Systematic and Applied Microbiology 42 (2019) 263–274

Contents lists available at ScienceDirect

Systematic and Applied Microbiology

journal homepage: http://www.elsevier.com/locate/syapm

Carbohydrate catabolic capability of a *Flavobacteriia* bacterium isolated from hadal water

Jiwen Liu[a,b,1], Chun-Xu Xue[a,1], Hao Sun[a], Yanfen Zheng[a], Zhe Meng[a], Xiao-Hua Zhang[a,b,*]

[a] *MOE Key Laboratory of Marine Genetics and Breeding, College of Marine Life Sciences, Ocean University of China, 5 Yushan Road, Qingdao 266003, China*
[b] *Laboratory for Marine Ecology and Environmental Science, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266071, China*

ARTICLE INFO

Article history:
Received 29 September 2018
Received in revised form
17 December 2018
Accepted 15 January 2019

Keywords:
Flavobacteriia
Hadal water
Carbohydrate catabolism
Organic matter

ABSTRACT

Flavobacteriia are abundant in many marine environments including hadal waters, as demonstrated recently. However, it is unclear how this flavobacterial population adapts to hadal conditions. In this study, extensive comparative genomic analyses were performed for the flavobacterial strain *Euzebyella marina* RN62 isolated from the Mariana Trench hadal water in low abundance. The complete genome of RN62 possessed a considerable number of carbohydrate-active enzymes with a different composition. There was a predominance of GH family 13 proteins compared to closely related relatives, suggesting that RN62 has preserved a certain capacity for carbohydrate utilization and that the hadal ocean may hold an organic matter reservoir distinct from the surface ocean. Additionally, RN62 possessed potential intracellular cycling of the glycogen/starch pathway, which may serve as a strategy for carbon storage and consumption in response to nutrient pulse and starvation. Moreover, the discovery of higher glycoside hydrolase dissimilarities among *Flavobacteriia*, compared to peptidases and transporters, suggested variation in polysaccharide utilization related traits as an important ecophysiological factor in response to environmental alterations, such as decreased labile organic carbon in hadal waters. The presence of abundant toxin exporting, transcription and signal transduction related genes in RN62 may further help to survive in hadal conditions, including high pressure/low temperature.

Introduction

Members of marine *Bacteroidetes*, especially the class *Flavobacteriia*, are distributed in a variety of environmental niches, including seawater and sediments [1]. *Flavobacteriia* are frequently documented as the dominant bacterial groups during and after phytoplankton blooms [57,74,75,80], and their abundance increases with enhanced primary production [81]. This suggests that species of *Flavobacteriia* prefer the "micro-eutrophic" environment characterized by concentrated nutrients and organic matter particles [54]. *Flavobacteriia* can directly act on algal cells and algal-derived detrital particles [74], and are specialized in degrading many high molecular weight compounds, such as polysaccharides and proteins [5,45,72,77].

The polymer-degradation capabilities of *Flavobacteriia* are executed by a suite of peptidases and carbohydrate-active enzymes (CAZymes), which consist of glycoside hydrolases (GHs), carbohydrate esterases (CEs), polysaccharide lyases (PLs) and glycoside transferases (GTs), with GHs mostly constituting the highest proportion. The relative abundance of peptidases and GHs varies among *Flavobacteriia* species [5–7,14,20,21,45,70,72,81]. For example, the planktonic *Gramella forsetii* KT0803 (40 GHs, 116 peptidases) [6], *Polaribacter* sp. MED152 (30 GHs, 93 peptidases) [20] and *Dokdonia donghaensis* MED134 (13 GHs, 106 peptidases) [21] have lower GH:peptidase ratios than algae-associated *Formosa agariphila* KMM 3901 (88 GHs, 129 peptidases) [45] and *Zobellia galactanivorans* Dsij (141 GHs, 101 peptidases) [5]. This low ratio seems to be attributed to a reduction in the number of GHs, which is frequently observed in species with relatively small genomes, although they contain proteorhodopsin for energy acquisition [20,82]. Additionally, marine *Bacteroidetes* have been reported to have a higher abundance of peptidases than GH genes when compared with non-marine *Bacteroidetes* [14]. These observations make it very likely that habitat condition is an important determinant of the GH:peptidase ratio.

The growth requirement for abundant organic matter may make *Flavobacteriia* less abundant in the deep sea prokaryotic community [64], where organic matter has a lower concentration and is

---

* Corresponding author at: MOE Key Laboratory of Marine Genetics and Breeding, College of Marine Life Sciences, Ocean University of China, 5 Yushan Road, Qingdao 266003, China.
E-mail address: xhzhang@ouc.edu.cn (X.-H. Zhang).

[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.syapm.2019.01.002

metabolically more refractory [25], compared to the surface ocean. However, recent studies have revealed a significantly distinct prokaryotic community in the hadal water (>6000 m below sea level), with the predominance of heterotrophic *Gammaproteobacteria* and *Bacteroidetes* (especially *Flavobacteriia*) bacteria, compared to that in bathypelagic and abyssopelagic waters [52,55,73]. This enrichment of heterotrophic bacteria may result from sporadic concentration of nutrients, facilitated by lateral transportation from trench rims and slopes that can be triggered by earthquakes [3]. To overcome this discrete nutrient supply, we hypothesized that certain metabolic strategies for carbon storage and utilization may have developed in this heterotrophic population. However, the microbial physiological and functional features in >6000 m hadal water are largely unknown. The Mariana Trench is the deepest ocean location and represents the most underexploited area. It serves as an excellent proxy of the hadal habitat for elucidating the microbial metabolic capability therein.

A hadal water sample was collected from the Challenger Deep of the Mariana Trench at a depth of 8727 m and spread on marine agar and R2A medium for bacterial isolation. Consistent with the previous results stated above, a high number of *Bacteroidetes* isolates were obtained, among which, a yellow-colored strain (RN62) was identified as *Euzebyella marina* based on 16S rRNA gene sequence similarity analysis. The genus *Euzebyella* was established in 2010 by Lucena et al. [43] and currently contains only three species [35,43,87]. No genomic sequence of *Euzebyella* is currently available. Therefore, this current study reports the complete genome of *E. marina* RN62 and compares it to the genomes of closely related species from ten genera. In-depth analyses of CAZymes and peptidases were performed with the aim of better understanding how hadal microbes utilize organic matter. The results revealed that *E. marina* RN62 lacked gliding ability, possessed a considerable capacity for carbohydrate metabolism, and had the potential for intracellular cycling of glycogen/starch.

Materials and methods

Bacterial isolation and identification

Water samples were collected from six depths (0, 1000, 1759, 3699, 5369 and 8727 m) in the Mariana Trench (142°30′E, 11°23.036′N) during a “Marathon Program” survey organized by the Ocean University of China from 2nd to 5th January 2016 by *R/V* Dong Fang Hong 2. The water sample was ten-fold serially diluted and spread on marine agar (MA) and R2A plates. Cultivation was conducted at both *in situ* (~2 °C) and room temperatures. After two days’ incubation at room temperature, a yellow colony was picked from the 8727 m plate, and purified by streaking three times on MA. The strain (RN62) was preserved in 0.85% (w/v) sterile sodium chloride solution supplemented with 15% (v/v) glycerol at −80 °C. The presence of gliding motility was investigated using the methods described by Ref. [8]. The genomic DNA of strain RN62 was extracted with the standard phenol–chloroform method [46]. The 16S rRNA gene was amplified with the universal primers B8F (5′-AGAGTTTGATCCTGGCTCAG-3′) and B1510R (5′-GGTTACCTTGTTACGACTT-3′). Amplicons were cloned into the pUCm-T vector and sequenced at BGI (Qingdao, China). Taxonomy and pairwise similarities were determined by using the EzBioCloud server (http://www.ezbiocloud.net/) [84].

Genome sequencing and annotation

The genome of strain RN62 was sequenced using the PacBio RS II Sequencing System (Pacific Biosciences, Menlo Park, CA) and Hiseq 2000 (Illumina) with a 20 kb and 270 bp paired-end library, respectively. The raw reads were *de novo* assembled using Canu 1.1 [32]. tRNAs and rRNAs were identified using the tRNAscan-SE 2.0 server [42] and RNAmmer 1.2 server [37], respectively. Putative genes were identified and annotated automatically in the RAST server with default parameters [53]. Clusters of orthologous groups of proteins (COG) [19] were predicted using RPS-BLAST [66] with an *E*-value cutoff of 10^{-5}. Metabolic pathways were predicted in the Kyoto encyclopedia of genes and genomes (KEGG) [27]. CAZymes were annotated against three databases and consistent annotations among them were considered as positive results. The CAZy database [41] was searched using BLASTp [2] with an *E*-value cutoff of 10^{-30}, whereas the dbCAN 5.0 [83] and Pfam 30.0 [17] databases were searched using HMMER 3.0 [16] with default parameters. Specific function or substrate specificity of members within a GH family were inferred by phylogenetic analysis with sequences derived from the CAZy database. Localization of GH proteins was predicted using SOSUI-GramN [30], PSORTb 3.0.2 [85], LipoP 1.0 [26] and SignalP 4.1 [56]. Subfamilies of sulfatases were classified in the SulfAtlas database [4]. Peptidases were annotated based on BLASTp against the MEROPS 11.0 database [60] using an *E*-value cutoff of 10^{-4}. Transporters were predicted based on BLASTp against the Transporter Classification Database (TCDB, [63]) using an *E*-value cutoff of 10^{-5}. Polysaccharide utilization loci (PUL) were defined as the presence of at least one CAZyme within ten genes downstream and/or upstream of a *susC-susD* gene. PUL were then extended by searching for genes potentially involved in carbohydrate degradation with a five gene sliding window beginning from the *susC–susD* and CAZyme genes.

