The metabolic mechanisms underlying zooplankton-derived dissolved organic matter's chemical properties

Muhammad Firman Nuruddin[1,2], Ding He[1,2], Longjun Wu[1,2*]

[1]Department of Ocean Science, The Hong Kong University of Science and Technology, Hong Kong SAR, China.

[2]Center for Ocean Research in Hong Kong and Macau, Hong Kong SAR, China.

Abstract

Dissolved organic matter (DOM), the largest reservoir of organic material in the ocean, plays a crucial role in the global nutrient cycle and the microbial loop. While existing studies have documented significant DOM release by zooplankton, the chemo diversity and properties of this DOM, along with the physiological mechanisms influencing these characteristics in the environment, remain inadequately explored. We conducted zooplankton sampling followed by onboard DOM release experiments in heterogeneous estuarine-coastal water systems, followed by molecular characterization of the DOM using Fourier-transform ion cyclotron resonance mass spectrometry. Additionally, we analyzed zooplankton metabolic activities through meta-transcriptomics to elucidate the relationship between the chemical properties of the released DOM and the underlying physiological processes of zooplankton. Our findings reveal substantial variations in the molecular diversity of DOM released by zooplankton, more specifically lipids-like, protein-like, and unsaturated hydrocarbon-like between mesotrophic and eutrophic coastal zooplankton communities. We found strong correlations between chemical composition of the DOM and zooplankton gene functions associated with metabolism processes such as carbohydrate metabolism, nucleotide processing, energy production, and coenzyme metabolism. Furthermore, the modified aromaticity indexes of the released DOM are also highly associated with metabolism-related gene functions such as amino acid, carbohydrate metabolism, lipid, energy production, as well as glycan biosynthesis, indicating that zooplankton metabolic processes significantly influence DOM aromaticity. This study enhances our understanding of how organism's metabolic processes shape the molecular characteristics of DOM they release, highlighting its implications for nutrient cycling in the environment.

Introduction

Dissolved organic matter (DOM) constitutes the largest pool of organic material in the ocean, driving global biogeochemical cycles and sustaining microbial communities through the microbial loop[1,2]. Comprising a complex mixture of bioavailable and refractory compounds,

DOM serves as energy source for microbial heterotrophs, modulates nutrient and trace metal bioavailability, and alters the optical characteristics of water bodies[3–6]. While autochthonous DOM derived from phytoplankton exudates has been extensively studied, it accounts for only ~50% of heterotrophic microbial carbon demand, suggesting critical gaps in identifying other key DOM sources within aquatic ecosystems[7–10].

Metazoan zooplankton serves as a link between primary producers and microbial loops, mediating DOM transfer via excretion, sloppy feeding, and digestion[1,11–14]. Unlike phytoplankton-derived DOM—which predominantly contributes to relatively more refractory pools[15–17]—zooplankton releases labile components such as amino acids and protein, which can serve as important nutritional sources for both autotrophic and heterotrophic microbes[8,18,19]. By modulating the availability of these labile substrates, zooplankton indirectly regulate microbial community dynamics, drive shifts in nutrient remineralization rates and alter the efficiency of nutrients transfer through aquatic food webs. On one hand, despite substantial advancements in our understanding of metazoan zooplankton's ecological functions, limited studies focus on characterized their derived DOM molecules composition[19,20]. On the other hand, although prior studies have focused on elucidating zooplankton gene expression under environmental heterogeneity[21–24] but have not explained the relation of these processes to molecular-level of the released DOM dynamics, leaving a critical disconnect between organismal biology and biogeochemical outcomes. Therefore, elucidating the physiological processes (gene expression) of zooplankton together with the variations of their derived DOM is critical for capturing a holistic understanding of the animal's role in biochemical cycle within aquatic ecosystems.

Advances in high-resolution mass spectrometry and community-level omics now enable unprecedented exploration of organism metabolic-DOM linkage. Techniques such as Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FTICR-MS) allows for the detailed characterization of DOM molecular formulae[30,31], while meta transcriptomics enable the examination of organism including zooplankton gene expression which elaborates their physiological and metabolism processes at the community level[32]. Here, we integrate these tools to test the hypothesis that zooplankton metabolism gene expression variabilities govern the molecular diversity and functional traits of released DOM. Through spatially resolved sampling across trophic level gradients, eutrophic to mesotrophic in pearl river estuary and adjacent coastal regions, the second largest river in China, followed with DOM release experiments, we uncover relationships between zooplankton metabolic pathways—including

amino acid, carbohydrate metabolism, lipid, energy production, as well as glycan biosynthesis and metabolism—and DOM chemical signatures. Our findings establish organism physiology as a key determinant of autochthonous marine DOM dynamics, advancing mechanistic understanding of animal physiological process role in marine biogeochemistry.

Result and Discussion

Zooplankton derived DOM overall variability

To characterize dissolved organic matter (DOM) produced by zooplankton physiological activity in heterogeneous coastal waters, we conducted controlled zooplankton release incubations targeting DOM from metabolic excretion and faecal pellet leaching. These experiments generated 15 DOM samples across seven sites spanning a trophic gradient from the eutrophic Pearl River Estuarine delta to adjacent mesotrophic coastal waters (Fig. 1a).

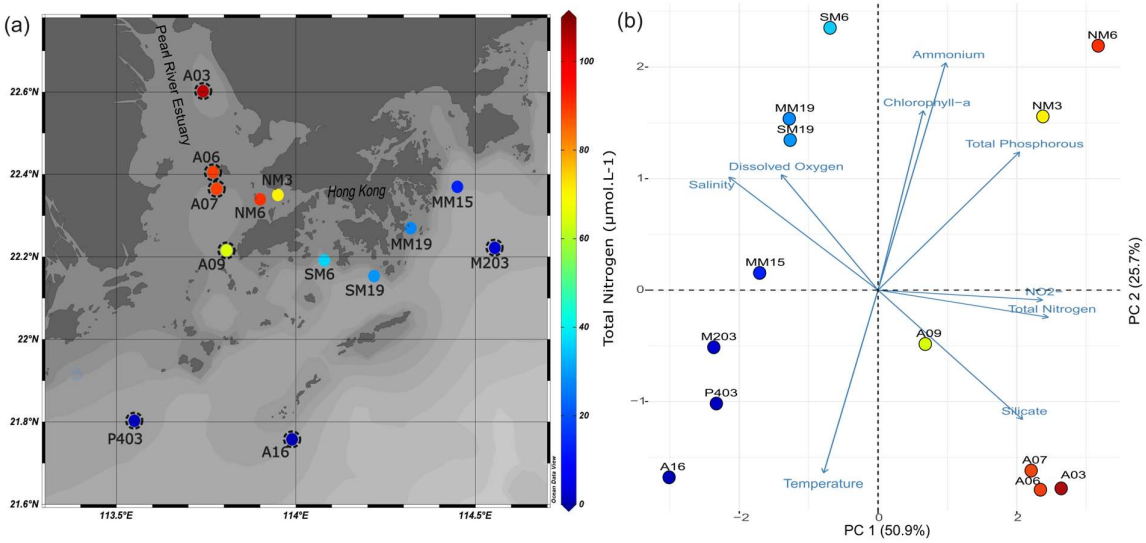

Fig 1. Environmental condition of the sampling sites. (a) Location of the zooplankton sampling and DOM released incubation along the Pearl River Estuary and adjacent coastal waters. (b) Principal component analysis (PCA) biplot of environmental parameters. The stations for zooplankton derived DOM release incubation (summer 2024 cruise) were indicated by dashed circles. The colours dots denote nutrient level (total nitrogen) while the arrow length represents the contributions of each environmental parameter toward the variance of principal component.

