

# UC Davis

## UC Davis Previously Published Works

### Title

Molecular signatures of soil-derived dissolved organic matter constrained by mineral weathering

### Permalink

<https://escholarship.org/uc/item/27z9c0g4>

### Journal

Fundamental Research, 3(3)

### ISSN

2096-9457

### Authors

Wang, Ying-Hui  
Zhang, Peng  
He, Chen  
et al.

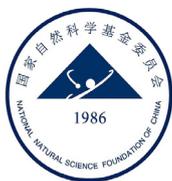
### Publication Date

2023-05-01

### DOI

10.1016/j.fmre.2022.01.032

Peer reviewed

Contents lists available at [ScienceDirect](#)

Fundamental Research

journal homepage: <http://www.keaipublishing.com/en/journals/fundamental-research/>

## Article

## Molecular signatures of soil-derived dissolved organic matter constrained by mineral weathering

Ying-Hui Wang<sup>a,b,c</sup>, Peng Zhang<sup>a,b</sup>, Chen He<sup>d</sup>, Jian-Chun Yu<sup>e</sup>, Quan Shi<sup>d</sup>, Randy A. Dahlgren<sup>f</sup>, Robert G.M. Spencer<sup>g</sup>, Zhi-Bing Yang<sup>c</sup>, Jun-Jian Wang<sup>a,b,\*</sup><sup>a</sup> State Environmental Protection Key Laboratory of Integrated Surface Water-Groundwater Pollution Control, School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China<sup>b</sup> Guangdong Provincial Key Laboratory of Soil and Groundwater Pollution Control, School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China<sup>c</sup> State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China<sup>d</sup> State Key Laboratory of Heavy Oil Processing, China University of Petroleum, Beijing 102249, China<sup>e</sup> Shanghai Engineering Research Center of Hadal Science and Technology, College of Marine Sciences, Shanghai Ocean University, Shanghai 201306, China<sup>f</sup> Department of Land, Air and Water Resources, University of California Davis, Davis 95616, United States<sup>g</sup> Department of Earth, Ocean and Atmospheric Science, Florida State University, Tallahassee 32306, United States

## ARTICLE INFO

## Article history:

Received 16 September 2021

Received in revised form 17 December 2021

Accepted 14 January 2022

Available online xxx

## Keywords:

Soil organic matter  
Dissolved organic matter  
Chemical composition  
Molecular signature  
Mineral weathering

## ABSTRACT

Dissolved organic matter (DOM) in soils drives biogeochemical cycling and soil functions in different directions depending on its molecular signature. Notably, there is a distinct paucity of information concerning how the molecular signatures of soil DOM vary with different degrees of weathering across wide geographic scales. Herein, we resolved the DOM molecular signatures from 22 diverse Chinese reference soils and linked them with soil organic matter and weathering-related mineralogical properties. The mixed-effects models revealed that the yields of DOM were determined by soil organic carbon content, whereas the molecular signature of DOM was primarily constrained by the weathering-related dimension. The soil weathering index showed a positive effect on the lability and a negative effect on the aromaticity of DOM. Specifically, DOM in highly weathered acidic soils featured more amino sugars, carbohydrates, and aliphatics, as well as less O-rich polyphenols and condensed aromatics, thereby conferring a higher DOM biolability and lower DOM aromaticity. This study highlights the dominance of the weathering-related dimension in constraining the molecular signatures and potential functions of DOM in soils across a wide geographic scale.

## 1. Introduction

Dissolved organic matter (DOM) in soil consists of a diverse mixture of water-soluble soil organic matter (SOM) molecules that are highly mobile and reactive [1,2]. These molecules play various roles with different capacities in biogeochemical processes involving carbon, nutrients and pollutants [3–6]. For example, polysaccharide-induced priming may stimulate microbial activity [7,8], whereas tannins may suppress the microbial decomposition process [9,10]. Theoretically, persistence of DOM molecules in soils is controlled by interactions between (i) DOM leaching and desorption that release DOM from plant residues and SOM and (ii) sorption and decomposition that remove DOM [2]. However, there is still no consensus on the dominant factor(s) or dimension(s) driving the yield and molecular signatures of DOM in soil environments.

