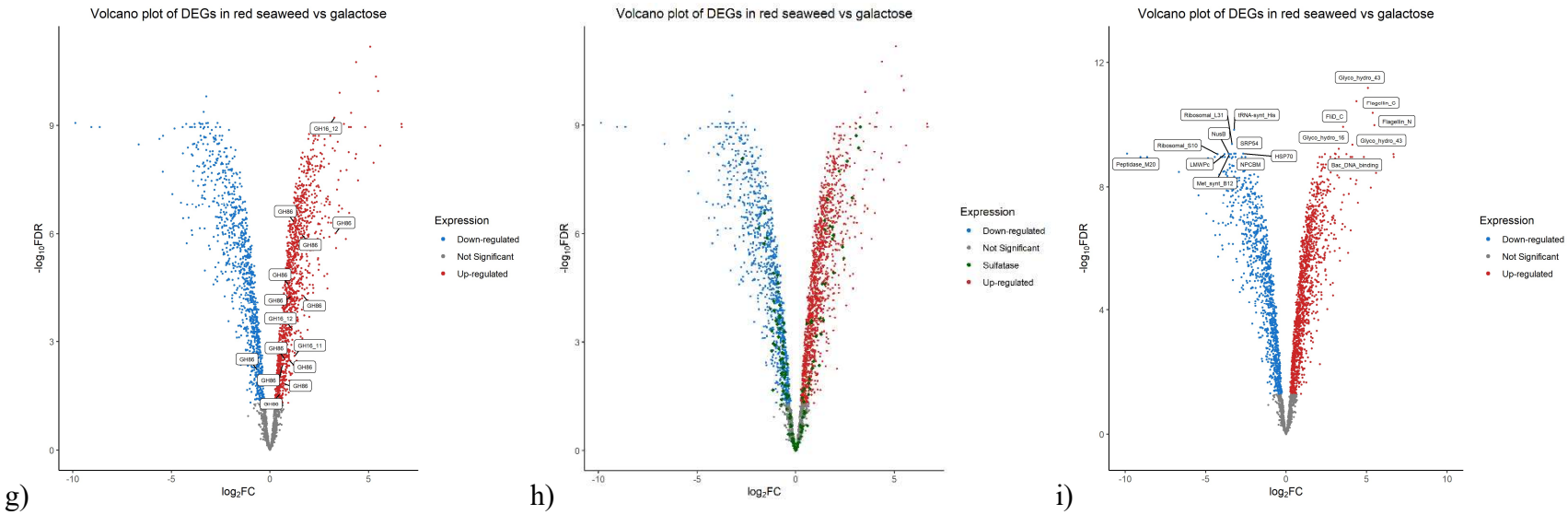
that were significantly expressed in brown seaweed samples to be up-regulated (**Figure 2.8d**). These 13 DEGs are all alpha-L-fucosidases (EC 3.2.1.51) which is an exo-acting hydrolase releasing monosaccharide L-fucose. In red seaweed *P. haitanensis*, 14 out of 15 beta-porphyranases (GH86, GH16_12, GH16_12) were significantly up-regulated (**Figure 2.8g**), which target the degradation of porphyran, the major polysaccharide component of *P. haitanensis*. A similar trend occurred in sulfatases - 73%, 55%, and 40% of significantly expressed sulfatase genes were up-regulated in agar (**Figure 2.8b**), brown seaweed (**Figure 2.8e**), and red seaweed (**Figure 2.8h**), respectively. These results agreed with our expectation that most GHs and sulfatases involved in macroalgal polysaccharide degradation exhibit greater expression when growing with algal polymers compared with D-galactose.

We further labeled the top 20 DEGs with the largest fold change and highest significance in agar (**Figure 2.8c**). The top 10 significantly up-regulated DEGs include beta-porphyranase, sugar transporter, sulfatase (S1_72), alpha-L-fucosidase etc., which are all enzymes related with sulfated polysaccharide degradation. Three out of 10 top up-regulated genes were without any annotations. A BLAST search of the protein sequence c_000000000020_63 with the longest length of 1015 amino acids showed that the top hits are all from *Kiritimatiellota*-related *Verrucomicrobia*. It is a hypothetical protein with PEP-CTERM sorting domain (hydrophobic, cross-membrane) and with a 30.66% amino acid sequence similarity with GH30. The GH30 hit was found in *Lentimonas* (a phylum in the PVC superphylum), and one of the relevant enzymes, endo-β-1,6-galactanase, performs endohydrolysis of (1→3):(1→6)-β-galactans to yield galactose and β-galactobiose as the final products. This indicates that c_000000000020_63 could be a new enzyme for hydrolyzing galactans in agar. The significantly up-regulated sulfatases with a high number

of counts belong to subfamilies S1_15 (n = 9), S1_24 (n = 9), S1_16 (n = 8), S1_19 (n = 8), S1_7 (n = 6), and S1_72 (n = 6). The sulfatases with known enzymes in these subfamilies include D-galactose-6-sulfate 6-O-sulfohydrolase (EC 3.1.6.4) and endo-/exo-galactose-4-sulfate-4-O-sulfohydrolase (EC 3.1.6.-). One of the sulfatases (c_000000000020_61) was among the top 10 significantly up-regulated genes with a 35-fold change of expression level. It is interesting to note that sulfatase subfamily S1_72 is currently without any known enzymatic activities in the SulfAtlas database, indicating that c_000000000020_61could be a new sulfatase that targets galactan substrates like agar and cleaves the sulfate ester group off agar.

Interestingly, besides GHs and sulfatases, genes relevant to flagellar formation increased expression by ~ 50-fold among the top 20 DEGs in both brown and red seaweed samples (**Figure 2.8f, 2.8i**), which may facilitate cellular movement toward nutrient resources (see more discussions in Section 2.3.6).


Volcano plot of DEGs in agar vs galactose
Volcano plot of DEGs in agar vs galactose
Volcano plot of DEGs in agar vs galactose
Volcano plot of DEGs in brown seaweed vs galactose
Volcano plot of DEGs in brown seaweed vs galactose
Volcano plot of DEGs in brown seaweed vs galactose
-log10FDR
log2FC
Expression
Down-regulated
Not Significant
Sulfatase
Up-regulated
a)
b)
c)
d)
e)
f)



g) h) i)

Figure 2.8. Volcano plots of DEGs in agar (a-c), brown seaweed (d-f), and red seaweed (g-i) versus D-galactose samples. Log fold change is on the x-axis and points with positive values (to the right of the x-axis) are up-regulated and points with negative values (to the left of the x-axis) are down-regulated. The false discovery rate (FDR, adjusted p-values) in its negative log form is on the y-axis. The higher the position of a point, the more significant its value is. Red dots are significantly up-regulated DEGs ($log_2FC > 0$ and $FDR < 0.05$) and blue dots are significantly down-regulated DEGs ($log_2FC < 0$ and $FDR < 0.05$). Grey dots are genes not significantly expressed ($FDR > 0.05$). a) Labels of GHs containing potential agarases (GH16, GH50, GH86, GH117) in agar samples; d) Labels of GHs containing potential fucoidanases (GH29, GH95, GH141) that are significantly expressed in brown seaweed samples; g) Labels of GHs containing porphyranases (GH86, GH16_11, GH16_12); b), e), h) green dots labeling all sulfatases; c), f), i) labels of the top 10 up- and down-regulated DEGs with the largest log fold change and highest significance in agar, brown seaweed, and red seaweed, respectively.

2.3.4 Polysaccharide utilization loci (PULs) for agar degradation

Putative polysaccharide utilization loci (PULs) were manually identified based on the CAZyme gene clusters (CGCs) predicted by dbCAN3 and homologous *susCD*-like gene pairs, as well as their expression levels in agar compared with D-galactose control condition. There were 12 PULs identified as potential gene clusters for agar degradation including significantly up-regulated glycoside hydrolases (GH16, GH50, GH86, GH117), sulfatases, transporters and/or regulatory elements with a co-regulation pattern (**Figure 2.9**).

The 12 PULs comprise 156 genes in total, including 65 GHs of 17 different families, 29 sulfatases of 14 different subfamilies, 8 transporters, and 5 pairs of two-component regulatory system genes. The GHs were most represented by GH86 (n = 13), GH117 (n = 10) and GH16 (n = 8) that contain homologs of agarases. The sulfatases were most represented by S1_19 (n = 6) and S1_15 (n = 5) which contain homologs of galactose sulfohydrolases. 76% of genes in these 12 PULs showed a significantly higher expression in agar samples and 63% were expressed by more than two folds compared to D-galactose control condition. Most PULs contain one or two two-component systems or transcription factors, including histidine kinase and response regulator, as well as AraC family and LacI family transcriptional regulators.

Though 10 genes of TonB-dependent transporter (susC-like proteins) were present in the genome of strain NLcol2, there were no significantly up-regulated TonB-dependent transporters included in the CGCs containing GHs relevant to agar degradation. There was no exact match of homologs of susD proteins found (PF07980, IPRIPR012944) in the genome either. However, four genes belonging to the tetratricopeptide-like helical domain superfamily (IPR011990) were found in gene pairs with susC-like proteins, which could be

potential susD-like proteins or other lipoproteins with similar function as susD. The non-canonical PULs in strain NLcol2, with a lack of *susCD* gene pairs, share similarities to those that occur in saccharolytic bacteria outside of *Bacteroidota* phylum (Ficko-Blean et al., 2017; Grondin et al., 2017).

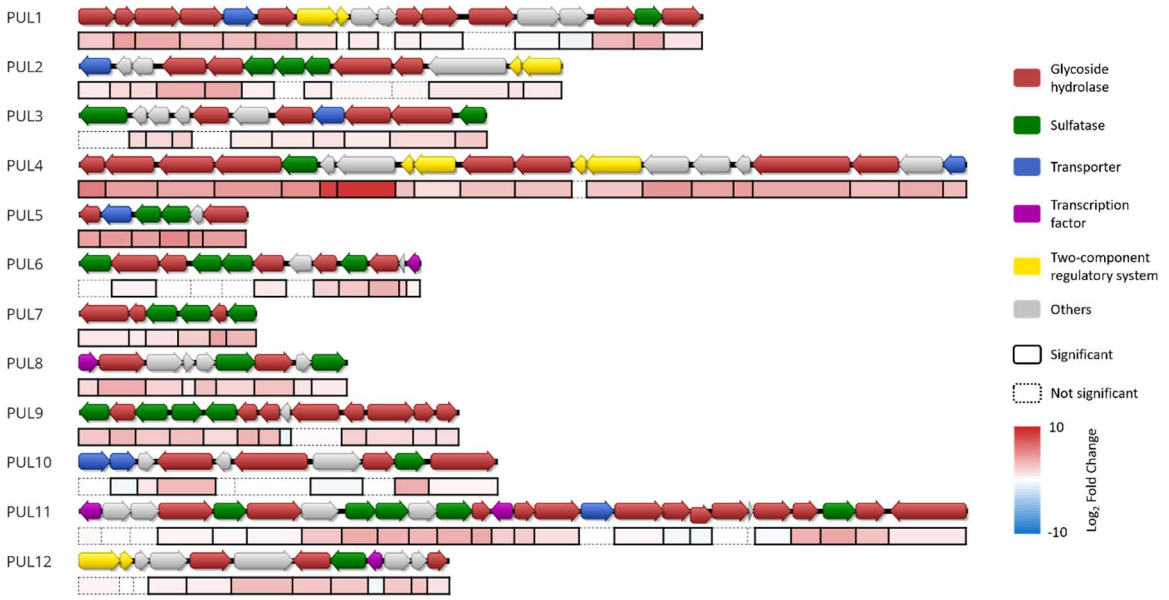


Figure 2.9. Putative polysaccharide utilization loci (PULs) for agar degradation found in the genome of strain NLcol2. Genes are shown as arrows and grouped in colors. The color boxes below the genes represent log2 fold change in the expression profile on agar compared with D-galactose. The box borders with solid lines represent significant (FDR < 0.05) differential gene expression in agar compared with D-galactose, and those in dashed lines represent not significantly expressed genes.

2.3.5 Microbial strategy for polysaccharide depolymerization in strain NLcol2

To finalize the catabolic pathway of agar degradation in strain NLcol2, the subcellular localization of CAZymes and sulfatases were predicted by PSORTb. Among the GHs containing potential agarases which were significantly up-regulated (logFC > 0 and FDR < 0.05) in agar samples compared with D-galactose, it predicted GH86 (n =13) to be extracellular, one GH16_12 and one GH117 to be periplasmic, GH117 (n = 5) and GH50 (n = 2) to be cytoplasmic. As for sulfatases that were significantly up-regulated (logFC > 0 and

FDR < 0.05) in agar samples compared with D-galactose, one was predicted as periplasmic and 28 as cytoplasmic.

As illustrated in **Figure 2.10**, agar could be initially hydrolyzed into oligosaccharides by endo-acting GH86 enzymes extracellularly. GH86 family contains endo-acting enzymes like β-agarases (EC 3.2.1.81) that hydrolyze (1→4)-β-D-galactosidic linkages in agarose, producing mainly tetramers of neoagarotetratose as products. The resulting oligosaccharides could be transported into the periplasmic space via TonB-dependent transporter (TBDT), which is driven by the proton motive force from associated ExbBD-TonB complex on the cytoplasmic membrane. No sulfatases that were significantly up-regulated in agar samples were predicted to be extracellular with confidence by PSORTb v.3.0, however, there were some predicted with a Sec- or lipoprotein-type signal peptide by SignalP v.5.0, which leaves open the possibility that sulfatases are translocated out of the cell. Thus, it remains uncertain whether desulfation occurs exclusively inside of the cell.

Following transport of oligosaccharides into the periplasm, other GHs and sulfatases there could break the glycosidic linkages and sulfate ester bonds to produce disaccharides of agarobiose and neoagarobiose. GH16_12 contains endo-acting β-porphyranase (EC 3.2.1.178) that cleaves the (1→4) linkages between β-D-galactopyranose and α-L-galactopyranose-6-sulfate in porphyran, forming mostly disaccharides. Though porphyranases were found to degrade porphyran, it has been shown that β-porphyranases can degrade the sulfated polymers in agar as well (Hehemann, Correc, et al., 2012). One sulfatase within S1_24 subfamily was predicted to be periplasmic, but unfortunately, we are aware of no enzymes biochemically characterized with known substrates. Again, as argued above, although most sulfatases were predicted to be cytoplasmic, many contain a Sec or Tat

signal peptide that could target the periplasmic area. It is possible that desulfation does not occur or occurs incompletely in the periplasm and that a fraction of the resulting disaccharides may remain sulfated. The disaccharides could possibly be further transported into the cytoplasm through the sugar transport system by major facilitator superfamily (MFS) transporter and sugar transporter proteins. A well-studied example is the *Escherichia coli* lactose permease that transports galactosides into the cell (Abramson et al., 2003). Sulfate ions could be transported by sulfate permease proteins on the inner membrane.

In the cytoplasm, disaccharides could undergo a final breakdown into monosaccharides and desulfation by more GHs and sulfatases. GH50 family contains both endo-acting β-agarases (EC 3.2.1.81) and exo-acting β-galactosidase (EC 3.2.1.23). Given the location of GH50 genes predicted to be in the cytoplasm, they would less likely be β-agarases that act on the agarose polymers, but more likely to be an exo-acting β-galactosidase that releases the terminal galactose residues in β-D-galactosides. GH117 family contains exo-acting α-agarases (EC 3.2.1.159) and β-galactofuranosidases (EC 3.2.1.146) that can release D-galactose and 3,6-anhydro-L-galactose. Concurrently, the remaining sulfate groups attached to the oligosaccharides could be released by sulfatases. One sulfate permease (c_000000000026_62) was significantly down-regulated when grown on agar, consistent with cellular sulfur acquisition from cleaved esters during growth on this sulfur-containing substrate.

Based on our results, we propose that strain NLcol2 adopted a selfish mechanism for agar depolymerization to maximize the uptake and utilization of oligosaccharide hydrolysates (**Figure 2.10**). In previous studies of starch degradation in *Bacteroides thetaiotaomicron* (Martens et al., 2009), agar degradation in *Zobellia galactanivorans*

(Hehemann, Correc, et al., 2012), and fucoidan degradation in *Lentimonas* strain CC4 (Sichert et al., 2020), selfish polysaccharide degraders were proposed to perform the degradation from extracellular depolymerization to the complete breakdown in the periplasm and transport monomers only into the cytoplasm. However, in our model we found that most expressed exo-acting GHs were cytoplasmic and would thus directly release monosaccharides in the cytoplasm. Additionally, there were no genes encoding galactose permease found in the genome. Our model suggests the uptake of disaccharides into the cytoplasm and degradation to be completed inside of the cytoplasm. This indicates a strengthened version of the selfish mechanism to prevent the loss of disaccharides, since disaccharides could be passively transported out of the cell through porins. For example, LamB proteins are thought to facilitate the diffusion of maltodextrins in *Escherichia coli* (Schirmer et al., 1995).

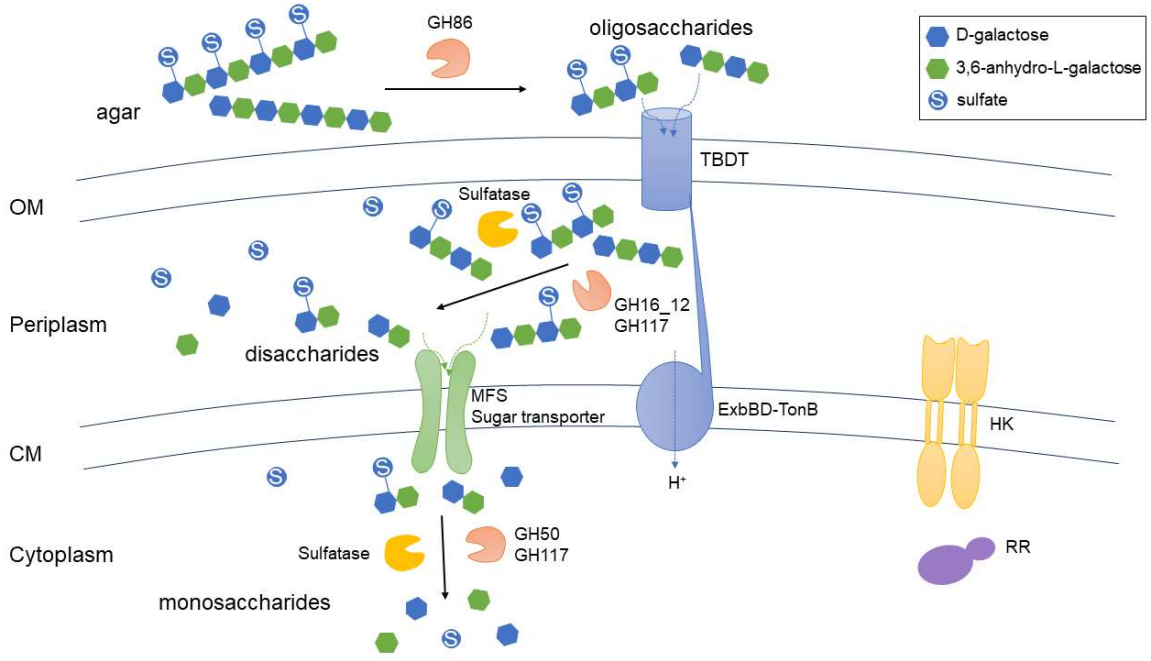


Figure 2.10. Schematic representation of the proposed "selfish" mechanism of agar degradation in strain NLcol2. Agar structure was illustrated schematically with a mixture of

agarose and sulfated polymers. The enzymes and their potential subcellular localizations were included. Abbreviations: OM, outer membrane; CM, cytoplasmic membrane; GH, glycoside hydrolase; MFS, major facilitator superfamily; TBDT, TonB-dependent transporter; ExbBD-TonB, biopolymer transport protein ExbB and ExbD; HK, histidine kinase; RR, response regulator.

2.3.6 Cell migration, biofilm formation and enzyme secretion when grown on seaweed

We found 29 genes involved in flagella formation significantly expressed with a greater fold change in brown and red seaweed samples (average logFC = 2.83 and 2.59, FDR < 0.05 respectively) than in agar (average logFC = 1.14, FDR < 0.05) compared to D-galactose (**Table S2.1a**). Both agar (0.2% w/v) and D-galactose were water soluble, but seaweeds were solids that were cut into small pieces in the culture media. This indicates that bacterial cells of strain NLcol2 could turn on more flagella motors and transition to a motile state to facilitate movement toward food resources like seaweed particles (**Figure 2.11**). Flagellar motility is highly regulated by environmental factors due to the expenditure of genes and energy required (Soutourina & Bertin, 2003). Therefore, when grown on D-galactose and agar, strain NLcol2 could save the energy of flagellar formation for other purposes, such as sugar metabolism and nitrogen fixation. However, when grown on seaweeds and substrates are not readily available, it is necessary for them to produce flagella and search for substrate.