We conducted a principal component analysis (PCA) on measured environmental parameters to assess the environmental heterogeneity across the sampling stations (Fig. 1 (b)). The first two principal components (PC1 and PC2) collectively explained 76.6% of the total variance

across sampling sites. PC1 was primarily driven by nutrient concentrations, including total nitrogen, nitrite (NO_2^-), and salinity, while PC2 reflected variations in ammonium, total phosphorus, and temperature. Given the distinct nutrient gradients observed, we classified regions using trophic state indices, revealing a clear separation into mesotrophic (blue dots) and eutrophic (yellow to red dots) regions (Fig. 1a).

DOM analysis identified a total of 4,945 molecular formulas, with an average of 900 formulas for each sample of zooplankton-derived DOM. These molecules were categorized into different classes, including lipids, proteins, amino sugars, carbohydrates, unsaturated hydrocarbons, lignin, tannins, and condensed hydrocarbon-like compounds, based on their H/C and O/C ratios threshold, as shown in a Van Krevelen diagram (Fig. 2a). Notably, nearly half (48.8%) of the DOM molecules were common to both eutrophic and mesotrophic zooplankton communities, while 30.8% and 20.5% were exclusive to eutrophic and mesotrophic systems, respectively (Fig. 2b). This overlap suggests a core set of metabolic byproducts shared across nutrient regimes, with unique molecular signatures may reflecting environment-specific physiological processes[33].

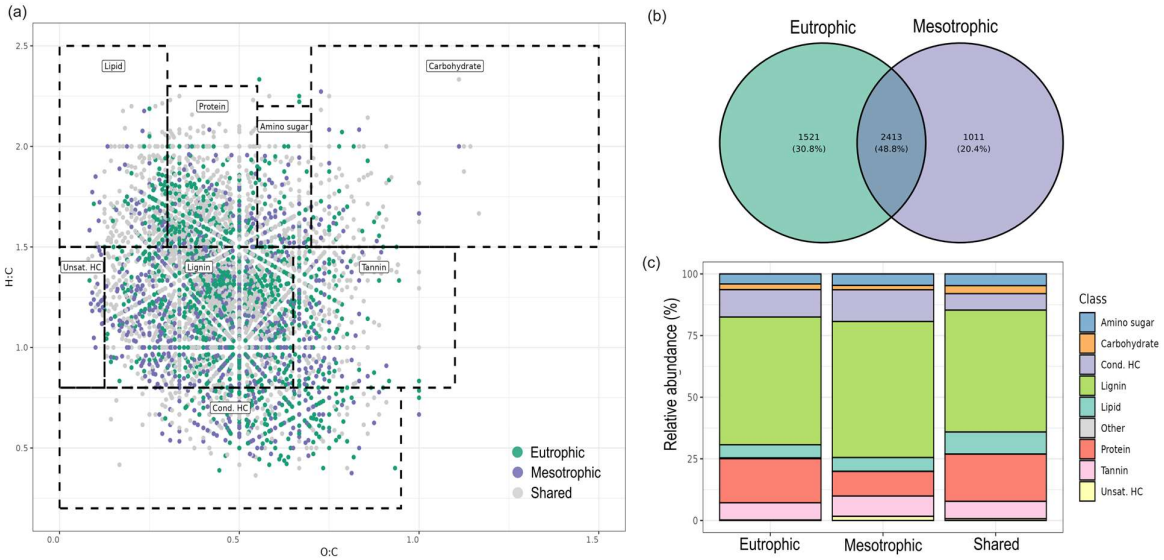

Fig 2. Distributions of zooplankton derived DOM molecules. (a) Van Kravelen diagram of the DOM molecules with the colours representing the molecule presences in which region trophic conditions. (b) Venn diagram of molecular formula numbers of presences in each region (mesotrophic, eutrophic, and shared). (c) Relative abundance of DOM compound like composition which uniquely presence in eutrophic, mesotrophic, and shared in both regions.

Lignin-like compounds dominated the DOM pool in all regions (55-62%), followed by protein-like (12-20%), condensed hydrocarbon-like (5-15%), and tannin-like (~5%) molecules (Fig. 2c). The comparative analysis revealed striking spatial differences in zooplankton derived DOM molecular composition. Mesotrophic zooplankton-derived DOM exhibited significantly higher proportions of lipid-like and unsaturated hydrocarbon-like compounds compared to eutrophic systems, whereas protein-like molecules were enriched in eutrophic DOM (Fig. 3a). These patterns align with the nutrient gradients where protein-like DOM in nutrient-rich systems likely reflects elevated nitrogen and phosphorous availability in the food sources (phytoplankton) and rapid zooplankton growth, driving nitrogen and phosphorous contained DOM release. Conversely, lipid- and unsaturated hydrocarbon-like compounds in mesotrophic regions may arise from lipid catabolism or structural tissue turnover under resource-limited conditions, where zooplankton rely on stored energy reserves.

Molecular traits further underscored spatial divergence. Mesotrophic DOM displayed significantly higher modified aromaticity indices and double bond equivalents ($p < 0.001$; Fig. 3b–c), indicative of structurally complex, aromatic, and unsaturated molecules. Such compounds are typically recalcitrant, aligning with relatively lower Gibbs free energy values (Fig. 3e), which suggest reduced bioavailability and slower microbial degradation rates. In contrast, eutrophic DOM showed a higher nominal oxidation state of carbon (NOSC; Fig. 3d), indicating greater oxidative potential and labile substrates.

