Vertical Stratification of Dissolved Organic Matter Linked to Distinct Microbial Communities in Subtropic Estuarine Sediments

Wenxiu Wang[1], Jianchang Tao[2], Ke Yu[3], Chen He[4], Jianjun Wang[5], Penghui Li[6,7], Hongmei Chen[8], Bu Xu[2], Quan Shi[4,9]* and Chuanlun Zhang[2,9,10,11]*

[1] State Key Laboratory of Marine Geology, Tongji University, Shanghai, China, [2] Department of Ocean Science and Engineering, Southern University of Science and Technology, Shenzhen, China, [3] School of Environment and Energy, Shenzhen Graduate School, Peking University, Shenzhen, China, [4] State Key Laboratory of Heavy Oil Processing, China University of Petroleum, Beijing, China, [5] State Key Laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, China, [6] School of Marine Science, Sun Yat-sen University, Guangzhou, China, [7] Southern Marine Science and Engineering Guangdong Laboratory, Zhuhai, China, [8] State Key Laboratory of Marine Environmental Science, College of Ocean and Earth Sciences, Xiamen University, Xiamen, China, [9] Shenzhen Key Laboratory of Marine Archaea Geo-Omics, Southern University of Science and Technology, Shenzhen, China, [10] Southern Marine Science and Engineering Guangdong Laboratory, Guangzhou, China, [11] Shanghai Sheshan National Geophysical Observatory, Shanghai, China

Dissolved organic matter (DOM) provides carbon substrates and energy sources for sediment microbes driving benthic biogeochemical processes. The interactions between microbes and DOM are dynamic and complex and require the understanding based on fine-scale microbial community and physicochemical profiling. In this study, we characterized the porewater DOM composition in a 300-cm sediment core from the Pearl River estuary, China, and examined the interactions between DOM and archaeal and bacterial communities. DOM composition were highly stratified and associated with changing microbial communities. Compared to bacteria, the amplicon sequence variants of archaea showed significant Pearson correlations ($r \geq 0.65$, $P < 0.01$) with DOM molecules of low H/C ratios, high C number and double bond equivalents, indicating that the distribution of archaea was closely correlated to recalcitrant DOM while bacteria were associated with relatively labile compounds. This was supported by the presence of auxiliary enzyme families essential for lignin degradation and *bcrABCD, UbiX* genes for anaerobic aromatic reduction in metagenome-assembled genomes of Bathyarchaeia. Our study demonstrates that niche differentiation between benthic bacteria and archaea may have important consequences in carbon metabolism, particularly for the transformation of recalcitrant organic carbon that may be predominant in aged marine sediments.

Keywords: FT-ICR MS, metagenomic, benthic archaea, recalcitrant DOM, estuarine sediment

INTRODUCTION

Dissolved organic matter (DOM) consists one of Earth's largest carbon exchangeable reservoirs, having major significance in biogeochemical cycles. In anaerobic sediments, the decomposition of particulate-attached organic matter to DOM by hydrolysis and depolymerization is a step-wise process and serves as the principal energy and carbon source for benthic organisms

(Burdige and Komada, 2015; Mahmoudi et al., 2017). However, mechanisms driving DOM-microbes coupling are largely unknown. DOM is highly complex in composition and consists of a mixture of organic compounds with different mass size and functionality (Burdige and Komada, 2015; Canuel and Hardison, 2016). The major composition of DOM is uncharacterized due to limitation of extraction and analytical methods. Nevertheless, advanced spectroscopic tools like Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) is capable of resolving thousands of DOM compounds with high mass resolution and accuracy and provides significant insights into the molecular composition of DOM (Oni et al., 2015; Schmidt et al., 2017).

The dynamics of DOM-microbe interactions could be investigated from geochemical and microbial aspects. Geochemical studies have examined vertical variation of porewater DOM in marine sediments, which showed the increase of recalcitrant carbon such as autochthonous humic-like fluorescent components, carboxyl-rich alicyclic molecules (CRAM) and aromatic compounds (Oni et al., 2015; Chen et al., 2017; Fox et al., 2018). The observed stratification of porewater DOM may attribute to selective biodegradation of labile DOM as well as diagenetic transformation of organic carbon (Chen et al., 2017; Schmidt et al., 2017). Consequently, refractory DOM constitutes the majority (>97%) of the total porewater DOM in anaerobic sediments (Burdige et al., 2016). However, refractory DOM could still be available for organisms in deep layers with regard to the consumption of O-rich molecules in sulfate-depleted conditions (Oni et al., 2015; Schmidt et al., 2017; Gan et al., 2020). Abiotic processes like adsorption/desorption of DOM compounds on redox-sensitive minerals may affect DOM composition but appear to be quite slow (Burdige, 2007; Chen and Hur, 2015). Thus, porewater DOM undergoing diverse interactions is dynamically changed in its composition and concentration, which requires further understanding.

Meanwhile, microbial studies have examined stratified distribution of microbial structures on a global scale (Parkes et al., 2014; Petro et al., 2017); for example, Gamma-, Delta- and Alphaproteobacteria account for a high fraction of the microbial community in surface sediments, while bacterial phyla including Chloroflexi and Planctomycetota, archaeal members Bathyarchaeia often dominate in subsurface sediments. Roughly 90% of the microbial activity may occur in subsurface sediments (Parkes et al., 2014), in which DOM-microbes interactions are diverse and fundamental for benthic carbon cycle (Kujawinski, 2011). Bathyarchaeia and Chloroflexi have been proposed to play a role in the degradation of recalcitrant organic carbon (Wasmund et al., 2014; Yu et al., 2018; Saw et al., 2020). Co-existence of highly unsaturated DOM molecules and Bathyarchaeia, Chloroflexi was also detected in subsurface sediments from Helgoland Mud Area (Oni et al., 2015). Moreover, a significant correlation between archaea and humic-like fluorescent DOM (FDOM) was observed in our previous work (Wang W. et al., 2020), indicating that archaea were potentially involved in sedimentary carbon transformation. While it is well known that DOM fuels the growth of microorganisms in sediments, specific interactions between microorganisms and DOM composition are poorly known.

Estuarine sediments consist a major reservoir of organic carbon sourced from terrestrial plants, autochthonous algae, macrophytes, microbial and anthropogenic activities (Burdige, 2007; Chen and Hur, 2015; Canuel and Hardison, 2016), in which about 45% of global carbon were estimated to be buried (Hedges and Keil, 1995). In this study, we used the state-of-art approaches of FT-ICR MS to characterize DOM molecular composition in a 300-cm sediment core from Pearl River Estuary, China. This technique was combined with high-throughput sequencing such as metagenomics to further examine the linkage between DOM and microbial community structure and function. We aimed to understand (1) the factors that influence vertical composition of sedimentary DOM and the DOM variations coupled with microbial community, (2) the characteristics of DOM associated with bacterial and archaeal communities in sediments, and (3) the metabolic potential of bacteria and archaea participating in sedimentary carbon degradation. Our results revealed the vertical stratification of DOM composition and the linkages between DOM and microbial composition, the latter of which was manifested by the different roles of bacterial and archaeal communities in sediment recalcitrant carbon transformation.

MATERIALS AND METHODS

Sample Collection

In October 2017, a sediment core of 300 cm was collected from near Guishan Island (Pearl River Estuary) (22.1315 N, 113.8055 E). This core was sectioned into 5-cm intervals for 5–100 cm and into 10-cm intervals for 100–300 cm below sediment surface. Bacterial and archaeal distributional patterns were examined by characterizing their 16S rRNA gene abundances and diversities along geochemical profile. Briefly, we detected geochemical species including NO_3^-, Mn(II), Fe(II), SO_4^{2-}, NH_4^+, dissolved inorganic carbon (DIC) and its stable carbon isotope ($\delta^{13}C_{DIC}$), dissolved organic carbon (DOC), total organic carbon (TOC) and its stable carbon isotope ($\delta^{13}C_{TOC}$), and FDOM (**Supplementary Table 1**). Amplicon sequencing were conducted using the universal prokaryotic primers 515FB (GTGYCAGCMGCCGCGGTAA) and 806RB (GGACTACNVGGGTWTCTAAT) (Caporaso et al., 2012). The used primers cover almost uniformly all major bacterial and archaeal phyla, with the coverage reaching 83.5% for archaea and 83.6% for bacteria[1]. Chemical analyses of environmental variables and 16S rRNA gene amplicon sequencing have been reported in Wang W. et al. (2020). In this study, taxonomy of amplicon sequence variant (ASV) was reassigned using the Silva 138 99% Operational Taxonomic Units database footnote[1] to update the taxonomic classification.

---

[1]https://www.arb-silva.de

Solid-Phase Extraction (SPE) of DOM

Sediment samples were stored at −80℃ before porewater extraction. Porewater samples was obtained by centrifuging ∼100 g sediments with the procedure set as 7000 rpm, 10 min and 4℃, and then filtering using a 0.22-μm Millipore PES membrane (Thermo Fisher Scientific). DOM was extracted using stryrene-divibyl-polymer type PPL cartridges (Agilent Bond Elut PPL, 200 mg, 3 mL) as described in Dittmar et al. (2008). Cartridges were activated using 6 mL MeOH (HPLC grade) and washed using at least 9 mL pH 2 ultrapure water. Porewater samples were acidified to pH 2 with HCl (G.R.) and then passed through the PPL cartridges by gravity at a flow rate of about 2 mL/min. Before DOM elution, the columns were rinsed with at least 9 mL pH 2 ultrapure water to remove the salt and dried under soft ultrapure nitrogen gas. The SPE-DOM was finally eluted with 5 mL MeOH (HPLC grade) and stored at −20℃ before MS analysis.

Molecular Analysis of DOM by FT-ICR MS

Mass spectra were obtained using a 9.4 T Apex Ultra FT-ICR MS instrument (Bruker) equipped with negative ion electrospray ionization (ESI) source at the State Key Laboratory of Heavy Oil Processing, China University of Petroleum, Beijing, China. The DOM samples were diluted to a final concentration of 100 mg/L and injected into the ESI source at a rate of 250 μL/h. The international standard substance, SRFA (Suwannee River fulvic acids) was used as quality control to optimize condition of FT-ICR MS before sample analysis (He et al., 2020). The typical operating condition for negative-ion ESI analysis was set as: 3.5 kV spray shield voltage, 4.0 kV capillary column introduced voltage, and −320 V capillary column end voltage. The ions accumulated in the collision cell for 0.2 s then transferred into the ICR cell with a 1.1 ms time-of-flight (ToF). The ion transformation parameter for the quadrupole (Q1) was optimized at m/z 300. A total of 128 scans with 4 M word size were accumulated to enhance the signal-to-noise ratio and mass spectra were evaluated in the range from 200 to 800 Da.

