Molecular properties of dissolved organic matter across Earth systems: A meta-analysis

Lei Han[1], Ang Hu[1], Hellen Lucas Mzuka[1], Xingting Chen[2], Ji Shen[2], and Jianjun Wang[1]

[1]Nanjing Institute of Geography and Limnology, State Key Laboratory of Lake Science and Environment, Chinese Academy of Science
[2]School of Geography and Ocean Science, Nanjing University

April 04, 2024

Molecular properties of dissolved organic matter across Earth systems: A meta-analysis

Lei Han[1], Ang Hu[1,*], Hellen Lucas Mzuka[1], Xingting Chen[2], Ji Shen[2], Jianjun Wang[1,*]

[1] State Key Laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology, Chinese Academy of Science, Nanjing 210008, China

[2] School of Geography and Ocean Science, Nanjing University, Nanjing 210023, China

* Correspondence: anghu@niglas.ac.cn (AH); jjwang@niglas.ac.cn (JW)

Abstract

Dissolved organic matter (DOM) represents the largest pool of reactive carbon on the Earth and plays a crucial role in various biogeochemical processes and ecosystem functions. However, it is understudied for a global understanding of DOM molecular properties such as molecular weight, stoichiometry, and oxidation state, and the linkages among them across Earth systems. Here, a meta-analysis of 2,707 sites in 204 literatures was conducted by synthesizing four representative molecular properties of DOM, i.e., mass, double bond equivalent (DBE), modified aromaticity index (AI_{mod}), and nominal oxidation state of carbon (NOSC). By exploring H/C and O/C ratios, we examined the relationships among these DOM properties across waters and land systems, and their geographical patterns and environmental drivers. We found that, compared to land system, the mass, DBE, and AI_{mod} were all significantly higher in water systems, with river sediments exhibiting the highest values. DOM oxidation state indicated by NOSC was greater on average in wastewater (NOSC = 0.226 ± 0.06) and marine water (NOSC = 0.133 ± 0.06) than in other habitats. Compared to waters, the mass in land system showed more strongly positive correlations with oxidation states such as NOSC and O/C, and the NOSC showed stronger relations to bioavailability

properties such as DBE, AI_{mod}, and H/C. Among all the properties, H/C and AI_{mod} contributed to the most variations in global DOM properties. In waters, NOSC monotonically increased towards high latitudes, while DBE and AI_{mod} showed significant hump-shaped patterns indicating peaked unsaturation and aromaticity at mid-latitudes of approximately 30°-50°. The variations in DOM properties were significantly correlated with environmental factors such as annual mean temperature and pH. Collectively, we revealed the spatial distribution and environmental drivers of DOM molecular properties across Earth ecosystems, which could shed light on our comprehensive understanding of DOM characteristics and its dynamics.

Highlights

1. DOM mass, DBE, and AI_{mod} were significantly higher in the waters than the land.
2. Across the habitats, the mass, DBE, and AI_{mod} were largest in river sediment, and the NOSC was largest in wastewater.
3. Compared to waters, mass showed more strongly positive correlations with oxidation states such as NOSC and O/C in the land system, and NOSC showed stronger correlations with bioavailability properties such as DBE, AI_{mod}, and H/C.
4. The H/C and AI_{mod} dominantly controlled for DOM properties, while the mass was less influenced in waters.
5. In the waters, NOSC linearly increased along latitude gradients, while DBE and AI_{mod} showed significant hump-shaped patterns, and AI_{mod} linearly decreased along latitude gradients in the land.
6. In the waters, mean annual temperature and pH were closely related to DOM properties.

Introduction

Dissolved organic matter (DOM) plays a crucial role in the global carbon cycle and participates in multiple physical, chemical, and biological processes across Earth systems (Schmidt et al. 2011, Dittmar and Stubbins 2014). The global DOM pool contains carbon, nitrogen, phosphorus, and elements essential to life (Creed et al. 2018, Kellerman et al. 2018). Marine DOM represents the largest reduced carbon reservoir (~662 Pg C) in the oceans and covers about 96% of the ocean's total organic carbon (Hansell et al. 2009, Wagner et al. 2020). Inland water is the recipient of approximately 5.1 Pg C $year^{-1}$ terrigenous carbon (Drake et al. 2017). Lake and river-derived DOM conversion releases 2.1 Pg C $year^{-1}$ as CO_2 into the atmosphere (Raymond et al. 2013, Moody 2020). The contributions of DOM properties to the global carbon cycle's processes have been extensively explored by focusing on one or a few ecosystems, such as glaciers (Singer et al. 2012, Wadham et al. 2019, Nagar et al. 2021), oceans (Jiao et al. 2010, Dittmar et al. 2021), inland water (Catalán et al. 2016, Begum et al. 2023), terrestrial soil (Schmidt et al. 2011, Kleber et al. 2021, Speetjens et al. 2022). However, the properties of DOM molecules and the drivers that control their recalcitrance and bioavailability remain poorly understudied across global scales.

DOM molecular indices have been developed to evaluate the bioavailability state of DOM such as mass, DBE, AI_{mod}, NOSC, H/C, and O/C (Kim et al. 2003, Koch and Dittmar 2006, LaRowe and Van Cappellen 2011, Koch and Dittmar 2016) using Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS). These indices can reflect molecule composition and biogeochemical reactions of DOM (Koch and Dittmar 2006, Cai and Jiao 2023). Mass, DBE, AI_{mod}, and H/C ratio are generally relevant to recalcitrance of DOM molecules; a higher mass, DBE, and AI_{mod} at the compositional level reflects more recalcitrant state in the DOM assemblage (Medeiros et al. 2015, Hildebrand et al. 2022, Zherebker et al. 2022). Based on molecular lability of organic matter, the environments generally ranked from most to least as glacial > marine > freshwater (D'Andrilli et al. 2015). NOSC and O/C ratio are relevant to oxidation state of DOM molecules (LaRowe and Van Cappellen 2011, Laszakovits and MacKay 2022), and NOSC could be also treated as a thermodynamic threshold to determine the decomposition of compounds (Boye et al. 2017, Bahureksa et al. 2021). Higher NOSC generally represents higher condensed hydrocarbons, lignin-like, and tannin-like compounds in DOM assemblages (Pracht et al. 2018). DOM properties are strongly depended on the original source and environmental conditions (Catalán et al. 2016, Ward and Cory 2016, Ward et al. 2017, Zark and Dittmar 2018), as well as on its susceptibility to microbial and abiotic transformations (Stegen et al. 2018, D'Andrilli et al. 2019, Hu et al. 2022a). For example, acidity enhancement causes an increase in the abundance of oxidized unsaturated compounds (e.g., aromatics and carboxylic-rich alicyclic molecules) (DiDonato et al. 2016, Zherebker et al. 2020). DOM properties are also largely controlled by climatic factors and extreme climate. For instance, the abundances of polyphenol, nitrogen-containing, and unsaturated oxygenated compounds are higher in the regions increasing in mean annual temperature and precipitation (Roth et al. 2013, Kellerman et al. 2014), and the organic matter concentrations and aromatic compounds abundance could be reduced in dry conditions (Szkokan-Emilson et al. 2017, Butturini et al. 2022). A recent global DOM meta-analysis indicates that there could be predictable spatial patterns of H/C and O/C ratios across Earth systems, and reveals the importance linkages between these two ratios and environmental conditions or extreme climates (Hu et al. 2024a). However, the other DOM properties such as mass, DBE, AI_{mod}, and NOSC are left understudied across various systems for a global scale.

Here, we compiled three categories of compositional-level chemical properties of DOM derived from 2,707 sites reported in 204 studies during 2003 - 2022 (Fig. 1, and Table S1). The properties included 1) sizes of molecules such as mass, 2) bioavailability such as DBE, AI_{mod}, and H/C, and 3) oxidation states such as NOSC and O/C. The datasets were mainly consisted of waters and land systems, and these habitats spanned ocean, river, lake, reservoir, engineered water, peatland, and soil. We aimed to characterize the global properties of DOM molecular indices. Specifically, we showed the distribution of DOM indices across Earth systems and the relationships among DOM indices as well as along latitudinal patterns, and illustrated the potential effects of environmental variables on DOM traits. Our results provide fundamental insights to better understand the overall processes of the global carbon cycling.

Materials and Methods

Data collection

We utilized the following keywords: "organic matter" AND "FT-ICR MS" AND "van Krevelen" to search for papers using the Google Scholar Database (http://scholar.google.com) and Web of Science (Core Collection; http://www.webofknowledge.com)up to the end of June 2022. We extracted studies based on the following criteria: (1) Weighted means of the H/C ratio, O/C ratio, mass, DBE, AI_{mod} (Koch and Dittmar 2006) and NOSC (LaRowe and Van Cappellen 2011) formula-based characteristics were calculated as the sum of the product of the individual information and relative intensity divided by the sum of all intensities. The H/C ratio, DBE, and AI_{mod} represent the saturation level of a molecule, whereas the O/C ratio and NOSC mainly represent the carbon oxidation state (Butturini et al. 2020). (2) There are raw FT-ICR MS data accessible, from which the molecular compositional-level parameters could be calculated. but only the DOM data with an electrospray ionization source in negative conditions for subsequent analysis. (3) They focused on the dissolved organic matter data in the natural ecosystems.

Following these criteria, a total of 2,707 samples from 204 articles, the range of MAT and MAP are mainly concentrated in 5-15℃ and 1500 mm, respectively (Fig. 1; Table S1). A database was designed to provide a comprehensive sampling of different systems and habitat types. We considered about 84% database of the database only those ESI sources in negative mode that had been measured to be relevant for statistical analysis. 15 environmental variables were collected for subsequent analysis. In addition, to explore major drivers of DOM properties at a global scale, we extracted bioclimatic variables based on latitude, longitude, and elevation from the WorldClim database (https://www.worldclim.org) with a spatial resolution of 0.5° for each sample (Delgado-Baquerizo et al. 2020). Further details on environmental and bioclimatic variables are described in Hu et al (Hu et al. 2024a).

