Water Research 308 (2027) 126771

Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres

Divergent molecular transformation networks converge on enhanced halogenated DOM formation in lakes under anthropogenic pressure: Insights across China and the United States

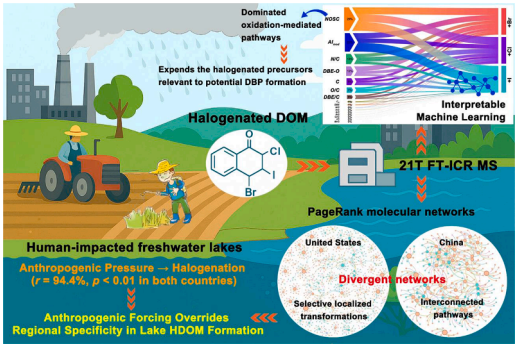
Chao Zhang [a,b,c], Junyu Zhu [a,b], Wenjie Mai [a,b], Yue Xie [a,b], Huan Chen [d], Harald Horn [c], Haijun He [e], Zhe Zhang [f,*], Mingzhi Huang [a,b,*], Fengchang Wu [g,*]

[a] Guangdong Provincial Engineering Research Center of Intelligent Low-carbon Pollution Prevention and Digital Technology & Guangdong Provincial Key Laboratory of Chemical Pollution and Environmental Safety & MOE Key Laboratory of Theoretical Chemistry of Environment, School of Environment, South China Normal University, Guangzhou, 510006, PR China
[b] SCNU (NAN'AN) Green and Low-Carbon Innovation Center & Nan'an SCNU Institute of Green and Low-Carbon Research, South China Normal University, Quanzhou, 362300, PR China
[c] Engler-Bunte-Institut, Water Chemistry and Water Technology, Karlsruhe Institute of Technology, Engler-Bunte-Ring 9, Karlsruhe, 76131, Germany
[d] National High Magnetic Field Laboratory, Florida State University, Tallahassee, FL, USA
[e] Guangzhou Marine Geological Survey, Guangzhou, 511458, PR China
[f] Department of Chemical and Environmental Engineering (ChEE), University of Cincinnati, Cincinnati, OH, 45221-0012, USA
[g] State Key Laboratory of Environmental Criteria and Risk Assessment, Chinese Research Academy of Environmental Sciences, Beijing, 100012, China

HIGHLIGHTS

- Anthropogenic pressure consistently enhances DOM halogenation and carbon cycling.
- Dominated oxidative halogenation expands DBP precursors beyond water treatment.
- U.S. lakes show localized HDOM networks, but Chinese lakes are more interconnected.
- Interpretable ML and molecular networks reveal mechanisms of HDOM transformation.
- DOM structure (aromaticity, aliphaticity, unsaturation) regulates halogenation reactivity.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Halogenated dissolved organic matter
Interpretable machine learning
Molecular transformation networks
Disinfection byproducts
Freshwater lakes

ABSTRACT

Halogenated dissolved organic matter (HDOM) is considered one of the most toxic fractions of DOM, as it potentially contributes to the formation of disinfection byproducts, which are documented to pose serious risks to water safety. Previous studies mainly focused on bulk DOM transformation and characterized its molecular diversity, while less attention has been given to the formation mechanisms and environmental drivers of HDOM that plays a key role in toxicity outcomes, particularly in lake systems relied upon as human water sources. Here, we elucidated molecular HDOM transformation dynamics in selected lakes in China and the U.S. using FT-ICR MS, interpretable machine learning, and PageRank molecular transformation networks. Results suggest that

* Corresponding authors.
E-mail addresses: zhang2z5@mail.uc.edu (Z. Zhang), mingzhi.huang@m.scnu.edu.cn (M. Huang), wufengchang@vip.skleg.cn (F. Wu).

https://doi.org/10.1016/j.watres.2026.126771
Received 9 March 2026; Received in revised form 19 July 2026; Accepted 20 August 2026
Available online 21 August 2026
0043-1354/© 2026 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

DOM halogenation in selected lakes from both countries is consistently dominated by oxidation-mediated pathways, enhancing the pool of halogenated precursors relevant to potential disinfection byproducts formation. Molecular networks reveal that dehalogenation and halogenation in U.S. lakes exhibit selective localized transformations, while Chinese lakes follow interconnected pathways, revealing region-specific biogeochemical behaviors of HDOM under anthropogenic pressures and molecular diversity. However, despite divergent molecular networks and pathways, lakes subjected to inevitably intense anthropogenic pressures in both countries exhibit a consistent tendency toward enhanced halogenation (increased toxicity) and accelerated carbon cycling. DOM structure (aromaticity, aliphaticity, and unsaturation) strongly regulates halogenation reactivity in lakes. Specifically, iodination preferentially targets electron-rich aromatic functional groups, whereas chlorination and bromination exhibit broader reactivity across both aromatic and aliphatic structures. This study provides cross-regional insights into how anthropogenic pressures regulate molecular pathways underlying HDOM transformation.

1. Introduction

Halogenated dissolved organic matter (HDOM), characterized by their covalent bonding with halogen atoms (i.e., Cl, Br, and I), poses a formidable threat to aquatic ecosystems and ultimately, human health, exhibiting pronounced toxicity (cellular, genetic, endocrine-disrupting, neurological), environmental persistence, and bioaccumulative properties (Li et al., 2023; Feng et al., 2025; Yang et al., 2026). The presence of HDOM in natural freshwater systems, particularly lakes that are subjected to intense human activities, leads to severe ecological imbalances, disruption of food web, and adverse health effects in exposed populations (Li et al., 2023; Wang et al., 2025). HDOM represents a critical reservoir of halogen-bearing functional groups within natural waters, and its presence substantially increases the potential for the formation of disinfection byproducts (DBPs) during subsequent chlorination or other oxidative disinfection processes, because these pre-halogenated and highly reactive molecular structures can serve as key precursors that facilitate or amplify DBP generation in drinking water treatment systems (Farinelli et al., 2022; Xiao et al., 2024). Therefore, elucidating the HDOM transformation processes within complex freshwater lake ecosystems is of paramount importance for ensuring the potable water safety.

Although HDOM poses significant threats to aquatic ecosystems and human health, the scientific understanding of its environmental behavior, particularly its transformation processes in real natural waters, remains limited. First, there is a lack of representative studies on real environmental samples. Most investigations into the molecular characteristics of HDOM rely on laboratory halogenation reactions using standardized natural organic matter (NOM) provided by the International Humic Substances Society (IHSS), such as Suwannee River NOM (Li et al., 2023). While these reference materials are chemically stable and analytically convenient, they cannot fully represent the compositional complexity of DOM in natural waters, which is shaped by diverse biogeochemical processes and anthropogenic influence (Wu et al., 2023; Ruan et al., 2021; Mitch et al., 2023). Considering that variations in DOM composition can substantially alter halogenation pathways and the HDOM formation spectrum, research based solely on reference materials may misestimate the actual HDOM formation potential in environmental systems. Second, molecular-level understanding of HDOM remains highly incomplete. The formation mechanisms of HDOM are intricately complex, and much of its molecular composition is still unknown. Although >800 HDOM species have been identified, most of them remain structurally unresolved (Čulin et al., 2015; Richardson et al., 2020). Consequently, the types, structures, toxicities, and environmental fates of HDOM molecules remain poorly characterized, leaving a substantial knowledge gap. Third, the complexity of environmental drivers further complicates HDOM research. The sources and characteristics of DOM, and consequently its halogenation reactivity, are governed by multiple interacting factors (Ji et al., 2024). Such spatiotemporal variability means that the HDOM transformation pathways and rates can differ dramatically across regions. This heterogeneity poses major challenges for systematic studies and interregional comparisons. Therefore, conducting cross-regional studies that encompass diverse anthropogenic contexts is essential to elucidate the universal and region-specific mechanisms of HDOM transformation. However, such efforts face substantial challenges in sample collection, data interpretation, and cross-comparison, underscoring the urgent need to integrate advanced analytical techniques with multidimensional data modeling approaches.

China and the United States, two of the largest economies on a planetary scale, face similar challenges in managing HDOM in their anthropogenically influenced freshwater lakes (Wang et al., 2024). However, the underlying mechanisms governing HDOM transformation may vary significantly between these two regions due to differences in environmental characteristics, climate, hydrology, and anthropogenic activities. Meanwhile, Unlike rivers, lakes exhibit longer water residence times that permit progressive halogenation and accumulation of recalcitrant HDOM, function as long-term sinks for anthropogenic inputs leading to decadal buildup of halogenated precursors, and serve directly as drinking water sources, making HDOM transformation in lakes a more immediate and cumulative risk for disinfection byproduct formation (Vörösmarty et al., 2010; Munthali et al., 2022; Hu et al., 2024). To this end, Fourier Transform–Ion Cyclotron Resonance Mass Spectrometry (FT-ICR MS) was employed to conduct a comparative analysis of the HDOM at molecular level in anthropogenically influenced freshwater lakes of both China and the United States. FT–ICR MS excels in delineating the molecular HDOM composition across diverse scenarios, offering the highest resolution available (Lu et al., 2025). This advanced analytical technique allowed for insights into the intricate molecular transformations that occur within the HDOM under these conditions. Simultaneously, interpretable machine learning (IML) has emerged as a powerful tool for gaining insights into complex environmental processes, including HDOM transformations (Zhang et al., 2025). Unlike traditional black-box machine learning, IML techniques provide not only predictive accuracy but also transparency and interpretability, allowing to understand the underlying mechanisms of HDOM transformations (Zhang et al., 2025). Meanwhile, to further deconvolute complex interdependence of HDOM molecules within this transformation network, we proposed PageRank molecular network analysis to operate primarily on feature-property relationships of HDOM. PageRank algorithm, grounded in graph theory, transcends this by mapping actual reaction pathways between HDOM precursors and products, thereby reconstructing functional transformation networks (Nykl et al., 2014). Its key advantage over traditional degree centrality lies in its recursive evaluation: product importance is weighted by the importance of its precursors. This approach robustly identifies critically important HDOM molecules that serve as hubs within the reaction network, even if their direct reaction count is low, while effectively mitigating misjudgment risks from incidental reactions. The synergy of IML and PageRank thus offers a multi-faceted dissection of HDOM dynamics from feature importance to network-driven transformation pathways.

In the present study, we conducted an in-depth investigation, leveraging FT–ICR MS at the molecular level, to discern both resemblances and differences in regional DOM distribution patterns in

selected human-impacted lakes between the U.S. and China. A comprehensive examination of the halogenation and dehalogenation reactivity of molecular formulas mediated by natural conditions in human-impacted lakes in the study areas of both the U.S. and China was undertaken in this study, as well as transformation processes of associated structural categories. Additionally, rigorous application of interpretable machine learning was undertaken to evaluate the impacts of molecular characteristics on halogenation and dehalogenation reactivity of individual DOM molecular classes, emphasizing distinctive features between the U.S. and China. Furthermore, an in-depth exploration of the plausible halogenation and dehalogenation pathways of DOM was conducted, leveraging PageRank molecular network analysis to scrutinize precursor-product relationships. This study provides molecular insights into molecular transformation dynamics of toxic HDOM in exchangeable organic carbon reservoirs of human-impacted lakes, elucidating their potential impacts on drinking water security and aquatic ecosystems.