Genome comparisons

Genomes of 26 *Flavobacteriia* species closest to RN62, according to the 16S rRNA gene phylogeny, were downloaded from the GenBank database (http://www.ncbi.nlm.nih.gov). They were distributed in ten genera (Table S1). The average nucleotide identity (ANI) between RN62 and the retrieved genomes was calculated using a Perl script (https://github.com/jhbadger/scripts/blob/master/ANI.pl). Orthologous proteins of the 27 genomes were identified using the GET_HOMOLOGUES package [10]. Orthologous families occurring in all three clustering algorithms embedded in the package, i.e., bidirectional best hit (BDBH), COGtriangles [34] and OrthoMCL [40], were considered as valid results. All the single-copy families were chosen for phylogenomic analysis, and sequences of each family were aligned using the MAFFT software [29]. The resulting alignments were concatenated and trimmed in TrimAl 1.2 [9]. The maximum-likelihood phylogenomic tree was constructed using the RAxML software [69] with the gamma and WAG + I + F model, as predicted by Prottest 3.4 [11]. The CAZymes, peptidases, transporters, COG category and GH localization of the relatives’ genomes were predicted using the same method described above.

Distribution matrices of CAZymes, peptidases and transporters were generated. Dissimilarities of each functional guild across the 27 genomes were calculated based on the Bray–Curtis distance, and their clustering relationship was illustrated using non-metric multidimensional scaling (NMDS). A total of 344 genomes covering most *Flavobacteriia* genera were downloaded from the GenBank database on Aug 25, 2018. The distribution matrices of CAZymes, peptidases and transporters among these genomes were constructed. The analyses were conducted in the R software package (http://www.r-project.org).

The genome sequence was deposited in NCBI’s GenBank (accession number CP032050).

Table 1
Genomic information for *Euzebyella marina* RN62 and closely related relatives.

Strains	Size (bp)	Contig no.	GC content (%)	Mean sequence size	N50	Isolation
Kriegella aquimaris DSM 19886	6,057,242	48	39.6	126,192.5	317,933	Seawater
Euzebyella marina RN62	4,707,469	1	40.2	4,707,469	NA	Hadal water
Pricia antarctica DSM 23421	4,851,997	52	43.8	93,307.6	450,975	Antarctic intertidal sediment
Pseudozobellia thermophila DSM 19858	5,036,817	36	47.1	139,911.6	475,669	Macroalga
Zobellia galactanivorans DsiJ	5,521,712	1	42.8	5,521,712	NA	Macroalga
Zobellia uliginosa MAR 2009 138	5,358,002	2	38.1	2,679,001	17,314	Seawater after phytoplankton bloom

Results and discussion

Genomic properties and phylogeny of E. marina *RN62*

In this study, the *E. marina* genome was analyzed as a representative species of hadal *Flavobacteriia*. This was because *E. marina* was only isolated from depths of 8727 m (three out of 63 sequenced strains) and 5367 m (one out of 42), and not from shallower depths, with the numbers of sequenced strains ranging from 48 to 70. In addition, mapping Mariana Trench metagenomic reads (0–9600 m; data not shown, Liu et al., submitted) to the *E. marina* RN62 genome showed that this species was more abundant in deep-sea water than surface water, although the overall abundance was low (Fig. S1).

Strain RN62 had a single chromosome with a genome size of 4,707,469 bp and G + C mol% content of 40.2% (Fig. S2, Table 1). The genome of RN62 contained 4215 ORFs, 38 tRNAs for all 20 standard amino acids, and a single set of 16S-23S-5S rRNA gene operons. To evaluate the phylogenetic position of strain RN62, a maximum likelihood tree was constructed with the core genome of RN62 and its close relatives. The tree identified four major clusters that shared comparable phylogenetic depth and were supported by high bootstrap values (>95%). One of them was comprised of RN62, *Pseudozobellia thermophile* DSM 19858, *Kriegella aquimaris* DSM 19886, *Pricia antarctica* DSM 23421, *Zobellia galactanivorans* DsiJ and Zobellia uliginosa MAR 2009 138 (Fig. 1). Most *Zobellia* strains are isolated from algae and *Z. galactanivorans* is well known as an algal polysaccharide degrader [5]. The genome size of RN62 was lower than those of the five closely related species that were either draft or complete, despite a similar G + C mol% content (Table 1). Strain RN62 showed an average ANI value of 70.7% to the five species, which was slightly higher than the average value of 68.8% to the remaining species (Fig. S3), but consistent with the phylogenomic clustering.

General features of energy metabolism

Strain RN62 contained a full set of genes for the citrate cycle, the Embden–Meyerhof–Parnas (EMP) and pentose phosphate pathways, including both the oxidative and non-oxidative phases. An incomplete Entner–Doudoroff (ED) pathway was present that included genes for 2-keto-3-deoxygluconokinase (EC 2.7.1.45) and 2-keto-3-deoxyphosphogluconate aldolase (EC 4.1.2.14), which catabolize the production of 2-keto-3-deoxyphosphogluconate (KDPG) and the subsequent transformation of KDPG to pyruvate, respectively. Pyruvate was converted to acetyl-CoA by the pyruvate dehydrogenase complex but not by the formate C-acetyltransferase (pyruvate formate-lyase). The electron transport chain contained a NADH dehydrogenase (complex I), succinate dehydrogenase (complex II), two cytochrome c oxidases (cbb3 and aa3 type; complex IV), cytochrome bd complex and an F-type H^+-transporting ATPase (complex V). RN62 also had the potential to gain energy from carbon monoxide, as its genome carried the aerobic carbon monoxide dehydrogenase (*coxL*, *coxM*, and *coxS* genes) and three accessory genes (*coxD, coxE* and *coxG*). Consistently, carbon monoxide dehydrogenase encoding genes have been found in deep oceans and are mainly derived from *Bacteroidetes* and *Chloroflexi* [38,48].

For nitrogen, the RN62 genome had a nitrate/nitrite transport system, assimilatory nitrate reductase gene *narB*, and dissimilatory nitrite reductase gene *nirB*, which indicated an ammonia production pathway from nitrate. Genes for nitrogen regulatory protein P-II (*glnB*) and ammonium transporter were concatenated, likely playing a role in sensing and uptake of environmental ammonium [18]. For sulfur, genes for sulfate permease (*sulP*), thiosulfate dehydrogenase (*doxA*), and all genes except the second step of assimilatory sulfate reduction were present.

For phosphorus, a low-affinity Na (+)-phosphate symporter was found, whereas the high-affinity *pst* transporter was not found. This was consistent with the observed high phosphate concentration in hadal water [52] from where RN62 was isolated. There were two copies of genes for polyphosphate kinase (*ppk*), indicating a polyphosphate related strategy for energy storage; however, the polyphosphate hydrolysis gene was absent. Instead, encoding genes for two alkaline phosphatases (*phoA*, *phoD*) and a phospholipase A1 were present; *phoD* was found exclusively in RN62 compared with its five closely related relatives. This, together with the presence of a *phoR*-*phoB* two-component regulatory system regulating the synthesis of alkaline phosphatase, suggested that RN62 had a high capability for organic phosphate utilization.