We further investigated biofilm formation of strain NLcol2 when grown on seaweed samples. Both glycoside transferases (GTs) involved in polysaccharide biosynthesis and enzymes of secretion systems were more abundant among the top 100 expressed genes in seaweed-grown cultures compared with growth on agar and galactose (**Figure 2.5**). Additionally, one putative PUL was identified by dbCAN3 (**Table S2.1b**) containing 19 consecutive genes, including 10 glycoside transferases (GT4, GT2, and GT8). It was

homologous to the PUL for capsule polysaccharide synthesis in *B. thetaiotaomicron* (PUL0257 in the dbCAN-PUL database), which has been verified experimentally to show greater expression in biofilm than planktonic *B. thetaiotaomicron,* and may provide advantages to bacterial retention in the gut and access to solid state food particles (TerAvest et al., 2014). There were five GTs significantly up-regulated in red seaweed, while all 10 GTs were significantly down-regulated in agar referenced to D-galactose condition. We did not find significant evidence for a higher expression of the PUL in brown seaweed samples, however, there are 13 significantly up-regulated GTs (mainly GT2 and GT4) in the whole genome suggesting other loci for polysaccharide synthesis (**Table S2.1c**). The observations above indicate that strain NLcol2 might produce extracellular polysaccharide for biofilm formation when grown on seaweeds (**Figure 2.11**), which may help with adherence of cells to one other, maintenance of mature biofilm, as well as attachment and colonization to surfaces like seaweed particles while simultaneously degrading the macroalgal polysaccharides (Limoli et al., 2015; H. Wang et al., 2015). The biofilm matrix also forms a diffusion barrier to prevent the loss of water, permeation of oxygen and antibiotics to benefit the inner layers (Flemming & Wingender, 2010; Nourbakhsh et al., 2022; Olivares et al., 2020). Therefore, biofilm formation might also attenuate the diffusive loss of oligosaccharide substrates and further add to selfishness of strain NLcol2.

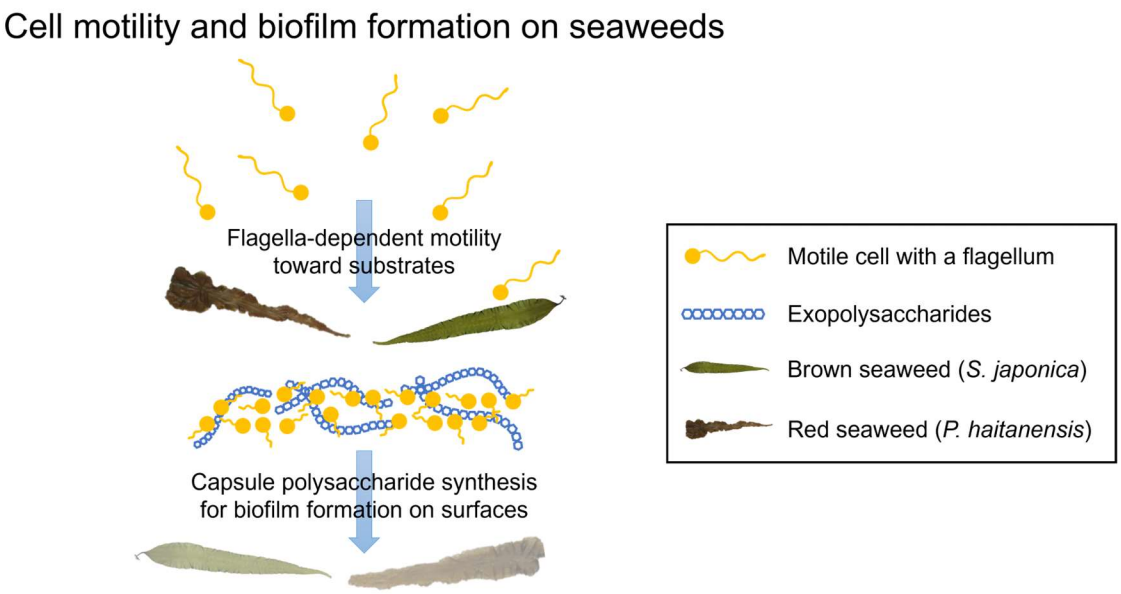


Figure 2.11. Schematic overview of cell motility and biofilm formation of strain NLcol2 when grown on brown seaweed (*S. japonica*) and red seaweed (*P. haitanensis*).

To investigate the relationship between secreting extracellular enzymes and energy conservation, the number of significantly up-regulated GHs and sulfatases with a signal peptide was counted for agar, brown seaweed, and red seaweed samples. We found that a smaller number of them were induced with significance on brown or red seaweeds compared with agar (**Figure 2.12**). This indicates that strain NLcol2 secretes fewer extracellular enzymes for polysaccharide degradation when they are given substrates as particulates like seaweed rather than dissolved compounds like agar. It has been shown that cells of marine *Vibrionaceae* strains secreting lower levels of extracellular alginate lyases aggregated more strongly and increased intercellular synergy to benefit from each other's degradation activities (D'Souza et al., 2023). Our finding agrees with this observation that cells may collaborate more by sharing extracellular enzymes and individual cells could spend less energy on secreting enzymes when given seaweeds as growing surfaces, so that they can grow in higher density and may eventually form biofilms.

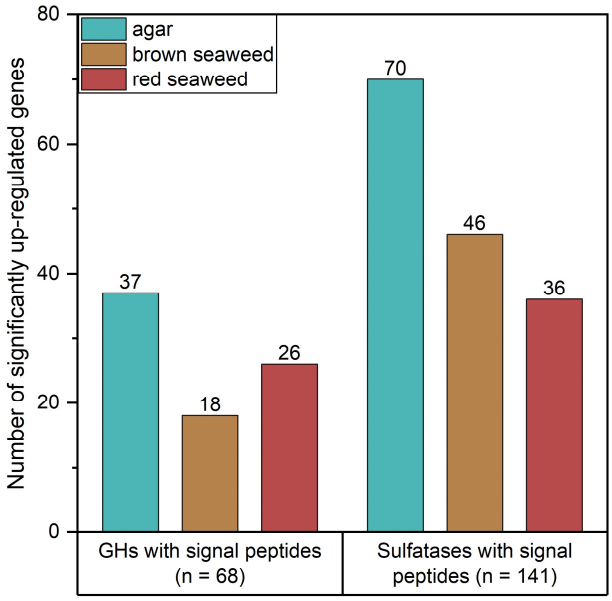


Figure 2.12. Number of significantly up-regulated GHs and sulfatases with a signal peptide sequence in agar, brown seaweed and red seaweed samples compared with D-galactose (logFC > 0 and FDR < 0.05).

2.4 Conclusion

This study analyzed transcriptomic profiles of *Pontiella agarivorans* strain NLcol2 grown on agar, brown seaweed (*Saccharina japonica*), and red seaweed (*Porphyra haitanensis*) in comparison to D-galactose as a control. Differential gene expression analysis identifies highly expressed carbohydrate-active enzymes (CAZymes) and sulfatases that play a crucial role in polysaccharide degradation. Potential polysaccharide utilization loci (PULs) responsible for agar degradation were identified, and a selfish mechanism for agar depolymerization was proposed in strain NLcol2. Finally, we found genes related to flagella formation and capsule polysaccharide synthesis were highly expressed during growth on seaweed, which together promote biofilm formation for attachment and access to seaweed particles. Overall, this study provides insights into the enzymatic systems and microbial strategies employed by *Pontiella agarivorans* strain NLcol2 for the degradation of

macroalgal polysaccharides, with a specific focus on agar. It also informs the eventual fate of macroalgal polysaccharides by anaerobic microbial degradation in marine sediments, which is an important aspect to consider in the turnover of macroalgae-derived carbon, as well as the examination of long-term storage of carbon in macroalgae as a CDR strategy.

Chapter 3. Microbial communities of kelp detritus in sediments of the Santa Barbara Basin

Abstract

Marine macroalgae, exemplified by the giant kelp *Macrocystis pyrifera*, play a crucial ecological role in the coastal ocean by forming extensive kelp forests. Detritus from macroalgae represents a significant but poorly defined source of organic carbon exported to the deep ocean. Furthermore, the decomposition of kelp detritus and associated microbes have been poorly studied. The Santa Barbara Basin is a unique environment for studying microbial interactions with kelp detritus due to seasonal deoxygenation events and sediments rich in organic matter. In this study, we investigated the microbial community structure in sediments with kelp detritus across various deep SBB stations. Analyses of alpha and beta diversity demonstrate a significant shift in microbial communities in kelp-laden sediments compared to background sediments. We found microbial communities enriched in kelp-laden sediments, primarily belonging to phyla *Bacteroidota*, *Spirochaetota*, *Desulfobacterota*, *Campylobacterota*, and *Verrucomicrobiota*. Specifically, *Bacteroidota* and *Spirochaetota* represent taxa capable of hydrolyzing complex polysaccharides in kelp detritus under low oxygen conditions. *Desulfobacterota* and *Campylobacterota* represent sulfur bacteria with a potential adaptation to the kelp detrital system at the sediment-water interface. Furthermore, our research delves into the environmental distribution of the Verrucomicrobial isolate *P. agarivorans* strain NLcol2, identified as *R76-B128* within the order *Kiritimatiellales*. Relatives of *P. agarivorans* were found to be more abundant in kelp-laden sediments and in shallower depths, suggesting their preference for carbon-rich

environments in which macroalgal biopolymers are accessible. The findings provide insights into the intricate relationships between microbial communities and kelp detritus, expanding our understanding of their ecological roles in biogeochemical cycling of carbon, sulfur, and nitrogen in benthic marine environments.

3.1 Introduction

Marine macroalgae are large photosynthetic algae widely distributed in the oceans. One classic example is the giant kelp *Macrocystis pyrifera*, which plays an important ecological role by forming extensive kelp forests on rocky reefs in temperate coastlines around the world (Steneck et al., 2002). The kelp canopy provides shelter to numerous invertebrates, fishes, and other animals, and its high productivity supports diverse communities in the food web (Kushner et al., 2013; Morton et al., 2021; Pedersen et al., 2020). Most of the kelp biomass form detritus and contribute to a significant amount of organic carbon exported to the deep ocean (Krause-Jensen & Duarte, 2016; Pedersen et al., 2020).

The decomposition of kelp can be performed by grazers like sea urchins on shallow reefs and detritivores like amphipods on sandy beaches (Filbee-Dexter et al., 2020; Lowman et al., 2019). They can also be broken down by heterotrophic microbes rapidly with the presence of oxygen (Delille & Perret, 1991). There is an increasing number of studies on the anaerobic digestion of marine macroalgae because it is considered as a renewable energy source of methane and alcohols (Chynoweth et al., 1981; McKennedy & Sherlock, 2015; Wei et al., 2013). In general, anaerobic digestion starts with the hydrolysis of complex organic molecules into simple sugars, amino acids, and fatty acids. These small organic compounds can be further fermented into acetic acid, which is utilized by methanogens to produce methane (McKennedy & Sherlock, 2015). Though macroalgae were studied due to

their ecological and economic importance, the decomposition of kelp detritus by anaerobic microbial communities remains largely unknown, especially in the benthic marine environment. A study on the microbial community involved in anaerobic decomposition of macroalgae found that phyla *Bacteroidetes* (now *Bacteroidota*), *Proteobacteria*, *Spirochaetes* (now *Spirochaetota*), and *Firmicutes* (now *Bacillota*) were enriched with inoculum sources of spring sediments, sludge from wastewater treatment plant and lake water (Morrison et al., 2017).

Nitrogen supply is another important factor in bacterial–macroalgal interaction. *Macrocystis* has limited ability to store nitrogen leading to a high C:N ratio (17:1) in its biomass (BRZEZINSKI et al., 2013; LTER & Reed, 2019). Diazotrophs including diverse *Cyanobacteria*, *Azotobacter* and other heterotrophs are important nitrogen sources on the living kelp thalli (Goecke et al., 2010; Hamersley et al., 2015). Similarly, nitrogen fixation is important for the microbiome during decomposition process as well. Increased nitrogen fixation rates were reported on decomposing *Macrocystis pyrifera* and *Sargassum horneri* primarily by heterotrophic bacteria under anaerobic conditions (Hamersley et al., 2015; Raut et al., 2018; Raut & Capone, 2021).

Marine sediments host abundant and highly diverse microbial communities, especially those involved in organic carbon remineralization. Community composition and diversity are largely controlled by oxygen level and organic carbon concentration (Hoshino et al., 2020). The Santa Barbara Basin (SBB) is one of the borderland basins in Southern California bounded by the California coastline to the north and the Channel Islands to the south, with a maximum water depth of ~ 590 m at the depocenter (**Figure 3.1**). The western sill is about 475m deep and meets the oxygen minimum zone (OMZ) of the Pacific Ocean

(Qin et al., 2022). Due to the high productivity in surface waters, high sedimentation rates, and restricted circulation of deep water below the sill, bottom waters are consistently low in dissolved oxygen (Hülsemann & Emery, 1961; E. Sholkovitz, 1973; SOUTAR & CRILL, 1977). This low oxygen condition inhibits the colonization of benthic macro-fauna and favors mat-forming bacteria utilizing nitrate as the alternative electron acceptor (Bernhard et al., 2000; Valentine et al., 2016). However, bottom water stratification can be interrupted by strong coastal upwelling of cold, dense, slightly-oxygenated, and nutrient rich Pacific deep water flushing into the basin usually in spring (Goericke et al., 2015; E. R. Sholkovitz & Gieskes, 1971). The concentrations of dissolved oxygen and nitrate in bottom waters showed a strong seasonal pattern with a peak in spring and gradual decrease in the rest of the year, varying between 0 and 16 μM for oxygen and 0-37 μM for nitrate (Bograd et al., 2002; Qin et al., 2022).

In this study, we identify and investigate the microbial community responsible for anaerobic decomposition of kelp and its ecological role in the kelp detrital system in the sediments of SBB.

3.2 Methods

3.2.1 Study sites and sediment sampling

Marine sediments were sampled from 11 stations along a north-south transect of the Santa Barbara Basin by remotely operated vehicle (ROV) *Jason*, from onboard R/V *Atlantis* during the AT42-19 cruise in November 2019 (**Figure 3.1**). Geographic locations of stations and a bathymetric map of SBB were visualized in Ocean Data View v.5.6.7 (Schlitzer, 2018) using the ETOPO1 map data in 2x2 min resolution. Latitude and longitude of each station were generated by the *Jason* data processor using navigation data from both the

Doppler velocity log system and the ultrashort baseline positioning system. Water depth was determined by adding depth and altitude records of the ROV. Seawater temperature and salinity were measured by a Seabird 19plus V2 SeaCAT Profiler CTD (Sea-Bird Scientific, USA) and oxygen concentration was measured by an Aanderaa 4831 oxygen optode (Aanderaa Instruments, Norway) equipped on *Jason*.

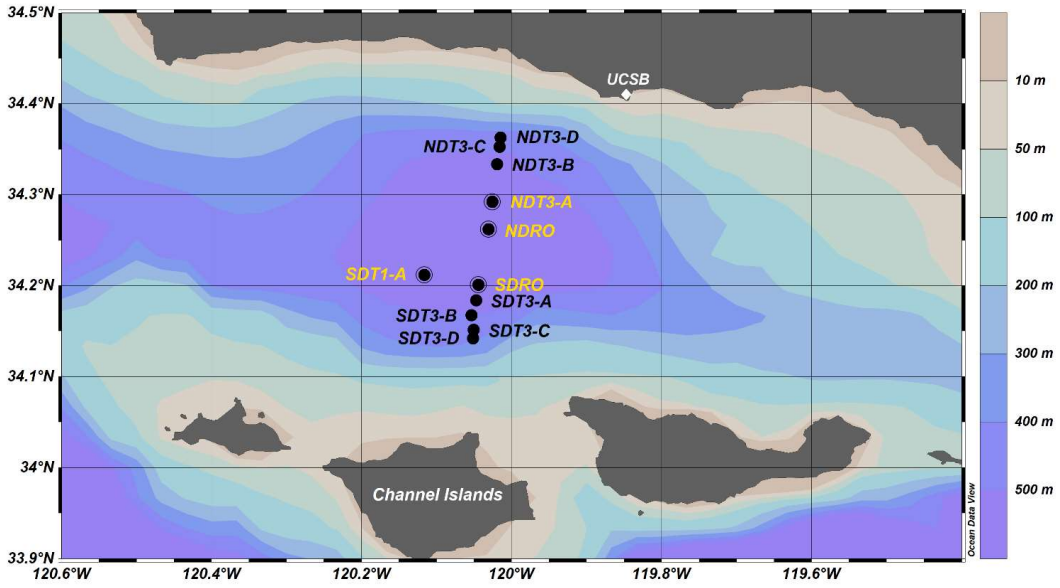


Figure 3.1. Bathymetric map of the Santa Barbara Basin with sampling stations visualized in Ocean Data View. The land topography was grey and colored contours were made for the bathymetry in SBB. Stations with kelp detritus sampled were circled and labeled in yellow.

Marine sediments were sampled by push cores (polycarbonate, 6.35 cm inner diameter, 30.5 cm length) operated by the manipulators of ROV *Jason*. Kelp-laden sediments were visually located and collected by push cores at stations NDT3-A, NDRO, and SDRO (**Figure 3.2c**). At station SDT1-A, kelp detritus and sediments were scooped into a fabric bag and stored in *Jason's* biobox (**Figure 3.2d**). The top 2cm of surface marine sediments adjacent to kelp falls were taken as background sediments for comparison with kelp-laden sediments. Sediment cores for DNA extraction were preserved in 2 cm depth intervals after

slicing under a constant argon flow in a cold room. All sediment samples were taken in triplicate and immediately preserved in a -80℃ freezer until DNA extraction.

3.2.2 DNA extraction, 16S rRNA gene amplification, and sequencing

DNA was extracted from ~ 0.5 ml of sediments using FastDNA Spin Kit for Soil (MP Biomedicals, OH, USA). Manufacturer's instructions were followed with the modification of 200 µl phenol:chloroform:isoamyl alcohol (25:24:1 v/v, for molecular biology, Sigma-Aldrich) added to increase DNA yield. DNA concentration was quantified with a Qubit fluorometer (Thermo Fisher Scientific, USA). The hypervariable V4 region of 16S rRNA gene was amplified using barcoded primers 515F (5'-GTGYCAGCMGCCGCGGTAA) - 806R (5'-GGACTACNVGGGTWTCTAAT) as described previously (Apprill et al., 2015; Kozich et al., 2013; Parada et al., 2016). The polymerase chain reaction (PCR) was performed with 1 µl DNA template (1-5 ng/µl), 1 µl forward primer (10 µM), 1 µl reverse primer (10 µM), and 17 µl AccuPrime Pfx SuperMix (Thermo Fisher Scientific, USA). The thermocycling program was set as follows: an initial denaturation at 95 ºC for 2 min, 35 cycles of denaturation at 95 ºC for 20s, annealing at 55 ºC for 15s, extension at 72 ºC for 5 min, and a final extension at 72 ºC for 10 min. PCR products were visualized by agarose gel electrophoresis to confirm successful amplification. DNA samples were normalized using the SequelPrep Normalization Kit (Thermo Fisher Scientific, USA) and further cleaned and concentrated using the Zymo DNA Clean & Concentrator kit (Zymo Research, USA). The purified PCR products were pooled and sequenced on Illumina MiSeq platform at University of California Davis Genome Center, generating paired end reads of 250 bp.

3.2.3 Microbial community analysis

Samples were demultiplexed and non-biological nucleotides like primers and adapters were removed before analysis. R package DADA2 v.1.16.0 (Callahan et al., 2016) was used for read processing following its pipeline tutorial (http://benjjneb.github.io/dada2/tutorial.html). In short, raw reads were quality-checked and filtered using standard parameters to remove Ns and low-quality sequences. Filtered and trimmed reads were error-corrected and merged to identify 16S rRNA amplicon sequence variants (ASVs). Taxonomy was assigned to the ASV sequences by referencing the SILVA Ref NR SSU r138.1 database (Quast et al., 2012).