2. Materials and methods

2.1. Site description and sample collection

Surface lake water samples were meticulously collected from locations within *Lake Erie, Lake St. Marys*, and *Lake Harsha* in the U.S., and from *Lake Liuhua, Lake Songshan*, and *Lake Wanlyu* in China (Fig. 1), during summer (July) and the following winter (January) to investigate seasonal variations. Based on extensive field investigations, these selected lakes are influenced by anthropogenic activities to varying extents, with potential sources contributing to the presence of halogenated contaminants, encompassing application of fertilizers and pesticides, domestic water discharge, the rainwater runoff containing dissolved halogenides from residential areas and roadways, and halogenated disinfection byproducts from water treatment. The information of watershed characteristics, water chemistry, and lake morphology of these selected lakes is presented in Table S9.

2.2. Sample extraction and analytical methods

The water samples underwent filtration via 0.22-μm membrane filters, and subsequently, its pH level was calibrated to 2.0 utilizing a 1 mol/L HCl solution, ensuring a precise adjustment. Solid-phase extraction utilizing Agilent Bond Elut PPL (500 mg per 6 mL) was employed to effectively desalinate and concentrate DOM present in the lake water samples. Following this extraction process, the concentrated DOM samples were subjected to analysis using Fourier Transform-Ion Cyclotron Resonance Mass Spectrometry (FT–ICR MS) equipped with negative electrospray ionization (ESI) sources. To ascertain the molecular identities of the DOM, peaks exhibiting a signal-to-noise ratio (S/N) of 4 or higher were identified using the rigorous criteria with a measurement m/z range of 100–1000 and elemental combinations of $^{12}C_{1-60}$, $^{1}H_{1-120}$, $^{16}O_{1-50}$, $^{14}N_{0-5}$, $^{32}S_{0-3}$, $^{31}P_{0-2}$, $^{35}Cl_{0-3}$, $^{37}Cl_{0-1}$, $^{79}Br_{0-2}$, $^{81}Br_{0-1}$, and $^{127}I_{0-2}$. The molecular formulas were determined based on permissible elemental combinations, ensuring that errors between the experimentally measured mass and the theoretically calculated mass for each identified chemical formula remained below the stringent threshold of 0.5 ppm. For sample preparation, detailed procedures of FT-ICR MS analysis and parameter calculation, please refer to Text S1.

2.3. Machine learning algorithms and interpretation strategy

In this study, a machine learning model was employed, leveraging molecular features derived from FT-ICR MS as descriptors, to analyze and predict the DOM properties pertinent to halogenation and dehalogenation reactivity. Specifically, *Transformer* model is harnessed to characterize the variety and intensity of halogenation and dehalogenation reactions that may occur during precursor-product pairing reactions during proliferation periods, thereby defining their respective reactivities. The descriptors utilized within the *Transformer* model consist of 15 molecular characteristics derived from FT-ICR MS analyses (Table S8). Interpretable Machine Learning is implemented to elucidate how the *Transformer* model predict the halogenation and dehalogenation reactivity. To interpret this process, an explanatory strategy incorporating feature importance and dependence plots is employed, utilizing SHapley Additive exPlanations (SHAP) method. The SHAP explanations harness Shapley values as proxies for the individual contributions of molecular property descriptors to halogenation and dehalogenation reactivity analyzed by the *Transformer* model. For each formula in the samples, the Shapley value of each descriptor (molecular property) is calculated. Subsequently, the absolute average of SHAP values across all formulas descriptors is computed, providing a metric for assessing the average contribution of each molecular property. All methods are implemented using the SHAP package in Python. For comprehensive understanding of objectives and methodology of the

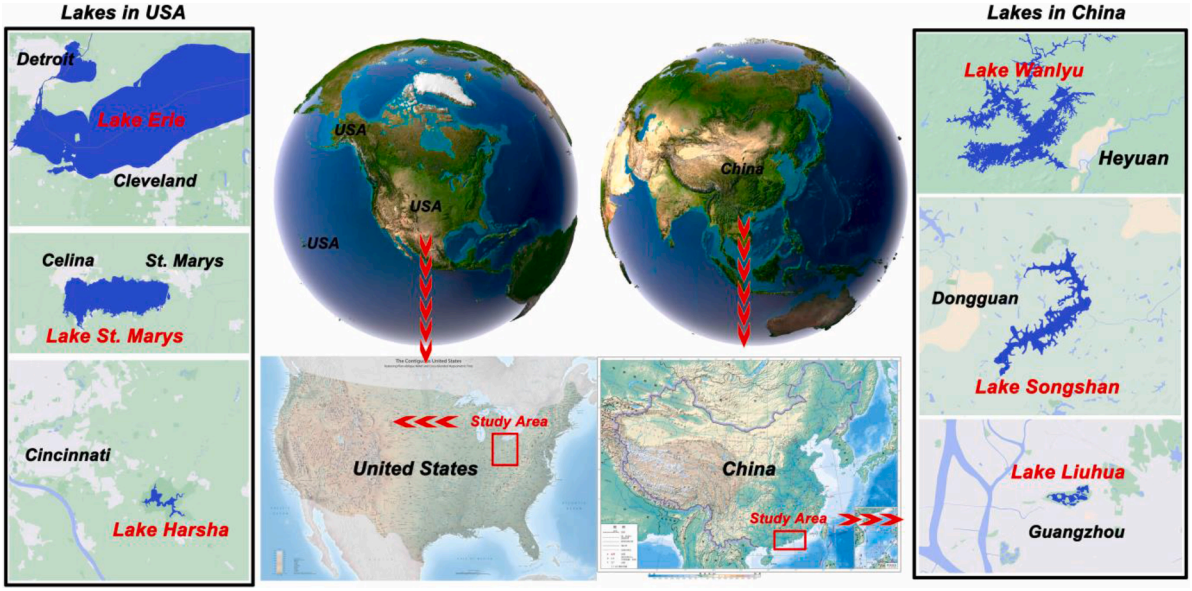


Fig. 1. Map of study areas and sampling lakes in the United States and China.

Interpretable Machine Learning analysis, please refer to the Text S2.

2.4. PageRank molecular network analysis

To clearly elucidate the potential key transformation processes of HDOM, this study constructs a molecular reaction network by applying graph theory to reactants, products, and their reactions, utilizing the PageRank algorithm to calculate node importance. It is based on the concept of link analysis, determining relative importance of nodes by analyzing the link structure between them (Nykl et al., 2014). Unlike traditional methods that use node degree to evaluate node importance, PageRank also considers the importance of incoming nodes as a factor in this evaluation. The PageRank value of each node can be iteratively calculated using the following formula:

$$PR(p) = (1-d) + d\sum_{q\in M(p)}\frac{PR(q)}{L(q)} \quad (1)$$

Where $PR(p)$ represents the PageRank value of node p, d denotes the damping factor, $M(p)$ is the set of nodes that point to node p, and $L(q)$ is the number of outbound links from node q. Compared to traditional methods that calculate product importance solely based on degree centrality, the PageRank algorithm provides significant advantages. Degree centrality considers only the number of potential precursor HDOM associated with a product, neglecting intrinsic importance of these precursors. In contrast, PageRank incorporates the importance of each precursor as a weighted factor when calculating the product importance, offering a more precise evaluation of its actual position in the network. By accounting for the global structure of the reaction network, PageRank propagates critical the influence of HDOM across the entire network, effectively capturing the importance transmission among HDOM. This approach distinguishes key products that, despite participating in fewer reactions, are connected to these highly influential precursors. Moreover, PageRank exhibits greater robustness against noise or incidental reactions within the network, reducing the risk of misjudgment caused by reliance on reaction counts alone. Consequently, PageRank is well-suited for analyzing the importance of complex interactions among HDOM in reaction networks.

3. Results and discussion

3.1. HDOM dynamics in the selected lakes in China vs. the U.S.

FT-ICR MS were deployed to characterize the dynamics in the molecular composition of HDOM in selected human-impacted lakes between China and the U.S., as illustrated in Fig. 2. As evident from Fig. 2a and b, under seasonal variation in the selected Chinese lakes, Lignin/CRAM-like substances dominated (65.44%), followed by proteins (10.72%), lipids (6.21%), tannins (6.02%), aromatics (3.21%), carbohydrates (2.11%), and unsaturated hydrocarbons (1.86%). In contrast, the U.S. lakes exhibited a lower lipid proportion (3.67%) with intermediate values between tannins (7.05%) and aromatics (2.93%) at this stage. It is noteworthy that, in both China and the U.S. lakes, the proportion of various HDOMs remains relatively stable during the course of seasonal variation, with variations within ±2%. Chinese lakes demonstrated divergent trends in the HDOM composition under seasonal variation: Lignin/CRAM-like substances (-1.31%), unsaturated hydrocarbons (-0.24%), and aromatics (-0.23%) generally decreased, while others increased (Fig. 2a). Conversely, the U.S. lakes exhibited elevated proportions of Lignin/CRAM-like compounds (+0.7%), aromatics (+1.22%), and tannins (+1.04%) with declines in remaining constituents (Fig. 2b). This contrast reflects fundamental disparities in human-impacted lakes between 2 countries, highlighting differential HDOM transformation pathways.

