

RESEARCH ARTICLE

Molecular activity mediates the composition and assembly of dissolved organic matter in lake sediments

Shuailong Wen ^{1,2}, Ang Hu,^{1,2} Francisco Dini-Andreote,^{3,4} Lei Han,^{1,2} Shuyu Jiang,^{1,2} Kyoung-Soon Jang,^{5*} Jianjun Wang ^{1,2*}

¹Key Laboratory of Lake and Watershed Science for Water Security, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, China; ²State Key Laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, China; ³Department of Plant Science & Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, Pennsylvania, USA; ⁴The One Health Microbiome Center, Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, Pennsylvania, USA; ⁵Bio-Chemical Analysis Team, Korea Basic Science Institute, Cheongju, Republic of Korea

Abstract

Lake sediments are hotspots for carbon transformation and burial, where the turnover of dissolved organic matter (DOM), influenced by molecular activity, regulates global carbon cycling. However, the composition of sediment DOM and how its assembly mechanisms are related to molecular activity remain poorly understood. Here, 63 freshwater sediments were collected from tropical to cold temperate climatic zones in China. We explored the molecular composition and assembly of sediment DOM and the underlying mechanisms driven by climate, physicochemical factors, and microbes along the gradient of molecular activity, indicated by the number of potential biochemical transformations in which a molecule is involved. Sediment DOM was dominated by lipid- (35%) and lignin-like compounds (33%), and the latter were enriched as the molecular activity of DOM increased. Besides, the DOM assemblage with higher potential biochemical transformations had greater compositional similarity across different climatic zones and tended to assemble deterministically. Specifically, as the average number of potential biochemical transformations of molecules increased from 0.4 to 14, the assembly of these molecules was structured by a shift from stochastic to deterministic processes, with the latter accounting for $\geq 75\%$ thereafter. Overall, DOM assemblages were primarily structured by physicochemical factors, including sediment total organic carbon and electrical conductivity. As molecular activity increased, however, assemblages were increasingly affected by climate and bacterial communities, consistent with the enhanced complexity of bacterial–molecular networks. Collectively, our study highlights that the turnover of DOM regulated by biotic and abiotic factors is further constrained by the intrinsic molecular activity.

As the most dynamic component for carbon cycling in waters and sediments, dissolved organic matter (DOM) exerts profound effects on the functionality of lakes and downstream aquatic ecosystems (Cui et al. 2024; Tanentzap and Fonvielle 2024). These effects include light attenuation, the supply of nitrogen and phosphorus to primary production, potential alteration of contaminant toxicity, and modulation of

biogeochemical cycles (e.g., nitrogen, carbon, sulfur) via shifts in microbial metabolisms (Brailsford et al. 2021; Chen et al. 2022; Gonsior et al. 2019). Therefore, DOM composition—which refers to the chemodiversity and structural properties of the organic compounds—has been broadly explored in diverse water bodies (e.g., rivers, lakes, and wetlands) (Lapierre et al. 2013; Tanentzap and Fonvielle 2024). However, knowledge of the composition of DOM in sediments, which represent one of the important active carbon reservoirs in freshwater ecosystems (Catalán et al. 2016), remains largely elusive. As such, understanding the underlying mechanisms structuring the composition of DOM in freshwater sediments currently represents a fundamental knowledge gap.

*Correspondence: ksjang@kbsi.re.kr; jjwang@niglas.ac.cn

Associate editor: Oscar Serrano

Data Availability Statement: Data available on request from the authors.

Dissolved organic matter composition in freshwaters is determined by a complex interplay among several forces, including the production, degradation, transport, and transformation of organic molecules (Danczak et al. 2020). Most interestingly, these interplays can occur through a balance of deterministic and stochastic mechanisms (Kajan et al. 2023). Analogous to the concept of ecological community assembly (Ning et al. 2020), deterministic processes are associated with the selective production, transformation, or loss of specific molecules through the interactions among biotic and abiotic factors, also known as environmental filtering (Hu et al. 2022b). For instance, microorganisms preferentially degrade lipid- and protein-like molecules, causing an enrichment of polyphenolic and highly unsaturated recalcitrant compounds in lakes (Catalán et al. 2024). Besides, DOM composition could be determined by stochastic processes, which involve random events (in ecology, dispersal and ecological drift)—here, the hydrological transport and vectorial movement of molecules and specific molecules' lifetimes (Hu et al. 2022b; Hubbell 2001).

Organic molecules undergo numerous biochemical transformations, with consequences for the composition and assembly of DOM that are highly related to molecular reactivity and activity (Kajan et al. 2023; Ryan et al. 2024). Molecular reactivity describes the ability of a molecule to be decomposed; for example, the recalcitrant compounds like lignin are considered to have low reactivity while labile amino acids are considered to have high reactivity, which has been extensively explored for revealing DOM dynamics in freshwater habitats (Berggren et al. 2022; Moran et al. 2016). Molecular activity—indicated by the number of potential biochemical transformations a molecule is involved in—is another trait that mediates the turnover of molecules in the environment (Hu et al. 2022b; Lau and Del Giorgio 2020). The potential biochemical transformations can be inferred from pairwise mass differences between molecules resolved by high-resolution mass spectrometry, and reflect the potential reactions and transformation network relationships between molecules (Danczak et al. 2020; Kajan et al. 2023). This provides valuable insights into the molecular activity associated with DOM dynamics, despite the potential not directly informing when or where a particular transformation occurs (Stegen et al. 2022). However, molecular activity has rarely been explored across divergent systems, and its effects on DOM composition and assembly are poorly understood. Here, we hypothesize that DOM would be assembled more deterministically as molecular activity increases. This is supported by the notion that the more biochemically active DOM assemblages tend to be more closely associated with microbial communities (Danczak et al. 2020). Besides, we hypothesize that the underlying drivers mediating DOM composition vary along the gradient of molecular activity, due to direct effects of microbes and indirect effects of physicochemical properties on the transformation of sediment DOM molecules.

To verify the above hypotheses, we studied variations in the composition and assembly mechanisms of continental-scale sediment DOM along the gradient of molecular activity

represented by the number of potential biochemical transformations. We collected 63 freshwater sediments from tropical to cold temperate climatic zones in China that span 31° of latitude and analyzed their DOM composition by ultra-high Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR MS). Potential drivers of DOM, including climatic parameters, bacterial community structure, and physicochemical factors, were collected. Collectively, this study has three major objectives: (1) determine the composition of DOM in sediments of lake systems; (2) explore the extent to which the composition and assembly of DOM vary with the activity of DOM molecules; and (3) elucidate the biotic and abiotic drivers of DOM composition along the gradient of molecular activity.

Materials and methods

Sediment sampling

In total, 63 sediments were collected from freshwater systems during the short period of July 9th to August 18th, 2020. These freshwater systems range from tropical to cold temperate climatic zones in China, with latitudes from 18°48' to 49°34'N and longitudes from 102°51' to 129°36'E. The mean annual temperature of these systems ranges from 2.85 to 25.65°C (Fig. 1a), and the mean annual precipitation ranges from 175 to 2362 mm. At each site, we collected the upper 5 cm of sediment and the associated surface 20 cm of water three times from the lakeshore at a water depth of 0.5 to 1 m. These samples were mixed to obtain one sediment and one water sample (Wen et al. 2024). All samples were kept in sterile polyethylene terephthalate bottles, wrapped in ice packs, and delivered rapidly to the laboratory for further analysis. Notably, 59 sediments were sampled from ecosystems such as natural lakes and artificial lakes and/or reservoirs, and an additional four sediments were collected from fluvial environments because of limited access to lake sediments at these sites. The fluvial sediments were also included in our analyses as their water depths were similar to the sampled lakes, and their inclusion would not confound the relationships between DOM composition, assembly processes, and molecular activity.

Climatic, physicochemical, and microbial variables

We explored a set of variables potentially associated with the composition and assembly processes of sediment DOM. Climatic variables (i.e., mean annual temperature and mean annual precipitation) were collected using WorldClim Version 2.1 (Fick and Hijmans 2017). Physicochemical variables consist of total phosphorus (TP), total nitrogen (TN), electrical conductivity (EC) and pH of water samples, and TP, TN, EC, pH, total organic carbon (TOC) and dissolved organic carbon (DOC) of sediment samples. Total phosphorus and TN were quantified using a continuous-flow analyzer (Skalar San++, Netherlands). Electrical conductivity and pH were measured directly with an EC (COM-100, HM Digital, USA) and pH probe (SX-610, China), respectively. Sediment TOC was