Capacity for polysaccharide catabolism and unusual glycogen intracellular cycling

The genome of strain RN62 encoded 122 CAZymes, including 62 GHs, 43 GTs, 1 PL, 10 CEs and 5 CBMs, 224 peptidases and 28 sulfatases (Table 2). Clearly, RN62 had a higher relative abundance of peptidases than GHs, and even than CAZymes. A similar phenomenon was seen for all 27 *Flavobacteriia* strains analyzed in the study (Fig. 2, $P < 0.001$). A previous study also demonstrated that marine *Bacteroidetes* possessed a higher number of peptidases than GHs [14]. Compared to the five closely related strains (Fig. 1), RN62 had the lowest CAZyme abundance, which was reflected by a reduction in the number of GH families, the relative abundance of GHs and PLs, as well as the proportion of GHs that were predicted to be periplasmic/membrane-bounded/extracellular (Table 2). This reduction in carbohydrate metabolism was accompanied by the lack of gliding ability. However, a comparison of the CAZyme abundance between strain RN62 and species from other phylogenetic clusters (Fig. 1) showed that the former had a higher abundance than 19 out of the 21 species (Fig. 2). Furthermore, the number of GH in the RN62 was higher than the average value in *Flavobacteriia* (47 GHs) [5], and was even higher than in other deep-sea bacteria

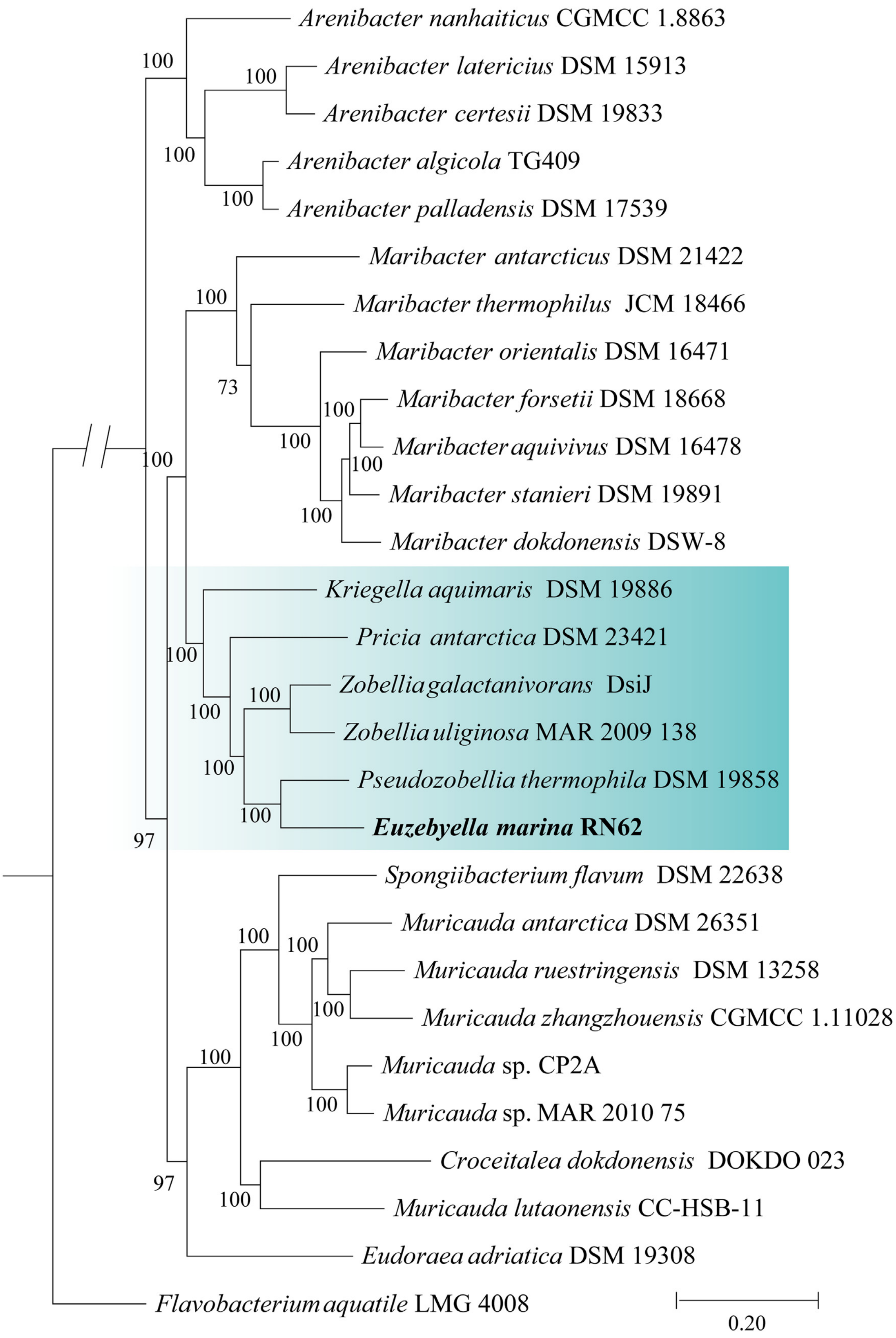

Fig. 1. The phylogenetic relationships of *E. marina* RN62 and its close relatives. The tree was constructed with a concatenated alignment of orthologous proteins using the maximum likelihood method. The genome of *Flavobacterium aquatile* LMG 4008 was used as an outgroup. Values at the nodes denote the bootstrap value based on 1,000 replicates, and the scale bar indicates the number of substitutions per site.

Table 2
The number of peptidases and CAZymes in *Euzebyella marina* RN62 and closely related relatives.

Strains	Peptidase	Peptidase family	CAZyme	PL	GT	GH	CE	CBM	GH family	Sulfatase	GH localization	
											PP/MB/EC	CP
Kriegella aquimaris DSM 19886	262	76	176	11	42	98	17	7	34	88	60	38
Euzebyella marina RN62	224	72	122	1	43	62	10	5	29	28	27	35
Pricia antarctica DSM 23421	233	79	135	1	46	64	14	7	33	36	30	34
Pseudozobellia thermophila DSM 19858	235	75	158	4	48	88	10	6	34	61	50	38
Zobellia galactanivorans DsiJ	223	78	185	7	44	117	9	8	38	75	70	47
Zobellia uliginosa MAR 2009 138	212	75	174	10	44	104	6	10	32	98	57	47

Note: PL, polysaccharide lyase; GT, glycoside transferase; GH, glycoside hydrolase; CE: carbohydrate esterase; CBM: carbohydrate-binding module; PP, periplasmic; MB, membrane-bounded; EC, extracellular; CP, cytoplasmic.

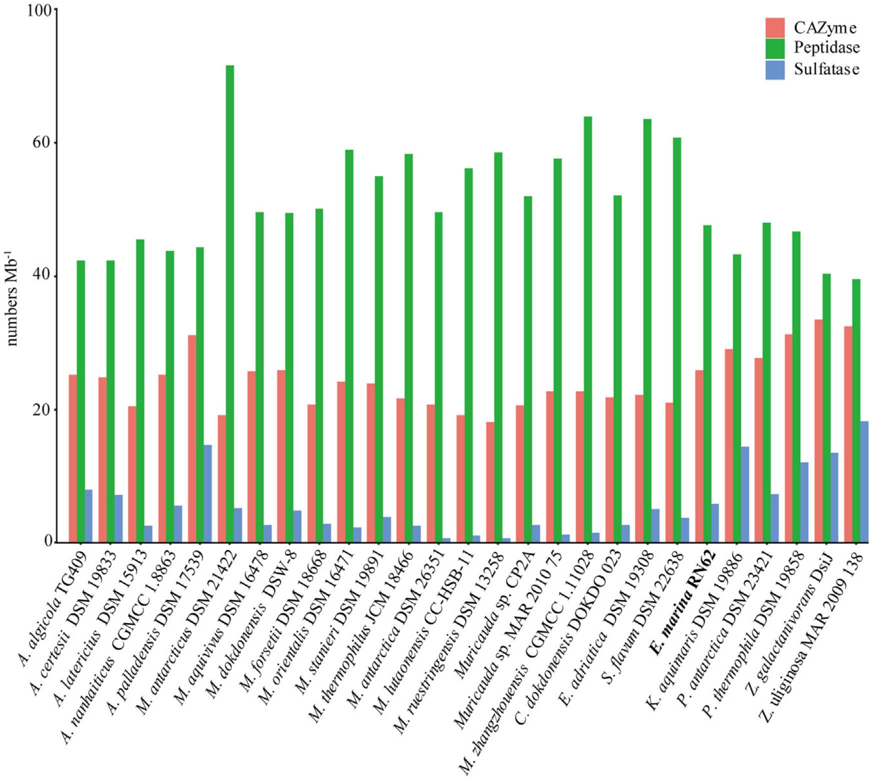

Fig. 2. Relative abundance of CAZyme, peptidase and sulfatase genes in the genomes of the 27 *Flavobacteriia* species shown in Fig. 1. The abundance was normalized by genome size.

[5], suggesting that strain RN62 preserved the capacity to utilize certain carbohydrates in hadal waters.