Statistical analysis

The pairwise comparison between water and land was analyzed by using the Wilcoxon test. Principal components were calculated by the six indices matrix of the water and land. The principal component analysis (PCA) to simplify the interpretation of principal components across different ecosystems by maximizing/minimizing the relevance between molecular indices and component axes. These analyses were conducted by R package stats V4.3.1. The Spearman analyses using the Hmisc (V 5.1.0) package in R were used to analyze the correlation among DOM molecular indices.

To explore correlation and latitudinal patterns of molecular indices across ecosystems, we utilized more suitable linear or quadratic model according to the lower value of Akaike's information criterion (Yamaoka et al. 1978). To deal with the limitation of incomplete dataset, we used the Euclidean distance as a dissimilarity index with the pcoa function in the vegan R package. We calculated DOM molecular indices in the waters to convert first and second principal co-ordinates component. To reduce collinearity among variables, we reduced the primary set of 36 drivers to 15 with variation inflation factor (VIF) below 5. This final set contained 10 variables database, such as latitude, elevation, BIO1, BIO2, BIO8, BIO9, BIO15, BIO18, BIO19, and pH. Then, we evaluate the relative importance of variables to DOM properties in waters using random forest analysis with the randomForest package V4.7-1.1 (Breiman 2001). The number of trees utilized in the random forest analysis was set as 500. According to above analyses, the impact of individual environmental variables on molecular indices in waters were further assessed by linear mixed-effects models (Galecki and Burzykowski 2013). We utilized literature identity as random factor and evaluated model significance by omnibus test (Nakagawa et al. 2013). Linear mixed-effects models were analyzed with R package lme4 V1.1.28.

Results

DOM properties across systems and habitats

The mass size and bioavailability properties (e.g., DBE and AI_{mod}) of DOM in the water system had higher values than in the land system, with mean values of mass-to-charge ratio $m/z = 425$, DBE = 10, and $AI_{mod} = 0.36$, respectively (Fig. 2a, and Table S2). For the individual habitats of waters, the mass, DBE and AI_{mod} values were highest in river sediment, with mean values of 447, 12, and 0.46, respectively, while were lower in wastewater, stream water and pond water than all the other habitats (Fig. 2b, and Table S3). The oxidation states of DOM, i.e. NOSC, varied from the highest to the lowest generally as wastewater > marine > freshwater in the habitats of waters, with the lowest average values in river sediment. For the land habitats, the mass, AI_{mod}, and NOSC showed the significantly ($P < 0.05$) lower values in peatland than in forest soil, and the DBE and AI_{mod} showed the largest values in cropland compared to other soil types (Figs. 2, S1).

The relationships among DOM properties

There were generally significant correlations within and between the three categories of chemical properties of DOM, and these correlations also depended on systems and habitats. Specifically, within the property categories, the DBE and AI_{mod} showed highly negative correlation with H/C, and the NOSC showed positive correlation with O/C in both systems, with higher correlations in the waters than land (Fig. 3). There is one exception to the relationship between NOSC and O/C in marine water, that is negative correlation with R^2 of 0.11 ($P < 0.05$) (Fig. 3a). In between the property categories, the mass strongly correlated with the bioavailability, with the positive correlations with DBE in both systems and the negative correlation with H/C especially in the land system. The mass also showed strongly positive correlations with the oxidation state properties of NOSC (Spearman $\rho = 0.45$) and O/C (Spearman $\rho = 0.46$) in the land system, but weak correlations in the waters. The bioavailability properties (i.e., DBE, AI_{mod}, and H/C) strongly correlated with NOSC in the land than in the waters, while they had relatively stronger correlations with O/C in the waters than in the land system.

Variations of DOM properties across systems and habitats

To identify the main indices contributed to variation in DOM properties, we applied the principal component analysis (PCA) of mass, DBE, AI_{mod}, NOSC, H/C, and O/C (Fig. 4; Fig. S2). Among all DOM samples, we found a clear separation of DOM properties between waters and land systems, and these differences were significant in the first two axes of principal component (Fig. 4c). The first principal component (PC1) accounted for 45% of variation in DOM properties, and had loadings of -0.59, 0.46, and 0.58 for H/C, DBE, and AI_{mod}, respectively (Fig. S2). This indicates that the first principal component could be interpreted as a dimension of bioavailability. The second principal component (PC2) accounted for 31% of variation in DOM properties, and had loadings of 0.62 and 0.70 for NOSC and O/C, respectively (Fig. S2). This indicates that the second principal component can be interpreted as an oxidation state dimension. For the mass, there were minor loadings in the both principal components, and few contributions to the variation in DOM properties.

Notably, the variations of DOM indices from waters were similarly to those from land (Fig. 4a, c). The PC1 and PC2 explained approximately 46% and 30% of the variance in waters, respectively (Fig. 4a). The DOM properties showed that the loadings of -0.59, 0.43, and 0.57 for H/C, DBE, and AI_{mod} were distributed along PC1, and the loadings of -0.17, 0.71, and 0.56 for mass, O/C, and NOSC were distributed along PC2 (Fig. S2). DOM properties also occurred in specific habitats such as riverine, lake water, and reservoir water; H/C, AI_{mod}, and NOSC generally had lager loadings along PC1 scores, whereas mass and O/C had major contribution to the variation along PC2 scores (Fig. 4d). The main variations of DOM indices in river sediment generally showed significantly differences relative to river water, lake water, and reservoir water, but were similar to land system (Fig. 4b, d). However, the variations of DOM indices from land distributed distinctly from waters. The PC1 and PC2 accounted for 46% and 34% of the variance in land, respectively (Fig. 4a). The DOM properties showed that the loadings of 0.56 and 0.03 for H/C and mass had positive PC1 scores, whereas the loadings of -0.37, -0.44, -0.26, and -0.53 for DBE, AI_{mod}, O/C, and NOSC had negative PC1 scores (Fig. 4b; Fig. S2). The DOM properties were also reported in forest soil (Fig. 4d).

Latitudinal patterns of DOM properties across systems and habitats

The DOM properties of mass, DBE, AI_{mod}, and NOSC generally showed predictable latitudinal patterns across systems and habitats (Fig. 5). In the waters, the DBE (R^2 = 0.54, P < 0.001) and AI_{mod} (R^2 = 0.077, P < 0.001) showed a significant hump-shaped pattern, with the highest values were observed in the latitudes of absolute 30° - 50° (Fig. 5a). The NOSC (R^2 = 0.019, P < 0.001) was significantly and monotonically increased along latitudinal gradients (Fig. 5a). These latitudinal patterns also showed in specific habitats especially oceans, river water, and reservoir water (Fig. 5b). For example, the DBE, AI_{mod}, and NOSC showed strongly hump-shaped pattern in the river water (Fig. 5b). The mass, DBE, and AI_{mod} showed significant U-shaped patterns in marine water, and the mass and DBE were monotonically decreasing in marine sediment (Fig. 5b). In the land, AI_{mod} showed a monotonically decreased pattern (R^2 = 0.129, P < 0.001) along the latitudinal gradient, while all the other properties had a nonsignificant (P > 0.05) pattern.

Drivers of DOM properties in waters and land systems

The distribution patterns of DOM properties of mass, DBE, AI_{mod}, and NOSC were significantly driven by geographical, environmental, and climatic variables in waters and land systems, indicated by random forest model (Figs. 6, S3). For the synthetic properties of DOM properties indicated by the first two axes of principal coordinates (PCoA1 and PCoA2), showed that mean annual temperature had a significant effect on DOM PCoA1, followed by geographic variables and pH (Fig. S3a). For the individual DOM properties, we found that DBE, AI_{mod}, and NOSC were most sensitive to geographical variables, while mass was major explained by mean annual temperature (Fig. 6a).

Linear mixed model further confirmed the relative importance of environmental (i.e., pH) and geographical variables on DOM properties, and extremes of climatic factors showed stronger affects compared to mean annual climates (Fig. 6b, S3, and Table S4). For the waters, pH and mean annual precipitation had the strongest effects on PCoA1 (R^2 = 0.806, P < 0.01) and PCoA2 (R^2 = 0.749, P < 0.05) of DOM properties, respectively (Fig. S3b). Geographical and climatic variables, such as elevation and mean temperature of warmest quarter, also strongly influenced on DOM properties, with the explained variations of 0.454 and 0.421, respectively (Fig. S3b). Moreover, extremes of climatic variables showed highly significant relationships with bioavailability properties, such as isothermality and precipitation of wettest month (Fig. 6b). The mass and NOSC were significantly related to pH (R^2 = 0.706, P < 0.05) and elevation (R^2 = 0.615, P < 0.05), followed by latitude and extremes of temperature variables, such as temperature seasonality, min temperature of coldest month, mean temperature of warmest quarter, and mean annual temperature. For the land, mass, DBE, AI_{mod}, and NOSC were most significantly related to (P < 0.05) mean annual temperature, precipitation seasonality, pH, and temperature annual range, respectively (Fig. 6b). For example, bioavailability properties (e.g., DBE and AI_{mod}) were dominantly affected by precipitation of bioclimatic variables, while oxidation state (e.g., NOSC) were mainly affected by temperature of bioclimatic variables (Fig. 6b).

Discussion

DOM properties across Earth ecosystems indicated by molecular indices

Our findings indicated the varying roles of DOM indices in defining the DOM properties across systems and habitats. Comparatively, mass size and bioavailability properties (e.g., DBE and AI_{mod}) of DOM in waters were significantly higher than in land, however NOSC values were similar between the two habitats (Fig. 2). This indicates that DOM contains less labile organic matter and a higher abundance of recalcitrant molecules in waters (D'Andrilli et al. 2015). When considering individual habitat of waters and land, we observed different DOM properties for the four molecular indices. For the waters, mass and bioavailability properties showed the greatest value in specific habitats as river sediment, indicating a lower abundance of more hydrogen saturated molecules (Hu et al. 2024a). This is mainly attributed to the fact that the primary source of DOM in rivers is terrigenous input (Wagner et al. 2015, Johnston et al. 2021, He et al. 2023). Oxidation state (e.g., NOSC) peaked at wastewater, which is consistent with previous works (Yang et al. 2022, Chen et al. 2023), as hydroxyl radical could enhance oxygenation, oxidative deamination, decyclopropyl, and deisopropyl reactions in wastewater treatment processes, reducing unsaturation and aromaticity of molecules behind and thus high oxidation and aliphatic content.

