Marine Life Science & Technology
https://doi.org/10.1007/s42995-025-00314-9

RESEARCH PAPER

Metagenomic analysis reveals genetic coupling between TonB-dependent transporters and extracellular enzymes in coastal bacterial communities

Shujing Liu[1] · Quanrui Chen[1] · Xuanyun Qiu[1] · Wenhao Li[1] · Kai Tang[1]

Received: 25 January 2025 / Accepted: 18 June 2025
© Ocean University of China 2025

Abstract
Marine heterotrophic prokaryotes initially release extracellular enzymes to cleave large organic molecules and then take up ambient substrates via transporters. Given the direct influence of extracellular enzymes on nutrient availability, understanding their diversity and dynamics is crucial in comprehending microbial interactions and organic matter cycling in aquatic ecosystems. In this study, metagenomics was employed to investigate the functional diversity and dynamics of extracellular enzymes and transporters in coastal waters over a 22-day period. The metagenome-derived gene pool of organic matter-degrading secretory enzymes and transporters was primarily contributed by three major bacterial classes. Bacteroidota were the primary contributors to the gene pool of secretory carbohydrate-active enzymes (CAZymes), whereas Gammaproteobacteria contribute more to secretory peptidases and TonB-dependent transporters (TBDTs), and Alphaproteobacteria to ATP-binding cassette (ABC) transporters. The distinct substrate targets of the enzymes and transporters combined with the unique dynamics of these taxa across depth layers suggest that organic matter degradation and uptake machinery played a role in ecological niche partitioning. At the community level, the abundance of TBDT genes was more positively correlated with extracellular enzymes than ABC transporters. To further explore taxon-specific differences, we reconstructed 163 bacterial and archaeal metagenome-assembled genomes (MAGs). Correlation patterns at the MAG level varied across taxa: Bacteroidota MAGs exhibited significant positive correlations between TBDTs and extracellular enzymes, whereas Gammaproteobacteria and Alphaproteobacteria MAGs showed weak or no significant correlations. These results suggest the diversity of ecological strategies among marine heterotrophic bacteria and highlight a potential coregulation or functional linkage between extracellular enzymes and TBDTs in the metabolism of marine heterotrophic prokaryotes. Our study advances the understanding of the microbial adaptations driving carbon and nutrient cycling.

Keywords Bacteria · CAZymes · Peptidases · Transporters · Organic matter assimilation

Introduction

Organic matter cycling in the ocean shapes the global climate and environmental systems while sustaining marine biological productivity and ecological balance. Marine organic matter is primarily derived from phytoplankton and is released into the environment as dissolved or particulate organic matter through processes, such as photosynthetic overflow, excretion, cell lysis, aging, and protective mucilage secretion (Boyd et al. 2019). Meanwhile, the assimilation, transformation, and redistribution of organic matter are predominantly mediated by heterotrophic prokaryotes (Azam and Malfatti 2007).

High-molecular-weight (HMW) organic matter, such as proteins, peptides, and humic acids, serves as an important

Edited by Chengchao Chen.

Special Topic: Ecology & Environmental Biology.

✉ Kai Tang
tangkai@xmu.edu.cn

[1] State Key Laboratory of Marine Environmental Science, Fujian Key Laboratory of Marine Carbon Sequestration, College of Ocean and Earth Sciences, Xiamen University, Xiamen 361005, China

Published online: 29 July 2025

carbon and energy source for heterotrophic prokaryotes. To cross cellular membranes, these compounds must first be hydrolyzed into molecules smaller than 600 Da (Arnosti 2011). Extracellular enzymatic hydrolysis is the primary mechanism through which microorganisms gain access to HMW organic matter (Chrost and Velimirov 2013). The substrate specificity of extracellular enzymes determines the type of organic matter that can be degraded and the rate of degradation (Sinsabaugh et al. 2008). Furthermore, the substrate specificity of membrane-bound transporters dictates the uptake of these substrates once they are small enough to be transported into the cell (Davidson Amy et al. 2008). Although extracellular enzymes and transporters are key drivers of organic matter cycling, they have been historically studied in isolation (Zhao et al. 2024, 2020). Minimal attention has been paid to the potential evolutionary or genomic linkages between them.

ATP-binding cassette (ABC) transporters and TonB-dependent transporters (TBDTs) are the primary transport systems mediating the active exchange of substrates between prokaryotic cells and their surrounding marine environment (Poretsky et al. 2010). ABC transporters facilitate the movement of diverse substrates, including amino acids, sugars, peptides, phosphonates, metal–chelate complexes, vitamins, and some ions, across cell membranes; they also play a crucial role in cellular nutrient and ion regulation (Francis et al. 2021). TBDTs specialize in transporting bulky compounds, such as mono- and oligosaccharides, siderophores, and vitamin B12 (Tang et al. 2012). The secretion of extracellular enzymes and the subsequent uptake of substrates by membrane transporters form the foundation of prokaryotic activity in marine ecosystems. However, a key unresolved question is whether marine microorganisms exhibit genetic coupling, such as coregulated gene clusters or phylogenetically conserved co-occurrence patterns, that synchronizes enzyme production with transporter deployment for optimal substrate acquisition.

Despite the crucial role of extracellular enzymes and transporters, their taxonomic affiliations and functional diversity throughout the coastal water column remain poorly understood. In this study, we analyzed a time-series collection of metagenomic samples obtained from the East China Sea within a 22-day period to explore the composition of the extracellular enzymes and transporters and the microbial taxa producing them. Owing to the prominent mixing of oceanic and coastal waters in this body of water, the East China Sea is characterized by a dynamic organic matter pool (Zhang 2015), making it an interesting region to investigate the adaptive response of microbial extracellular enzymes and transporters. We investigated the abundance, diversity, and functional classification of genes encoding secretory carbohydrate-active enzymes (CAZymes), peptidases, ABC transporters, and TBDTs across 93 metagenomes. To gain deep insights into genomic traits, we conducted metagenomic binning to recover metagenome-assembled genomes (MAGs) from these samples and then performed a comprehensive characterization of extracellular enzymes and transporters at the population level. Our findings reveal a positive relationship between the abundance of extracellular enzymes and TBDTs in a dynamic marine environment, providing new insights into their coupling in coastal ecosystems.

Materials and methods

Sample collection, DNA extraction, and metagenomic sequencing

Metagenomic samples were collected at two sites in the East China Sea at eight different time points between July 11 and August 4, 2022 (Fig. 1). At site A1 (30.12079°N, 122.72947°E), water samples were collected from three layers: surface (3 m), middle (17 m), and bottom (40 m). At site C4 (30.1282°N, 122.74599°E), samples were obtained from the surface (5 m), middle (25 m), and bottom (60 m) layers. Seawater samples were collected using 10-L Niskin bottles attached to a conductivity–temperature–depth (CTD) rosette sampler. Environmental parameters, including temperature, pH, dissolved oxygen, and salinity, were measured using CTD sensors. For each water sample, approximately

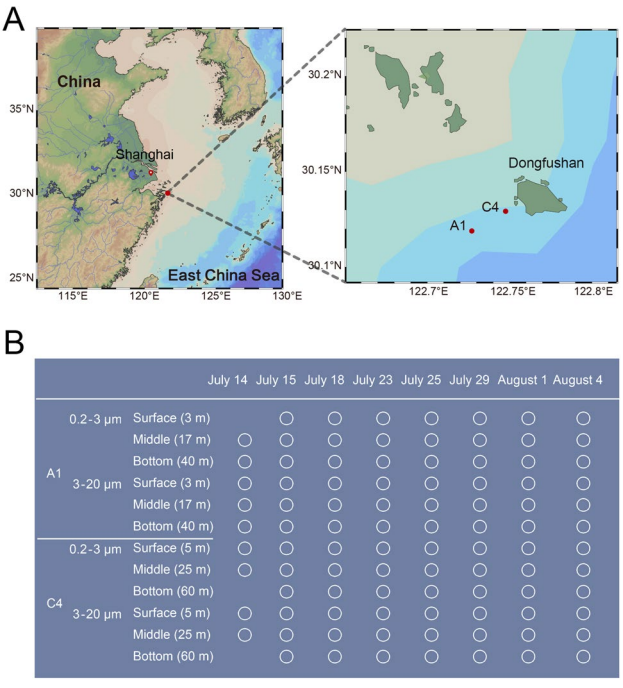


Fig. 1 Sites and dates of metagenomic sampling in the East China Sea. The dots represent the sampling sites. The circles in the image represent the sampling events

15 L of seawater was prefiltered through a 20 μm mesh and then sequentially filtered using 3 and 0.2 μm polycarbonate membranes (Millipore, USA) to obtain size-fractionated biomass. Microbial DNA was extracted individually from 3 to 0.2 μm filters using the DNeasy Powersoil Kit (MoBio, USA). Metagenomic sequencing was performed on the NovaSeq 6000 platform with Illumina PE150 chemistries at Hanyu Bio-Tech (Shanghai, China).