GH13 was the most abundant GH family in all the genomes analyzed (Fig. 3A). GH13 (12.9%), GH43 (9.7%) and GH3 (6.5%) constituted the most abundant GH families in RN62 (Fig. 3A). It was noticeable that the proportion of GH13 in the RN62 genome (12.9%) was higher than those in other strains of the same cluster (4.5–7.8%; Fig. 3A). Specifically, four of the eight GH13 proteins belonged to putative α-amylase; three of which were predicted to be cytoplasmic and the other periplasmic (Fig. S4). These results may indicate a high potential of RN62 for intracellular utilization of starch or glycogen. Correspondingly, the strain had the potential for glycogen biosynthesis because it possessed two copies of the gene encoding starch synthase (*glgA*) and one copy of the gene encoding 1,4-α-glucan (glycogen) branching enzyme (*glgB*), although the gene for glucose-1-phosphate adenylyltransferase (*glgC*) was missing. A similar glycogen biosynthetic pathway has also been found in *Z. galactanivorans* [5]. However, RN62 seemed to have a more pronounced glycogen recycling pathway in addition to α-amylase because of the presence of genes encoding glycogen phosphorylase (*glgP*) that breaks glycogen into glucose-1-phosphate and phosphorylase-limited dextrin, and maltodextrin glucosidase (*malZ*) that converts maltodextrin into glucose, and they were located next to each other in the genome, although a gene for glycogen de-branching enzyme (pullulanase) was missing (Fig. S5). The *glgP* gene is rarely present in *Flavobacteriia* [5] but it was found in *P. antarctica* isolated from a cold and oligotrophic Antarctic intertidal sediment [86], although it was not found in the other four closely related strains. Moreover, the *malZ* gene was specific to RN62. The results provided strong evidence for the existence of an unusual intracellular glycogen recycling pathway in RN62 as an important strategy for carbon storage and consumption in order

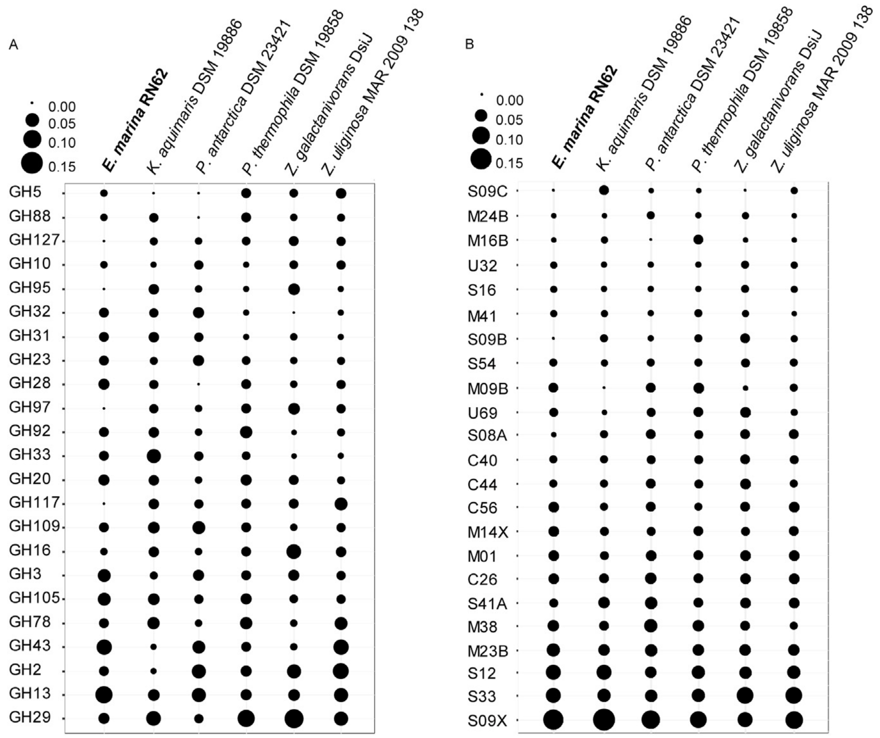

Fig. 3. Distribution of the top abundant glycoside hydrolases and peptidases in *E. marina* RN62 and its five closely related relatives. Dot size in the panel is in proportion to relative abundance.

to overcome the discontinuous organic matter supply in the hadal environment.

The GH43 and GH3 families comprise a functionally diverse set of enzymes. Enzymes from the GH43 family may help RN62 to utilize xylans derived mainly from plant cell walls, as well as some green and red macroalgae [12,58]. Four of them were predicted to be periplasmic/membrane-bound. RN62 also harbored a complete pathway for converting D-xylose to L-arabinose, including xylose isomerase (*xylA*), xylulose kinase (*xylB*), L-arabinose isomerase (*araA*), L-ribulose-5-phosphate 4-epimerase (*araD*) and L-ribulokinase (*araB*). The GH3 family is represented mainly by β-hexosaminidase that may be involved in processing oligosaccharides [78].

On the other hand, GH29 and GH16 were less represented in RN62 than other strains of the cluster, except for *P. antarctica* (Fig. 3A). GH29 is represented mainly by α-L-fucosidase, which can act on fucoidan, a sulfated polysaccharide produced mainly by brown algae [50]. GH16 includes diverse hydrolases toward plant and algal polysaccharides, including laminarin, agar, κ-carrageenan and xyloglucan. One GH16 family enzyme of *Z. galactanivorans* has been found to exhibit a high specificity for algal β-1,3-glucans [36]. The only GH16 family protein in RN62 belonged to endo-1,3-β-glucanase (laminarinase), whereas the number of sulfatases in RN62 (28) was comparable to *P. antarctica* and lower than other close relatives. However, this number was higher than the average value in *Flavobacteriia* (10) and other deep-sea bacteria [5]. All sulfatases in the RN62 genome belonged to family S1 (formyl-glycine dependent sulfatases), which can be further classified into 14 subfamilies. Predominant amongst these were S1_8 (4), SI_11 (4), S1_15 (3) and S1_20 (3) (Table S2). S1_8 and S1_11 were *N*-sulfoglucosamine sulfohydrolase and heparin/heparan sulfate 6-*O*-sulfatase, respectively, and both of them are involved in heparan sulfate degradation. This indicates the presence of animal polysaccharides in hadal waters, which may be derived from a diverse array of metazoan organisms that have been found in deep trenches [23], and would serve as carbon and energy sources for RN62.

In contrast to GH families, the number of peptidase families and the relative abundance of each peptidase family were similar between RN62 and its closely related relatives (Fig. 3B). Serine peptidases were predominantly represented, with families S9 (oligopeptidase), S12 (carboxypeptidase B) and S33 (prolyl aminopeptidase) being the most abundant ones. The predicted numbers of CAZymes, peptidases and sulfatases of *Z. galactanivorans* in this study were fewer than previously reported [5], which may relate to the stringent assignment strategy applied.

Transporter classification

The most abundant transporters in the RN62 genome annotated in TCDB were the TonB-dependent receptor/transporter (TBDR, 9.6%), followed by the ATP binding cassette (ABC) transporter (9.2%), solute:sodium symporter (SSS, 5.2%), and the resistance-nodulation-cell division (RND, 5.2%) superfamily (Fig. S6). These were also the most abundant transporters across all 27 genomes analyzed (Fig. S6). There were two types of TBDR that were SusC-like and FhuA-like proteins. The *susC*-like TBDR is associated with a *susD*-like gene encoding outer membrane lipoproteins, which was first identified as part of the *sus* cluster responsible for starch utilization in an intestinal *Bacteroides thetaiotaomicron* [61,62]. Among the 74 predicted TBDR in the RN62 genome, 34 were SusC-like and 40 were FhuA-like. The SusC-like/FhuA-like TBDR ratio of RN62 was 0.85, which was lower than in the other five strains (all greater than 1) in the same cluster (Table S3). This may imply that the strain possessed more TBDR specific to small molecules than to polymers [5], since the SusC-like and FhuA-like proteins in *Bacteroidetes* have been documented to be specific for polymers and small molecules, respectively [13,33]. However, this conclusion needs further verification, as non-SusC-like TBDRs can also be used for polymer uptake by members of phyla other than *Bacteroidetes* and transportation of other substrates, such as vitamins and siderophores [71] that are unrelated to carbon utilization.

Fig. 4. Putative polysaccharide utilization loci of *E. marina* RN62. Polysaccharide utilization loci were defined by possessing a *susC-susD* gene cluster surrounded by at least one CAZyme gene within ten genes downstream and/or upstream (in the case of PULs10 and 12, CAZymes are located 12-13 genes away from *susC-susD*). PUL was then extended by searching for genes potentially involved in carbohydrate degradation with a five-gene sliding window beginning from the *susC-susD* and CAZyme genes.

The six strains shared a comparable abundance of ABC and SSS. However, RN62 harbored a higher abundance of the RND superfamily, which plays a role in substrate efflux *via* an H^+ antiport mechanism, than the other five strains. Specifically, there were more multidrug exporters in the RN62 genome than the other five strains; the heavy metal exporters of RN62 were comparable with *P. antarctica* and were doubled compared with the other strains. Although the Mariana Trench is a long way from land and human disturbance, the deep trenches have been considered as sinks for contaminants, such as persistent organic pollutants [67]. Extraordinary levels of persistent organic pollutants have been identified in amphipod fauna collected from two trenches [24]. Similarly, the location where *P. antarctica* was isolated is close to the Chinese Antarctic Zhongshan Station, and is thus subjected to human disturbances. In this context, the two strains may use the excess RND exporters to resist toxic compounds.