For the land, the highest bioavailability properties were observed in the order of cropland > forest soil/grassland > peatland/paddy soil, where disparate soils generally tended towards a consistent ranking of bioavailability properties (He et al. 2023). DOM associated with cropland and grassland inputs was dominated by lignin-like species, and forest soil DOM primarily originates from the degradation products of woody plants, indicating these habitats contain more high recalcitrant molecules (O'Donnell et al. 2016, Ge et al. 2022, Sheng et al. 2023). Paddy soil DOM show the lowest DBE, where fertilization causes more bioavailable fractions with low condensation (Li et al. 2018). We integrated datasets of primary DOM indices such as mass, DBE, AI_{mod}, and NOSC as well as H/C and O/C ratios on a global scale which provides a comprehensive project for comparing DOM properties from molecular perspective along the aquatic-terrestrial continuum.

Investigation within the properties revealed correlations between mass size, bioavailability properties, and oxidation states of DOM (Fig. 3). Firstly, DBE and AI_{mod} showed highly negative correlation with H/C, and NOSC showed positive correlation with O/C in both systems and individual habitats, indicating that molecular properties were closely related to the bio-recalcitrance of DOM (Hildebrand et al. 2022, Cai and Jiao 2023). Secondly, mass was strongly positively correlated with DBE in both systems, and negatively correlated with H/C, indicating that the most recalcitrant DOM generally accompanies with a low molecular weight, which is consistent with previous observations (Benner and Amon 2015, Li et al. 2019, Zheng et al. 2019). Lastly, compared to water systems, land systems exhibited higher oxidation states (e.g., O/C and NOSC) with higher molecular mass, suggesting a size-reactivity continuum, which might emerge from a complex reaction cascade, such as litter decomposition to microbial biomass and DOM (Hertkorn et al. 2006, Roth et al. 2014).

Variations of molecular indices controlling global DOM

Molecular indices covaried in predictable ways within systems despite high variability in DOM properties among the systems (Figs. 4, S2). This finding also further validates the linkages between molecular weight, bioavailability properties, and oxidation states. Principal components analysis showed that the correlation matrix of DOM properties was dominantly explained (R^2 = 0.76) by the first two axes. Collectively, the variables H/C, DBE, and AI_{mod} were separated on the first dimension, which related to the unsaturation degree (or aromaticity), indicating that differences in DOM properties between the systems was mainly caused by the bioavailability dimension. The bioavailability dimension is associated with the change of recalcitrant molecules, which is caused by carbon metabolism and microbial processing of organic matter (Valle et al. 2020, Wang et al. 2021, Hu et al. 2022b). The variables O/C, NOSC, and mass were separated on the second dimension, suggesting that the oxidation state also contributed to the variations in DOM properties among systems. DOM properties in waters are dominated by high molecular weight and highly aromaticity (Frey et al. 2016). Notably, mass size and oxidation states showed the greater loadings in land in comparison to waters, and were higher in river sediment than river water. The shift of molecular mass intensities is strongly related to the degradation state of DOM in grassland soils (Roth et al. 2019) and forest soils (Benk et al. 2018). DOM is degraded at the molecular level along a gradient from high to low NOSC (Kellerman et al. 2015). DOM aromaticity significantly improves through degradation processes as mass decreases (Chen et al. 2021). Thus, although we utilized different ecosystems with molecular indices, the DOM properties are clearly covariant within the DOM samples.

Relationships of DOM properties to latitudes

DOM properties had predictable latitudinal patterns across Earth systems especially in waters. Mass, DBE, and AI_{mod} showed the highest values in mid-latitude waters regions, but a nonsignificant ($P > 0.05$) pattern was observed for mass (Fig. 5). These patterns also showed in individual habitats as ocean, river, and reservoir. The importance by latitude driven trend in DOM properties has been reported such as for lakes (Roth et al. 2013). The magnitude of aromatic functional groups for DOM shows an increasing pattern along latitudinal gradients in the Baltic sea, which implies that more organic contaminants at the high-latitude region (Ripszam et al. 2015). The abundance of protein-like components was relatively lower towards higher latitudes lakes, but humic-like component show positively pattern (Zhu et al. 2019). By contrast, DOM oxidation states significantly increase towards high latitudes in waters, indicating the conversion of natural aliphatic compounds to aromatic compounds from low to high latitudes (Kellerman et al. 2014).

For land, DOM bioavailability properties (i.e., AI_{mod}) showed a significantly negative correlation with latitude, while oxidation state (i.e., NOSC) had a weak positive correlation. Previous studies show that soil organic carbon generally enriches in shallow soil horizons of high-latitude areas (Hugelius et al. 2014, Krachler and Krachler 2021), with a monotonically latitudinal patterns on DOM properties (Li et al. 2022, Lin et al. 2023). For instance, DOM in polar regions generally show the low contribution of aromatic species due to high preservation degree of organic matter (Ward and Cory 2015, Zherebker et al. 2020). Our results reveal predictable patterns of DOM properties across ecosystems, and DOM had more unsaturated and aromatic in the mid-latitudes of 30°-50°, but less oxidation at the equator regions.

Linkage between environmental variables and DOM properties across systems

Association between DOM properties and environmental variables and climate, especially pH and mean annual temperature was observed (Fig. 6, S3). This is supported by previous studies cementing the importantly effect of ecosystem in DOM properties and biodegradation, such as geographical locations (Zhou et al. 2018, Peralta-Maraver et al. 2021), climate (Kellerman et al. 2014, Follstad Shah et al. 2017, Hu et al. 2024b), salinity (Yang et al. 2020), and acidity (Hyung and Kim 2008, Groeneveld et al. 2022). Climate changes could cause temperature, precipitation, and dust flux to shift DOM properties, including lakes (Mladenov et al. 2011), river (van Vliet et al. 2023), and groundwater (McDonough et al. 2020). DOM properties sensitively respond to pH as it consists of various functional groups susceptive to protons (Chen et al. 2019).

Besides, we also found that the importance of climate extremes on DOM properties. Compared to land, mass and bioavailability properties (e.g., DBE and AI_{mod}) were more closely related to extremes of precipitation, while oxidation state (e.g., NOSC) was dominantly affected by extremes of temperature. The molecular complexity, aromaticity, and highly unsaturated compounds of DOM are greater during the flood period compared with the dry period (Pang et al. 2021). Drought conditions can intensity DOM transformation by delaying residence time of water systems especially summer (Ejarque et al. 2018), river water quality deteriorates under hydroclimate extremes (van Vliet et al. 2023). Extremes of climatic variables are key to comprehending the effect of climate change on organic carbon characteristics (Hu et al. 2024a). Our findings indicated the roles of climatic factors in explaining DOM properties especially extreme climatic conditions under future global carbon processes.

Conclusion

Based on a global survey, we comprehensively revealed the bioavailability and environmental drivers of DOM properties across Earth systems. Mass size and bioavailability properties were found to be significantly higher in waters, with the largest values in specific water habitats such as river sediment, while oxidation state peaked for wastewater. Compared to waters, mass displayed strong positive correlations with oxidation state in land system, while NOSC showed stronger relations to bioavailability properties. Bioavailability indices especially H/C and AI_{mod} contributed most to variations in DOM properties. DBE and AI_{mod} showed significant hump-shaped patterns indicating peaked unsaturation and aromaticity at mid-latitudes of approximately 30°-50° in waters, while NOSC monotonically increased towards high latitudes. Climate and environmental variables such as annual mean temperature and pH showed significant correlations with DOM properties. These findings are pivotal in our understanding of the characteristics and distribution patterns of DOM at a global scale and the roles of climatic and environmental variables underlying global carbon across systems.

Author contributions

LH: Writing – original draft, Data extraction, Software, Visualization. **HL**: Writing review & editing. **XC**: Data extraction. **JS**: Writing review & editing. **AH:** Conceptualization, Methodology, Writing review & editing. **JW**: Conceptualization, Supervision, Writing review & editing, Funding acquisition.

Conflict of interests

The authors declare no conflict of interest.

References

[1] Bahureksa, W., Tfaily, M. M., Boiteau, R. M., Young, R. B., Logan, M. N., McKenna, A. M., and Borch, T. 2021. Soil Organic Matter Characterization by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FTICR MS): A Critical Review of Sample Preparation, Analysis, and Data Interpretation. Environ Sci Technol **55**:9637-9656.

[2] Begum, M. S., Park, J. H., Yang, L., Shin, K. H., and Hur, J. 2023. Optical and molecular indices of dissolved organic matter for estimating biodegradability and resulting carbon dioxide production in inland waters: A review. Water Res **228**:119362.

[3] Benk, S. A., Li, Y., Roth, V.-N., and Gleixner, G. 2018. Lignin Dimers as Potential Markers for 14C-young Terrestrial Dissolved Organic Matter in the Critical Zone. Frontiers in Earth Science **6**.

[4] Benner, R., and Amon, R. M. 2015. The size-reactivity continuum of major bioelements in the ocean. Ann Rev Mar Sci **7**:185-205.

[5] Boye, K., Noël, V., Tfaily, M. M., Bone, S. E., Williams, K. H., Bargar, John R., and Fendorf, S. 2017. Thermodynamically controlled preservation of organic carbon in floodplains. Nature Geoscience **10**:415-419.

[6] Breiman, L. 2001. Random Forests. Machine Learning **45**:5-32.