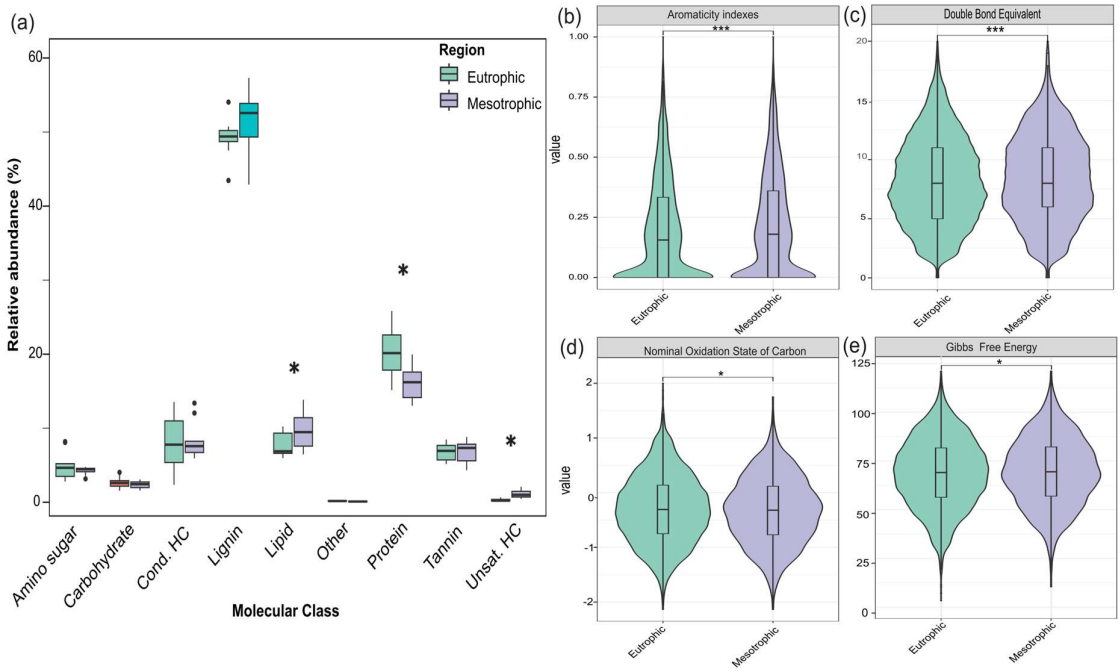

Fig 3. Comparison of zooplankton derived DOM molecules properties between eutrophic and mesotrophic. (a) Relative abundance (%) of DOM molecular classes with the clear difference's molecule classes marked by asterisk. DOM chemical traits (b) Aromaticity indexes, (c) double bond equivalent, (d) Nominal Oxidation State of Carbon, and (e) Gibbs Free Energy. The colours represent the eutrophic and mesotrophic regions.

Zooplankton community gene expression profiles highly associated with environmental factors

To investigate physiological variability in zooplankton communities across sampling regions, we conducted a meta-transcriptomic analysis. Taxonomic classification revealed 20 zooplankton classes (Fig. S1), with the top 20 classes representing ~80% of total relative abundance. Dominant taxa included *Calanoida, Harpacticoida*, and *Siphonostomatoida*, consistent across all regions. Functional annotation identified 12,663 transcripts with Kyoto Encyclopaedia of Genes and Genomes (KEGG) orthologs (KOs).

To resolve metabolic heterogeneity in zooplankton communities, we analyzed KEGG-annotated gene functional profiles and their association with environmental variables using Weighted Gene Co-Expression Network Analysis (WGCNA; Fig. 4a). Principal Component Analysis (PCA) of environmental parameters revealed nutrient concentrations (e.g., total phosphorus, ammonium) as primary drivers of variance, prompting focused analysis of nutrient-associated gene modules. WGCNA identified two gene modules highly associated with nutrient concentrations, module 4 (2,096 KOs) exhibited positive correlations with phosphorus and ammonium concentrations and module 7 (341 KOs) showed inverse relationships (Fig. 4a).

Gene enrichment analysis of these modules revealed distinct transcriptional strategies tied to metabolic adaptation. The module 4 associated positively with nutrient level, exhibited enrichment in pathways associated with amino acid metabolism, energy production, cofactor and vitamin biosynthesis, and terpenoid and polyketide metabolism (Fig. 4b). This profile suggests a metabolic emphasis on biosynthesis and energy-intensive processes, consistent with environments nutrients level which support growth and compound synthesis. In contrast, the module 7 inversely linked to nutrient levels, displayed some enrichment for pathways such as retinol metabolism, drug metabolism–cytochrome P450 (both subsets of cofactor/vitamin metabolism), and linoleic acid metabolism (lipid metabolism). The latter may reflect zooplankton lipid metabolism or detoxification processes under different trophic conditions,

where metabolic activity shifts toward conservation rather than biosynthesis. These findings highlight how zooplankton communities dynamically modulate core metabolic pathways—such as energy optimization under nutrient scarcity—in response to shifting trophic conditions. The results further underscore nutrient availability as a key driver of transcriptional plasticity, shaping metabolic resource allocation strategies across environmental gradients which potentially influence the physiology of zooplankton and their derived DOM molecules characteristics.

To further understand the potential zooplankton metabolic processes variabilities among different environmental conditions, we look at all the genes annotated to KEGG subcategories metabolism function. Notably, energy metabolism genes correlated positively with total phosphorus (PO_4^{3-}) and ammonium (NH_4^+) (Fig. 4c), indicating that zooplankton in eutrophic (nutrient-rich) environments may prioritize energy (ATP) generation, which potentially elevating DOM excretion rates of phosphorus and nitrogen-rich compounds (e.g., phospholipids, nucleic acid derivatives). Similarly, the positive association between lipid metabolism genes and NH_4^+ suggests nitrogen availability enhances lipid turnover in zooplankton, which will potentially increase the release of lipid-derived DOM (e.g., fatty acids, sterols). Conversely, negative correlations between dissolved oxygen (DO) and amino acid as well as energy metabolism genes imply oxygen level affect respiration and protein catabolism in zooplankton (Fig. 4c). This may reduce the release of proteinaceous DOM (e.g., amino acids, peptides) while favouring alternative pathways (e.g., anaerobic fermentation), altering DOM composition to include relatively more refractory metabolites.

To compare metabolic processes across trophic gradients, we also quantified the relative expression ratios of KEGG-annotated metabolism genes between zooplankton communities in eutrophic and mesotrophic environments. Strikingly, zooplankton in eutrophic regions exhibited reduced expression of metabolism-related genes compared to their mesotrophic counterparts (Fig. 4d). This pattern can reflect the observed physiological heterogeneity in zooplankton communities, where mesotrophic populations—facing relatively nutrient scarcity—appear to upregulate metabolic processes critical for nutrient acquisition, stoichiometric homeostasis, and survival under resource constraints. Specifically, heightened metabolic activity in mesotrophic systems likely reflects acclimatization strategies to compensate for limited nutrient availability[24,34], such as enhanced biosynthesis of cofactors, energy-efficient substrate utilization, and recycling of essential biomolecules. Conversely, the lower metabolic gene expression in eutrophic zooplankton suggests reduced pressure to

optimize nutrient uptake or balance elemental ratios, as abundant resources may alleviate the need for stringent metabolic regulation. These findings underscore how nutrient level heterogeneity may governs transcriptional pattern in metabolic pathways, with nutrient-poor environments driving physiological processes that prioritize resource optimization overgrowth.