The mass spectrometer was calibrated using sodium formate and then recalibrated with a known homologous mass series of the SRFA. Molecular formulae were assigned to masses with s/n > 4 and mass accuracy < 1 ppm using in-house software (He et al., 2020), and those with outlier values or identified as contaminations were manual removed. Relative abundances of DOM formulae were calculated by normalizing the peak intensities to the sum of FT-ICR MS peak intensities. For each assigned formula, the double bond equivalent [DBE = 1+1/2(2C-H+N+P)] (Koch and Dittmar, 2006) and modified aromaticity index [AI-mod = (1+C-1/2O-S-1/2H)/(C-1/2O-S-N-P)] (Koch and Dittmar, 2016) were calculated as a measure for the degree of unsaturation and aromaticity. For each sample, bulk molecular parameters were calculated based on magnitude-weighted average (Sleighter and Hatcher, 2007); for example, the nominal oxidation state of carbon [NOSC = 4-(4C+H-3N-2O-2S)/C] was calculated to indicate the average oxidation state (Riedel et al., 2012) and marked as $NOSC_{wa}$. The degradation index (I_{DEG}) was calculated to assess the degradation state based on intensities of ten molecular formulae (Flerus et al., 2012). In addition, the assigned molecular formulae were further classified to seven compound categories based on their unique O/C and H/C ratios according to Yuan et al. (2017). The following ranges were used: lipids (0 < O/C ≤ 0.3, 1.5 ≤ H/C ≤ 2.0), aliphatic/proteins (0.3 ≤ O/C < 0.67, 1.5 ≤ H/C ≤ 2.2), lignins/CRAM-like structures (0.1 ≤ O/C ≤ 0.67, 0.7 ≤ H/C ≤ 1.5), carbohydrates (0.67 < O/C ≤ 1.2, 1.5 ≤ H/C ≤ 2.4), unsaturated hydrocarbons (0.7 ≤ O/C ≤ 1.5, 0 < H/C ≤ 0.1), aromatic structures (0 < O/C ≤ 0.67, 0.2 ≤ H/C < 0.7), and tannins (0.67 < O/C ≤ 1.0, 0.6 ≤ H/C ≤ 1.5). The limitation of the categorization of compound groups from FT-ICR MS should be noted, as it is only based on the elemental ratios without structural information of DOM (Li et al., 2019).

Metagenomic Sequencing

Six samples from the depths of 5, 20, 70, 130, 200, and 300 cm were selected for metagenomic sequencing. Total genome DNA was extracted as detailed previously (Wang W. et al., 2020). DNA was sheared into 350 bp using the Covaris M220 Focused-ultrasonicator (Covaris, Woburn, MA, United States), purified using MinkaGene Gel Extraction Kit (mCHIP, Guangzhou, China) and quantified using Qubit 3.0 Fluorometer (Thermo Fisher Scientific). Illumina libraries were constructed from about 100 ng DNA using NEB Next® Ultra™ DNA Library Prep Kit for Illumina® (New England Biolabs, United States) according to the manufacturer's instructions and index codes were added to sequences to each sample. Sequencing was performed on an Illumina Hiseq X-ten platform using the paired end 2 × 150 bp run-type mode at the Guangzhou Magigene Biotechnology (Guangzhou, China).

Sequence Assembly, Genomic Binning, and Functional Annotation

Paired-end reads were processed using Trim galore (v0.5.0) (Krueger, 2012) with default settings. After removing adapter sequences and reads with a length shorter than 20 bp or with average quality score lower than Q20, about 77 ∼ 96 G clean bases per sample were generated. Trimmed reads were then assembled using megahit (v.1.1.3) (Li et al., 2016) with the following options: –presets meta-sensitive for individual assembly and –min-count 1 –k-list 25, 29, 39, 49, 59, 69, 79, 89, 99, 109, 119, 129, and 141 for mixed assembly which combine high-quality reads from six depths. The quality of the metagenomic assembly was evaluated using Quast (v.4.6.3) (Mikheenko et al., 2016) and the summary statistics are provided in **Supplementary Table 2**.

Individual assemblies and mix assembly with contigs longer than 1,000 bp were provided as input for metagenomic binning. BASALT (Binning Across a Series of AssembLies Toolkit) (Yu et al., 2021) was used to bin contigs from all seven assemblies via MetaBAT (v.2.12.1) (Kang et al., 2019), Maxbin (v.2.2.4) (Wu et al., 2016), and CONCOCT (v.1.1.0) (Alneberg et al., 2014) with different sensitivity parameters. Genomic dereplication and reassembly modules were conducted using SPAdes (v.3.13.1) (Bankevich et al., 2012) with careful mode employed to ensure higher quality and accuracy of the individual reconstructed metagenome-assembled genomes

(MAGs, simplified as "genome" below). The completeness and contaminatio n of the genomes were then estimated using CheckM (v.1.0.11) (Parks et al., 2015) and high-quality genomes (completeness ≥ 70% and contamination ≤ 5%) were retained in this study. Phylogenetic analysis of reconstructed genomes is detailedly described in **Supplementary Text 1**.

Putative coding DNA sequences were predicted for each genome using prodigal (v.2.6.3, -meta) (Hyatt et al., 2012). Averagely 1,660 predicated proteins were obtained per archaeal genomes (min = 745 and max = 4,220) and 2,781 per bacterial genomes (min = 721 and max = 5,790). Quantification of these proteins across six metagenomes were performed using Salmon (v.0.7.2) (Patro et al., 2017) by calculating reads per million, which is similar to the calculation of TPM (Transcripts Per Million) in metatranscriptome. Carbohydrate-active enzymes (CAZymes) were identified using HMM models from dbCAN (Huang et al., 2018) with an e-value threshold of 1e-15 and coverage of 0.35. Predicated proteins were further blastp (e-value ≤ 1e-5, bit score ≥ 50, sequence similarity ≥ 30%) against KEGG and Uniprot databases for searching matches to aromatic degradation genes.

Statistical Analysis

The relationships between sediment depth and indices of SPE-DOM were explored with linear and exponential models using the "basicTrendline" package (Mei et al., 2018). Linear discriminant analysis effect size (LEfSe) (Segata et al., 2011) was used to obtain DOM formulae significantly distinct between the upper and deep layers. The samples between the depth of 250 and 270 cm were unsuccessfully in sequencing and thus excluded from subsequent analysis. Non-metric multidimensional scaling (NMDS) and principal coordinates analysis (PCoA) were used to examine the changes of DOM or microbial compositions based on Bray-Curtis dissimilarity (Bray and Curtis, 1957). All available environmental parameters and the first two principal coordinate axes (that is, PCo1 and PCo2) of microbial community composition were fit to ordination of NMDS of DOM composition using envfit function with Monte Carlo permutation test (permutation = 999). Canonical correlation analysis (CCorA) was performed using the first five PCoA axes (PCo!-5) of of microbial community and DOM compositions, which explained 64.7 and 82.1% variations, respectively, as described in Osterholz et al. (2016). Optimal model of distance-based redundancy analysis (dbRDA) was obtained with variables selected through forward step-wise selection based on 999 Monte Carlo permutations (Parvin et al., 2019). Influence of environmental drivers and microbial community on DOM composition were then estimated using dbRDA variance partitioning (Zhao W. et al., 2019). The above analyses were conducted using vegan packages (Oksanen et al., 2013). Moreover, Pearson's correlation analysis was performed between microbial ASVs and DOM molecules occurring in over one-third of the samples. Network analysis was conducted based on significant positive correlations ($r \geq 0.65$, $P < 0.01$) between microbial ASVs and DOM molecules having a relative abundance (among total DOM molecules) greater than 0.05%. Topological properties of network were calculated in Gephi (v0.9.2) (Bastian

et al., 2009). Network was generated by grouping the same taxonomic or DOM groups to provide better visualization. Modules of highly interconnected nodes were identified using MCODE (v.1.6) (Bader and Hogue, 2003) Cytoscape plug-in with default parameters in Cytoscape (v.3.6.1) (Shannon et al., 2003).