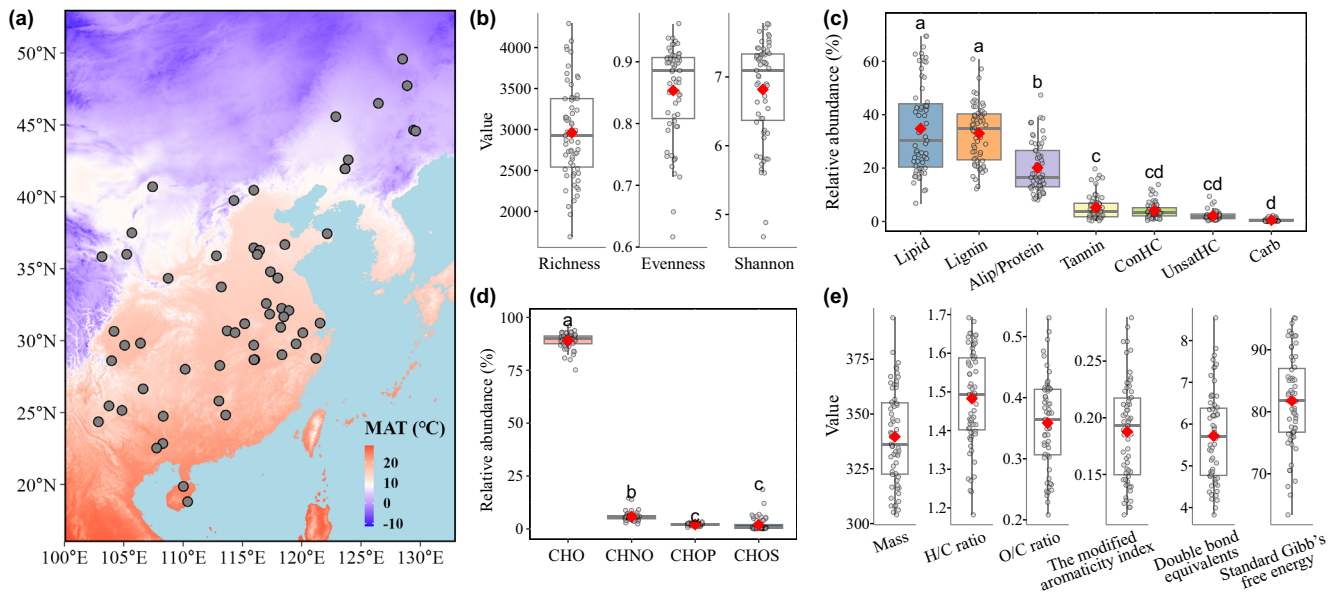


Fig. 1. Chemodiversity of sediment DOM in freshwater ecosystems. **(a)** Sampling map. **(b)** Alpha diversity of DOM. **(c)** Relative abundance of DOM compounds. **(d)** Relative abundance of element combinations of DOM. **(e)** Intensity-weighted mean values of functional traits of DOM. Lowercase letters in panels (c) and (d) represent group differences determined by the LSD test after the ANOVA. Red dots in panels b–e represent mean values. MAT: mean annual temperature. Alip/Protein: aliphatic/protein-like compounds. ConHC: condensed aromatic-like compounds. UnsathC: unsaturated hydrocarbon-like compounds. Carb: carbohydrate-like compounds.

measured with a solid TOC analyzer (SSM-5000A, Shimadzu, Japan). Dissolved organic carbon was analyzed with a Torch TOC analyzer (Teledyne Tekmar, USA). Sediment bacterial communities were profiled via high-throughput 16S rRNA gene sequencing performed on the Illumina MiSeq platform. Data analysis based on amplicon sequence variants was processed using the DADA2 package V1.18 (Callahan et al. 2016). Sequences were rarefied to an equal depth of 32,718 per sample with the “rarefy” function from the R package “vegan” V2.4.6 (Dixon 2003). Further details on the bacterial community analyses are provided in our previous study (Wen et al. 2024).

Fourier transform ion cyclotron resonance mass spectrometer analyses of sediment-dissolved organic matter

First, we extracted DOM from the sediment using ultra-pure water (0.7 g/30 mL) after ultrasonication for 2 h (Hu et al. 2022a). The supernatant was obtained by centrifugation, filtration, and acidification (pH = 2). Second, we extracted the organic matter from the supernatant using pre-activated Oasis HLB cartridges with methanol (ULC-MS grade) and acidic water (pH = 2). Third, DOM in cartridges was eluted into pre-combusted glass vials using methanol, then dried by purging with high-purity nitrogen, and finally stored at -20°C for further analysis. Dissolved organic matter was measured by a 15 Tesla FT-ICR MS (Bruker Daltonics, Billerica, MA) in negative ion mode, and all spectra were scanned 100 times from m/z 150 to 1200. Detailed test and calibration parameters are available in Wen et al.

(2024). Mass spectrum peaks with signal-to-noise ratios > 7 were used to assign molecular formulae according to rigorous elemental combination rules (Hu et al. 2022a; Tolic et al. 2017). According to the stoichiometry of molecules, seven classes were categorized as lipid- ($\text{H/C} = 1.5\text{--}2.0$, $\text{O/C} = 0\text{--}0.3$), lignin- ($\text{H/C} = 0.7\text{--}1.5$, $\text{O/C} = 0.1\text{--}0.67$), aliphatic/protein- (Alip/protein; $\text{H/C} = 1.5\text{--}2.2$, $\text{O/C} = 0.3\text{--}0.67$), condensed aromatic- (ConHC; $\text{H/C} = 0.2\text{--}0.7$, $\text{O/C} = 0\text{--}0.67$), tannin- ($\text{H/C} = 0.5\text{--}1.5$, $\text{O/C} = 0.67\text{--}1.2$), carbohydrate- (Carb; $\text{H/C} = 1.5\text{--}2$, $\text{O/C} = 0.67\text{--}1.2$) and unsaturated hydrocarbon-like compounds (UnsathC; $\text{H/C} = 0.7\text{--}1.5$, $\text{O/C} = 0\text{--}0.1$) (Kim et al. 2003).

The intensity-weighted traits were calculated to characterize DOM properties such as bioavailability and degree of saturation, using the “molTrait” function from the R package iDOM (Meng et al. 2024). These traits included molecular mass, stoichiometry ratios (e.g., H/C, O/C, P/C, S/C, and N/C ratios), the modified aromaticity index, double bond equivalents, and standard Gibb’s free energy of carbon oxidation (Koch and Dittmar 2015; LaRowe and Van Cappellen 2011). The detailed description of these molecular traits was summarized in Supporting Information Table S1. Lower values of H/C ratio and high values of the modified aromaticity index and standard Gibb’s free energy are interpreted as greater recalcitrance of DOM, and high double bond equivalents values imply unsaturation (D’Andrilli et al. 2015).

Molecular activity

Molecular activity, represented by the number of potential biochemical transformations in which a molecule is involved,

was inferred from pairwise mass differences among molecules within sediment DOM and is a unitless measure. Firstly, we paired all molecules to construct a molecular mass difference matrix. Secondly, these pairwise mass differences were compared with a common transformation database within 1 ppm of error (Danczak et al. 2020). A match in mass suggests a potential biochemical transformation pathway between the paired molecules. For instance, a mass difference of 14.01565 between two molecules could presumably indicate a gain or loss of the $-\text{CH}_2$ group. Finally, we quantified the number of potential biochemical transformations associated with each molecule, representing its molecular activity. Molecules were parsed as putatively more biochemically active when they had a larger number of potential biochemical transformations, otherwise as more biochemically inactive (Hu et al. 2022b; Koch and Dittmar 2015). The FT-ICR MS enables us to accurately infer such potential biochemical transformations by detecting pairwise mass differences with a deviation of less than 0.0001 mass units (Breitling et al. 2006). It should be noted that it is challenging to associate these transformations with specific microbial taxa or metabolic processes. This is largely because multiple distinct biochemical pathways exist for the transformation of one molecule to another. The transformation is also constrained by a variety of environmental factors, including temperature, pH, and the composition of microbial communities (Hu et al. 2022b). Consequently, the potential reactions among molecules may not always occur naturally in freshwater ecosystems, and the rates of these reactions may not be directly relevant to molecular activity. Despite these limitations, this methodology remains one of the most effective approaches for inferring transformation networks among DOM molecules and for enhancing the interpretability of currently unidentified biochemical reactions (Ryan et al. 2024). It also provides a comprehensive assessment of the activity of thousands of molecules in diverse environmental DOM samples.

Statistical analyses

To evaluate the effects of molecular activity on DOM composition and assembly, we developed a metabolome binning approach to classify DOM molecules into distinct groups (or bins) based on their number of potential biochemical transformations. The number of molecules in each bin was set as equal, with the first bin having the minimum molecular activity and the last the maximum. Specifically, a moving-window approach was used to categorize all molecules equally into bins based on the rank order of molecular activity. We then used 800 molecules as the window size (i.e., 1–800, 201–1000, ..., 8001–8800, and 8094–8893). This approach generated a total of 42 bins, with bin 1 and bin 42 containing the lowest and highest average number of potential biochemical transformations, that is, the most inactive and most active molecules, respectively. This window size provides enough molecules for the analysis of assembly processes within each bin. In addition, the total number of bins enabled us to

effectively test for relationships between molecular activity and chemodiversity and/or the assembly processes at high resolution. To further validate our results, we also performed additional analyses using molecular window sizes of 600 and 1000. The composition and assembly mechanisms of DOM assemblages within each bin and their association with biotic and abiotic variables were explored to test our hypotheses.