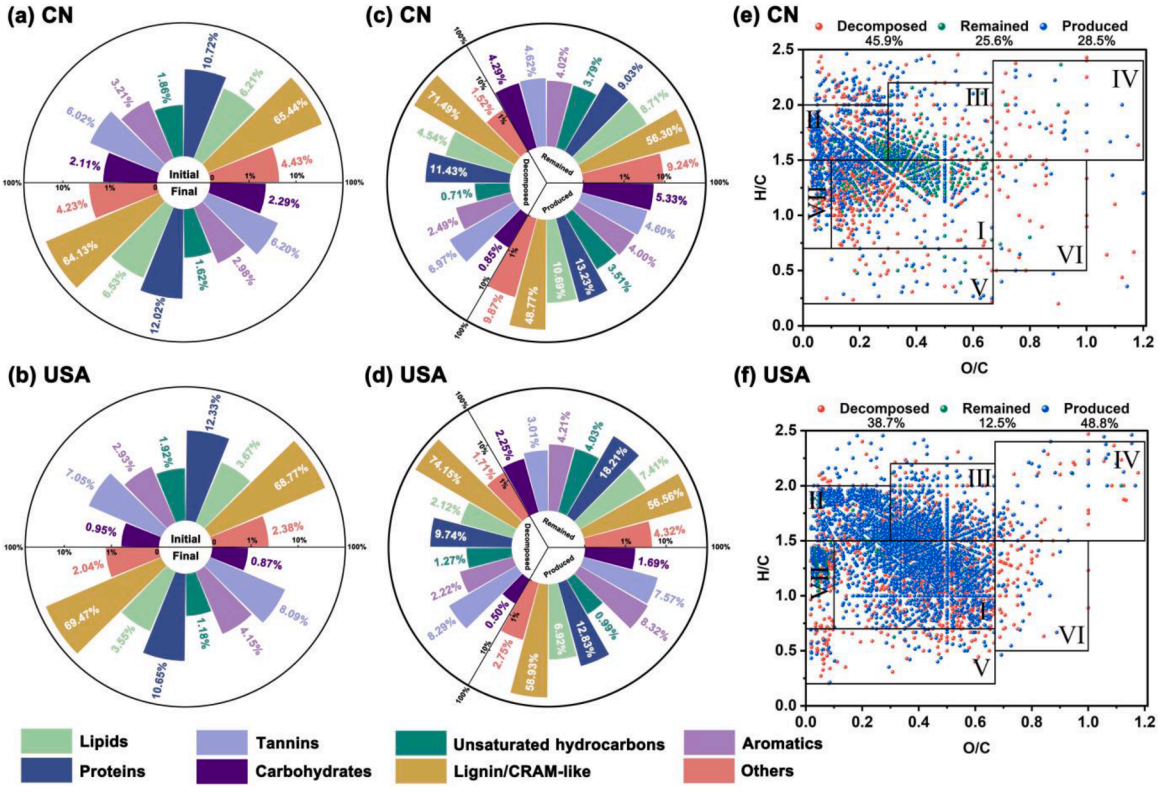


Fig. 2. Dynamics in the molecular composition of HDOM during the course of seasonal variation in human-impacted lakes between China and the United States. (a, b) Proportions of different types of HDOM substances across seasonal transitions; (c, d) The percentages of HDOM substances remaining, decomposed, and produced; (e, f) Van Krevelen diagrams of HDOM remaining, decomposed, and produced across seasonal transitions. **I**, Lignin/CRAM-like; **II**, Lipids; **III**, Proteins; **IV**, Unsaturated hydrocarbons; **V**, Aromatics; **VI**, Tannins; and **VII**, Carbohydrates.

The significant proportion and dynamic variations of Lignin/CRAM-like substances (a characteristic terrestrial humic substance marker) in Chinese and U.S. lakes corroborate the dominant terrestrial halogenated organic matter input regime driven by intensive agricultural activities and urbanization, demonstrating terrestrial HDOM fluxes amplified by anthropogenic watershed modification (Hedges et al., 1997). Furthermore, lakes subjected to inevitably intensive anthropogenic activities, such as Lake Liuhua in China and Lake Erie in the U.S. (Table S2), exhibit features indicative of an accelerated carbon cycle, as evidenced by decreased molecular weights and elevated H/C ratios (Tables S5 and S6). A growing body of literature using FT-ICR MS has established robust relationships between these molecular parameters and carbon mineralization dynamics: (1) H/C ratio as a proxy for DOM biodegradability: High H/C ratios are characteristic of more saturated, aliphatic, and hydrogen-rich molecular structures (e.g., lipids, aliphatic/proteins). Multiple FT-ICR MS-based studies have demonstrated that DOM with higher H/C is preferentially removed by microorganisms during degradation, whereas DOM with low H/C (e.g., lignin/CRAM-like structures, tannins) exhibits recalcitrance (Huang et al., 2026; Yu et al., 2024; Chen et al., 2023). Therefore, an increase in H/C in bulk DOM indicates a relative enrichment of microbially labile components, implying enhanced microbial metabolic activity and accelerated DOM turnover. (2) Molecular weight as a proxy for bioavailability: Low-molecular-weight DOM (especially below ~1000 Da) is generally more bioavailable, as smaller molecules can be directly taken up by microbial cells without extracellular enzymatic cleavage. Conversely, high-molecular-weight DOM tends to be more aromatic and recalcitrant. Thus, a decrease in molecular weight suggests microbial fragmentation and processing of DOM, which directly corresponds to faster carbon mineralization (Davenport et al., 2025). (3) Cumulative interpretation: When both trends occur simultaneously, the most parsimonious interpretation of lower molecular weight and higher H/C is that the DOM pool has shifted toward more aliphatic, less aromatic, and more microbially accessible compounds. This shift is consistent with enhanced microbial metabolic activity and shortened carbon turnover times, i.e., an accelerated carbon cycle.

This indicates that anthropogenic disturbances can profoundly reshape the source–sink dynamics of lake carbon cycling on a global scale (Liu et al., 2022). This process has, to varying degrees, elevated average levels of carbon cycle acceleration in the selected lakes in both countries, as reflected by observed reductions in aromatics, unsaturated hydrocarbons, and lignin/CRAM-like substances (Fig. 2a and b) (Raymond et al., 2013; Catalán et al., 2016). This may be related to significant carbon reduction pressures faced by urban areas with intensive anthropogenic activities, e.g., infrastructure development and demographic change, supporting theoretical framework of anthropogenic disturbances reshaping greenhouse gas source-sink dynamics in human-impacted lakes (Catalán et al., 2016).

Fig. 2c and d illustrated the percentages of HDOM remained, decomposed, and produced across seasonal transitions in the human-impacted lakes between China and the U.S., respectively. Halogenated lignin/CRAM-like and protein compounds exhibited marked activity, demonstrating their significant decomposition and production in lakes of the U.S. and China. Based on the Van Krevelen diagrams of HDOM remained, decomposed, and produced across seasonal transitions presented in Fig. 2e and f, the decomposition and production effects of HDOM in U.S. lakes are stronger (48.8% produced and 38.7% decomposed), and the proportion of the remained part is smaller (12.5% remained), compared to Chinese lakes (28.5% produced, 45.9% decomposed, and 25.6% remained). This indicates that HDOM in U.S. lakes exhibits higher reactivity. Meanwhile, significant overlap was found when comparing the decomposition and production of HDOM molecules during transformation process, with the lakes in China primarily concentrated in regions I, II, and VII (Lignin/CRAM-like substances, Lipids, and Carbohydrates), and the U.S. lakes in regions I, II, III, and VII (Lignin/CRAM-like substances, Lipids, Proteins, and Carbohydrates). These results indicate the bidirectionality of the transformation processes that molecules of the identical type may undergo processes of decomposition and reproduction. Meanwhile, it also reveals that the interplay between terrestrial inputs and microbial metabolism serves as an important driving force for the HDOM transformation in lakes of both the U.S. and China. Furthermore, carbohydrates in region VII may originate from microbial extracellular polymeric substances or plant polysaccharides, suggesting high biological activity (Tiwari et al., 2020; Mohammed et al., 2021; Benalaya et al., 2024). Protein (III) response signals in U.S. lakes are significantly pronounced, reflecting its stronger characteristics of endogenous productivity or eutrophication (Rogers et al., 2021).

3.2. Halogenation and dehalogenation pathways of DOM

Fig. 3 illustrates the matched pairs of HDOM compounds (i.e., Cl-DOM, Br-DOM, and I-DOM) identified by mass difference analysis through transformation pathways during seasonal variation in selected human-impacted lakes between China and the U.S., and percentages of halogenation and dehalogenation pathways of DOM are presented in Table S1. The area and proportion of regions potentially affected by human activities within catchment areas of the lakes in China and the U.S. are listed in Table S2 (The data is sourced from the open-source database of Google, ESA WorldCover 10 m v100 dataset). Table S3 outlines the potential transformation pathways and corresponding mass differences of HDOM in water, while Table S4 details the possible structures of products derived from HDOM.

The most pronounced halogenation and dehalogenation effects were observed in Lake Erie in the U.S. and Lake Wanlyu in China, respectively, presented in Fig. 3. Lake Wanlyu is a crucial water source in the Guangdong-Hong Kong-Macao Greater Bay Area in China, one of the largest bay areas in the world (Wanlv Lake, 2023). Its water quality is crucial for maintaining the ecological balance and sustaining the livelihoods of surrounding areas (Wanlv Lake, 2023), while the region affected by anthropogenic activities only accounts for 1.78% (Table S2). The proportions of debromination, dechlorination, and deiodination reactions in Lake Wanlyu are as high as 96.8%, 63.8%, and 62.5%, respectively. In contrast, due to its central location of Guangzhou, the largest city in the southern China, Lake Liuhua has a total proportion of anthropogenically influenced land types reaching 79.4%, with the proportions of the three halogenation reactions being 75.8%, 96.1%, and 59.5%, respectively. In the case of Songshan Lake in China, the proportion of areas affected by anthropogenic activities has decreased to 46.0%, resulting in a relatively balanced distribution between halogenation and dehalogenation reactions, with chlorination at 58.5%, dechlorination at 41.5%, iodination at 53.2%, and deiodination at 46.8%. In the U.S., due to the proximity of Lake Erie to the urban agglomeration of the Great Lakes and the fact that 58.8% of its approximate catchment area of the sampling point is influenced by urbanization and agriculture, the proportions of its three of halogenation reactions are remarkably high, reaching 78.6%, 83.5%, and 92.7%, respectively. A similar phenomenon has been observed in Lake St. Marys and Lake Harsha in the U.S.