Enzyme and transporter annotation

Raw reads were processed to eliminate adaptors and low-quality sequences using Trimmomatic v0.36 for quality control (Bolger et al. 2014). The clean reads were assembled using Megahit v1.1.3 (Li et al. 2016), and the genes in the contigs were predicted using Prodigal v2.6.3 (-p meta) (Hyatt et al. 2010). A nonredundant gene catalog was constructed using CD-HIT v4.8.1 (-c 0.95, -aS 0.9) (Fu et al. 2012). The genes were annotated using DIAMOND v2.0.7 (Buchfink et al. 2015) by aligning them against the Kyoto Encyclopedia of Genes and Genomes database (Kanehisa et al. 2016). Gene abundances were quantified using Samtools v1.7 (Li et al. 2009). The raw read counts were normalized using transcripts per million to account for the differences in sequencing depth and gene length (Wagner et al. 2012). This normalization approach adjusts for variations in sequencing effort and enables direct comparisons of gene abundances across samples (Wagner et al. 2012). CAZymes were annotated using hmmsearch against the CAZymes database, applying a threshold of e value $< 1 \times 10^{-10}$ and coverage > 0.3 (Drula et al. 2021). Peptidases were annotated using DIAMOND v2.0.7 (Buchfink et al. 2015) by searching against the MEROPS database using cutoffs of e value $< 1 \times 10^{-10}$ (Rawlings et al. 2017). SignalP v6.0 was employed to identify the signal peptides in the predicted gene sequences (Teufel et al. 2022). Transporters were identified with DIAMOND v2.0.7 (Buchfink et al. 2015) by searching against the transporter classification database (TCDB) using cutoffs of e value $< 1 \times 10^{-10}$ (3.A.1 for ABC transporters and 2.C.1 for TBDTs) (Saier et al. 2021). The taxonomy of gene sequences was assigned by searching against the nonredundant database using DIAMOND v2.0.7; the single best hit was selected as the taxonomic annotation for each query sequence (Buchfink et al. 2015). Substrate predictions of ABC transporters were retrieved from TCDB (Saier et al. 2021), and substrate targets of TBDTs were predicted by searching a curated TBDT sequence database (Tang et al. 2012; Zhao et al. 2024). The same analytical pipeline was applied to two additional marine datasets: transcriptomic data from our Bohai Sea dataset (Yuan et al. 2021) and publicly available metagenomic data from the northeastern Indian Ocean (Wang et al. 2024b).

MAG assembly, annotation, and abundance

MAGs from the metagenomic data were recovered using previously described methods (Liu et al. 2021). Clean reads were aligned to contigs longer than 1,500 bp using Bowtie2 v2.3.2 (Langmead and Salzberg 2012), and the resulting BAM files were sorted with Samtools v1.7 (Li et al. 2009). Binning analysis was performed using MetaBAT2 v2.10.2 by applying eight combinations of specificity and sensitivity parameters (–maxP set to either 60 or 95, –minS set to either 60 or 95, and –maxEdges set to either 200 or 500) (Kang et al. 2019). The binning results were refined and combined using DAS_Tool v1.0 (Sieber et al. 2018). CheckM2 v.1.0.1 was then used to evaluate the quality of these MAGs (Chklovski et al. 2023), and MAGs with completeness $\geq 50\%$ and contamination $\leq 10\%$ were selected to retain medium- and high-quality MAGs, according to the MIMAG standards (Bowers et al. 2017). The nonredundant MAG dataset was generated using dRep v3.4.2 at 95% average nucleotide identity (–completeness 75) (Olm et al. 2017). The genes of MAGs were predicted using Prodigal v2.6.3 (Hyatt et al. 2010). The CAZymes and peptidases of MAGs were also annotated using the CAZymes database (Drula et al. 2021) and MEROPS database (Rawlings et al. 2017), respectively. The taxonomic classification of the MAGs was determined using GTDB-Tk v2.1.1 (Chaumeil et al. 2019) with the GTDB r2.1.4 reference database (Rinke et al. 2021). A phylogenetic tree of the MAGs was constructed from the GTDB conserved marker genes (120 for bacteria and 122 for archaea) using FastTree v2.1.10 (Price et al. 2010) and visualized with ChiPlot (https://www.chiplot.online/) (Xie et al. 2023).

Statistical analysis and data visualization

A sampling map was generated using Ocean Data View v5.7.2, a widely used software for oceanographic data visualization (Schlitzer 2002). Other statistics and figures were generated in the R language. Statistical analyses and data visualization were conducted using the R packages "stats", "ggplot2", "ggpubr", and "car".

Results

Temporal variations in the metagenomic functional gene composition

To assess the effect of sampling time, depth, and site on metagenomic composition, we performed principal coordinate analysis (PCoA) based on functional genes (KO numbers) and nonredundant genes (Fig. S1). The results indicate that sampling time had the most significant effect, with the

functional gene composition showing the highest variation (adonis $R^2 = 0.318$, $p < 0.001$). By comparison, depth (adonis $R^2 = 0.06$, $p < 0.001$) and site (adonis $R^2 = 0.026$, $p < 0.026$) had smaller but still significant impacts. Analysis of the nonredundant gene set revealed that sampling time had a stronger effect (adonis $R^2 = 0.1655$, $p < 0.001$) than depth (adonis $R^2 = 0.0938$, $p < 0.001$) and site (adonis $R^2 = 0.0183$, $p < 0.05$). These findings suggest that temporal fluctuations shape microbial functional profiles. Spearman correlation analysis was performed among environmental factors (temperature, salinity, pH, and dissolved oxygen), major families associated with metabolism-related functional genes, and corresponding gene abundances to explore the influence of environmental parameters on microbial metabolic potential (Fig. S1). Among the metabolism-associated families, the relative abundances of Paracoccaceae, Flavobacteriaceae, Halieaceae, Nitrosopumilaceae, Cryomorphaceae, and Synechococcaceae showed significant correlations with salinity, temperature, and dissolved oxygen ($p < 0.05$). pH was negatively correlated with the abundances of Nitrosopumilaceae and Erythrobacteraceae. At the functional level, only the genes involved in xenobiotic biodegradation and metabolism exhibited significant correlations with salinity, temperature, and dissolved oxygen ($p < 0.05$). No significant associations were observed between the abundances of other metabolic gene categories and environmental variables.