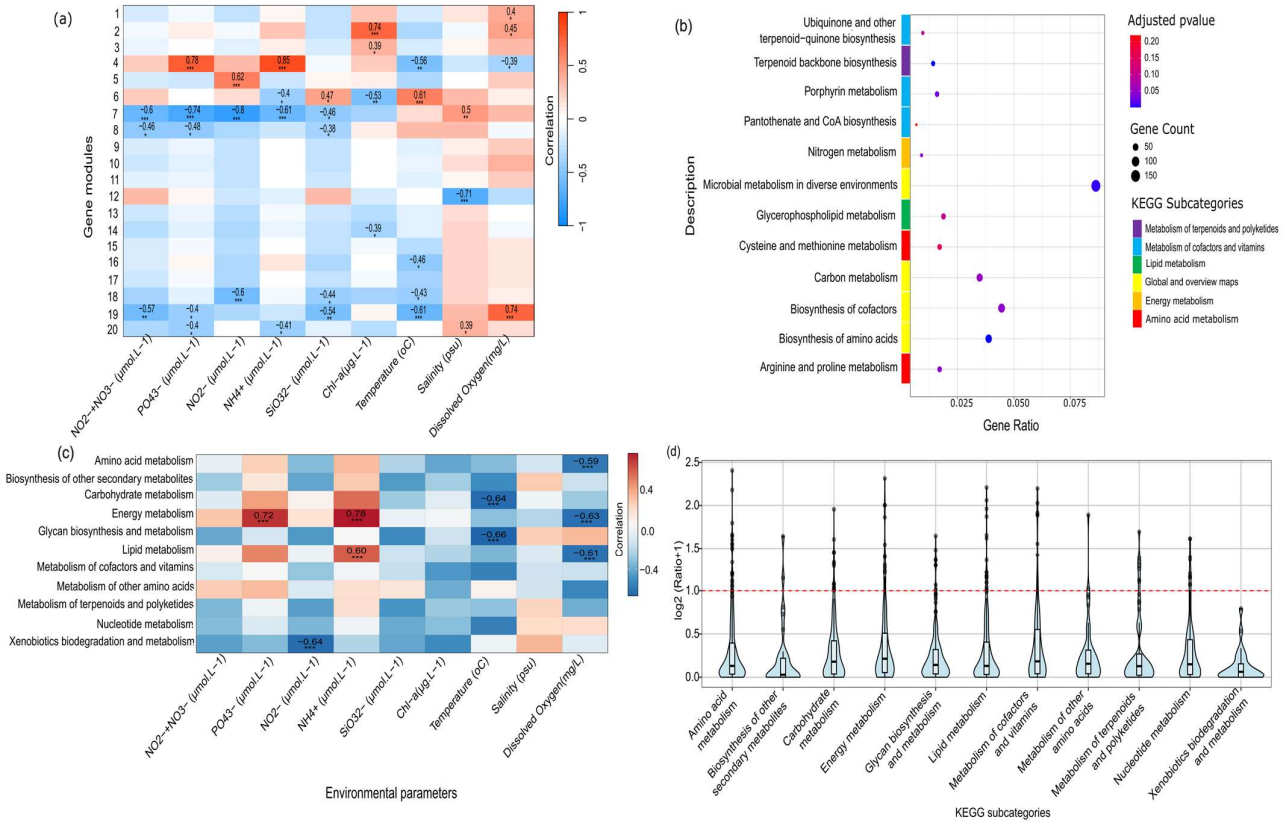

Fig 4. (a) Weighted gene co-expression network analysis illustrates the correlation between environmental parameters (x-axis) and groups of similarly expressed genes, assigned to gene modules (ME) (y-axis). The top numbers in each cell represent Pearson's correlation coefficients, while the bottom numbers indicate the p-values from the correlation test. The color of each cell reflects the strength of the correlation between the ME and environmental parameters, with red indicating a positive correlation and blue indicating a negative correlation. (b) Dot plot of KEGG enrichment analysis for module 4. Enrichment is calculated based on the number of KOs overrepresented in KEGG pathways compared to all annotated KOs. Significance is determined using false discovery rate-adjusted p-values of <0.05, with darker colors indicating higher significance. The gene count represents the number of KOs detected in each pathway, and different color annotations correspond to KEGG subcategories. (c) Correlation between environmental parameters (x-axis) and KEGG metabolism subcategories (y-axis). The top numbers in each cell represent Pearson's correlation coefficients, while the

bottom asterisk indicates the significant p value (* = p value < 0.05, ** = p value < 0.01, and *** = p value < 0.001) from the correlation test. (d) Ratios of KEGG subcategory profiles between eutrophic and mesotrophic nutrient regions. The red dashed line indicates a ratio of 1; values below 1 signify lower expression of metabolic gene functions in eutrophic regions relative to mesotrophic regions, while values above 1 indicate the opposite.

Zooplankton DOM molecules chemical properties highly correlated with metabolic gene functions

To investigate the molecular mechanism behind the compositional and functional properties variability of zooplankton-derived (DOM), we conducted Spearman correlation analyses between DOM molecular properties and metabolic gene functions. Our results demonstrate significant linkages between specific DOM chemical traits and zooplankton metabolic pathways (Fig. 5a). Specifically, unsaturated hydrocarbons showed positive correlations with amino acid and terpenoid/polyketides metabolism-related gene functions, while protein-like DOM molecules exhibited negative associations with carbohydrate, energy, and lipid metabolism genes. Meanwhile Lipid-like molecules demonstrated negative correlations with both energy metabolism and glycan biosynthesis/metabolism pathways. Furthermore, condensed hydrocarbon-like molecules were inversely related to terpenoid and polyketide metabolism genes. In contrast, carbohydrate and amino sugar-like compounds displayed positive correlations with energy metabolism and lipid metabolism genes, respectively.

Notably, we also observed significant positive correlations between the weighted average modified aromaticity index (AI_w_mean) and three metabolic pathways: carbohydrate metabolism, glycan biosynthesis and metabolism, and terpenoid/polyketide metabolism (Fig. 5b). The observed patterns indicate that specific metabolic processes may selectively influence the chemical properties of excreted DOM, particularly its aromatic content and molecular diversity. These findings suggest that zooplankton metabolic activity plays a crucial role in determining the molecular composition and aromaticity characteristics of DOM they released into the surrounding environment.

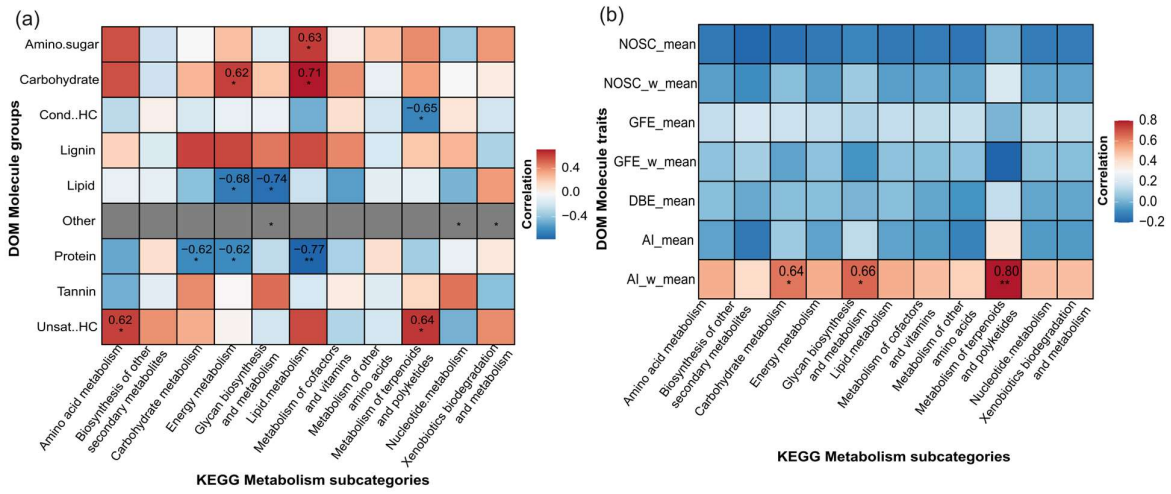

Fig 5. Association between DOM chemical properties with zooplankton gene functions. Spearman's rank correlations between zooplankton's KEGG metabolism subcategories and (a) DOM molecule groups, (b) DOM molecule traits. Significant correlations (* = p value < 0.05, ** = p value < 0.01) are indicated by the asterisks with the value inside the boxes represent correlation coefficient.