It was found that, the intensity ratio of halogenation under seasonal variation in various lakes across China and the U.S. shows a significant and strong positive correlation with the proportion of areas experiencing intense human activities (agricultural and urbanized areas; $r = 94.4\%$, $p < 0.01$). This clearly confirms that anthropogenic pressure is an essential driving factor in the halogenation process. The commonalities observed in lakes across China and the U.S. suggest that the association between halogenation and human activities may be universal. This indicates the widespread risk of water halogenation in the context of global urbanization and agricultural expansion, consequently threatening drinking water safety and aquatic ecosystems (Shi et al., 2024). This may be related to the contaminants introduced by human activities, e.g., pesticides, fertilizers, industrial chemicals, and domestic sewage

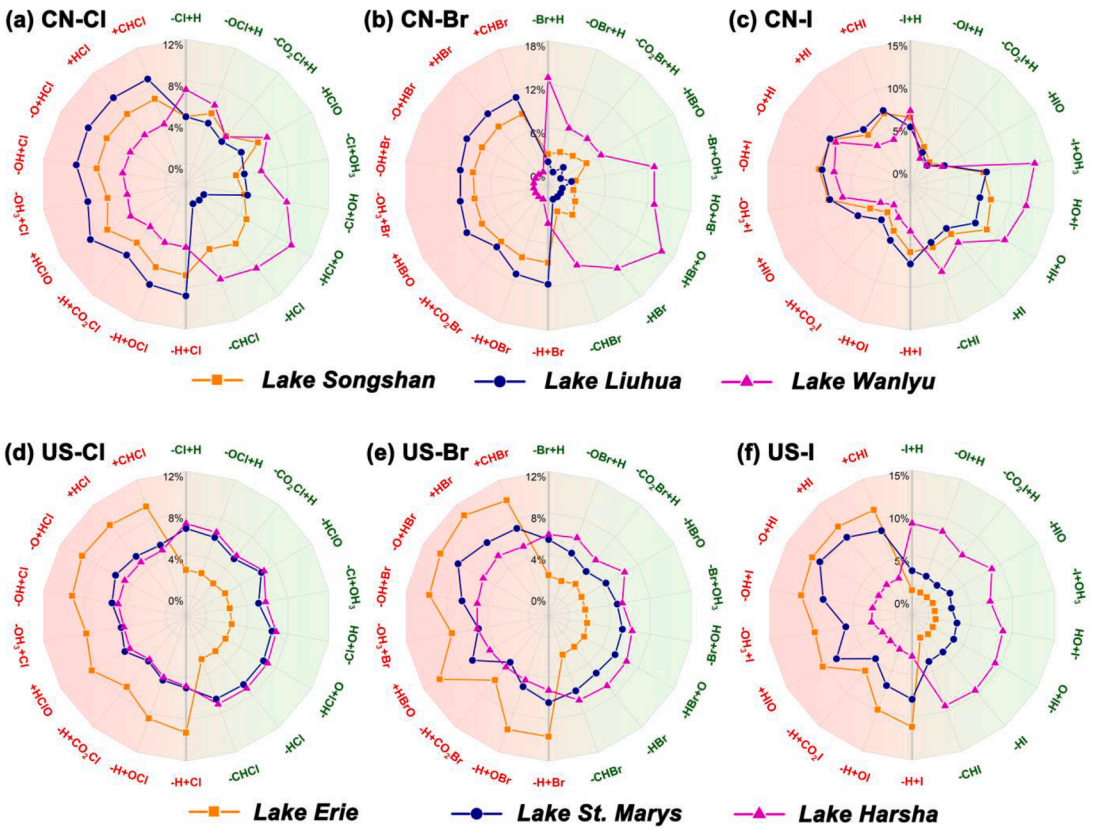


Fig. 3. Halogenation and dehalogenation pathways of DOM in lakes in China and the U.S.: (a) Chlorination and dechlorination in China; (b) Bromination and debromination in China; (c) Iodization and deiodination in China; (d) Chlorination and dechlorination reactions in the U.S.; (e) Bromination and debromination reactions in the U.S.; and (f) Iodization and deiodination reactions in the U.S. In each small diagram, the red section on the left represents halogenation reactions, while the green section on the right represents dehalogenation reaction.

(Shi et al., 2024). These contaminants can affect water bodies by providing halogen precursors (Zhao et al., 2022).

Halogens and their derivatives in water bodies exhibit significant biotoxic effects on humans and aquatic communities (Li et al., 2023; Feng et al., 2025). Notably, the halogenation reaction between halogens and DOM produces HDOM, which presents multidimensional ecological risks: (1) Amplified typical toxic effects (carcinogenicity, teratogenicity, and mutagenicity): the strong electronegativity of halogen atoms induce DNA alkylation and inhibit the function of key enzyme systems (Kodavanti et al., 2023); (2) Increased environmental persistence: high bond energy of C-X bonds results in the recalcitrant nature of HDOM such as the resistance to biodegradation (e.g., hydrolytic enzyme tolerance) and photolytic stability (Feng et al., 2023); (3) Enhanced bioaccumulation: the strong hydrophobicity of halogen substituents (increased log K_{ow}) promotes transmembrane penetration and biomagnification of HDOM in lipid tissues (Feng et al., 2025). Based on these mechanisms, the environmental risks of halogenation far exceed those of dehalogenation. Although dehalogenation reactions may release inorganic halide ions, the acute toxicity and bioaccumulation are significantly lower than those of HDOM (Feng et al., 2023). Only intermediate organics produced during dehalogenation exhibit secondary toxicity, whereas the combined pollution effects of HDOM lead to synergistic toxicity amplification (Kodavanti et al., 2023).

3.3. The impact of microorganisms on HDOM transformation

Cluster analysis based on 16S rRNA gene sequencing in this study established that, alongside Cyanobacteria, the primary microbial communities at the phylum level under seasonal variation encompass Proteobacteria, Actinobacteria, Verrucomicrobiota, Bacteroidota, Armatimonadota, Acidobacteriota, Patescibacteria, Sva0485 and Nitrospirota between lakes in China and the U.S. (Fig. 4). Based on halogenation and dehalogenation pathways, the relative reaction abundance was systematically quantified and Mantel tests ware further performed to identify critical reactions and their associated microbial taxa driving the biogeochemical cycling of HDOM (Fig. 4a). This analytical framework was implemented to elucidate mechanistic linkages between reaction prevalence patterns and microbial phylogenetic assemblages.

The quantification results of the relative abundance of different reaction types (Fig. 4a) reveal a clear trend: Cl-related reactions, encompassing both chlorination and dechlorination, exhibit significantly higher prevalence compared to Br- and I-related reactions ($p < 0.01$). This pattern is consistently observed across the datasets of both China and the U.S., suggesting a shared chemical or environmental preference for Cl-based transformations. Such dominance may be attributed to the higher reactivity, availability, or industrial usage of Cl compounds, aligning with previous finding or theoretical expectation in halogen chemistry (Ilunga et al., 2021). Simultaneously, both halogenation and dehalogenation reactions exhibit greater intensities for lakes located in U.S. compared to those in China. Among the lakes studied, Lake Wanlyu demonstrates the lowest relative intensity across all reaction categories. This observation provides further evidence supporting the point that anthropogenic activities exert a substantial influence on the transformation processes of HDOM under seasonal variation.

The microorganisms of *Cyanobacteria, Proteobacteria, Bacteroidota, Verrucomicrobiota*, and *Actinobacteriota* exhibit significant correlations with various categories of dehalogenation reactions ($p < 0.05$), indicating their metabolic versatility in HDOM transformation. Concurrently, as illustrated in the right panel of Fig. 4a, the members affiliated

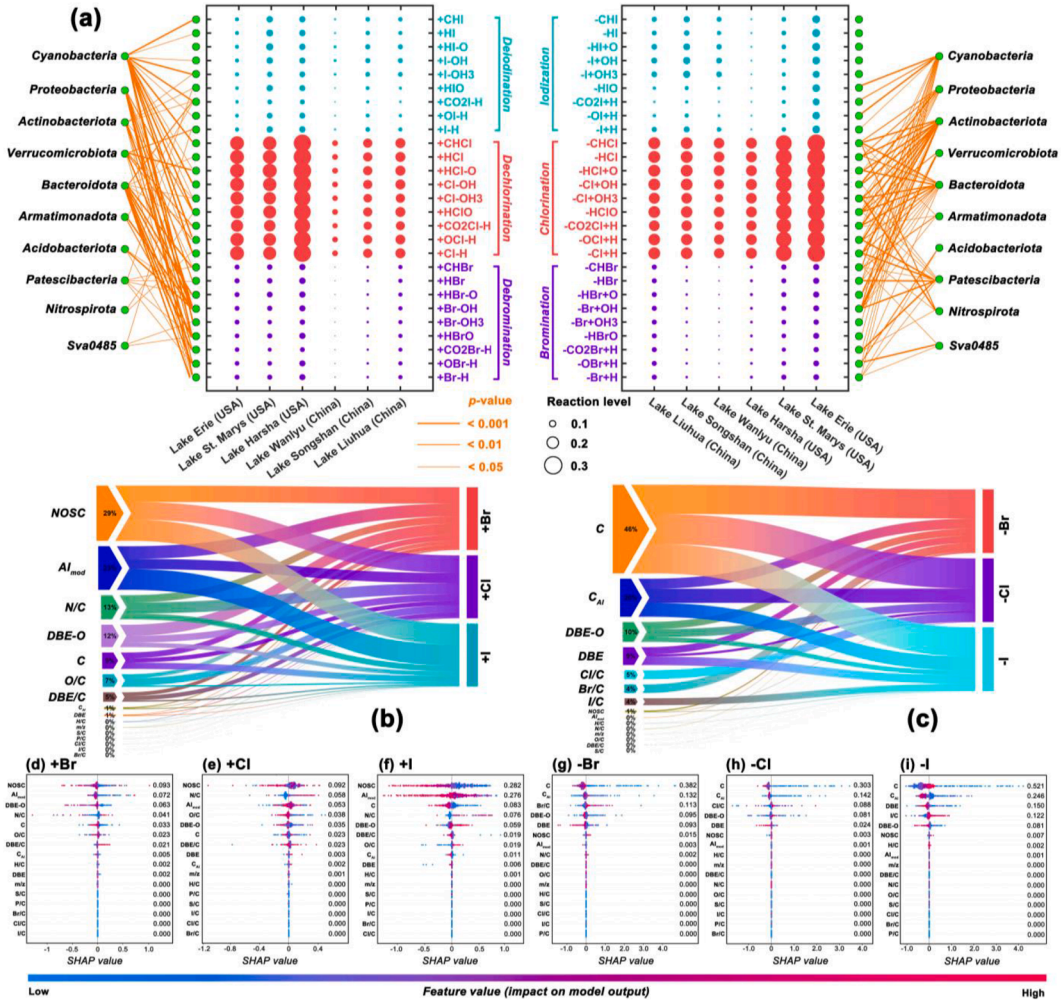


Fig. 4. (a) Prevalent microbially-driven processes associated with transformation pathways (halogenation and dehalogenation) of halogenated dissolved organic matter (HDOM) and their relationships to microbial taxonomic groups. Sankey diagram for the normalized SHAP metrics illustrating the influence of individual features on formula reactivity prediction, derived from (b) halogenation and (c) dehalogenation pathways, along with element-specific SHAP importance in (d) Bromination, (e) Chlorination, (f) Iodization, (g) Debromination, (h) Dechlorination, and (i) Deiodination. NOSC, nominal oxidation state of carbon; DBE, Double bond equivalence; AI_{mod}, modified aromaticity index; C, number of carbon atoms; C_{AI}, the number of aromatic carbon atoms; DBE-O, the oxygen-adjusted double bond equivalents. We emphasize that our analysis is based on single-step mass difference matching and does not infer a sequential process.

with *Cyanobacteria, Actinobacteriota, Bacteroidota*, and *Patescibacteria* demonstrate notable influence across multiple types of DOM halogenation reactions ($p < 0.05$). Importantly, among the 10 primary microbial communities identified as having the greatest impact on HDOM halogenation and dehalogenation processes, only *Cyanobacteria* show a highly significant association ($p < 0.01$) with specific iodination and deiodination reactions. These findings collectively underscore the substantial role of microorganisms in modulating both halogenation and dehalogenation dynamics of HDOM. In contrast, within the scope of sample validation conducted in this study, microorganisms belonging to *Acidobacteriota* and *Sva0485* exhibited a statistically significant influence ($p < 0.01$) specifically on the bromination and debromination of HDOM. Meanwhile, *Nitrospirota* showed a significant correlation ($p < 0.05$) exclusively with the chlorination and dechlorination processes of HDOM.