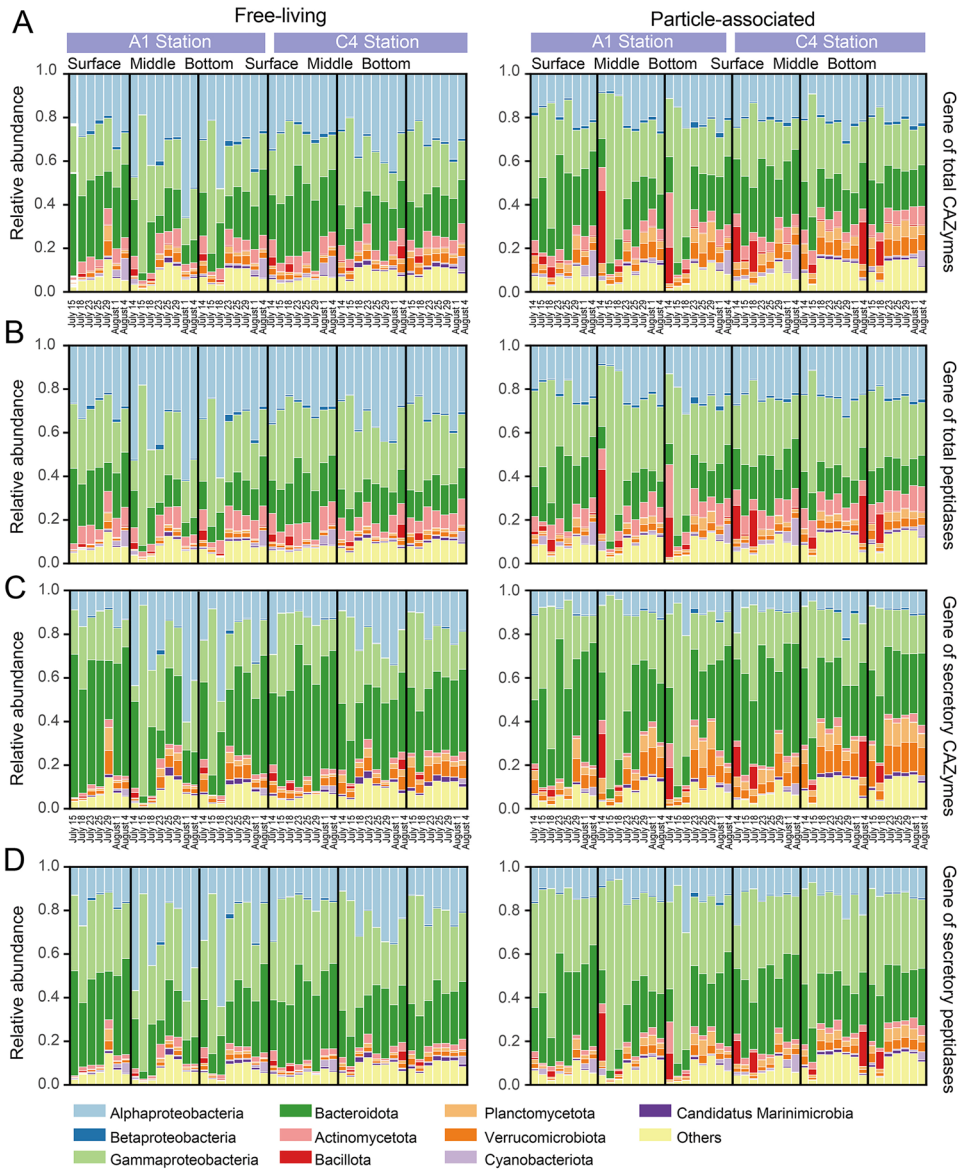


Fig. 2 Taxonomic breakdown at the phylum level (class level for Pseudomonadota) of total secretory carbohydrate-active enzymes (CAZymes) (**A**), total peptidases (**B**), secretory CAZymes (**C**), and secretory peptidases (**D**)

CAZyme and peptidase profiles revealed by metagenomic analysis

Analysis of the metagenomic contigs revealed 83,423 bacterial and 2,799 archeal CAZyme sequences and 188,107 bacterial and 7,610 archeal peptidase sequences. The dominant taxonomic groups harboring CAZymes and peptidases were Gammaproteobacteria (27.5% and 32.2%), Alphaproteobacteria (25.4% and 27.6%), and Bacteroidota (22.1% and 16.3%) (Fig. 2A and B). We identified the secretory fraction of these enzymes according to the presence of signal peptides. A total of 17,506 bacterial and 160 archaeal CAZyme sequences and 39,122 bacterial and 893 archaeal peptidase sequences with secretory potential were identified. Bacteroidota (33.6%) were the primary contributors to the gene pool of secretory CAZymes, followed by

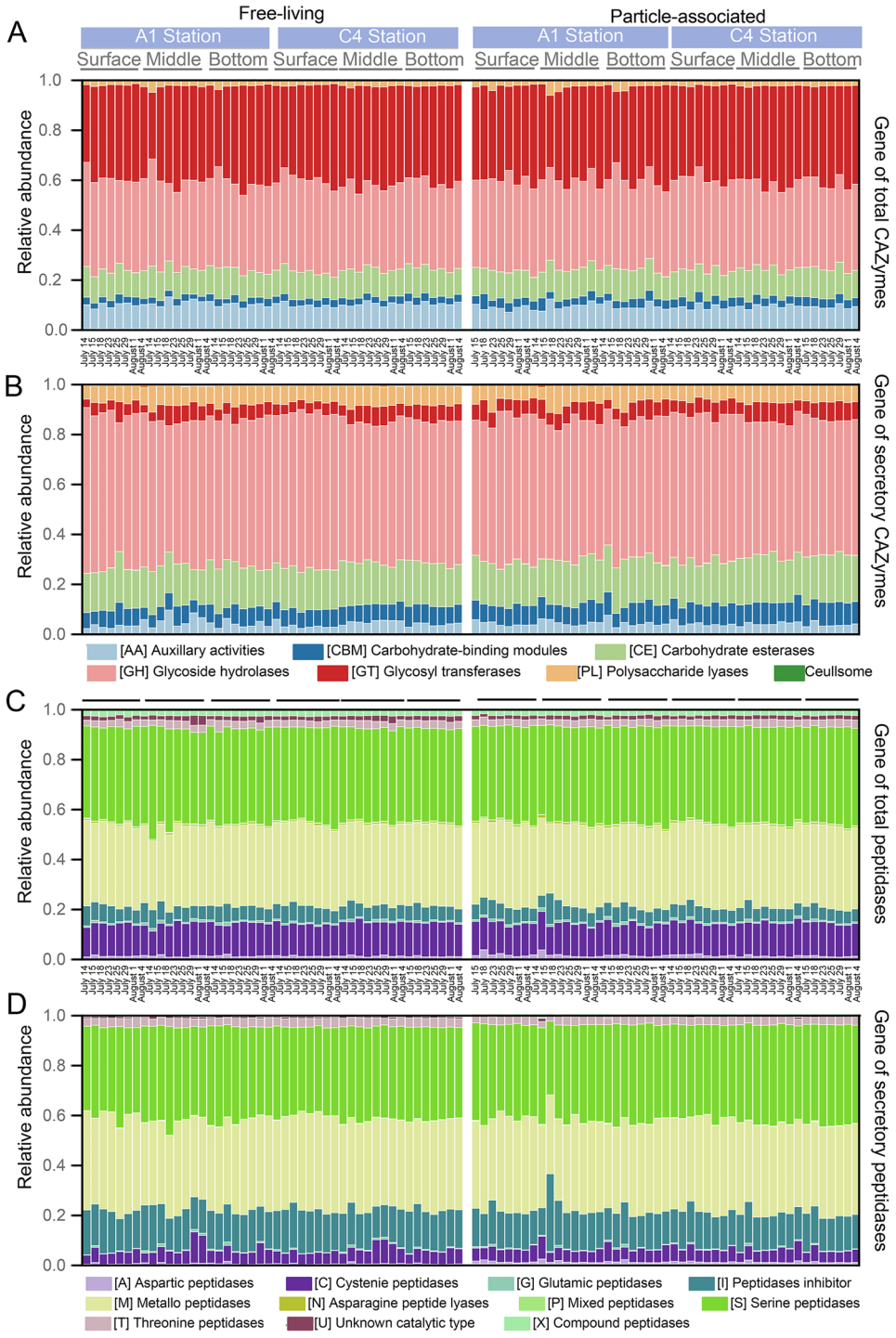


Fig. 3 Functional classification of genes encoding total secretory carbohydrate-active enzymes (CAZymes) (**A**), total peptidases (**B**), secretory CAZymes (**C**), and secretory peptidases (**D**) at the community level

Gammaproteobacteria (31.8%) and Alphaproteobacteria (14.3%) (Fig. 2C). Meanwhile, Gammaproteobacteria (40.0%) dominated the secretory peptidase pool, followed by Bacteroidota (25.5%) and Alphaproteobacteria (19.1%) (Fig. 2D). A clear size-fractionated stratification was found in the taxonomic assignments of the genes encoding CAZymes and peptidases (Fig. 2). Gammaproteobacteria were highly abundant in the particle-associated (PA) fraction, whereas Alphaproteobacteria were highly abundant in the free-living (FL) fraction (Fig. 2). The relative contribution of Bacteroidota to the genes encoding CAZymes and peptidases decreased with depth, whereas those from Alphaproteobacteria and Gammaproteobacteria increased with depth (Fig. 2).

In the total CAZyme gene pool, the dominant CAZyme classes were glycosyl transferases (37.6%), glycoside hydrolases (35.7%), carbohydrate esterases (11.3%), auxiliary activities (9.6%), and carbohydrate-binding modules (9.2%) (Fig. 3A). However, in the secretory CAZyme fraction, glycosyl transferases (5.8%) and auxiliary activities (4.4%) were less abundant, whereas glycoside hydrolases were highly abundant (57.8%) (Fig. 3B). The proportion of cysteine peptidases was considerably lower in the secretory peptidase pool (5.5%) than in the total peptidase pool (13.0%) (Fig. 3C and D). By contrast, serine peptidases, metallopeptidases, and peptidase inhibitors were predominant in the secretory peptidase gene pool throughout the coastal water column, accounting for 37.9%, 36.0%, and 15.5%, respectively.