First, the assembly of DOM assemblages in each bin was evaluated using the beta nearest taxon index (βNTI) ecological model (Dini-Andreote et al. 2015). Initially, we multiplied the scaled stepwise transformation distance matrix by the Euclidean distance matrix of 16 molecular traits (Supporting Information Table S1). The above step generated a transformation-weighted molecular characteristics distance matrix, which was used to perform hierarchical clustering analysis by the unweighted pair-cluster method with arithmetic averages to construct the transformation-weighted characteristic dendrogram (Danczak et al. 2020). These analyses were performed using the “hclust” function from the R package stats V4.3.1. Then, βNTI analysis for each bin was performed according to the abundance table. Specifically, the transformation-weighted characteristic dendrogram and a subset of the abundance table for each bin were used to calculate the beta mean nearest taxon distance (βMNTD), which assesses the average phylogenetic distance to the nearest relative between pairs of assemblages (Hu et al. 2022b). The βNTI was calculated as follows:

$$\beta\text{NTI} = \frac{\beta\text{MNTD}_{\text{obs}} - \overline{\beta\text{MNTD}_{\text{null}}}}{SD(\beta\text{MNTD}_{\text{null}})}$$

where $\beta\text{MNTD}_{\text{obs}}$ represents the βMNTD of observed assemblages, $\overline{\beta\text{MNTD}_{\text{null}}}$ and $SD(\beta\text{MNTD}_{\text{null}})$ represent the average and standard deviation βMNTD of null assemblages, respectively (Dini-Andreote et al. 2015). We performed 999 randomized null assemblages for βNTI analyses, using the “qpen” function from the R package iCAMP V1.5.12 (Ning et al. 2020). Dissolved organic matter assemblage was interpreted to be structured by deterministic processes when $|\beta\text{NTI}| > 2$, indicating either variable selection ($\beta\text{NTI} > 2$) or homogeneous selection ($\beta\text{NTI} < -2$). If $|\beta\text{NTI}| \leq 2$, the assemblage was interpreted to be structured by stochastic processes. To improve the robustness of our analyses, we also constructed a molecular characteristic dendrogram and a transformation-based dendrogram, which were used as inputs for βNTI analyses. The molecular characteristic dendrogram and transformation-based dendrogram were constructed using 16 molecular traits in Supporting Information Table S1 and the number of potential biochemical transformations of molecules, respectively. The scripts used to construct these dendrograms can be accessed freely at https://github.com/danczakre/Meta-Metabolome_Ecology.

Second, we employed redundancy analysis to quantify the variables associated with variation in DOM assembly along

the gradient of molecular activity. This would reveal the relative contribution of climatic, physicochemical, and bacterial variables to the DOM assemblage in each bin. Climatic variables consisted of mean annual temperature and mean annual precipitation. Physicochemical variables included TOC, DOC, TOC/TN ratio, pH, and EC of sediment samples, and EC of water samples. Total nitrogen in sediment samples was excluded in these analyses due to its high Pearson's correlation ($P < 0.001$) with sediment TOC. Likewise, TN, TP, and pH of water samples were excluded as they did not significantly contribute to the variation observed in the assemblages. We performed non-metric multidimensional scaling to analyze variations in bacterial communities according to their Bray–Curtis dissimilarity, and then used the first two axes of the non-metric multidimensional scaling as explanatory variables. For response variables, we used the first eight axes of the non-metric multidimensional scaling according to Bray–Curtis dissimilarities of DOM assemblages. The individual contribution of each predictor to the total variance of the assemblages was determined using the hierarchical partitioning method performed by the R package *rdacca.hp* V1.1.0 (Lai et al. 2022). A detailed explanation of this method can be found in Lai et al. (2022).

Third, to explore variations in associations between bacterial taxa and DOM along the molecular activity gradient, we constructed co-occurrence bipartite networks. This was performed using data on bacterial amplicon sequence variants and molecules subjected to Spearman's rank correlations. To ensure the reliability of these correlations, only the paired species and molecules that appeared in more than 20 samples were used to generate the network. We retained correlation matrices with significance levels below 0.05 (de Vries et al. 2018). As we focused on correlations between bacterial taxa and molecules, bacteria–bacteria and molecule–molecule associations were removed from the final network. Furthermore, we extracted sub-networks based on the molecules within the 42 bins and calculated topological parameters for each sub-network that include the number of nodes, number of links, average degree, diameter, and average path length. We used the “*net_properties*” function from the R package *ggClusterNet* V0.1.0 to compute these topological parameters (Wen et al. 2022) and visualized these networks using Gephi (V0.10.1).

We calculated alpha diversity of DOM, including richness, evenness, and Shannon index, as described in Supporting Information Text S1. Relationships between compound abundances and/or DOM traits with molecular activity were examined using linear or quadratic models, depending on the smaller value of Akaike's information criterion (Akaike 1974). Regression analysis was used to model the variations in traits and assembly processes of DOM assemblages for each bin along the molecular activity gradient. Least squares algorithms were used to test the relationships between the relative contribution of potential predictors or network parameters and

molecular activity. All of the above statistics were conducted in R version 4.3.1 (R Core Team 2023).

Results

Chemodiversity of sediment dissolved organic matter

A total of 8893 formulae were identified in sediment DOM, with the molecular richness ranging from 1690 to 4293 (Fig. 1b). The evenness and Shannon diversity of sediment DOM ranged from 0.6 to 1 and 4.7 to 7.8, respectively. We found DOM to be dominated by lipid- (34.8% \pm 16.9%) and lignin-like (33% \pm 11%) compounds, followed by aliphatic/protein-like (20.1% \pm 9.4%), while the other compounds, such as tannin- or condensed aromatic-like, contributed less than 5.3% (Fig. 1c). With respect to elemental combinations, sediment DOM was primarily composed of CHO compounds (89.1% \pm 3.6%, Fig. 1d). Molecules containing heteroatoms such as CHNO, CHOP, and CHOS accounted for 5.8% \pm 2%, 2.1% \pm 0.5%, and 1.9% \pm 3.2%, respectively (Fig. 1d). The analysis of functional traits of DOM at the compositional level revealed high spatial heterogeneity (Fig. 1e). Specifically, molecular mass ranged from 304 to 393.9 Da, with a mean of 339.6 Da. H/C and O/C ratios ranged from 1.2 to 1.7 and 0.2 to 0.5, with a mean of 1.48 and 0.35, respectively. The modified aromaticity index (0.1–0.3), double bond equivalents (3.8–8.6) and standard Gibb's free energy (63.4–95.2) showed means of 0.19, 5.7, and 81.8, respectively.

Varying sediment dissolved organic matter chemodiversity with molecular activity

For the 8893 unique molecules detected in sediment DOM, the potential biochemical transformations ranged from 0 to 79. Toward high biochemical transformations, molecular frequency showed a decreasing trend (Fig. 2a), while the molecular relative abundances increased significantly ($P < 0.001$, Fig. 2b). As for the molecules in each bin, lignin-, lipid-, and aliphatic/protein-like compounds dominated but varied across the potential biochemical transformations (average number for bins, the same below). Specifically, the molecular percentage of lignin-like compounds increased significantly with increasing potential biochemical transformations, with R^2_{adj} of 0.98 ($P < 0.001$, Fig. 2c; Supporting Information Fig. S1). Conversely, the percentage of aliphatic/protein-, condensed aromatic-, unsaturated hydrocarbon-, and carbohydrate-like compounds decreased significantly and reached zero, with R^2_{adj} values ranging from 0.46 to 0.91 (Supporting Information Fig. S1). The relative abundance of lipid-like substances exhibited a hump-shaped pattern, in contrast to the pattern of lignin-like compounds (Fig. 2d; Supporting Information Fig. S2). Furthermore, we found a decreasing trend in the Bray–Curtis dissimilarities of molecules in the bins as the potential biochemical transformations increased ($R^2_{\text{adj}} = 0.65$, $P < 0.001$, Fig. 2e). Taken together, these results demonstrate that DOM assemblages with higher potential biochemical