To analyze microbial community structure in kelp and background sediments, R package phyloseq v.1.44.0 was used for data analysis and visualization of community diversity and ASV abundances (McMurdie & Holmes, 2013). Alpha diversity was measured and visualized for all samples with triplicates by the observed number of species (Observed) to estimate richness of true species diversity, as well as Shannon index to combine both richness and evenness. The analysis of variance (ANOVA) test of alpha diversity between kelp and background sediments was performed using the aov() function on all samples with triplicates. Mean comparison p-values were added to the boxplot using stat_compare_means() function in ggpubr v.0.6.0. Beta diversity was estimated by calculating Bray-Curtis dissimilarities between samples and was visualized in non-metric multidimensional scaling (NMDS) ordination plots. Read counts were transformed to proportions of total number of reads for each sample prior to the calculation of Bray-Curtis dissimilarities. To statistically test the significance of the effect of kelp detritus versus background sediments on community structure, we performed a permutational multivariate analysis of variance (PERMANOVA) test using the adonis() function in vegan v.2.6.4. To

analyze the composition of microbial communities, triplicates were merged and counts were transformed to percentages to represent relative abundances of ASVs in each sample.

EdgeR v.3.42.4 was used to study the differential abundance of ASVs in kelp versus background sediments (Robinson et al., 2010). Read counts were normalized to the median of total number of reads for each sample. To focus on taxa with big differences between kelp and background samples, ASVs with low variance (variance of normalized counts < 100) were filtered out. The phyloseq data object was converted to an edgeR object and a binary test was performed to find the ASVs with a significant difference of abundances in kelp-laden sediments compared with background sediments (FDR < 0.05).

All community analyses were performed in R v.4.3.1 and ggplot2 v.3.4.2 was used for visualization of most plots. The heatmap with contours of relative abundance of *R76-B128* clade in sediments across the north-south transect of SBB was made in Ocean Data View v.5.6.7.

3.2.4 Phylogenetic tree reconstruction of ASVs in the order *Kiritimatiellales*

A phylogenetic tree of 16S rRNA sequences was built to visualize the phylogeny of five ASVs belonging to the order Kiritimatiellales with a positive log fold change (logFC > 0) and FDR < 0.05 in the above differential abundance analysis. 106 representative 16S rRNA gene sequences within the order Kiritimatiellales were selected from SILVA RefNR SSU database release r138.1 (Quast et al., 2012). Sequences also include four previously isolated and characterized species *K. glycovorans* (CP010904), *P. desulfatans* (LS453292), *P. sulfatireligans* (LS453290), *T. aerotolerans* (MN909984), and *P. agarivorans* strain NLcol2 (OQ749723). Two Verrucomicrobia sequences (ABEA03000104, AF075271) were taken as the outgroups and the tree was rooted there. The partial 16S rRNA gene sequences (V4

region only) of the five selected ASVs were aligned with near full-length sequences of other 16S sequences using Geneious Prime 2023.1.2 (https://www.geneious.com). The alignment was further trimmed by removing hanging ends outside of V4 region, as well as columns with > 90% of gaps in Jalview v.2.11.3 (Waterhouse et al., 2009). A maximum-likelihood tree was constructed by RAxML v.8.2.9 (Stamatakis, 2014) with GTR substituion model plus GAMMA model or rate heterogeneity. Rapid bootstrap search was stopped after 650 replicates with MRE-based criterion. The best-scoring maximum-likelihood tree with bootstrap values was visualized in the iTOL server (Letunic & Bork, 2019).

3.3 Results and Discussion

3.3.1 Study sites and kelp detritus

The water depth of all stations surveyed was over ~ 450 m, below the western sill of SBB, with a maximum of ~ 590 m at the two depocenter stations: NDRO and SDRO (**Table 3.1**). The bottom water temperature (~ 6.5 ℃) and salinity (~ 34.25 PSU) differed only slightly across the transect and dissolved oxygen concentration was consistently below 10 μM. SBB was within the classical deoxygenation period during our visit in November 2019 and typically bottom water was hypoxic (< 20 μM) if not completely anoxic (Qin et al., 2022). In general, the oxygen levels were higher in shallower C and D stations than deep A, B, and depocenter stations. However, an inverted oxygen profile was observed at NDT3-A (10 μM at 573 m, 5 μM at 450 m, 28 μM at 350 m) which could be explained by transient mixing of an oxygenated water mass (Qin et al., 2022).

Table 3.1. Metadata of stations across the north-south transect of SBB. Sediment type listed includes kelp detritus (KD) and typical (kelp free) marine sediments referred to here as background sediments (BG) and used as reference. *Bottom water oxygen concentrations at stations NDRO, SDRO, and SDT1-A were regarded as 0 because the measured values from the optode were below 0.

Station Name	Sediment Type	Latitude (degrees North)	Longitude (degrees East)	Water Depth (m)	Temperature (°C)	Salinity (PSU)	Oxygen (μM)
NDT3-D	BG	34.36270	-120.01475	448	6.68	34.235	7.97
NDT3-C	BG	34.35261	-120.01617	499	6.58	34.236	4.98
NDT3-B	BG	34.33337	-120.01952	538	6.48	34.232	6.51
NDT3-A	KD, BG	34.29191	-120.02555	574	6.39	34.236	9.90
NDRO	KD, BG	34.26207	-120.03074	583	6.56	34.264	0*
SDRO	KD, BG	34.20102	-120.04414	587	6.56	34.264	0*
SDT1-A	KD, BG	34.21174	-120.11647	574	6.56	34.264	0*
SDT3-A	BG	34.18379	-120.04721	572	6.57	34.261	0.05
SDT3-B	BG	34.16772	-120.05342	537	6.57	34.260	1.65
SDT3-C	BG	34.15159	-120.05048	495	6.63	34.256	2.93
SDT3-D	BG	34.14209	-120.05154	449	6.69	34.239	9.65

Large piles of kelp detritus fallen onto the seafloor were found near stations NDT3-A, NDRO, SDRO, and SDT1-A in the deep SBB (**Figure 3.2**). Whitish microbial mats grew densely over the kelp piles and along the edges of the kelp blades. Fauna including snails and pyrosomes were often observed around kelp falls. The kelp falls were found at water depths greater than 570 m with low dissolved oxygen concentrations (**Table 3.1**). Most deep stations were completely anoxic ($O_2 \approx 0$ μM) except station NDT3-A which showed a slightly higher O_2 concentration of ~ 10 μM at the bottom interpreted as a transient oxygen introduction by water mass mixing. Therefore, kelp falls are present mostly in anoxic conditions with potential transient exposure to hypoxic conditions during periodic reoxygenation events in SBB (Qin et al., 2022; Yousavich et al., 2023).

a)  b) 

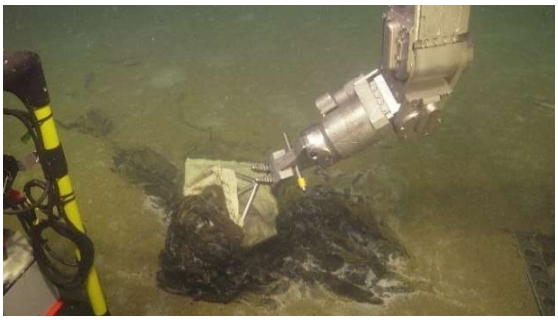
c)  d) 

Figure 3.2. Photos of kelp detritus found on the seafloor in the Santa Barbara Basin at stations NDT3-A (a), NDRO (b), SDRO (c), and SDT1-A (d). Kelp-laden sediments were collected by either a push core (c) or by a scoop operated by the manipulators of ROV *Jason* (d).

3.3.2 Diversity of microbial communities in kelp-laden sediments

The microbial community showed a lower alpha diversity in kelp-laden sediments than in background marine sediments at all four stations except for the measure of observed number of species at station NDRO (**Figure 3.3a**). Moreover, the analysis of variance (ANOVA) tests showed a significant difference in alpha diversity between kelp-laden sediments and background sediments for both Observed and Shannon indices ($p < 0.01$, ANOVA) (**Figure 3.3b**). This indicates that specific taxa could be enriched by kelp detritus on the seafloor regardless of their locations, shift the microbial community composition, and thus lead to a decrease in diversity. If we group the samples by different stations, SDRO surface sediments showed the highest diversity and the other three deep stations showed similar diversity (**Figure 3.3c**). There seems to be significant difference of alpha diversity between stations by Shannon index ($p < 0.05$, ANOVA), but not by observed number of species ($p > 0.05$, ANOVA). The alpha diversity difference between kelp-laden sediments at different stations could be due to the amount of kelp, status of degradation, and the type of kelp.

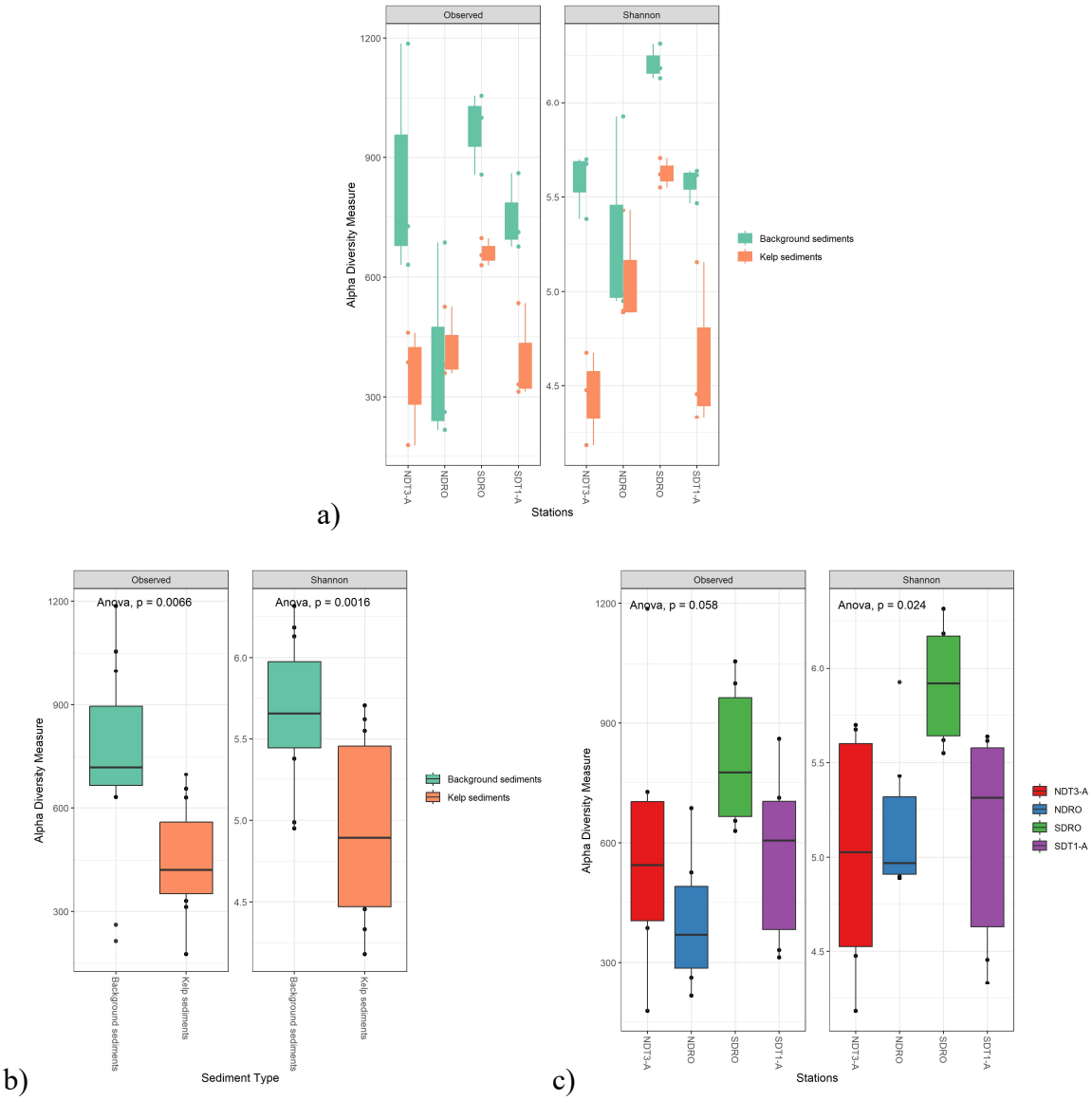


Figure 3.3. Alpha diversity of kelp-laden sediments and background sediments measured by observed number of species (Observed) and Shannon index. a) Boxplot of all samples in triplicates (n = 24) with sediment type labeled by color. b) Boxplot of samples grouped by sediment type with p-values labeled on top, resulted from ANOVA tests between sediments with and without kelp detritus for both indices. c) Boxplot of samples grouped by stations with p-values labeled on top, resulted from ANOVA tests between sediments from different stations for both indices.

As for beta diversity, non-metric multidimensional scaling (NMDS) plots were made with Bray-Curtis dissimilarities to compare differences of microbial community structure between samples (**Figure 3.4a**). Similar to alpha diversity, there are significant differences

of beta diversity between kelp and background marine sediments ($p < 0.001$, PERMANOVA) (**Figure 3.4b**), as well as between four deep stations ($p < 0.01$, PERMANOVA) (**Figure 3.4c**). Kelp-laden sediments at SDRO seem to be more closely related to background sediments than at other stations. And NDRO sediments seem to have a much more different microbial community structure from those at the other three stations.

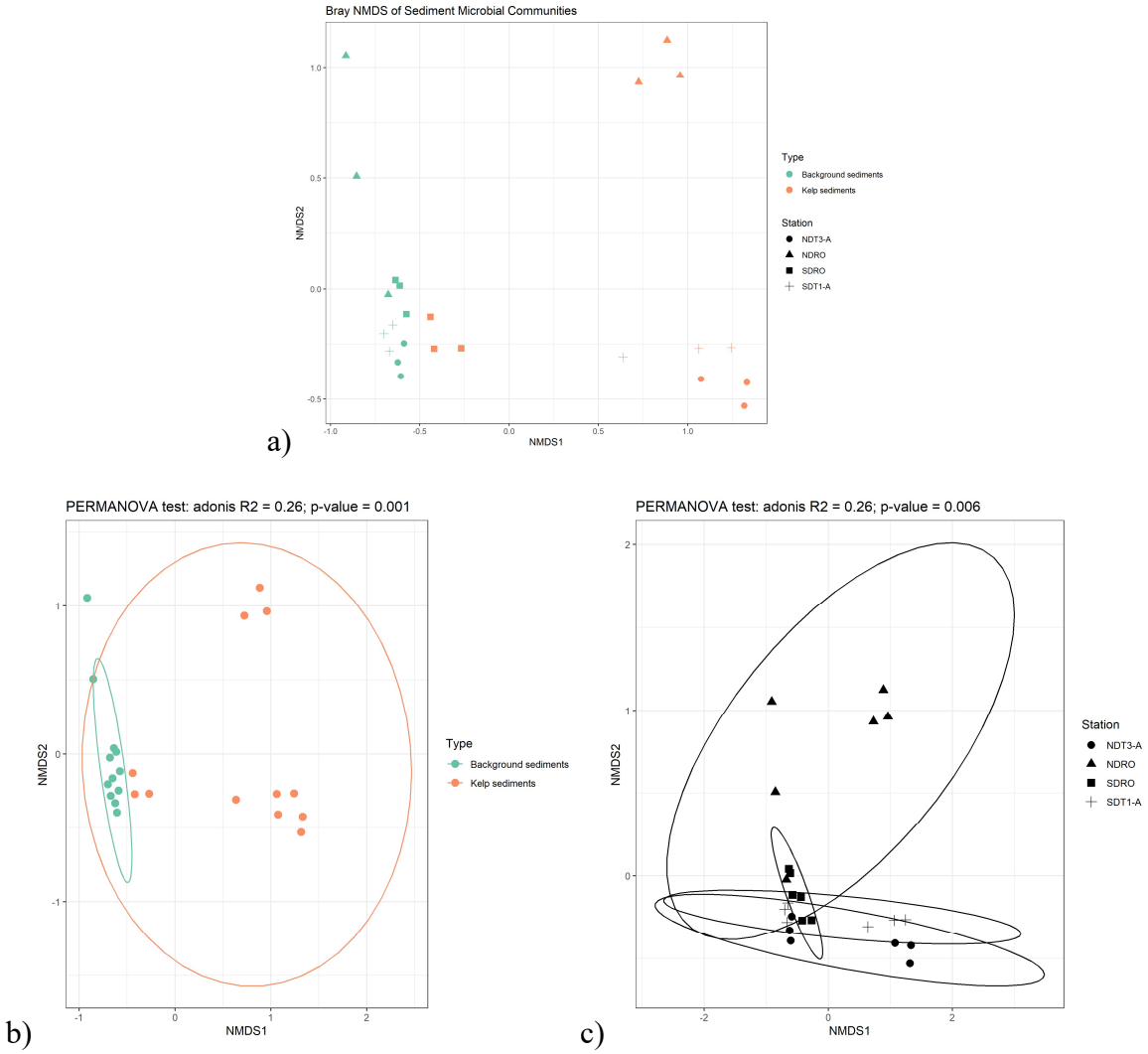


Figure 3.4. Beta diversity of kelp-laden sediments and background sediments estimated by Bray-Curtis dissimilarities and visualized in NMDS plots. a) NMDS plot of all samples in triplicates (n = 24). Ellipses were calculated with a confidence level of 0.95 for samples grouped by b) sediment type and c) station. Sediment type was labeled in colors and station was drawn with different shapes.

3.3.3 Microbial community composition

In general, the microbial community composition in kelp-laden sediments and background sediments were similar at the phylum level with some phyla differing in relative abundance, but it was quite different at the ASV level (**Figure 3.5**). Among the top 10 phyla in sediments, the most abundant phylum was *Bacteroidota* which constituted 20.1% of the total community and were more abundant in kelp-laden sediments (23.2%) than in background sediments (17.0%). *Desulfobacterota*, *Spirochaetota*, and *Campylobacterota* showed similar patterns as *Bacteroidota* that they were relatively enriched in kelp-laden sediments compared with background sediments (**Figure 3.5a, Table 3.2**).

Bacteroidota and *Spirochaetota* are often found in the gut microbiota of animals and include saccharolytic bacteria specialized in degrading high molecular weight compounds. *Bacteroidota* are important and dominant residents of the human microbiome and harbor a large number of carbohydrate-active enzymes (CAZymes) in their genomes for the degradation of various glycopolymers (Lapébie et al., 2019; Thomas et al., 2011). *Spirochaetota* were found to be a major degrader of pectin and xylan with a variety of related CAZymes hosted in their genomes in the gut microbiota of cattle, moose, and termites (Gharechahi et al., 2021; Svartström et al., 2017; Tokuda et al., 2018). Although *Bacteroidota* are well-studied in mammalian gut systems, they have colonized almost all types of habitats on Earth (Thomas et al., 2011). Many marine species have shown specialty in degrading polymers rather than monomers and have important functional roles in the biogeochemical cycling of carbon in the marine environment. They prefer to grow attached to particles, surfaces or algal cells, and have the ability to degrade an array of complex polymers derived from plants and animals (such as pectins, agars, cellulose, and chitin) using glycosidic hydrolases and peptidases (Fernández-Gómez et al., 2013; Thomas et al.,

2011). This indicates that species belonging to these two phyla enriched by the kelp detritus could be the dominant taxa degrading kelp detritus in SBB sediments.

Desulfobacterota includes many sulfate-reducing bacteria (SRB) that use sulfate as electron acceptor and reduce it to sulfide (Garrity et al., 2005). Members of *Campylobacterota* (previously classified as *Epsilonproteobacteria*) were shown to be sulfide-oxidizing bacteria that could reoxidize sulfide using oxygen or nitrate as electron acceptor (Garrity et al., 2005; Waite et al., 2017). These two phyla of sulfur bacteria could play an important role in sulfur cycling at the sediment-water interface of kelp detritus.