ABC transporters and TBDT profiles revealed by metagenomic analysis

We identified 108,557 ABC transporter and 2,007 TBDT sequences in the metagenomic contigs. Alphaproteobacteria (47.4%), Gammaproteobacteria (20.5%), Actinomycetota (10.4%), and Bacteroidota (6.8%) dominated the gene pool of ABC transporters (Fig. 4A). Among the genes encoding TBDTs, Gammaproteobacteria (38.1%) and Alphaproteobacteria (35.2%) were the major contributors, followed by Bacteroidota (17.2%) (Fig. 4B). Although the transporters from Alphaproteobacteria and Gammaproteobacteria remained predominant, differences in size fractions for transporters were observed between these two groups. Alphaproteobacteria were highly abundant in the FL fraction, whereas Gammaproteobacteria were highly prevalent in the PA fraction.

Substrate specificity analysis revealed the distinct targets of the ABC transporters and TBDTs. The primary substrate targets of the ABC transporters were sugars, peptides, amino acids, phosphorus compounds, metal ions, and glycine betaine/proline, the composition of which was largely consistent throughout the coastal water column (Fig. 5A). Within Alphaproteobacteria, SAR11 and Rhodobacterales transporters were primarily associated with taurine uptake and polyamine transport, respectively. The substrate targets of TBDTs included chitooligosaccharides, arabinose, digested proteins, starch, vitamin B12, and iron-containing siderophores and thiamin (Fig. 5B). The substrate specificity of TBDTs did not vary substantially between the size fractions or sampling layers but was influenced by the sampling dates. Alphaproteobacterial and gammaproteobacterial TBDTs showed significant differences in substrate preference. In the FL and PA fractions, approximately 70% of alphaproteobacterial TBDTs targeted heme (41.2%), siderophore (19.5%), and vitamin B12 (10.0%) (Fig. 5B). By contrast, gammaproteobacterial TBDTs primarily targeted siderophores (38.6%), followed by heme (15.5%) and vitamin B12 (SiderophoreVitamin_Cluster_973, 12.5%) (Fig. 5B). For Bacteroidota, 50% of the TBDTs were related to siderophores (36.2%) and vitamin B12 (14.5%) (Fig. 5B).

Correlation between enzymes and transporters at the community level

To examine the relationship between enzymes and transporters at the community level, we conducted linear regression and Pearson's correlation analyses. The abundances of total CAZyme and peptidase genes were positively correlated with those of ABC transporters ($r = 0.35$, 0.60, respectively; $p < 0.001$) and TBDTs ($r = 0.87$, 0.88, respectively; $p < 0.001$) (Fig. S2). For secretory gene pools, no significant correlation was observed with ABC transporters ($p > 0.05$), whereas secretory CAZyme ($r = 0.68$, $p < 0.001$) and peptidase genes ($r = 0.76$, $p < 0.001$) were positively correlated with TBDTs (Fig. 6). This trend was further supported by two additional marine datasets. In the transcriptomic data from the Bohai Sea (Fig. S3), transcript abundance correlations revealed strong positive associations between TBDTs and secretory CAZymes ($r = 0.72$, $p < 0.001$) and secretory peptidases ($r = 0.80$, $p < 0.001$). By contrast, ABC transporters showed weaker or nonsignificant correlations with secretory CAZymes ($r = 0.34$, $p > 0.05$) and secretory peptidases ($r = 0.43$, $p < 0.05$). Similarly, gene abundance correlations in the metagenomic data from the northeastern Indian Ocean (Fig. S4) indicated significant relationships of TBDTs with secretory CAZymes ($r = 0.70$, $p < 0.001$) and secretory peptidases ($r = 0.66$, $p < 0.001$), whereas ABC transporters exhibited nonsignificant correlations (secretory CAZymes: $r = -0.28$, $p > 0.05$; secretory peptidases: $r = -0.14$, $p > 0.05$).

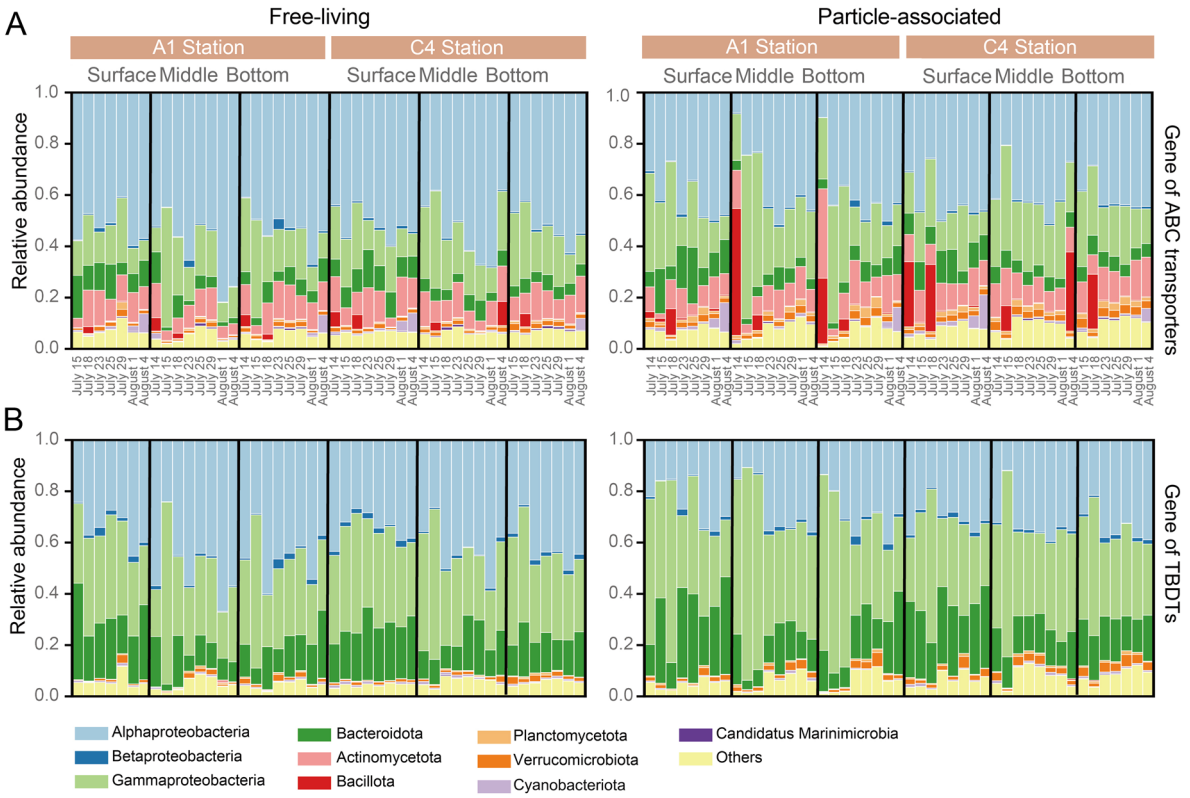


Fig. 4 Taxonomic breakdown at the phylum level (class level for Pseudomonadota) of ATP-binding cassette (ABC) transporters (**A**) and TonB-dependent transporters (TBDTs) (**B**)

MAGs of prokaryotes

To investigate the carbohydrate and protein utilization potential at a high resolution, we constructed 163 nonredundant MAGs with ≥ 75% completeness and ≤ 10% contamination at a 95% average nucleotide identity (Fig. 7). These MAGs captured a substantial fraction of microbial metagenomic reads, ranging from 7.6 to 53.7%. In accordance with the GTDB taxonomy, 154 bacterial MAGs were taxonomically assigned to 12 phyla, and 9 archaeal MAGs were assigned to 4 phyla. Among the bacterial MAGs, Pseudomonadota ($n = 74$) was the most prevalent, followed by Bacteroidota ($n = 43$) and Actinomycetota ($n = 14$). For the archaeal MAGs, Thermoplasmatota ($n = 5$) had the most representatives.

Correlation between enzymes and transporters at the MAG level

Within the nonredundant MAG catalog, the largest secretory CAZyme and peptidase gene repertoires were found in the MAGs of Gemmatimonadota and Bacteroidia, with averages of 24 and 21 CAZymes per Mbp, respectively (Fig. 7). For transporters, Bacillota, Asgardarchaeota, and Actinomycetota MAGs had high ABC transporter gene contents, whereas Gemmatimonadota and Bacteroidia MAGs had high TBDT gene contents. To further explore taxon-specific trends, we conducted a detailed analysis of MAGs from the three major groups: Alphaproteobacteria, Gammaproteobacteria, and Bacteroidota (Fig. 8). In Bacteroidota MAGs, the TBDT gene counts showed a significant positive correlation with the extracellular enzyme gene counts, including secretory CAZymes ($r = 0.32$, $p < 0.05$) and secretory peptidases ($r = 0.54$, $p < 0.001$). Meanwhile, the correlations in Gammaproteobacteria MAGs were not significant (secretory CAZymes: $r = 0.34$, $p > 0.05$; secretory peptidases: $r = 0.26$, $p > 0.05$). For Alphaproteobacteria MAGs, the correlation coefficients were low with poor model fit (secretory CAZymes: $r = 0.009$, $p > 0.05$; secretory peptidases: $r = -0.059$, $p > 0.05$), indicating the lack of significant relationship in this group.