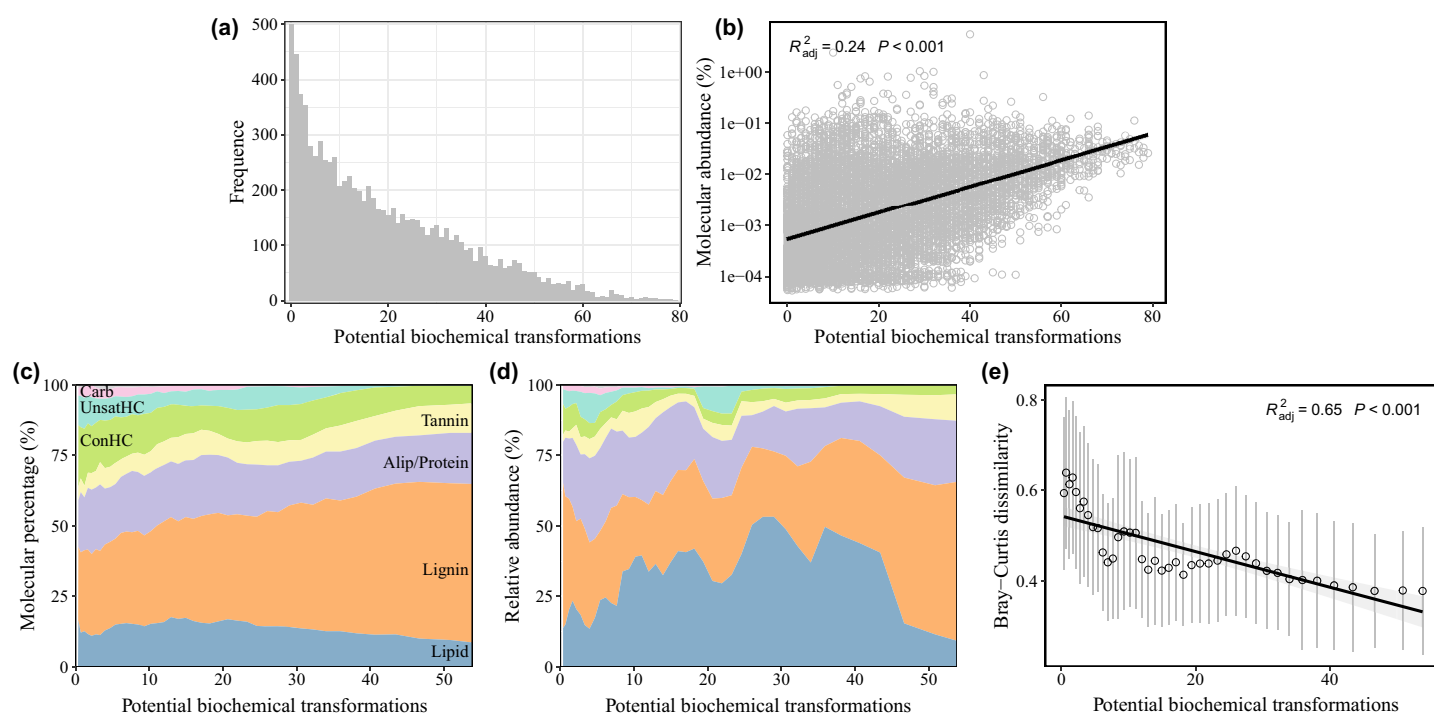


Fig. 2. Variations in chemodiversity of sediment DOM along the gradient of molecular activity represented by the number of potential biochemical transformations. Panels (a) and (b) depict variations in the frequency and relative abundance of molecules in DOM. Panels (c) and (d) display variations in molecular percentage and relative abundance of different classes within each bin. Gray areas indicate other molecules that are not classified as specific compounds. (e) Bray–Curtis dissimilarity of the DOM assemblage within each bin as a function of the potential biochemical transformations, with dots representing means and error bars representing standard deviations. For the visualization of the data in the bins, the x-axis denotes the average number of potential biochemical transformations of all molecules in each bin. Alip/Protein: aliphatic/protein-like compounds. ConHC: condensed aromatic-like compounds. UnsathC: unsaturated hydrocarbon-like compounds. Carb: carbohydrate-like compounds.

transformations are increasingly dominated by recalcitrant lignin-like compounds and are more compositionally similar.

Variations in the intensity-weighted traits of molecules in bins, including molecular mass, elemental stoichiometric ratios, the modified aromaticity index, double bond equivalents, and standard Gibb's free energy, exhibited complex and distinct patterns with increasing potential biochemical transformations (Supporting Information Fig. S3). The molecular mass was found to decrease rapidly from 482 ± 50 Da to 324 ± 23 Da as the potential biochemical transformations increased from 0.4 to 10, remaining stable thereafter (Supporting Information Fig. S3a). The modified aromaticity index and double bond equivalents followed a parabolic pattern with a minimum around 16, while the H/C ratio followed the opposite pattern (Supporting Information Fig. S3). Standard Gibb's free energy and O/C ratio had opposite patterns, with extremes observed at the potential biochemical transformations of 29 (Supporting Information Fig. S3c,i). For features characterizing heteroatoms, such as N/C, P/C, and S/C ratios, it was found that they decrease and converge to zero with increasing potential biochemical transformations (Supporting Information Fig. S3d–f). This indicates that active compounds are less likely to contain heteroatoms.

Sediment-dissolved organic matter assembly integrated with molecular activity

For the assembly of DOM, we found that the interplay of stochastic and deterministic processes varies across the potential biochemical transformations. Specifically, the relative contribution of stochastic processes sharply dropped from 93% to 5.1% as the potential biochemical transformations increased from 0.4 to 14, having a slight increase or being relatively stable thereafter (Fig. 3). Conversely, the contribution of variable selection changed from 6.2% to 95% and remained relatively stable thereafter. The contribution of homogeneous selection was found to be < 1.6%. The robustness of these findings was corroborated by similar analyses applied to the molecular characteristic dendrogram or transformation-based dendrogram with window sizes of 800 molecules and the transformation-weighted characteristic dendrogram with window sizes of 600 and 1000 molecules (Supporting Information Fig. S4).

Variations in the assembly processes of DOM assemblages were also well-predicted by other functional traits. For instance, we found the relative importance of variable selection to decrease monotonically with mass and double bond equivalents ($P < 0.001$), with R^2_{adj} of 0.89 and 0.82,

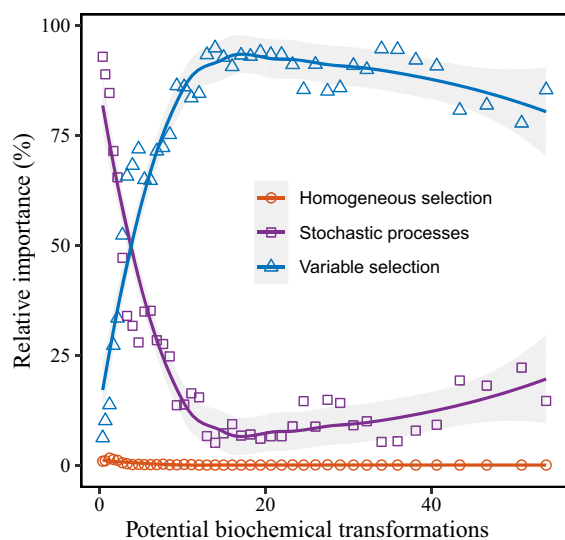


Fig. 3. Relative importance of assembly processes structuring the DOM assemblages within bins. The x-axis denotes the average number of potential biochemical transformations of all molecules in each bin. Shadings represent 95% confidence intervals. The processes that structure DOM assemblages are analogous to the ecological processes that govern the assembly of biological communities, with homogeneous selection and variable selection considered deterministic processes, and dispersal and drift as stochastic processes.

respectively (Fig. 4a,e). Even though the relative importance of variable selection had significant correlations ($P < 0.01$) with traits of H/C ratio, the modified aromaticity index, and standard Gibb's free energy, the R^2_{adj} values were ≤ 0.23 (Fig. 4b,d,f). Furthermore, DOM assemblages with higher potential biochemical transformations displayed greater variability in H/C ratio, O/C ratio, the modified aromaticity index, and standard Gibb's free energy, but converged at lower molecular mass and double bond equivalents (Fig. 4). Thus, we considered mass and double bond equivalents as potentially important traits associated with molecular activity to determine the assembly processes of DOM.

We found that the composition of DOM assemblages was mostly associated with sediment physicochemical properties, while the contributions of climatic and bacterial variables increased with increasing potential biochemical transformations. Specifically, the total variance (R^2_{adj}) of DOM assemblages within bins explained by these variables ranged from 16.5% to 24.9% (Fig. 5a), with greater contributions from physicochemical variables (i.e., sediment TOC, EC, and pH) (Fig. 5b; Supporting Information Fig. S5). Moreover, the contributions of climate and bacterial composition were found to increase with increasing potential biochemical transformations, while the contribution of physicochemical variables exhibited a decreasing trend (Fig. 5b). Although the two lines for bacteria and climate appeared to have similar slopes, there was no significant correlation between them (Supporting Information Fig. S6). The increased effect of bacterial

communities was further supported by the increasing complexity of DOM–bacteria bipartite networks (Fig. 6a). Specifically, the number of nodes, links, and average degree of networks were found to increase, while the average path length and diameter decreased significantly ($P < 0.001$; Fig. 6). In addition, we found that the number of molecular nodes increased significantly (Fig. 6b), while the average degree of molecules remained relatively stable (Fig. 6d). Conversely, bacterial taxa exhibited an opposite pattern. This indicated that the increase in the number of nodes was mostly attributed to DOM molecules (especially lignin-like compounds, Fig. 6b; Supporting Information Fig. S7), while the increase in average degree was mostly attributed to bacterial taxa (especially *Chloroflexi*, *Bacteroidota*, and *Actinobacteriota*, Fig. 6d; Supporting Information Figs. S8, S9). Therefore, the strengthening of the network relationships was primarily due to the increase in the number of molecules associated with bacteria. Bacterial taxa in the network may be comparable across DOM assemblages with differing potential biochemical transformations (Supporting Information Fig. S7b), but the frequency of bacterial–molecular interactions is likely to vary.

Discussion

Sediment DOM is transported and transformed at the sediment–water interface and plays a crucial role in the regional and global carbon cycling (Drake et al. 2018; Moras et al. 2024; Wen et al. 2024). In the present study, we investigated changes in DOM composition and assembly processes in transcontinental lake sediments along the gradient of molecular activity, using the proposed high-performance metabolome binning approach. The findings support our hypotheses that DOM assemblages with higher molecular activity are assembled more deterministically and are increasingly influenced by microbes, as more active molecules are associated with bacteria. Our results emphasize that the composition and assembly mechanisms of sediment DOM are regulated by environmental factors including climate, physicochemical factors, and microbial communities, which are further constrained by intrinsic molecular traits such as molecular activity. The approach we used, along with the large-scale study of lake sediments, is innovative and contributes to a deeper understanding of DOM turnover and assembly mechanisms, as well as the impact of molecular activity.

