

---

# Metagenomic Insights Reveal the Microbial Diversity and Associated Algal Polysaccharide-Degrading Enzymes on the Surface of Red Algae among Remote Regions

---

Xiaoqian Gu , Zhe Cao , Luying Zhao , Dewi Seswita-Zilda , Qian Zhang , Liping Fu , [Jiang Li](#) \*

Posted Date: 30 May 2023

doi: 10.20944/preprints202305.2067.v1

Keywords: Metagenomics; red algae; microbial diversity; algal-polysaccharide-degrading enzymes



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

# Metagenomic Insights Reveal the Microbial Diversity and Associated Algal Polysaccharide-Degrading Enzymes on the Surface of Red Algae among Remote Regions

Xiaoqian Gu <sup>1,2</sup>, Zhe Cao <sup>1</sup>, Luying Zhao <sup>1</sup>, Dewi Seswita-Zilda <sup>3</sup>, Qian Zhang <sup>1</sup>, Liping Fu <sup>1</sup> and Jiang Li <sup>1,\*</sup>

<sup>1</sup> Key Lab of Ecological Environment Science and Technology, First Institute of Oceanography, Ministry of Natural Resources, Qingdao, 266061, China;

<sup>2</sup> CAS and Shandong Province Key Laboratory of Experimental Marine Biology, Center for Ocean Mega-Science, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, 266071, China

<sup>3</sup> Research Center for Deep Sea, Earth Sciences and Maritime Research Organization, National Research and Innovation Agency (BRIN), Jl. Pasir Putih Raya, Pademangan, North Jakarta City, Jakarta, 14430, Indonesia

\* Correspondence: lijiang@fio.org.cn

**Abstract:** Macroalgae and macroalgae-associated bacteria together constitute the most efficient metabolic cycling system in the ocean. Their interactions, especially the responses of macroalgae-associated bacteria communities to algae in different geographical locations, is mostly unknown. In this study, metagenomics was used to analyze the microbial diversity and associated algal polysaccharide-degrading enzymes on the surface of red algae among three remote regions. There were significant differences in the macroalgae-associated bacteria community composition and diversity among the different regions. At the phylum level, Proteobacteria, Bacteroidetes, and Actinobacteria had a significantly high relative abundance among the regions. From the perspective of species diversity, samples from China had the highest macroalgae-associated bacteria diversity, followed by those from Antarctica and Indonesia. In addition, in the functional prediction of the bacterial community, genes associated with amino acid metabolism, carbohydrate metabolism, energy metabolism, metabolism of cofactors and vitamins, and membrane transport had a high relative abundance. Canonical correspondence analysis and redundancy analysis of environmental factors showed that, without considering algae species and composition, pH and temperature were the main environmental factors affecting bacterial community structure. Furthermore, there were significant differences in algal polysaccharide-degrading enzymes among the regions. Samples from China and Antarctica had high abundances of algal polysaccharide-degrading enzymes, while those from Indonesia had extremely low abundances. The environmental differences between these three regions may impose a strong geographic differentiation regarding the biodiversity of algal microbiomes and their expressed enzyme genes. This work expands our knowledge of algal microbial ecology, and contributes to an in-depth study of their metabolic characteristics, ecological functions, and applications.

**Keywords:** metagenomics; red algae; microbial diversity; algal-polysaccharide-degrading enzymes

## 1. Introduction

Marine algae produce approximately a half of the global primary productivity, and the algal polysaccharides produced by these algae, together with polysaccharide-degrading heterotrophic microorganisms and their polysaccharide-degrading enzymes, form the largest and most dynamic carbon cycle on Earth [1, 2]. However, little is known about this microbial-driven system of marine carbon metabolism.

Marine microorganisms play important roles throughout all marine ecological processes; therefore, an increasing amount of attention is being paid to the study of their populations and functions [3,4,5]. However, the microbial communities on the surface of algae are still not fully

understood, despite their high biodiversity; furthermore, they are markedly different from those that live freely in seawater [3]. Microorganisms on the surface of algae, through a series of complex interactions with their hosts, constitute a rich source of bioactive compounds and specific polysaccharide active enzymes [6]. It is generally believed that the secondary metabolites of microorganisms depend on a range of available organic carbon sources produced by the host algae [7]. Red macroalgae are the source of nutrition for many marine microorganisms. Carrageenan and agar are types of sulphated galactose, which are the main cell wall components of red algae [8]. The study of macroalgae-associated bacterial communities and their relationship with algal hosts will provide important information on the ecological function of algae in aquatic ecosystems [8]. In addition, because these microbes interact with algae in a number of complex ways, together they form an important source of novel bioactive compounds such as agarase and carrageenase, with potential applications [6]. Agarose and carrageen, and their oligosaccharides, are widely used in food, medicine, agriculture, cosmetics, and other fields, and have huge application potential and market prospects [9, 10].

Algae host a large number of bacteria on their surface (up to  $10^6$  cells·cm<sup>-2</sup>), which varies greatly according to species location and climatic conditions [4, 6, 11,12]. In addition, bacteria can become endosymbionts in algal cells and freely grow in the “phycosphere”. These microorganisms have a very complex relationship with algae [13, 14]. Attached bacteria play a crucial role in the normal growth and development of algae [15]. However, some bacteria can also become pathogenic [16]. Due to the different thallus components, macroalgae can also regulate the attached bacteria in various ways, and the community structure of macroalgae-associated bacteria on the surface of macroalgae also differs [17]. In addition, seasonal variations, spatial differences, and environmental factors can also affect the composition of macroalgae-associated bacteria [18, 19]. The microbial community composition and diversity on the same algal species from different sea areas can be significantly different. The functional selection of different seaweed species from the same sea area is also closely related to their habitat, and there are large differences in the dominant species and diversity [20, 21]. The community composition of macroalgae-associated bacteria is not only related to the environment, but is also affected by the host, which is in accordance with natural selection.

In recent years, an increasing amount of research has investigated macroalgae-associated bacteria and algal polysaccharide-degrading enzymes [22, 23]. Numerous macroalgae-associated marine microorganisms contain undeveloped enzymes such as glycoside hydrolase (GH) and polysaccharide lyase (PL) [24], which are involved in important metabolic pathways and their characterization has the potential to provide new insights into the marine carbon cycle.

The application of metagenomics will significantly promote the study of marine microbial diversity and biological function, and greatly increase the opportunities for discovery of potential new enzymes [25, 26]. Metagenomics not only provides more information for understanding the microbial diversity that flourishes in marine systems, but can also further reveal the taxonomic composition and functional potential of microbial communities, and provide genomic information for understanding the structure and biogeochemical characteristics of macroalgae ecosystems [6, 27].

This study reports on the metagenomic sequencing and analysis of red algae samples of a macroalgae-associated bacteria system to explore microbial diversity, functional diversity, and the number of marine polysaccharide-degrading enzyme genes. The findings will effectively extend our understanding of the diversity of algal surface microorganisms and their polysaccharide-degrading enzymes. They will also deepen our understanding of how microorganisms process the hundreds of millions of tons of polysaccharides produced by algae each year and their role in the marine carbon cycling system.