Polysaccharide utilization loci characterization

The SusD-like protein, since its first discovery, has been documented to bind various polysaccharides, including starch [61,62], fructan [68], cellulose [44], chitin [39], glucan [76], carrageenan [15] and xyloglucan [65], and thus exhibits a universal role as a polysaccharide capturer [28,77]. It can facilitate the SusC-like TBDR in taking up substrates after polysaccharide is degraded extracellularly. The *susC-susD* gene cluster, in most cases, is co-located with genes encoding GH, PL and sulfatase, which together organize into

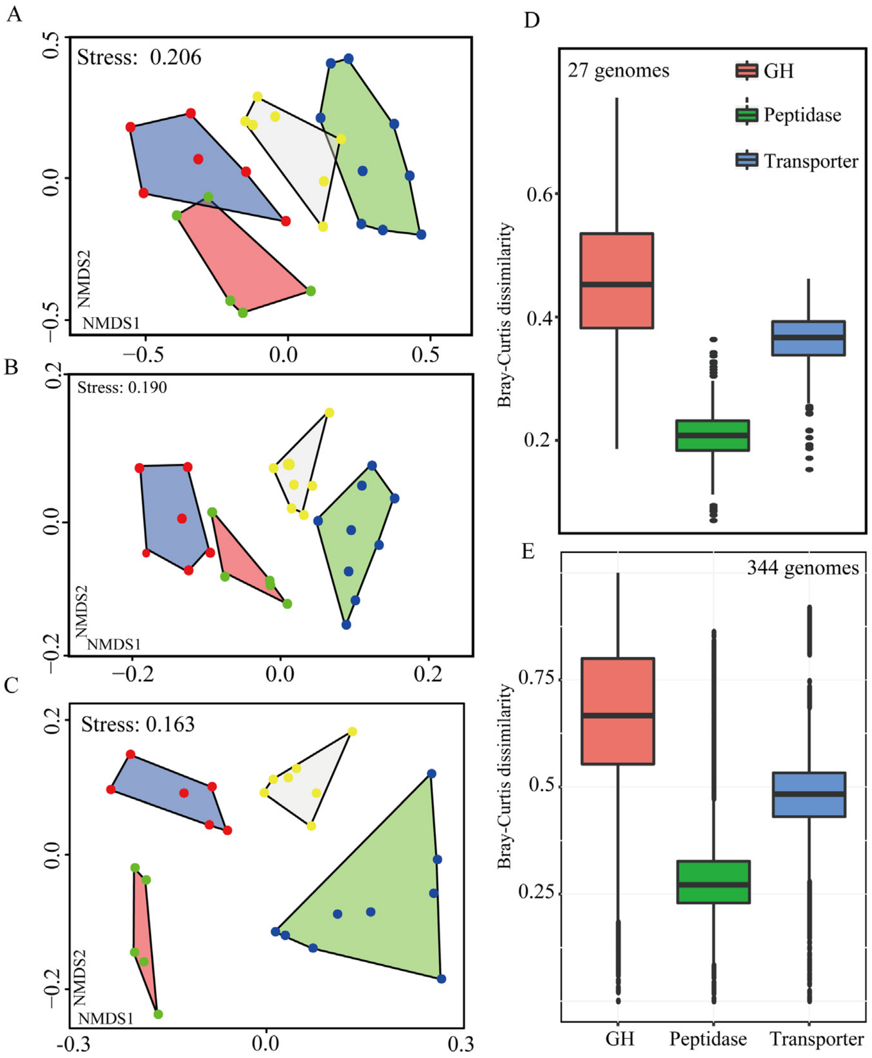

Fig. 5. Distribution and dissimilarity level of different functional groups across *Flavobacteriia* genomes. Panels A, B and C show the distribution of CAZymes, peptidases and transporters, respectively, across the 27 *Flavobacteriia* genomes. Dots with different colors indicate the four phylogenetic clusters as shown in Fig. 1, with green representing *Arenibacter* species, yellow representing *Maribacter* species, blue representing *Muricauda*, *Croceitalea* and *Eudoraea* species, and red representing *Euzebyella*, *Kriegella*, *Pseudozobellia*, *Zobellia* and *Pricia* species. Panels D and E show the dissimilarity levels of the three functional groups across the 27 and 344 *Flavobacteriia* genomes, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

operons or regulons described as polysaccharide utilization loci (PUL) [47,68]. RN62 contained 15 PULs that were defined here as containing at least a *susC-susD* gene cluster and a CAZyme gene. Predicted functions of the PULs were described where possible: PUL2 may target α-L-rhamnoside; PUL3 contained the most GHs and may target sulfated rhamnogalacturonan and xylan; PUL5 may target sulfated fucoside or fucoidan; PUL6 may target laminarin; PUL7 may target sulfated xylan; PUL10 may target sulfated starch and/or α-glucan; PUL11 may target sulfated pectin. PULs 8, 12 and 15 contained only β-hexosaminidase, but from different families, and they may target oligosaccharides. Many of the targeted substrates were sulfated, indicating that the strain retained several PULs for utilizing algal derived polysaccharides, although the substrate is scarce. The only PL was not clustered in a PUL, but 37 out of 62 GHs and 16 out of 28 sulfatases were distributed in the PULs (Fig. 4). However, 9 out of 15 PULs contained only a single GH and another one contained two GHs from the same family. Two of the four putative α-amylases in the GH13 family were present and distributed in different PULs, indicating that starch utilization of RN62 could occur outside PULs. The PUL diversity and complexity of strain RN62 was lower than the algal-associated *Algibacter alginolytica* [70] and *Z. galactanivorans* [5], the latter of which was closely related to RN62.

Divergent evolution of carbohydrate metabolism in Flavobacteriia *species*

To elucidate the evolutionary pathway of substrate utilization traits in *E. marina* RN62, the level of dissimilarly was compared for the compositions of GH, peptidase and transporter families in the 27 genomes. GHs were used as a proxy of CAZymes as they were significantly represented in the genomes. All three functional components showed similar clustering relationships in NMDS, with strains in the same phylogenetic cluster (Fig. 1) being much closer to each other (Fig. 5A–C). This indicated that the substrate utilization traits of *Flavobacteriia* species were concordant with the phylogeny. Previous studies also demonstrated that microbial functional capabilities, including oxygenic photosynthesis, methane oxidation and sulfate reduction, were phylogenetically conserved [49,88]. It was noticeable that the genomes were more scattered in the GH NMDS plot compared to peptidase and transporter, which was consistent with higher Bray-Curtis dissimilarities of GH than

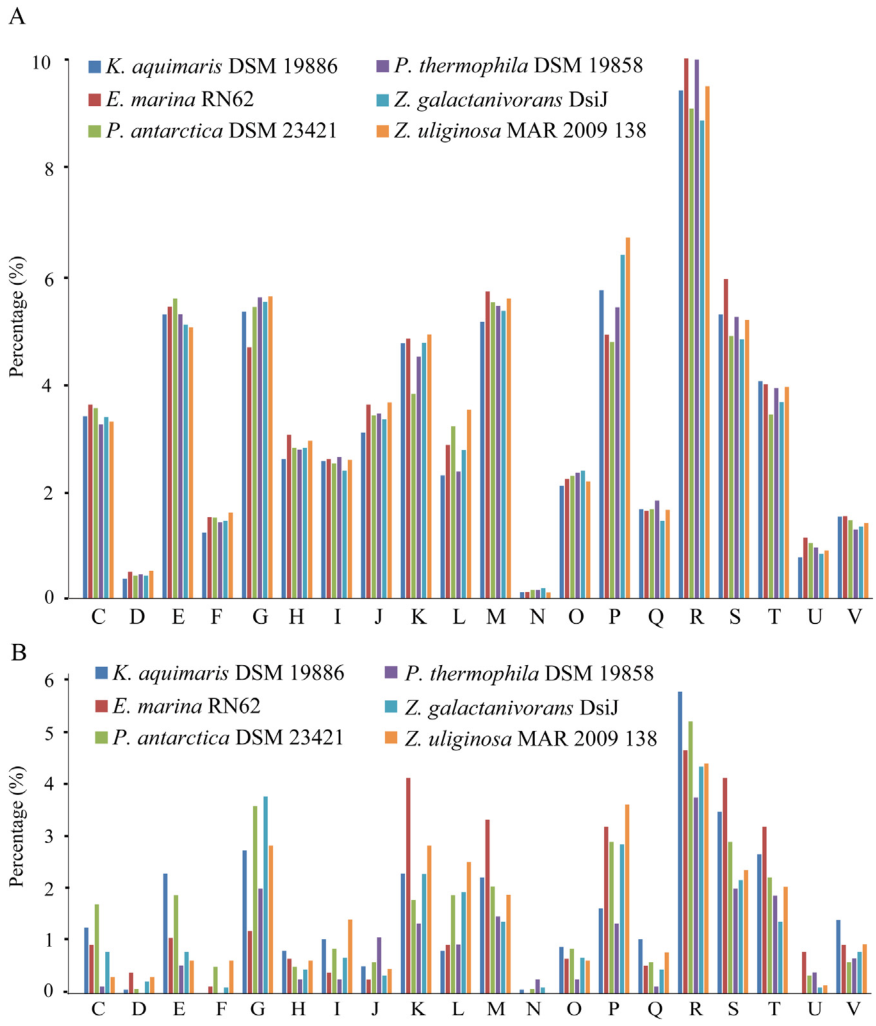

Fig. 6. Comparison of the COG categories in *E. marina* RN62 and its five closely related strains based on the complete genome (A) and specific genes (B). COG categories: (C) energy production and conversion; (D) cell cycle control, cell division, chromosome partitioning; (E) amino acid transport and metabolism; (F) nucleotide transport and metabolism; (G) carbohydrate transport and metabolism; (H) coenzyme transport and metabolism; (I) lipid transport and metabolism; (J) translation, ribosomal structure and biogenesis; (K) transcription; (L) replication, recombination and repair; (M) cell wall/membrane/envelope biogenesis; (N) cell motility; (O) post-translational modification, protein turnover, and chaperones; (P) inorganic ion transport and metabolism; (Q) secondary metabolites biosynthesis, transport, and catabolism; (R) general function prediction only; (S) function unknown; (T) signal transduction mechanisms; (U) intracellular trafficking, secretion, and vesicular transport; (V) defense mechanisms.