Dissolved organic matter in lake sediments exhibits distinctive chemodiversity. The Shannon diversity of sediment DOM is lower than that observed in 45 lake waters (7.2–8.4) (Luo et al. 2022) and in 11 lakeshore surface sediments (7–8.2) (Li et al. 2023b). This is consistent with the understanding of higher alpha diversity of DOM in waters than in sediments (Cui et al. 2024) because freshwater DOM originates from a variety of sources, including terrestrial, anthropogenic, and autochthonous contributions, and is profoundly influenced by hydrological processes such as vertical and longitudinal

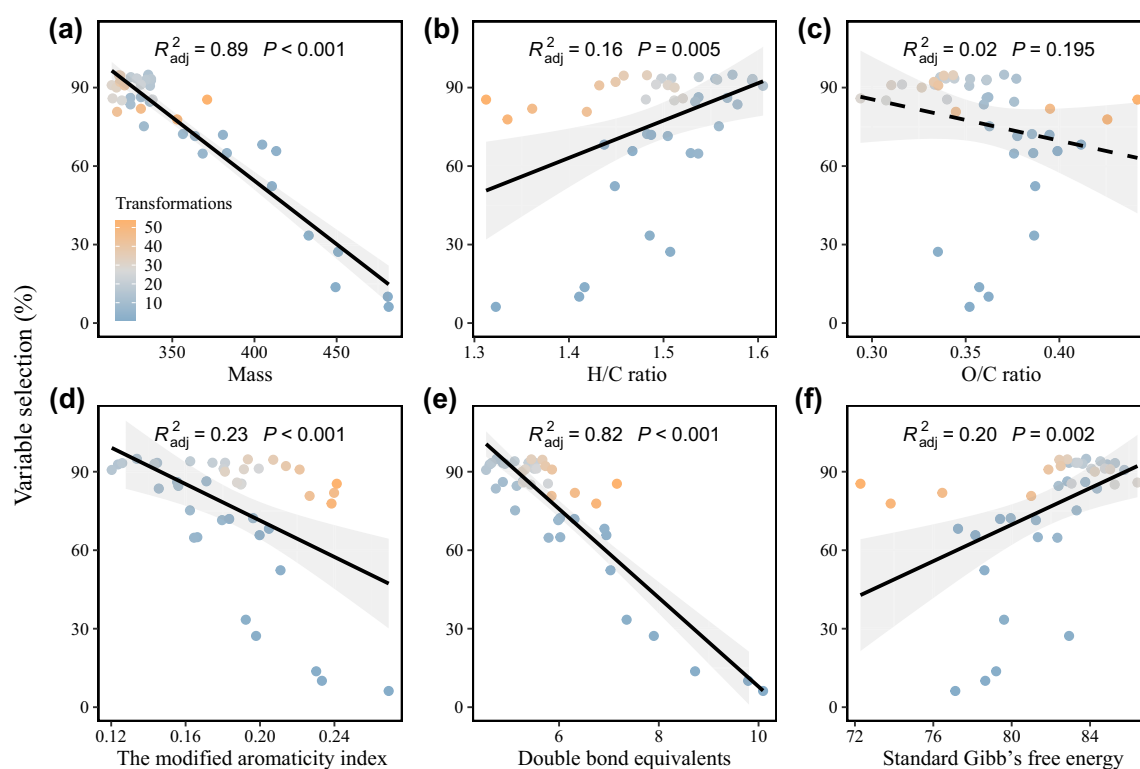


Fig. 4. Relationships between the relative importance of variable selection and functional traits of sediment DOM. Lines and shadings indicate linear regression and 95% confidence intervals. The colors of the dots in the panels represent the average number of potential biochemical transformations of molecules in each bin.

migration (Kellerman et al. 2014; Lynch et al. 2019). The discrepancy in sediment DOM diversity is largely attributed to the different study areas and scales, suggesting the necessity for research conducted at broader geographic scales that could provide more comprehensive and representative insights (Cui et al. 2024; Ma et al. 2024).

The dominance of lignin-like compounds in sediment DOM is consistent with previous studies in water systems and inland sediments (Kajan et al. 2023; Li et al. 2023b), and coastal wetlands (Li et al. 2022). This could be attributed to the accumulation of soluble compounds released from the complexly transformed plant biomass and soil organic matter, as well as their low bioavailability. A distinct finding, however, is the high abundance of lipid-like substances detected in lake sediments (34.8%, Fig. 1c), which is much higher than that in lake waters (< 5%), soils (< 5%) and coastal sediments (< 10%) but similar to that in mudflat intertidal sediments (32.1%) (Ding et al. 2020; Kajan et al. 2023; Li et al. 2022; Ma et al. 2024). This implies that microorganisms may contribute a lot to lake sediment DOM by utilizing a higher abundance of CHO compounds (Fig. 1d) to produce lipid-like substances with high hydrogen to carbon ratios and low oxygen to carbon ratios (Luo et al. 2022; Valle et al. 2020). Functional traits of lake sediment DOM also exhibit a relatively unique range

compared to other habitats. The molecular mass of sediment DOM (304–393.9 Da, Fig. 1e) is consistent with the mass range observed in lake waters (Kajan et al. 2023) but lower than that reported for riverine waters (518.6–567.1 Da) (Li et al. 2023a), and for riverine sediments with a mean value of 511 Da (He et al. 2016). Sediment DOM of lakes also showed a lower modified aromaticity index and double bond equivalents, and a greater H/C ratio compared to riverine sediments (Han et al. 2024). Despite the variations in climatic and anthropogenic variables across large spatial scales playing a fundamental role in shaping the chemodiversity of DOM (Kida et al. 2023; Lehmann et al. 2020), the distinctive diversity, functional traits, and molecular composition of DOM in lake sediments are partially attributed to unique habitat characteristics, such as relative hypoxia, darkness, and limited hydraulic transport.

The percentage and relative abundance of lignin-like recalcitrant compounds in sediment DOM increased with increasing molecular activity (Fig. 2). While some molecules exhibit higher activity, they can still resist transformation (Hu et al. 2022b). This is because a molecule's activity—represented by the number of potential biochemical transformations that the molecule is involved in—does not directly determine its bioavailability or degradability. Instead, the

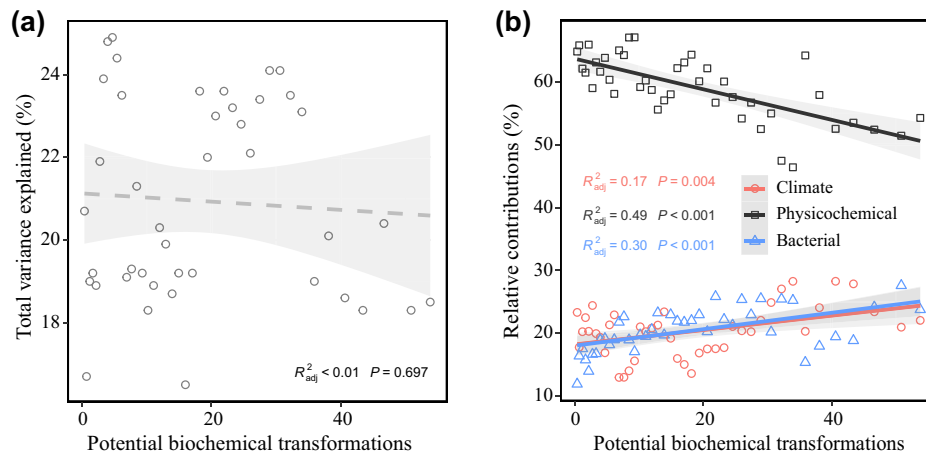


Fig. 5. Variances of DOM assemblages within bins explained by climatic, physicochemical, and bacterial variables. The x-axis denotes the average number of potential biochemical transformations of all molecules in each bin. **(a)** Total variance of the DOM assemblages explained by explanatory variables. **(b)** Relative contributions of climatic, physicochemical, and bacterial variables to the variances of the DOM assemblages. Lines and shadings indicate linear regression and 95% confidence intervals. Climatic variables include mean annual temperature and mean annual precipitation. Physicochemical variables include electrical conductivity of water, carbon to nitrogen ratio, electrical conductivity, pH, dissolved organic carbon, and total organic carbon of sediment.

activity merely reflects the possible transformation pathways of the molecule (Breitling et al. 2006). Whether and how fast a biochemical reaction occurs depends on various conditions, including molecular reactivity, microbial communities, the

concentration of reactants, temperature, and light (Berg et al. 2023; Hu et al. 2022b; Pearson et al. 2024). For instance, sunlight irradiation converts high molecular weight aromatic compounds, which are difficult for microorganisms to