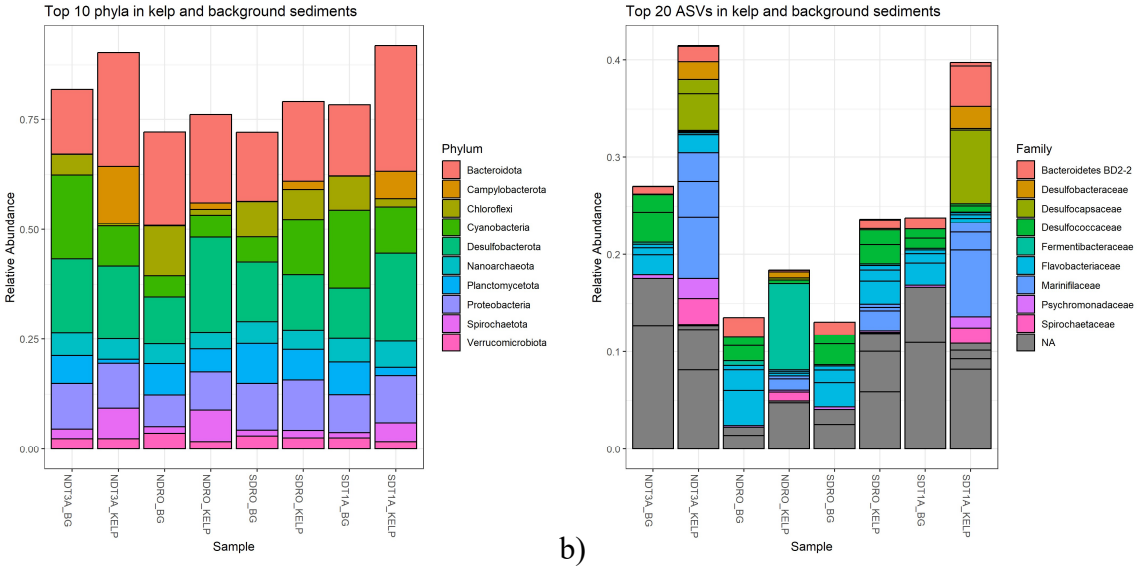


Figure 3.5. Stacked bar plots of the most abundant 10 phyla (a) and 20 ASVs (b) in kelp and background (BG) sediments at stations NDT3-A, NDRO, SDRO, and SDT1-A. The relative abundance was the proportion of certain taxa of the total number of read counts taken from triplicates. The colors were filled by phylum for the top 10 phyla and by family for the top 20 ASVs.

Table 3.2. Relative abundances of the most abundant 10 phyla in kelp and background sediments.

Phylum	Average % in all samples	Average % in kelp-laden sediments	Average % in background sediments

Bacteroidota	20.1%	23.2%	17.0%
Desulfobacterota	15.4%	17.7%	13.1%
Cyanobacteria	10.6%	9.4%	11.9%
Proteobacteria	9.7%	10.2%	9.2%
Planctomycetota	5.6%	3.7%	7.5%
Chloroflexi	5.3%	2.6%	7.9%
Nanoarchaeota	4.9%	4.7%	5.1%
Spirochaetota	3.3%	5.0%	1.5%
Campylobacterota	2.9%	5.6%	0.1%
Verrucomicrobiota	2.5%	2.2%	2.9%

At the ASV level, the community composition showed greater differences in sediments with and without kelp detritus (**Figure 3.5b**). The ASVs with a more abundance in kelp-laden sediments belong to the phyla *Bacteroidota*, *Desulfobacterota*, *Fermentibacterota*, *Proteobacteria*, *Spirochaetota* etc. It is notable that ASVs not present in the background sediments were enriched in kelp-laden sediments. They belong to the families *Marinifilaceae* (4.1%, 1.5%, 1.1%), *Bacteroidetes BD2-2* (1.4%), *Spirochaetaceae* (1.3%), *Desulfobacteraceae* (1.2%), and *Desulfocapsaceae* (1.0%). The enrichment of specific taxa of *Bacteroidota* and *Spirochaetota* in kelp-laden sediments could potentially contribute to the degradation of kelp detritus. There was also a shift of SRB from *Desulfococcaceae* to *Desulfobacteraceae* and *Desulfocapsaceae*, especially at shallower stations NDT3-A and SDT1-A. They belong to the same order *Desulfobacterales* and can all perform dissimilatory sulfate reduction (Garrity et al., 2005).

To quantitively investigate the differential abundance of ASVs between kelp-laden sediments and background sediments, normalized counts were compared and tested between the two types of sediments using the edgeR pipeline (Robinson et al., 2010). We found a

total number of 270 ASVs to be significantly more abundant in kelp-laden sediments compared with background sediments (FDR < 0.05), and 120 of these ASVs were with known taxonomic classification at the genus level belonging to 10 different phyla (**Figure 3.6**). The most enriched taxa include various genera of *Bacteroidota* (*Marinifilum, Labilibacter, Lutibacter* etc.), *Spirochaetota* (*Spirochaeta, Sediminispirochaeta* etc.), *Desulfobacterota* (*Desulforhopalus, Desulfoconvexum, Desulfuromusa, Desulfobacterium, Desulfobacter* etc.), and *Campylobacterota* (*Halarcobacter, Sulfurospirillum, Sulfurimonas* etc.).

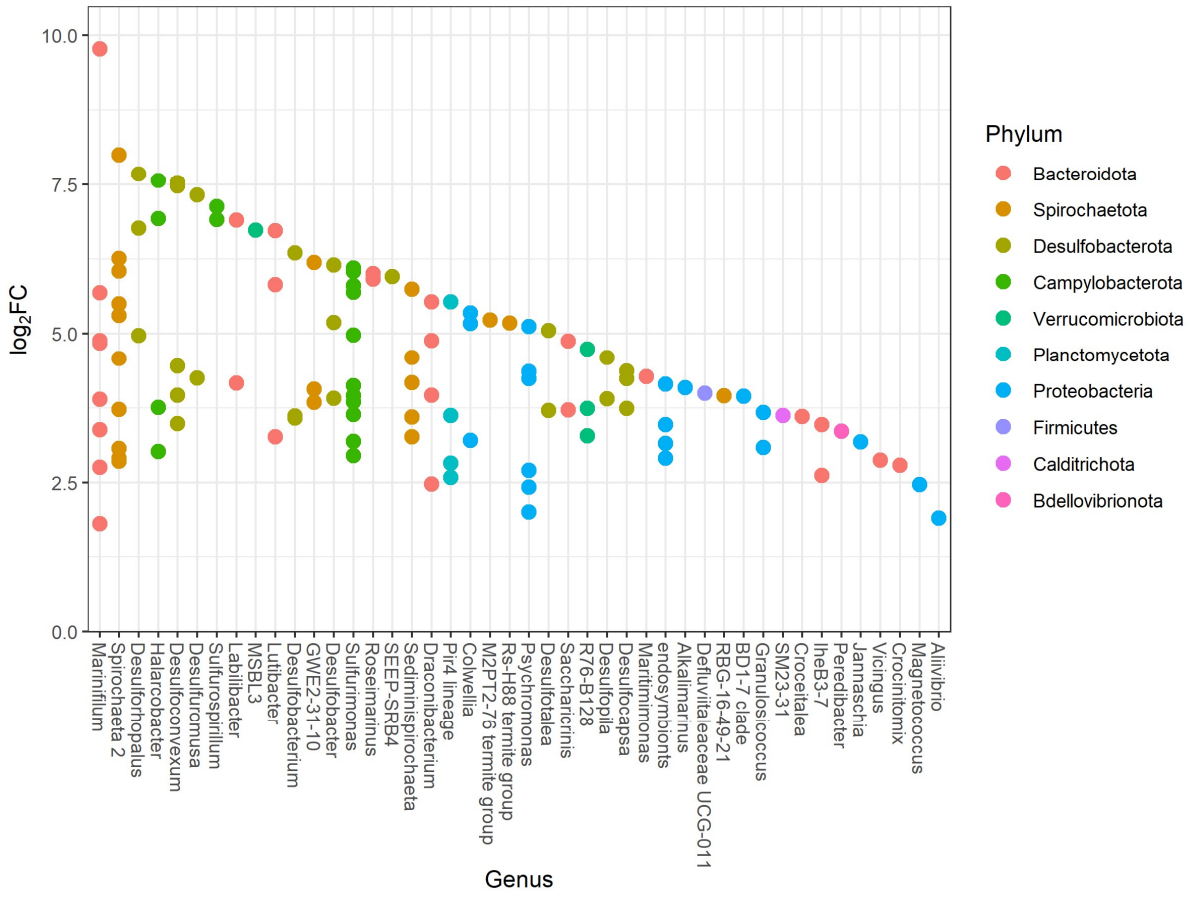


Figure 3.6. Enrichment of ASVs in kelp-laden sediments compared to background sediments organized by genus on the x axis and colored by phylum. Taxa were filtered by a positive log fold change of normalized counts for each taxon with significance (logFC > 0, FDR < 0.05). Both genus and phylum were ordered by enrichment level. ASVs with unknown taxonomic classification at genus level (NAs) were removed for simplicity.

Several species of *Marinifilum* have been isolated from coastal sediments and sea water, and were found to be facultative anaerobic, filamentous, and chemoorganotrophic bacteria fermenting sugar (Na et al., 2009; Ruvira et al., 2013; Xu et al., 2016). Starch, gelatin, and casein can be hydrolyzed by many species of *Marinifilum*, but not cellulose, alginate, or agar (Dai et al., 2022; Fu et al., 2018; Na et al., 2009; Ruvira et al., 2013; Xu et al., 2016). Species of *Labilibacter* were also isolated from marine environments (marine sediments and sea squirt) with the ability to hydrolyze agar (*L. marinus*, *L. aurantiacus*, *L. sediminis*), gelatin, alginate (*L. aurantiacus* and *L. sediminis*), starch, and cellulose (*L. sediminis*) (Liu et al., 2015; Lu et al., 2017; F.-Q. Wang et al., 2020). Over 10 species of *Lutibacter* have been isolated from marine sediments and coastal seawater as heterotorphic bacteria utilizing pyruvate and succinate (Du et al., 2020; Krieg, 2011). *Lutibacter* species can utilize a series of polymers including aesculin, agar, casein, gelatin, and starch (Du et al., 2020; Le Moine Bauer et al., 2016; Park et al., 2013; Sung et al., 2015). The genus *Spirochaeta* represents a group of free-living, saccharolytic, obligate or facultative anaerobic helical shaped bacteria (Krieg, 2011). *S. psychrophila* was isolated from the subseafloor sediment and grew chemooranotrophically with mono-, di- and polysaccharides (Miyazaki et al., 2014). Similarly, *S. bajacaliforniensis*, the type species of *Sedimispirochaeta*, was a chemoorganotrophic anaerobe isolated from a microbial mat in lagoon sediments that grows by fermenting carbohydrates (Fracek & Stolz, 1985). Various genera in the phyla *Bacteroidota* and *Spirocheates* were enriched in kelp-laden sediments and their culture representatives have been isolated from different marine environments and shown to utilize carbohydrates as their carbon and energy sources. These observations are consistent with the

organisms represented by these ASVs utilizing and degrading various polysaccharides in kelp detritus under anaerobic conditions.

The shift of SRB among *Desulfobacterota* and the emergence of *Campylobacterota* raise the question of why these sulfur bacteria would be more enriched in kelp detritus compared to background sediments. It is possible that these sulfur bacteria could benefit more by growing heterotrophically on the hydrolyzed products of polysaccharides by degraders belonging to *Bacteroidota* and *Spirocheates*. Another possible explanation would be the limitation of bioavailable nitrogen in kelp detritus. Previous studies showed that biological nitrogen fixation rates increased during decomposition of detrital kelp, especially for those with high C:N ratios (> 10) (Hamersley et al., 2015; Raut et al., 2018; Raut & Capone, 2021). It has also been shown that members of *Desulfobacterales* and *Desulfuromonadales* are important diazotrophs in deep-sea sediments (Kapili et al., 2020). ASVs of genera *Desulfoconvexum, Desulfuromusa, Desulfobacterium,* and *Desulfobacter* belonging to these two orders could be enriched in kelp-laden sediments due to their nitrogen-fixing ability. Sulfur bacteria using nitrate as electron acceptor may also be favored over those using oxygen, resulting in a shift of the community. For example, *Desulforhopalus singaporensis* and various species in genera *Sulfurospirillum* and *Sulfurimonas* can use nitrate as the alternative electron acceptor and undergo dissimilatory nitrate reduction to ammonium which could be an advantage when living on kelp detritus (Garrity et al., 2005; Han & Perner, 2015; Lie et al., 1999).

3.3.4 Distribution of *R76-B128* in SBB sediments

To investigate the ecological role of *P. agarivorans* in the marine environment, we surveyed the distribution of its close relatives in kelp-laden sediments as well as along the

N-S transect across the whole basin. Stain NLcol2 was classified as *R76-B128*, which is the sister group of *MSBL3* in the order *Kiritimatiellales*. In general, ASVs of *R76-B128* and *MSBL3* were relatively low in abundance in sediments (**Figure 3.7**). The average relative abundance of *R76-B128* was 0.38% in kelp-laden sediments and 0.19% in background sediments, and the order *Kiritimatiellales* accounted for only 0.94% and 0.38% of the microbial community in kelp and background sediments respectively. However, at all four stations, they were more abundant in kelp-laden sediments than in background sediments. As indicated by the differential abundance analysis above, five ASVs classified in the *Kiritimatiellales* order were found to be significantly more abundant in kelp-laden sediments than in background sediments - three as *R76-B128*, one as *MSBL3*, and one with unknown taxonomy at the genus level (**Figure 3.6**). The phylogeny of V4 region of 16S rRNA gene sequences from the five ASVs showed that four of them were grouped into the *R76-B128* clade (corresponding to the GTDB family *Pontiellaceae*) together with three cultivated *Pontiella* species. The other one ASV was classified as *MSBL3* (corresponding to the GTDB family *Tichowtungiaceae*) with *T. aerotolerans* as the closest related isolate (**Figure 3.8**). The observations above indicate that the environmental relatives of *P. agarivorans* (from family *Pontiellaceae* up to order *Kiritimatiellales*) could be favored by kelp detritus in marine sediments. Kelp detritus provides rich carbon sources that support their growth, and in turn, their potential to degrade macroalgal polysaccharides could contribute to the turnover of complex organic carbon molecules in kelp detritus sunk to the seafloor. They may also take advantage of the ability to fix nitrogen to survive better while degrading kelp detritus with a high C:N ratio.

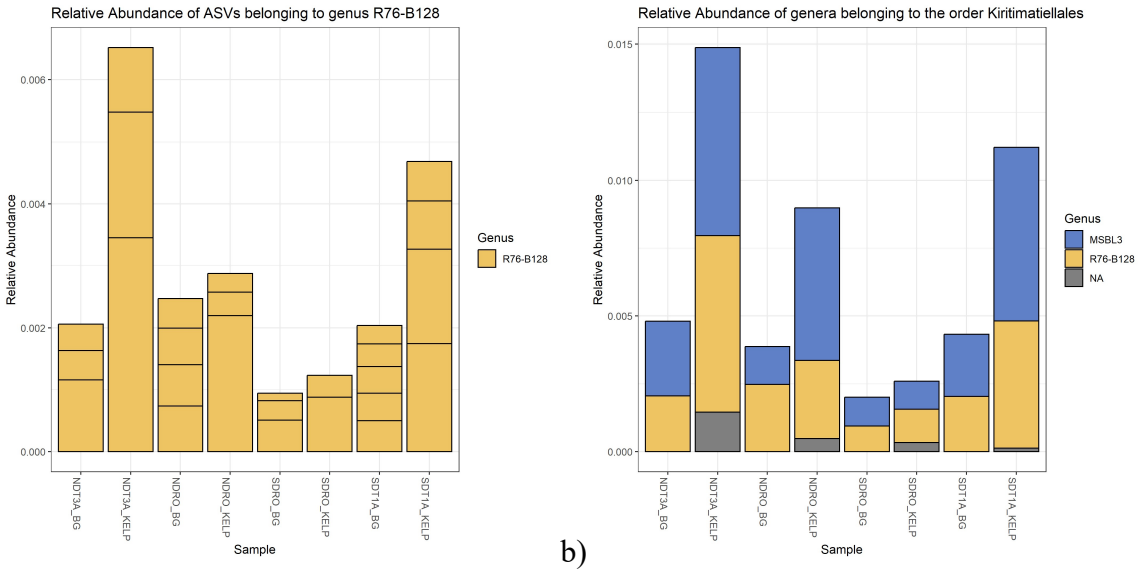


Figure 3.7. Relative abundance of taxa related to *P. agarivorans* in kelp and background sediments, including a) ASVs belonging to *R76-B128*, and b) genera belonging to the order *Kiritimatiellales*.

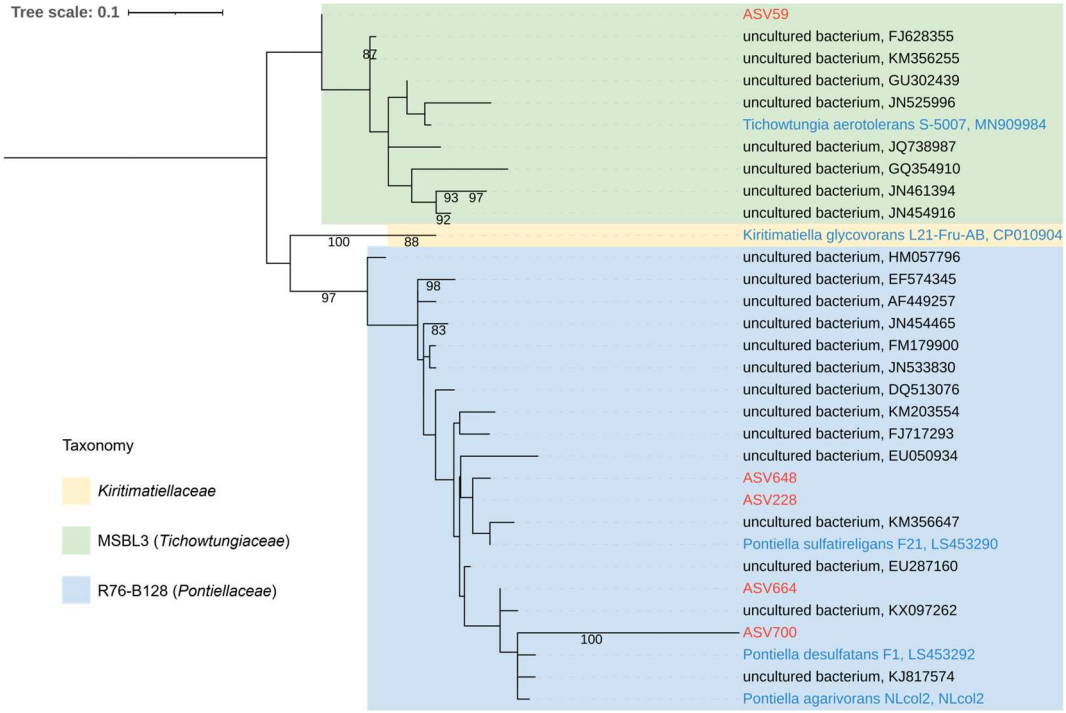


Figure 3.8. Maximum-likelihood phylogenetic tree of 16S rRNA genes (V4 hypervariable region) of the *Kiritimatiellales* order. Five ASVs from the current study are labeled in red, and the cultivated strains are labeled in blue. The tree was rooted with two *Verrucomicrobia* sequences (not shown) and pruned with 27 selected representative sequences. Bootstrap values over 80 are shown on the nodes.

R76-B128 represents the closest relatives of *P. agarivorans*, and they were widely distributed in natural marine habitats in the deep SBB present in the upper 20 cm of the sediments. However, they were low in abundance and constituted only a small portion (0.21% on average) of the microbial community in sediments across the N-S transect of the basin (**Figure 3.9**). They were relatively more abundant within the top 2 cm toward the sediment-water interface as well as in shallower D stations (1.17% at NDT3-D and 1.06% at SDT3-D). The flux of particulate organic carbon (POC) in the ocean decreases with depth below the euphotic zone due to depth-dependent remineralization of sinking particles by heterotrophic bacteria in the water column (Omand et al., 2020; Suess, 1980). Carbohydrates are important components of the POC pool in the ocean, constituting 3-18% of sinking POC (Kharbush et al., 2020). *R76-B128* contains bacteria specialized in degrading glycan polymers, therefore, would be favored at shallower depths with more POC available.

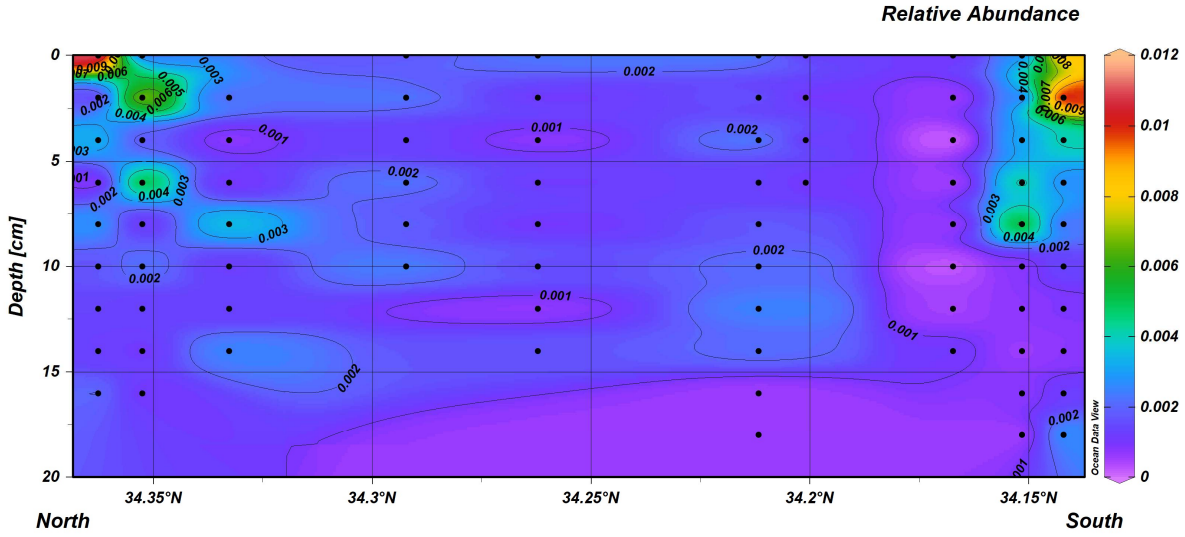


Figure 3.9. Heatmap of relative abundance of *R76-B128* in the top 20 cm of sediments across the north-south transect of the SBB. Stations from left (north) to right (south) are NDT3-D, NDT3-C, NDT3-B, NDT3-A, NDRO, SDT1-A, SDRO, SDT3-B, SDT3-C, SDT3-D. Data were taken from the average of relative abundance in triplicates and labeled as black dots. Contour lines were drawn by an increment of 0.1%.