It is noteworthy that the core metabolic functions of certain microbial taxa show limited association with HDOM. For instance, the primary ecological role of *Nitrospirota* lies in its function as chemolithoautotrophic bacteria involved in second stage of nitrification, oxidizing nitrite to nitrate (Lawson et al., 2021; Daims et al., 2015). These microorganisms derive their energy and carbon predominantly from inorganic sources (e.g., nitrite/ammonia and CO_2), rather than utilizing complex organic carbon compounds such as DOM as their main energy or carbon source (Lawson et al., 2021). Although *Nitrospirota* do not exert a direct biochemical effect on the chlorination of HDOM, they may have an indirect influence on the extent and distribution of Cl-DOM products. This influence likely arises from their role in driving the nitrification process and contributing to biofilm formation, which in turn alters the aquatic chemical environment, particularly the nitrogen cycle, the speciation and concentration of chlorine (Zhou et al., 2024). Accordingly, this underscores the importance of investigating the interactions between microbial taxa and HDOM molecules to gain a more comprehensive understanding of the specific roles of microorganisms play in HDOM transformation processes. Meanwhile, the six-month sampling interval encompasses multiple seasonal changes. The observed molecular differences likely reflect the combined influence of various environmental drivers, including but not limited to microorganism dynamics.

3.4. Contributions of HDOM features to halogenation and dehalogenation reactions

To elucidate feature-specific propensity for halogenation and dehalogenation under seasonal variation, we employed interpretable

machine learning to quantify the contribution of HDOM characteristics to these transformation processes. Fig. 4b-i represents Sankey diagrams for normalized Shapley Additive exPlanations (SHAP) values illustrating influence of individual features on the formula reactivity, derived from halogenation and dehalogenation pathways. The primary objective of SHAP is to provide an interpretable, consistent, and game-theory-based quantification of feature contributions toward machine learning model predictions.

Fig. 4b and d-f show the contribution and its proportion of individual DOM features on halogenation reactivity. Interestingly, the most influential individual feature of DOM affecting the 3 halogenation reactions (+Cl, +Br, and +I) consistently converge to the nominal oxidation state of carbon (NOSC), which accounts for 29% of the total contribution. Following NOSC, the next most impactful descriptors are the modified aromaticity index (AI_{mod}, 23%), the nitrogen-to-carbon ratio (N/C, 13%), and the oxygen-adjusted double bond equivalents (DBE-O, 12%). More importantly, during the complex reaction processes of DOM, a consistent trend emerges in which lower NOSC values are more prone to promoting halogenation (Fig. 4d-f). This finding suggests that halogen incorporation into DOM predominantly occurs through oxidative halogenation mechanism in human-impacted lakes, in contrast to classical halogenation pathways such as addition reactions (e.g., halogen addition to alkenes or alkynes). These results reveal that HDOM formation is typically driven by oxidative processes that generate reactive halogen species capable of electrophilic substitution or addition to DOM, introducing halogenated functional groups and thereby expanding the pool of halogenated organic structures that can act as potential precursors of disinfection byproducts during subsequent oxidative disinfection (Ding et al., 2019). Additionally, the oxidative pathway occupies a central role in HDOM generation under seasonal variation, indicating that its influence is not confined to water treatment processes alone. This process not only leads to formation of toxic disinfection byproducts in freshwater lakes, but depletes low-NOSC DOM through halogenation, thereby suppressing essential anaerobic respiratory pathways, as low-NOSC DOM serves as a critical electron donor for denitrifying bacteria (Ren et al., 2022). Therefore, this mechanism is expected to exacerbate eutrophication and potentially propagate its impacts to ecosystems (Ren et al., 2022).

Similarly, interpretable SHAP analysis indicates that DOM with higher AI_{mod} tends to inhibit halogen addition reactions. This finding is consistent with the theoretical framework of DOM chemical reactivity: high-AI_{mod} DOM is typically characterized by the condensed aromatic structures, representing the stable conjugated systems that are resistant to microbial utilization under seasonal variation and less prone to halogen binding (Patel et al., 2020). In contrast, low-AI_{mod} aliphatic DOM exhibits higher reactivity and bioavailability. This also explains the observed relationships between N/C, DBE-O, and the halogenation behavior of DOM: a higher N/C ratio and a lower DBE-O value respectively indicate greater aliphatic character and lower unsaturation degrees, which are associated with enhanced bioavailability and reactivity (Hu et al., 2024). Interestingly, in contrast to chlorination and bromination, iodination is markedly more favorable toward high-DBE-O DOM. This finding suggests that the iodination preference for high-DBE-O DOM arises from the fact that iodine, being the weakest and most selective electrophile among halogens, targets the most electron-dense sites. High-DBE-O DOM provides such electron-rich aromatic ring sites bearing strong electron-donating substituents, thereby facilitating highly efficient and selective electrophilic iodination (Hua et al., 2006). By contrast, chlorine and bromine are stronger electrophiles, capable of reacting not only with moderately electron-rich aromatic rings (e.g., unsubstituted benzene) but also with certain electron-rich aliphatic sites, resulting in a much broader reaction spectrum.

Fig. 4c and g-i present the contribution proportion of individual DOM features on dehalogenation reactivity. For dehalogenation reactions, the individual DOM features exerting the strongest influence across all three pathways consistently converge on the number of carbon atoms (C), following by the number of aromatic carbon atoms (C_{AI}). For dechlorination and debromination reactions, an increase in the number of carbon atoms tends to attenuate reactivity (Fig. 4g-i). This may be attributed to steric hindrance and reduced accessibility, whereby a chlorine or bromine atom may be embedded within the interior of a large conjugated aromatic DOM framework with a complex 3D configuration (Li et al., 2016). As DOM molecules grow in size, the proportion of reactive sites decreases, and the electron density becomes more delocalized across the structure, thereby diminishing the presence of distinct reactive hot spots (Li et al., 2016). However, the case of deiodination is markedly different: as the number of carbon atoms increases, its influence on this reaction tends to diminish (Fig. 4i). When considered together with the earlier conclusion on iodination, this result suggests that iodine–DOM interactions in human-impacted lakes are predominantly governed by the characteristics of local functional groups, e.g., aromatic structures rich in phenolic hydroxyl groups (high NOSC and high DBE-O) (Xue et al., 2025). The finding that an increase in C_{AI} tends to suppress deiodination corroborates this interpretation (Fig. 4i). This implies that aromatic structures enhance the stability of iodine within DOM, thereby reducing its reactivity.

A limitation of this study is that the microbial analysis was based on 16S rRNA amplicon sequencing, which does not directly provide the information on functional genes involved in halogenation or dehalogenation processes. While Mantel tests revealed significant correlations between certain microbial taxa (e.g., Cyanobacteria, Nitrospirota, Acidobacteriota) and HDOM transformation pathways, the underlying genetic mechanisms, e.g., presence of halogenase or dehalogenase genes, remain unknown. Deep metagenomic or metatranscriptomic sequencing would be required to establish causal functional links, which represents an important direction for future research.

3.5. *Decoding halogenation and dehalogenation via PageRank molecular network*

PageRank molecular transformation networks of halogenation and dehalogenation reactions under seasonal variation in selected human-impacted lakes in the United States and China are presented in Figs. 5 and S1, based on the fundamental reaction types listed in Table S4. Relative distribution of these reactions are further summarized in Figure S2-S5. The PageRank algorithm surpasses traditional degree centrality by introducing a more nuanced and powerful framework for evaluating product importance in reaction networks (Zhou et al., 2019). Its principal advantages are threefold: (1) It incorporates weighted influence mechanism, where the intrinsic importance of each precursor HDOM is factored into the calculation, moving beyond the simplistic count of connections; (2) By accounting for the global network structure, it dynamically propagates the influence throughout the system, enabling the identification of crucial products that serve as hubs connected to highly influential precursors, even if their direct reaction count is low; and (3) this holistic approach grants PageRank molecular network superior robustness against noise and incidental reactions, mitigating misjudgment risks inherent in methods reliant solely on the local connection counts. Consequently, PageRank is exceptionally well-suited for delivering a precise and reliable assessment of complex interactions in HDOM reaction networks.

As illustrated in Figs. 5 and S1, both halogenation and dehalogenation reactions predominantly occur among Lignin/CRAM-like, Lipids, and Proteins, regardless of whether in the U.S. or China. This distribution pattern is consistent with their relative abundance in HDOM and further corroborates the results shown in Fig. 2. In contrast, the involvement of other classes in halogenation and dehalogenation is relatively minor, with only a small fraction of unsaturated hydrocarbons participating in these reactions in U.S. lakes. The predominance of halogenation and dehalogenation reactions among Lignin/CRAM-like compounds, Lipids, and Proteins can be attributed to both their molecular characteristics and relative abundance in HDOM (Fig. 2).