## 2. Material and Methods

### 2.1. Macroalgae Sampling and Characterization

Three macroalgae samples (NJDZ01, NJDZ02, and NJDZ03) were collected from the Great Wall Station in Antarctica (Lat: 62° 12', Long: 58° 57') during China's 34<sup>th</sup> Antarctic expedition (Gui et al.,

2019), another three (YNDZ01, YNDZ02, and YNDZ03) from Halmahera in Indonesia (Lat: 0° 36', Long: 127° 52'), and a final three (WHDZ01, WHDZ02, and WHDZ03) from Weihai in China (Lat: 37° 25', Long: 122° 17'). The nine macroalgae samples, identified as *Palmaria decipiens* (NJDZ01), *Curdiea racovitzae* (NJDZ02), *Iridaea cordata* (NJDZ03), *Laurencia japonica* (YNDZ01), Coralline algae (YNDZ02), *Porphyra* (YNDZ03), *Grateloupia filicina* (WHDZ01), *Chondrus ocellatus* (WHDZ02), and *Hyalosiphonia caespitosa* (WHDZ03), were collected from the seashore. The algal samples were washed with sterile seawater and brought back to the laboratory for collection of surface-associated bacteria.

Samples were then washed again with sterile seawater in the laboratory. After washing, sterile scissors were used to cut the algae into smaller chunks. They were then placed in a sterile tube with an appropriate amount of sterilized seawater and vortexed three times for 40 s each. The liquid was filtered onto 0.22  $\mu\text{m}$  filter membranes (Merck Millipore, Darmstadt, Germany) to collect the associated bacteria.

## 2.2. DNA Extraction and Sequencing

The FastDNA Spin Kit for Soil (MP Bio, USA) was used to extract genomic DNA from the macroalgae-associated bacteria. After genomic DNA extraction was completed, the extracted genomic DNA purity and integrity were analyzed by agarose gel electrophoresis. Qualified DNA samples were detected, library construction and library detection were performed, and qualified libraries were sequenced by Illumina PE150 (Novogene, Beijing, China).

The original data obtained by sequencing was the image file. After base calling, the result file stored in FASTQ format was obtained, called a raw read. Since the original dismounting data may contain data with joints (introduced during database construction) and a certain proportion of low-quality data (generated by sequencing reading), in order to ensure the accuracy and reliability of subsequent analysis results, Trim galore was used to preprocess raw data and obtain clean data for subsequent analysis. In the pre-processing, the junction sequence in sequencing reads, sequences with average mass value < 25, truncated sequences with a length < 70 bp, and host sequence contaminations were removed. Due to the impact of the high-throughput sequencing error rate on the results, it was necessary to evaluate the quality of the optimized data.

## 2.3. Metagenome Assembly

After pretreatment, clean data were obtained, and assembly analysis was performed using MEGAHIT assembly software (<http://www.l3-bioinfo.com/products/megahit.html>). The assembled scaffolds were broken at the N join to produce N-free sequence fragments called scaftigs. Clean data of samples after quality control were compared to scaftigs after sample assembly by Bowtie2 software (<https://bowtie-bio.sourceforge.net/bowtie2/index.shtml>), and unused pair-end reads were obtained. The unused reads of each sample were combined together for mixed assembly with the same assembly parameters as those of a single sample. The scaffolds of the hybrid assembly were broken at the N junction to obtain N-free scaftigs sequences. For scaftigs generated by single sample and mixed assembly, fragments < 500 bp were filtered out, and statistical analysis and subsequent gene prediction were performed.

## 2.4. Gene Prediction and Abundance Analysis

MetaGeneMark ([http://topaz.gatech.edu/GeneMark/meta\\_gmhmp.cgi](http://topaz.gatech.edu/GeneMark/meta_gmhmp.cgi)) was utilized to predict the open reading frame (ORF) of each sample and mixed assembled scaftig ( $\geq 500$  bp). Based on the predicted result, information < 100 nt was filtered out. For ORF prediction results of each sample and mixed assembly, CD-HIT software (<http://www.bioinformatics.org/cd-hit/>) was used to remove redundancy in order to obtain a non-redundant initial gene catalogue. By default, a 95% identify and 90% coverage were used for clustering, and the longest sequence was selected as the representative sequence. Clean data of each sample was compared to the initial gene catalogue with Bowtie2, and the number of gene reads in each sample was calculated. The number of reads  $\leq 2$  in each sample

were filtered out, and the final gene catalogue (unigenes) for subsequent analysis was obtained. Based on the number of reads and gene length, the abundance of each gene in each sample was calculated.

### 2.5. Taxonomy Prediction

DIAMOND software (Version 0.9.9, <https://github.com/bbuchfink/diamond/>) compared unigenes with bacteria, fungi, archaea, and viruses extracted from the National Center for Biotechnology Information non-redundant (NR) database (Version 2018.01). For the comparison results of each sequence,  $e$  values  $\leq$  the smallest  $e$  value\* 10 were selected for subsequent analysis. After filtering, since each sequence may have multiple comparison results, multiple different species classification information could be obtained. In order to ensure its biological significance, the lowest common ancestor algorithm (systematic classification applied to MEGAN software) was adopted to apply the classification level before the first branch as the species annotation information of the sequence. Based on the results of lowest common ancestor annotation and gene abundance table, the abundance of each sample in each taxonomic level (kingdom, phylum, class, order, family, genus, and species) was obtained.

### 2.6. Functional Annotations

DIAMOND software (Version 0.9.9, <https://github.com/bbuchfink/diamond/>) was used to compare unigenes with various functional databases. For the comparison results of each sequence, the results with the highest score were selected for subsequent analysis. Based on the comparison results, the relative abundance of different functional levels was calculated (the relative abundance of each functional level was equal to the sum of the relative abundance of the genes annotated as this functional level). The Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Version 2018-01-01, <http://www.kegg.jp/kegg/>) was divided into six levels, the eggNOG database (Version 4.5, <http://eggnogdb.embl.de/#/app/home>) into three levels, and the CAZy database (Version 201801, <http://www.cazy.org/>) into three levels. Based on the results of functional annotation and gene abundance table, the gene number table of each sample at each classification level was obtained.

## 3. Results and Discussion

### 3.1. Characteristics of the Nine Macroalgae Samples

The physicochemical parameters of the seawater overlying the red algae among the three remote regions differed, with the following ranges (Table 1): water temperature  $-1.34$ – $27.9$  °C, salinity  $32.4\%$ – $34.7\%$ , pH  $8.01$ – $8.79$ , dissolved oxygen (DO)  $8.07$ – $8.79$  mg/L, latitude  $0.6^{\circ}$ – $62.2^{\circ}$ , and longitude  $58.95^{\circ}$ – $127.86^{\circ}$ . These physicochemical parameters of water samples can serve as an imperfect proxy for the properties of macroalgae samples, due to wave and current distributions.