3.4 Conclusion

The microbial community structure in kelp-laden sediments was different from background marine sediments in the deep Santa Barbara Basin with a decrease in diversity and an enrichment of specific taxa: *Bacteroidota*, *Spirochaetota*, *Desulfobacterota*, *Campylobacterota*, and *Verrucomicrobiota*. Relatives of *P. agarivorans* also showed more abundance in kelp-laden sediments, and at the surface of shallower stations in the basin, consistent with the ecology of *P. agarivorans*. These environmental representatives of *Kiritimatiellales* demonstrated their wide distribution in low abundance in natural marine sediments. Taxa enriched in kelp-laden sediments have a potential to be involved in the degradation of kelp detritus on the seafloor while performing nitrogen fixation at the same time. The overlooked macroalgal detrital system could be an important ecological niche for diverse microbial community contributing to carbon, sulfur, and nitrogen cycling in marine sediments.

References

Abramson, J., Smirnova, I., Kasho, V., Verner, G., Kaback, H. R., & Iwata, S. (2003). Structure and Mechanism of the Lactose Permease of Escherichia coli. *Science*, *301*(5633), 610–615. https://doi.org/10.1126/science.1088196

Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., Von Heijne, G., & Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology*, *37*(4), Article 4. https://doi.org/10.1038/s41587-019-0036-z

Anderson, N. S., Dolan, T. C. S., & Rees, D. A. (1965). Evidence for a common structural pattern in the polysaccharide sulphates of the rhodophyceae. *Nature*, *205*(4976), Article 4976. https://doi.org/10.1038/2051060a0

Appel, M. J., & Bertozzi, C. R. (2015). Formylglycine, a post-translationally generated residue with unique catalytic capabilities and biotechnology applications. *ACS Chemical Biology*, *10*(1), Article 1. https://doi.org/10.1021/cb500897w

Apprill, A., McNally, S., Parsons, R., & Weber, L. (2015). Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquatic Microbial Ecology*, *75*(2), Article 2. https://doi.org/10.3354/ame01753

Araki, C. (1956). Structure of the agarose constituent of agar-agar. *Bulletin of the Chemical Society of Japan*, *29*(4), Article 4. https://doi.org/10.1246/bcsj.29.543

Arcondéguy, T., Jack, R., & Merrick, M. (2001). P_{II} signal transduction proteins, pivotal players in microbial nitrogen control. *Microbiology and Molecular Biology Reviews*, *65*(1), Article 1. https://doi.org/10.1128/mmbr.65.1.80-105.2001

Arnosti, C., Wietz, M., Brinkhoff, T., Hehemann, J.-H., Probandt, D., Zeugner, L., & Amann, R. (2021). The Biogeochemistry of Marine Polysaccharides: Sources, Inventories, and Bacterial Drivers of the Carbohydrate Cycle. *Annual Review of Marine Science*, *13*(1), Article 1. https://doi.org/10.1146/annurev-marine-032020-012810

Barbeyron, T., Brillet-Guéguen, L., Carré, W., Carrière, C., Caron, C., Czjzek, M., Hoebeke, M., & Michel, G. (2016). Matching the diversity of sulfated biomolecules: Creation of a classification database for sulfatases reflecting their substrate specificity. *PLOS ONE*, *11*(10), Article 10. https://doi.org/10.1371/journal.pone.0164846

Bernhard, J. M., Buck, K. R., Farmer, M. A., & Bowser, S. S. (2000). The Santa Barbara Basin is a symbiosis oasis. *Nature*, *403*(6765), Article 6765. https://doi.org/10.1038/47476

Berteau, O., Guillot, A., Benjdia, A., & Rabot, S. (2006). A New Type of Bacterial Sulfatase Reveals a Novel Maturation Pathway in Prokaryotes *. *Journal of Biological Chemistry*, *281*(32), Article 32. https://doi.org/10.1074/jbc.M602504200

Bograd, S. J., Schwing, F. B., Castro, C. G., & Timothy, D. A. (2002). Bottom water renewal in the Santa Barbara Basin. *Journal of Geophysical Research: Oceans*, *107*(C12), Article C12. https://doi.org/10.1029/2001JC001291

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), Article 15. https://doi.org/10.1093/bioinformatics/btu170

Brodie, J., & Lewis, J. (Eds.). (2007). *Unravelling the algae: The past, present, and future of algal systematics*. CRC ; Taylor & Francis [distributor].

BRZEZINSKI, M. A., REED, D. C., HARRER, S., RASSWEILER, A., MELACK, J. M., GOODRIDGE, B. M., & DUGAN, J. E. (2013). MULTIPLE SOURCES AND FORMS OF NITROGEN SUSTAIN YEAR-ROUND KELP GROWTH: On the Inner Continental Shelf of the Santa Barbara Channel. *Oceanography*, *26*(3), 114–123. JSTOR.

Buchko, G. W., Robinson, H., & Addlagatta, A. (2009). Structural characterization of the protein cce_0567 from *Cyanothece 51142*, a metalloprotein associated with nitrogen fixation in the DUF683 family. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, *1794*(4), Article 4. https://doi.org/10.1016/j.bbapap.2009.01.002

Cabello-Yeves, P. J., Ghai, R., Mehrshad, M., Picazo, A., Camacho, A., & Rodriguez-Valera, F. (2017). Reconstruction of Diverse Verrucomicrobial Genomes from Metagenome Datasets of Freshwater Reservoirs. *Frontiers in Microbiology*, *8*, 2131. https://doi.org/10.3389/fmicb.2017.02131

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 581–583. https://doi.org/10.1038/nmeth.3869

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*(1), Article 1. https://doi.org/10.1186/1471-2105-10-421

Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, *25*(15), Article 15. https://doi.org/10.1093/bioinformatics/btp348

Cardman, Z., Arnosti, C., Durbin, A., Ziervogel, K., Cox, C., Steen, A. D., & Teske, A. (2014). *Verrucomicrobia* Are Candidates for Polysaccharide-Degrading Bacterioplankton in an Arctic Fjord of Svalbard. *Applied and Environmental Microbiology*, *80*(12), Article 12. https://doi.org/10.1128/AEM.00899-14

Caspi, R., Billington, R., Keseler, I. M., Kothari, A., Krummenacker, M., Midford, P. E., Ong, W. K., Paley, S., Subhraveti, P., & Karp, P. D. (2020). The MetaCyc database of metabolic pathways and enzymes—A 2019 update. *Nucleic Acids Research*, *48*(D1), Article D1. https://doi.org/10.1093/nar/gkz862

Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P., & Parks, D. H. (2020). GTDB-Tk: A toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics*, *36*(6), Article 6. https://doi.org/10.1093/bioinformatics/btz848

Chen, Y., Lun, A. T. L., & Smyth, G. K. (2016). From reads to genes to pathways: Differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. *F1000Research*, *5*, 1438. https://doi.org/10.12688/f1000research.8987.2

Chynoweth, D. P., Ghosh, S., & Klass, D. L. (1981). Anaerobic Digestion of Kelp. In *Biomass Conversion Processes for Energy and Fuels* (pp. 315–338). Springer US. https://doi.org/10.1007/978-1-4757-0301-6_17

Cong, Q., Chen, H., Liao, W., Xiao, F., Wang, P., Qin, Y., Dong, Q., & Ding, K. (2016). Structural characterization and effect on anti-angiogenic activity of a fucoidan from *Sargassum fusiforme*. *Carbohydrate Polymers*, *136*, 899–907. https://doi.org/10.1016/j.carbpol.2015.09.087

Dai, W., Sun, W., Fu, T., Jia, C., Cui, H., Han, Y., Shi, X., & Zhang, X.-H. (2022). *Marinifilum caeruleilacunae* sp. Nov., isolated from Yongle Blue Hole in the South China Sea. *International Journal of Systematic and Evolutionary Microbiology*, *72*(5). https://doi.org/10.1099/ijsem.0.005358

Delille, D., & Perret, E. (1991). The influence of giant kelp Macrocystis pyrifera on the growth of subantarctic marine bacteria. *Journal of Experimental Marine Biology and Ecology*, *153*(2), Article 2. https://doi.org/10.1016/0022-0981(91)90227-N

Delmont, T. O., Quince, C., Shaiber, A., Esen, Ö. C., Lee, S. T., Rappé, M. S., McLellan, S. L., Lücker, S., & Eren, A. M. (2018). Nitrogen-fixing populations of Planctomycetes and Proteobacteria are abundant in surface ocean metagenomes. *Nature Microbiology*, *3*(7), Article 7. https://doi.org/10.1038/s41564-018-0176-9

Ding, H., & Valentine, D. L. (2008). Methanotrophic bacteria occupy benthic microbial mats in shallow marine hydrocarbon seeps, Coal Oil Point, California. *Journal of Geophysical Research: Biogeosciences*, *113*(G1), Article G1. https://doi.org/10.1029/2007JG000537

Dolliver, J., & O'Connor, N. (2022). Whole System Analysis Is Required To Determine The Fate Of Macroalgal Carbon: A Systematic Review. *Journal of Phycology*, *58*(3), 364–376. https://doi.org/10.1111/jpy.13251

Drula, E., Garron, M.-L., Dogan, S., Lombard, V., Henrissat, B., & Terrapon, N. (2022). The carbohydrate-active enzyme database: Functions and literature. *Nucleic Acids Research*, *50*(D1), Article D1. https://doi.org/10.1093/nar/gkab1045

D'Souza, G., Ebrahimi, A., Stubbusch, A., Daniels, M., Keegstra, J., Stocker, R., Cordero, O., & Ackermann, M. (2023). Cell aggregation is associated with enzyme secretion

strategies in marine polysaccharide-degrading bacteria. *The ISME Journal*, *17*(5), Article 5. https://doi.org/10.1038/s41396-023-01385-1

Du, Z.-Z., Zhou, L.-Y., Wang, T.-J., Li, H.-R., & Du, Z.-J. (2020). *Lutibacter citreus* sp. Nov., isolated from Arctic surface sediment. *International Journal of Systematic and Evolutionary Microbiology*, *70*(5), 3154–3161. https://doi.org/10.1099/ijsem.0.004146

Duarte, C. M., Gattuso, J., Hancke, K., Gundersen, H., Filbee-Dexter, K., Pedersen, M. F., Middelburg, J. J., Burrows, M. T., Krumhansl, K. A., Wernberg, T., Moore, P., Pessarrodona, A., Ørberg, S. B., Pinto, I. S., Assis, J., Queirós, A. M., Smale, D. A., Bekkby, T., Serrão, E. A., & Krause-Jensen, D. (2022). Global estimates of the extent and production of macroalgal forests. *Global Ecology and Biogeography*, *31*(7), 1422–1439. https://doi.org/10.1111/geb.13515

Duckworth, M., & Yaphe, W. (1971). The structure of agar: Part I. Fractionation of a complex mixture of polysaccharides. *Carbohydrate Research*, *16*(1), Article 1. https://doi.org/10.1016/S0008-6215(00)86113-3

Eddy, S. R. (2011). Accelerated Profile HMM Searches. *PLoS Computational Biology*, *7*(10), Article 10. https://doi.org/10.1371/journal.pcbi.1002195

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), Article 5. https://doi.org/10.1093/nar/gkh340

Eichhubl, P., Greene, H. G., Naehr, T., & Maher, N. (2000). Structural control of fluid flow: Offshore fluid seepage in the Santa Barbara Basin, California. *Journal of*

Geochemical Exploration, *69–70*, 545–549. https://doi.org/10.1016/S0375-6742(00)00107-2

El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., Qureshi, M., Richardson, L. J., Salazar, G. A., Smart, A., Sonnhammer, E. L. L., Hirsh, L., Paladin, L., Piovesan, D., Tosatto, S. C. E., & Finn, R. D. (2019). The Pfam protein families database in 2019. *Nucleic Acids Research*, *47*(D1), Article D1. https://doi.org/10.1093/nar/gky995

Engel, A., & Händel, N. (2011). A novel protocol for determining the concentration and composition of sugars in particulate and in high molecular weight dissolved organic matter (HMW-DOM) in seawater. *Marine Chemistry*, *127*(1–4), Article 1. https://doi.org/10.1016/j.marchem.2011.09.004

Eren, A. M., Esen, Ö. C., Quince, C., Vineis, J. H., Morrison, H. G., Sogin, M. L., & Delmont, T. O. (2015). Anvi'o: An advanced analysis and visualization platform for 'omics data. *PeerJ*, *3*, e1319. https://doi.org/10.7717/peerj.1319

Fernández-Gómez, B., Richter, M., Schüler, M., Pinhassi, J., Acinas, S. G., González, J. M., & Pedrós-Alió, C. (2013). Ecology of marine Bacteroidetes: A comparative genomics approach. *The ISME Journal*, *7*(5), 1026–1037. https://doi.org/10.1038/ismej.2012.169

Ficko-Blean, E., Préchoux, A., Thomas, F., Rochat, T., Larocque, R., Zhu, Y., Stam, M., Génicot, S., Jam, M., Calteau, A., Viart, B., Ropartz, D., Pérez-Pascual, D., Correc, G., Matard-Mann, M., Stubbs, K. A., Rogniaux, H., Jeudy, A., Barbeyron, T., … Michel, G. (2017). Carrageenan catabolism is encoded by a complex regulon in

marine heterotrophic bacteria. *Nature Communications*, *8*(1), Article 1. https://doi.org/10.1038/s41467-017-01832-6

Filbee-Dexter, K., Pedersen, M. F., Fredriksen, S., Norderhaug, K. M., Rinde, E., Kristiansen, T., Albretsen, J., & Wernberg, T. (2020). Carbon export is facilitated by sea urchins transforming kelp detritus. *Oecologia*, *192*(1), 213–225. https://doi.org/10.1007/s00442-019-04571-1

Flemming, H.-C., & Wingender, J. (2010). The biofilm matrix. *Nature Reviews Microbiology*, *8*(9), 623–633. https://doi.org/10.1038/nrmicro2415

Folch, J., Lees, M., & Stanley, G. H. S. (1957). A simple method for the isolation and purification of total lipides from animal tissues. *Journal of Biological Chemistry*, *226*(1), Article 1. https://doi.org/10.1016/S0021-9258(18)64849-5

Fracek, S. P., & Stolz, J. F. (1985). *Spirochaeta bajacaliforniensis* sp. N. From a microbial mat community at Laguna Figueroa, Baja California Norte, Mexico. *Archives of Microbiology*, *142*(4), 317–325. https://doi.org/10.1007/BF00491897

Fu, T., Jia, C., Fu, L., Zhou, S., Yao, P., Du, R., Sun, H., Yang, Z., Shi, X., & Zhang, X.-H. (2018). *Marinifilum breve* sp. Nov., a marine bacterium isolated from the Yongle Blue Hole in the South China Sea and emended description of the genus *Marinifilum*. *International Journal of Systematic and Evolutionary Microbiology*, *68*(11), 3540–3545. https://doi.org/10.1099/ijsem.0.003027

Garrity, G. M., Brenner, D. J., Krieg, N. R., & Staley, J. T. (2005). *Bergey's manual of systematic bacteriology. Vol. 2: The Proteobacteria.* (Second ed). Springer.

Gharechahi, J., Vahidi, M. F., Bahram, M., Han, J.-L., Ding, X.-Z., & Salekdeh, G. H. (2021). Metagenomic analysis reveals a dynamic microbiome with diversified

adaptive functions to utilize high lignocellulosic forages in the cattle rumen. *The ISME Journal*, *15*(4), 1108–1120. https://doi.org/10.1038/s41396-020-00837-2

Gilmore, S. P., Lankiewicz, T. S., Wilken, St. E., Brown, J. L., Sexton, J. A., Henske, J. K., Theodorou, M. K., Valentine, D. L., & O'Malley, M. A. (2019). Top-Down Enrichment Guides in Formation of Synthetic Microbial Consortia for Biomass Degradation. *ACS Synthetic Biology*, *8*(9), Article 9. https://doi.org/10.1021/acssynbio.9b00271

Glöckner, F. O., Kube, M., Bauer, M., Teeling, H., Lombardot, T., Ludwig, W., Gade, D., Beck, A., Borzym, K., Heitmann, K., Rabus, R., Schlesner, H., Amann, R., & Reinhardt, R. (2003). Complete genome sequence of the marine planctomycete Pirellula sp. Strain 1. *Proceedings of the National Academy of Sciences*, *100*(14), Article 14. https://doi.org/10.1073/pnas.1431443100

Glover, J. S., Ticer, T. D., & Engevik, M. A. (2022). Characterizing the mucin-degrading capacity of the human gut microbiota. *Scientific Reports*, *12*(1), Article 1. https://doi.org/10.1038/s41598-022-11819-z

Goecke, F., Labes, A., Wiese, J., & Imhoff, J. F. (2010). Chemical interactions between marine macroalgae and bacteria. *Marine Ecology Progress Series*, *409*, 267–299. https://doi.org/10.3354/meps08607

Goericke, R., Bograd, S. J., & Grundle, D. S. (2015). Denitrification and flushing of the Santa Barbara Basin bottom waters. *Deep Sea Research Part II: Topical Studies in Oceanography*, *112*, 53–60. https://doi.org/10.1016/j.dsr2.2014.07.012

Gong, G., Zhao, J., Wang, C., Wei, M., Dang, T., Deng, Y., Sun, J., Song, S., Huang, L., & Wang, Z. (2018). Structural characterization and antioxidant activities of the

degradation products from Porphyra haitanensis polysaccharides. *Process Biochemistry*, *74*, 185–193. https://doi.org/10.1016/j.procbio.2018.05.022

Grondin, J. M., Tamura, K., Déjean, G., Abbott, D. W., & Brumer, H. (2017). Polysaccharide Utilization Loci: Fueling Microbial Communities. *Journal of Bacteriology*, *199*(15), 10.1128/jb.00860-16. https://doi.org/10.1128/jb.00860-16

Haft, D. H. (2001). TIGRFAMs: A protein family resource for the functional identification of proteins. *Nucleic Acids Research*, *29*(1), Article 1. https://doi.org/10.1093/nar/29.1.41

Hamersley, M. R., Sohm, J. A., Burns, J. A., & Capone, D. G. (2015). Nitrogen fixation associated with the decomposition of the giant kelp Macrocystis pyrifera. *Aquatic Botany*, *125*, 57–63. https://doi.org/10.1016/j.aquabot.2015.05.003

Han, Y., & Perner, M. (2015). The globally widespread genus Sulfurimonas: Versatile energy metabolisms and adaptations to redox clines. *Frontiers in Microbiology*, *6*. https://doi.org/10.3389/fmicb.2015.00989

Hehemann, J.-H., Correc, G., Barbeyron, T., Helbert, W., Czjzek, M., & Michel, G. (2010). Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature*, *464*(7290), Article 7290. https://doi.org/10.1038/nature08937

Hehemann, J.-H., Correc, G., Thomas, F., Bernard, T., Barbeyron, T., Jam, M., Helbert, W., Michel, G., & Czjzek, M. (2012). Biochemical and Structural Characterization of the Complex Agarolytic Enzyme System from the Marine Bacterium *Zobellia galactanivorans*. *Journal of Biological Chemistry*, *287*(36), Article 36. https://doi.org/10.1074/jbc.M112.377184

Hehemann, J.-H., Kelly, A. G., Pudlo, N. A., Martens, E. C., & Boraston, A. B. (2012). Bacteria of the human gut microbiome catabolize red seaweed glycans with carbohydrate-active enzyme updates from extrinsic microbes. *Proceedings of the National Academy of Sciences*, *109*(48), Article 48. https://doi.org/10.1073/pnas.1211002109