[7] Butturini, A., Herzsprung, P., Lechtenfeld, O. J., Alcorlo, P., Benaiges-Fernandez, R., Berlanga, M., Boadella, J., Freixinos Campillo, Z., Gomez, R. M., Sanchez-Montoya, M. M., Urmeneta, J., and Romani, A. M. 2022. Origin, accumulation and fate of dissolved organic matter in an extreme hypersaline shallow lake. Water Res **221**:118727.

[8] Butturini, A., Herzsprung, P., Lechtenfeld, O. J., Venturi, S., Amalfitano, S., Vazquez, E., Pacini, N., Harper, D. M., Tassi, F., and Fazi, S. 2020. Dissolved organic matter in a tropical saline-alkaline lake of the East African Rift Valley. Water Res **173**:115532.

[9] Cai, R., and Jiao, N. 2023. Recalcitrant dissolved organic matter and its major production and removal processes in the ocean. Deep Sea Research Part I: Oceanographic Research Papers **191**.

[10] Catalán, N., Marcé, R., Kothawala, D. N., and Tranvik, L. J. 2016. Organic carbon decomposition rates controlled by water retention time across inland waters. Nature Geoscience **9**:501-504.

[11] Chen, Q., Chen, F., Gonsior, M., Li, Y., Wang, Y., He, C., Cai, R., Xu, J., Wang, Y., Xu, D., Sun, J., Zhang, T., Shi, Q., Jiao, N., and Zheng, Q. 2021. Correspondence between DOM molecules and microbial community in a subtropical coastal estuary on a spatiotemporal scale. Environ Int **154**:106558.

[12] Chen, W., Gu, Z., He, C., and Li, Q. 2023. Molecular Characteristics and Formation Mechanisms of Unknown Ozonation Byproducts during the Treatment of Flocculated Nanofiltration Leachate Concentrates Using O(3) and UV/O(3) Processes. Environ Sci Technol **57**:20349-20359.

[13] Chen, W., Teng, C. Y., Qian, C., and Yu, H. Q. 2019. Characterizing Properties and Environmental Behaviors of Dissolved Organic Matter Using Two-Dimensional Correlation Spectroscopic Analysis. Environ Sci Technol **53**:4683-4694.

[14] Creed, I. F., Bergstrom, A. K., Trick, C. G., Grimm, N. B., Hessen, D. O., Karlsson, J., Kidd, K. A., Kritzberg, E., McKnight, D. M., Freeman, E. C., Senar, O. E., Andersson, A., Ask, J., Berggren, M., Cherif, M., Giesler, R., Hotchkiss, E. R., Kortelainen, P., Palta, M. M., Vrede, T., and Weyhenmeyer, G. A. 2018. Global change-driven effects on dissolved organic matter composition: Implications for food webs of northern lakes. Glob Chang Biol **24**:3692-3714.

[15] D'Andrilli, J., Cooper, W. T., Foreman, C. M., and Marshall, A. G. 2015. An ultrahigh-resolution mass spectrometry index to estimate natural organic matter lability. Rapid Commun Mass Spectrom **29**:2385-2401.

[16] D’Andrilli, J., Junker, J. R., Smith, H. J., Scholl, E. A., and Foreman, C. M. 2019. DOM composition alters ecosystem function during microbial processing of isolated sources. Biogeochemistry **142**:281-298.

[17] Delgado-Baquerizo, M., Guerra, C. A., Cano-Díaz, C., Egidi, E., Wang, J.-T., Eisenhauer, N., Singh, B. K., and Maestre, F. T. 2020. The proportion of soil-borne pathogens increases with warming at the global scale. Nature Climate Change **10**:550-554.

[18] DiDonato, N., Chen, H., Waggoner, D., and Hatcher, P. G. 2016. Potential origin and formation for molecular components of humic acids in soils. Geochimica et Cosmochimica Acta **178**:210-222.

[19] Dittmar, T., Lennartz, S. T., Buck-Wiese, H., Hansell, D. A., Santinelli, C., Vanni, C., Blasius, B., and Hehemann, J.-H. 2021. Enigmatic persistence of dissolved organic matter in the ocean. Nature Reviews Earth & Environment.

[20] Dittmar, T., and Stubbins, A. 2014. 12.6—Dissolved Organic Matter in Aquatic Systems. Pages 125-156 *in* T. Dittmar and Stubbins, A., editors. Treatise on Geochemistry.

[21] Drake, T. W., Raymond, P. A., and Spencer, R. G. M. 2017. Terrestrial carbon inputs to inland waters: A current synthesis of estimates and uncertainty. Limnology and Oceanography Letters **3**:132-142.

[22] Ejarque, E., Khan, S., Steniczka, G., Schelker, J., Kainz, M. J., and Battin, T. J. 2018. Climate-induced hydrological variation controls the transformation of

dissolved organic matter in a subalpine lake. Limnology and Oceanography **63**:1355-1371.

[23] Follstad Shah, J. J., Kominoski, J. S., Ardon, M., Dodds, W. K., Gessner, M. O., Griffiths, N. A., Hawkins, C. P., Johnson, S. L., Lecerf, A., LeRoy, C. J., Manning, D. W. P., Rosemond, A. D., Sinsabaugh, R. L., Swan, C. M., Webster, J. R., and Zeglin, L. H. 2017. Global synthesis of the temperature sensitivity of leaf litter breakdown in streams and rivers. Glob Chang Biol **23**:3064-3075.

[24] Frey, K. E., Sobczak, W. V., Mann, P. J., and Holmes, R. M. 2016. Optical properties and bioavailability of dissolved organic matter along a flow-path continuum from soil pore waters to the Kolyma River mainstem, East Siberia. Biogeosciences **13**:2279-2290.

[25] Galecki, A., and Burzykowski, T. 2013. Linear Mixed Effects Models Using R. : A Step-by-Step Approach.

[26] Ge, J., Qi, Y., Li, C., Ma, J., Yi, Y., Hu, Q., Mostofa, K. M. G., Volmer, D. A., and Li, S. L. 2022. Fluorescence and molecular signatures of dissolved organic matter to monitor and assess its multiple sources from a polluted river in the farming-pastoral ecotone of northern China. Sci Total Environ **837**:154575.

[27] Groeneveld, M., Catalan, N., Einarsdottir, K., Bravo, A. G., and Kothawala, D. N. 2022. The influence of pH on dissolved organic matter fluorescence in inland waters. Anal Methods **14**:1351-1360.

[28] Hansell, D., Carlson, C., Repeta, D., and Schlitzer, R. 2009. Dissolved Organic Matter in the Ocean: A Controversy Stimulates New Insights. Oceanography **22**:202-211.

[29] He, C., Yi, Y., He, D., Cai, R., Chen, C., and Shi, Q. 2023. Molecular composition of dissolved organic matter across diverse ecosystems: Preliminary implications for biogeochemical cycling. J Environ Manage **344**:118559.

[30] Hertkorn, N., Benner, R., Frommberger, M., Schmitt-Kopplin, P., Witt, M., Kaiser, K., Kettrup, A., and Hedges, J. I. 2006. Characterization of a major refractory component of marine dissolved organic matter. Geochimica et Cosmochimica Acta **70**:2990-3010.

[31] Hildebrand, T., Osterholz, H., Bunse, C., Grotheer, H., Dittmar, T., and Schupp, P. J. 2022. Transformation of dissolved organic matter by two Indo-Pacific sponges. Limnology and Oceanography **67**:2483-2496.

[32] Hu, A., Choi, M., Tanentzap, A. J., Liu, J., Jang, K. S., Lennon, J. T., Liu, Y., Soininen, J., Lu, X., Zhang, Y., Shen, J., and Wang, J. 2022a. Ecological networks of dissolved organic matter and microorganisms under global change. Nat Commun **13**:3600.

[33] Hu, A., Han, L., Lu, X., Zhang, G., and Wang, J. 2024a. Global patterns and drivers of dissolved organic matter across Earth systems: Insights from H/C and O/C ratios. Fundamental Research.

[34] 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., and Wang, J. 2024b. Thermal responses of dissolved organic matter under global change. Nature Communications **15**.

[35] Hu, A., Jang, K. S., Meng, F. F., Stegen, J., Tanentzap, A. J., Choi, M., Lennon, J.

T., Soininen, J., and Wang, J. J. 2022b. Microbial and Environmental Processes Shape the Link between Organic Matter Functional Traits and Composition. Environmental Science & Technology **56**:10504-10516.

[36] Hugelius, G., Strauss, J., Zubrzycki, S., Harden, J. W., Schuur, E. A. G., Ping, C. L., Schirrmeister, L., Grosse, G., Michaelson, G. J., Koven, C. D., O'Donnell, J. A., Elberling, B., Mishra, U., Camill, P., Yu, Z., Palmtag, J., and Kuhry, P. 2014. Estimated stocks of circumpolar permafrost carbon with quantified uncertainty ranges and identified data gaps. Biogeosciences **11**:6573-6593.

[37] Hyung, H., and Kim, J. H. 2008. Natural organic matter (NOM) adsorption to multi-walled carbon nanotubes: effect of NOM characteristics and water quality parameters. Environ Sci Technol **42**:4416-4421.

[38] Jiao, N., Herndl, G. J., Hansell, D. A., Benner, R., Kattner, G., Wilhelm, S. W., Kirchman, D. L., Weinbauer, M. G., Luo, T., Chen, F., and Azam, F. 2010. Microbial production of recalcitrant dissolved organic matter: long-term carbon storage in the global ocean. Nat Rev Microbiol **8**:593-599.

[39] Johnston, S. E., Carey, J. C., Kellerman, A., Podgorski, D. C., Gewirtzman, J., and Spencer, R. G. M. 2021. Controls on Riverine Dissolved Organic Matter Composition Across an Arctic-Boreal Latitudinal Gradient. Journal of Geophysical Research: Biogeosciences **126**.