those of transporter and peptidase ($P < 0.001$, Wilcox ranking test, Fig. 5D). The analysis was further expanded to 344 *Flavobacteriia* genomes comprising most genera (Fig. 5E). The results showed similar variations, with GH being more divergent between different species than peptidase and transporter. Therefore, it could be inferred that *Flavobacteriia* species had conserved similar protein utilization properties during evolution, and that variations in polysaccharide utilization may be an important strategy utilized by *Flavobacteriia* species to accommodate environmental changes, in this case by RN62, in order to deal with the compositional alterations of organic carbon in the hadal ocean.

Other features for deep-sea adaptation

A proportion of 70.3% of the genes in RN62 could be classified into a COG category, which was comparable with the other five strains (67.1–71.2%). RN62 clearly had a lower relative proportion of COG G (carbohydrate transport and metabolism) than the other five strains, which was consistent with the lower abundance of CAZymes, especially GHs. In contrast, *P. antarctica* had a similar abundance of CAZymes to RN62 but possessed a high abundance of genes for COG G, indicating different carbohydrate utilization pathways in RN62 and *P. antarctica*. A total of 747 genes were RN62 specific, of which 31.7% could be classified in the COG category, which was a little higher than other strains that ranged from 18.9-31.4%. The few RN62-specific COG genes included a galactarate dehydratase (*garD*), which together with its unique 5-dehydro-4-deoxyglucarate dehydratase (*kdgD*) and the common 2,5-dioxopentanoate dehydrogenase (*araE*), may facilitate a unique pathway for converting D-galactarate to 2-oxoglutarate in the citrate cycle. On the other hand, the RN62-specific genes were more enriched in COG M (cell wall/membrane/envelope biogenesis), COG K (transcription) and COG T (signal transduction mechanisms) (Fig. 6, Table S4). The increase of COG M genes has also been reported in an ecotype of *Alteromonas macleodii* specific to deep water [22] and alphaproteobacterial SAR11 deep subclade Ic [79] compared to their respective surface counterparts, implying a universal role for COG M in deep-water inhabitation. The specific COG M included eight glycosyltransferases, two polysaccharide biosynthetic proteins and one polysaccharide exporter, which suggested enhanced capability of polysaccharide synthesis in RN62. Polysaccharide plays vital roles in microbial survival in the deep sea by serving as a cryoprotectant [51]. COG K is also more abundant in bacterial species isolated from deep water and sediment than from shallow water [59], as well as in an ALOHA metagenome at a water depth of 4000 m compared to surface water [31]. In addition, the potentially enhanced transcription and signal transduction performances may relate to *in situ* bacterial growth and substrate utilization.

Conclusion

In summary, this study reports the first genome of a hadal water-derived *Flavobacteriia* bacterium, and the first genome sequence of the genus *Euzebyella*. The genome of *E. marina* RN62 showed abundant GHs and sulfatases in a higher number than many other *Flavobacteriia* species and deep-sea bacteria, demonstrating its capacity to utilize carbohydrates. RN62 harbored a high abundance of putative α-amylases targeting glycogen, and sulfatases for heparan sulfate degradation, which may suggest that the growth of strain RN62 is supported by animal polysaccharides in hadal water. This carbohydrate metabolism capacity of RN62 may reflect changes in the chemical composition of organic matter between the surface and hadal ocean. Strikingly, RN62 had the potential for intracellular glycogen biosynthesis and degradation in addition to α-amylases, which may provide a means for carbon storage and utilization when organic matter is repleted and depleted, respectively. Additionally, the genetic potential for carbohydrate metabolism, represented by the GH compositions, varied among *Flavobacteriia* species to a certain degree, which were higher than the peptidase and transporter compositions, suggesting that alternation in carbohydrate metabolism may be a preferential strategy for *Flavobacteriia* to occupy new ecological niches. The study raised important questions that warrant further investigation, such as whether or not the intracellular cycling of organic matter represents a common strategy used by microbes to survive in the hadal environment or whether habitat type and phylogeny interact to drive the variation of carbohydrate metabolism in *Flavobacteriia*.

Acknowledgements

We would like to thank all the scientists and crews on the *R/V* Dong Fang Hong 2 for their assistance with sampling during the cruise. This work was funded by the National Natural Science Foundation of China (91751202, 41730530, 41476112, 41506154), Fundamental Research Funds from the Central Universities (201762009), and the Scientific and Technological Innovation Project of Qingdao National Laboratory for Marine Science and Technology (2016ASKJ14).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.syapm.2019.01.002.

References

[1] Alonso, C., Warnecke, F., Amann, R., Pernthaler, J. (2007) High local and global diversity of *Flavobacteria* in marine plankton. Environ. Microbiol. 9, 1253–1266.

[2] Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. (1990) Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

[3] Bao, R., Strasser, M., McNichol, A.P., Haghipour, N., McIntyre, C., Wefer, G., Eglinton, T.I. (2018) Tectonically-triggered sediment and carbon export to the Hadal zone. Nat. Commun. 9, 121.

[4] 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, e0164846.

[5] Barbeyron, T., Thomas, F., Barbe, V., Teeling, H., Schenowitz, C., Dossat, C., Goesmann, A., Leblanc, C., Oliver Glöckner, F., Czjzek, M. (2016) Habitat and taxon as driving forces of carbohydrate catabolism in marine heterotrophic bacteria: example of the model algae-associated bacterium *Zobellia galactanivorans* Dsij^T. Environ. Microbiol. 18, 4610–4627.

[6] Bauer, M., Kube, M., Teeling, H., Richter, M., Lombardot, T., Allers, E., Würdemann, C.A., Quast, C., Kuhl, H., Knaust, F. (2006) Whole genome analysis of the marine *Bacteroidetes* '*Gramella forsetii*' reveals adaptations to degradation of polymeric organic matter. Environ. Microbiol. 8, 2201–2213.

[7] Bennke, C.M., Krüger, K., Kappelmann, L., Huang, S., Gobet, A., Schüler, M., Barbe, V., Fuchs, B.M., Michel, G., Teeling, H. (2016) Polysaccharide utilisation loci of *Bacteroidetes* from two contrasting open ocean sites in the North Atlantic. Environ. Microbiol. 18, 4456–4470.

[8] Bernardet, J., Nakagawa, Y., Holmes, B. (2002) Proposed minimal standards for describing new taxa of the family *Flavobacteriaceae* and emended description of the family. Int. J. Syst. Evol. Microbiol. 52, 1049–1070.

[9] Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T. (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973.

[10] Contreras-Moreira, B., Vinuesa, P. (2013) GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. Appl. Environ. Microb. 79, 7696–7701.

[11] Darriba, D., Taboada, G.L., Doallo, R., Posada, D. (2011) ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics. 27, 1164–1165.

[12] Deniaud, E. (2003) Interactions of the mix-linked β-(1,3)/β-(1,4)-D-xylans in the cell walls of *Palmaria palmata* (rhodophyta). J. Phycol. 39, 74–82.

[13] Ferguson, A.D., Hofmann, E., Coulton, J.W., Diederichs, K., Welte, W. (1998) Siderophore-mediated iron transport: crystal structure of FhuA with bound lipopolysaccharide. Science 282, 2215–2220.

[14] Fernández-Gomez, B., Richter, M., Schüler, M., Pinhassi, J., Acinas, S.G., González, J.M., Pedros-Alio, C. (2013) Ecology of marine *Bacteroidetes*: a comparative genomics approach. ISME J. 7, 1026–1037.

[15] 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., Médigue, C., Czjzek, M., Vallenet, D., McBride, M.J., Duchaud, E., Michel, G. (2017) Carrageenan catabolism is encoded by a complex regulon in marine heterotrophic bacteria. Nat. Commun. 8, 1685.

[16] Finn, R.D., Clements, J., Eddy, S.R. (2011) HMMER web server: interactive sequence similarity searching. Nucleic Acids Res. 39, W29–W37.

[17] Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A. (2015) The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. 44, D279–D285.

[18] Forchhammer, K. (2008) P_{II} signal transducers: novel functional and structural insights. Trends Microbiol. 16, 65–72.