Data Availability

Source sequencing data, reconstructed genomes, and sample information are available in the NCBI under the BioProject ID PRJNA575161.
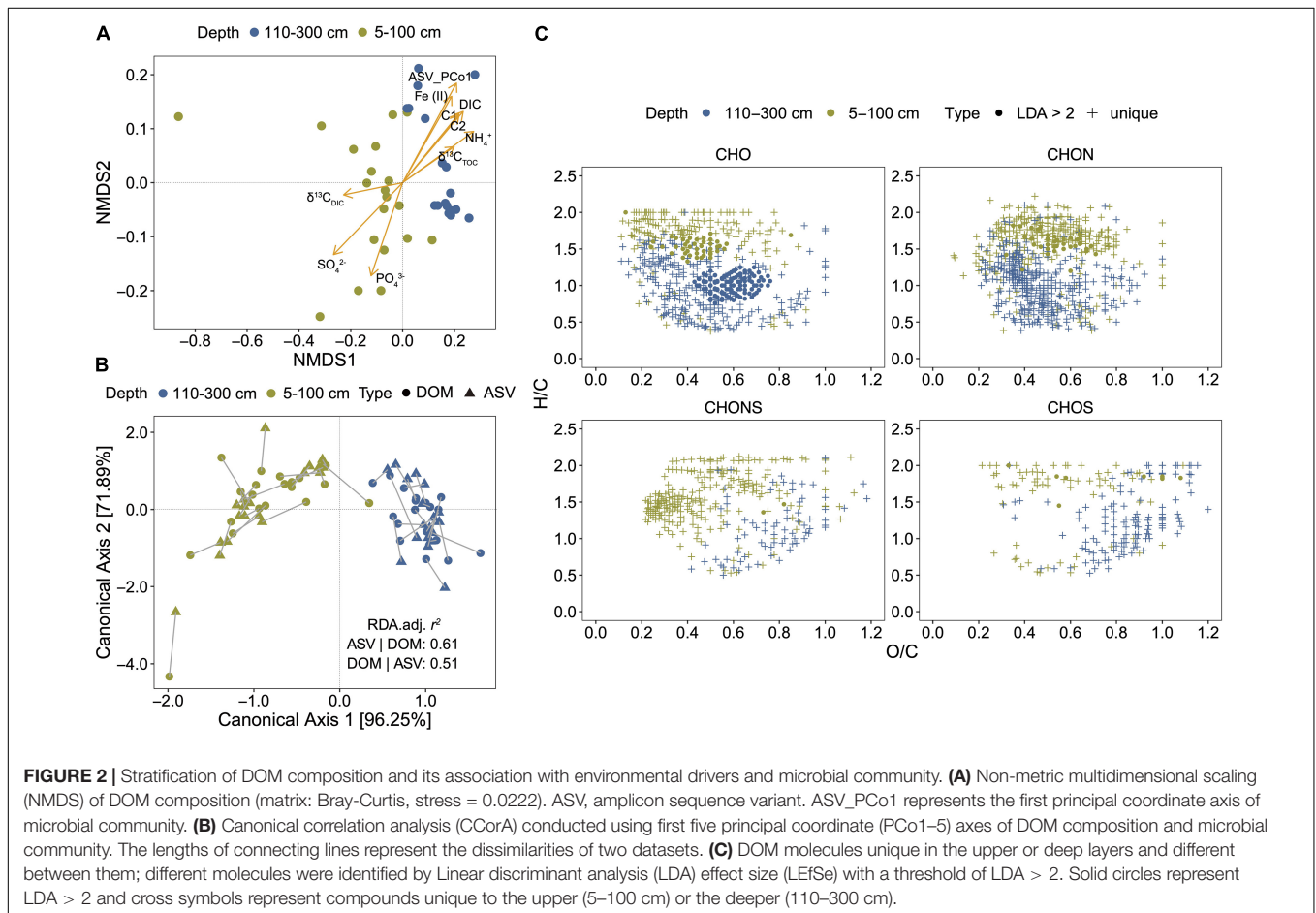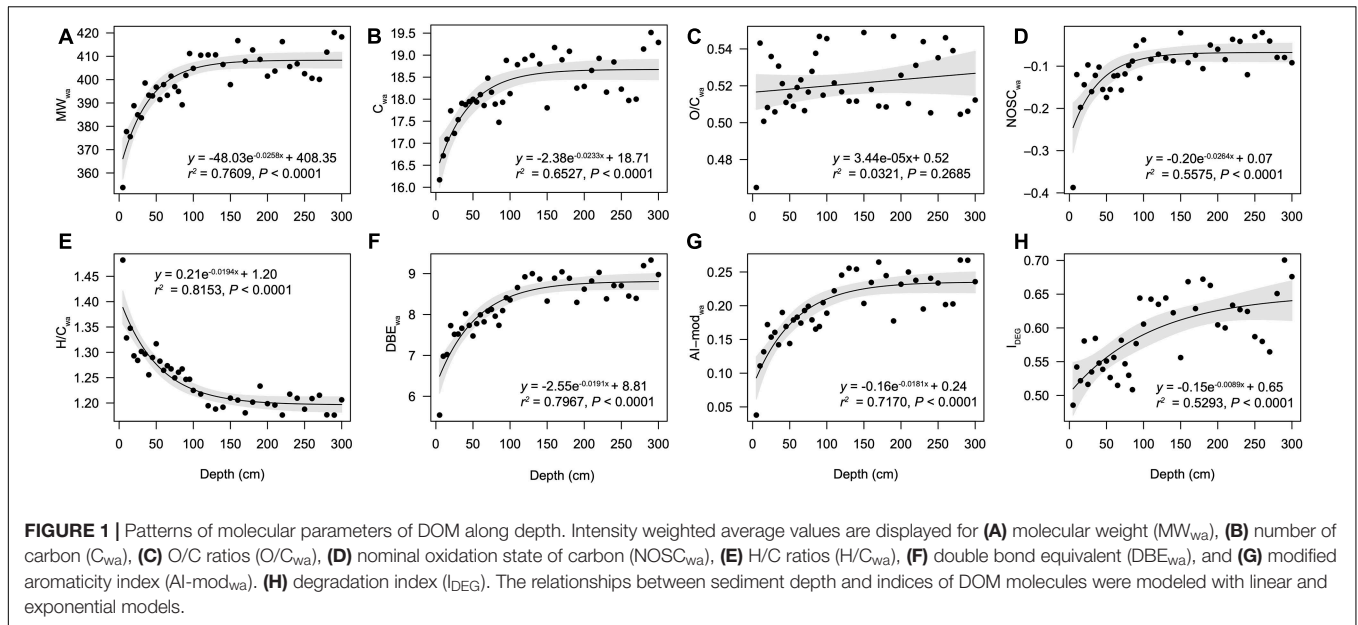
RESULTS

Depth Profile of DOM Molecular Features

A total of 7,830 unique molecular formulae were identified via FT-ICR MS (**Supplementary Table 3**). The formulae number of each sample ranged from 2,840 to 5,252, and significantly increased toward deep sediments ($r = 0.91$, $P < 0.01$). Generally, the DOM compositional formulae in all but one samples were represented primarily by CHO (58.1 ± 5.5%), followed by CHON (28.2 ± 3.0%), CHOS (11.1 ± 5.9%) and CHONS (2.6 ± 1.0%), and except for the 5-cm depth, where CHOS was the most abundant compounds (36.0%) (**Supplementary Figure 1A**). Bulk parameters showed tremendous changes in the surface sample (**Figures 1A–H**) and exhibited exponential-like profiles expect for the oxygen saturation index (O/C_{wa}). The molecular weight (MW_{wa}) increased from 353.8 Da at surface to 420.2 Da in the deep layers and the number of carbon (C_{wa}) increased from 16.2 to 19.5 downcore. The O/C_{wa} varied randomly while $NOSC_{wa}$ showed a slight increase below 5-cm depth. The hydrogen saturation index (H/C_{wa}) dropped from 1.48 at the surface to 1.19 at the 120-cm depth, and stayed relatively constant in deeper layers. The variations of these indices imply the accumulation of relatively aged and recalcitrant molecules toward the deep layers. This is further supported by the depth profiles of DBE_{wa}, $AI\text{-}mod_{wa}$, and I_{DEG}, which showed the increasing depth patterns in the unsaturated, aromatic and refractory compounds. When all the molecules were classified by specific elemental compositions, lignins/CRAM-like structures dominated at all depths (**Supplementary Figure 1B**), averagely accounting for 69.5% (7.5%). The decreases of aliphatic/proteins (12.7 ± 5.0%), lipids (3.2 ± 3.9%), and carbohydrates (3.5 ± 1.7%) were accompanied by the increases of aromatic structures (2.1 ± 0.6%) and tannins (8.9 ± 2.0%) (**Supplementary Figure 1B**). Specifically, aromatic structures and tannins increased from 0.66 and 4.99% at the surface to 3.24 and 11.89% at 250-cm depth, respectively.
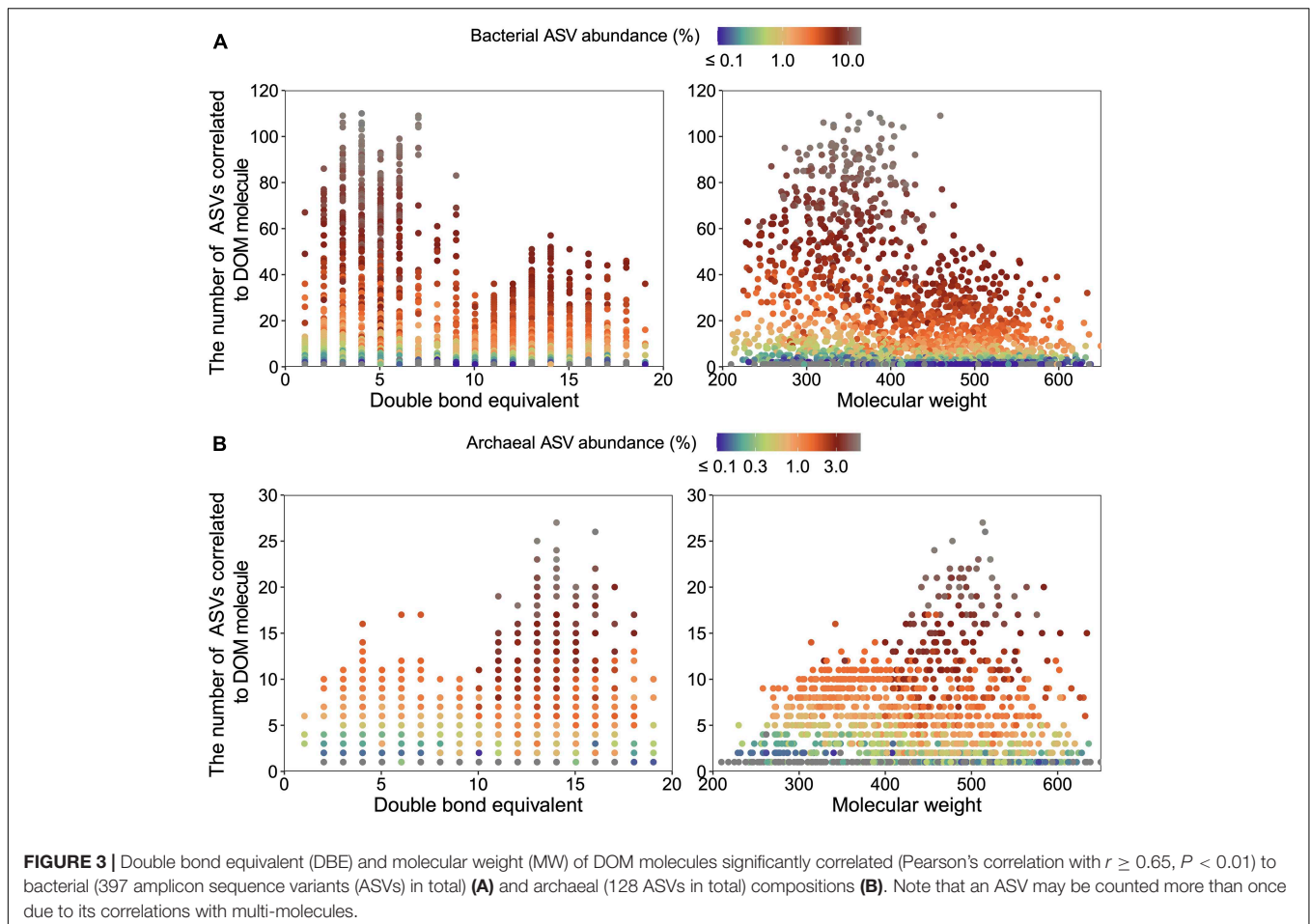
Stratification of DOM Molecular Composition and Its Relationship With Microbial Community

The NMDS based on the relative abundance of DOM molecules revealed a separation of samples with depth (5–100 cm and 110–300 cm) (**Figure 2A**). The explanatory variables such as SO_4^{2-}, NH_4^+, DIC, humic-like FDOM (C1 and C2), Fe (II), and

FIGURE 1 | Patterns of molecular parameters of DOM along depth. Intensity weighted average values are displayed for **(A)** molecular weight (MW$_{wa}$), **(B)** number of carbon (C$_{wa}$), **(C)** O/C ratios (O/C$_{wa}$), **(D)** nominal oxidation state of carbon (NOSC$_{wa}$), **(E)** H/C ratios (H/C$_{wa}$), **(F)** double bond equivalent (DBE$_{wa}$), and **(G)** modified aromaticity index (AI-mod$_{wa}$). **(H)** degradation index (I$_{DEG}$). The relationships between sediment depth and indices of DOM molecules were modeled with linear and exponential models.