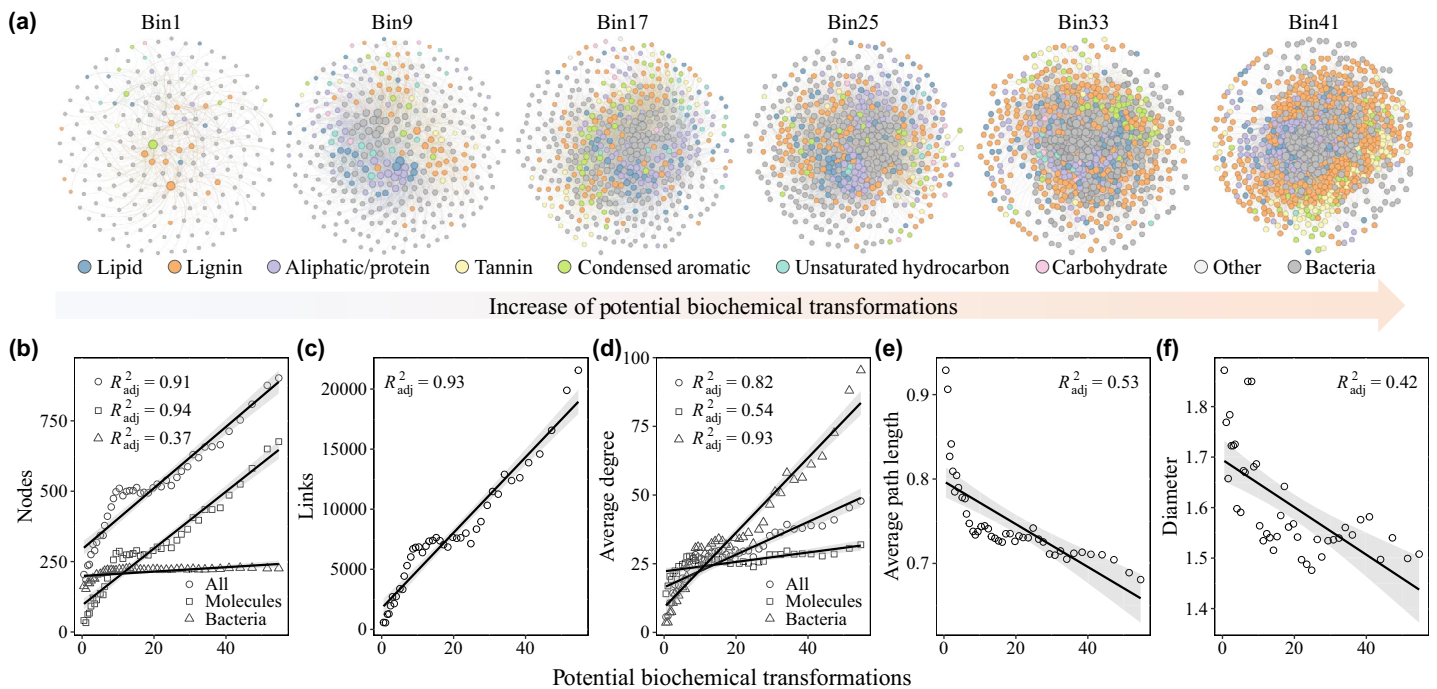


Fig. 6. Bacteria-molecules bipartite networks along the activity gradient. **(a)** Only six bins with evenly spaced intervals are presented to illustrate their network changes across the molecular activity gradient. The network graphs of all remaining bins are shown in Supporting Information Fig. S10. **(b-f)** Topology properties of networks, including the network number of nodes (b), number of links (c), average degree (d), average path length (e), and diameter (f). Lines and shadings represent linear regression ($P < 0.001$) and 95% confidence intervals. The x-axis denotes the average number of potential biochemical transformations of all molecules in each bin.

degrade, into small aromatic molecules that are easily utilized (Hu et al. 2023). Recalcitrant molecules such as lignin-like compounds are more sensitive to temperature and therefore could be degraded more rapidly in warmer conditions (Craine et al. 2010; Davidson and Janssens 2006).

We further elucidate how the assembly processes of DOM assemblages change across the gradient of molecular activity, using the metabolome binning approach that classifies molecules into distinct bins based on their activity. For instance, when the molecular activity of an assemblage—quantified by the average number of potential biochemical transformations—is less than 2.8, the assemblage is predominantly structured by stochastic processes (Fig. 3). This indicates that these inactive molecules within the assemblage are rarely produced or transformed and are more influenced by random events such as hydrological transport and vector movement (Hu et al. 2022b). Conversely, when molecular activity exceeds 14, the relative importance of deterministic processes maintains a relatively stable contribution of approximately 80–90% (Fig. 3). The predominance of deterministic processes in structuring DOM assemblages with higher molecular activity is partly because active molecules are more closely associated with microorganisms and greater influenced by biochemical reactions (Danczak et al. 2020).

The metabolome binning approach enhances our understanding of the mechanisms that structure DOM assemblages within specific activity segments, particularly low-activity segments (e.g., < 10). This is largely because DOM samples contain molecules with potential biochemical transformations ranging from zero to hundreds; yet the mean or intensity-weighted mean number of the transformations is typically greater than 10 (Hu et al. 2022b; Ryan et al. 2024). For instance, the analysis of our 63 sediment DOM samples at the compositional level revealed that they were structured by deterministic processes with a contribution of 99.6% (Supporting Information Fig. S11). The mean or intensity-weighted mean numbers of potential biochemical transformations of these DOM samples ranged from 23.5 to 29.3 and 21.9 to 31, respectively. Consequently, detailed information regarding low-activity DOM assemblages could be overlooked when analyzing bulk DOM at the compositional level. In summary, the influence of molecular traits such as activity on DOM assembly mechanisms can be more effectively revealed using the metabolome binning approach based on trait gradients at the molecular level.

The strong linear relationships between assembly processes and both molecular mass and double bond equivalents suggest that these two molecular traits are functionally linked to molecular activity in mediating the assembly of sediment DOM. Molecular mass and double bond equivalents are important properties of DOM that govern its selective adsorption by minerals (Lv et al. 2016). Additionally, molecular mass is critical to the photo- and bio-degradation processes of DOM (Catalán et al. 2017). Molecules with low mass are known to

exhibit high photo-reactivity, as indicated by high quantum yields (Zhang et al. 2022). These molecules are also preferentially metabolized by bacteria due to their greater labile nature and bioavailability (Roth et al. 2019). Thus, variable selection dominates the assembly processes of assemblages with low molecular mass or double bond equivalents (Fig. 4). The above factors are likely to contribute to the higher activity of the low molecular mass and double bond equivalents fractions (Supporting Information Fig. S3). Consequently, changes in molecular activity are accompanied by synchronized alterations in mass and double bond equivalents, collectively mediating the turnover and composition of sediment DOM.

Both external environmental conditions and the nature of DOM molecules are known to influence DOM turnover and shape its diversity and composition. However, the interaction between molecular activity and environmental variables remains unclear. Our findings indicated that the total variance in DOM assemblages within bins explained by climatic, physicochemical, and microbial variables was insignificantly correlated with molecular activity (Fig. 5a). This seems to contradict the positive relationship between the importance of deterministic processes and molecular activity (Fig. 3). The discrepancy may arise because the composition of sediment DOM is dynamically regulated by various factors, including benthic fauna, plant composition, and fungi (Taube et al. 2018; Wetzel and Søndergaard 1998), which were not included in our study. Nevertheless, we observed that different environmental variables responded variably to molecular activity in regulating DOM composition. Specifically, as molecular activity increased, the contribution of physicochemical factors decreased, while the influence of microorganisms and climate increased (Fig. 5b). This is attributed to the fact that more active molecules possess greater potential for biochemical transformations and are more closely associated with microorganisms, which aligns with our findings on the DOM–bacterial networks (Fig. 6). Additionally, climatic factors such as rising temperatures may reduce the activation energy required for reactions or transformations, thereby accelerating the decomposition of certain recalcitrant compounds (Davidson and Janssens 2006), which are abundant in sediment DOM, such as lignin-like compounds with higher molecular activity (Fig. 2). In summary, our study highlights that the role of biotic and abiotic factors in regulating DOM turnover is further constrained by the intrinsic activity of molecules.

Conclusions

Our study across large spatial scales provides evidence for a strong coupling between the intrinsic activity of DOM molecules and their composition or assembly in lake sediments. Dissolved organic matter molecules with higher activity were more enriched in lignin-like compounds and exhibited greater compositional similarity across diverse climatic zones. As the potential biochemical transformations of DOM assemblages

increased from 0.4 to 14, their assembly processes shifted from stochastically to deterministically dominant, indicating that active molecules are more prone to deterministic assembly. While physicochemical variables emerged as the primary drivers of DOM composition, the influence of climatic and bacterial variables increased with increasing molecular activity, consistent with the more complex network relationships between the active molecules and bacterial communities. The leveraging of a metabolome binning approach at the molecular level provides an important tool to explore the processes and mechanisms structuring DOM that would be less possible via compositional-level analyses. In summary, our findings elucidate the composition, assembly processes, and structuring mechanisms of DOM in lake systems. This represents an important step toward advancing our understanding of DOM dynamics in freshwaters and provides new insights into the global carbon cycle. Future studies are encouraged to further describe and compare the DOM compositions of the lakes across such large spatial scales and to support our findings of molecular traits in determining DOM turnover.

Author Contributions

Shuailong Wen and Ang Hu contributed equally to this study. Jianjun Wang conceptualized the study. Lei Han, Shuyu Jiang, Ang Hu, and Jianjun Wang collected the samples. Kyoung-Soon Jang analyzed the DOM with the FT-ICR MS. Shuailong Wen conducted the statistical analyses and wrote the first draft with the guidance of Jianjun Wang and Ang Hu. Shuailong Wen revised the manuscript based on suggestions from Jianjun Wang, Francisco Dini-Andreote, and Ang Hu. All authors contributed to the intellectual development of this study.

Acknowledgments

We thank the editors and reviewers for their valuable comments. We appreciate Fanfan Meng, Jinfu Liu, Minglei Ren, Hao Wu, Yanan Zhou, Jianing Xu, Jiyi Wang, Pubo Chen, Huilin Liu, Zhenghua Liu, Weizhen Zhang, Xu Ma, Chaoxun Guo, and Yunfeng Wang for their assistance with field sampling. This work was financially supported by research grants from the National Natural Science Foundation of China (42225708, U24A0578, 42377122, 42307323, 92251304), the Second Tibetan Plateau Scientific Expedition and Research Program (STEP, 2019QZKK0503), the Basic Research Program of Jiangsu (BK20240111), the Key Laboratory of Lake and Watershed Science for Water Security (NKL2023-QN04), the China Postdoctoral Science Foundation (2024M753322), and the Science and Technology Planning Project of NIGLAS (NIGLAS2022GS09).