[40] Kellerman, A. M., Dittmar, T., Kothawala, D. N., and Tranvik, L. J. 2014. Chemodiversity of dissolved organic matter in lakes driven by climate and hydrology. Nat Commun **5**:3804.

[41] Kellerman, A. M., Guillemette, F., Podgorski, D. C., Aiken, G. R., Butler, K. D., and Spencer, R. G. M. 2018. Unifying Concepts Linking Dissolved Organic Matter Composition to Persistence in Aquatic Ecosystems. Environ Sci Technol **52**:2538-2548.

[42] Kellerman, A. M., Kothawala, D. N., Dittmar, T., and Tranvik, L. J. 2015. Persistence of dissolved organic matter in lakes related to its molecular characteristics. Nature Geoscience **8**:454-457.

[43] Kim, S., Kramer, R. W., and Hatcher, P. G. 2003. Graphical method for analysis of ultrahigh-resolution broadband mass spectra of natural organic matter, the van Krevelen diagram. Anal Chem **75**:5336-5344.

[44] Kleber, M., Bourg, I. C., Coward, E. K., Hansel, C. M., Myneni, S. C. B., and Nunan, N. 2021. Dynamic interactions at the mineral–organic matter interface. Nature Reviews Earth & Environment **2**:402-421.

[45] Koch, B. P., and Dittmar, T. 2006. From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter. Rapid Communications in Mass Spectrometry **20**:926-932.

[46] Koch, B. P., and Dittmar, T. 2016. From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter. Rapid Communications in Mass Spectrometry **30**:250-250.

[47] Krachler, R., and Krachler, R. F. 2021. Northern High-Latitude Organic Soils As a Vital Source of River-Borne Dissolved Iron to the Ocean. Environ Sci Technol **55**:9672-9690.

[48] LaRowe, D. E., and Van Cappellen, P. 2011. Degradation of natural organic matter: A thermodynamic analysis. Geochimica et Cosmochimica Acta **75**:2030-2042.
[49] Laszakovits, J. R., and MacKay, A. A. 2022. Data-Based Chemical Class Regions for Van Krevelen Diagrams. J Am Soc Mass Spectrom **33**:198-202.
[50] Li, J., Wang, B., Yang, M., Li, W., Liu, N., Qi, Y., and Liu, C.-Q. 2022. Geographical constraints on chemodiversity of sediment dissolved organic matter in China's coastal wetlands. Applied Geochemistry **147**.
[51] Li, P., Tao, J., Lin, J., He, C., Shi, Q., Li, X., and Zhang, C. 2019. Stratification of dissolved organic matter in the upper 2000m water column at the Mariana Trench. Sci Total Environ **668**:1222-1231.
[52] Li, X. M., Sun, G. X., Chen, S. C., Fang, Z., Yuan, H. Y., Shi, Q., and Zhu, Y. G. 2018. Molecular Chemodiversity of Dissolved Organic Matter in Paddy Soils. Environ Sci Technol **52**:963-971.
[53] Lin, Q., Tian, Q., Liao, C., Yuan, X., Lu, M., and Liu, F. 2023. Accumulation of microbial residuals and lignin phenols in forest soils along the latitude. **PREPRINT (Version 1) available at Research Square**
[54] McDonough, L. K., Santos, I. R., Andersen, M. S., O'Carroll, D. M., Rutlidge, H., Meredith, K., Oudone, P., Bridgeman, J., Gooddy, D. C., Sorensen, J. P. R., Lapworth, D. J., MacDonald, A. M., Ward, J., and Baker, A. 2020. Changes in global groundwater organic carbon driven by climate change and urbanization. Nat Commun **11**:1279.
[55] Medeiros, P. M., Seidel, M., Powers, L. C., Dittmar, T., Hansell, D. A., and Miller, W. L. 2015. Dissolved organic matter composition and photochemical transformations in the northern North Pacific Ocean. Geophysical Research Letters **42**:863-870.
[56] Mladenov, N., Sommaruga, R., Morales-Baquero, R., Laurion, I., Camarero, L., Dieguez, M. C., Camacho, A., Delgado, A., Torres, O., Chen, Z., Felip, M., and Reche, I. 2011. Dust inputs and bacteria influence dissolved organic matter in clear alpine lakes. Nat Commun **2**:405.
[57] Moody, C. S. 2020. A comparison of methods for the extraction of dissolved organic matter from freshwaters. Water Res **184**:116114.
[58] Nagar, S., Antony, R., and Thamban, M. 2021. Extracellular polymeric substances in Antarctic environments: A review of their ecological roles and impact on glacier biogeochemical cycles. Polar Science **30**.
[59] Nakagawa, S., Schielzeth, H., and O'Hara, R. B. 2013. A general and simple method for obtainingR2from generalized linear mixed-effects models. Methods in Ecology and Evolution **4**:133-142.
[60] O'Donnell, J. A., Aiken, G. R., Butler, K. D., Guillemette, F., Podgorski, D. C., and Spencer, R. G. M. 2016. DOM composition and transformation in boreal forest soils: The effects of temperature and organic-horizon decomposition state. Journal of Geophysical Research: Biogeosciences **121**:2727-2744.
[61] Pang, Y., Wang, K., Sun, Y., Zhou, Y., Yang, S., Li, Y., He, C., Shi, Q., and He, D. 2021. Linking the unique molecular complexity of dissolved organic matter to flood period in the Yangtze River mainstream. Sci Total Environ **764**:142803.

[62] Peralta-Maraver, I., Stubbington, R., Arnon, S., Kratina, P., Krause, S., de Mello Cionek, V., Leite, N. K., da Silva, A. L. L., Thomaz, S. M., Posselt, M., Milner, V. S., Momblanch, A., Moretti, M. S., Nobrega, R. L. B., Perkins, D. M., Petrucio, M. M., Reche, I., Saito, V., Sarmento, H., Strange, E., Taniwaki, R. H., White, J., Alves, G. H. Z., and Robertson, A. L. 2021. The riverine bioreactor: An integrative perspective on biological decomposition of organic matter across riverine habitats. Sci Total Environ **772**:145494.

[63] Pracht, L. E., Tfaily, M. M., Ardissono, R. J., and Neumann, R. B. 2018. Molecular characterization of organic matter mobilized from Bangladeshi aquifer sediment: tracking carbon compositional change during microbial utilization. Biogeosciences **15**:1733-1747.

[64] Raymond, P. A., Hartmann, J., Lauerwald, R., Sobek, S., McDonald, C., Hoover, M., Butman, D., Striegl, R., Mayorga, E., Humborg, C., Kortelainen, P., Durr, H., Meybeck, M., Ciais, P., and Guth, P. 2013. Global carbon dioxide emissions from inland waters. Nature **503**:355-359.

[65] Ripszam, M., Paczkowska, J., Figueira, J., Veenaas, C., and Haglund, P. 2015. Dissolved organic carbon quality and sorption of organic pollutants in the Baltic Sea in light of future climate change. Environ Sci Technol **49**:1445-1452.

[66] Roth, V.-N., Dittmar, T., Gaupp, R., and Gleixner, G. 2013. Latitude and pH driven trends in the molecular composition of DOM across a north south transect along the Yenisei River. Geochimica et Cosmochimica Acta **123**:93-105.

[67] Roth, V.-N., Dittmar, T., Gaupp, R., and Gleixner, G. 2014. Ecosystem-Specific Composition of Dissolved Organic Matter. Vadose Zone Journal **13**:1-10.

[68] Roth, V.-N., Lange, M., Simon, C., Hertkorn, N., Bucher, S., Goodall, T., Griffiths, R. I., Mellado-Vázquez, P. G., Mommer, L., Oram, N. J., Weigelt, A., Dittmar, T., and Gleixner, G. 2019. Persistence of dissolved organic matter explained by molecular changes during its passage through soil. Nature Geoscience **12**:755-761.

[69] Schmidt, M. W., Torn, M. S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I. A., Kleber, M., Kogel-Knabner, I., Lehmann, J., Manning, D. A., Nannipieri, P., Rasse, D. P., Weiner, S., and Trumbore, S. E. 2011. Persistence of soil organic matter as an ecosystem property. Nature **478**:49-56.

[70] Sheng, M., Chen, S., Liu, C. Q., Fu, Q., Zhang, D., Hu, W., Deng, J., Wu, L., Li, P., Yan, Z., Zhu, Y. G., and Fu, P. 2023. Spatial and molecular variations in forest topsoil dissolved organic matter as revealed by FT-ICR mass spectrometry. Sci Total Environ **895**:165099.

[71] Singer, G. A., Fasching, C., Wilhelm, L., Niggemann, J., Steier, P., Dittmar, T., and Battin, T. J. 2012. Biogeochemically diverse organic matter in Alpine glaciers and its downstream fate. Nature Geoscience **5**:710-714.

[72] Speetjens, N. J., Tanski, G., Martin, V., Wagner, J., Richter, A., Hugelius, G., Boucher, C., Lodi, R., Knoblauch, C., Koch, B. P., Wünsch, U., Lantuit, H., and Vonk, J. E. 2022. Dissolved organic matter characterization in soils and streams in a small coastal low-Arctic catchment. Biogeosciences **19**:3073-3097.

[73] Stegen, J. C., Johnson, T., Fredrickson, J. K., Wilkins, M. J., Konopka, A. E.,

Nelson, W. C., Arntzen, E. V., Chrisler, W. B., Chu, R. K., Fansler, S. J., Graham, E. B., Kennedy, D. W., Resch, C. T., Tfaily, M., and Zachara, J. 2018. Influences of organic carbon speciation on hyporheic corridor biogeochemistry and microbial ecology. Nat Commun **9**:585.

[74] Szkokan-Emilson, E. J., Kielstra, B. W., Arnott, S. E., Watmough, S. A., Gunn, J. M., and Tanentzap, A. J. 2017. Dry conditions disrupt terrestrial-aquatic linkages in northern catchments. Glob Chang Biol **23**:117-126.