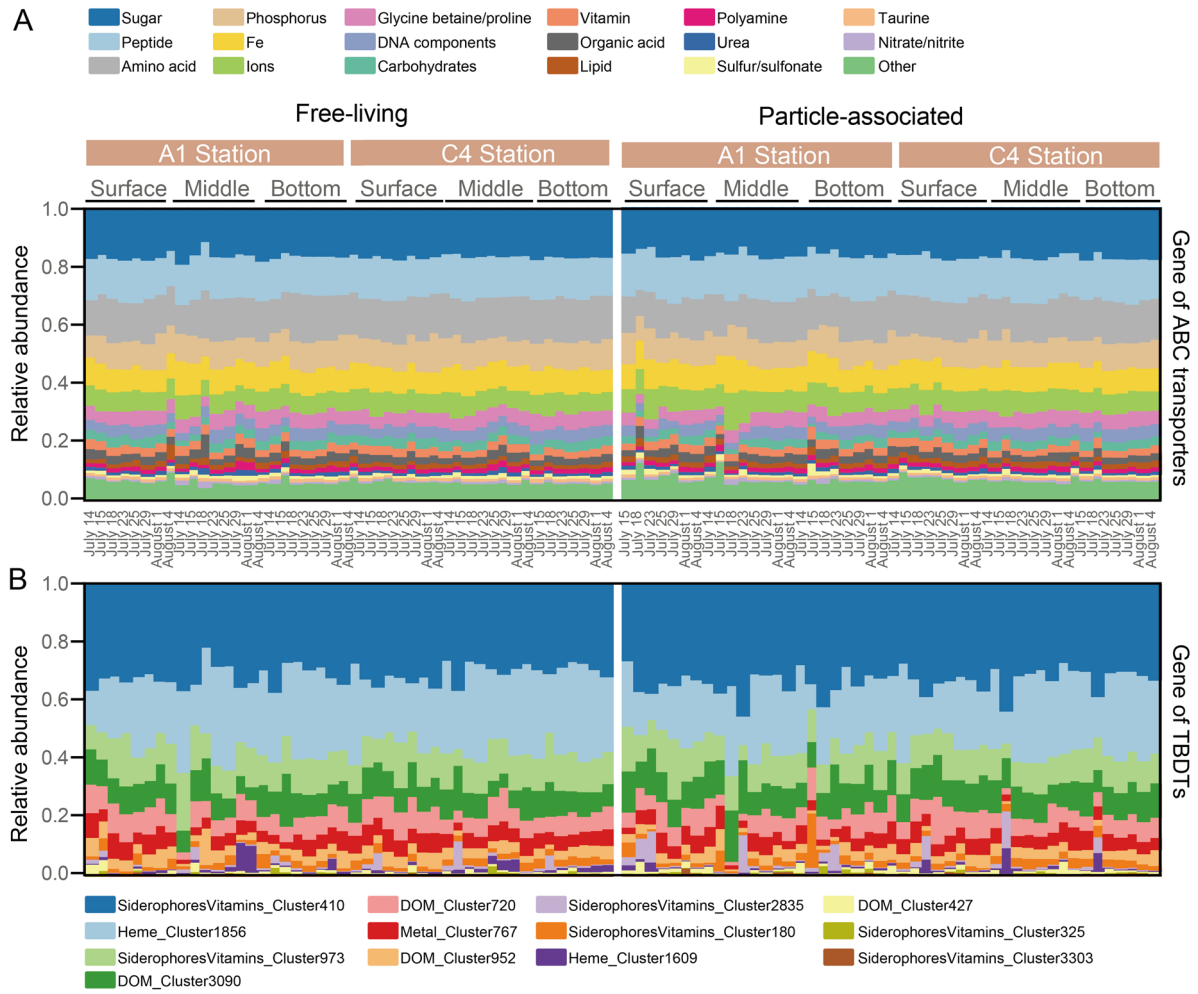


Fig. 5 Substrate specificity of ATP-binding cassette (ABC) transporters (**A**) and TonB-dependent transporters (TBDTs) (**B**) in the metagenomic dataset

Discussion

Core microbial population involved in organic matter metabolism in coastal seawater

This study revealed that Alphaproteobacteria, Gammaproteobacteria, and Bacteroidota were the major clades of heterotrophic bacteria responsible for organic matter degradation in coastal environments (Figs. 2 and 4). These groups were the dominant contributors to extracellular enzymes and transporters (Figs. 2 and 4), but their roles differed across size fractions and organic matter classes. For example, the genes associated with organic matter degradation were predominantly observed in Gammaproteobacteria within the PA fraction; such genes were considerably enriched in Alphaproteobacteria inhabiting the FL fraction (Fig. 2). This finding is consistent with a previous observation during spring phytoplankton blooms (Wang et al. 2024a): Gammaproteobacteria efficiently degraded complex organic matter in the PA fraction, whereas Alphaproteobacteria primarily contributed to organic matter processing in the FL fraction, reflecting their adaptation to small genome sizes and streamlined metabolisms (Crespo et al. 2013). Bacteroidota were the primary contributors of secretory CAZymes (Fig. 2D), highlighting their specialization in HMW polysaccharide degradation, particularly during marine algae blooms (Beidler et al. 2023). Their glycan-degrading systems involve the secretion of CAZymes to partially degrade polysaccharides into oligomers outside the cell before being transported into the periplasm, providing a competitive advantage by taking the substrates out of the dissolved pool (Anderson and Salyers 1989). Overall, the distinct contributions of these bacterial taxa to the secretory enzyme pools across different size fractions emphasize their specialized adaptations

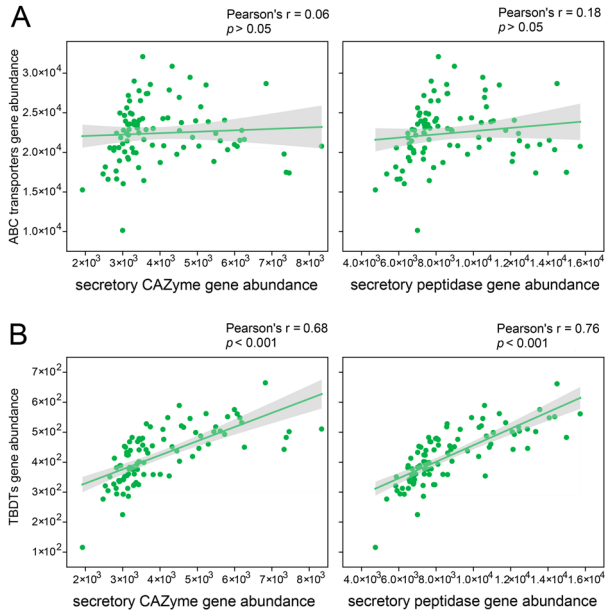


Fig. 6 Correlation among the abundances of ATP-binding cassette (ABC) transporters (**A**), TonB-dependent transporters (TBDTs) (**B**), and extracellular enzymes at the community level

to utilize organic compounds of diverse types and sizes. Furthermore, the extracellular release of hydrolysis products likely fosters crossfeeding interactions among microbial populations, promoting complementary roles in the overall degradation and recycling of organic matter in marine environments.

Transporter genes, which are the key indicators of organic matter uptake patterns (Green et al. 2014; Debeljak et al. 2023), also exhibited distinct taxonomic contributions. Alphaproteobacteria dominated the ABC transporter gene pool (47.4%) (Fig. 4A), with SAR11 playing a significant role in taurine uptake and Rhodobacterales specializing in polyamine transport. Both are critical for nutrient cycling (Clifford et al. 2019). Despite the higher relative abundance of ABC transporter genes compared with TBDTs (Fig. 6), metaproteomic evidence suggests that TBDTs are more actively utilized in marine ecosystems (Zhao et al. 2024). We found that TBDTs were primarily of Gammaproteobacteria and Bacteroidota origin (Fig. 4B) and were predicted to target complex organic molecules, such as chitooligosaccharides (DOM_Cluster_3090) and digested proteins and starch (DOM_Cluster_720) using TBDTs (Fig. 5B). Despite the heterogeneous composition of the organic matter supply (Mou et al. 2008), the diversity in extracellular enzymes and transporters contribute to ecological niche partitioning among bacteria throughout the water column.