Conflicts of Interest

None declared.

References

- Akaike, H. 1974. "A New Look at the Statistical Model Identification." *IEEE Transactions on Automatic Control* 19, no. 6: 716–723. <https://doi.org/10.1109/TAC.1974.1100705>.
- Berg, S. M., K. H. Wammer, and C. K. Remucal. 2023. "Dissolved Organic Matter Photoreactivity Is Determined by its Optical Properties, Redox Activity, and Molecular Composition." *Environmental Science & Technology* 57, no. 16: 6703–6711. <https://doi.org/10.1021/acs.est.3c01157>.
- Berggren, M., F. Guillemette, M. Bieroza, et al. 2022. "Unified Understanding of Intrinsic and Extrinsic Controls of Dissolved Organic Carbon Reactivity in Aquatic Ecosystems." *Ecology* 103, no. 9: e3763. <https://doi.org/10.1002/ecy.3763>.
- Brailsford, F. L., H. C. Glanville, M. R. Marshall, et al. 2021. "Land Cover and Nutrient Enrichment Regulates Low-Molecular Weight Dissolved Organic Matter Turnover in Freshwater Ecosystems." *Limnology and Oceanography* 66, no. 8: 2979–2987. <https://doi.org/10.1002/lno.11852>.
- Breitling, R., S. Ritchie, D. Goodenowe, M. L. Stewart, and M. P. Barrett. 2006. "Ab Initio Prediction of Metabolic Networks Using Fourier Transform Mass Spectrometry Data." *Metabolomics* 2, no. 3: 155–164. <https://doi.org/10.1007/s11306-006-0029-z>.
- Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes. 2016. "DADA2: High-Resolution Sample Inference From Illumina Amplicon Data." *Nature Methods* 13, no. 7: 581–583. <https://doi.org/10.1038/nmeth.3869>.
- Catalán, N., J. P. Casas-Ruiz, D. von Schiller, et al. 2017. "Biodegradation Kinetics of Dissolved Organic Matter Chromatographic Fractions in an Intermittent River." *Journal of Geophysical Research: Biogeosciences* 122, no. 1: 131–144. <https://doi.org/10.1002/2016JG003512>.
- Catalán, N., R. Marcé, D. N. Kothawala, and L. J. Tranvik. 2016. "Organic Carbon Decomposition Rates Controlled by Water Retention Time Across Inland Waters." *Nature Geoscience* 9, no. 7: 501–504. <https://doi.org/10.1038/ngeo2720>.
- Catalán, N., C. Rofner, C. Verpoorter, et al. 2024. "Treeline Displacement May Affect Lake Dissolved Organic Matter Processing at High Latitudes and Altitudes." *Nature Communications* 15, no. 1: 2640. <https://doi.org/10.1038/s41467-024-46789-5>.
- Chen, Q., C. Lønborg, F. Chen, et al. 2022. "Increased Microbial and Substrate Complexity Result in Higher Molecular Diversity of the Dissolved Organic Matter Pool." *Limnology and Oceanography* 67, no. 11: 2360–2373. <https://doi.org/10.1002/lno.12206>.
- Craine, J. M., N. Fierer, and K. K. McLauchlan. 2010. "Widespread Coupling Between the Rate and Temperature Sensitivity of Organic Matter Decay." *Nature Geoscience* 3, no. 12: 854–857. <https://doi.org/10.1038/ngeo1009>.
- Cui, Y., S. Wen, J. C. Stegen, A. Hu, and J. Wang. 2024. "Chemodiversity of Riverine Dissolved Organic Matter:

- Effects of Local Environments and Watershed Characteristics.” *Water Research* 250: 121054. <https://doi.org/10.1016/j.watres.2023.121054>.
- D’Andrilli, J., W. T. Cooper, C. M. Foreman, and A. G. Marshall. 2015. “An Ultrahigh-Resolution Mass Spectrometry Index to Estimate Natural Organic Matter Lability.” *Rapid Communications in Mass Spectrometry* 29, no. 24: 2385–2401. <https://doi.org/10.1002/rcm.7400>.
- Danczak, R. E., R. K. Chu, S. J. Fansler, et al. 2020. “Using Metacommunity Ecology to Understand Environmental Metabolomes.” *Nature Communications* 11, no. 1: 6369. <https://doi.org/10.1038/s41467-020-19989-y>.
- Davidson, E. A., and I. A. Janssens. 2006. “Temperature Sensitivity of Soil Carbon Decomposition and Feedbacks to Climate Change.” *Nature* 440, no. 7081: 165–173. <https://doi.org/10.1038/nature04514>.
- de Vries, F. T., R. I. Griffiths, M. Bailey, et al. 2018. “Soil Bacterial Networks Are Less Stable under Drought Than Fungal Networks.” *Nature Communications* 9, no. 1: 3033. <https://doi.org/10.1038/s41467-018-05516-7>.
- Ding, Y., Z. Shi, Q. Ye, et al. 2020. “Chemodiversity of Soil Dissolved Organic Matter.” *Environmental Science & Technology* 54, no. 10: 6174–6184. <https://doi.org/10.1021/acs.est.0c01136>.
- Dini-Andreote, F., J. C. Stegen, J. D. van Elsas, and J. F. Salles. 2015. “Disentangling Mechanisms That Mediate the Balance Between Stochastic and Deterministic Processes in Microbial Succession.” *Proceedings of the National Academy of Sciences of the United States of America* 112, no. 11: E1326–E1332. <https://doi.org/10.1073/pnas.1414261112>.
- Dixon, P. 2003. “VEGAN, a Package of R Functions for Community Ecology.” *Journal of Vegetation Science* 14, no. 6: 927–930. <https://doi.org/10.1111/j.1654-1103.2003.tb02228.x>.
- Drake, T. W., P. A. Raymond, and R. G. M. Spencer. 2018. “Terrestrial Carbon Inputs to Inland Waters: A Current Synthesis of Estimates and Uncertainty.” *Limnology and Oceanography Letters* 3, no. 3: 132–142. <https://doi.org/10.1002/lo2.10055>.
- Fick, S. E., and R. J. Hijmans. 2017. “WorldClim 2: New 1-Km Spatial Resolution Climate Surfaces for Global Land Areas.” *International Journal of Climatology* 37, no. 12: 4302–4315. <https://doi.org/10.1002/joc.5086>.
- Gonsior, M., L. C. Powers, E. Williams, et al. 2019. “The Chemodiversity of Algal Dissolved Organic Matter From Lysed *Microcystis Aeruginosa* Cells and its Ability to Form Disinfection by-Products during Chlorination.” *Water Research* 155: 300–309. <https://doi.org/10.1016/j.watres.2019.02.030>.
- Han, L., A. Hu, L. H. Mzuka, X. Chen, J. Shen, and J. Wang. 2024. “Molecular Properties of Dissolved Organic Matter Across Earth Systems: A Meta-Analysis.” *Journal of Earth Science*. <https://doi.org/10.1007/s12583-024-0061-9>.
- He, W., M. Chen, J. E. Park, and J. Hur. 2016. “Molecular Diversity of Riverine Alkaline-Extractable Sediment Organic Matter and its Linkages With Spectral Indicators and Molecular Size Distributions.” *Water Research* 100: 222–231. <https://doi.org/10.1016/j.watres.2016.05.023>.
- Hu, A., M. Choi, A. J. Tanentzap, et al. 2022a. “Ecological Networks of Dissolved Organic Matter and Microorganisms Under Global Change.” *Nature Communications* 13, no. 1: 3600. <https://doi.org/10.1038/s41467-022-31251-1>.
- Hu, A., K. S. Jang, F. Meng, et al. 2022b. “Microbial and Environmental Processes Shape the Link Between Organic Matter Functional Traits and Composition.” *Environmental Science & Technology* 56, no. 14: 10504–10516. <https://doi.org/10.1021/acs.est.2c01432>.
- Hu, J., L. Kang, Z. Li, et al. 2023. “Photo-Produced Aromatic Compounds Stimulate Microbial Degradation of Dissolved Organic Carbon in Thermokarst Lakes.” *Nature Communications* 14, no. 1: 3681. <https://doi.org/10.1038/s41467-023-39432-2>.
- Hubbell, S. P. 2001. *The Unified Neutral Theory of Biodiversity and Biogeography* (MPB-32). Princeton: Princeton University Press. <https://doi.org/10.1515/9781400837526>.
- Kajan, K., H. Osterholz, J. Stegen, M. Gligora Udovic, and S. Orlic. 2023. “Mechanisms Shaping Dissolved Organic Matter and Microbial Community in Lake Ecosystems.” *Water Research* 245: 120653. <https://doi.org/10.1016/j.watres.2023.120653>.
- Kellerman, A. M., T. Dittmar, D. N. Kothawala, and L. J. Tranvik. 2014. “Chemodiversity of Dissolved Organic Matter in Lakes Driven by Climate and Hydrology.” *Nature Communications* 5: 3804. <https://doi.org/10.1038/ncomms4804>.
- Kida, M., J. Merder, N. Fujitake, et al. 2023. “Determinants of Microbial-Derived Dissolved Organic Matter Diversity in Antarctic Lakes.” *Environmental Science & Technology* 57, no. 13: 5464–5473. <https://doi.org/10.1021/acs.est.3c00249>.
- Kim, S., R. W. Kramer, and P. G. Hatcher. 2003. “Graphical Method for Analysis of Ultrahigh-Resolution Broadband Mass Spectra of Natural Organic Matter, the Van Krevelen Diagram.” *Analytical Chemistry* 75, no. 20: 5336–5344. <https://doi.org/10.1021/ac034415p>.
- Koch, B. P., and T. Dittmar. 2015. “From Mass to Structure: An Aromaticity Index for High-Resolution Mass Data of Natural Organic Matter.” *Rapid Communications in Mass Spectrometry* 30, no. 1: 250. <https://doi.org/10.1002/rcm.7433>.
- Lai, J., Y. Zou, J. Zhang, and P. R. Peres-Neto. 2022. “Generalizing Hierarchical and Variation Partitioning in Multiple Regression and Canonical Analyses Using the Rdacca.Hp R Package.” *Methods in Ecology and Evolution* 13, no. 4: 782–788. <https://doi.org/10.1111/2041-210X.13800>.
- Lapierre, J. F., F. Guillemette, M. Berggren, and P. A. del Giorgio. 2013. “Increases in Terrestrially Derived Carbon Stimulate Organic Carbon Processing and CO₂ Emissions in Boreal Aquatic Ecosystems.” *Nature Communications* 4: 2972. <https://doi.org/10.1038/ncomms3972>.