**Table 1.** The physicochemical parameters of overlying seawater among three remote space.

Location	Lat	Long	DO	pH	Salinity	Temperature
Great Wall Station	$62.2^{\circ}$	$58.95^{\circ}$	10.836 mg/L	8.01	34.07‰	$-1.34^{\circ}\text{C}$
Halmahera	$0.6^{\circ}$	$127.86^{\circ}$	5.44 mg/L	8.79	34.7‰	$27.9^{\circ}\text{C}$
Weihai	$37.42^{\circ}$	$122.28^{\circ}$	6.65 mg/L	8.43	32.4‰	$8^{\circ}\text{C}$

### 3.2. Sequencing and Metagenomic Assembly

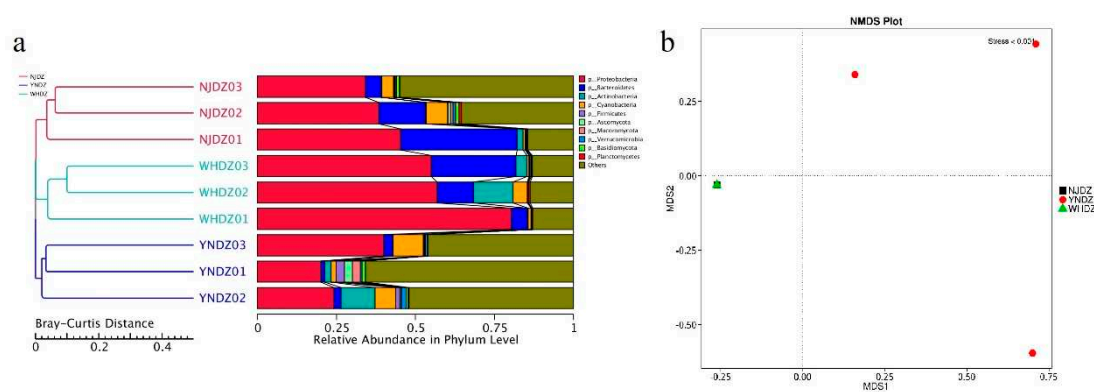
The Illumina HiSeq sequencing platform produced a total of 81,611.46 Mbp of raw data (average data volume was 9,067.94 Mbp). After quality control, 81,245.68 Mbp of clean data were obtained (average data volume was 9,027.30 Mbp), and the effective data rate of the quality control was 99.55% (Table S1). After single sample assembly and mixed assembly, a total of 2,155,261,173 bp scaffolds were obtained with an average length of 1,178.16 bp, maximum length of 514,363 bp, N50 length of 1,230.89 bp, and N90 length of 573.11 bp. Scaffolds were interrupted at N to generate scaffolds, yielding a total

of 2,155,261,173 bp scaffigs with an average length of 1,178 bp; 1,231 bp for N50 and 573 bp for N90 (Table S2).

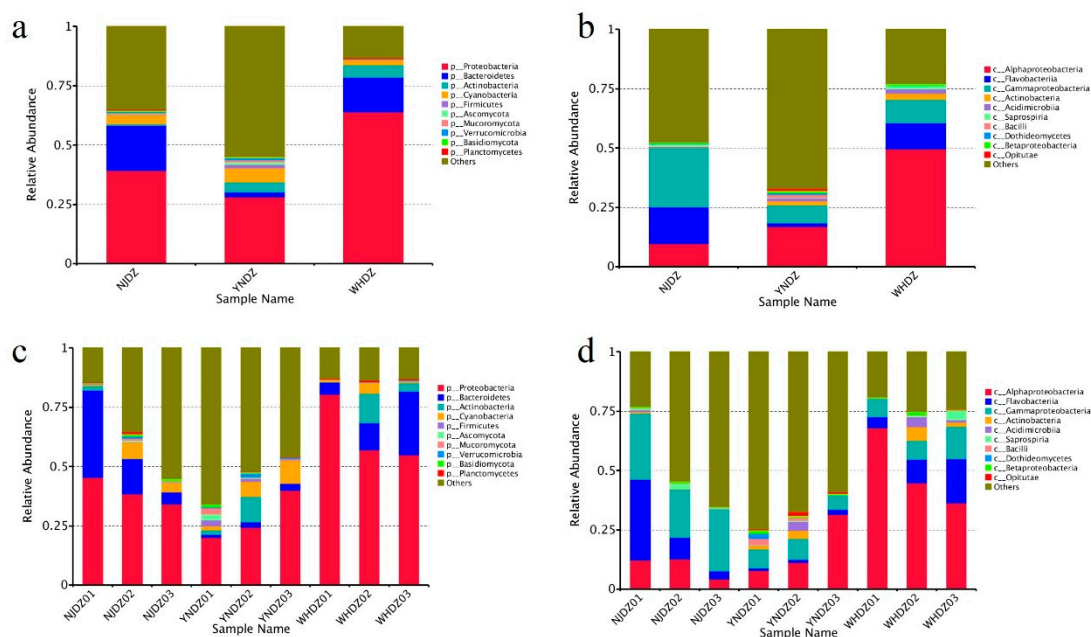
### 3.3. Microbial Diversity and Community Composition

There were 2,485,201 prediction genes after the original redundancy removal, among which the number of ORFs annotated in the NR database was 1,781,793 (71.70%), and 83.52% of ORFs annotated in NR database reached the boundary level. The proportions were 79.82%, 74.49%, 69.68%, 66.09%, 54.98%, and 44.15% at the phylum, class, order, family, genus, and species levels, respectively.

Among all seawater microorganisms, the community structure of macroalgae-associated bacteria has certain specificity, and the two can be bidirectionally selected through interactions [28]. Therefore, studying the macroalgae-associated bacteria community is a necessary prerequisite to understanding the relationship between algae and bacteria. At the phylum level, Proteobacteria, Bacteroidetes, and Actinobacteria had a significantly high relative abundance in NJDZ, YNDZ, and WHDZ samples (Figure 1a, Figure 2a). However, Proteobacteria, Bacteroidetes, and Actinobacteria had a significantly higher relative abundance in WHDZ samples than in NJDZ and YNDZ samples. Among them, the maximum abundances of Proteobacteria, Bacteroidetes, and Actinobacteria were recorded in WHDZ01, NJDZ01, and WHDZ02, respectively (Figure 2c).



**Figure 1.** Relative abundance of species at phylum level. **a:** Sample clustering analysis, on the left is Bray-Curtis distance clustering tree structure; on the right is the relative abundance distribution of each sample at phylum level. **b:** NMDS analysis, each point in the figure represents a sample, the distance between points represents the degree of difference, and samples in the same group are represented by the same color; the stress is less than 0.2 indicating that the NMDS analysis has a certain reliability.



**Figure 2.** Relative species abundance at phylum and class levels **a** and **c**: relative species abundance at phylum levels; **b** and **d**: relative species abundance at class levels.

Proteobacteria is the most dominant seawater phylum [29]. Bacteroidetes can degrade some biological macromolecules such as chitin and cellulose [30]. Actinobacteria can degrade organic pollutants and play an important role in marine pollution remediation [31, 32]. In addition, some phylum with a low abundance were also recorded, such as Cyanobacteria and Firmicutes, among which Cyanobacteria had a high abundance in YNDZ02 and YNDZ03, while Firmicutes had a high abundance in YNDZ01 and YNDZ02. Vieira et al. pointed out that Firmicutes may be related to sea pollution; therefore, sea areas with high Firmicutes abundance may indicate pollution [32]. Non-metric multidimensional scaling showed that the microbial communities of the NJDZ and WHDZ samples had a similar species composition, and the microbial community of YNDZ samples had a different species composition (Figure 1b). Environmental and ecological selection may play a role in generating and maintaining microbial diversity in a geographically defined and seemingly unstructured marine ecosystem. Algal microbiomes are complex and dynamic, and their diversity may be driven by ecological or environmental selection to generate and maintain these intimate relationships over space and evolutionary time [33, 34].