Potential functional associations between TBDTs and extracellular enzymes

Our PCoA results suggest that the sampling time has the most significant impact on the metagenomic functional composition (Fig. S1). Our previous study found that during this sampling period, the composition of dissolved organic matter substantially fluctuated (Liu et al. 2024). This dynamic organic matter pool provides an excellent opportunity to investigate the relationship between extracellular enzymes and transporters under varying substrate conditions. In the present work, a significant positive relationship was observed between the abundance of extracellular enzymes and TBDTs, not ABC transporters (Fig. 6). This difference can be attributed to several factors. First, despite being one of the most prevalent and essential transport proteins in marine bacteria, ABC transporters exhibit relative specificity, with each substrate corresponding to a specific transporter (Winkelmann 2002). They primarily facilitate the transport of amino acids, carbohydrates, peptides, vitamins, and metal–ion complexes (Poretsky et al. 2010; Davidson and Chen 2004; Sowell et al. 2009). In contrast to ABC transporters being situated in the inner membrane, TBDTs are located on the outer membrane (Moeck and Coulton 1998). TBDTs were initially identified as a facilitator of the uptake of scarce nutrients, with a particular affinity for iron complexes, such as siderophores, citrate, heme, hemoglobin, transferrin, and lactoferrin (Moeck and Coulton 1998; Hopkinson and Barbeau 2012). They were also associated with the transport of vitamins B12 and B1 and metal ions (Shultis et al. 2006; Schauer et al. 2007). However, further studies revealed that TBDTs can also mediate the transport and absorption of carbohydrates, amino acids, and organic acids (Schauer et al. 2008; Blanvillain et al. 2007) and participate in the translocation of large organic molecules, including aromatic compounds (Miller Todd et al. 2010). Although TBDTs do not exhibit the high specificity of ABC transporters, they can handle a broader and more diverse range of substrates. The increased breakdown of HMW organic matter by extracellular enzymes generates a diversity of complex small molecules, necessitating the use of additional TBDTs to efficiently import these substrates into bacterial cells for utilization. In addition to their roles in import processes, TBDTs are involved in protein secretion. A previous study highlighted the essential role of TBDT Oar for the extracellular secretion of the protease PopC in bacteria (Gómez-Santos et al. 2019). Many TBDTs remain uncharacterized, and some of these could play roles in both the import and secretion—or exclusively in the secretion—of proteins across the outer membrane (Gómez-Santos et al. 2019). Therefore, the relationship between extracellular enzymes and TBDTs, as opposed to ABC transporters, is likely attributed to the broad substrate range of TBDTs, their

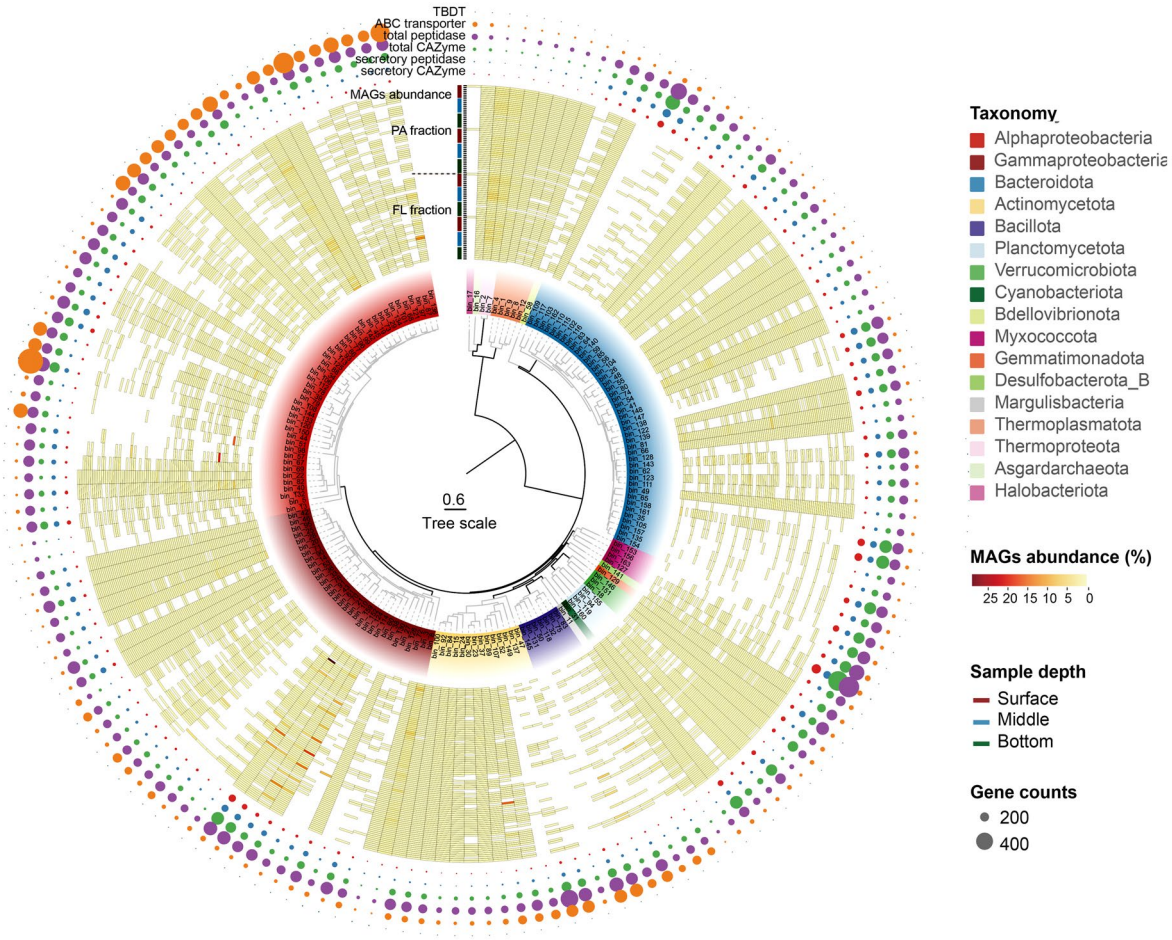


Fig. 7 Phylogenetic trees of the 164 dereplicated metagenome-assembled genomes (MAGs) inferred using maximum-likelihood analysis with a concatenated set of 122 archaeon-specific and 120 bacterium-specific marker genes

essentiality in importing complex small molecules created through the actions of extracellular enzymes, and their additional role in protein secretion. This genetic correlation suggests a potential coregulation or functional linkage between extracellular enzymes and TBDTs in the metabolism of marine heterotrophic prokaryotes.

To further enhance the generalizability of our conclusions, we utilized the transcriptomic data from our own Bohai Sea dataset (Han et al. 2022) and the publicly available metagenomic data from the northeastern Indian Ocean (Wang et al. 2024b). In both datasets, we observed a significant positive correlation between extracellular enzymes and TBDTs, not ABC transporters, which is consistent with our primary findings (Fig. S3 and S4). These additional datasets, representing distinct marine environments, further validate our results. Our conclusions are also consistent with our previous laboratory experiments, where sole carbon source cultivation assays revealed high levels of TBDTs and extracellular enzymes in proteomic analyses (Liu et al. 2023). Overall, these findings support the association between TBDTs and extracellular enzymes across diverse marine environments. The consistent coupling of these gene groups across multiple datasets suggests their functional linkage, which potentially reflects coordinated microbial strategies for resource acquisition and organic matter utilization.

The functional coupling between extracellular enzymes and TBDTs plays a pivotal role in microbial organic matter utilization by facilitating the simultaneous breakdown of complex macromolecules and the efficient uptake of the resulting degradation products. Extracellular enzymes, such as glycoside hydrolases and proteases, break down polysaccharides and proteins into oligomers or small monomers (Arnosti 2011). TBDTs specialize in the transmembrane transport of these degradation products, ensuring their rapid internalization for further metabolism (Tang et al. 2012). This synergistic mechanism enhances the microbial resource acquisition efficiency, providing a competitive advantage in nutrient-limited environments. In marine ecosystems, the