[75] Valle, J., Harir, M., Gonsior, M., Enrich-Prast, A., Schmitt-Kopplin, P., Bastviken, D., and Hertkorn, N. 2020. Molecular differences between water column and sediment pore water SPE-DOM in ten Swedish boreal lakes. Water Res **170**:115320.

[76] van Vliet, M. T. H., Thorslund, J., Strokal, M., Hofstra, N., Flörke, M., Ehalt Macedo, H., Nkwasa, A., Tang, T., Kaushal, S. S., Kumar, R., van Griensven, A., Bouwman, L., and Mosley, L. M. 2023. Global river water quality under climate change and hydroclimatic extremes. Nature Reviews Earth & Environment **4**:687-702.

[77] Wadham, J. L., Hawkings, J. R., Tarasov, L., Gregoire, L. J., Spencer, R. G. M., Gutjahr, M., Ridgwell, A., and Kohfeld, K. E. 2019. Ice sheets matter for the global carbon cycle. Nat Commun **10**:3567.

[78] Wagner, S., Riedel, T., Niggemann, J., Vahatalo, A. V., Dittmar, T., and Jaffe, R. 2015. Linking the Molecular Signature of Heteroatomic Dissolved Organic Matter to Watershed Characteristics in World Rivers. Environ Sci Technol **49**:13798-13806.

[79] Wagner, S., Schubotz, F., Kaiser, K., Hallmann, C., Waska, H., Rossel, P. E., Hansman, R., Elvert, M., Middelburg, J. J., Engel, A., Blattmann, T. M., Catalá, T. S., Lennartz, S. T., Gomez-Saez, G. V., Pantoja-Gutiérrez, S., Bao, R., and Galy, V. 2020. Soothsaying DOM: A Current Perspective on the Future of Oceanic Dissolved Organic Carbon. Frontiers in Marine Science **7**.

[80] Wang, W., Tao, J., Yu, K., He, C., Wang, J., Li, P., Chen, H., Xu, B., Shi, Q., and Zhang, C. 2021. Vertical Stratification of Dissolved Organic Matter Linked to Distinct Microbial Communities in Subtropic Estuarine Sediments. Front Microbiol **12**:697860.

[81] Ward, C. P., and Cory, R. M. 2015. Chemical composition of dissolved organic matter draining permafrost soils. Geochimica et Cosmochimica Acta **167**:63-79.

[82] Ward, C. P., and Cory, R. M. 2016. Complete and Partial Photo-oxidation of Dissolved Organic Matter Draining Permafrost Soils. Environ Sci Technol **50**:3545-3553.

[83] Ward, C. P., Nalven, S. G., Crump, B. C., Kling, G. W., and Cory, R. M. 2017. Photochemical alteration of organic carbon draining permafrost soils shifts microbial metabolic pathways and stimulates respiration. Nat Commun **8**:772.

[84] Yamaoka, K., Nakagawa, T., and Uno, T. 1978. Application of Akaike's information criterion (AIC) in the evaluation of linear pharmacokinetic equations. J Pharmacokinet Biopharm **6**:165-175.

[85] Yang, J., Jiang, H., Liu, W., Huang, L., Huang, J., Wang, B., Dong, H., Chu, R. K.,

and Tolic, N. 2020. Potential utilization of terrestrially derived dissolved organic matter by aquatic microbial communities in saline lakes. ISME J **14**:2313-2324.

[86] Yang, X., Rosario-Ortiz, F. L., Lei, Y., Pan, Y., Lei, X., and Westerhoff, P. 2022. Multiple Roles of Dissolved Organic Matter in Advanced Oxidation Processes. Environ Sci Technol.

[87] Zark, M., and Dittmar, T. 2018. Universal molecular structures in natural dissolved organic matter. Nat Commun **9**:3178.

[88] Zheng, Q., Chen, Q., Cai, R., He, C., Guo, W., Wang, Y., Shi, Q., Chen, C., and Jiao, N. 2019. Molecular characteristics of microbially mediated transformations of Synechococcus-derived dissolved organic matter as revealed by incubation experiments. Environ Microbiol **21**:2533-2543.

[89] Zherebker, A., Rukhovich, G. D., Sarycheva, A., Lechtenfeld, O. J., and Nikolaev, E. N. 2022. Aromaticity Index with Improved Estimation of Carboxyl Group Contribution for Biogeochemical Studies. Environ Sci Technol **56**:2729-2737.

[90] Zherebker, A., Shirshin, E., Rubekina, A., Kharybin, O., Kononikhin, A., Kulikova, N. A., Zaitsev, K. V., Roznyatovsky, V. A., Grishin, Y. K., Perminova, I. V., and Nikolaev, E. N. 2020. Optical Properties of Soil Dissolved Organic Matter Are Related to Acidic Functions of Its Components as Revealed by Fractionation, Selective Deuteromethylation, and Ultrahigh Resolution Mass Spectrometry. Environ Sci Technol **54**:2667-2677.

[91] Zhou, Y., Davidson, T. A., Yao, X., Zhang, Y., Jeppesen, E., de Souza, J. G., Wu, H., Shi, K., and Qin, B. 2018. How autochthonous dissolved organic matter responds to eutrophication and climate warming: Evidence from a cross-continental data analysis and experiments. Earth-Science Reviews **185**:928-937.

[92] Zhu, L., Zhao, Y., Bai, S., Zhou, H., Chen, X., and Wei, Z. 2019. New insights into the variation of dissolved organic matter components in different latitudinal lakes of northeast China. Limnology and Oceanography **65**:471-481.

Figures

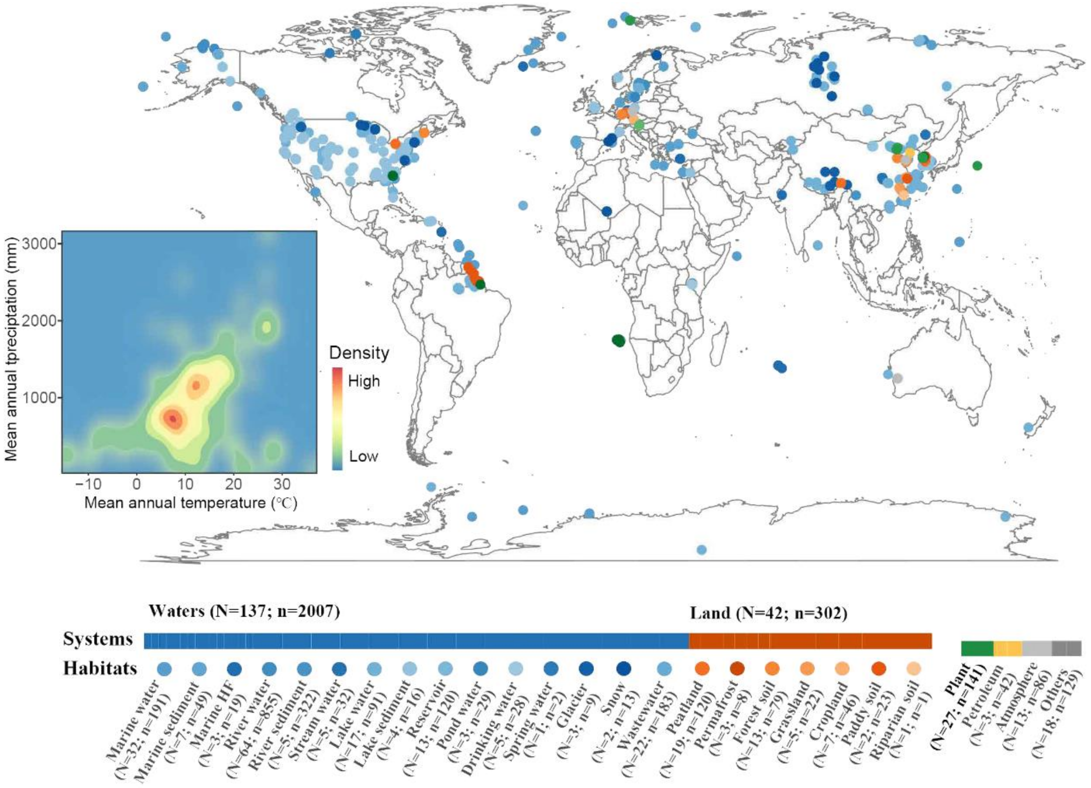

Figure 1

Map of samples included in the compiled dataset of mass, DBE, AI_{mod}, and NOSC of DOM. We collected a total of 2707 sites from 204 studies that measure four indices across Earth systems. Colored dots represent the geographic locations of DOM samples. Numbers of studies and samples correspond to (N) and (n), respectively. Insert figure shows the distribution density of DOM samples within temperature and precipitation.