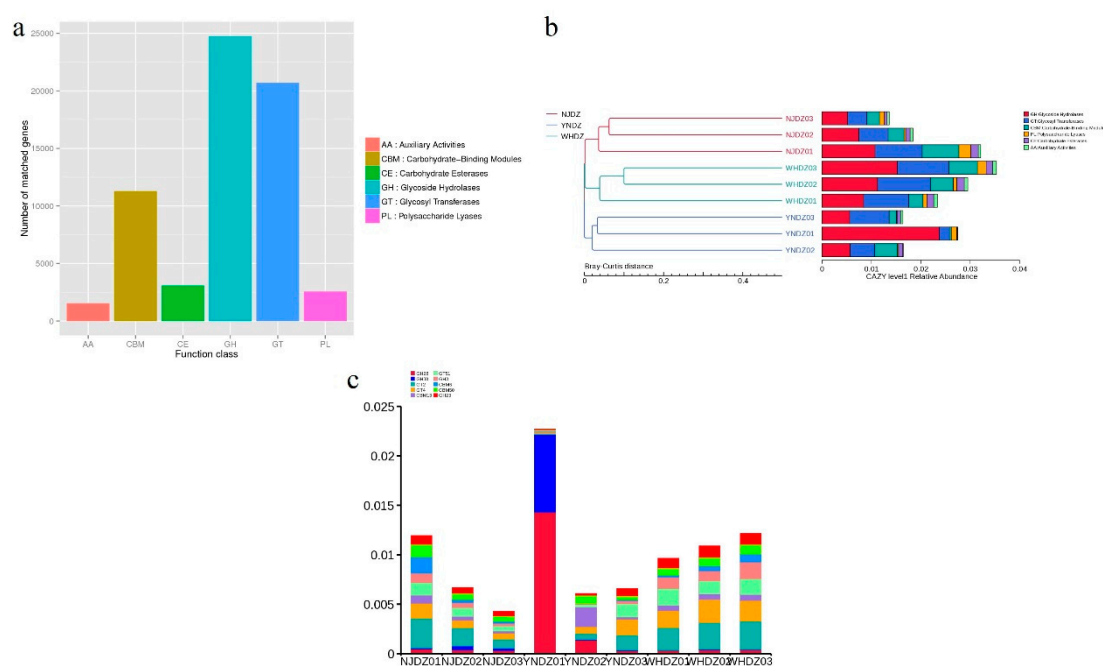
The most abundant microbial class among the nine macroalgae samples also significantly differed according to region. The dominant macroalgae-associated bacteria of NJDZ samples were mostly Gammaproteobacteria, Flavobacteriia, Alphaproteobacteria, and Saprospiria; the dominant macroalgae-associated bacteria of YNDZ samples were mostly Alphaproteobacteria, Gammaproteobacteria, Actinobacteria, Flavobacteriia, Acidimicrobiia, and Bacilli; and the dominant macroalgae-associated bacteria of WHDZ samples were mostly Alphaproteobacteria, Flavobacteriia, Gammaproteobacteria, Actinobacteria, and Acidimicrobiia. It is worth noting that Gammaproteobacteria was the dominant class in NJDZ samples, while Alphaproteobacteria was the dominant class in WHDZ and YNDZ samples (Figure 2b, Figure 2d). These results indicate that the community composition of the macroalgae-associated bacteria is not only related to the environment, but also affected by the host, in accordance with natural selection.

### 3.4. Functional Prediction of Bacterial Communities

DIAMOND software was used to annotate the database of common functions of non-redundant gene sets ( $e\text{-value} \leq 10^{-5}$ ); there were 2,485,201 prediction genes after the original redundancy was removed, and 1,493,359 (60.09%) genes could be compared with the KEGG database, among which,



be obtained. Comparing the various strains of the six families to the CAZy database (Figure 4a) revealed 2,151,233 carbohydrate active enzyme genes, among which the GH family had the most genes (24,756), followed by the family of glycosyltransferases (GTs) genes (20,694), carbohydrate-binding modules (CBMs) family genes (11,291), carbohydrate esterases family genes (3,079), PL family genes (2,564), and auxiliary activities family genes (1,539). The strains with more GH family genes also had more GT and CBM family genes (Figure 4b), which might be related to the functional correlation of enzymes in each family. The most abundant families predicted in this genome were GH28, GH38, GT2, GT4, CBM13, GT51, GH3, CBM6, CBM50, and GH23 from level 2 of the CAZymes database (Figure 4c). Further analysis showed that the strains with a high carbohydrate active enzyme gene richness were mainly from Pseudoalteromonadaceae of Proteobacteria and Flavobacteriaceae of Bacteroidetes.



**Figure 4.** Annotation of CAZy enzyme. **a:** notes on CAZy Unigenes; **b:** CAZy level 1 relative abundance; **c:** CAZy level 2 relative abundance.

Marine algae are the most promising raw material for the replacement of land plants, and algal oligosaccharides are degraded and utilized by marine heterotrophic microorganisms [35]. Many enzymes in algae-related microorganisms, such as GH and PL, are involved in this important metabolic process [6, 35]. The GH family can hydrolyze the glycosidic bonds between two or more carbohydrates or between carbohydrate and non-carbohydrate components, and are essential in the red algae polysaccharide hydrolyzing process [36]. The GT family is involved in the biosynthesis of disaccharides, oligosaccharides, and polysaccharides by catalyzing the transfer of glycosylates from activated donor molecules to specific receptors to form glycosidic bonds. The most widely distributed GT2 family proteins have the activities of cellulose synthetase, chitin synthetase, hyaluronic acid synthetase, glucan synthetase, and mannan synthetase [37]. These enzymes are related to the synthesis of algae polysaccharides, indicating that these macroalgae-associated bacteria can not only degrade algae polysaccharides, but also synthesize related polysaccharides themselves.

The main components of red algae are agar and carrageen. Research on carrageenase is the earliest and most extensive;  $\kappa$ -carrageenase belongs to the GH16 family and  $\tau$ -carrageenase belongs to the GH82 family.  $\beta$ -Agarase belongs to the GH16, GH39, GH50, GH86, and GH118 families, while  $\alpha$ -agarase belongs to the GH96 and GH117 families [38].  $\beta$ -agarase in the GH16 and GH86 families is an endonuclease,  $\beta$ -agarase in the GH50 family has exonuclease activity, and  $\alpha$ -agarase in the GH117 family is an exonuclease [39]. Therefore, when macroalgae-associated bacteria communities contain

a variety of GH family enzyme coding genes, there is an opportunity for the emergence of endo-cut  $\beta$ -agarase, exo-cut  $\beta$ -agarase, and exo-cut  $\alpha$ -agarase systems, which eventually degrade polysaccharides into oligosaccharides.

Annotation of the functional genes of the macroalgae-associated bacteria genome revealed that the bacteria genome isolated from the surface of red algae contained a large number of genes related to the degradation of alginate. This indicated that most of the strains on the surface of algae had the basic function of degrading alginate, which was related to their living environment and their own carbon source utilization function [40, 41]. Therefore, the diversity of these functional genes was further analyzed, revealing the presence of the PL7 family in strains distributed in Proteobacteria and Bacteroidetes. These results indicate that the microorganisms on the surface of algae are rich in alginate lyase, and that algae degradation requires the coordination of enzymes with multiple functions. These enzymes work together to utilize algae polysaccharides.