- LaRowe, D. E., and P. Van Cappellen. 2011. "Degradation of Natural Organic Matter: A Thermodynamic Analysis." *Geochimica et Cosmochimica Acta* 75, no. 8: 2030–2042. <https://doi.org/10.1016/j.gca.2011.01.020>.
- Lau, M. P., and P. Del Giorgio. 2020. "Reactivity, Fate and Functional Roles of Dissolved Organic Matter in Anoxic Inland Waters." *Biology Letters* 16, no. 2: 20190694. <https://doi.org/10.1098/rsbl.2019.0694>.
- Lehmann, J., C. M. Hansel, C. Kaiser, et al. 2020. "Persistence of Soil Organic Carbon Caused by Functional Complexity." *Nature Geoscience* 13, no. 8: 529–534. <https://doi.org/10.1038/s41561-020-0612-3>.
- Li, J., B. Wang, M. Yang, et al. 2022. "Geographical Constraints on Chemodiversity of Sediment Dissolved Organic Matter in China's Coastal Wetlands." *Applied Geochemistry* 147: 105506. <https://doi.org/10.1016/j.apgeochem.2022.105506>.
- Li, S., L. Meng, C. Zhao, et al. 2023a. "Spatiotemporal Response of Dissolved Organic Matter Diversity to Natural and Anthropogenic Forces Along the Whole Mainstream of the Yangtze River." *Water Research* 234: 119812. <https://doi.org/10.1016/j.watres.2023.119812>.
- Li, W., N. Liu, J. Li, et al. 2023b. "Chemodiversity of Dissolved Organic Matter Is Governed by Microbial Biogeography in Inland Waters." *Environmental Science & Technology* 57, no. 20: 7753–7763. <https://doi.org/10.1021/acs.est.3c00896>.
- Luo, J., Q. Zhou, X. Hu, et al. 2022. "Lake Chemodiversity Driven by Natural and Anthropogenic Factors." *Environmental Science & Technology* 56, no. 9: 5910–5919. <https://doi.org/10.1021/acs.est.1c08148>.
- Lv, J., S. Zhang, S. Wang, L. Luo, D. Cao, and P. Christie. 2016. "Molecular-Scale Investigation With ESI-FT-ICR-MS on Fractionation of Dissolved Organic Matter Induced by Adsorption on Iron Oxyhydroxides." *Environmental Science & Technology* 50, no. 5: 2328–2336. <https://doi.org/10.1021/acs.est.5b04996>.
- Lynch, L. M., N. A. Sutfin, T. S. Fegler, C. M. Boot, T. P. Covino, and M. D. Wallenstein. 2019. "River Channel Connectivity Shifts Metabolite Composition and Dissolved Organic Matter Chemistry." *Nature Communications* 10, no. 1: 459. <https://doi.org/10.1038/s41467-019-08406-8>.
- Ma, K., Y. Li, W. Song, et al. 2024. "Disentangling Drivers of Mudflat Intertidal DOM Chemodiversity Using Ecological Models." *Nature Communications* 15, no. 1: 6620. <https://doi.org/10.1038/s41467-024-50841-9>.
- Meng, F., A. Hu, K. S. Jang, and J. Wang. 2024. "iDOM: Statistical Analysis of Dissolved Organic Matter Based on High-Resolution Mass Spectrometry." *mLife*. <https://doi.org/10.1002/mlf2.70002>.
- Moran, M. A., E. B. Kujawinski, A. Stubbins, et al. 2016. "Deciphering Ocean Carbon in a Changing World." *Proceedings of the National Academy of Sciences* 113, no. 12: 3143–3151. <https://doi.org/10.1073/pnas.1514645113>.
- Moras, S., U. R. Zellmer, E. Hiltunen, C. Grasset, and S. Sobek. 2024. "Predicting Methane Formation Rates of Freshwater Sediments in Different Biogeographic Regions." *Journal of Geophysical Research: Biogeosciences* 129: e2023JG007463. <https://doi.org/10.1029/2023JG007463>.
- Ning, D., M. Yuan, L. Wu, et al. 2020. "A Quantitative Framework Reveals Ecological Drivers of Grassland Microbial Community Assembly in Response to Warming." *Nature Communications* 11, no. 1: 4717. <https://doi.org/10.1038/s41467-020-18560-z>.
- Pearson, A. R., B. R. S. Fox, J. C. Hellstrom, et al. 2024. "Warming Drives Dissolved Organic Carbon Export From Pristine Alpine Soils." *Nature Communications* 15, no. 1: 3522. <https://doi.org/10.1038/s41467-024-47706-6>.
- R Core Team. 2023. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.
- Roth, V.-N., M. Lange, C. Simon, et al. 2019. "Persistence of Dissolved Organic Matter Explained by Molecular Changes during its Passage through Soil." *Nature Geoscience* 12, no. 9: 755–761. <https://doi.org/10.1038/s41561-019-0417-4>.
- Ryan, A. K., A. V. Garayburu-Caruso, C. B. Crump, et al. 2024. "Riverine Dissolved Organic Matter Transformations Increase with Watershed Area, Water Residence Time, and Damköhler Numbers in Nested Watersheds." *Biogeochemistry* 167: 1203–1224. <https://doi.org/10.1007/s10533-024-01169-5>.
- Stegen, J. C., S. J. Fansler, M. M. Tfaily, et al. 2022. "Organic Matter Transformations Are Disconnected between Surface Water and the Hyporheic Zone." *Biogeosciences* 19, no. 12: 3099–3110. <https://doi.org/10.5194/bg-19-3099-2022>.
- Tanentzap, A. J., and J. A. Fonvielle. 2024. "Chemodiversity in Freshwater Health." *Science* 383, no. 6690: 1412–1414. <https://doi.org/10.1126/science.adg8658>.
- Taube, R., L. Ganzert, H.-P. Grossart, G. Gleixner, and K. Premke. 2018. "Organic Matter Quality Structures Benthic Fatty Acid Patterns and the Abundance of Fungi and Bacteria in Temperate Lakes." *Science of the Total Environment* 610–611: 469–481. <https://doi.org/10.1016/j.scitotenv.2017.07.256>.
- Tolic, N., Y. Liu, A. Liyu, et al. 2017. "Formularity: Software for Automated Formula Assignment of Natural and Other Organic Matter From Ultrahigh-Resolution Mass Spectra." *Analytical Chemistry* 89, no. 23: 12659–12665. <https://doi.org/10.1021/acs.analchem.7b03318>.
- Valle, J., M. Harir, M. Gonsior, et al. 2020. "Molecular Differences between Water Column and Sediment Pore Water SPE-DOM in Ten Swedish Boreal Lakes." *Water Research* 170: 115320. <https://doi.org/10.1016/j.watres.2019.115320>.
- Wen, S., A. Hu, S. Jiang, et al. 2024. "Temperature Sensitivity of Organic Carbon Decomposition in Lake Sediments Is Mediated by Chemodiversity." *Global Change Biology* 30, no. 2: e17158. <https://doi.org/10.1111/gcb.17158>.

- Wen, T., P. Xie, S. Yang, et al. 2022. “ggClusterNet: An R Package for Microbiome Network Analysis and Modularity-Based Multiple Network Layouts.” *iMeta* 1, no. 3: e32. <https://doi.org/10.1002/imt2.32>.
- Wetzel, R. G., and M. Søndergaard. 1998. The Structuring Role of Submerged Macrophytes in Lakes, edited by E. Jeppesen, M. Søndergaard, M. Søndergaard, and K. Christoffersen, 133–148. New York, NY: Springer New York.
- Zhang, T., H. Ma, Z. Hong, et al. 2022. “Photo-Reactivity and Photo-Transformation of Algal Dissolved Organic Matter Unraveled by Optical Spectroscopy and High-Resolution Mass Spectrometry Analysis.” *Environmental Science &*

Technology 56, no. 18: 13439–13448. <https://doi.org/10.1021/acs.est.2c03524>.

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Submitted 22 July 2024
Revised 24 January 2025
Accepted 16 February 2025