FIGURE 2 | Stratification of DOM composition and its association with environmental drivers and microbial community. **(A)** Non-metric multidimensional scaling (NMDS) of DOM composition (matrix: Bray-Curtis, stress = 0.0222). ASV, amplicon sequence variant. ASV_PCo1 represents the first principal coordinate axis of microbial community. **(B)** Canonical correlation analysis (CCorA) conducted using first five principal coordinate (PCo1–5) axes of DOM composition and microbial community. The lengths of connecting lines represent the dissimilarities of two datasets. **(C)** DOM molecules unique in the upper or deep layers and different between them; different molecules were identified by Linear discriminant analysis (LDA) effect size (LEfSe) with a threshold of LDA > 2. Solid circles represent LDA > 2 and cross symbols represent compounds unique to the upper (5–100 cm) or the deeper (110–300 cm).

PCo1 of microbial Bray-Curtis distances (ASV_PCo1) ($r > 0.7$, $P < 0.01$) may have significantly shaped the DOM composition (**Figure 2A**). dbRDA were conducted to estimate the contribution of environmental variables and microbial community in DOM composition. Three environmental factors (SO$_4^{2-}$, δ^{13}C$_{DIC}$, and δ^{13}C$_{TOC}$) and one microbial feature (ASV_PCo1) were

FIGURE 3 | Double bond equivalent (DBE) and molecular weight (MW) of DOM molecules significantly correlated (Pearson's correlation with $r \geq 0.65$, $P < 0.01$) to bacterial (397 amplicon sequence variants (ASVs) in total) **(A)** and archaeal (128 ASVs in total) compositions **(B)**. Note that an ASV may be counted more than once due to its correlations with multi-molecules.

selected for optimal dbRDA model (**Supplementary Figure 2A**). Totally 44.9% of DOM variance could be explained, in which microbial feature and environmental factors accounted for 3.15 and 15.96% (**Supplementary Figure 2B**), respectively; and 25.81% together. To further characterize the association between DOM and microbial compositions, CCorA was then performed using PCo1-5 axes of DOM and microbial Bray-Curtis matrices (82.1 and 64.7% of the total variance, respectively) (**Figure 2B**). Highly significant relationships between DOM and microbial compositions were observed by CCorA (Pillai's trace statistic = 1.704, $P < 0.001$) with high canonical correlations found on the first two axes (0.963 for canonical axis 1, 0.719 for canonical axis 2).

Based on the depth separation of DOM composition, we divided all samples into two groups (5–100 cm and 110–300 cm). LEfSe analysis determined 210 molecules significantly contributing to the differences in DOM composition between the two groups (**Figure 2C** and **Supplementary Table 4**). Most of them belonged to CHO compounds with decreased H/C ratios and increased O/C ratios in deep layers. CHON compounds occurred only in the upper layers. In addition, there were 749 molecules unique in the upper layers, which belonged to aliphatic/proteins, lignins/CRAM-like structures or carbohydrates; and 805 molecules unique in the deep

layers, which belonged to lignins/CRAM-like structures, tannins or aromatic structures (**Figure 2C** and **Supplementary Table 4**). Unique CHOS and CHONS compounds in deep layers were characterized by higher O/C ratios than CHO and CHON compounds.

DOM Molecules Correlated With Bacterial and Archaeal Communities

Microbial ASVs (1,031, accounting for 75.72% of the total abundance) and DOM molecules (4,994, accounting for 97.96% of the total abundance) occurring over one-third of the samples were selected for Pearson's correlation analysis; molecular features (e.g., DBE and MW) of formulae correlated with bacterial and archaeal communities were then demonstrated. Pearson's correlation analysis determined 397 bacterial and 128 archaeal ASVs significantly positively correlated with DOM molecules ($r \geq 0.65$, $P < 0.01$). Most of bacterial ASVs were associated with molecules whose DBE was less than 10 and MW less than 450 Da, approximately (**Figure 3A** and **Supplementary Table 5**). In contrast, archaeal ASVs with increasing abundance were associated with molecules of high DBE (in the range of 11–18) and high MW (500 Da), approximately (**Figure 3B** and **Supplementary Table 5**). These results indicated distinct

ecological niches of bacterial and archaeal communities though more experimental evidence are needed. Furthermore, the average values of DBE, MW, AI-mod, H/C, and O/C ratios of 2,038 bacterial-correlated and 1,593 archaeal-correlated molecules were comparable while the weighted average had substantial difference except for O/C ratios (**Table 1**). Moreover, DOM molecules correlated to archaea mainly belonged to the compounds of N_1O_{10}-N_1O_{13} and O_{11}-O_{14} (**Figure 4B** and **Supplementary Table 5**), while those correlated to bacteria were more distributed in the compounds of $N_1O_{11}S_1$, N_1O_5-N_1O_{12}, O_6-O_9, N_2O_X-N_3O_X, and O_XS_X (**Figure 4A** and **Supplementary Table 5**). Overall, bacteria-correlated DOM was more widely distributed and archaea-correlated DOM was more concentrated with the characteristics of O-enrichment.

Bacterial and Archaeal Groups Differentially Associated With DOM Molecules

The DOM molecules with relative abundance over 0.05% were considered for network analysis to further explore the interconnections between DOM composition and microbial communities (**Supplementary Table 6**). The original network was composed of 457 nodes and 3,397 edges with an average number of neighbors of 14.85 (**Supplementary Figure 3**). Network was then generated by grouping categories for better visualization (**Figure 5A**). The lignins/CRAM-like structures consist of 120 DOM molecules showed most correlations with Bathyarchaeia, followed by Dehalococcoidia, Anaerolineae, Planctomycetota, and Desulfatiglandales, whereas the aliphatic/proteins largely associated with Gammaproteobacteria, Desulfobulbia and Desulfobacterales. Tight connections between diverse species and tannins were also found, especially in Bathyarchaeia, Anaerolinaea, and Dehalococcoidia. The remaining DOM belonging to lipids exhibited less associations with microbial species compared to others. In addition, although DOM molecules belonging to aromatic structures were not included in the network, they exhibited close correlations with archaeal groups (70 DOM molecules significantly correlated to 59 ASVs) such as Bathyarchaeia, and bacterial groups (90

DOM molecules significantly correlated to 114 ASVs) such as Dehalococcoidia.

Four significant clusters with the scores ranging between 6.36 and 3.20 (**Figures 5B–E**) were identified by MCODE analysis. Cluster ranks 1–3 consisted of 21 DOM molecules and 18 microbial ASVs. These molecules were enriched in the 110–300 cm depth interval, which had H/C ratios less than 1.05 and were associated with ASVs affiliated to Bacthyarchaeia, Dehalococcoidia, Desulfatiglandales and Sva0485 (**Figures 5B–D**). Two DOM molecules ($C_{15}H_{20}O_6$ and $C_{15}H_{22}O_6$) belonging to Cluster rank 4 were enriched in the 5–110 cm depth interval and correlated to microbial ASVs affiliated to Alphaproteobacteria, Desulfobacterales as well as Anaerolinaea and Zixibacteria (**Figure 5E**). Details of H/C and O/C ratios of different DOM molecules were showed in **Supplementary Figure 4**. Overall, archaeal groups including Bathyarchaeia and Lokiarchaeia were predominately associated to molecules belonging to lignins/CRAM-like structures having H/C ratios < 1.5, while MBG-D was associated to alphatic/proteins. Bacterial groups such as Gammaproteobacteria, Desulfobulbia and Desulfobacterales predominately represented correlations to molecules belonging to alphatic/proteins having H/C ratios > 1.5; however, Dehalococcoidia, Anaerolinaea, Planctomycetota, and Desulfatiglandales were exceptionally linked with lignins/CRAM-like structures and tannins.