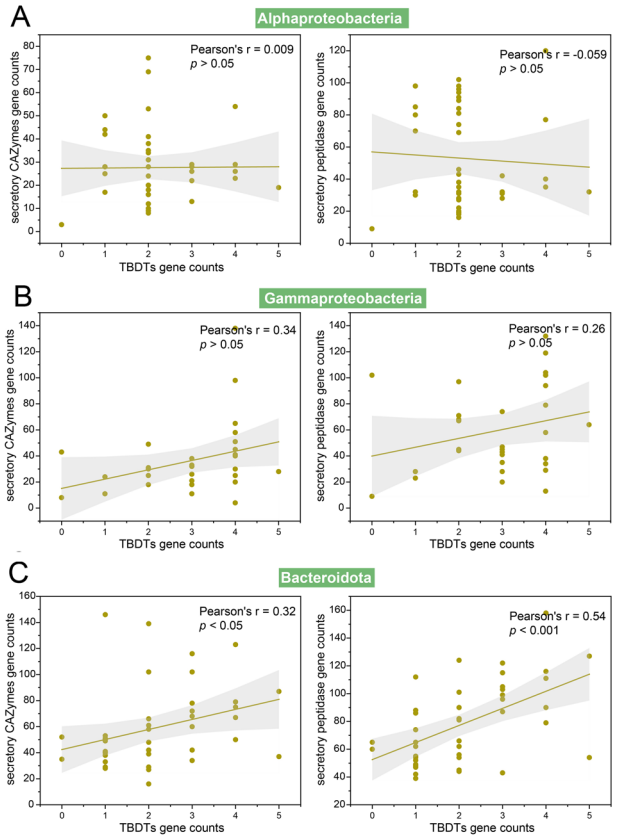


Fig. 8 Correlation between TonB-dependent transporter (TBDT) gene counts and secretory extracellular enzyme gene counts in metagenome-assembled genomes (MAGs) from Alphaproteobacteria (**A**), Gammaproteobacteria (**B**), and Bacteroidota (**C**)

secretion of extracellular enzymes promotes nutrient sharing within microbial communities, and the specificity of TBDTs ensures that degraders benefit from their own enzymatic activity. This finding aligns with the previous reports, proposing that "selfish" substrate acquisition strategies reduce carbon leakage to the environment (Reintjes et al. 2017, 2019). Our MAG analysis revealed distinct patterns across taxonomic groups. Bacteroidia MAGs exhibited a significant positive correlation between extracellular enzymes and TBDT gene abundances (Fig. 8), aligning with their role as "selfish" degraders that prioritize enzymatic investment for private metabolite harvest (Reintjes et al. 2019). By contrast, Gammaproteobacteria MAGs showed nonsignificant correlations (Fig. 8), suggesting a mixed ecological strategy where some members engage in selfish uptake and others act as scavengers reliant on exogenous hydrolysis products (Reintjes et al. 2019). Alphaproteobacteria MAGs, including oligotrophic specialists such as SAR11, displayed negligible correlations (Fig. 8), consistent with their known scavenging lifestyle of being dependent on low-molecular-weight dissolved organic matter (Reintjes et al. 2019). These findings suggest that enzyme–transporter coupling enhances the efficiency of organic matter degradation while reducing carbon loss to the surrounding microbes. Moreover, the coupling of extracellular enzymes and TBDTs accelerates the bacterial intracellular turnover of organic matter, thereby enhancing the efficiency of the microbial carbon pump by promoting the transformation of labile carbon into refractory dissolved organic carbon for long-term sequestration (Jiao et al. 2024). This coordinated action is a crucial upstream process that directly influences long-term carbon storage in marine systems.

Limitations and future directions

This study identified potential functional associations between TBDTs and extracellular enzymes at the metagenomic level. Although our findings provide a strong foundation for this hypothesis, experimental validation is still warranted to elucidate the underlying biochemical mechanisms. Conducting targeted TBDT gene knockout experiments in a specific bacterial strain and analyzing their impact on extracellular enzyme activity are essential. Future studies employing metaproteomics under varying environmental conditions and coculture experiments will further refine our understanding of these interactions. These strategies will elucidate the regulatory mechanisms governing extracellular enzyme secretion and transporter function, providing deep insights into microbial substrate degradation.

Conclusion

In this study, we used metagenomics to explore the diversity and distribution of extracellular enzymes and transporters in coastal waters, revealing distinct patterns in their taxonomic affiliations. Our results highlight the critical roles of Gammaproteobacteria, Alphaproteobacteria, and Bacteroidota in organic matter degradation, each with unique strategies for processing and assimilating organic compounds. Gammaproteobacteria were predominantly associated with PA fractions and contributed notably to the secretion of peptidases. Alphaproteobacteria were prevalent in FL fractions and dominated the gene pool of ABC transporters. Bacteroidota were the major contributors to the secreted CAZyme pool. Our study uncovered a significant positive correlation between extracellular enzymes and TBDTs at the metagenomic and MAG levels, particularly in Bacteroidota. This result suggests a functional linkage between the hydrolysis of HMW organic matter and the subsequent uptake of the cleaved compounds through TBDTs. Despite their abundance, ABC transporters showed weaker correlations with extracellular enzyme activities, possibly due to their substrate specificity and involvement in the uptake of small

organic molecules. These findings provide new insights into the metabolic strategies employed by marine heterotrophic prokaryotes, emphasizing the role of extracellular enzymes and transporters in shaping the communities of microbes and their ability to occupy distinct ecological niches in dynamic marine environments. This work contributes to a comprehensive understanding of organic matter cycling in coastal ecosystems and highlights the importance of considering enzymatic and transporter-mediated processes in future microbial ecology studies.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s42995-025-00314-9.

Acknowledgements This study was supported by the National Natural Science Foundation of China Project (42141003, 42188102, and 42076160) and the Ministry of Science and Technology (MOST) ONCE Project.

Author contributions LSJ and TK conceived and designed the experiments. LSJ, CQR, and QXY performed sample collection. LSJ and LWH conducted data analysis. LSJ and TK interpreted the data and wrote the manuscript. All authors critically reviewed, edited, and approved the final manuscript.

Data availability Metagenomics raw data have been uploaded to the National Omics Data Encyclopedia (NODE, https://www.biosino.org/node/) database with the identifier OEX00024201.

Declarations

Conflict of interest The authors declare no conflict of interest.

Animal and human rights statement We declare that this study does not include any research involving human participants. All the experiments were conducted following standard procedures.

References

Anderson KL, Salyers AA (1989) Biochemical evidence that starch breakdown by *Bacteroides thetaiotaomicron* involves outer membrane starch-binding sites and periplasmic starch-degrading enzymes. J Bacteriol 171:3192–3198

Arnosti C (2011) Microbial extracellular enzymes and the marine carbon cycle. Annu Rev Mar Sci 3:401–425

Azam F, Malfatti F (2007) Microbial structuring of marine ecosystems. Nat Rev Microbiol 5:782–791

Beidler I, Robb CS, Vidal-Melgosa S, Zühlke MK, Bartosik D, Solanki V, Markert S, Becher D, Schweder T, Hehemann JH (2023) Marine bacteroidetes use a conserved enzymatic cascade to digest diatom β-mannan. ISME J 17:276–285

Blanvillain S, Meyer D, Boulanger A, Lautier M, Guynet C, Denancé N, Vasse J, Lauber E, Arlat M (2007) Plant carbohydrate scavenging through TonB-dependent receptors: a feature shared by phytopathogenic and aquatic bacteria. PLoS ONE 2:e224

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120

Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, Schulz F, Jarett J, Rivers AR, Eloe-Fadrosh EA, Tringe SG, Ivanova NN, Copeland A, Clum A, Becraft ED, Malmstrom RR, Birren B, Podar M, Bork P, Weinstock GM et al (2017) Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat Biotechnol 35:725–731

Boyd PW, Claustre H, Levy M, Siegel A, Weber T (2019) Multi-faceted particle pumps drive carbon sequestration in the ocean. Nature 568:327–335

Buchfink B, Xie C, Huson DH (2015) Fast and sensitive protein alignment using DIAMOND. Nat Methods 12:59–60

Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH (2019) GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics 36:1925–1927

Chklovski A, Parks DH, Woodcroft BJ, Tyson GW (2023) CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. Nat Methods 20:1203–1212

Chrost R, Velimirov B (2013) Measurement of enzyme kinetics in water samples: Effect of freezing and soluble stabilizer. Mar Ecol Prog Ser 70:93–100

Clifford EL, Varela MM, De Corte D, Bode A, Ortiz V, Herndl GJ, Sintes E (2019) Taurine is a major carbon and energy source for marine prokaryotes in the North Atlantic Ocean off the Iberian Peninsula. Microb Ecol 78:299–312

Crespo BG, Pommier T, Fernández-Gómez B, Pedrós-Alió C (2013) Taxonomic composition of the particle-attached and free-living bacterial assemblages in the Northwest Mediterranean Sea analyzed by pyrosequencing of the 16S rRNA. MicrobiologyOpen 2:541–552

Davidson AL, Chen J (2004) ATP-binding cassette transporters in bacteria. Annu Rev Biochem 73:241–268

Davidson AL, Dassa E, Orelle C, Chen J (2008) Structure, function, and evolution of bacterial ATP-binding cassette systems. Microbiol Mol Biol Rev 72:317–364

Debeljak P, Bayer B, Sun Y, Herndl GJ, Obernosterer I (2023) Seasonal patterns in microbial carbon and iron transporter expression in the Southern Ocean. Microbiome 11:187