Helbert, W. (2017). Marine Polysaccharide Sulfatases. *Frontiers in Marine Science*, *4*, 6. https://doi.org/10.3389/fmars.2017.00006

Hoshino, T., Doi, H., Uramoto, G.-I., Wörmer, L., Adhikari, R. R., Xiao, N., Morono, Y., D'Hondt, S., Hinrichs, K.-U., & Inagaki, F. (2020). Global diversity of microbial communities in marine sediment. *Proceedings of the National Academy of Sciences*, *117*(44), 27587–27597. https://doi.org/10.1073/pnas.1919139117

Hülsemann, J., & Emery, K. O. (1961). Stratification in Recent Sediments of Santa Barbara Basin as Controlled by Organisms and Water Character. *The Journal of Geology*, *69*(3), Article 3. https://doi.org/10.1086/626742

Hyatt, D., Chen, G.-L., LoCascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, *11*(1), Article 1. https://doi.org/10.1186/1471-2105-11-119

Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., & Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nature Communications*, *9*(1), Article 1. https://doi.org/10.1038/s41467-018-07641-9

Joshi, N., & Fass, J. (2011). *Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) [Software].* (1.33) [C]. https://github.com/najoshi/sickle (Original work published 2011)

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., & Tanabe, M. (2012). KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Research*, *40*(D1), Article D1. https://doi.org/10.1093/nar/gkr988

Kanehisa, M., Sato, Y., & Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. *Journal of Molecular Biology*, *428*(4), Article 4. https://doi.org/10.1016/j.jmb.2015.11.006

Kapili, B. J., Barnett, S. E., Buckley, D. H., & Dekas, A. E. (2020). Evidence for phylogenetically and catabolically diverse active diazotrophs in deep-sea sediment. *The ISME Journal*, *14*(4), 971–983. https://doi.org/10.1038/s41396-019-0584-8

Kawai, H., & Henry, E. C. (2016). Phaeophyta. In J. M. Archibald, A. G. B. Simpson, C. H. Slamovits, L. Margulis, M. Melkonian, D. J. Chapman, & J. O. Corliss (Eds.), *Handbook of the Protists* (pp. 1–38). Springer International Publishing. https://doi.org/10.1007/978-3-319-32669-6_31-1

Khadem, A. F., Pol, A., Jetten, M. S. M., & Op Den Camp, H. J. M. (2010). Nitrogen fixation by the verrucomicrobial methanotroph *'Methylacidiphilum fumariolicum'* SolV. *Microbiology*, *156*(4), Article 4. https://doi.org/10.1099/mic.0.036061-0

Kharbush, J. J., Close, H. G., Van Mooy, B. A. S., Arnosti, C., Smittenberg, R. H., Le Moigne, F. A. C., Mollenhauer, G., Scholz-Böttcher, B., Obreht, I., Koch, B. P., Becker, K. W., Iversen, M. H., & Mohr, W. (2020). Particulate Organic Carbon

Deconstructed: Molecular and Chemical Composition of Particulate Organic Carbon in the Ocean. *Frontiers in Marine Science*, *7*, 518. https://doi.org/10.3389/fmars.2020.00518

Kim, J. W., Brawley, S. H., Prochnik, S., Chovatia, M., Grimwood, J., Jenkins, J., LaButti, K., Mavromatis, K., Nolan, M., Zane, M., Schmutz, J., Stiller, J. W., & Grossman, A. R. (2016). Genome Analysis of Planctomycetes Inhabiting Blades of the Red Alga *Porphyra umbilicalis*. *PLOS ONE*, *11*(3), Article 3. https://doi.org/10.1371/journal.pone.0151883

Konstantinidis, K. T., Rosselló-Móra, R., & Amann, R. (2017). Uncultivated microbes in need of their own taxonomy. *The ISME Journal*, *11*(11), Article 11. https://doi.org/10.1038/ismej.2017.113

Konstantinidis, K. T., & Tiedje, J. M. (2005). Genomic insights that advance the species definition for prokaryotes. *Proceedings of the National Academy of Sciences*, *102*(7), Article 7. https://doi.org/10.1073/pnas.0409727102

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform. *Applied and Environmental Microbiology*, *79*(17), Article 17. https://doi.org/10.1128/AEM.01043-13

Krause-Jensen, D., & Duarte, C. M. (2016). Substantial role of macroalgae in marine carbon sequestration. *Nature Geoscience*, *9*(10), Article 10. https://doi.org/10.1038/ngeo2790

Krieg, N. R. (Ed.). (2011). *Bergey's manual of systematic bacteriology. Vol. 4: The Bacteroidetes, Spirochaetes, Tenericutes (Mollicutes), Acidobacteria, Fibrobacteres, Fusobacteria, Dictyoglomi, Gemmatimonadetes, Lentisphaerae, Verrucomicrobia, Chlamydiae, and Planctomycetes* (Second ed). Springer.

Kushner, D. J., Rassweiler, A., McLaughlin, J. P., & Lafferty, K. D. (2013). A multi-decade time series of kelp forest community structure at the California Channel Islands: Ecological Archives E094-245. *Ecology*, *94*(11), 2655–2655. https://doi.org/10.1890/13-0562R.1

Lagesen, K., Hallin, P., Rødland, E. A., Stærfeldt, H.-H., Rognes, T., & Ussery, D. W. (2007). RNAmmer: Consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Research*, *35*(9), Article 9. https://doi.org/10.1093/nar/gkm160

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), Article 4. https://doi.org/10.1038/nmeth.1923

Lapébie, P., Lombard, V., Drula, E., Terrapon, N., & Henrissat, B. (2019). Bacteroidetes use thousands of enzyme combinations to break down glycans. *Nature Communications*, *10*(1), 2043. https://doi.org/10.1038/s41467-019-10068-5

Lapointe, B. E., Brewton, R. A., Herren, L. W., Wang, M., Hu, C., McGillicuddy, D. J., Lindell, S., Hernandez, F. J., & Morton, P. L. (2021). Nutrient content and stoichiometry of pelagic *Sargassum* reflects increasing nitrogen availability in the Atlantic Basin. *Nature Communications*, *12*(1), 3060. https://doi.org/10.1038/s41467-021-23135-7

Le Moine Bauer, S., Roalkvam, I., Steen, I. H., & Dahle, H. (2016). *Lutibacter profundi* sp. Nov., isolated from a deep-sea hydrothermal system on the Arctic Mid-Ocean Ridge

and emended description of the genus *Lutibacter*. *International Journal of Systematic and Evolutionary Microbiology*, *66*(7), 2671–2677. https://doi.org/10.1099/ijsem.0.001105

Leimbach, A. (2016). *bac-genomics-scripts: Bovine E. coli mastitis comparative genomics edition* (bovine_ecoli_mastitis) [Computer software]. Zenodo. https://doi.org/10.5281/ZENODO.215824

Letunic, I., & Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Research*, *47*(W1), Article W1. https://doi.org/10.1093/nar/gkz239

Li, D., Liu, C.-M., Luo, R., Sadakane, K., & Lam, T.-W. (2015). MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct *de Bruijn* graph. *Bioinformatics*, *31*(10), Article 10. https://doi.org/10.1093/bioinformatics/btv033

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and samtools. *Bioinformatics*, *25*(16), Article 16. https://doi.org/10.1093/bioinformatics/btp352

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics (Oxford, England)*, *30*(7), 923–930. https://doi.org/10.1093/bioinformatics/btt656

Lie, T. J., Clawson, M. L., Godchaux, W., & Leadbetter, E. R. (1999). Sulfidogenesis from 2-Aminoethanesulfonate (Taurine) Fermentation by a Morphologically Unusual

Sulfate-Reducing Bacterium, *Desulforhopalus singaporensis* sp. Nov. *Applied and Environmental Microbiology*, *65*(8), 3328–3334. https://doi.org/10.1128/AEM.65.8.3328-3334.1999

Limoli, D. H., Jones, C. J., & Wozniak, D. J. (2015). Bacterial Extracellular Polysaccharides in Biofilm Formation and Function. *Microbiology Spectrum*, *3*(3), 10.1128/microbiolspec.mb-0011–2014. https://doi.org/10.1128/microbiolspec.mb-0011-2014

Liu, Q.-Q., Li, J., Xiao, D., Lu, J.-X., Chen, G.-J., & Du, Z.-J. (2015). *Saccharicrinis marinus* sp. Nov., isolated from marine sediment. *International Journal of Systematic and Evolutionary Microbiology*, *65*(Pt_10), 3427–3432. https://doi.org/10.1099/ijsem.0.000436

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., & Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Research*, *42*(D1), Article D1. https://doi.org/10.1093/nar/gkt1178

Lowe, T. M., & Chan, P. P. (2016). tRNAscan-SE On-line: Integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Research*, *44*(W1), Article W1. https://doi.org/10.1093/nar/gkw413

Lowman, H. E., Emery, K. A., Kubler-Dudgeon, L., Dugan, J. E., & Melack, J. M. (2019). Contribution of macroalgal wrack consumers to dissolved inorganic nitrogen concentrations in intertidal pore waters of sandy beaches. *Estuarine, Coastal and Shelf Science*, *219*, 363–371. https://doi.org/10.1016/j.ecss.2019.02.004

LTER, S. B. C., & Reed, D. C. (2019). *SBC LTER: Reef: Macrocystis pyrifera CHN content (carbon, hydrogen, nitrogen)* [dataset]. Environmental Data Initiative. https://doi.org/10.6073/PASTA/F2C945E3FB8559FA436DAB9D9CCE3491

Lu, D.-C., Zhao, J.-X., Wang, F.-Q., Xie, Z.-H., & Du, Z.-J. (2017). *Labilibacter aurantiacus* gen. Nov., sp. Nov., isolated from sea squirt (*Styela clava*) and reclassification of *Saccharicrinis marinus* as *Labilibacter marinus* comb. Nov. *International Journal of Systematic and Evolutionary Microbiology*, *67*(2), 441–446. https://doi.org/10.1099/ijsem.0.001649

Macreadie, P. I., Costa, M. D. P., Atwood, T. B., Friess, D. A., Kelleway, J. J., Kennedy, H., Lovelock, C. E., Serrano, O., & Duarte, C. M. (2021). Blue carbon as a natural climate solution. *Nature Reviews Earth & Environment*, *2*(12), 826–839. https://doi.org/10.1038/s43017-021-00224-1

Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., Lu, S., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, D., … Bryant, S. H. (2017). CDD/SPARCLE: Functional classification of proteins via subfamily domain architectures. *Nucleic Acids Research*, *45*(D1), Article D1. https://doi.org/10.1093/nar/gkw1129

Martens, E. C., Koropatkin, N. M., Smith, T. J., & Gordon, J. I. (2009). Complex Glycan Catabolism by the Human Gut Microbiota: The Bacteroidetes Sus-like Paradigm *. *Journal of Biological Chemistry*, *284*(37), Article 37. https://doi.org/10.1074/jbc.R109.022848

Martinez-Garcia, M., Brazel, D. M., Swan, B. K., Arnosti, C., Chain, P. S. G., Reitenga, K. G., Xie, G., Poulton, N. J., Gomez, M. L., Masland, D. E. D., Thompson, B., Bellows, W. K., Ziervogel, K., Lo, C.-C., Ahmed, S., Gleasner, C. D., Detter, C. J., & Stepanauskas, R. (2012). Capturing Single Cell Genomes of Active Polysaccharide Degraders: An Unexpected Contribution of *Verrucomicrobia*. *PLoS ONE*, *7*(4), Article 4. https://doi.org/10.1371/journal.pone.0035314

McKee, L. S., La Rosa, S. L., Westereng, B., Eijsink, V. G., Pope, P. B., & Larsbrink, J. (2021). Polysaccharide degradation by the Bacteroidetes: Mechanisms and nomenclature. *Environmental Microbiology Reports*, *13*(5), 559–581. https://doi.org/10.1111/1758-2229.12980

McKennedy, J., & Sherlock, O. (2015). Anaerobic digestion of marine macroalgae: A review. *Renewable and Sustainable Energy Reviews*, *52*, 1781–1790. https://doi.org/10.1016/j.rser.2015.07.101

McMurdie, P. J., & Holmes, S. (2013). phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE*, *8*(4), Article 4. https://doi.org/10.1371/journal.pone.0061217

Mitchell, A. L., Attwood, T. K., Babbitt, P. C., Blum, M., Bork, P., Bridge, A., Brown, S. D., Chang, H.-Y., El-Gebali, S., Fraser, M. I., Gough, J., Haft, D. R., Huang, H., Letunic, I., Lopez, R., Luciani, A., Madeira, F., Marchler-Bauer, A., Mi, H., … Finn, R. D. (2019). InterPro in 2019: Improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Research*, *47*(D1), Article D1. https://doi.org/10.1093/nar/gky1100

Miyazaki, M., Sakai, S., Yamanaka, Y., Saito, Y., Takai, K., & Imachi, H. (2014). *Spirochaeta psychrophila* sp. Nov., a psychrophilic spirochaete isolated from subseafloor sediment, and emended description of the genus *Spirochaeta*. *International Journal of Systematic and Evolutionary Microbiology*, *64*(Pt_8), 2798–2804. https://doi.org/10.1099/ijs.0.062463-0

Morrison, J. M., Murphy, C. L., Baker, K., Zamor, R. M., Nikolai, S. J., Wilder, S., Elshahed, M. S., & Youssef, N. H. (2017). Microbial communities mediating algal detritus turnover under anaerobic conditions. *PeerJ*, *5*, e2803. https://doi.org/10.7717/peerj.2803

Morton, D. N., Antonino, C. Y., Broughton, F. J., Dykman, L. N., Kuris, A. M., & Lafferty, K. D. (2021). A food web including parasites for kelp forests of the Santa Barbara Channel, California. *Scientific Data*, *8*(1), 99. https://doi.org/10.1038/s41597-021-00880-4

Mu, D.-S., Zhou, L.-Y., Liang, Q.-Y., Chen, G.-J., & Du, Z.-J. (2020). *Tichowtungia aerotolerans* gen. Nov., sp. Nov., a novel representative of the phylum *Kiritimatiellaeota* and proposal of *Tichowtungiaceae* fam. Nov., *Tichowtungiales* ord. Nov. And *Tichowtungiia* class. Nov. *International Journal of Systematic and Evolutionary Microbiology*, *70*(9), Article 9. https://doi.org/10.1099/ijsem.0.004370

Na, H., Kim, S., Moon, E. Y., & Chun, J. (2009). *Marinifilum fragile* gen. Nov., sp. Nov., isolated from tidal flat sediment. *INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY*, *59*(9), 2241–2246. https://doi.org/10.1099/ijs.0.009027-0

National Academies of Sciences, E., and Medicine. (2022). *A Research Strategy for Ocean-based Carbon Dioxide Removal and Sequestration*. The National Academies Press. https://doi.org/10.17226/26278

Nourbakhsh, F., Nasrollahzadeh, M. S., Tajani, A. S., Soheili, V., & Hadizadeh, F. (2022). Bacterial biofilms and their resistance mechanisms: A brief look at treatment with natural agents. *Folia Microbiologica*, *67*(4), 535–554. https://doi.org/10.1007/s12223-022-00955-8

Olivares, E., Badel-Berchoux, S., Provot, C., Prévost, G., Bernardi, T., & Jehl, F. (2020). Clinical Impact of Antibiotics for the Treatment of Pseudomonas aeruginosa Biofilm Infections. *Frontiers in Microbiology*, *10*, 2894. https://doi.org/10.3389/fmicb.2019.02894

Omand, M. M., Govindarajan, R., He, J., & Mahadevan, A. (2020). Sinking flux of particulate organic matter in the oceans: Sensitivity to particle characteristics. *Scientific Reports*, *10*(1), 5582. https://doi.org/10.1038/s41598-020-60424-5

Oren, A., & Garrity, G. M. (2021). Valid publication of the names of forty-two phyla of prokaryotes. *International Journal of Systematic and Evolutionary Microbiology*, *71*(10), Article 10. https://doi.org/10.1099/ijsem.0.005056

Ortega, A., Geraldi, N. R., Alam, I., Kamau, A. A., Acinas, S. G., Logares, R., Gasol, J. M., Massana, R., Krause-Jensen, D., & Duarte, C. M. (2019). Important contribution of macroalgae to oceanic carbon sequestration. *Nature Geoscience*, *12*(9), Article 9. https://doi.org/10.1038/s41561-019-0421-8

Parada, A. E., Needham, D. M., & Fuhrman, J. A. (2016). Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock communities, time

series and global field samples. *Environmental Microbiology*, *18*(5), 1403–1414. https://doi.org/10.1111/1462-2920.13023

Park, S. C., Choe, H. N., Hwang, Y. M., Baik, K. S., & Seong, C. N. (2013). *Lutibacter agarilyticus* sp. Nov., a marine bacterium isolated from shallow coastal seawater. *International Journal of Systematic and Evolutionary Microbiology*, *63*(Pt_7), 2678–2683. https://doi.org/10.1099/ijs.0.047670-0

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015). CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Research*, *25*(7), Article 7. https://doi.org/10.1101/gr.186072.114

Patel, A. K., Vadrale, A. P., Singhania, R. R., Michaud, P., Pandey, A., Chen, S.-J., Chen, C.-W., & Dong, C.-D. (2023). Algal polysaccharides: Current status and future prospects. *Phytochemistry Reviews*, *22*(4), 1167–1196. https://doi.org/10.1007/s11101-021-09799-5

Pedersen, M. F., Filbee-Dexter, K., Norderhaug, K. M., Fredriksen, S., Frisk, N. L., Fagerli, C. W., & Wernberg, T. (2020). Detrital carbon production and export in high latitude kelp forests. *Oecologia*, *192*(1), 227–239. https://doi.org/10.1007/s00442-019-04573-z

Peng, X., Wilken, St. E., Lankiewicz, T. S., Gilmore, S. P., Brown, J. L., Henske, J. K., Swift, C. L., Salamov, A., Barry, K., Grigoriev, I. V., Theodorou, M. K., Valentine, D. L., & O'Malley, M. A. (2021). Genomic and functional analyses of fungal and bacterial consortia that enable lignocellulose breakdown in goat gut microbiomes. *Nature Microbiology*, *6*(4), Article 4. https://doi.org/10.1038/s41564-020-00861-0

Popper, Z. A., Michel, G., Hervé, C., Domozych, D. S., Willats, W. G. T., Tuohy, M. G., Kloareg, B., & Stengel, D. B. (2011). Evolution and Diversity of Plant Cell Walls: From Algae to Flowering Plants. *Annual Review of Plant Biology*, *62*(1), Article 1. https://doi.org/10.1146/annurev-arplant-042110-103809

Pruesse, E., Peplies, J., & Glöckner, F. O. (2012). SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics*, *28*(14), Article 14. https://doi.org/10.1093/bioinformatics/bts252

Qin, Q., Kinnaman, F. S., Gosselin, K. M., Liu, N., Treude, T., & Valentine, D. L. (2022). Seasonality of water column methane oxidation and deoxygenation in a dynamic marine environment. *Geochimica et Cosmochimica Acta*, *336*(1), 219–230. https://doi.org/10.1016/j.gca.2022.09.017

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2012). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, *41*(D1), Article D1. https://doi.org/10.1093/nar/gks1219

Raut, Y., & Capone, D. G. (2021). Macroalgal detrital systems: An overlooked ecological niche for heterotrophic nitrogen fixation. *Environmental Microbiology*, *23*(8), 4372–4388. https://doi.org/10.1111/1462-2920.15622

Raut, Y., Morando, M., & Capone, D. G. (2018). Diazotrophic Macroalgal Associations With Living and Decomposing Sargassum. *Frontiers in Microbiology*, *9*. https://www.frontiersin.org/articles/10.3389/fmicb.2018.03127

Reintjes, G., Arnosti, C., Fuchs, B., & Amann, R. (2019). Selfish, sharing and scavenging bacteria in the Atlantic Ocean: A biogeographical study of bacterial substrate

utilisation. *The ISME Journal*, *13*(5), Article 5. https://doi.org/10.1038/s41396-018-0326-3

Reintjes, G., Arnosti, C., Fuchs, B. M., & Amann, R. (2017). An alternative polysaccharide uptake mechanism of marine bacteria. *The ISME Journal*, *11*(7), Article 7. https://doi.org/10.1038/ismej.2017.26

Reisky, L., Préchoux, A., Zühlke, M.-K., Bäumgen, M., Robb, C. S., Gerlach, N., Roret, T., Stanetty, C., Larocque, R., Michel, G., Song, T., Markert, S., Unfried, F., Mihovilovic, M. D., Trautwein-Schult, A., Becher, D., Schweder, T., Bornscheuer, U. T., & Hehemann, J.-H. (2019). A marine bacterial enzymatic cascade degrades the algal polysaccharide ulvan. *Nature Chemical Biology*, *15*(8), Article 8. https://doi.org/10.1038/s41589-019-0311-9