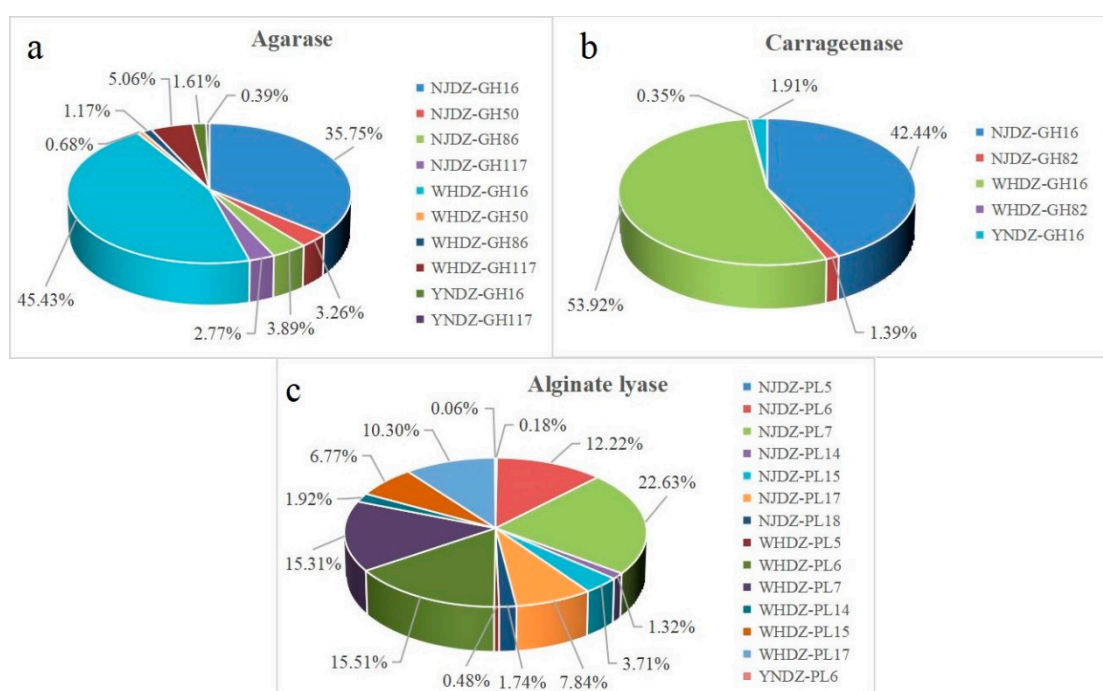
### 3.6. Mining of Enzymes for Algal Polysaccharide Degradation

Increasing attention has been paid to marine oligosaccharides due to their bioactivity, solubility, and bioavailability. Studies have shown that algae oligosaccharides have many potential applications [42]. The in-depth functional analysis of nine macroalgae metagenomic data was performed to identify new algal polysaccharide-degrading enzymes for biotechnological purposes (Table 2 and Figure 5). Metagenomic data were then further analyzed to explore the number of algal polysaccharide-degrading enzyme genes. There were significant differences in algal polysaccharide-degrading enzymes on the surface of red algae among the remote regions. The number of alginate lyase, agarase, and carrageenase genes on the surface of red algae among the regions is shown in Table 3. The number of agarase genes from WHDZ was 1,076, and that from NJDZ was 939, while YNDZ had only 41. The number of carrageenase genes from WHDZ was 940, and that from NJDZ was 759, with YNDZ only having 33. The number of alginate lyase genes from WHDZ was 840, and that from NJDZ was 829, while YNDZ had only one. The agarase gene was mainly distributed in WHDZ03, followed by NJDZ01. There were very few agarase genes in YNDZ, with only three alginate lyase genes in YNDZ01 and 38 in YNDZ03. The carrageenase gene was mainly distributed in WHDZ03, followed by NJDZ01. Similarly, there were very few carrageenase genes in YNDZ, with only three carrageenase lyase genes in YNDZ01 and 30 in YNDZ03. The alginate lyase gene was mainly distributed in NJDZ01, followed by WHDZ03. There was also very few alginate lyase genes in YNDZ, with only one in YNDZ03.

**Table 2.** The predicted proteins containing CAZymes modules acting on algal-polysaccharide.

Locations	Enzymes Number	E.C. Number	CAZy Family
Agarase			
NJDZ	939	EC 3.2.1.81	GH16
		EC 3.2.1.81	GH50
		EC 3.2.1.81	GH86
		EC 3.2.1.-	GH117
WHDZ	1076	EC 3.2.1.81	GH16
		EC 3.2.1.81	GH50
		EC 3.2.1.81	GH86
YNDZ	41	EC 3.2.1.-	GH117
		EC 3.2.1.81	GH16
		EC 3.2.1.-	GH117
Carrageenase			
NJDZ	759	EC 3.2.1.83	GH16

		EC 3.2.1.157	GH82
WHDZ	940	EC 3.2.1.83	GH16
		EC 3.2.1.157	GH82
YNDZ	33	EC 3.2.1.83	GH16
Alginate lyase			
		EC 4.2.2.3	PL5
		EC 4.2.2.3	PL6
		EC 4.2.2.-	PL7
NJDZ	829	EC 4.2.2.3	PL14
		EC 4.2.2.3	PL15
		EC 4.2.2.3	PL17
		EC 4.2.2.3	PL18
		EC 4.2.2.3	PL5
		EC 4.2.2.3	PL6
HDZ	840	EC 4.2.2.-	PL7
		EC 4.2.2.3	PL14
		EC 4.2.2.3	PL15
		EC 4.2.2.3	PL17
YNDZ	1	EC 4.2.2.3	PL6



**Figure 5.** Relative abundances of agarase (a), carrageenase (b) and alginate lyase (c) in CAZy family from the metagenomic sample.

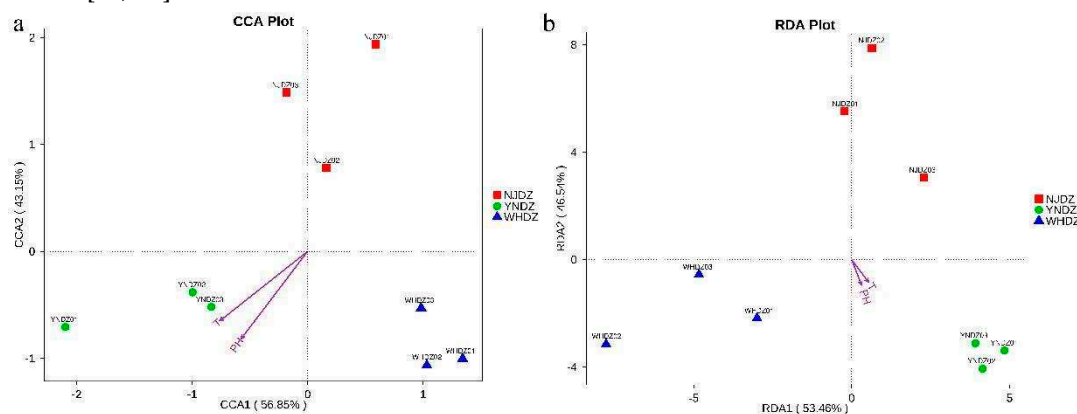
These results suggest that WHDZ and NJDZ had a high abundance of algal polysaccharide-degrading enzymes, while YNDZ had an extremely low abundance. There is evidence that the characteristics of the algae itself, secretions at different growth stages (small organic molecules such as early amino acids and organic acids, and large organic molecules such as algal polysaccharides and lipids), and environmental conditions for algal growth (such as pH, water flow, light,

temperature, and nutrients) affect the composition of the macroalgae-associated bacterial community, and thus enzyme abundance [44, 45, 46]. Of note, there were significant differences in the algal polysaccharide-degrading enzymes on the surface of red algae among the remote regions, especially in the NJDZ samples. The environmental differences between NJDZ, WHDZ, and YNDZ may impose a strong geographic differentiation in the biodiversity of algal microbiomes and their expressed enzyme genes.