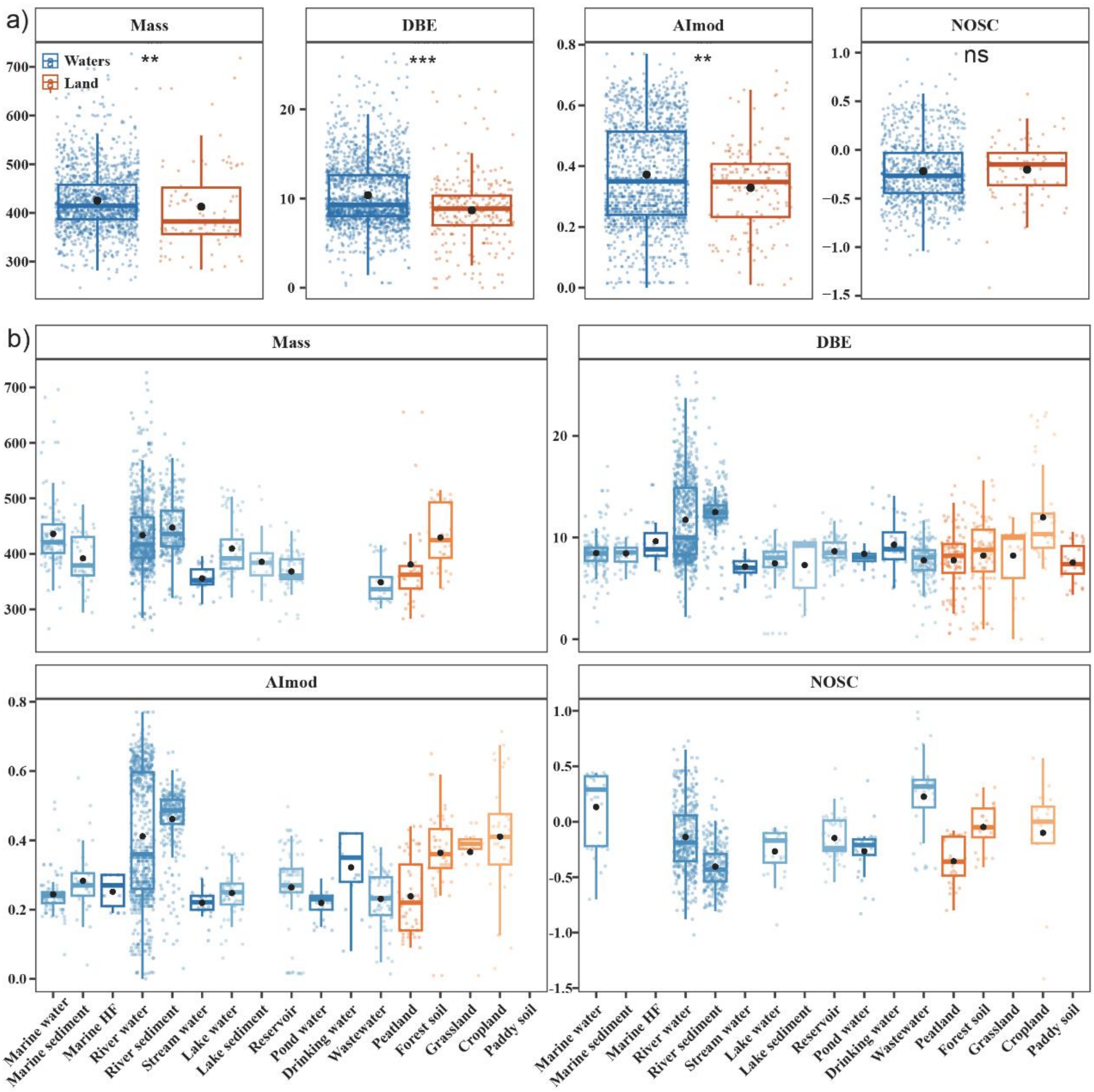

Figure 2

Variation of mass, DBE, AI_{mod}, and NOSC of DOM across systems and habitats. Boxplot shows the compositional-level mass, DBE, AI_{mod}, and NOSC for better comparisons among the systems (a) and habitats (b). Significances for pairwise comparisons by a Wilcoxon test are shown in Fig. S1. Colored dots and black dots of boxplots represent individual samples and average values in DOM indices, respectively. We included the habitats with the samples size over 15.

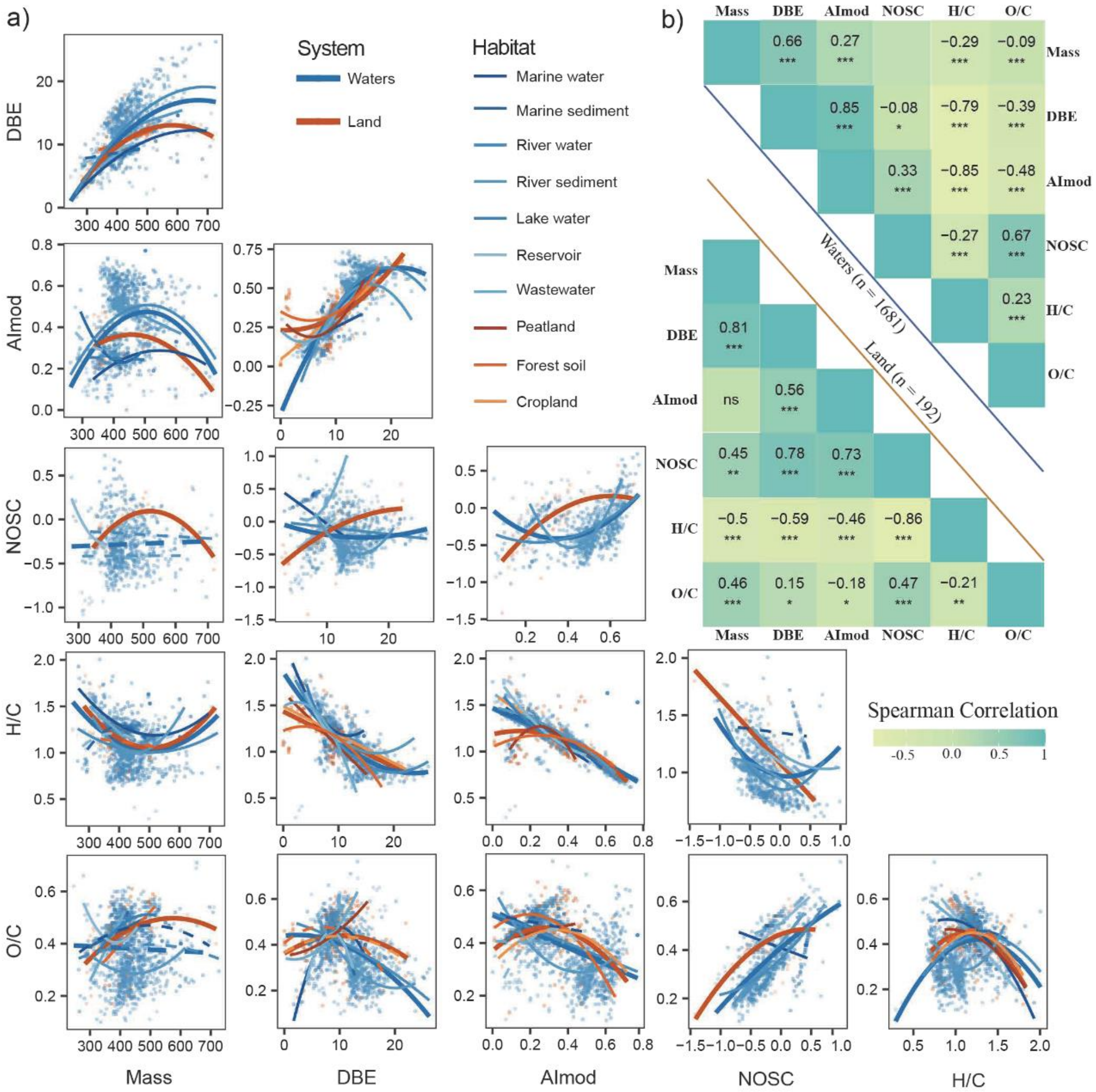

Figure 3

Indices relation of DOM across systems and habitats. (a) Linear regression analysis shows the linkages between mass size, bioavailability, and oxidation state by the significance of the correlation, and their relationships are estimated by solid ($P \leq 0.05$) or dotted ($P > 0.05$) lines. (b) The colors from yellow to green represent the Spearman correlation varying. We included the habitats with the samples size over 30.

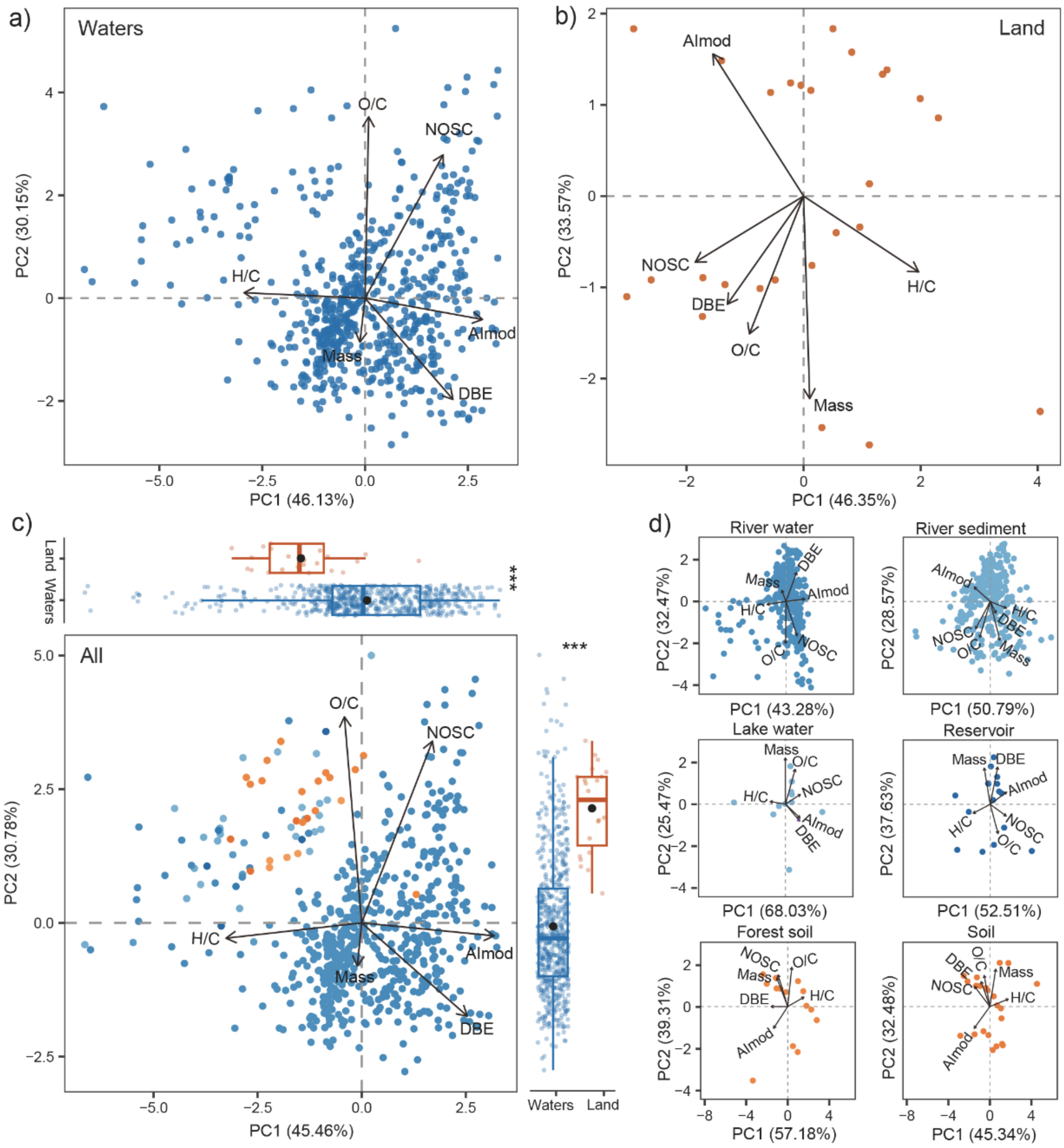

Figure 4

Association between six DOM indices at the molecular level. Principal component analysis plots show the variations of DOM indices across systems (a, b, and c) and habitats (d). Insert boxplot indicated the differences between waters and land by a Wilcox test, colored dots and black dots represent individual samples and average values in DOM properties, respectively. We included the habitats with the samples size over 15.