[19] Galperin, M.Y., Makarova, K.S., Wolf, Y.I., Koonin, E.V. (2014) Expanded microbial genome coverage and improved protein family annotation in the COG database. Nucleic Acids Res. 43, D261–D269.

[20] González, J.M., Fernández-Gómez, B., Fernàndez-Guerra, A., Gómez-Consarnau, L., Sánchez, O., Coll-Lladó, M., Del Campo, J., Escudero, L., Rodríguez-Martínez, R., Alonso-Sáez, L. (2008) Genome analysis of the proteorhodopsin-containing marine bacterium *Polaribacter* sp. MED152 (Flavobacteria). Proc. Natl. Acad. Sci. U. S. A. 105, 8724–8729.

[21] González, J.M., Pinhassi, J., Fernández-Gómez, B., Coll-Lladó, M., González-Velázquez, M., Puigbò, P., Jaenicke, S., Gómez-Consarnau, L., Fernàndez-Guerra, A., Goesmann, A. (2011) Genomics of the proteorhodopsin-containing marine flavobacterium *Dokdonia* sp strain MED134. Appl. Environ. Microb. 77, 8676–8686.

[22] Ivars-Martinez, E., Martin-Cuadrado, A., D'Auria, G., Mira, A., Ferriera, S., Johnson, J., Friedman, R., Rodriguez-Valera, F. (2008) Comparative genomics of two ecotypes of the marine planktonic copiotroph *Alteromonas macleodii* suggests alternative lifestyles associated with different kinds of particulate organic matter. ISME J. 2, 1194–1212.

[23] Jamieson, A.J., Fujii, T., Mayor, D.J., Solan, M., Priede, I.G. (2010) Hadal trenches: the ecology of the deepest places on Earth. Trends Ecol. Evol. 25, 190–197.

[24] Jamieson, A.J., Malkocs, T., Piertney, S.B., Fujii, T., Zhang, Z. (2017) Bioaccumulation of persistent organic pollutants in the deepest ocean fauna. Nat. Ecol. Evol. 1, 0051.

[25] Jiao, N., Cai, R., Zheng, Q., Tang, K., Liu, J., Jiao, F., Wallace, D., Chen, F., Li, C., Amann, R. (2018) Unveiling the enigma of refractory carbon in the ocean. Natl. Sci. Rev. 5, 459–463.

[26] Juncker, A.S., Willenbrock, H., Von Heijne, G., Brunak, S., Nielsen, H., Krogh, A. (2003) Prediction of lipoprotein signal peptides in Gram-negative bacteria. Protein Sci. 12, 1652–1662.

[27] Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., Morishima, K. (2016) KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 45, D353–D361.

[28] Kappelmann, L., Krüger, K., Hehemann, J., Harder, J., Markert, S., Unfried, F., Becher, D., Shapiro, N., Schweder, T., Amann, R.I. (2019) Polysaccharide utilization loci of North Sea *Flavobacteriia* as basis for using SusC/D-protein expression for predicting major phytoplankton glycans. ISME J. 13, 76–91.

[29] Katoh, K., Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.

[30] Kenichiro, I., Naoyuki, A., Toshiyuki, T., Fumitsugu, A., Ayano, I., Masashi, S., Shigeki, M. (2008) SOSUI-GramN: high performance prediction for subcellular localization of proteins in Gram-negative bacteria. Bioinformation 2, 417–421.

[31] Konstantinidis, K.T., Braff, J., Karl, D.M., DeLong, E.F. (2009) Comparative metagenomic analysis of a microbial community residing at a depth of 4000 meters at station ALOHA in the North Pacific subtropical gyre. Appl. Environ. Microb. 75, 5345–5355.

[32] Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H., Phillippy, A.M. (2017) Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. 27, 722–736.

[33] Koropatkin, N.M., Cameron, E.A., Martens, E.C. (2012) How glycan metabolism shapes the human gut microbiota. Nat. Rev. Microbiol. 10, 323–335.

[34] Kristensen, D.M., Kannan, L., Coleman, M.K., Wolf, Y.I., Sorokin, A., Koonin, E.V., Mushegian, A. (2010) A low-polynomial algorithm for assembling clusters of orthologous groups from intergenomic symmetric best matches. Bioinformatics 26, 1481–1487.

[35] Kwon, Y.M., Kim, K.W., Kim, J.Y.H., Choi, T., Yang, S., Oh, C.H., Kwon, K.K., Kim, S. (2018) *Euzebyella algicola* sp. nov., a marine bacterium of the family *Flavobacteriaceae*, isolated from green algae. Int. J. Syst. Evol. Microbiol. 68, 764–768.

[36] Labourel, A., Jam, M., Jeudy, A., Hehemann, J., Czjzek, M., Michel, G. (2014) The β-glucanase ZgLamA from *Zobellia galactanivorans* evolved a bent active site adapted for efficient degradation of algal laminarin. J. Biol. Chem. 289, 2027–2042.

[37] Lagesen, K., Hallin, P., Rødland, E.A., Stærfeldt, H., Rognes, T., Ussery, D.W. (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. 35, 3100–3108.

[38] Landry, Z., Swan, B.K., Herndl, G.J., Stepanauskas, R., Giovannoni, S.J. (2017) SAR202 genomes from the dark ocean predict pathways for the oxidation of recalcitrant dissolved organic matter. mBio 8, e00413–17.

[39] Larsbrink, J., Zhu, Y., Kharade, S.S., Kwiatkowski, K.J., Eijsink, V.G., Koropatkin, N.M., McBride, M.J., Pope, P.B. (2016) A polysaccharide utilization locus from *Flavobacterium johnsoniae* enables conversion of recalcitrant chitin. Biotechnol. Biofuels 9, 260.

[40] Li, L., Stoeckert, C.J., Roos, D.S. (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res. 13, 2178–2189.

[41] Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 42, D490–D495.

[42] Lowe, T.M., Eddy, S.R. (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25, 955–964.

[43] Lucena, T., Pascual, J., Giordano, A., Gambacorta, A., Garay, E., Arahal, D.R., Macián, M.C., Pujalte, M.J. (2010) *Euzebyella saccharophila* gen. nov., sp. nov., a marine bacterium of the family *Flavobacteriaceae*. Int. J. Syst. Evol. Microbiol. 60, 2871–2876.

[44] Mackenzie, A.K., Pope, P.B., Pedersen, H.L., Gupta, R., Morrison, M., Willats, W.G.T., Eijsink, V. (2012) Two SusD-like proteins encoded within a polysaccharide utilization locus of an uncultured ruminant Bacteroidetes phylotype bind strongly to cellulose. Appl. Environ. Microb. 78, 5935–5937.

[45] Mann, A.J., Hahnke, R.L., Huang, S., Werner, J., Xing, P., Barbeyron, T., Huettel, B., Stüber, K., Reinhardt, R., Harder, J. (2013) The genome of the alga-associated marine *Flavobacterium Formosa agariphila* KMM 3901^T reveals a broad potential for degradation of algal polysaccharides. Appl. Environ. Microb. 79, 6813–6822.

[46] Marmur, J. (1961) A procedure for the isolation of deoxyribonucleic acid from micro-organisms. J. Mol. Biol. 3, 275–282.

[47] Martens, E.C., Lowe, E.C., Chiang, H., Pudlo, N.A., Wu, M., McNulty, N.P., Abbott, D.W., Henrissat, B., Gilbert, H.J., Bolam, D.N. (2011) Recognition and degradation of plant cell wall polysaccharides by two human gut symbionts. PLoS Biol. 9, e1001221.

[48] Martin-Cuadrado, A., Ghai, R., Gonzaga, A., Rodriguez-Valera, F. (2009) CO dehydrogenase genes found in metagenomic fosmid clones from the deep Mediterranean Sea. Appl. Environ. Microb. 75, 7436–7444.

[49] Martiny, A.C., Treseder, K., Pusch, G. (2013) Phylogenetic conservatism of functional traits in microorganisms. ISME J. 7, 830–838.

[50] Morya, V.K., Kim, J., Kim, E. (2012) Algal fucoidan: structural and size-dependent bioactivities and their perspectives. Appl. Microbiol. Biot. 93, 71–82.

[51] Nichols, C.M., Guezennec, J., Bowman, J.P. (2005) Bacterial exopolysaccharides from extreme marine environments with special consideration of the southern ocean, sea ice, and deep-sea hydrothermal vents: a review. Mar. Biotechnol. 7, 253–271.

[52] Nunoura, T., Takaki, Y., Hirai, M., Shimamura, S., Makabe, A., Koide, O., Kikuchi, T., Miyazaki, J., Koba, K., Yoshida, N. (2015) Hadal biosphere: insight into the microbial ecosystem in the deepest ocean on Earth. Proc. Natl. Acad. Sci. U. S. A. 112, E1230–E1236.

[53] Overbeek, R., Olson, R., Pusch, G.D., Olsen, G.J., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Parrello, B., Shukla, M. (2013) The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res. 42, D206–D214.