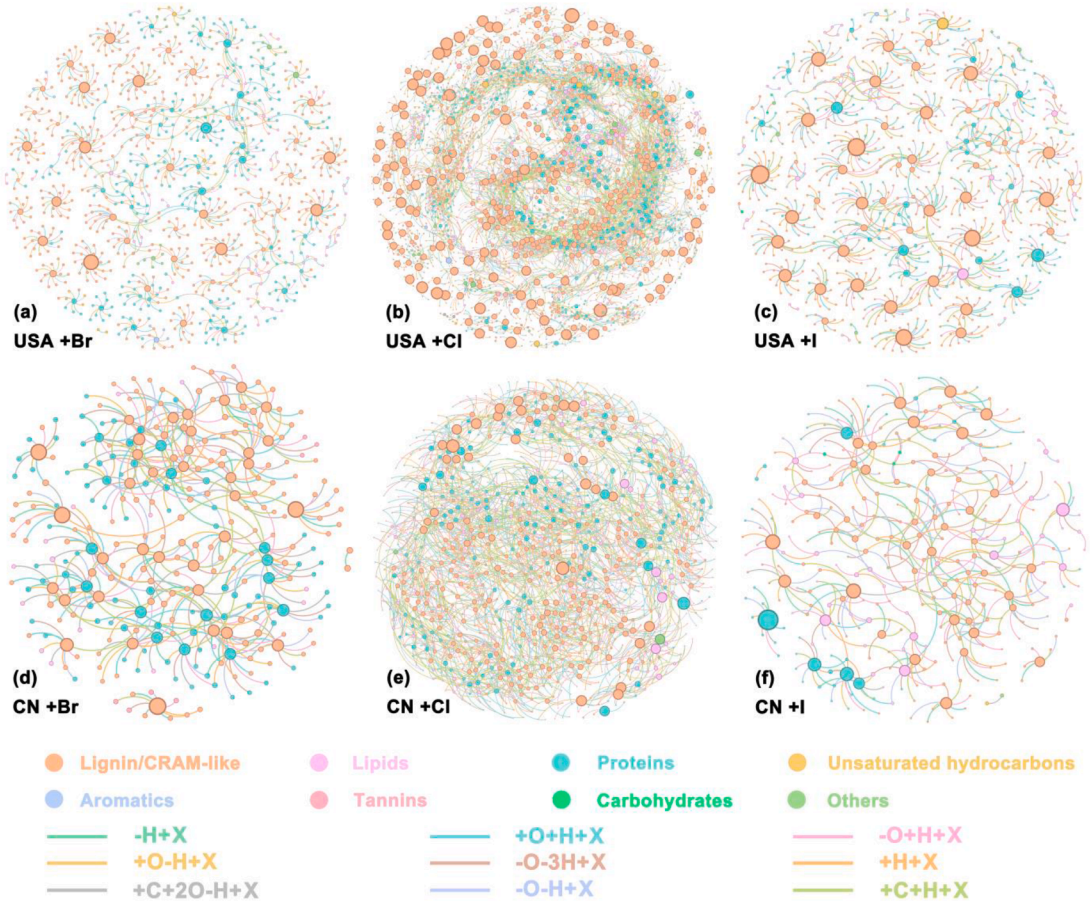


Fig. 5. PageRank molecular networks indicating some possible halogenation reactions for the DOM classes under seasonal variation in human-impacted lakes in the U.S. and China. (a) Bromination in three U.S. lakes, (b) chlorination in three U.S. lakes, (c) Iodization in three U.S. lakes, (d) bromination in three lakes in China, (e) chlorination in three lakes in China, and (f) Iodization in three lakes in China.

Structurally, lignin/CRAM-like compounds are enriched in phenolic hydroxyl groups and aromatic moieties, which serve as reactive sites for electrophilic substitution and radical-mediated halogenation (Liu et al., 2025). Lipids contain unsaturated bonds and hydroxyl functionalities that are prone to halogen addition or substitution (Crowe et al., 2021). Similarly, proteins provide multiple reactive amino acid residues (e.g., tyrosine, tryptophan, cysteine) that readily undergo halogenation and subsequent transformations (Sana et al., 2022). From a compositional perspective, these classes constitute a dominant fraction of HDOM, thereby increasing the likelihood of their involvement in halogen exchange reactions. Furthermore, environmental conditions, e.g., photochemical processes, redox dynamics, and microbial enzymatic activity preferentially facilitate the halogenation and dehalogenation of these reactive substrates.

Meanwhile, the observation that the halogenation and dehalogenation networks in U.S. lakes exhibit a more localized pattern, whereas those in Chinese lakes show greater overall connectivity, points to region-specific driving factors beyond mere compositional differences. (Figs. 5 and S1). In the U.S. lakes (e.g., Lake Erie, Lake St. Marys), anthropogenic pressure is mainly agricultural (Table S2). Agricultural runoff introduces a relatively homogeneous pool of halogenated precursors and DOM with limited functional group diversity, as elevated nitrate and acidic pH from fertilization have been shown to reduce molecular chemodiversity and restrict transformation pathways to a few dominant compound classes (Gad et al., 2025). This chemical homogeneity likely results in a localized network with fewer cross-class connections. In contrast, Chinese lakes (e.g., Lake Liuhua, Lake Songshan) receive a more diverse mixture of urban wastewater, industrial effluents, and agricultural runoff (Table S2). Urban eutrophic lakes contain DOM derived from multiple autochthonous sources, e.g., microbial metabolites, which introduce a broader range of reactive molecular structures (Wu et al., 2025). This creates a more heterogeneous DOM pool with higher molecular diversity, as reflected by the wider participation of proteins and lipids in halogenation reactions (Fig. 5d–f). Furthermore, studies on DOM chemodiversity across Chinese lakes have shown that land cover and human activity exert strong direct effects on DOM composition, with more diverse anthropogenic inputs leading to greater molecular heterogeneity (Luo et al., 2022). Additionally, differences in lake hydrology may contribute. Lakes with longer water residence time (WRT) exhibit more stable DOM composition and deeper processing of organic matter, which can restrict reaction diversity (Abbasi et al., 2024). The U.S. lakes studied (Erie, St. Marys, Harsha) are larger with longer WRT, potentially allowing spatial segregation of reaction zones. Conversely, Chinese lakes are smaller and often artificially managed, promoting faster mixing and greater interaction among different DOM pools (Jones et al., 2018). Together, these factors, such as anthropogenic source diversity, molecular heterogeneity, and hydrological regime, shape the contrasting network architectures observed between the two countries.

The halogenation and dehalogenation reactions of DOM exhibit substantial similarities in lakes in China and the U.S. (Figures S2–S5). Specifically, for Lignin/CRAM-like compounds, as the most dominant fraction of DOM, the reactions tend to favor halogenation processes coupled with carbon hydroxylation, with the highest relative contributions observed for the +H+X, +O+H+X, and -O+H+X pathways. It is plausible that the reactive halogen species (e.g., HOCl and HOBr) first

oxidize carbon moieties, thereby activating the molecular framework for subsequent halogen substitution. This oxidation–halogenation sequence could account for the higher proportions of coupled reactions, e.g., +O+H+X and -O+H+X, since oxidation introduces hydroxyl or carbonyl functionalities that facilitate electrophilic halogen attack. Strong oxidative capacity of halogen species, particularly under sunlight and in the presence of transition metals, further supports the feasibility of this pathway in natural lake environments (Fig. 4d-f). This corroborates the previous conclusion that oxidation-mediated reactions between halogens and Lignin/CRAM-like DOM remain active in human-impacted lakes, which may be attributed to the continuous environmental release of oxidative halogen species and formation of disinfection byproducts (Ding et al., 2019). In contrast, lipids and proteins exhibit greater propensity for halogen substitution reactions involving hydrogen (e.g., -H+Br and +O-H+Br) and hydroxyl groups (e.g., -O-H+Cl and -O-3H+Cl), respectively. The preferential halogen substitution of hydrogen and hydroxyl groups in lipids and proteins can be attributed to their molecular structures. Lipids, with abundant aliphatic C–H bonds and unsaturated carbon chains, provide readily accessible sites for electrophilic halogen attack (Duché et al., 2024). Proteins, on the other hand, contain amino acid residues with hydroxyl, amino, or thiol functionalities, such as those in serine, threonine, and tyrosine, which are particularly susceptible to halogen substitution (Sana et al., 2022). The inherent reactivity of these functional groups, coupled with the oxidative nature of halogen species (e.g., HOCl and HOBr), facilitates direct substitution rather than complex oxidation–halogenation coupling pathways observed in Lignin/CRAM-like DOM.

In contrast, the distribution of dehalogenation reaction types in HDOM appears more balanced, with the occurrence of different reaction pathways being relatively even (Figures S3 and S5). It likely reflects the combined effects of the heterogeneous molecular structure of HDOM, the stochastic distribution of halogen substituents across various functional sites, and environmental drivers such as sunlight and radical processes, which collectively allow multiple dehalogenation pathways to occur more evenly. Conversely, halogenation tends to target specific reactive sites, resulting in a less uniform distribution of reaction types. Notably, lipids in lakes of China and the U.S. were not observed to participate in the reactions of deiodination and debromination, respectively (Figures S3h and S5b), likely reflecting a combination of molecular and environmental constraints. The intrinsic properties of iodine and bromine require specific activation conditions, e.g., oxidation or radical generation, that may not be sufficiently met in these natural lake environments.

4. Conclusion

This study elucidates molecular transformation pathways and networks of halogenated DOM (HDOM) and its associated mechanism in human-impacted freshwater lakes in China and the U.S., considering that HDOM poses a formidable threat to drinking water safety due to its pronounced toxicity. At macroscopic compositional level, HDOM in both countries is dominated by lignin/CRAM-like substances, lipids, and proteins. Molecular transformation networks reveal that halogenation and dehalogenation pathways exhibit pronounced regional differences. HDOM in U.S. lakes displays higher reactivity and more localized transformation networks, whereas Chinese lakes exhibit stronger molecular connectivity, revealing region-specific biogeochemical behaviors of HDOM under anthropogenic pressures. However, through molecular-level and cross-regional analysis between the U.S. and China, we found that anthropogenic activities are a significant driving factor in the halogenation process ($r = 94.4\%$, $p < 0.01$), which is associated with DOM toxicity increase. These findings indicate that despite structural divergence, molecular transformation networks converge toward enhanced HDOM formation in lakes subjected to anthropogenic pressure. Meanwhile, our results demonstrate that anthropogenic pressures play significant roles in accelerating carbon cycling processes in these lakes.

At the molecular transformation level, interpretable machine learning analysis using the SHAP model confirmed that halogen-driven oxidative mechanisms dominate in natural lakes. This mechanism not only potentially enhances the pool of halogenated precursors relevant to highly toxic disinfection by-products formation in freshwater systems but may also suppress denitrification by depleting key electron donors (low-NOSC DOM), thereby further exacerbating eutrophication. Meanwhile, our results indicate that DOM molecular structure is a key determinant of halogenation reactivity in lakes. Structural characteristics such as aromaticity, aliphaticity, and unsaturation strongly influence the probability of halogen addition. Highly aromatic DOM generally shows lower halogenation reactivity due to the stability of condensed conjugated systems, whereas more aliphatic DOM tends to be more reactive. Distinct mechanistic differences also occur among halogens: iodination preferentially targets electron-rich aromatic functional groups, while chlorination and bromination exhibit broader reaction spectra involving both aromatic and certain aliphatic structures. In dehalogenation processes, DOM molecular size primarily controls dechlorination and debromination, likely due to steric hindrance and reduced accessibility of reactive sites, whereas the stability of iodine within DOM is more strongly governed by the local aromatic functional group environment. Aromatic structures, particularly those enriched with phenolic hydroxyl groups, can stabilize iodine substitutions through conjugated electronic systems and enhanced electron donation.

At the compositional level, halogenation and dehalogenation reactions predominantly occur within lignin/CRAM-like compounds, lipids, and proteins, indicating that both molecular structure and relative abundance determine the main targets of halogen exchange reactions. Lignin/CRAM-like compounds mainly undergo oxidative halogenation mechanism, in which oxidation activates DOM by introducing oxygen-containing functional groups that facilitate electrophilic halogen substitution. In contrast, lipids primarily experience halogen substitution on aliphatic C–H bonds, while proteins undergo substitution at hydroxyl-containing functional groups associated with reactive amino acid residues. Compared with halogenation, dehalogenation reactions show a more even distribution across pathways, likely reflecting the structural heterogeneity of HDOM, the stochastic distribution of halogen substituents, and the combined influence of photochemical and radical-mediated processes.