Numerous macroalgae-associated marine microorganisms contain undeveloped enzymes such as GH and PL, which are involved in important metabolic pathways and whose characterization has the potential to provide new insights into the marine carbon cycle [24]. This study effectively extends our understanding of the diversity of algae surface microorganisms and their polysaccharide-degrading enzymes. It also deepens our understanding of how microorganisms process the hundreds of millions of tons of polysaccharides produced by algae each year and their role in the marine carbon cycling system.

### 3.7. Environmental Factor Analysis

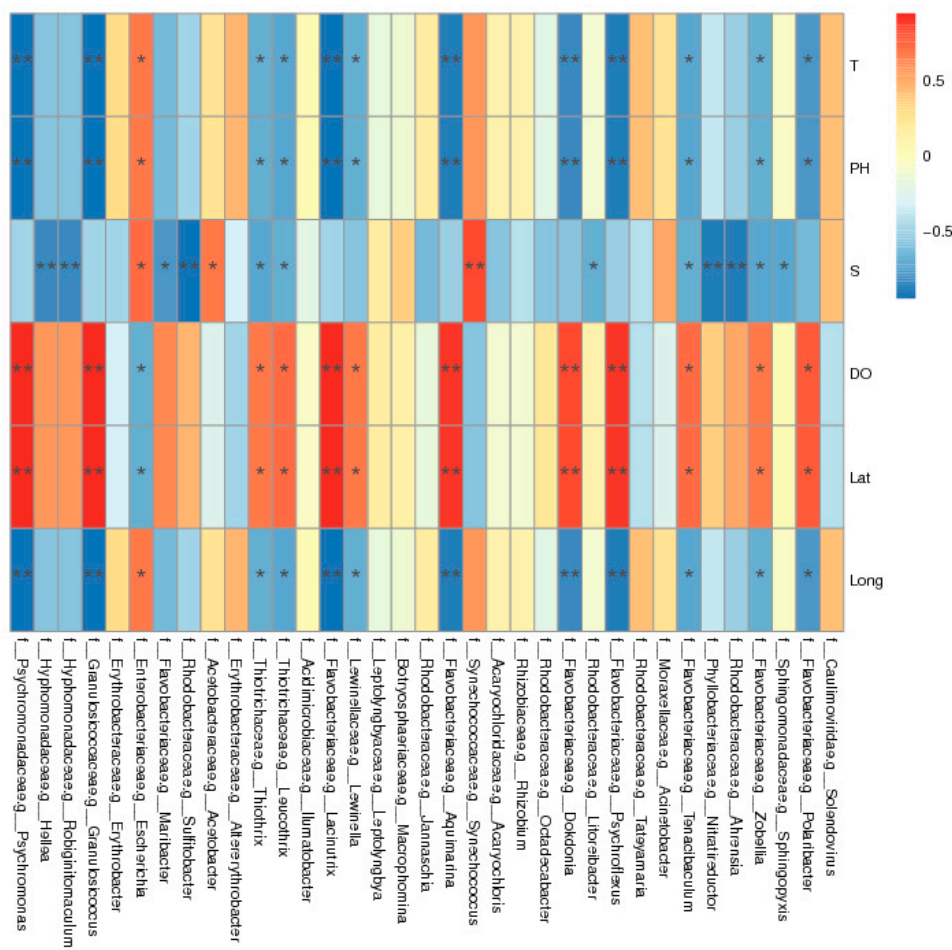
Canonical correspondence analysis (CCA) and redundancy analysis (RDA) analyses were performed using R software (Version 2.15.3), and relationships between bacterial communities and environmental variables were constructed at the genus level. As shown in Figure 6a and Figure 6b, the first and second spindles of the CCA accounted for 56.85% and 43.15% of the variance of the relative abundance of the bacterial community, respectively, and the first and second spindles of RDA accounted for 53.46% and 46.54%, respectively. This indicated that the relationship between the bacterial community and environmental variables was reliable. The CCA and RDA analyses of environmental factors showed that, without considering the species and composition of algae, pH and temperature were the main environmental factors affecting bacterial community structure. Most previous studies have compared native species diversity (alpha diversity) at different latitudes [46]. However, recent studies have shown that temperature is also one of the most important variables explaining the differences in local community species composition over large-scale latitudinal gradients [34, 47].



**Figure 6.** CCA (a) and RDA (b) of bacterial and environmental factors at the genus level.

Salinity, temperature, pH, and DO are the main environmental factors affecting macroalgae-associated bacteria communities. They not only affect the community structure and distribution of bacteria, but also can be used to indicate the environmental status of marine ecosystems [44, 45]. In order to further explore the influence of environmental factors on the macroalgae-associated bacteria communities, a correlation heat map analysis was conducted, and the results are shown in Figure 7. *Limnothrix* and *Leptolyngbya* were positively correlated with temperature and pH, but negatively correlated with DO. *Psychroflexus* and *Psychromonas* were positively correlated with DO, but negatively correlated with temperature and pH. In addition, *Maribacter*, *Nitratireductors*, and *Robiginitomaculum* were negatively correlated with salinity, and *Acinetobacter* was positively

correlated with salinity. This showed that environmental factors had varying effects on different macroalgae-associated bacteria.



**Figure 7.** Correlation analysis between macroalgae-associated bacteria and environmental factors.

### 3.8. Diverse Bacterial Lineages Potentially Harbor Antibiotic Resistance Genes

This study further screened the diversity and abundance of the antibiotic resistance genes (ARGs) in the microbial gene catalog. The ARG-like genes were assigned to 20 ARG classes, with high proportions being unclassified (Figure S2a). The total numbers of ARG classes in YNZD02 were surprisingly higher than in other samples, suggesting that human activities may also lead to ARG enrichment [48, 49]. YNZD had a higher relative abundance of ARG classes compared with NJDZ and YNDZ (Figure S2b). These findings suggest that macroalgae are a neglected but potentially important reservoir of ARGs. The main ARG classes detected in YNZD02 were *TEM-171* and *tetG* (tetracycline resistance) genes (Figure S2a). These findings suggest that macroalgae in Halmahera, Indonesia are possibly contaminated by anthropogenic ARGs to a certain extent. The diversity of ARG hosts with multiple antibiotic potential resistance suggests that these bacteria are ARG hosts and may play a key role in the acquisition and spread of antibiotic resistance in macroalgae.

ARGs are present in almost all environments. They are either endemic to the natural environment or derived from human-dominated ecosystems [50, 51]. However, the diversity and hosts of ARGs in macroalgae remain unclear. Their abundance of microbes indicates that macroalgae could be a potential reservoir for ARGs. If these microbes that are resistant to various antibiotics are algal pathogens, this could worsen an outbreak of macroalgal disease. In addition, macroalgal ecosystems may be contaminated by antibiotics and ARGs from human and agricultural wastes [52]. Residues of antibiotics in aquatic systems can be deposited in macroalgae, and may eventually

threaten their growth and affect microbial ecology [53]. Therefore, studying ARGs is an important component of evaluating macroalgae ecosystem health.