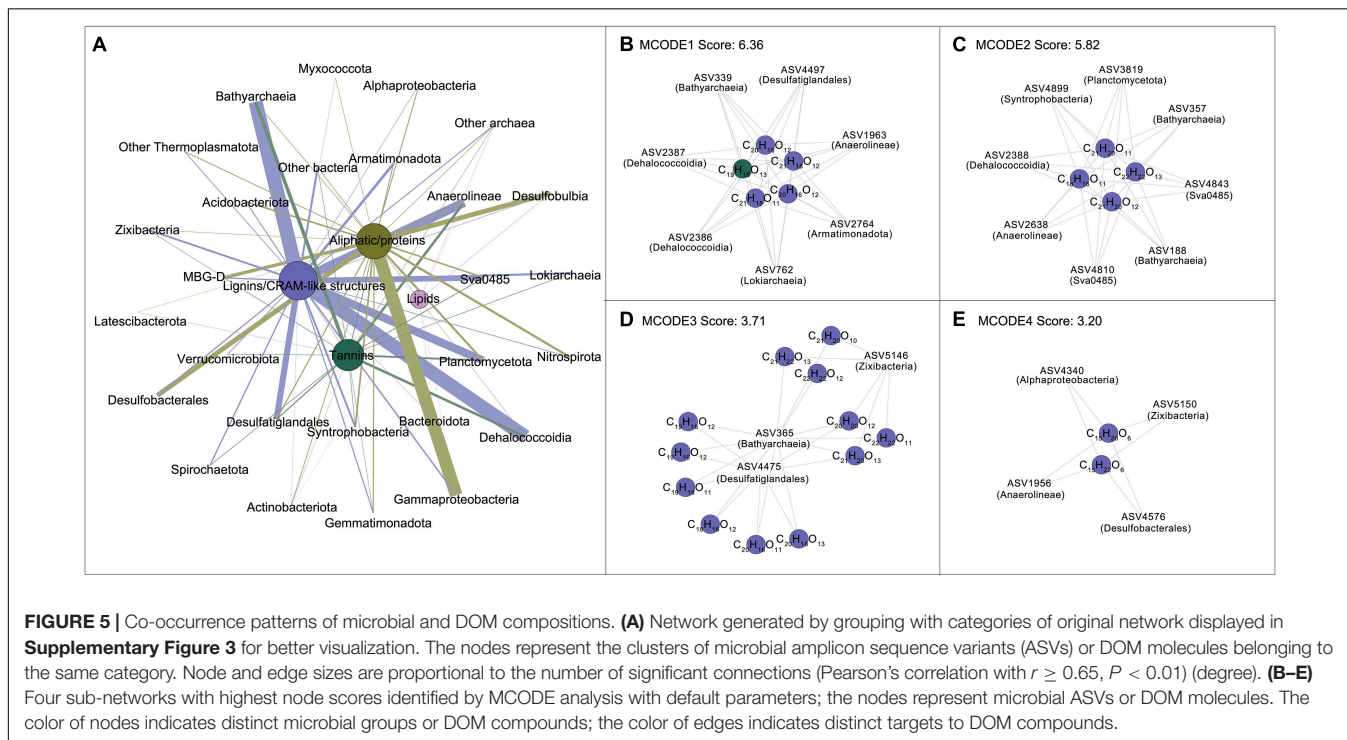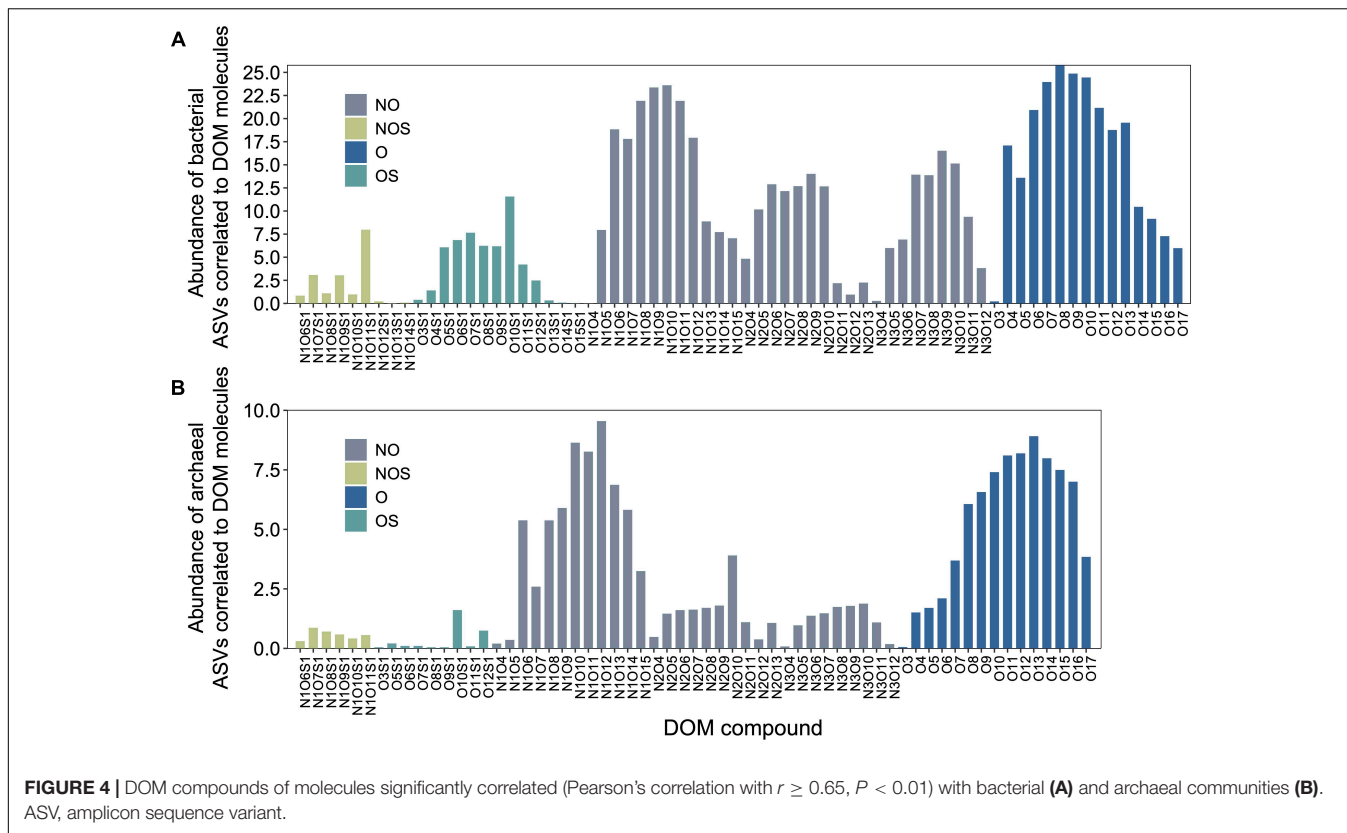## Potential Functions of Bacteria and Archaea in Organic Matter Degradation

A total of 51 archaeal and 166 bacterial genomes with completeness more than 70% and contamination lower than 5% were chosen for function predictions (**Supplementary Table 7**, **Supplementary Figures 5, 6**, and **Supplementary Text 2**). Metabolic potentials for metabolizing complex carbohydrates in archaeal and bacterial communities were investigated by searching CAZymes within each of the genomes. In total, we detected 1,219 (24 per genome) archaeal and 11,834 (72 per genome) bacterial potential CAZymes. The abundances of CAZymes in archaeal genomes generally increased along the depth profile and approached that of bacteria in the 200 and 300-cm depths. Significant difference ($P < 0.05$, Nonparametric Mann-Whitney test) existed between the bacterial and archaeal communities in upper layers (**Supplementary Figure 7A** and **Supplementary Table 7**). Glycosyltransferases (GT) and glycosyl-hydrolases (GH) catalyzing the synthesis and breakage of glycosidic bonds were the predominant CAZymes families in the archaeal and bacterial enzyme pools (**Supplementary Figure 7B**), which is consistent with observation in other environments (Jimenez et al., 2016; Dombrowski et al., 2018). Auxiliary activities (AA) were associated with redox enzymes essential for lignin degradation, including ligninolytic enzymes and lytic polysaccharide mono-oxygenases (Levasseur et al., 2013). Archaea encoded higher fraction (15.76%) of AA families and resulted in higher AA/GH ratios (0.86 ∼ 1.04) than bacteria (**Supplementary Figures 7B,C**).

In general, bacteria encoded a broader range of CAZymes than archaea, similar with the findings in the Guaymas Basin

TABLE 1 | Average and weighted average of indices of DOM molecules significantly correlated (Pearson's correlation with $r \geq 0.65$, $P < 0.01$) to bacterial and archaeal communities.

Taxonomy		DBE	MW	AI-mod	H/C	O/C
Bacteria	Average	8.89	415.99	0.18	1.26	0.51
	Average SD	4.58	97.19	0.30	0.37	0.14
	Weighted average	6.83	376.30	0.07	1.42	0.52
	Weighted average SD	0.10	1.85	0.01	0.01	0.00
Archaea	Average	9.02	416.24	0.19	1.25	0.52
	Average SD	4.80	95.88	0.30	0.38	0.14
	Weighted average	10.00	429.76	0.24	1.18	0.53
	Weighted average SD	0.12	2.16	0.01	0.01	0.00

Weighted average was calculated by weighting with number of correlated amplicon sequence variants (ASVs). DBE, double bond equivalent; MW, molecular weight; AI-mod, modified aromaticity index; H/C, H/C ratios; O/C, O/C ratios.

FIGURE 4 | DOM compounds of molecules significantly correlated (Pearson's correlation with $r \geq 0.65$, $P < 0.01$) with bacterial **(A)** and archaeal communities **(B)**. ASV, amplicon sequence variant.

FIGURE 5 | Co-occurrence patterns of microbial and DOM compositions. **(A)** Network generated by grouping with categories of original network displayed in **Supplementary Figure 3** for better visualization. The nodes represent the clusters of microbial amplicon sequence variants (ASVs) or DOM molecules belonging to the same category. Node and edge sizes are proportional to the number of significant connections (Pearson's correlation with $r \geq 0.65$, $P < 0.01$) (degree). **(B–E)** Four sub-networks with highest node scores identified by MCODE analysis with default parameters; the nodes represent microbial ASVs or DOM molecules. The color of nodes indicates distinct microbial groups or DOM compounds; the color of edges indicates distinct targets to DOM compounds.

(Dombrowski et al., 2018). Most CAZymes identified in archaeal genomes were AA4, AA6, AA7, CBM44, GH109, and GH57 families, while AA3, CBM48, GH3, GH23, and GH77 were more common in bacteria (**Figure 6A** and **Supplementary Table 8**). These enzymes involving in carbon degradation showed obvious depth distribution accompanied with the changing of

FIGURE 6 | Metabolic potential involved in organic degradation of genomes. **(A)** Percentage of carbohydrate-active enzymes (CAZymes) and aromatic degradation genes encoded in genomes summarized for each cluster. Each group of enzymes were set as the same color codes strategies. Brackets: Total number of genomes in each phylogenetic cluster. AA, auxiliary activities; CBM, cellulose-binding motifs; GH, glycosyl-hydrolases; Aerobic AD, aerobic aromatic degradation; Anaerobic AD, anaerobic aromatic degradation. **(B)** Reconstructed anaerobic aromatic degradation pathway and reactions participating in lignin degradation. Percentage of genes were summarized for archaea and bacteria. Brackets: Total number of genomes in each domain.

microbial community composition (**Supplementary Figure 8** and **Supplementary Table 9**). CAZymes involving cleavages of peptidoglycan (GH13, GH23, and GH102), xyloglucan (GH16), cellobiose (AA3), glucooligosaccharide (AA7), glycogen- and chitopentaose-binding motifs (CBM48 and CBM50)

were predominant in the upper layers, mainly derived from Desulfobulbia, Gammaproteobacteria, and Desulfatiglanales. CAZymes enriched in deep layers were encoded by Planctomycetota (most of GH families), Bathyarchaeia (AA6, CBM44, GH57, and GH133), and Anaerolineae (CAZymes

enriched in 200-cm depth). AA6 (1,4-benzoquinone reductase) and AA4 (vanillyl-alcohol oxidase) involved in degradation of aromatic compounds in subsurface were mainly present in Bathyarchaeia and Desulfatiglanales (**Supplementary Figure 9**). Accordingly, cellulose-binding motifs CBM44 and CBM9 were also enriched in deep layers.