Implications for Nutrient Cycling and DOM Dynamics

While prior studies have extensively documented organismal physiological responses to environmental variabilities, the ecosystem-level consequences of such plasticity remain poorly resolved. Here, we bridge this gap by linking zooplankton-derived dissolved organic matter (DOM) molecular traits to transcriptional activity, revealing how metabolic gene functions shape DOM chemical properties and its role in coastal biogeochemical cycles.

Our findings demonstrate that spatial variability in zooplankton-derived DOM composition strongly correlated with transcriptional heterogeneity across sampling stations. Strong correlations between DOM molecular classes—particularly protein-like (labile, nitrogen-rich), lipid-like (semi-labile, carbon-dense), and unsaturated hydrocarbon-like (potentially more recalcitrant, reduced-carbon) compounds—and zooplankton metabolic gene expression profiles suggest that physiological responses to environmental gradients directly modulate DOM stoichiometry and bioavailability (Fig. 5a). This regulatory mechanism holds critical implications for coastal ecosystems as DOM molecular composition governs the bioavailability of bound macronutrients (e.g., nitrogen, phosphorus), which regulate microbial and phytoplankton productivity[35]. Higher or lower content of N and P within the zooplankton derived DOM compound will subsequently affect the availability of N and P in their living

habitat[36]. For instance, protein-like DOM, characterized by high nitrogen content and low C:N ratios, exhibited the highest inferred bioavailability. Its production correlated negatively with carbohydrate, energy, and lipid metabolism genes, suggesting preferential resource allocation to nitrogen-rich DOM under nutrient stress. These compounds likely fuel rapid microbial assimilation and short-term nutrient remineralization, enhancing localized nitrogen/phosphorus availability. Whereas lipid-like DOM showed intermediate bioavailability, reflecting its hydrophobic properties and higher C:N ratios. Its abundance correlated negatively with energy metabolism and glycan biosynthesis pathways, implying reduced allocation in lipid content under high-nutrient conditions. Unsaturated hydrocarbon-like DOM, linked positively to amino acid and terpenoid/polyketide metabolism genes, displayed the lowest bioavailability due to structural complexity (e.g., conjugated double bonds) and enzymatic recalcitrance. These compounds may accumulate as semi-refractory carbon pools, with degradation energetics governed by their elemental composition. For instance, DOM with higher oxygen-to-carbon (O:C) ratios (e.g., protein-like molecules) requires less energy to degrade, promoting microbial growth, while low O:C compounds (e.g., unsaturated hydrocarbons) increase respiration costs and limit carbon use efficiency[37].

Notably, zooplankton metabolic processes—particularly amino acid, carbohydrate, and terpenoid/polyketide metabolism—emerged as potential key drivers of DOM aromaticity (Fig. 5 b). This was shown by the high correlation between zooplankton metabolism gene functions activity level and the modified aromaticity indexes of the derived DOM, suggesting that zooplankton metabolism processes potentially facilitate the transformation of the released DOM into more or less aromatic compounds depending on the condition of their living environment. The strong correlation between transcriptional activity of these pathways and DOM modified aromaticity indices (AImod) implies a mechanistic link such as upregulated terpenoid/polyketide metabolism under mesotrophic conditions (relatively low nutrients) may enhance biosynthesis of aromatic secondary metabolites, increasing DOM aromaticity. Conversely, suppressed lipid or carbohydrate metabolism in eutrophic systems could favor less recalcitrant DOM production. This transcriptional tuning of DOM aromaticity has cascading effects as aromatic compounds exhibit greater biogeochemical recalcitrance due to stabilized conjugated π-electron systems[38], potentially extending DOM residence times in the water column and altering carbon export fluxes.

Conclusion

We integrated mass spectra and meta transcriptomics data to characterize the association between zooplankton derived DOM molecular properties and zooplankton physiological processes, and applied gene-centric analysis to elucidate the zooplankton metabolic potential that potentially shapes DOM heterogeneity. DOM chemical properties were significantly and broadly correlated with metabolism profiles of the zooplankton. These results provide evidence of how zooplankton active roles—not only intermediaries—in DOM cycling by modulating derived DOM molecules composition and aromaticity signatures through their metabolic plasticity which is strongly driven by the environmental parameters such as nutrient level.

Methods

Zooplankton field sampling

We collected zooplankton samples during summer 2023 (August) and 2024 (July) cruises in Pearl River Estuary-Hong Kong coastal waters from the surface (2 m depth, neustonic layer) layers using a custom-made plankton pump equipped with a 150 μm mesh sieve. The zooplankton samples were collected by deploying a plankton pump for 15 minutes with pumping rate 105 m^3/h. The sampling sites were selected based on the long-term monitoring stations of the Environmental Protection Department Hong Kong Monitoring Program for the Summer 2023 and Earth HK project Cruise stations for Summer 2024.We categorized the sampling sites into two distinct physicochemical zones (Fig. 1a) — mesotrophic and eutrophic — based on the TRIX trophic index, a well-established metric for evaluating aquatic trophic status[39–41].

Zooplankton derived dissolved organic matters release experiment

We conducted incubation experiments to evaluate the release of dissolved organic matter (DOM) by zooplankton during the Summer Cruise 2024. Initially, the collected zooplankton samples were rinsed with 0.22 μm-filtered ambient seawater. Subsequently, we incubated the zooplankton in a 2-liter Watson bottle containing 1 liter of 0.22 μm-filtered ambient seawater. The control treatment consisted of 0.22 μm-filtered ambient seawater without zooplankton. The incubation lasted for eight hours in the dark at in-situ temperature. Following the incubation period, we separated the zooplankton from the released DOM using a 0.22 μm Millipore membrane filter. The collected DOM was preserved at -20°C. This incubation experiment was conducted in triplicate to ensure the reliability of the results.

Zooplankton RNA preservation

We performed filtration using 10 μm nylon membrane to separate the collected zooplankton from both 2023 and 2024 summer cruises samples from the seawater. The zooplankton-containing filters were then preserved in Invitrogen™ RNAlater™ Stabilization Solution and stored at -80°C until processing for RNA extraction in the laboratory. Concurrently, water quality parameters including temperature (°C), salinity (PSU), and dissolved oxygen (%) were measured at each sampling station using a multi-parameter CTD-DO prior to zooplankton sampling. Additionally, nutrient concentration such as total nitrogen ($mmol.m^{-3}$), total phosphorus ($mmol.m^{-3}$), and silicate ($mmol.m^{-3}$), and chlorophyll-a ($\mu g.L^{-1}$) were also measured by laboratory analysis for the 2024 samples while for the 2023 we utilized complementary environmental data collected in the same sampling stations at the same month by the Environmental Protection Department (EPD) which are publicly available through the EPD data repository at Environmental Protection Interactive Centre : Marine Water Quality Monitoring Data.