#### 4. Conclusions

Among the abundance of living organisms in the ocean, algae are the main source of carbon and carry out photosynthesis for microorganisms. Marine algae contain rich polysaccharide substances, and their polysaccharide degradation products are widely used in medicine, food, and cosmetics industries. Therefore, it is very important to understand the degradation strains and enzymes of algae-related microbes. The results of this study showed that the microbial community composition was closely related to the living environment of the algae, with obvious regional differences, and was greatly affected by temperature and DO. By analyzing the genomic data of isolated strains on red algae surfaces, the diversity of related bacteria and polysaccharide-degrading enzyme systems were discussed. These findings broaden the research into marine polysaccharide-degrading enzymes, and provide reference for the subsequent development of macroalgae polysaccharide-degrading bacteria and enzymes.

**Author Contributions:** JL collected the samples, designed research, supervised the project, and analyzed data. XG conducted experimental work and drafted the manuscript. ZC, LZ, DS-Z, QZ and LF assisted with experiments. All authors contributed to the article and approved the submitted version.

**Funding:** This research was financially supported by Impact and Response of Antarctic Seas to Climate Change (RFSOCC2023-2025)

**Data Availability Statement:** The dataset of the metagenome supporting the conclusions of this article is available in the NCBI repository: <http://www.ncbi.nlm.nih.gov/bioproject/974727> (the BioProject ID PRJNA974727).

**Acknowledgments:** We are grateful to all the laboratory members for their technical advice and helpful suggestions. We thank International Science Editing (<http://www.internationalscienceediting.com>) for editing this manuscript.

**Conflicts of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

#### References

1. Field: C.B.; Behrenfeld, M.J.; Randerson, J.T.; Falkowski, P. Primary production of the biosphere: integrating terrestrial and oceanic components. *Science*. 1998, 281, 237-240.
2. Hedges, J.I.; Baldock, J.A.; Ge'linas, Y.; Lee, C.; Peterson, M.; Wakeham, S.G. Evidence for non-selective preservation of organic matter in sinking marine particles. *Nature*. 2001, 409, 801-804.
3. Martin, M.; Portetelle, D.; Michel, G.; Vandenbol, M. Microorganisms living on macroalgae: diversity, interactions, and biotechnological applications. *Appl. Microbiol. Biot.* 2014, 98, 2917-2935.
4. Cundell, A.M.; Sleeter, T.D.; Mitchell, R. Microbial populations associated with the surface of the brown alga *Ascophyllum nodosum*. *Microb. Ecol.* 1997, 4, 81-91.
5. Martin, M.; Portetelle, D.; Michel, G.; Vandenbol, M. Microorganisms living on macroalgae: diversity, interactions, and biotechnological applications. *Appl. Microbiol. Biot.* 2014, 98, 2917-2935.
6. Gu, X.Q.; Zhao, L.Y.; Tan, J.J.; Zhang, Q.; Fu, L.P.; Li, J. Characterization of a novel  $\beta$ -agarase from Antarctic macroalgae-associated bacteria metagenomic library and anti-inflammatory activity of the enzymatic hydrolysates. *Front. Microbiol.* doi:10.3389/fmicb.2022.972272.
7. Hanson, A.D.; Pribat, A.; Waller, J.C.; Cre'cy-Lagard, V. "Unknown" proteins and "orphan" enzymes: the missing half of the engineering parts list — and how to find it. *Biochem. J.* 2010, 425, 1-11.
8. Popper, Z.A.; Michel, G.; Hervé, C.; Domozych, D.S.; Willats, W.G.T.; Tuohy, M.G.; Kloareg, B.; Stengel, D.B.; Evolution and diversity of plant cell walls: from algae to flowering plants. *Annu. Rev. Plant. Biol.* 2011, 62, 567-590.
9. Chiovitti, A.; Kraft, G.T.; Bacic, A.; Liao, M.L. Review: Gelling polysaccharides from Australian seaweeds: research and potential. *Mar. Freshwater Res.* 2001, 52:917-935.
10. Sarwar, G.; Matayoshi, S.; Oda, H. Purification of a kappacarrageenase from marine Cytophaga species. *Microbiol. Immunol.* 2013, 31, 869-877.
11. Blackburn, N.; Fenchel, T.; Mitchell, J. Microscale nutrient patches in planktonic habitats shown by chemotactic bacteria. *Science*. 1998, 282, 254-2256.