Drula E, Garron M-L, Dogan S, Lombard V, Henrissat B, Terrapon N (2021) The carbohydrate-active enzyme database: functions and literature. Nucleic Acids Res 50:D571–D577

Francis TB, Bartosik D, Sura T, Sichert A, Hehemann J-H, Markert S, Schweder T, Fuchs BM, Teeling H, Amann RI, Becher D (2021) Changing expression patterns of TonB-dependent transporters suggest shifts in polysaccharide consumption over the course of a spring phytoplankton bloom. ISME J 15:2336–2350

Fu L, Niu B, Zhu Z, Wu S, Li W (2012) CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 28:3150–3152

Gómez-Santos N, Glatter T, Koebnik R, Świątek-Połatyńska WA, Søgaard-Andersen L (2019) A TonB-dependent transporter is required for secretion of protease PopC across the bacterial outer membrane. Nat Commun 10:1360

Green NW, Perdue EM, Aiken GR, Butler KD, Chen H, Dittmar T, Niggemann J, Stubbins A (2014) An intercomparison of three methods for the large-scale isolation of oceanic dissolved organic matter. Mar Chem 161:14–19

Han Y, Zhang M, Chen X, Zhai W, Tan E, Tang K (2022) Transcriptomic evidences for microbial carbon and nitrogen cycles in the deoxygenated seawaters of Bohai Sea. Environ Int 158:106889

Hopkinson BM, Barbeau KA (2012) Iron transporters in marine prokaryotic genomes and metagenomes. Environ Microbiol 14:114–128

Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinf 11:119

Jiao N, Luo T, Chen Q, Zhao Z, Xiao X, Liu J, Jian Z, Xie S, Thomas H, Herndl GJ, Benner R, Gonsior M, Chen F, Cai W-J, Robinson C (2024) The microbial carbon pump and climate change. Nat Rev Microbiol 22:408–419

Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M (2016) KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res 44:D457-462

Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z (2019) MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ 7:e7359

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359

Li D, Luo R, Liu C-M, Leung C-M, Ting H-F, Sadakane K, Yamashita H, Lam T-W (2016) MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. Methods 102:3–11

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079

Liu Y, Makarova KS, Huang W-C, Wolf YI, Nikolskaya AN, Zhang X, Cai M, Zhang C-J, Xu W, Luo Z, Cheng L, Koonin EV, Li M (2021) Expanded diversity of Asgard archaea and their relationships with eukaryotes. Nature 593:553–557

Liu L, Chen X, Ye J, Ma X, Han Y, He Y, Tang K (2023) Sulfoquinovose is a widespread organosulfur substrate for Roseobacter clade bacteria in the ocean. ISME J 17:393–405

Liu S, Chen Q, Liu L, Dong C, Qiu X, Tang K (2024) Organic matter composition fluctuations disrupt free-living bacterial communities more than particle-associated bacterial communities in coastal waters. Sci Total Environ 949:174845

Miller TR, Delcher AL, Salzberg SL, Saunders E, Detter JC, Halden RU (2010) Genome sequence of the dioxin-mineralizing bacterium *Sphingomonas wittichii* RW1. J Bacteriol 192:6101–6102

Moeck GS, Coulton JW (1998) TonB-dependent iron acquisition: mechanisms of siderophore-mediated active transport. Mol Microbiol 28:675–681

Mou X, Sun S, Edwards RA, Hodson RE, Moran MA (2008) Bacterial carbon processing by generalist species in the coastal ocean. Nature 451:708–711

Olm MR, Brown CT, Brooks B, Banfield JF (2017) dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. ISME J 11:2864–2868

Poretsky RS, Sun S, Mou X, Moran MA (2010) Transporter genes expressed by coastal bacterioplankton in response to dissolved organic carbon. Environ Microbiol 12:616–627

Price MN, Dehal PS, Arkin AP (2010) FastTree 2 – approximately maximum-likelihood trees for large alignments. PLoS ONE 5:e9490

Rawlings ND, Barrett AJ, Thomas PD, Huang X, Bateman A, Finn RD (2017) The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. Nucleic Acids Res 46:D624–D632

Reintjes G, Arnosti C, Fuchs BM, Amann R (2017) An alternative polysaccharide uptake mechanism of marine bacteria. ISME J 11:1640–1650

Reintjes G, Arnosti C, Fuchs B, Amann R (2019) Selfish, sharing and scavenging bacteria in the Atlantic Ocean: a biogeographical study of bacterial substrate utilisation. ISME J 13:1119–1132

Rinke C, Chuvochina M, Mussig AJ, Chaumeil P-A, Davín AA, Waite DW, Whitman WB, Parks DH, Hugenholtz P (2021) A standardized archaeal taxonomy for the Genome Taxonomy Database. Nat Microbiol 6:946–959

Saier MH, Reddy VS, Moreno-Hagelsieb G, Hendargo KJ, Zhang Y, Iddamsetty V, Lam Katie Jing K, Tian N, Russum S, Wang J, Medrano-Soto A (2021) The Transporter Classification Database (TCDB): 2021 update. Nucleic Acids Res 49:D461–D467

Schauer K, Gouget B, Carrière M, Labigne A, De Reuse H (2007) Novel nickel transport mechanism across the bacterial outer membrane energized by the TonB/ExbB/ExbD machinery. Mol Microbiol 63:1054–1068

Schauer K, Rodionov DA, De Reuse H (2008) New substrates for TonB-dependent transport: do we only see the 'tip of the iceberg'? Trends Biochem Sci 33:330–338

Schlitzer R (2002) Interactive analysis and visualization of geoscience data with Ocean Data View. Comput Geosci 28:1211–1218

Shultis DD, Purdy MD, Banchs CN, Wiener MC (2006) Outer membrane active transport: structure of the BtuB:TonB complex. Science 312:1396–1399

Sieber CMK, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG, Banfield JF (2018) Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. Nat Microbiol 3:836–843

Sinsabaugh RL, Lauber CL, Weintraub MN, Ahmed B, Allison SD, Crenshaw C, Contosta AR, Cusack D, Frey S, Gallo ME, Gartner TB, Hobbie SE, Holland K, Keeler BL, Powers JS, Stursova M, Takacs-Vesbach C, Waldrop MP, Wallenstein MD, Zak DR et al (2008) Stoichiometry of soil enzyme activity at global scale. Ecol Lett 11:1252–1264

Sowell SM, Wilhelm LJ, Norbeck AD, Lipton MS, Nicora CD, Barofsky DF, Carlson CA, Smith RD, Giovanonni SJ (2009) Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. ISME J 3:93–105

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

Teufel F, Almagro Armenteros JJ, Johansen AR, Johansen AR, Gíslason MH, Pihl SI, Tsirigos KD, Winther O, Brunak S, von Heijne G, Nielsen H (2022) SignalP 6.0 predicts all five types of signal peptides using protein language models. Nat Biotechnol 40:1023–1025

Wagner GP, Kin K, Lynch VJ (2012) Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. Theory Biosci 131:281–285

Wang F-Q, Bartosik D, Sidhu C, Siebers R, Lu D-C, Trautwein-Schult A, Becher D, Huettel B, Rick J, Kirstein IV, Wiltshire KH, Schweder T, Fuchs BM, Bengtsson MM, Teeling H, Amann RI (2024a) Particle-attached bacteria act as gatekeepers in the decomposition of complex phytoplankton polysaccharides. Microbiome 12:32

Wang X, Arifeen MZU, Hou S, Zheng Q (2024b) Depth-dependent microbial metagenomes sampled in the northeastern Indian Ocean. Sci Data 11:88

Xie J, Chen Y, Cai G, Cai R, Hu Z, Wang H (2023) Tree visualization by one table (tvBOT): a web application for visualizing, modifying and annotating phylogenetic trees. Nucleic Acids Res 51:W587–W592

Yuan H, Li T, Li H, Wang C, Li L, Lin X, Lin S (2021) Diversity distribution, driving factors and assembly mechanisms of free-living and particle-associated bacterial communities at a Subtropical Marginal Sea. Microorganisms 9:2445

Zhang J (2015) Ecological continuum from the Changjiang (Yangtze River) watersheds to the East China Sea continental margin. Springer Cham, Switzerland

Zhao Z, Baltar F, Herndl GJ (2020) Linking extracellular enzymes to phylogeny indicates a predominantly particle-associated lifestyle of deep-sea prokaryotes. Sci Adv 6:eaaz4354

Zhao Z, Amano C, Reinthaler T, Orellana MV, Herndl GJ (2024) Substrate uptake patterns shape niche separation in marine prokaryotic microbiome. Sci Adv 10:eadn5143

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.