Rivas-Marín, E., & Devos, D. P. (2018). The Paradigms They Are a-Changin': Past, present and future of PVC bacteria research. *Antonie van Leeuwenhoek*, *111*(6), Article 6. https://doi.org/10.1007/s10482-017-0962-z

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, *26*(1), 139–140. https://doi.org/10.1093/bioinformatics/btp616

Rodriguez-R, L. M., & Konstantinidis, K. T. (2016). *The enveomics collection: A toolbox for specialized analyses of microbial genomes and metagenomes*. https://doi.org/10.7287/PEERJ.PREPRINTS.1900V1

Ruvira, M. A., Lucena, T., Pujalte, M. J., Arahal, D. R., & Macián, M. C. (2013). *Marinifilum flexuosum* sp. Nov., a new Bacteroidetes isolated from coastal Mediterranean Sea water and emended description of the genus *Marinifilum* Na et

al., 2009. *Systematic and Applied Microbiology*, *36*(3), 155–159. https://doi.org/10.1016/j.syapm.2012.12.003

Santarella-Mellwig, R., Franke, J., Jaedicke, A., Gorjanacz, M., Bauer, U., Budd, A., Mattaj, I. W., & Devos, D. P. (2010). The Compartmentalized Bacteria of the Planctomycetes-Verrucomicrobia-Chlamydiae Superphylum Have Membrane Coat-Like Proteins. *PLoS Biology*, *8*(1), Article 1. https://doi.org/10.1371/journal.pbio.1000281

Schirmer, T., Keller, T. A., Wang, Y.-F., & Rosenbusch, J. P. (1995). Structural Basis for Sugar Translocation Through Maltoporin Channels at 3.1 Å Resolution. *Science*, *267*(5197), 512–514. https://doi.org/10.1126/science.7824948

Schlitzer, R. (2018). *Ocean Data View* [Computer software]. https://odv.awi.de/

Seefeldt, L. C., Hoffman, B. M., & Dean, D. R. (2009). Mechanism of Mo-Dependent Nitrogenase. *Annual Review of Biochemistry*, *78*(1), Article 1. https://doi.org/10.1146/annurev.biochem.78.070907.103812

Sholkovitz, E. (1973). Interstitial water chemistry of the Santa Barbara Basin sediments. *Geochimica et Cosmochimica Acta*, *37*(9), Article 9. https://doi.org/10.1016/0016-7037(73)90008-2

Sholkovitz, E. R., & Gieskes, J. M. (1971). A Physical-Chemical Study of the Flushing of the Santa Barbara Basin. *Limnology and Oceanography*, *16*(3), Article 3. https://doi.org/10.4319/lo.1971.16.3.0479

Sichert, A., Corzett, C. H., Schechter, M. S., Unfried, F., Markert, S., Becher, D., Fernandez-Guerra, A., Liebeke, M., Schweder, T., Polz, M. F., & Hehemann, J.-H. (2020). Verrucomicrobia use hundreds of enzymes to digest the algal polysaccharide

fucoidan. *Nature Microbiology*, *5*(8), Article 8. https://doi.org/10.1038/s41564-020-0720-2

Skerker, J. M., Prasol, M. S., Perchuk, B. S., Biondi, E. G., & Laub, M. T. (2005). Two-Component Signal Transduction Pathways Regulating Growth and Cell Cycle Progression in a Bacterium: A System-Level Analysis. *PLoS Biology*, *3*(10), Article 10. https://doi.org/10.1371/journal.pbio.0030334

SOUTAR, A., & CRILL, P. A. (1977). Sedimentation and climatic patterns in the Santa Barbara Basin during the 19th and 20th centuries. *GSA Bulletin*, *88*(8), Article 8. https://doi.org/10.1130/0016-7606(1977)88<1161:SACPIT>2.0.CO;2

Soutourina, O. A., & Bertin, P. N. (2003). Regulation cascade of flagellar expression in Gram-negative bacteria. *FEMS Microbiology Reviews*, *27*(4), 505–523. https://doi.org/10.1016/S0168-6445(03)00064-0

Spring, S., Brinkmann, N., Murrja, M., Spröer, C., Reitner, J., & Klenk, H.-P. (2015). High Diversity of Culturable Prokaryotes in a Lithifying Hypersaline Microbial Mat. *Geomicrobiology Journal*, *32*(3–4), Article 3–4. https://doi.org/10.1080/01490451.2014.913095

Spring, S., Bunk, B., Spröer, C., Schumann, P., Rohde, M., Tindall, B. J., & Klenk, H.-P. (2016). Characterization of the first cultured representative of *Verrucomicrobia* subdivision 5 indicates the proposal of a novel phylum. *The ISME Journal*, *10*(12), Article 12. https://doi.org/10.1038/ismej.2016.84

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), Article 9. https://doi.org/10.1093/bioinformatics/btu033

Steneck, R. S., Graham, M. H., Bourque, B. J., Corbett, D., Erlandson, J. M., Estes, J. A., & Tegner, M. J. (2002). Kelp forest ecosystems: Biodiversity, stability, resilience and future. *Environmental Conservation*, *29*(4), 436–459. https://doi.org/10.1017/S0376892902000322

Suess, E. (1980). Particulate organic carbon flux in the oceans—Surface productivity and oxygen utilization. *Nature*, *288*(5788), 260–263. https://doi.org/10.1038/288260a0

Sung, H.-R., Shin, K.-S., & Ghim, S.-Y. (2015). *Lutibacter oricola* sp. Nov., a marine bacterium isolated from seawater. *International Journal of Systematic and Evolutionary Microbiology*, *65*(Pt_2), 485–490. https://doi.org/10.1099/ijs.0.067132-0

Svartström, O., Alneberg, J., Terrapon, N., Lombard, V., De Bruijn, I., Malmsten, J., Dalin, A.-M., El Muller, E., Shah, P., Wilmes, P., Henrissat, B., Aspeborg, H., & Andersson, A. F. (2017). Ninety-nine de novo assembled genomes from the moose (Alces alces) rumen microbiome provide new insights into microbial plant biomass degradation. *The ISME Journal*, *11*(11), 2538–2551. https://doi.org/10.1038/ismej.2017.108

Tatusov, R. L., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Kiryutin, B., Koonin, E. V., Krylov, D. M., Mazumder, R., Mekhedov, S. L., Nikolskaya, A. N., Rao, B. S., Smirnov, S., Sverdlov, A. V., Vasudevan, S., Wolf, Y. I., Yin, J. J., & Natale, D. A. (2003). The COG database: An updated version includes eukaryotes. *BMC Bioinformatics*, *4*(1), Article 1. https://doi.org/10.1186/1471-2105-4-41

TerAvest, M. A., He, Z., Rosenbaum, M. A., Martens, E. C., Cotta, M. A., Gordon, J. I., & Angenent, L. T. (2014). Regulated expression of polysaccharide utilization and

capsular biosynthesis loci in biofilm and planktonic *Bacteroides thetaiotaomicron* during growth in chemostats. *Biotechnology and Bioengineering*, *111*(1), 165–173. https://doi.org/10.1002/bit.24994

Thomas, F., Hehemann, J.-H., Rebuffet, E., Czjzek, M., & Michel, G. (2011). Environmental and Gut Bacteroidetes: The Food Connection. *Frontiers in Microbiology*, *2*. https://doi.org/10.3389/fmicb.2011.00093

Tokuda, G., Mikaelyan, A., Fukui, C., Matsuura, Y., Watanabe, H., Fujishima, M., & Brune, A. (2018). Fiber-associated spirochetes are major agents of hemicellulose degradation in the hindgut of wood-feeding higher termites. *Proceedings of the National Academy of Sciences*, *115*(51). https://doi.org/10.1073/pnas.1810550115

Usov, A. I. (2011). Polysaccharides of the red algae. In D. Horton (Ed.), *Advances in Carbohydrate Chemistry and Biochemistry* (Vol. 65, pp. 115–217). Academic Press. https://doi.org/10.1016/B978-0-12-385520-6.00004-2

V, G. S., M, D. K., Pugazhendi, A., Bajhaiya, A. K., Gugulothu, P., & J, R. B. (2021). Biofuel production from Macroalgae: Present scenario and future scope. *Bioengineered*, *12*(2), 9216–9238. https://doi.org/10.1080/21655979.2021.1996019

Valentine, D. L., Fisher, G. B., Pizarro, O., Kaiser, C. L., Yoerger, D., Breier, J. A., & Tarn, J. (2016). Autonomous Marine Robotic Technology Reveals an Expansive Benthic Bacterial Community Relevant to Regional Nitrogen Biogeochemistry. *Environmental Science & Technology*, *50*(20), Article 20. https://doi.org/10.1021/acs.est.6b03584

van Vliet, D. M., Lin, Y., Bale, N. J., Koenen, M., Villanueva, L., Stams, A. J. M., & Sánchez-Andrea, I. (2020). *Pontiella desulfatans* gen. Nov., sp. Nov., and *Pontiella*

sulfatireligans sp. Nov., Two Marine Anaerobes of the *Pontiellaceae* fam. Nov. Producing Sulfated Glycosaminoglycan-like Exopolymers. *Microorganisms*, *8*(6), Article 6. https://doi.org/10.3390/microorganisms8060920

van Vliet, D. M., Palakawong Na Ayudthaya, S., Diop, S., Villanueva, L., Stams, A. J. M., & Sánchez-Andrea, I. (2019). Anaerobic Degradation of Sulfated Polysaccharides by Two Novel *Kiritimatiellales* Strains Isolated From Black Sea Sediment. *Frontiers in Microbiology*, *10*, 253. https://doi.org/10.3389/fmicb.2019.00253

Vishchuk, O. S., Ermakova, S. P., & Zvyagintseva, T. N. (2011). Sulfated polysaccharides from brown seaweeds Saccharina japonica and Undaria pinnatifida: Isolation, structural characteristics, and antitumor activity. *Carbohydrate Research*, *346*(17), 2769–2776. https://doi.org/10.1016/j.carres.2011.09.034

Wagner, M., & Horn, M. (2006). The *Planctomycetes*, *Verrucomicrobia*, *Chlamydiae* and sister phyla comprise a superphylum with biotechnological and medical relevance. *Current Opinion in Biotechnology*, *17*(3), Article 3. https://doi.org/10.1016/j.copbio.2006.05.005

Waite, D. W., Vanwonterghem, I., Rinke, C., Parks, D. H., Zhang, Y., Takai, K., Sievert, S. M., Simon, J., Campbell, B. J., Hanson, T. E., Woyke, T., Klotz, M. G., & Hugenholtz, P. (2017). Comparative Genomic Analysis of the Class Epsilonproteobacteria and Proposed Reclassification to Epsilonbacteraeota (phyl. Nov.). *Frontiers in Microbiology*, *8*, 682. https://doi.org/10.3389/fmicb.2017.00682

Wang, F.-Q., Chen, Z.-J., Yang, J.-M., Wang, W.-J., Feng, Y.-W., Li, Z., & Sun, G.-H. (2020). *Labilibacter sediminis* sp. Nov., isolated from marine sediment.

International Journal of Systematic and Evolutionary Microbiology, *70*(1), 321–326. https://doi.org/10.1099/ijsem.0.003758

Wang, H., Wilksch, J. J., Strugnell, R. A., & Gee, M. L. (2015). Role of Capsular Polysaccharides in Biofilm Formation: An AFM Nanomechanics Study. *ACS Applied Materials & Interfaces*, *7*(23), 13007–13013. https://doi.org/10.1021/acsami.5b03041

Wang, J., Zhang, Q., Zhang, Z., Zhang, H., & Niu, X. (2010). Structural studies on a novel fucogalactan sulfate extracted from the brown seaweed Laminaria japonica. *International Journal of Biological Macromolecules*, *47*(2), 126–131. https://doi.org/10.1016/j.ijbiomac.2010.05.010

Wang, M., Hu, C., Barnes, B. B., Mitchum, G., Lapointe, B., & Montoya, J. P. (2019). The great Atlantic *Sargassum* belt. *Science*, *365*(6448), 83–87. https://doi.org/10.1126/science.aaw7912

Washburn, L., Clark, J. F., & Kyriakidis, P. (2005). The spatial scales, distribution, and intensity of natural marine hydrocarbon seeps near Coal Oil Point, California. *Marine and Petroleum Geology*, *22*(4), Article 4. https://doi.org/10.1016/j.marpetgeo.2004.08.006

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., & Barton, G. J. (2009). Jalview Version 2—A multiple sequence alignment editor and analysis workbench. *Bioinformatics*, *25*(9), Article 9. https://doi.org/10.1093/bioinformatics/btp033

Wei, N., Quarterman, J., & Jin, Y.-S. (2013). Marine macroalgae: An untapped resource for producing fuels and chemicals. *Trends in Biotechnology*, *31*(2), Article 2. https://doi.org/10.1016/j.tibtech.2012.10.009

Wertz, J. T., Kim, E., Breznak, J. A., Schmidt, T. M., & Rodrigues, J. L. M. (2012). Genomic and Physiological Characterization of the *Verrucomicrobia* Isolate *Diplosphaera colitermitum* gen. Nov., sp. Nov., Reveals Microaerophily and Nitrogen Fixation Genes. *Applied and Environmental Microbiology*, *78*(5), Article 5. https://doi.org/10.1128/AEM.06466-11

Wirén, N. V., & Merrick, M. (2004). Regulation and function of ammonium carriers in bacteria, fungi, and plants. In *Molecular Mechanisms Controlling Transmembrane Transport* (Vol. 9, pp. 95–120). Springer Berlin Heidelberg. https://doi.org/10.1007/b95775

Wu, Y.-T., Huo, Y.-F., Xu, L., Xu, Y.-Y., Wang, X.-L., & Zhou, T. (2020). Purification, characterization and antioxidant activity of polysaccharides from Porphyra haitanensis. *International Journal of Biological Macromolecules*, *165*, 2116–2125. https://doi.org/10.1016/j.ijbiomac.2020.10.053

Xu, Z.-X., Mu, X., Zhang, H.-X., Chen, G.-J., & Du, Z.-J. (2016). *Marinifilum albidiflavum* sp. Nov., isolated from coastal sediment. *International Journal of Systematic and Evolutionary Microbiology*, *66*(11), 4589–4593. https://doi.org/10.1099/ijsem.0.001395

Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F. O., Ludwig, W., Schleifer, K.-H., Whitman, W. B., Euzéby, J., Amann, R., & Rosselló-Móra, R. (2014). Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nature Reviews Microbiology*, *12*(9), Article 9. https://doi.org/10.1038/nrmicro3330

Yousavich, D. J., Robinson, D., Peng, X., Krause, S. J. E., Wenzhoefer, F., Janßen, F., Liu, N., Tarn, J., Kinnaman, F., Valentine, D. L., & Treude, T. (2023). Marine anoxia initiates giant sulfur-bacteria mat proliferation and associated changes in benthic nitrogen, sulfur, and iron cycling in the Santa Barbara Basin, California Borderland. *EGUsphere*, *2023*, 1–48. https://doi.org/10.5194/egusphere-2023-1198

Yu, N. Y., Wagner, J. R., Laird, M. R., Melli, G., Rey, S., Lo, R., Dao, P., Sahinalp, S. C., Ester, M., Foster, L. J., & Brinkman, F. S. L. (2010). PSORTb 3.0: Improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics*, *26*(13), 1608–1615. https://doi.org/10.1093/bioinformatics/btq249

Zehr, J. P., Jenkins, B. D., Short, S. M., & Steward, G. F. (2003). Nitrogenase gene diversity and microbial community structure: A cross-system comparison. *Environmental Microbiology*, *5*(7), Article 7. https://doi.org/10.1046/j.1462-2920.2003.00451.x

Zhang, H., Yohe, T., Huang, L., Entwistle, S., Wu, P., Yang, Z., Busk, P. K., Xu, Y., & Yin, Y. (2018). dbCAN2: A meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Research*, *46*(W1), Article W1. https://doi.org/10.1093/nar/gky418

Zhang, Q., Li, N., Liu, X., Zhao, Z., Li, Z., & Xu, Z. (2004). The structure of a sulfated galactan from *Porphyra haitanensis* and its in vivo antioxidant activity. *Carbohydrate Research*, *339*(1), Article 1. https://doi.org/10.1016/j.carres.2003.09.015

Zheng, J., Ge, Q., Yan, Y., Zhang, X., Huang, L., & Yin, Y. (2023). dbCAN3: Automated carbohydrate-active enzyme and substrate annotation. *Nucleic Acids Research*, *51*(W1), W115–W121. https://doi.org/10.1093/nar/gkad328

Zheng, M., Ma, M., Yang, Y., Liu, Z., Liu, S., Hong, T., Ni, H., & Jiang, Z. (2023). Structural characterization and antioxidant activity of polysaccharides extracted from Porphyra haitanensis by different methods. *International Journal of Biological Macromolecules*, *242*, 125003. https://doi.org/10.1016/j.ijbiomac.2023.125003

Zhu, X.-Y., Li, Y., Xue, C.-X., Lidbury, I. D. E. A., Todd, J. D., Lea-Smith, D. J., Tian, J., Zhang, X.-H., & Liu, J. (2023). Deep-sea Bacteroidetes from the Mariana Trench specialize in hemicellulose and pectin degradation typically associated with terrestrial systems. *Microbiome*, *11*(1), 175. https://doi.org/10.1186/s40168-023-01618-7

Zhu, Y., Chen, P., Bao, Y., Men, Y., Zeng, Y., Yang, J., Sun, J., & Sun, Y. (2016). Complete genome sequence and transcriptomic analysis of a novel marine strain Bacillus weihaiensis reveals the mechanism of brown algae degradation. *Scientific Reports*, *6*(1), Article 1. https://doi.org/10.1038/srep38248

Appendix

Table S1.1. Comparison of percentages of total fatty acids between cultivated strains in the *Kiritimatiellota* phylum. Average % of fatty acids of strain NLcol2 was calculated from six replicates. Major fatty acids (>5%) are labeled in bold. i- indicates iso- branched fatty acids.

FAME compound	*P. agarivorans* NLcol2 (%)	*P. desulfatans* F1 (%)	*P. sulfatireligans* F21 (%)	*K. glycovorans* L21-Fru-AB (%)	*T. aerotolerans* S-5007 (%)
i-C12:0	**29.1**	**21.5**	**27.5**	-	**19.8**
C12:0	0.3	0.3	0.2	2.7	-
i-C14:0	**5.8**	**10.3**	**6.0**	**47.0**	**8**
C14:0	0.2	0.2	0.2	0.7	-
i-C16:0	2.3	2.5	**5.3**	0.7	**10.2**
C16:0	4.8	2.9	3.5	1.9	**6.6**
i-C18:0	**11.8**	**7.2**	**13.1**	4.3	**10.5**
C18:0	**42.6**	**40.1**	**35.3**	**37.6**	**30.4**
i-C20:0	0.7	0.9	0.2	-	-
C20:0	2.3	2.4	0.5	2.1	1.3

Table S1.2a. List of genes potentially involved in agarose degradation.

Gene ID	SignalP	dbCAN
c_000000000020_87	N	GH16
c_000000000020_42	Y(1-19)	GH16
c_000000000027_439	Y(1-23)	GH16
c_000000000030_907	Y(1-25)	GH16
c_000000000026_443	Y(1-19)	GH16
c_000000000029_59	Y(1-23)	GH16
c_000000000018_53	Y(1-25)	GH16
c_000000000029_62	Y(1-39)	GH16
c_000000000030_1001	Y(1-30)	GH16
c_000000000027_632	N	GH86
c_000000000030_799	Y(1-21)	GH86
c_000000000027_710	Y(1-20)	GH86
c_000000000020_67	Y(1-19)	GH86
c_000000000020_66	Y(1-23)	GH86
c_000000000020_60	Y(1-22)	GH86
c_000000000030_1009	N	GH86
c_000000000027_530	Y(1-23)	GH86
c_000000000030_1011	N	GH86
c_000000000027_541	N	GH86
c_000000000030_773	N	GH86

c_000000000020_73	Y(1-25)	GH86
c_000000000030_804	Y(1-28)	GH86
c_000000000018_55	Y(1-18)	GH86
c_000000000027_545	N	GH86
c_000000000027_438	Y(1-11)	GH86
c_000000000028_672	Y(1-24)	GH86
c_000000000020_86	N	GH50
c_000000000028_627	Y(1-23)	GH50
c_000000000027_577	Y(1-22)	GH50
c_000000000028_646	Y(1-27)	GH50
c_000000000027_533	Y(1-24)	GH50
c_000000000027_709	N	GH50
c_000000000028_611	N	GH117
c_000000000030_1003	N	GH117
c_000000000027_531	N	GH117
c_000000000030_913	N	GH117
c_000000000027_36	Y(1-22)	GH117
c_000000000020_81	N	GH117
c_000000000030_1008	N	GH117
c_000000000018_60	N	GH117
c_000000000030_910	N	GH117
c_000000000030_912	N	GH117
c_000000000020_57	Y(1-24)	GH117
c_000000000030_1012	Y(1-22)	GH117

Table S1.2b. List of genes potentially involved in iota-carrageenan degradation.