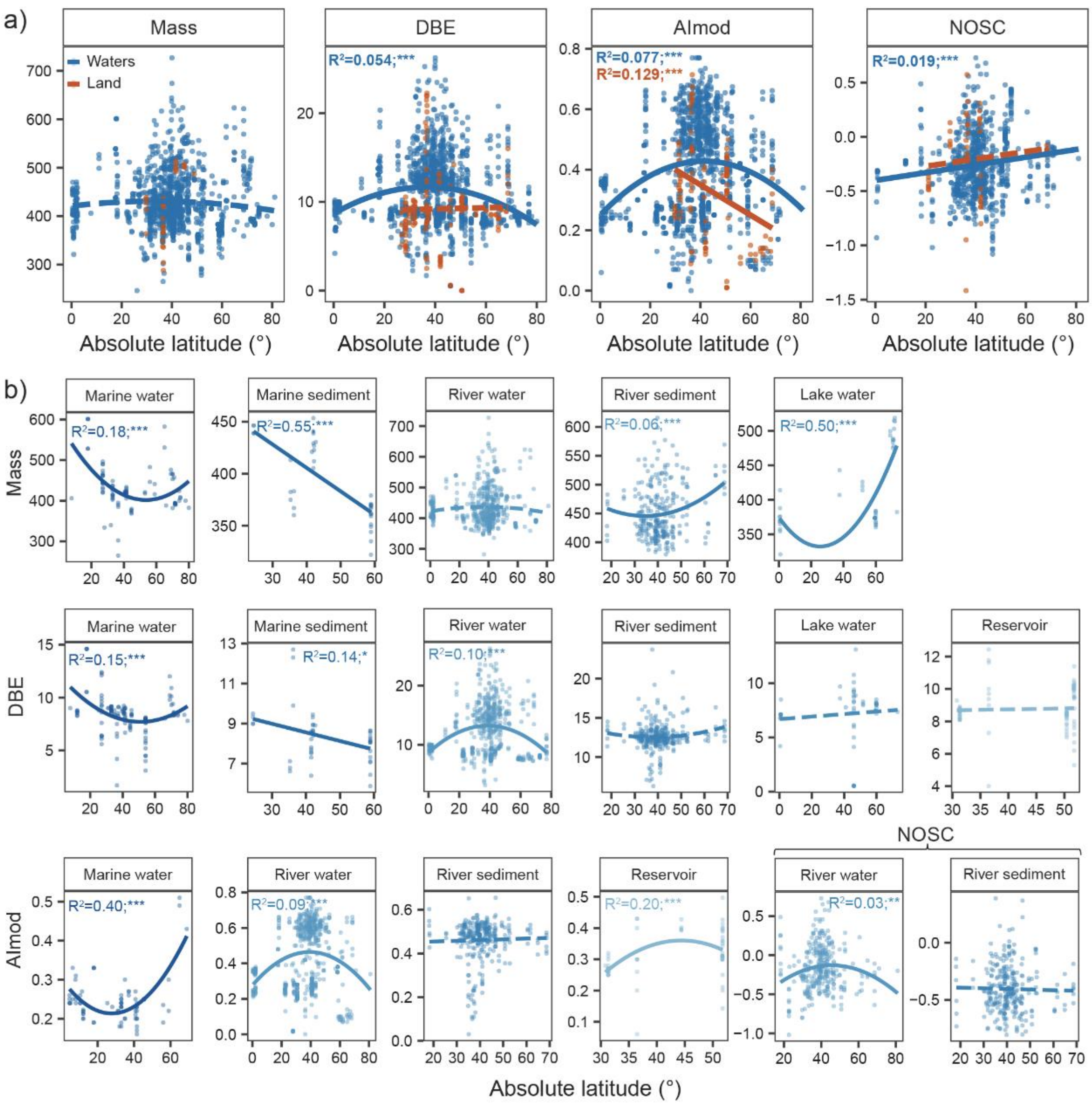

Figure 5

The distribution patterns of mass, DBE, AI_{mod}, and NOSC of DOM along latitudinal gradients. Latitudinal patterns are visualized with linear models across systems (a) and habitats (b), and the significant properties are indicated by asterisks (***, $P \leq 0.001$; **, $P \leq 0.01$; *, $P \leq 0.05$). North and South latitudes were converted as absolute latitudes.

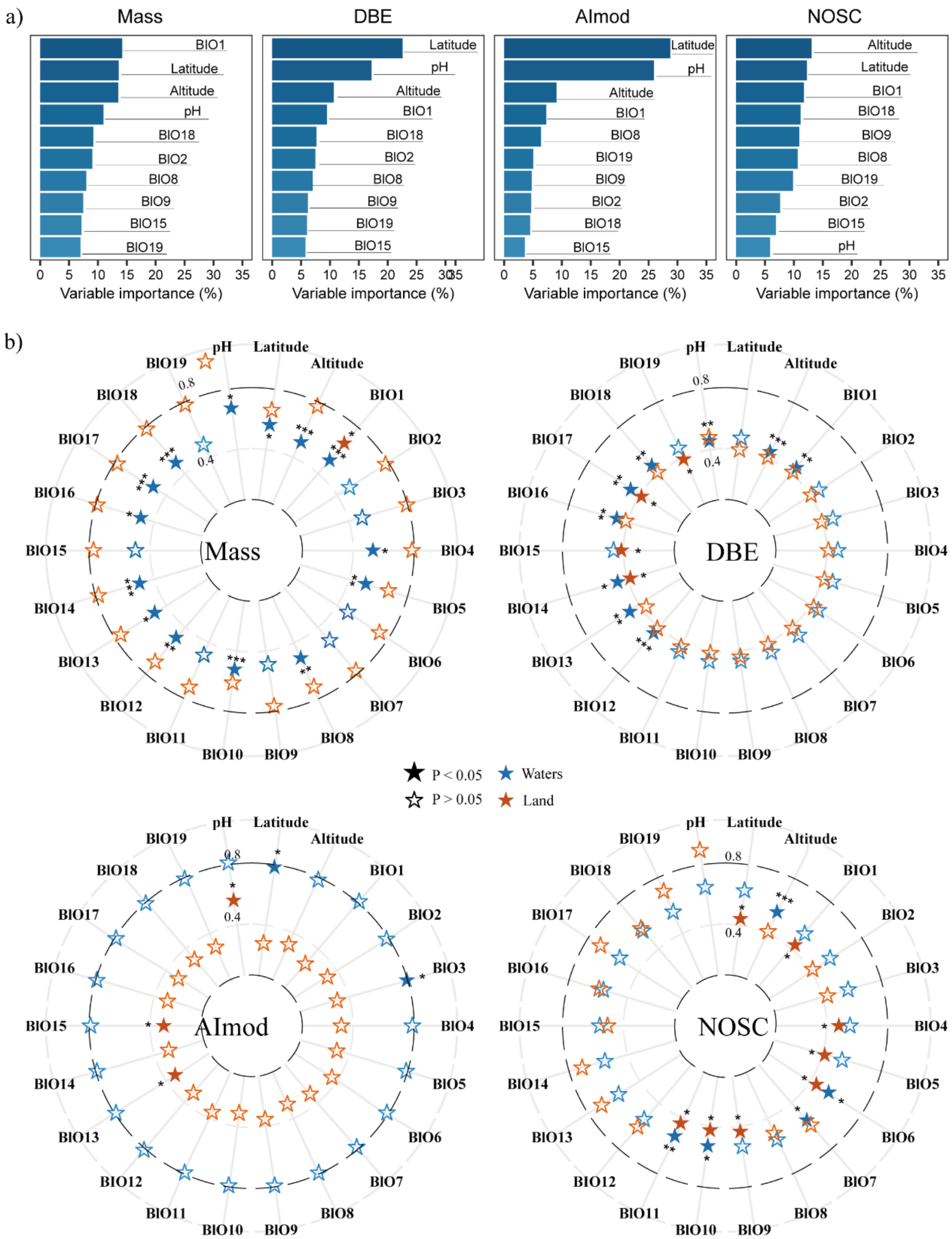

Figure 6

The influences of climatic and environmental variables on molecular indices of DOM. (a) Random forest analysis shows the relative importance of each explanatory variable on DOM properties at the molecular level. (b) DOM properties were examined with conditional explained heterogeneity (R^2) of linear mixed models for waters and land. Solid ($P \leq 0.05$) and open ($P > 0.05$) pentagram were indicated by the significant. We included the habitats with the samples size over 30.