What needs to be highlighted is the diversity of aromatic metabolisms involved by both archaeal and bacterial populations from sediments. Genes involving in aerobic aromatic degradation pathways were strictly found in bacteria; for example, *pcaC*, *pcaD,* and *pcaH* genes were exclusively present in bacteria for the degradation of terephthalate, *hcaC* gene for the degradation of phenylpropanoate and *gnl* gene for the degradation of caprolactone. Genes that are commonly present in both archaeal and bacterial genomes include the *paaK* and *padI* genes related to the phenylacetic acid pathway and the genes encoding enzymes for hydrolytic cleavage of the ring *(dch, had, and oah)*. The *bcrABCD* genes targeting anaerobic aromatics degradation, however, were identified in most archaeal lineages, especially Bathyarchaeia and other Thermoplasmatota (MBG-D was excluded), suggesting the increasing importance of archaea in the utilization of ATP-dependent benzoyl-CoA reductase (BCR, class I) (**Figure 6B** and **Supplementary Table 8**). Also, the gene encoding enzyme for carboxylation (*UbiX*) was frequently detected in 50.9% of archaeal genomes and only 32.5% of the bacterial genomes (**Figure 6B**). Lastly, though bacteria harbored more broad aromatic degradation genes (i.e., near complete pathway for phenol metabolism in bacterial linages like Desulfatiglanales, Zixibacteria, and Gemmatimonadota), archaea especially Bathyarchaeia contained greater abundance of genes (*UbiX* and *bcrABCD*) for key processes (carboxylation and dearomatization) (**Supplementary Figure 9**).

DISCUSSION

Dissolved organic matter in sediment porewater is primarily derived from enzyme-mediated hydrolysis and depolymerization of POM and/or lysis of microbial cell components (Burdige and Komada, 2015; Liang et al., 2017). Porewater DOC concentrations generally increase with depth, which in anoxic sediments show particular exponential-like profiles (Komada et al., 2013; Chen et al., 2017; Schmidt et al., 2017). This study first characterized the exponential-like profiles of the properties of DOM molecules (MW_{wa}, C_{wa}, $NOSC_{wa}$, H/C_{wa}, $AI\text{-}mod_{wa}$, DBE_{wa}, and I_{DEG}) in the Pearl River estuarine sediments. The stable values of these parameters below 100 cm could be explained by the balance of DOM production and consumption. It could also indicate a combination of net DOM production and decreasing bulk reactivity of DOM (Burdige and Komada, 2015), as evidenced by the accumulation of non-reactive DOM in this study (**Figures 1E–H**). Increasing MW as well as AI-mod, DBE, and I_{DEG} of SPE-DOM molecules down core indicated accumulation of refractory DOM as reported elsewhere (Li et al., 2019).

While carbon degradation continues with sediment depth, the spectrum of substrates becomes narrow (Jørgensen, 2000;

Orsi et al., 2020) and becomes increasingly important in shaping the microbial community structure. Our study demonstrated that bacterial and archaeal communities were separated based on their correlations to different DOM compositions. Gammaproteobacteria, Desulfobacterales, and Desulfobulbia occupied their ecological niches in upper sediments and may exhibit preference for alphatic/protein groups (see also references Teeling et al., 2012; Nikrad et al., 2014; Zhao Z. et al., 2019). Dehalococcoidia and Bathyarchaeia have been thought to be important for organic matter degradation in deep sediments (Oni et al., 2015; Lazar et al., 2017; Vuillemin et al., 2020); Bathyarchaeia, in particular, were capable of lignin utilization in enrichment cultivation (Yu et al., 2018) and significantly correlated with terrestrially derived organic matter in Pearl River estuary (Wang P. et al., 2020). This is consistent with our observation that DOM molecules belonging to lignins/CRAM-like structures were correlated to Dehalococcoidia, Bathyarchaeia and other deep-sediment enrich groups like Planctomycetota and Desulfatiglanales (**Figure 5** and **Supplementary Figure 4**). These microorganisms in deep subsurface may continuously utilize refractory DOM. Recent studies showed that CRAM organic matter such as benzoic acid and lignin were able to rapidly or slowly stimulate microbial growth (Liu S. et al., 2020). This was also supported by increased archaeal biomass especially Bathyarchaeia (Wang W. et al., 2020) and their correlations to O-rich molecules (**Figure 4B**).

The increasing recalcitrance of DOM represented by lignins/CRAM-like structures, tannins and aromatic structures in sediments apparently affect the strategies for substrate utilization by bacteria and archaea. Aromatic metabolism is commonly observed to be performed by sedimentary bacteria. For example, the sulfate/nitrate-reducing bacteria such as *Desulfobacula* and *Desulfobacterium* via benzoyl-CoA degradation pathway (Wöhlbrand et al., 2012; Boll et al., 2020). In recent years, aromatic degradation pathways were extended to archaeal members like MBG-D that have the phenylacetic acid pathway (Liu W.W. et al., 2020), Bathyarchaeia the protocatechuate degradation pathway (Meng et al., 2014; Zhou et al., 2018) and a few Asgard members the benzoyl-CoA degradation pathway (Farag et al., 2020). In this study, though bacteria encode a near complete pathway of aromatic degradation, archaea were possible to mediate the key processes like dearomatization and carboxylation (**Figures 6A,B**). BCR encoded for benzoyl-CoA pathway was frequently identified in archaeal genomes especially in Bathyarchaeia and other Thermoplasmatota (**Figure 6A**). The central aromatic intermediate benzoyl-CoA in benzoyl-CoA degradation pathway could derive from various monocyclic aromatic compounds including benzene, toluene, phthalates, phenylacetic acid, O-xylene, and more (Boll et al., 2020) and BCR were targeted to initiate the key reaction that dearomatizes the benzoyl-CoA to 1,5-dienoyl-CoA, which further involves a series of beta-oxidation like reactions to produces acetyl-CoA (Boll et al., 2020; Farag et al., 2020). Our study points out the prevalence of BCR in archaea, especially in Bathyarchaeia, expanding the spectrum of archaeal lineages capable of utilizing aromatic compounds as carbon or energy source. Another potential key reaction in the anaerobic degradation pathways

found here could be C-H bond activation of non-substituted benzene, naphthalene or phenol by *UbiD/UbiX*-like genes-induced carboxylation, which were identified in benzene degrading *Ferroglobus placidus* (Holmes et al., 2011) and Fe (III)-reducing or methanogenic enrichment (Kunapuli et al., 2008; Luo et al., 2014). Archaea containing genes involving carboxylation (*UbiX*) and dearomatization (*bcrABCD*) were both of high frequency and abundance in deep layers (**Figure 6B** and **Supplementary Figure 9**), demonstrating potential importance of archaea in anaerobic aromatic degradation.

Benthic archaea may be particularly important in utilization of recalcitrant organic matter in low energy environments (see below). So far, archaea are known to degrade detrital proteins (Lloyd et al., 2013; Lazar et al., 2017), polymeric carbohydrates (Lazar et al., 2017), alkane or methylated compounds (Wang et al., 2019), aromatics (Meng et al., 2014; Farag et al., 2020; Liu W.W. et al., 2020, this study) and lignin (Yu et al., 2018). Besides, AA6 families involving in intracellular aromatic biodegradation (Levasseur et al., 2013) are identified in lignocellulolytic or cellulolytic bacterium dominated consortia (Jimenez et al., 2016; Kanokratana et al., 2018) and were in high abundance in Bathyarchaeia genomes (**Supplementary Figure 9**). Linolytic enzymes in the AA class generally cooperate with classical polysaccharide depolymerases targeting the plant biomass (Levasseur et al., 2013). High potential of lignin degradation may exist in archaeal communities (Yu et al., 2018), which is also supported by high AA/GH ratios and tight associations to lignins/CRAM structures in this study. It has been noted that "we are woefully unaware of DOM production (or assimilation) mechanisms in the Archaea." (Kujawinski, 2011). A recent study has suggested that the metabolites of marine ammonia-oxidizing archaea like CRAM-like compounds may contribute to the composition of ocean DOM (Bayer et al., 2019). In this study, we highlighted the role of archaea in interacting with sedimentary non-reactive DOM considering their considerable biomass in subseafloor sediments globally (Lipp et al., 2008; Hoshino and Inagaki, 2019). Future work is needed to develop models based on cultivation combined with isotope and metatranscriptomic methods to quantify the contribution of archaea to carbon cycle in the sediments.

In summary, the molecular composition of porewater DOM showed stratification along depth in the Pearl River estuarine sediments. The recalcitrance of DOM increased below 100 cm, which was characterized by higher MW, DBE and lower H/C. Most bacteria were found to be correlated to DOM molecules with DBE < 10 and MW < 450 Da while archaea were closely correlated to molecules with the DBE between 11 and 18 and the MW of 500 Da, indicating greater correlations between archaea and the recalcitrant DOM. In particular, the archaeal group Bathyarhcaeia along with the bacterial groups Dehalococcoidia, Anaerolineae, Planctomycetota, and Desulfatiglandales were tightly correlated to the lignins/CRAM-like structures. In contrast, bacterial groups such as Gammaproteobacteria, Desulfobulbia, and Desulfobacterales were closely associated to alphatic/proteins that are more labile for degradation. These results allowed for distinction of carbon metabolizing genes from microbial genomes, with bacteria encoding a wider range of GH families of CAZymes involving glycolytic pathways while archaea a higher proportion of AA families and key genes in anaerobic aromatic degradation of predominantly recalcitrant DOM. These findings expand our knowledge about microbe-DOM relationship in sediments, particularly the role of archaea in sedimentary carbon transformation and/or perseveration.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/genbank/, PRJNA575161.