Gene ID	Signal P	dbCAN/SulfAtlas
c_000000000027_83	Y(1-23)	GH82
c_000000000027_68	Y(1-22)	GH82
c_000000000029_21	N	GH82
c_000000000030_507	Y(1-23)	GH82
c_000000000020_44	N	GH127
c_000000000027_615	Y(1-19)	GH127
c_000000000027_575	N	GH127
c_000000000027_61	Y(1-20)	GH127
c_000000000018_56	Y	S1_17
c_000000000030_479	Y	S1_17
c_000000000027_10	Y	S1_17
c_000000000027_67	Y	S1_17
c_000000000028_621	Y	S1_17
c_000000000030_477	Y	S1_17
c_000000000027_82	Y	S1_17
c_000000000028_607	Y	S1_17

c_000000000027_14	Y	S1_17
c_000000000027_492	N	S1_19
c_000000000027_546	Y	S1_19
c_000000000027_582	Y	S1_19
c_000000000020_84	Y	S1_19
c_000000000029_63	Y	S1_19
c_000000000018_39	Y	S1_19
c_000000000028_682	Y	S1_19
c_000000000027_74	Y	S1_19
c_000000000028_657	Y	S1_19
c_000000000020_38	Y	S1_19
c_000000000018_18	Y	S1_19
c_000000000020_48	Y	S1_19
c_000000000018_17	Y	S1_19
c_000000000020_83	Y	S1_19

Table S1.2c. List of genes potentially involved in fucoidan degradation.

Gene ID	Signal P	dbCAN/SulfAtlas
c_000000000020_59	N	GH141
c_000000000018_63	N	GH29
c_000000000018_64	N	GH29
c_000000000018_65	Y(1-22)	GH29
c_000000000018_66	Y(1-27)	GH29
c_000000000018_67	Y(1-25)	GH29
c_000000000020_51	Y(1-19)	GH29
c_000000000027_242	N	GH29
c_000000000027_487	Y(1-20)	GH29
c_000000000027_535	N	GH29
c_000000000027_566	Y(1-26)	GH29
c_000000000027_627	Y(1-22)	GH29
c_000000000027_647	Y(1-19)	GH29
c_000000000027_715	Y(1-27)	GH29
c_000000000028_324	Y(1-23)	GH29
c_000000000028_602	Y(1-21)	GH29
c_000000000028_603	N	GH29
c_000000000028_642	Y(1-20)	GH29
c_000000000028_656	Y(1-21)	GH29
c_000000000028_670	Y(1-24)	GH29
c_000000000029_1	Y(1-27)	GH29
c_000000000029_129	Y(1-12)	GH29
c_000000000029_16	N	GH29
c_000000000029_23	N	GH29

c_000000000030_1007	N	GH29
c_000000000030_1014	Y(1-19)	GH29
c_000000000030_383	N	GH29
c_000000000030_384	Y(1-21)	GH29
c_000000000030_385	Y(1-26)	GH29
c_000000000030_906	N	GH29
c_000000000027_542	Y(1-21)	GH95
c_000000000027_80	Y(1-22)	GH95
c_000000000029_15	Y(1-23)	GH95
c_000000000027_2	Y	S1_15
c_000000000027_211	Y	S1_15
c_000000000027_516	Y	S1_15
c_000000000027_52	Y	S1_15
c_000000000027_571	Y	S1_15
c_000000000027_635	Y	S1_15
c_000000000027_711	Y	S1_15
c_000000000027_8	Y	S1_15
c_000000000027_9	Y	S1_15
c_000000000029_14	Y	S1_15
c_000000000030_203	Y	S1_15
c_000000000030_730	Y	S1_15
c_000000000030_901	Y	S1_15
c_000000000030_982	Y	S1_15
c_000000000030_989	Y	S1_15
c_000000000030_997	Y	S1_15
c_000000000026_442	Y	S1_16
c_000000000027_28	Y	S1_16
c_000000000027_35	Y	S1_16
c_000000000027_366	Y	S1_16
c_000000000027_41	Y	S1_16
c_000000000027_51	Y	S1_16
c_000000000027_517	Y	S1_16
c_000000000027_556	Y	S1_16
c_000000000027_56	Y	S1_16
c_000000000027_643	Y	S1_16
c_000000000027_89	Y	S1_16
c_000000000028_313	N	S1_16
c_000000000029_22	Y	S1_16
c_000000000030_166	Y	S1_16
c_000000000030_503	Y	S1_16
c_000000000030_505	Y	S1_16
c_000000000030_508	Y	S1_16
c_000000000018_56	Y	S1_17

c_000000000027_10	Y	S1_17
c_000000000027_14	Y	S1_17
c_000000000027_67	Y	S1_17
c_000000000027_82	Y	S1_17
c_000000000028_607	Y	S1_17
c_000000000028_621	Y	S1_17
c_000000000030_477	Y	S1_17
c_000000000030_479	Y	S1_17
c_000000000027_4	Y	S1_25
c_000000000027_76	Y	S1_25
c_000000000028_622	Y	S1_25
c_000000000030_491	Y	S1_25
c_000000000030_903	Y	S1_25

Table S1.3. List of *nif* genes annotated by Tigrfam and Pfam databases.

Gene ID	Tigrfam ID	Tigrfam annotation	Pfam ID	Pfam annotation
c_000000000018_26	TIGR01817	nifA, a DNA-binding regulatory protein for nitrogen fixation	PF00158	Sigma-54 interaction domain
c_000000000026_50	NA	NA	PF02579	Dinitorgenase FeMo cofactor
c_000000000026_55	NA	NA	PF02579	Dinitorgenase FeMo cofactor
c_000000000027_177	TIGR02933	nifM, accessory protein for NifH	PF00693	PPIC-type PPIASE domain
c_000000000027_466	TIGR03402	nifS, cysteine desulfurase involved in iron-sulfur (FeS) cluster biosynthesis	PF00266	Aminotransferase class-V
c_000000000027_623	TIGR01581	nifC-like ABC-type molybdate porter	PF00528	Binding-protein-dependent transport system inner membrane component
c_000000000028_195	NA	NA	PF00543	Nitrogen regulatory protein P-II
c_000000000028_221	NA	NA	PF00543	Nitrogen regulatory protein P-II
c_000000000028_411	TIGR02933	nifM, nitrogen fixation protein	PF13616	PPIC-type PPIASE domain
c_000000000028_484	NA	NA	PF00543	Nitrogen regulatory protein P-II
c_000000000028_865	TIGR02933	nifM, nitrogen fixation protein	PF00693	PPIC-type PPIASE domain
c_000000000028_912	TIGR01290	nifB, nitrogenase cofactor biosynthesis protein	PF04055	Radical SAM superfamily
c_000000000028_913	TIGR01286	nifK, nitrogenase molybdenum-iron protein beta chain	PF00148	nifDK, MoFe nitrogenase complex
c_000000000028_915	TIGR01283	nifE, nitrogenase MoFe cofactor biosysnthesis	PF00148	nifDK, MoFe nitrogenase complex
c_000000000028_920	NA	NA	PF05082	Rop-like protein with unknown function
c_000000000028_921	TIGR01616	uncharacterized protein associated with nitrogenase	NA	NA
c_000000000028_924	TIGR01286	nifK, nitrogenase molybdenum-iron protein beta chain	PF00148	nifDK, MoFe nitrogenase complex
c_000000000028_925	TIGR01282	nifD, nitrogenase molybdenum-iron protein alpha chain	PF00148	nifDK, MoFe nitrogenase complex
c_000000000028_926	NA	NA	PF00543	Nitrogen regulatory protein P-II
c_000000000028_927	NA	NA	PF00543	Nitrogen regulatory protein P-II
c_000000000028_928	TIGR01287	nifH, nitrogenase iron protein	PF00142	nifH

c_000000000028_1065	TIGR01817	nifA, a DNA-binding regulatory protein for nitrogen fixation	PF00158	Sigma-54 interaction domain
c_000000000030_132	TIGR01817	nifA, a DNA-binding regulatory protein for nitrogen fixation	PF00158	Sigma-54 interaction domain
c_000000000030_136	TIGR01817	nifA, a DNA-binding regulatory protein for nitrogen fixation	PF00158	Sigma-54 interaction domain
c_000000000030_137	TIGR02660	nifV, homocitrate synthase	PF00682	HMGL-like pyruvate carboxylase
c_000000000030_215	TIGR02933	nifM, accessory protein for NifH	PF13145	PPIC-type PPIASE domain
c_000000000030_754	TIGR03341	IscR-regulated protein YhgI, an iron-sulfur cluster biosynthesis protein	PF01106	nifU-like C terminal domain
c_000000000030_868	TIGR02000	nifU, FeS cluster assembly protein	PF01592	nifU-like N terminal domain
c_000000000030_869	TIGR03402	nifS, cysteine desulfurase involved in iron-sulfur (FeS) cluster biosynthesis	PF00266	Aminotransferase class-V

Table S2.1a. List of genes involved in flagella formation and their expression level as logFC when growth on polymers compared to growth on D-galactose.

Gene ID	Symbol	Pfam	Description	KO	KO_annotation	agar_logFC	agar_FDR	brown_seaweed_logFC	brown_seaweed_FDR	red_seaweed_logFC	red_seaweed_FDR
c_000000000027_245	Flagellin_N	PF00669.20	Bacterial flagellin N-terminal helical region	K02397	flagellar hook-associated protein 3 FlgL	1.25812	3.01E-05	0.746777	0.003113	1.447611	3.78E-06
c_000000000027_249	Flagellin_N	PF00669.20	Bacterial flagellin N-terminal helical region	K02406	flagellin	1.431979	0.000267	5.998051	6.79E-10	4.836845	1.11E-09
c_000000000027_250	Flagellin_C	PF00700.21	Bacterial flagellin C-terminal helical region	K02406	flagellin	1.66315	1.58E-06	6.013493	3.65E-11	5.375555	4.36E-11
c_000000000027_251	Flagellin_N	PF00669.20	Bacterial flagellin N-terminal helical region	K02406	flagellin	1.781382	1.26E-05	6.121947	4.84E-11	5.481002	1.08E-10

c_000000000027_252	Flagellin_N	PF00669.20	Bacterial flagellin N-terminal helical region	K02406	flagellin	0.674555	0.010008	5.160066	6.12E-10	4.10243	1.11E-09
c_000000000027_253	FliD_C	PF07195.12	Flagellar hook-associated protein 2 C-terminus	K02407	flagellar hook-associated protein 2	1.03903	1.70E-05	3.892192	5.74E-11	3.536067	1.21E-10
c_000000000027_255	FliW	PF02623.15	FliW protein	K13626	flagellar assembly factor FliW	1.823282	0.001344	3.056681	6.39E-06	3.657032	2.04E-07
c_000000000027_257	Flagellin_N	PF00669.20	Bacterial flagellin N-terminal helical region	K02406	flagellin	1.7293	5.92E-05	2.667413	8.43E-07	2.770806	1.64E-07
c_000000000027_265	FliS	PF02561.14	Flagellar protein FliS	K02422	flagellar secretion chaperone FliS	-0.59599	0.035753	0.004663	0.989119	-0.10837	0.674342
c_000000000027_268	FliP	PF00813.20	FliP family	K02419	flagellar biosynthesis protein FliP	-0.0222	0.961175	1.067482	0.009393	1.454581	0.000502
c_000000000027_272	FHIPEP	PF00771.20	FHIPEP family	K02400	flagellar biosynthesis protein FlhA	0.426214	0.085076	1.332182	4.51E-05	2.124455	1.26E-07
c_000000000027_273	CbiA	PF01656.23	CobQ/CobB/MinD/ParA nucleotide binding domain	K04562	flagellar biosynthesis protein FlhG	0.298412	0.289657	1.582761	1.02E-05	2.044752	1.94E-07
c_000000000027_280	Flg_bb_rod	PF00460.20	Flagella basal body rod protein	K02387	flagellar basal-body rod protein FlgB	1.07572	0.063791	2.35639	0.000218	2.754327	1.49E-05
c_000000000027_281	Flg_bbr_C	PF06429.13	Flagellar basal body rod FlgEFG protein C-terminal	K02388	flagellar basal-body rod protein FlgC	1.661042	0.001836	3.118481	2.50E-06	3.019801	1.22E-06
c_000000000027_283	YscJ_FliF	PF01514.17	Secretory protein of YscJ/FliF family	K02409	flagellar M-ring protein FliF	1.191372	0.001245	2.267169	2.70E-06	2.537094	2.48E-07

c_000000000027_284	FliG_C	PF01706.16	FliG C-terminal domain	K02410	flagellar motor switch protein FliG	0.56568	0.068456	2.048504	4.20E-06	2.016682	1.64E-06
c_000000000027_285	FliH	PF02108.16	Flagellar assembly protein FliH	K02411	flagellar assembly protein FliH	0.439041	0.346538	2.733732	7.46E-06	2.494705	7.21E-06
c_000000000027_286	ATP-synt_ab	PF00006.25	ATP synthase alpha/beta family, nucleotide-binding domain	K02412	flagellum-specific ATP synthase [EC:7.4.2.8]	1.080432	0.011475	2.503056	7.42E-06	1.944369	4.38E-05
c_000000000027_289	Flg_hook	PF02120.16	Flagellar hook-length control protein FliK	K02414	flagellar hook-length control protein FliK	0.670801	0.266052	3.315358	7.50E-06	2.899997	1.30E-05
c_000000000027_290	FlgD	PF03963.14	Flagellar hook capping protein - N-terminal region	K02389	flagellar basal-body rod modification protein FlgD	1.086199	0.010054	3.215267	3.80E-07	3.403093	5.72E-08
c_000000000027_292	FlbD	PF06289.11	Flagellar and Swarming motility proteins	K02385	flagellar protein FlbD	0.890896	0.065194	3.307522	7.69E-07	3.10157	4.97E-07
c_000000000027_296	FliL	PF03748.14	Flagellar basal body-associated protein FliL	K02415	flagellar protein FliL	0.115962	0.756217	2.395398	4.47E-06	2.38957	1.57E-06
c_000000000027_297	FliM	PF02154.15	Flagellar motor switch protein FliM	K02416	flagellar motor switch protein FliM	0.169009	0.588943	2.523309	4.58E-07	1.854255	3.68E-06
c_000000000027_298	FliMN_C	PF01052.20	Type III flagellar switch regulator (C-ring) FliN C-term	K02417	flagellar motor switch protein FliN	0.254806	0.606479	2.421545	3.35E-05	1.35791	0.003353
c_000000000027_300	Flg_bbr_C	PF06429.13	Flagellar basal body rod FlgEFG protein C-terminal	K02392	flagellar basal-body rod protein FlgG	0.934324	0.000902	2.054419	5.91E-07	0.881193	0.000936
c_000000000027_301	ChapFlgA	PF13144.6	Chaperone for flagella	K02386	flagellar basal body	0.964931	0.000268	1.971917	3.45E-07	1.255797	1.14E-05

			basal body P-ring formation		P-ring formation protein FlgA						
c_00000000 0027_302	FlgH	PF02107.16	Flagellar L-ring protein	K02393	flagellar L-ring protein FlgH	0.555132	0.045969	1.560296	2.69E-05	0.805944	0.004096
c_00000000 0027_303	FlgI	PF02119.16	Flagellar P-ring protein	K02394	flagellar P-ring protein FlgI	0.220245	0.370609	1.798033	2.41E-06	1.235781	3.99E-05
c_00000000 0027_306	Flg_bbr_C	PF06429.13	Flagellar basal body rod FlgEFG protein C-terminal	K02396	flagellar hook-associated protein 1	0.2411	0.51815	2.1435	3.18E-05	1.703946	0.000115

Table S2.1b. List of genes involved in capsule polysaccharide synthesis and their expression level as logFC when growth on polymers compared to growth on D-galactose.

Gene ID	CAZyme (>=2 tools)	agar_logFC	agar_FDR	brown_seaweed_logFC	brown_seaweed_FDR	red_seaweed_logFC	red_seaweed_FDR
c_000000000020_5		-0.75057	6.23E-05	-0.28804	0.045139	-0.11994	0.359286
c_000000000020_6	GT2(4-129)	-0.57068	0.006326	-0.00296	0.989514	0.06986	0.694636
c_000000000020_7		-0.57278	0.000841	0.008394	0.959552	0.397718	0.006604
c_000000000020_8	GT2	-0.30223	0.044103	0.135439	0.356664	0.43512	0.004438
c_000000000020_9		-0.84954	6.75E-05	-0.25892	0.097532	0.239502	0.094076
c_000000000020_10		-0.63913	0.000412	-0.12844	0.379611	0.837436	1.46E-05
c_000000000020_11		-0.37314	0.008922	-0.09165	0.496543	0.3352	0.011719
c_000000000020_12	GT4(185-323)	-0.30993	0.031238	0.050276	0.735661	0.159938	0.21854
c_000000000020_13	GT4(193-340)	-0.90035	7.93E-06	-0.31973	0.021439	-0.01977	0.87919
c_000000000020_14		-1.08606	2.04E-06	-0.53914	0.00115	-0.2239	0.09036
c_000000000020_15	GT4(202-347)	-0.53741	0.000409	-0.25141	0.048884	-0.02272	0.852477
c_000000000020_16	GT4(179-292)	-0.96954	2.29E-06	-0.3243	0.015472	0.010811	0.929599
c_000000000020_17		-0.86892	7.20E-06	-0.45526	0.001977	0.103252	0.380581
c_000000000020_18	GT2(10-181)	-0.32619	0.019318	-0.06465	0.640914	0.43133	0.002531
c_000000000020_19		-0.02219	0.870018	0.391152	0.005413	0.550722	0.000252

c_000000000020_20	GT4(219-356)	-0.32213	0.014126	0.154137	0.212278	0.383361	0.003362
c_000000000020_21		-0.66128	0.000173	-0.15924	0.250433	0.02145	0.875727
c_000000000020_22	GT4(182-324)	-0.57109	0.000383	-0.01167	0.93793	0.373941	0.005988
c_000000000020_23	GT8(3-245)	-0.37758	0.008273	0.258095	0.055841	0.262247	0.040741

Table S2.1c. Significantly up-regulated (FDR < 0.05) glycoside transferases (GTs) in brown seaweed compared to D-galactose.

Gene ID	CAZyme (>=2 tools)	agar_logFC	agar_FDR	brown_seaweed_logFC	brown_seaweed_FDR	red_seaweed_logFC	red_seaweed_FDR
c_000000000027_329	GT2(8-119)	-0.35092	0.018666	0.578125	0.000704	-0.66194	0.000152
c_000000000027_331	GT4(198-349)	0.11413	0.466245	0.939898	2.77E-05	-0.72568	0.000204
c_000000000027_332	GT2(26-153)	0.133174	0.342788	1.175631	1.26E-06	-0.21707	0.111713
c_000000000027_337	GT4(183-338)	0.161783	0.231051	0.779794	4.30E-05	-0.66767	0.000127
c_000000000027_340	GT26(358-529)+GT26(624-796)	0.530475	0.001605	0.693026	0.000237	-0.66331	0.000236
c_000000000027_394	GT28(189-350)	1.022102	1.76E-05	0.959549	4.21E-05	0.600614	0.001229
c_000000000027_689	GT2(4-164)	1.490393	3.41E-05	1.445511	6.17E-05	0.706548	0.009735
c_000000000026_391	GT19(16-369)	0.146035	0.533318	0.692122	0.005612	1.138804	4.70E-05
c_000000000028_239	GT2	0.401697	0.045964	0.60811	0.005259	1.534402	5.70E-07
c_000000000028_243	GT4(222-372)	-0.40154	0.031866	0.482916	0.010907	1.406018	5.98E-07
c_000000000030_53	GT2(8-172)	1.407298	8.42E-07	1.511476	6.82E-07	1.720193	5.82E-08
c_000000000030_211	GT89(22-301)	0.370392	0.109655	0.653608	0.008682	0.106756	0.636067

c_000000000030_292	GT4(179-323)	-0.30394	0.113864	0.987568	8.07E-05	1.57891	3.06E-07