[54] Pedrotti, M.L., Beauvais, S., Kerros, M., Iversen, K., Peters, F. (2009) Bacterial colonization of transparent exopolymeric particles in mesocosms under different turbulence intensities and nutrient conditions. Aquat. Microb. Ecol. 55, 301–312.

[55] Peoples, L.M., Donaldson, S., Osuntokun, O., Xia, Q., Nelson, A., Blanton, J., Allen, E.E., Church, M.J., Bartlett, D.H. (2018) Vertically distinct microbial communities in the Mariana and Kermadec trenches. PLoS One 13, e0195102.

[56] Petersen, T.N., Brunak, S., von Heijne, G., Nielsen, H. (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods 8, 785–786.

[57] Pinhassi, J., Sala, M.M., Havskum, H., Peters, F., Guadayol, O., Malits, A., Marrasé, C. (2004) Changes in bacterioplankton composition under different phytoplankton regimens. Appl. Environ. Microbiol. 70, 6753–6766.

[58] Popper, Z.A., Michel, G., Hervé, C., Domozych, D.S., Willats, W.G., Tuohy, M.G., Kloareg, B., Stengel, D.B. (2011) Evolution and diversity of plant cell walls: from algae to flowering plants. Annu. Rev. Plant Biol. 62, 567–590.

[59] Qin, Q., Li, Y., Zhang, Y., Zhou, Z., Zhang, W., Chen, X., Zhang, X., Zhou, B., Wang, L., Zhang, Y. (2011) Comparative genomics reveals a deep-sea sediment-adapted life style of *Pseudoalteromonas* sp. SM9913. ISME J. 5, 274–284.

[60] Rawlings, N.D., Barrett, A.J., Finn, R. (2016) Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res. 44, D343–D350.

[61] Reeves, A.R., D'Elia, J.N., Frias, J., Salyers, A.A. (1996) A *Bacteroides thetaiotaomicron* outer membrane protein that is essential for utilization of maltooligosaccharides and starch. J. Bacteriol. 178, 823–830.

[62] Reeves, A.R., Wang, G.R., Salyers, A.A. (1997) Characterization of four outer membrane proteins that play a role in utilization of starch by *Bacteroides thetaiotaomicron*. J. Bacteriol. 179, 643–649.

[63] Saier, M.H., Jr., Reddy, V.S., Tsu, B.V., Ahmed, M.S., Li, C., Moreno-Hagelsieb, G. (2016) The transporter classification database (TCDB): recent advances. Nucleic Acids Res. 44, D372–D379.

[64] Salazar, G., Cornejo-Castillo, F.M., Benítez-Barrios, V., Fraile-Nuez, E., Álvarez-Salgado, X.A., Duarte, C.M., Gasol, J.M., Acinas, S.G. (2016) Global diversity and biogeography of deep-sea pelagic prokaryotes. ISME J. 10, 596.

[65] Salmeán, A.A., Guillouzo, A., Duffieux, D., Jam, M., Matard-Mann, M., Larocque, R., Pedersen, H.L., Michel, G., Czjzek, M., Willats, W.G.T., Hervé, C. (2018) Double blind microarray-based polysaccharide profiling enables parallel identification of uncharacterized polysaccharides and carbohydrate-binding proteins with unknown specificities. Sci. Rep. 8, 2500.

[66] Schäffer, A.A., Aravind, L., Madden, T.L., Shavirin, S., Spouge, J.L., Wolf, Y.I., Koonin, E.V., Altschul, S.F. (2001) Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements. Nucleic Acids Res. 29, 2994–3005.

[67] Smith, K.L., Ruhl, H.A., Bett, B.J., Billett, D., Lampitt, R.S., Kaufmann, R.S. (2009) Climate, carbon cycling, and deep-ocean ecosystems. Proc. Natl. Acad. Sci. U. S. A., 19211–19218.

[68] Sonnenburg, E.D., Zheng, H., Joglekar, P., Higginbottom, S.K., Firbank, S.J., Bolam, D.N., Sonnenburg, J.L. (2010) Specificity of polysaccharide use in intestinal *Bacteroides* species determines diet-induced microbiota alterations. Cell 141, 1241–1252.

[69] Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.

[70] Sun, C., Fu, G., Zhang, C., Hu, J., Xu, L., Wang, R., Su, Y., Han, S., Yu, X., Cheng, H. (2016) Isolation and complete genome sequence of *Algibacter alginolytica* sp. nov., a novel seaweed-degrading *Bacteroidetes* bacterium with diverse putative polysaccharide utilization loci. Appl. Environ. Microb. 82, 2975–2987.

[71] Tang, K., Jiao, N., Liu, K., Zhang, Y., Li, S. (2012) Distribution and functions of TonB-dependent transporters in marine bacteria and environments: implications for dissolved organic matter utilization. PLoS One 7, e41204.

[72] Tang, K., Lin, Y., Han, Y., Jiao, N. (2017) Characterization of potential polysaccharide utilization systems in the marine *Bacteroidetes Gramella flava* JLT2011 using a multi-omics approach. Front. Microbiol. 8, 220.

[73] Tarn, J., Peoples, L.M., Hardy, K., Cameron, J., Bartlett, D.H. (2016) Identification of free-living and particle-associated microbial communities present in hadal regions of the Mariana Trench. Front. Microbiol. 7, 665.

[74] Teeling, H., Fuchs, B.M., Becher, D., Klockow, C., Gardebrecht, A., Bennke, C.M., Kassabgy, M., Huang, S., Mann, A.J., Waldmann, J. (2012) Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. Science 336, 608–611.

[75] Teeling, H., Fuchs, B.M., Bennke, C.M., Krueger, K., Chafee, M., Kappelmann, L., Reintjes, G., Waldmann, J., Quast, C., Gloeckner, F.O. (2016) Recurring patterns in bacterioplankton dynamics during coastal spring algae blooms. ELife 5, e11888.

[76] Temple, M.J., Cuskin, F., Baslé, A., Hickey, N., Speciale, G., Williams, S.J., Gilbert, H.J., Lowe, E.C. (2017) A Bacteroidetes locus dedicated to fungal 1,6-β-glucan degradation: unique substrate conformation drives specificity of the key endo-1,6-β-glucanase. J. Biol. Chem. 292, 10639–10650.

[77] Thomas, F., Hehemann, J., Rebuffet, E., Czjzek, M., Michel, G. (2011) Environmental and gut *Bacteroidetes*: the food connection. Front. Microbiol. 2, 93.

[78] Thompson, J.N., Stoolmiller, A.C., Matalon, R., Dorfman, A. (1973) *N*-acetyl-β-hexosaminidase: role in the degradation of glycosaminoglycans. Science 181, 866–867.

[79] Thrash, J.C., Temperton, B., Swan, B.K., Landry, Z.C., Woyke, T., DeLong, E.F., Stepanauskas, R., Giovannoni, S.J. (2014) Single-cell enabled comparative genomics of a deep ocean SAR11 bathytype. ISME J. 8, 1440–1451.

[80] West, N.J., Obernosterer, I., Zemb, O., Lebaron, P. (2008) Major differences of bacterial diversity and activity inside and outside of a natural iron-fertilized phytoplankton bloom in the Southern Ocean. Environ. Microbiol. 10, 738–756.

[81] Williams, T.J., Wilkins, D., Long, E., Evans, F., DeMaere, M.Z., Raftery, M.J., Cavicchioli, R. (2013) The role of planktonic *Flavobacteria* in processing algal organic matter in coastal East Antarctica revealed using metagenomics and metaproteomics. Environ. Microbiol. 15, 1302–1317.

[82] Xing, P., Hahnke, R.L., Unfried, F., Markert, S., Huang, S., Barbeyron, T., Harder, J., Becher, D., Schweder, T., Glöckner, F.O. (2015) Niches of two polysaccharide-

degrading *Polaribacter* isolates from the North Sea during a spring diatom bloom. ISME J. 9, 1410–1422.

[83] Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F., Xu, Y. (2012) dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. 40, W445–W451.

[84] Yoon, S., Ha, S., Kwon, S., Lim, J., Kim, Y., Seo, H., Chun, J. (2017) Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. Int. J. Syst. Evol. Microbiol. 67, 1613–1617.

[85] 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. (2010) PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics 26, 1608–1615.

[86] Yu, Y., Li, H., Zeng, Y., Sun, K., Chen, B. (2012) *Pricia antarctica* gen. nov., sp. nov., a member of the family *Flavobacteriaceae*, isolated from Antarctic intertidal sediment. Int. J. Syst. Evol. Microbiol. 62, 2218–2223.

[87] Zhang, Y., Zhao, J., Zhang, X., Chen, G., Zhou, M., Mo, X., He, H., Chen, S. (2017) *Euzebyella marina* sp. nov., isolated from seawater. Int. J. Syst. Evol. Microbiol. 67, 920–924.

[88] Zimmerman, A.E., Martiny, A.C., Allison, S.D. (2013) Microdiversity of extracellular enzyme genes among sequenced prokaryotic genomes. ISME J. 7, 1187–1199.