12. Arnosti, C.; Wietz, M.; Brinkhoff, T.; et al. The Biogeochemistry of Marine Polysaccharides: Sources, Inventories, and Bacterial Drivers of the Carbohydrate Cycle. *Annu. Rev. Mar. Sci.* 2021, 13, 1-28.
13. Lewis, J.; Kennaway, G.; Franca, S.; et al. Bacterium-dinoflagellate interactions: investigative microscopy of *Alexandrium* spp. (Gonyaulacales, Dinophyceae). *Phycologia*. 2001, 40, 280-285.
14. Bell, W.; Mitchell, R. Chemotactic and growth responses of marine bacteria to algal extracellular products. *Biol. Bull-us.* 1972, 143, 265-277.
15. Ohta, S.; Chang, T.; Ikegami, N.; et al. Antibiotic substance produced by a newly isolated marine microalga, *Chlorococcum* HS-101. *B. Environ. Contam. Tox.* 1993, 50, 171-178.
16. Kubanek, J.; Jensen, P.R.; Keifer, P.A.; et al. Seaweed resistance to microbial attack: a targeted chemical defense against marine fungi. *P. Natl. Acad. Sci. USA.* 2003, 100, 6916-6921.
17. Lachnit, T.; Blümel, M.; Imhoff, J.F.; et al. Specific epibacterial communities on macroalgae: phylogeny matters more than habitat. *Aquat. Biol.* 2009, 5, 181-186.
18. Tujula, N.A.; Crocetti, G.R.; Burke, C.; et al. Variability and abundance of the epiphytic bacterial community associated with a green marine Ulvacean alga. *ISME. J.* 2010, 4, 301-311.
19. Hengst, M.B.; Andrade, S.; Gonzalez, B.; et al. Changes in epiphytic bacterial communities of intertidal seaweeds modulated by host, temporality, and copper enrichment. *Microb. Ecol.* 2010, 60, 282-290.
20. Campbell, A.H.; Marzinelli, E.M.; Gelber, J.; Steinberg, P.D. Spatial variability of microbial assemblages associated with a dominant habitat-forming seaweed. *Front. Microbiol.* 2015, 6, 230.
21. Joke, H.; Frederik, L.; Olivier, D.C.; Anne, W. What we can learn from sushi: a review on seaweed-bacterial associations. *Fems. Microbiology. Ecol.* 2012, 83, 1-16.
22. Gui, Y.Y.; Gu, X.Q.; Fu, L.P.; Zhang, Q.; Fu, L.P.; Li, J. Expression and Characterization of a Thermostable Carrageenase From an Antarctic *Polaribacter* sp. NJDZ03 Strain. *Front. Microbiol.* 2021, 12.
23. Li, J.; Pan, A.H.; Xie, M.S.; Zhang, P.P.; Gu, X.Q. Characterization of a thermostable  $\kappa$ -carrageenase from a hot spring bacterium and plant protection activity of the oligosaccharide enzymolysis product. *J. Sci. Food. Agr.* 2019, 99, 1812-1819.
24. Hehemann, J.H.; Boraston, A.B.; Czjzek, M. A sweet new wave: structures and mechanisms of enzymes that digest polysaccharides from marine algae. *Curr. Opin. Struc. Biol.* 2014, 28, 7-86.
25. Bhattacharyya, P.N.; Tanti, B.; Barman, P.; Jhn, D.K. Culture-independent metagenomic approach to characterize the surface and subsurface soil bacterial community in the Brahmaputra valley, Assam, North-East India, an Indo-Burma mega-biodiversity hotspot. *World. J. Microb. Biot.* 2014, 30, 519-528.
26. Tian, X.P.; Long, L.J.; Wang, F.Z.; Xu, Y.; Li, J.; Zhang, J.; Zhang, C.S.; Zhang, S.; Li, W.J.; *Streptomyces nanhaiensis* sp. nov. a marine streptomycete isolated from a deep-sea sediment. *Int. J. Syst. Evol. Micr.* 2012, 62, 864-868.
27. Tringe, S.G.; Rubin, E.M. Metagenomics: DNA sequencing of environmental samples. *Nat. Rev. Genet.* 2005, 6, 805-814.
28. Pasmore, M.; Costerton, J.W. Biofilms, bacterial signaling, and their ties to marine biology. *J. Ind. Microbiol. Biotechnol.* 2003, 30, 407-413.
29. Cottrell, M.T.; Kirchman, D.L. Natural Assemblages of Marine Proteobacteria and Members of the Cytophaga-Flavobacter Cluster Consuming Low- and High-Molecular-Weight Dissolved Organic Matter. *Appl. Environ. Microb.* 2000, 66, 1692-1697.
30. Segawa, T.; Miyamoto, K.; Ushida, K.; Agata, K.; Okada, N.; Kohshima, S. Seasonal change in bacterial flora and biomass in mountain snow from the Tateyama Mountains, Japan, analyzed by 16S rRNA gene sequencing and real-time PCR. *Appl. Environ. Microbiol.* 2005, 71, 123-130.
31. Amato, P.; Hennebelle, R.; Magand, O.; Sancelme, M.; Delort, A.M.; Barbante, C. Bacterial characterization of the snow cover at Spitzberg, Svalbard. *FEMS Microbiol. Ecol.* 2007, 59, 255-264.
32. Viera, R.P.; Ggonzalez, A.M.; Cardoso, A.M. Relationships between bacterial diversity and environmental variables in a tropical marine environment, Rio de Janeiro. *Environ. Microbiol.* 2008, 10, 189-199.
33. Lozada, M.; Diéguez, M.C.; García, P.E.; Dionisi, H.M. Microbial communities associated with kelp detritus in temperate and subantarctic intertidal sediments. *Sci. Total. Environ.* 2023, 857, 159392.
34. Martin, K.; Schmidt, K.; Toseland, A.; Boulton, C.A.; et al. The biogeographic differentiation of algal microbiomes in the upper ocean from pole to pole. *Nat. Commun.* 2022, <https://doi.org/10.1038/s41467-021-25646-9>.
35. Martin, M.; Barbeyron, T.; Martin, R.; et al. The Cultivable Surface Microbiota of the Brown Alga *Ascophyllum nodosum* is Enriched in Macroalgal-Polysaccharide-Degrading Bacteria. *Front. Microbiol.* 2015. doi: 10.3389/fmicb.2015.01487
36. Longford, S.; Tujula, N.; Crocetti, G.; Holmes, A.; Holmström, C.; Kjelleberg, S.; Steinberg, P.; Taylor, M. Comparisons of diversity of bacterial communities associated with three sessile marine eukaryotes. *Aquat. Microb. Ecol.* 2007, 48, 217-229.
37. Egelund, J.; Ellis, M.; Doblin, M.; et al. Genes and Enzymes of the GT31 Family: Towards Unravelling the Function(s) of the Plant Glycosyltransferase Family Members. Wiley-Blackwell, 2010.

38. Matard-Mann, M.; Bernard, T.; Leroux, C.; et al. Structural insights into marine carbohydrate degradation by family GH16  $\kappa$ -carrageenases. *J. Biol. Chem.* 2017, 292.
39. Pomin, V.H. Structural and functional insights into sulfated galactans: a systematic review. *Glycoconjugate J.* 2010.
40. Lavin, P.; Gallardo-Cerda, J.; Torres-Diaz, C.; Asencio, G.; Gonzalez, M. Antarctic strain of *Bacillus* sp. with extracellular agarolytic and alginate-lyase activities. *Gayana.* 2013, 77, 75–82.
41. Dhargalkar, V.K.; Verlecar, X.N. Southern Ocean seaweeds: a resource for exploration in food and drugs. *Aquaculture.* 2009, 287, 229–242.
42. Zhu, B.; Ni, F.; Xiong, Q.; Yao, Z. Marine oligosaccharides originated from seaweeds: Source, preparation, structure, physiological activity and applications. *Crit. Rev. Food Sci.* 2021, 61, 60–74.
43. Herrmann, M.; Saunders, A.M.; Schramm, A. Archaea Dominate the Ammonia-Oxidizing Community in the Rhizosphere of the Freshwater Macrophyte *Littorella uniflora*. *Appl. Environ. Microb.* 2008, 74, 3279–3283.
44. Hao, B.; Wu, H.; Cao, Y.; et al. Comparison of periphyton communities on natural and artificial macrophytes with contrasting morphological structures. *Freshwater. Biol.* 2017, 62, 1783–1793.
45. Bell, W.; Mitchell, R. Chemotactic and growth responses of marine bacteria to algal extracellular products. *Biol. Bull-us.* 1972, 143, 265–277.
46. Ibarbalz, F.M.; et al. Global trends in marine plankton diversity across kingdoms of life. *Cell.* 2019, 179, 1084–1097.
47. Mori, A.S.; Isbell, F.; Seidl, R. Beta-diversity, community assembly, and ecosystem functioning. *Trends. Ecol. Evol.* 2018, 33, 549–564.
48. Li, Y.; Chen, H.; Song, L.; et al. Effects on microbiomes and resistomes and the source-specific ecological risks of heavy metals in the sediments of an urban river. *J. Hazard. Mater.* 2021, 409, 124472.
49. Imchen, M.; Kumavath, R. Metagenomic insights into the antibiotic resistome of mangrove sediments and their association to socioeconomic status. *Environ. Pollut.* 2021, 268, 115795.
50. Zhao, R.; Yu, K.; Zhang, J.; et al. Deciphering the mobility and bacterial hosts of antibiotic resistance genes under antibiotic selection pressure by metagenomic assembly and binning approaches. *Water. Res.* 2020, 186, 116318.
51. Cuadrat, R.R.C.; Sorokina, M.; Andrade, B.G.; et al. Global ocean resistome revealed: exploring antibiotic resistance gene abundance and distribution in TARA Oceans samples. *GigaScience.* 2020, 9, 1–12.
52. Chen, J.; McIlroy, S.E.; Archana, A.; et al. A pollution gradient contributes to the taxonomic, functional, and resistome diversity of microbial communities in marine sediments. *Microbiome.* 2019, 7, 104.
53. Dong, X.Y.; Lan, H.Y.; Huang, L.G. Metagenomic Views of Microbial Communities in Sand Sediments Associated with Coral Reefs. *Microb. Ecol.* 2022, <https://doi.org/10.1007/s00248-021-01957-8>

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.