DOM extraction and characterization by FT-ICR-MS

We use solid-phase extraction (SPE) to extract dissolved organic matter (DOM) from the 0.22 µm-filtered incubation water samples following well established protocols[42] with a recovery efficiency of around 40% for marine water DOC, and approximately 50% for freshwater water DOC. In summary, 500 mL of samples were filtered through GTTP filters (Millipore, 0.22 µm pore size, which were prerinsed three times with MilliQ water). The samples were then acidified to a pH of 2 using analytical-grade HCl and passed through a 1-g SPE cartridge (which had been prerinsed twice with methanol, MilliQ water, and MilliQ water at pH 2) (200 mg, Bond Elut, PPL, Agilent). The cartridges were desalted using 0.01 M HCl and dried with pure nitrogen gas. Finally, the SPE-DOM samples were eluted from the cartridge with 2 mL of 99.9% HPLC-grade methanol.

To characterize the molecular formulae of the SPE-DOM, methanol PPL extracts were analyzed with a 9.4 T Apex-ultra X FT-ICR MS (China University of Petroleum, CUP) under the negative mode with an electrospray ionization source (Bruker Apollo II). 128 scans were used in the mass range m/z 200 to 800 Da with the calibration of a series of established formula mass peaks (Data Analysis 3.4 software package)[43]. Peaks were selected when the signal/noise ratio (S/N) was greater than 4. Formula were assigned according to the formula rules[44], allowing for elemental compositions $C_{1-60}H_{1-120}O_{1-40}N_{0-3}S_{0-3}P_{0-1}$ with an error range of ± 1 ppm.

The annotated DOM molecules were assigned to biochemical compound categories based on the stoichiometry of C, H, and O for the following H:C and O:C ranges; lipids ($0 < O:C \leq 0.3$, $1.5 \leq H:C \leq 2.5$), unsaturated hydrocarbons ($0 \leq O:C \leq 0.125$, $0.8 \leq H:C < 2.5$), proteins ($0.3 < O:C \leq 0.55$, $1.5 \leq H:C \leq 2.3$), amino sugars ($0.55 < O:C \leq 0.7$, $1.5 \leq H:C \leq 2.2$), carbohydrates ($0.7 < O:C \leq 1.5$, $1.5 \leq H:C \leq 2.5$), lignin ($0.125 < O:C \leq 0.65$, $0.8 \leq H:C < 1.5$), tannins ($0.65 < O:C \leq 1.1$, $0.8 \leq H:C < 1.5$), and condensed hydrocarbons ($0 \leq O:C \leq 0.95$, $0.2 \leq H:C < 0.8$). Unnamed compounds (others) are calculated as the proportion of identified C that does not fit within the above defined H:C and O:C ranges[45]. Relative peak intensities were calculated based on the sum-normalized intensities of all assigned peaks in each sample[46].

RNA extraction, library preparation, and sequencing

The collected zooplankton samples preserved in RNAlater were used for zooplankton total RNA extraction by using FastPure® Cell/Tissue Total RNA Isolation Kit V2 following the manufacturer's protocol. For the purpose of RNA quantity and quality checking, Qubit, Biodrop, and Gel Electrophoresis were used prior to library creation and sequencing. In the RNA sequencing step, each sample was processed by isolating mRNA from one microgram of total RNA through poly-T oligo-attached magnetic bead selection. Subsequently, the mRNAs were reverse transcribed into cDNAs to construct RNA sequencing libraries with average insert sizes ranging from 250 to 300 bp, utilizing TruSeq™ RNA Sample Prep Kits (Illumina) in accordance with the manufacturer's guidelines. To ensure robustness and account for biological variability, two to three biological replicates were included for each sample, resulting in a total of 34 libraries. The metatranscriptomes were subjected to pair-end sequencing, producing 150 bp reads using the HiSeq™ 2000 sequencing platform (Illumina) following standard protocols.

Meta transcriptome de novo assembly, transcript abundance estimation, and functional annotation

The raw RNA sequence data quality was evaluated using FastQC Version 0.11.5[47], followed by quality trimming of the raw reads for poor quality bases and RNA-Seq adapters using Trimmomatic. The resulting paired-end clean sequence from all 34 libraries were simultaneously utilized to construct the meta-assembly using Trinity software v2.15.2 with default setting[48]. Raw reads were quality-filtered and assembled de novo, yielding 12,975,434 contigs. The assembled contigs then being used for open reading frame prediction by using

transdecoder against pfam database. Subsequently, the reading alignment and transcript abundance estimation was performed by using Salmon-based approach. In brief, the reads are first aligned to the assembled contigs using Bowtie2. The output BAM file is then used as the input for Salmon tool. Normalization of transcript per million (TPM) was performed based on library sizes using the “calcNormFactors (method = "TMM")” function to ensure accurate quantification of gene expression levels across the samples[49].

To obtain functional annotations for the assembled contigs, we utilized the Trinotate pipeline[50]. This process also needs protein sequence as one of the inputs for the functional annotations, consequently we used Transdecoder software packaged with Trinity to predict the protein translations from the contigs. Subsequently, the unigenes obtained were queried against several public databases, including SwissProt, Pfam-A, and Eggnog[51], with an E-value threshold of less than $1E^{-5}$ using the Trinotate pipeline. In this annotation process, BLAST 2.9.0+ was employed for gene assignments, using diamond BLASTp for protein queries and diamond BLASTx for nucleotide queries[52]. Both query sets were compared against the Uniprot/SwissProt database of reference proteins[53]. To identify conserved domains within the protein models, we utilized HMMER v.3.3 and the Pfam-A database[54,55]. The results were then loaded into Trinotate's SQL database. Trinotate used the top BLAST hits to populate the database with gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) terms.

Statistical analyses and data visualization

We employed the MetaboDirect pipeline[56] to conduct data exploration and statistical analysis, including the generation of van Krevelen diagrams, Venn diagrams, and visualizations of dissolved organic matter (DOM) relative abundance. The input data consisted of preprocessed DOM molecular formulas along with their corresponding monoisotopic peak intensities and *m/z* values. To evaluate molecular properties and potential decomposability, we computed several thermodynamic and molecular indices based on each peak’s elemental composition (equations provided in Supplementary Table 2). These indices included the Nominal Oxidation State of Carbon (NOSC), which describes the average carbon oxidation state of the DOM molecules; Gibbs free energy ($\Delta G°C$-ox or GFE), indicating the thermodynamic favorability of degradation[57];the Modified Aromaticity Index (AImod), reflecting aromaticity and carbon double-bond density[58];and the double bond equivalent (DBE), quantifying molecular unsaturation as a proxy for DOM aromatic structures[59].