This study demonstrates that the composition and transformation of HDOM in lakes are co-shaped by terrestrial inputs and microorganisms within the context of global anthropogenic activities (agriculture and urbanization). Anthropogenic pressures not only directly drive this process by introducing precursors but also indirectly reshape HDOM transformation pathways and networks by altering the physicochemical aquatic environment and microbial community structure. Our findings suggest several practical strategies to reduce HDOM-related risks in drinking water source lakes: (1) control anthropogenic inputs (agricultural/urban runoff) at the catchment level, as land use correlates strongly with halogenation intensity; (2) enhance monitoring of low-NOSC DOM as an early warning indicator for increased DBP precursor formation; (3) adopt halogen-specific treatment approaches (e.g., targeted aromatic removal for iodinated DBP control); (4) avoid pre-chlorination and instead apply disinfectants after DOM removal to limit oxidative halogenation; and (5) tailor management strategies to regional pollution patterns: targeted removal for agricultural catchments, broad-spectrum DOM reduction for mixed-source catchments.

This study examined six lakes (three in China and three in the U.S.) that span a gradient of anthropogenic pressure. While the sample size is small and does not allow statistical generalization to all lakes in either country, the consistency of the observed pressure-halogenation relationship across two geographically and culturally distinct regions, together with the strong correlation, suggests that the underlying mechanism may be broadly relevant. Larger and more systematic lake selections are needed to confirm and extend these findings. Future

studies should employ metagenomic or metatranscriptomic approaches to directly identify functional genes and metabolic pathways involved in HDOM transformation, thereby establishing causal links between specific microbial taxa and HDOM dynamics observed in the present study. Stable isotope probing combined with metagenomics could further help identify active functional populations incorporating halogenated substrates.

CRediT authorship contribution statement

Chao Zhang: Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Junyu Zhu:** Software, Methodology, Formal analysis. **Wenjie Mai:** Software, Formal analysis. **Yue Xie:** Formal analysis. **Huan Chen:** Methodology, Data curation. **Harald Horn:** Writing – review & editing. **Haijun He:** Resources, Methodology. **Zhe Zhang:** Writing – original draft, Methodology, Formal analysis, Data curation. **Mingzhi Huang:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Fengchang Wu:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 41977300 and 42277053), Science and Technology Program of Quanzhou (China) (No. 2022C003R), Science and Technology Program of Qingyuan (China) (No. 2022KJJH020). The author C. Zhang acknowledges the support from Alexander von Humboldt Stiftung, National Natural Science Foundation of China (No. 42407308), the China Postdoctoral Science Foundation (No. 2024M760988), and Young Teachers Research Cultivation Project of South China Normal University (No. 23KJ02). A portion of this work was performed at the National High Magnetic Field Laboratory ICR User Facility, which is supported by the National Science Foundation Division of Chemistry and Division of Materials Research through DMR-2128556, (previous grant number DMR-1644779) and the State of Florida. We would like to express our sincere gratitude to the late Dr. Dionysios (Dion) Dionysiou at the University of Cincinnati for his valuable guidance in the design of this study, as well as his supervision during the experiments and data analysis. We gratefully acknowledge Dr. Amy M. McKenna at the National High Magnetic Field Laboratory, Florida State University, for their expert assistance and insightful contributions to the FT-ICR MS analysis of our samples.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.watres.2026.126771.

Data availability

Data will be made available on request.

References

Abbasi, M., Peacock, M., Drakare, S., Hawkes, J., Jakobsson, E., Kothawala, D., 2024. Water residence time is an important predictor of dissolved organic matter composition and drinking water treatability. Water Res. 260, 121910. https://doi.org/10.1016/j.watres.2024.121910.

Benalaya, I., Alves, G., Lopes, J., Silva, L.R., 2024. A review of natural polysaccharides: sources, characteristics, properties, food, and pharmaceutical applications. Int. J. Mol. Sci. 25 (2), 1322. https://doi.org/10.3390/ijms25021322.

Catalán, N., Marcé, R., Kothawala, D., Tranvik, L.J., 2016. Organic carbon decomposition rates controlled by water retention time across inland waters. Nat. Geosci. 9, 501–504. https://doi.org/10.1038/ngeo2720.

Chen, W., Hu, H., Heal, K., Sohi, S., Tigabu, M., Qiu, W., Zhou, C., 2023. Linking microbial decomposition to dissolved organic matter composition in the revegetation of the red soil erosion area. Forests 14 (2), 270. https://doi.org/10.3390/f14020270.

Crowe, C., Molyneux, S., Sharma, S.V., Zhang, Y., Gkotsi, D.S., Connaris, H., Goss, R.J.M., 2021. Halogenases: a palette of emerging opportunities for synthetic biology–synthetic chemistry and C–H functionalisation. Chem. Soc. Rev. 50, 9443–9481. https://doi.org/10.1039/D0CS01551B.

Čulin, J., Mustać, B., 2015. Environmental risks associated with ballast water management systems that create disinfection by-products (DBPs). Ocean Coast. Manage. 105, 100–105. https://doi.org/10.1016/j.ocecoaman.2015.01.004.

Daims, H., Lebedevam, E.V., Pjevac, P., Han, P., Herbold, C., Albertsen, M., Jehmlich, N., Palatinszky, M., Vierheilig, J., Bulaev, A., Kirkegaard, R.H., von Bergen, M., Rattei, T., Bendinger, B., Nielsen, P.H., Wagner, M., 2015. Complete nitrification by Nitrospira bacteria. Nature 528 (7583), 504–509. https://doi.org/10.1038/nature16461.

Davenport, R.E., Kroeger, M.E., Tolić, N., Toyoda, J.G., Chu, R.K., Lehmann, S., Clendinen, C.S., Bowen, B.P., Rossiter, D.G., Buckley, D.H., Lehmann, J., 2025. Decomposition causes short-term increases in functional molecular diversity of dissolved organic matter. Nat. Commun. 16 (1), 11155. https://doi.org/10.1038/s41467-025-65990-8.

Ding, S.K., Deng, Y., Bond, T., Fang, C., Cao, Z.Q., Chu, W.H., 2019. Disinfection byproduct formation during drinking water treatment and distribution: a review of unintended effects of engineering agents and materials. Water Res. 160, 313–329. https://doi.org/10.1016/j.watres.2019.05.024.

Duché, G., Sanderson, J.M., 2024. The chemical reactivity of membrane lipids. Chem. Rev. 124 (6), 3284–3330. https://doi.org/10.1021/acs.chemrev.3c00608.

Farinelli, G., Coha, M., Vione, D., Minella, M., Tiraferri, A., 2022. Formation of halogenated byproducts upon water treatment with peracetic acid. Environ. Sci. Technol. 56 (8), 5123–5131. https://doi.org/10.1021/acs.est.1c06118.

Feng, H.L., Zhang, Z.J., Kuang, Q.Y., Chen, S.H., Huang, D., Zhou, X.W., 2023. The transformation of dissolved organic matter and formation of halogenated by-products during electrochemical advanced oxidation pretreatment for shale gas produced water. J. Hazard. Mater. 455, 131614. https://doi.org/10.1016/j.jhazmat.2023.131614.

Feng, Y.M., Qian, Y.F., Liu, C., Fu, Q.L., 2025. Molecular characteristics of halogenated disinfection byproducts elucidated by Fourier transform ion cyclotron resonance mass spectrometry. Environ. Pollut. 385, 127071. https://doi.org/10.1016/j.envpol.2025.127071.

Gad, M., Khomami, N.T.S., Krieg, R., Schor, J., Philippe, A., Lechtenfeld, O.J., 2025. Environmental drivers of dissolved organic matter composition across central European aquatic systems: a novel correlation-based machine learning and FT-ICR MS approach. Water Res. 273, 123018. https://doi.org/10.1016/j.watres.2024.123018.

Hedges, J.I., Keil, R.G., Benner, R., 2013. What happens to terrestrial organic matter in the ocean? Org. Geochem. 503 (5–6), 355–359. https://doi.org/10.1038/nature12760.

Hu, A., Jang, K.S., Tanentzap, A.J., Zhao, W., Lennon, J.T., Liu, J., Li, M., Stegen, J., Choi, M., Lu, Y., Feng, X., Wang, J., 2024. Thermal responses of dissolved organic matter under global change. Nat. Commun. 15 (1), 576. https://doi.org/10.1038/s41467-024-44813-2.

Hu, S., Zhang, Y., et al., 2024. Response of dissolved organic matter and disinfection by-product precursors to algal blooms and thermal stratification in deep reservoirs. Chemosphere 368, 143757. https://doi.org/10.1016/j.chemosphere.2024.143757.

Hua, G.H., Reckhow, D.A., Kim, J., 2006. Effect of bromide and iodide ions on the formation and speciation of disinfection byproducts during chlorination. Environ. Sci. Technol. 40 (9), 3050–3056. https://doi.org/10.1021/es0519278.

Huang, Z., Liu, H., Wang, C., Wang, J., Tian, C., Feng, J., Shen, J., Wang, X., 2026. Regulation of carbon cycling in plateau lakes by trophic states and seasonal variations: a focus on dissolved organic matter and microbial interactions. Water Res. 292, 125312. https://doi.org/10.1016/j.watres.2025.125312.

Ilunga, A.K., Mamba, B.B., Nkambule, T.T.L., 2021. Catalytic hydrodehalogenation of halogenated disinfection byproducts for clean drinking water production: a review. J. Water Process Eng. 44, 102402. https://doi.org/10.1016/j.jwpe.2021.102402.

Introduction to Wanlv Lake, 2023. Wanlvhu.cn. https://www.wanlvhu.cn/en/introduct-33.html (Accessed 12 October 2025).

Ji, H.K., Wang, H., Wu, Z.P., Wang, D.F., Wang, X., Fu, P.J., Li, C.S., Deng, W.G., 2024. Source, composition and molecular diversity of dissolved and particulate organic matter varied with riparian land use in tropical coastal headstreams. *Sci. Total Environ.* 908 (1), 168577. https://doi.org/10.1016/j.scitotenv.2023.168577.

Jones, S.E., Zwart, J.A., Kelly, P.T., Solomon, C.T., 2018. Hydrologic setting constrains lake heterotrophy and terrestrial carbon fate. Limnol. Oceanogr. Lett. 3, 256–264. https://doi.org/10.1002/lol2.10054.