AUTHOR CONTRIBUTIONS

WW carried out the experimental work, data analysis, and wrote the manuscript. JT carried out sample collecting and assisted in the laboratory work. KY, CH, JW, PL, HC, and BX helped in data analysis and the writing of the manuscript. QS and CZ supervised and contributed to the design of the experiments and revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.697860/full#supplementary-material

REFERENCES

Alneberg, J., Bjarnason, B. S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U. Z., et al. (2014). Binning metagenomic contigs by coverage and composition. *Nat. Methods* 11, 1144–1146. doi: 10.1038/nmeth.3103

Bader, G. D., and Hogue, C. W. (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* 4:2. doi: 10.1186/1471-2105-4-2

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.10..9..cmb.2012.0021

Bastian, M., Heymann, S., and Jacomy, M. (2009). Gephi: an open source software for exploring and manipulating networks. *ICWSM* 8, 361–362.

Bayer, B., Hansman, R. L., Bittner, M. J., Noriega-Ortega, B. E., Niggemann, J., Dittmar, T., et al. (2019). Ammonia-oxidizing archaea release a suite of organic compounds potentially fueling prokaryotic heterotrophy in the ocean. *Environ. Microbiol.* 21, 4062–4075. doi: 10.1111/1462-2920.14755

Boll, M., Estelmann, S., and Heider, J. (2020). "Catabolic Pathways and Enzymes Involved in the Anaerobic Degradation of Monocyclic Aromatic Compounds," in *Anaerobic Utilization of Hydrocarbons, Oils, and Lipids*, ed. M. Boll (Cham: Springer), 85–133.

Bray, J. R., and Curtis, J. T. (1957). An ordination of the upland forest communities of southern Wisconsin. *Ecol. Monogr.* 27, 326–349. doi: 10.2307/1942268

Burdige, D. J. (2007). Preservation of organic matter in marine sediments: controls, mechanisms, and an imbalance in sediment organic carbon budgets? *Chem. Rev.* 107, 467–485. doi: 10.1021/cr050347q

Burdige, D. J., and Komada, T. (2015). "Sediment pore waters," in *Biogeochemistry of Marine Dissolved Organic Matter*, eds D. A. Hansen and C. A. Carlson (Cambridge, MA: Academic Press), 535–577. doi: 10.1016/B978-0-12-405940-5.00012-1

Burdige, D. J., Komada, T., Magen, C., and Chanton, J. P. (2016). Modeling studies of dissolved organic matter cycling in Santa Barbara Basin (CA, USA) sediments. *Geochim. Cosmochim. Acta* 195, 100–119. doi: 10.1016/j.gca.2016.09.007

Canuel, E. A., and Hardison, A. K. (2016). Sources, ages, and alteration of organic matter in estuaries. *Ann. Rev. Mar. Sci.* 8, 409–434. doi: 10.1146/annurev-marine-122414-034058

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., et al. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* 6, 1621–1624. doi: 10.1038/ismej.2012.8

Chen, M., and Hur, J. (2015). Pre-treatments, characteristics, and biogeochemical dynamics of dissolved organic matter in sediments: a review. *Water Res.* 79, 10–25. doi: 10.1016/j.watres.2015.04.018

Chen, M., Kim, J. H., Choi, J., Lee, Y. K., and Hur, J. (2017). Biological early diagenesis and insolation-paced paleoproductivity signified in deep core sediment organic matter. *Sci. Rep.* 7:1581. doi: 10.1038/s41598-017-01759-4

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

Dombrowski, N., Teske, A. P., and Baker, B. J. (2018). Expansive microbial metabolic versatility and biodiversity in dynamic Guaymas Basin hydrothermal sediments. *Nat. Commun.* 9:4999. doi: 10.1038/s41467-018-07418-0

Farag, I. F., Biddle, J. F., Zhao, R., Martino, A. J., House, C. H., and Leon-Zayas, R. I. (2020). Metabolic potentials of archaeal lineages resolved from metagenomes of deep Costa Rica sediments. *ISME J.* 14, 1345–1358. doi: 10.1038/s41396-020-0615-5

Flerus, R., Lechtenfeld, O. J., Koch, B. P., McCallister, S. L., Schmitt-Kopplin, P., Benner, R., et al. (2012). A molecular perspective on the ageing of marine dissolved organic matter. *Biogeosciences* 9, 1935–1955. doi: 10.5194/bg-9-1935-2012

Fox, C. A., Abdulla, H. A., Burdige, D. J., Lewicki, J. P., and Komada, T. (2018). Composition of dissolved organic matter in pore waters of anoxic marine sediments analyzed by 1H nuclear magnetic resonance spectroscopy. *Front. Mar. Sci.* 5:172. doi: 10.3389/fmars.2018.00172

Gan, S., Schmidt, F., Heuer, V. B., Goldhammer, T., Witt, M., and Hinrichs, K.-U. (2020). Impacts of redox conditions on dissolved organic matter (DOM) quality in marine sediments off the River Rhône, Western Mediterranean Sea. *Geochim. Cosmochim. Acta* 276, 151–169. doi: 10.1016/j.gca.2020.02.001

He, C., Zhang, Y., Li, Y., Zhuo, X., Li, Y., Zhang, C., et al. (2020). In-house standard method for molecular characterization of dissolved organic matter by FT-ICR mass spectrometry. *ACS Omega* 5, 11730–11736. doi: 10.1021/acsomega.0c01055

Hedges, J. I., and Keil, R. G. (1995). Sedimentary organic matter preservation: an assessment and speculative synthesis. *Mar. Chem.* 49, 81–115. doi: 10.1016/0304-4203(95)00008-F

Holmes, D. E., Rsso, C., Smith, J. A., and Lovley, D. R. (2011). Anaerobic oxidation of benzene by the hyperthermophilic archaeon *Ferroglobus placidus*. *Appl. Environ. Microbiol.* 77(17), 5926–5933. doi: 10.1128/AEM.05452-11

Hoshino, T., and Inagaki, F. (2019). Abundance and distribution of archaea in the subseafloor sedimentary biosphere. *ISME J.* 13, 227–231. doi: 10.1038/s41396-018-0253-3

Huang, L., Zhang, H., Wu, P., Entwistle, S., Li, X., Yohe, T., et al. (2018). dbCAN-seq: a database of carbohydrate-active enzyme (CAZyme) sequence and annotation. *Nucleic Acids Res.* 46, D516–D521. doi: 10.1093/nar/gkx894

Hyatt, D., LoCascio, P. F., Hauser, L. J., and Uberbacher, E. C. (2012). Gene and translation initiation site prediction in metagenomic sequences. *Bioinformatics* 28, 2223–2230. doi: 10.1093/bioinformatics/bts429

Jimenez, D. J., de Lima Brossi, M. J., Schuckel, J., Kracun, S. K., Willats, W. G., and van Elsas, J. D. (2016). Characterization of three plant biomass-degrading microbial consortia by metagenomics- and metasecretomics-based approaches. *Appl. Microbiol. Biotechnol.* 100, 10463–10477. doi: 10.1007/s00253-016-7713-3

Jørgensen, B. B. (2000). "Bacteria and marine biogeochemistry," in *Marine Geochemistry*, eds H. D. Schulz and M. Zabel (Berlin: Springer), 173–207.

Kang, D. D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., et al. (2019). MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 7:e7359. doi: 10.7717/peerj.7359

Kanokratana, P., Wongwilaiwalin, S., Mhuantong, W., Tangphatsornruang, S., Eurwilaichitr, L., and Champreda, V. (2018). Characterization of cellulolytic microbial consortium enriched on Napier grass using metagenomic approaches. *J. Biosci. Bioeng.* 125, 439–447. doi: 10.1016/j.jbiosc.2017.10.014

Koch, B. P., and Dittmar, T. (2006). From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter. *Rapid Commun. Mass Spectrom.* 20, 926–932. doi: 10.1002/rcm.2386

Koch, B. P., and Dittmar, T. (2016). From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter. *Rapid Commun. Mass Spectrom.* 30, 250–250. doi: 10.1002/rcm.7433

Komada, T., Burdige, D. J., Crispo, S. M., Druffel, E. R. M., Griffin, S., Johnson, L., et al. (2013). Dissolved organic carbon dynamics in anaerobic sediments of the Santa Monica Basin. *Geochim. Cosmochim. Acta* 110, 253–273. doi: 10.1016/j.gca.2013.02.017

Krueger, F. (2012). *Trim Galore: A Wrapper Tool Around Cutadapt and FastQC to Consistently Apply Quality and Adapter Trimming to FastQ files, with Some Extra Functionality for MspI-Digested RRBS-type (Reduced Representation Bisufite-Seq) Libraries.* Available online at: http://www. bioinformatics. babraham. ac. uk/projects/trim_galore/.(accessed April 28, 2016).

Kujawinski, E. B. (2011). The impact of microbial metabolism on marine dissolved organic matter. *Ann. Rev. Mar. Sci.* 3, 567–599. doi: 10.1146/annurev-marine-120308-081003

Kunapuli, U., Griebler, C., Beller, H. R., and Meckenstock, R. U. (2008). Identification of intermediates formed during anaerobic benzene degradation by an iron-reducing enrichment culture. *Environ. Microbiol.* 10, 1703–1712. doi: 10.1111/j.1462-2920.2008.01588.x

Lazar, C. S., Baker, B. J., Seitz, K. W., and Teske, A. P. (2017). Genomic reconstruction of multiple lineages of uncultured benthic archaea suggests distinct biogeochemical roles and ecological niches. *ISME J.* 11, 1118–1129. doi: 10.1038/ismej.2016.189

Levasseur, A., Drula, E., Lombard, V., Coutinho, P. M., and Henrissat, B. (2013). Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol. Biofuels* 6:41. doi: 10.1186/1754-6834-6-41

Li, D., Luo, R., Liu, C. M., Leung, C. M., Ting, H. F., Sadakane, K., et al. (2016). MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* 102, 3–11. doi: 10.1016/j.ymeth.2016.02.020

Li, P., Tao, J., Lin, J., He, C., Shi, Q., Li, X., et al. (2019). Stratification of dissolved organic matter in the upper 2000m water column at the Mariana Trench. *Sci. Total Environ.* 668, 1222–1231. doi: 10.1016/j.scitotenv.2019.03.094