During data exploration, MetaboDirect pipeline generated van Krevelen diagrams[60] to visualize molecular formula distributions, along with violin plots of the calculated thermodynamic indices. Statistical differences between eutrophic and mesotrophic DOM release were assessed using Tukey post hoc tests, while bar plots summarized molecular and elemental compositions across different regions. Additionally, pairwise comparisons were performed to identify unique and shared molecules between mesotrophic and eutrophic groups, employing van Krevelen diagrams, Venn diagrams, and stacked bar plots for visualization.

To investigate zooplankton gene functional profiles and their relationship with environmental variables, we conducted a weighted gene co-expression network analysis (WGCNA) using the WGCNA package (version 1.72-1) in R[60]. The analysis utilized a trimmed mean of M-values (TMM) normalized expression matrix of Kyoto Encyclopedia of Genes and Genomes (KEGG)-annotated transcripts as the input to identify gene modules exhibiting strong co-expression patterns. Network construction employed an unsigned topological overlap matrix (TOM), with module detection based on TOM-derived dissimilarity measures (1-TOM). Parameters were optimized for biological network properties, including a soft threshold power (β) of 30 to satisfy scale-free topology criteria, minimum module size of 70 genes, and a merge cut height of 0.4[22]. Module eigengenes were computed and correlated with both environmental parameters and zooplankton taxonomic group relative abundances. Gene enrichment analysis was performed using the clusterProfiler package in R to identify metabolic pathways associated with significant module eigengenes. Resulting p-values were adjusted for multiple comparisons using the Benjamini-Hochberg false discovery rate (FDR) correction[22,61].

We employed Spearman rank correlation analysis to examine relationships between KEGG metabolic subcategories and both environmental factors and DOM chemical characteristics, including molecular class compositions and molecular traits. To compare zooplankton metabolic processes between trophic states, we calculated eutrophic-to-mesotrophic ratios of total transcripts per million (TPM) values for each KEGG metabolic subcategory.

References

1. Carlson, C. A., Liu, S., Stephens, B. M. & English, C. J. Chapter 5 - DOM production, removal, and transformation processes in marine systems. in *Biogeochemistry of Marine Dissolved Organic Matter (Third Edition)* (eds. Hansell, D. A. & Carlson, C. A.) 137–246 (Academic Press, 2024). doi:10.1016/B978-0-443-13858-4.00013-7.
2. Kujawinski, E. B. The Impact of Microbial Metabolism on Marine Dissolved Organic Matter. *Annual Review of Marine Science* **3**, 567–599 (2011).
3. Dittmar, T. & Stubbins, A. 12.6 - Dissolved Organic Matter in Aquatic Systems. in *Treatise on Geochemistry (Second Edition)* (eds. Holland, H. D. & Turekian, K. K.) 125–156 (Elsevier, Oxford, 2014). doi:10.1016/B978-0-08-095975-7.01010-X.
4. Bertilsson, S. & Jones, J. B. 1 - Supply of Dissolved Organic Matter to Aquatic Ecosystems: Autochthonous Sources. in *Aquatic Ecosystems* (eds. Findlay, S. E. G. & Sinsabaugh, R. L.) 3–24 (Academic Press, Burlington, 2003). doi:10.1016/B978-012256371-3/50002-0.
5. Osterholz, H. *et al.* Deciphering associations between dissolved organic molecules and bacterial communities in a pelagic marine system. *ISME J* **10**, 1717–1730 (2016).
6. Wetzel, R. Detrital dissolved and particulate organic carbon functions in aquatic ecosystems. *Bulletin of Marine Science* (1984).
7. Cole, J. J., Findlay, S. & Pace, M. L. Bacterial production in fresh and saltwater ecosystems: a cross-system overview. *Marine Ecology Progress Series* **43**, 1–10 (1988).
8. Hygum, B. H., Petersen, J. W. & Søndergaard, M. Dissolved organic carbon released by zooplankton grazing activity-a high-quality substrate pool for bacteria. *Journal of Plankton Research* **19**, 97–111 (1997).
9. Møller, E. F. Production of dissolved organic carbon by sloppy feeding in the copepods Acartia tonsa, Centropages typicus, and Temora longicornis. *Limnology and Oceanography* **52**, 79–84 (2007).

10. Baines, S. B. & Pace, M. L. The production of dissolved organic matter by phytoplankton and its importance to bacteria: Patterns across marine and freshwater systems. *Limnology and Oceanography* **36**, 1078–1090 (1991).
11. Jumars, P. A., Penry, D. L., Baross, J. A., Perry, M. J. & Frost, B. W. Closing the microbial loop: dissolved carbon pathway to heterotrophic bacteria from incomplete ingestion, digestion and absorption in animals. *Deep Sea Research Part A. Oceanographic Research Papers* **36**, 483–495 (1989).
12. Steinberg, D. K. & Landry, M. R. Zooplankton and the Ocean Carbon Cycle. *Annu. Rev. Mar. Sci.* **9**, 413–444 (2017).
13. Lampert, W. Release of dissolved organic carbon by grazing zooplankton. *Limnology and Oceanography* **23**, 831–834 (1978).
14. Johannes, R. E. & Webb, K. L. Release of Dissolved Amino Acids by Marine Zooplankton. *Science* **150**, 76–77 (1965).
15. Aoki, S. *et al.* Characterization of fluorophores released from three kinds of lake phytoplankton using gel chromatography and fluorescence spectrophotometry. *Analytical Sciences* **24**, 1461–1467 (2008).
16. Wen, Z. *et al.* Composition of dissolved organic matter (DOM) in lakes responds to the trophic state and phytoplankton community succession. *Water Research* **224**, 119073 (2022).
17. Mangal, V., Stock, N. L. & Guéguen, C. Molecular characterization of phytoplankton dissolved organic matter (DOM) and sulfur components using high resolution Orbitrap mass spectrometry. *Analytical and bioanalytical chemistry* **408**, 1891–1900 (2016).
18. Maas, A. E. *et al.* Migratory Zooplankton Excreta and Its Influence on Prokaryotic Communities. *Front. Mar. Sci.* **7**, (2020).

19. De Corte, D. *et al.* Zooplankton-derived dissolved organic matter composition and its bioavailability to natural prokaryotic communities. *Limnology and Oceanography* **68**, 336–347 (2023).
20. Johnston, S. E. *et al.* Zooplankton release complex dissolved organic matter to aquatic environments. *Biogeochemistry* **157**, 313–325 (2022).
21. Mojib, N., Thimma, M., Kumaran, M., Sougrat, R. & Irigoien, X. Comparative metatranscriptomics reveals decline of a neustonic planktonic population. *Limnology and Oceanography* **62**, 299–310 (2017).
22. Perneel, M., Lagaisse, R., Mortelmans, J., Maere, S. & Hablützel, P. I. Seasonal metabolic dynamics of microeukaryotic plankton: a year-long metatranscriptomic study in a temperate sea. *mBio* **15**, e00383-24 (2024).
23. Berger, C. A., Steinberg, D. K. & Tarrant, A. M. Nutritional condition drives spatial variation in physiology of Antarctic lipid-storing copepods. *Ecology and Evolution* **14**, e70210 (2024).
24. Roncalli, V. *et al.* Physiological acclimatization in high-latitude zooplankton. *Molecular Ecology* **31**, 1753–1765 (2022).
25. Lai, R. W. S. *et al.* Hong Kong's marine environments: History, challenges and opportunities. *Regional Studies in Marine Science* **8**, 259–273 (2016).
26. Inter Research » MEPS » v245 » p111-122. https://www.int-res.com/abstracts/meps/v245/p111-122/.
27. Cai, Z. Spatiotemporal variability of water exchanges in the Pearl River Estuary by interactive multiscale currents. (2022).
28. Tao, W., Niu, L., Dong, Y., Fu, T. & Lou, Q. Nutrient Pollution and Its Dynamic Source-Sink Pattern in the Pearl River Estuary (South China). *Front. Mar. Sci.* **8**, (2021).