Kodavanti, P.R.S., Costa, L.G., Aschner, M., 2023. Perspective on halogenated organic compounds. Adv. Neurotoxicol. 10, 1–25. https://doi.org/10.1016/bs.ant.2023.06.001.

Lawson, C.E., Mundinger, A.B., Koch, H., Jacobson, T.B., Weathersby, C.A., Jetten, M.S. M., Pabst, M., Amador-Noguez, D., Noguera, D.R., McMahon, K., Lücker, S., 2021. Investigating the chemolithoautotrophic and formate metabolism of nitrospira

moscoviensis by constraint-based metabolic modeling and 13C-tracer analysis. mSystems 6 (4), e0017321. https://doi.org/10.1128/mSystems.00173-21.
Li, L., Hu, J., Shi, X., Ruan, W., Luo, J., Wei, X., 2016. Theoretical studies on structures, properties and dominant debromination pathways for selected polybrominated diphenyl ethers. Int. J. Mol. Sci. 17, 927. https://doi.org/10.3390/ijms17060927.
Li, T.Q., Lü, F., Zhang, H., Xu, Q.Y., He, P.J., 2023. Nontarget insights into the fate of Cl-/Br-containing DOM in leachate during membrane treatment. Environ. Sci. Technol. 57 (42), 16033–16042. https://doi.org/10.1021/acs.est.3c04422.
Liu, Y., Li, Y., He, Z., et al., 2025. Producing aryl halides from lignin. Nat. Commun. 16, 3673. https://doi.org/10.1038/s41467-025-59054-0.
Liu, Z., Deng, Z., He, G., Wang, H.L., Zhang, X., Lin, J., Qi, Y., Liang, X., 2022. Challenges and opportunities for carbon neutrality in China. Nat. Rev. Earth Environ. 3, 141–155. https://doi.org/10.1038/s43017-021-00244-x.
Lu, W.C., Babcock-Adams, L., Sharifan, H., Kornuc, J.J., Hao, S.L., Higgins, C.P., Young, R.B., McKenna, A.M., Blotevogel, J., 2026. Exploring the fluorinome of PFAS-impacted groundwater using 21 tesla FT-ICR mass spectrometry. Water Res. 288, 124698. https://doi.org/10.1016/j.watres.2025.124698.
Luo, J.W., Zhou, Q.X., Hu, X.G., Zeng, H., Deng, P., He, C., Shi, Q., 2022. Lake chemodiversity driven by natural and anthropogenic factors. Environ. Sci. Technol. 56 (9), 5910–5919. https://doi.org/10.1021/acs.est.1c08148.
Mitch, W.A., Richardson, S.D., Zhang, X., Gonsior, M., 2023. High-molecular-weight by-products of chlorine disinfection. Nat. Water 1, 336–347. https://doi.org/10.1038/s44221-023-00064-x.
Mohammed, A.S.A., Naveed, M., Jost, N., 2021. Polysaccharides: classification, chemical properties, and future perspective applications in fields of pharmacology and biological medicine (a review of current applications and upcoming potentialities). J. Polym. Environ. 29 (8), 2359–2371. https://doi.org/10.1007/s10924-021-02052-2.
Munthali, E., Marcé, R., Farré, M.J., 2022. Drivers of variability in disinfection by-product formation potential in a chain of thermally stratified drinking water reservoirs. Environ. Sci.: Water Res. Technol. 8 (4), 968–980. https://doi.org/10.1039/D1EW00788B.
Nykl, M., Ježek, K., Fiala, D., Dostal, M., 2014. PageRank variants in the evaluation of citation networks. J. Informetr. 8 (3), 683–692. https://doi.org/10.1016/j.joi.2014.06.005.
Patel, A.B., Shaikh, S., Jain, K.R., Desai, C., Madamwar, D., 2020. Polycyclic aromatic hydrocarbons: sources, toxicity, and remediation approaches. Front. Microbiol. 11, 562813. https://doi.org/10.3389/fmicb.2020.562813.
Raymond, P., Hartmann, J., Lauerwald, R., Sobek, S., McDonald, C., Hoover, M., Butman, D., Striegl, R., Mayorga, E., Humborg, C., Kortelainen, P., Dürr, H., Meybeck, M., Ciais, P., Guth, P., 2013. Global carbon dioxide emissions from inland waters. Nature 503, 355–359. https://doi.org/10.1038/nature12760.
Ren, X., Wang, Y., Wan, J., Yan, Z., Ma, Y., Zhang, G., Zhu, B., 2022. The nitrogen removal performance and functional bacteria in heterotrophic denitrification and mixotrophic denitrification process. Water 14, 3603. https://doi.org/10.3390/w14223603.
Richardson, S.D., Kimura, S.Y., 2020. Water analysis: emerging contaminants and current issues. Anal. Chem. 92 (1), 473–505. https://doi.org/10.1021/acs.analchem.9b05269.
Rogers, J.E., Devereux, R., James, J.B., George, S.E., Forshay, K.J., 2021. Seasonal distribution of cyanobacteria in three urban eutrophic lakes results from an epidemic-like response to environmental conditions. Curr. Microbiol. 78 (6), 2298–2316. https://doi.org/10.1007/s00284-021-02498-6.
Ruan, X.X., Xiang, Y.Y., Shang, C., Cheng, S.S., Liu, J.F., Hao, Z.N., Yang, X., 2021. Molecular characterization of transformation and halogenation of natural organic matter during the UV/chlorine AOP using FT-ICR mass spectrometry. J. Environ. Sci. 102, 24–36. https://doi.org/10.1016/j.jes.2020.08.028.
Sana, B., Ke, D., Li, E.H.Y., Ho, T., Seayad, J., Duong, H.A., Ghadessy, F.J., 2022. Halogenation of peptides and proteins using engineered tryptophan halogenase enzymes. Biomolecules 12 (12), 1841. https://doi.org/10.3390/biom12121841.
Shi, X., Mao, D., Song, K., Xiang, H., Li, S., Wang, Z., 2024. Effects of landscape changes on water quality: a global meta-analysis. Water Res. 260, 121946. https://doi.org/10.1016/j.watres.2024.121946.
Tiwari, O.N., Sasmal, S., Kataria, A.K., Devi, I., 2020. Application of microbial extracellular carbohydrate polymeric substances in food and allied industries. 3 Biotech 10 (5), 221. https://doi.org/10.1007/s13205-020-02200-w.
Vörösmarty, C.J., McIntyre, P.B., Gessner, M.O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S.E., Sullivan, C.A., Liermann, C.R., Davies, P.M., 2010. Global threats to human water security and river biodiversity. Nature 467 (7315), 555–561. https://doi.org/10.1038/NATURE09440.
Wang, S.L., Ran, F.W., Li, Z.W., Yang, C.R., Xiao, T., Liu, Y.J., Nie, X.D., 2024. Coupled effects of human activities and river–lake interactions evolution alter sources and fate of sedimentary organic carbon in a typical river–Lake system. Water Res. 255 (17), 121509. https://doi.org/10.1016/j.watres.2024.121509.
Wang, X.G., He, W., He, X.S., Cao, X., Sun, Y., Zeng, X.J., 2025. Occurrence, removal and ecotoxicological assessment of halogenated dissolved organic matter in waste leachates. J. Hazard. Mater. 497, 139640. https://doi.org/10.1016/j.jhazmat.2025.139640.
Wu, S.X., Fu, Q.L., Fujii, M., Liu, J.B., Elreedy, A., Yang, X., Fu, P.Q., Wang, Y.X., 2025. Qualitatively elucidating the molecular characteristics of precursors for saturated halogenated disinfection byproducts in chlorinated urban eutrophic lake water by ultrahigh-resolution mass spectrometry. J. Hazard. Mater. 489, 137709. https://doi.org/10.1016/j.jhazmat.2025.137709.
Wu, S.X., Fujii, M., Yang, X., Fu, Q.L., 2023. Characterization of halogenated organic compounds by the Fourier transform ion cyclotron resonance mass spectrometry: a critical review. Water Res. 246, 120694. https://doi.org/10.1016/j.watres.2023.120694.
Xiao, R., Deng, Y., Xu, Z.X., Chu, W.H., 2024. Disinfection byproducts and their precursors in drinking water sources: origins, influencing factors, and environmental insights. Engineering 36, 36–50. https://doi.org/10.1016/j.eng.2023.08.017.
Xue, J.K., Deng, Y.M., Fu, Q.L., Du, Y., Shi, J.B., Wang, Y.X., 2025. Unveiling molecular constraints on the formation and fate of organoiodine compounds under aquifer conditions. Environ. Sci. Technol. 59 (34), 18283–18293. https://doi.org/10.1021/acs.est.5c05558.
Yang, B., Xu, W.L., Zhao, W.K., 2026. Multi-scale mechanisms and environmental implications of dissolved organic matter-metal ions interactions in aquatic environments: a review. Water Res. 288 (A), 124563. https://doi.org/10.1016/j.watres.2025.124563.
Yu, W., Yang, H., Chen, J., et al., 2024. Molecular insights into the microbial degradation of sediment-derived DOM in a macrophyte-dominated lake under aerobic and hypoxic conditions. Sci. Total Environ. 916, 170257. https://doi.org/10.1016/j.scitotenv.2024.170257.
Zhang, C., Zhu, J.Y., Mai, W.J., Chen, Z.G., Xie, Y., Fu, S.N., Xia, D., Cai, C., Zheng, W.B., Liu, J.X., Yang, L.M., Zhang, Z., Huang, M.Z., Wu, F.C., 2025. Satellite remote sensing-implemented nontargeted screening of emerging contaminant fingerprints in a river-to-ocean continuum through interpretable machine learning: the pivotal intermediary role of dissolved organic matter. Environ. Sci. Technol. 59 (17), 8714–8726. https://doi.org/10.1021/acs.est.4c14425.
Zhao, W.L., Cai, M.G., Adelman, D., Khairy, M., Lin, Y., Li, Z.H., Liu, H.J., Lohmann, R., 2022. Legacy halogenated organic contaminants in urban-influenced waters using passive polyethylene samplers: emerging evidence of anthropogenic land-use-based sources and ecological risks. Environ. Pollut. 298, 118854. https://doi.org/10.1016/j.envpol.2022.118854.
Zhou, L., Wu, Y.H., Zhou, Y.Q., Zhang, Y.L., Xu, H., Jang, K.S., Dolfing, J., Spencer, R.G.M., Jeppesen, E., 2024. Terrestrial dissolved organic matter inputs drive the temporal dynamics of riverine bacterial ecological networks and assembly processes. Water Res. 249, 120955. https://doi.org/10.1016/j.watres.2023.120955.
Zhou, T.C., Martinez-Baez, E., Schenter, G., Clark, A.E., 2019. PageRank as a collective variable to study complex chemical transformations and their energy landscapes. J. Chem. Phys. 150 (13), 134102. https://doi.org/10.1063/1.5082648.