Liang, C., Schimel, J. P., and Jastrow, J. D. (2017). The importance of anabolism in microbial control over soil carbon storage. *Nat. Microbiol.* 2:17105. doi: 10.1038/nmicrobiol.2017.105

Lipp, J. S., Morono, Y., Inagaki, F., and Hinrichs, K. U. (2008). Significant contribution of archaea to extant biomass in marine subsurface sediments. *Nature* 454, 991–994. doi: 10.1038/nature07174

Liu, S., Parsons, R., Opalk, K., Baetge, N., Giovannoni, S., Bolaños, L. M., et al. (2020). Different carboxyl−rich alicyclic molecules proxy compounds select distinct bacterioplankton for oxidation of dissolved organic matter in the mesopelagic Sargasso Sea. *Limnol. Oceanogr.* 65, 1532–1553. doi: 10.1002/lno.11405

Liu, W. W., Pan, J., Feng, X., Li, M., Xu, Y., Wang, F., et al. (2020). Evidences of aromatic degradation dominantly via the phenylacetic acid pathway in marine benthic Thermoprofundales. *Environ. Microbiol.* 22, 329–342. doi: 10.1111/1462-2920.14850

Lloyd, K. G., Schreiber, L., Petersen, D. G., Kjeldsen, K. U., Lever, M. A., Steen, A. D., et al. (2013). Predominant archaea in marine sediments degrade detrital proteins. *Nature* 496, 215–218. doi: 10.1038/nature12033

Luo, F., Gitiafroz, R., Devine, C, E., Gong, Y., Hug, L. A., Raskin, L., et al. (2014). Metatranscriptome of an anaerobic benzene-degrading, nitrate-reducing enrichment culture reveals involvement of carboxylation in benzene ring activation. *Appl. Environ. Microbiol.* 80, 4095–4107. doi: 10.1128/AEM.00717-14

Mahmoudi, N., Beaupre, S. R., Steen, A. D., and Pearson, A. (2017). Sequential bioavailability of sedimentary organic matter to heterotrophic bacteria. *Environ. Microbiol.* 19, 2629–2644. doi: 10.1111/1462-2920.13745

Mei, W., Yu, G., Lai, J., Rao, Q., and Umezawa, Y. (2018). *Package 'basicTrendline'*.

Meng, J., Xu, J., Qin, D., He, Y., Xiao, X., and Wang, F. (2014). Genetic and functional properties of uncultivated MCG archaea assessed by metagenome and gene expression analyses. *ISME J.* 8, 650–659. doi: 10.1038/ismej.2013.174

Mikheenko, A., Saveliev, V., and Gurevich, A. (2016). MetaQUAST: evaluation of metagenome assemblies. *Bioinformatics* 32, 1088–1090. doi: 10.1093/bioinformatics/btv697

Nikrad, M. P., Cottrell, M. T., Kirchman, D. L., and Kostka, J. E. (2014). Uptake of dissolved organic carbon by gammaproteobacterial subgroups in Coastal Waters of the West Antarctic Peninsula. *Appl. Environ. Microbiol.* 80, 3362–3368. doi: 10.1128/aem.00121-14

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'hara, R., et al. (2013). Package 'vegan'. *Commun. Ecol. Package Version* 2, 1–295.

Oni, O. E., Schmidt, F., Miyatake, T., Kasten, S., Witt, M., Hinrichs, K. U., et al. (2015). Microbial communities and organic matter composition in surface and subsurface sediments of the Helgoland Mud Area, North Sea. *Front. Microbiol.* 6:1290. doi: 10.3389/fmicb.2015.01290

Orsi, W. D., Schink, B., Buckel, W., and Martin, W. F. (2020). Physiological limits to life in anoxic subseafloor sediment. *FEMS Microbiol. Rev.* 44, 219–231. doi: 10.1093/femsre/fuaa004

Osterholz, H., Singer, G., Wemheuer, B., Daniel, R., Simon, M., Niggemann, J., et al. (2016). Deciphering associations between dissolved organic molecules and bacterial communities in a pelagic marine system. *ISME J.* 10, 1717–1730. doi: 10.1038/ismej.2015.231

Parkes, R. J., Cragg, B., Roussel, E., Webster, G., Weightman, A., and Sass, H. (2014). A review of prokaryotic populations and processes in sub-seafloor sediments, including biosphere:geosphere interactions. *Mar. Geol.* 352, 409–425. doi: 10.1016/j.margeo.2014.02.009

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. doi: 10.1101/gr.186072.114

Parvin, S., Van Geel, M., Yeasmin, T., Lievens, B., and Honnay, O. (2019). Variation in arbuscular mycorrhizal fungal communities associated with lowland rice (*Oryza sativa*) along a gradient of soil salinity and arsenic contamination in Bangladesh. *Sci. Total Environ.* 686, 546–554. doi: 10.1016/j.scitotenv.2019.05.450

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* 14, 417–419. doi: 10.1038/nmeth.4197

Petro, C., Starnawski, P., Schramm, A., and Kjeldsen, K. U. (2017). Microbial community assembly in marine sediments. *Aquat. Microb. Ecol.* 79, 177–195. doi: 10.3354/ame01826

Riedel, T., Biester, H., and Dittmar, T. (2012). Molecular fractionation of dissolved organic matter with metal salts. *Environ. Sci. Technol.* 46, 4419–4426. doi: 10.1021/es203901u

Saw, J. H. W., Nunoura, T., Hirai, M., Takaki, Y., Parsons, R., Michelsen, M., et al. (2020). Pangenomics analysis reveals diversification of enzyme families and niche specialization in globally abundant SAR202 Bacteria. *mBio* 11:e02975-19. doi: 10.1128/mBio.02975-19

Schmidt, F., Koch, B. P., Goldhammer, T., Elvert, M., Witt, M., Lin, Y.-S., et al. (2017). Unraveling signatures of biogeochemical processes and the depositional setting in the molecular composition of pore water DOM across different marine environments. *Geochim. Cosmochim. Acta* 207, 57–80. doi: 10.1016/j.gca.2017.03.005

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., et al. (2011). Metagenomic biomarker discovery and explanation. *Genome Biol.* 12, 1–18. doi: 10.1186/gb-2011-12-6-r60

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303

Sleighter, R. L., and Hatcher, P. G. (2007). 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. doi: 10.1002/jms.1221

Teeling, H., Fuchs, B. M., Becher, D., Klockow, C., Gardebrecht, A., Bennke, C. M., et al. (2012). Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* 336, 608–611. doi: 10.1126/science.1218344

Vuillemin, A., Kerrigan, Z., D'Hondt, S., and Orsi, W. D. (2020). Chloroflexi persisting for millions of years in oxic and anoxic deep-sea clay. *bioRxiv* [Preprint] doi: 10.1101/2020.05.26.116590

Wang, P., Zhang, T., Chen, S., Li, X., Lai, D., Gao, S., et al. (2020). Niche specificity and potential terrestrial organic carbon utilization of benthic bathyarchaeota in a eutrophic subtropic estuarine system. *Chem. Geol.* 556:119839. doi: 10.1016/j.chemgeo.2020.119839

Wang, W., Tao, J., Liu, H., Li, P., Chen, S., Wang, P., et al. (2020). Contrasting bacterial and archaeal distributions reflecting different geochemical processes in a sediment core from the Pearl River Estuary. *AMB Express* 10:16. doi: 10.1186/s13568-020-0950-y

Wang, Y., Wegener, G., Hou, J., Wang, F., and Xiao, X. (2019). Expanding anaerobic alkane metabolism in the domain of archaea. *Nat. Microbiol.* 4, 595–602. doi: 10.1038/s41564-019-0364-2

Wasmund, K., Schreiber, L., Lloyd, K. G., Petersen, D. G., Schramm, A., Stepanauskas, R., et al. (2014). Genome sequencing of a single cell of the widely distributed marine subsurface *Dehalococcoidia*, phylum Chloroflexi. *ISME J.* 8, 383–397. doi: 10.1038/ismej.2013.143

Wöhlbrand, L., Jacob, J. H., Kube, M., Mussmann, M., Jarling, R., Beck, A., et al. (2012). Complete genome, catabolic sub-proteomes and key-metabolites of *Desulfobacula toluolica* Tol2, a marine, aromatic compound-degrading, sulfate-reducing bacterium. *Environ. Microbiol.* 15, 1334–1355. doi: 10.1111/j.1462-2920.2012.02885.x

Wu, Y.-W., Simmons, B. A., and Singer, S. W. (2016). MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* 32, 605–607. doi: 10.1093/bioinformatics/btv638

Yu, K., Qiu, Z., Mu, R., Qiao, X., Zhang, L., Lian, C.-A., et al., (2021). Recovery of high-qualitied Genomes from a deep-inland Salt Lake Using BASALT. *bioRxiv* doi: 10.1101/2021.03.05.434042

Yu, T., Wu, W., Liang, W., Lever, M. A., Hinrichs, K.-U., and Wang, F. (2018). Growth of sedimentaryBathyarchaeotaon lignin as an energy source. *Proc. Natl. Acad. Sci. U.S.A.* 115, 6022–6027. doi: 10.1073/pnas.1718854115

Yuan, Z., He, C., Shi, Q., Xu, C., Li, Z., Wang, C., et al. (2017). Molecular insights into the transformation of dissolved organic matter in landfill leachate concentrate during biodegradation and coagulation processes using ESI FT-ICR MS. *Environ. Sci. Technol.* 51, 8110–8118. doi: 10.1021/acs.est.7b02194

Zhao, W., Hu, A., Ni, Z., Wang, Q., Zhang, E., Yang, X., et al. (2019). Biodiversity patterns across taxonomic groups along a lake water-depth gradient: effects of abiotic and biotic drivers. *Sci. Total Environ.* 686, 1262–1271. doi: 10.1016/j.scitotenv.2019.05.381

Zhao, Z., Gonsior, M., Schmitt-Kopplin, P., Zhan, Y., Zhang, R., Jiao, N., et al. (2019). Microbial transformation of virus-induced dissolved organic matter from picocyanobacteria: coupling of bacterial diversity and DOM chemodiversity. *ISME J.* 13, 2551–2565. doi: 10.1038/s41396-019-0449-1

Zhou, Z., Pan, J., Wang, F., Gu, J. D., and Li, M. (2018). Bathyarchaeota: globally distributed metabolic generalists in anoxic environments. *FEMS Microbiol. Rev.* 42, 639–655. doi: 10.1093/femsre/fuy023

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