29. Yin, K. Monsoonal influence on seasonal variations in nutrients and phytoplankton biomass in coastal waters of Hong Kong in the vicinity of the Pearl River estuary. *Marine Ecology Progress Series* **245**, 111–122 (2002).

30. Qi, Y. *et al.* Deciphering dissolved organic matter by Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS): from bulk to fractions and individuals. *carbon res* **1**, 3 (2022).

31. Koch, B. P., Dittmar, T., Witt, M. & Kattner, G. Fundamentals of Molecular Formula Assignment to Ultrahigh Resolution Mass Data of Natural Organic Matter. *Anal. Chem.* **79**, 1758–1763 (2007).

32. Shakya, M., Lo, C.-C. & Chain, P. S. Advances and challenges in metatranscriptomic analysis. *Frontiers in genetics* **10**, 904 (2019).

33. Li, H.-Y. *et al.* The chemodiversity of paddy soil dissolved organic matter correlates with microbial community at continental scales. *Microbiome* **6**, 187 (2018).

34. Roncalli, V., Cieslak, M. C., Germano, M., Hopcroft, R. R. & Lenz, P. H. Regional heterogeneity impacts gene expression in the subarctic zooplankter Neocalanus flemingeri in the northern Gulf of Alaska. *Commun Biol* **2**, 1–13 (2019).

35. Tanentzap, A. J. & Fonvielle, J. A. Chemodiversity in freshwater health. *Science* **383**, 1412–1414 (2024).

36. Granéli, E., Carlsson, P. & Legrand, C. The role of C, N and P in dissolved and particulate organic matter as a nutrient source for phytoplankton growth, including toxic species. *Aquatic Ecology* **33**, 17–27 (1999).

37. Guillemette, F. & del Giorgio, P. A. Reconstructing the various facets of dissolved organic carbon bioavailability in freshwater ecosystems. *Limnology and Oceanography* **56**, 734–748 (2011).

38. Keiluweit, M. & Kleber, M. Molecular-Level Interactions in Soils and Sediments: The Role of Aromatic π-Systems. *Environ. Sci. Technol.* **43**, 3421–3429 (2009).

39. Vollenweider, R. A., Giovanardi, F., Montanari, G. & Rinaldi, A. Characterization of the trophic conditions of marine coastal waters with special reference to the NW Adriatic Sea: proposal for a trophic scale, turbidity and generalized water quality index. *Environmetrics* **9**, 329–357 (1998).

40. Primpas, I. & Karydis, M. Scaling the trophic index (TRIX) in oligotrophic marine environments. *Environ Monit Assess* **178**, 257–269 (2011).

41. Kang, J. *et al.* Biogeochemical factors regulating photosynthetically dissolved organic carbon produced by phytoplankton in the Taiwan Strait. *Environmental Research* **252**, 119090 (2024).

42. Dittmar, T., Koch, B., Hertkorn, N. & Kattner, G. A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnology and Oceanography: Methods* **6**, 230–235 (2008).

43. He, C. *et al.* In-house standard method for molecular characterization of dissolved organic matter by FT-ICR mass spectrometry. *ACS omega* **5**, 11730–11736 (2020).

44. Kujawinski, E. B. & Behn, M. D. Automated analysis of electrospray ionization Fourier transform ion cyclotron resonance mass spectra of natural organic matter. *Analytical chemistry* **78**, 4363–4373 (2006).

45. Sleighter, R. L. & Hatcher, P. G. The application of electrospray ionization coupled to ultrahigh resolution mass spectrometry for the molecular characterization of natural organic matter. *J. Mass Spectrom.* **42**, 559–574 (2007).

46. Yi, Y. *et al.* Unraveling the Linkages between Molecular Abundance and Stable Carbon Isotope Ratio in Dissolved Organic Matter Using Machine Learning. *Environ. Sci. Technol.* **57**, 17900–17909 (2023).

47. Andrews, S. FastQC: a quality control tool for high throughput sequence data. 2010. (2017).

48. Haas, B. J. *et al.* De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature protocols* **8**, 1494–1512 (2013).

49. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome biology* **11**, 1–9 (2010).

50. Bryant, D. M. *et al.* A tissue-mapped axolotl de novo transcriptome enables identification of limb regeneration factors. *Cell reports* **18**, 762–776 (2017).

51. Huerta-Cepas, J. *et al.* eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic acids research* **47**, D309–D314 (2019).

52. McGinnis, S. & Madden, T. L. BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic acids research* **32**, W20–W25 (2004).

53. UniProt: the universal protein knowledgebase in 2021. *Nucleic acids research* **49**, D480–D489 (2021).

54. Mistry, J. *et al.* Pfam: The protein families database in 2021. *Nucleic acids research* **49**, D412–D419 (2021).

55. Finn, R. D., Clements, J. & Eddy, S. R. HMMER web server: interactive sequence similarity searching. *Nucleic acids research* **39**, W29–W37 (2011).

56. Ayala-Ortiz, C. *et al.* MetaboDirect: an analytical pipeline for the processing of FT-ICR MS-based metabolomic data. *Microbiome* **11**, 28 (2023).

57. LaRowe, D. E. & Van Cappellen, P. Degradation of natural organic matter: a thermodynamic analysis. *Geochimica et Cosmochimica Acta* **75**, 2030–2042 (2011).

58. Chassé, A. W. *et al.* Chemical Force Spectroscopy Evidence Supporting the Layer-by-Layer Model of Organic Matter Binding to Iron (oxy)Hydroxide Mineral Surfaces. *Environ. Sci. Technol.* **49**, 9733–9741 (2015).

59. Koch, B. P. & Dittmar, T. From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter. *Rapid Communications in Mass Spectrometry* **20**, 926–932 (2006).

60. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).

61. Wright, R. M., Aglyamova, G. V., Meyer, E. & Matz, M. V. Gene expression associated with white syndromes in a reef building coral, Acropora hyacinthus. *BMC Genomics* **16**, 371 (2015).

Acknowledgements

The work described in this paper was substantially supported by the funding from Center for Ocean Research in Hong Kong and Macau (CORE), the joint research centre for ocean research between Qingdao National Laboratory for Marine Science and Technology (QNLM) and Hong Kong University of Science and Technology (HKUST). This study is also partially supported by Research Grants Council of the Hong Kong Special Administrative Region, China (project reference no. AoE/P-601/23-N). We thank J. Poon, L. Yuanmei, Y.yan for assistance with the fieldwork. We also thank M. Ye for lab works and J. Li for the bioinformatics assistance.