Recent studies revealed that the microbe-mediated transformation and assembling processes constrained the vertical variations in DOM signatures within soil profiles [11,12] and attributed soil DOM with more recalcitrant signature [13], thereby maintaining relatively persistent DOM signatures in soils [11]. However, molecular variations of soil DOM from distinct geographical regions are primarily attributed to variations in geographical climate conditions and soil clay content [14]. Also, both geographical and edaphic factors can drive soil microbial diversity [15]. Soil weathering condition is highly related to geographical climate conditions, such as temperature and precipitation, and could be a reflection of the soil mineral characteristics [16,17]. Therefore, soil weathering condition could be a promising factor that might simplify the investigation of the mobilization process of DOM in soils. To date, it remains uncertain whether and how mineralogical properties (e.g., mineral elemental composition and weathering indicators) and SOM (e.g., SOM content and composition) may interact to constrain DOM yield and signatures across a diversity of soils. Although optical parameters of DOM based on absorbance and fluorescence spectroscopy

\* Corresponding author.

E-mail address: [wangjj@sustech.edu.cn](mailto:wangjj@sustech.edu.cn) (J.-J. Wang).<https://doi.org/10.1016/j.fmre.2022.01.032>2667-3258/© 2022 The Authors. Publishing Services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

have recently been linked to DOM molecular formulae in aquatic environments [18–20], their linkages to DOM molecular signatures in soil environments have not been rigorously investigated. Thus, improving our understanding of the molecular and optical signatures of DOM from diverse soils across a wide geographic scale, and their constraints, are important goals for unifying conceptual models of DOM dynamics and functions across global terrestrial ecosystems.

Herein, we studied the molecular and optical properties of soil-derived DOM from a set of reference soils with diverse climatic/weathering gradients across China. The reference soils were specifically selected, instead of randomly collected, to represent a range of soil characteristics representative of the large geographical region of China. As the reference soils are commercially available, they allow future studies of DOM signatures and functions on a consistent set of soils. Those reference soils were classified into six soil taxonomic orders based on USDA Soil Taxonomy and represent a soil developmental/weathering sequence: Entisols < Inceptisols  $\approx$  Aridisols < Mollisols < Alfisols < Ultisols [21]. The detailed mineral elemental composition combined with the SOM composition provided a powerful approach to quantitatively examine how weathering-defined soil mineralogy and SOM properties affect the molecular signatures of soil-derived DOM and assess the molecular-level interpretation of the commonly used optical parameters for soil-derived DOM. Here, we aim to identify the hitherto unknown association between specific molecular classes and soil-derived DOM fluorescence, and to distinguish the different regulating effects of SOM and mineral elemental properties on DOM quantity and quality in soils across a large weathering gradient.

## 2. Methods

### 2.1. Soil preparation and characterization

We used 22 Chinese national standard soils (GBW07401–GBW07385) purchased from <https://www.gbw-china.com>. These soil samples were collected from 15 different provinces (Fig. S1) and had been homogenized by gridding for commercial supplies. Standard soil characterization metrics and analytical methodologies are provided in the standard soil certificates, and the selected metrics used in this paper are reported in Table S1. The selected standard soils had wide ranges of pH (5.4–9.4), soil organic carbon (SOC) content (0.3–1.8%), mineral elemental composition (e.g., SiO<sub>2</sub>, 32.7–74.7%; Al<sub>2</sub>O<sub>3</sub>, 10.3–29.3%; Fe<sub>2</sub>O<sub>3</sub>, 2.0–18.8%), and weathering degree (Si/Al ratio [22]: 1.9–12.1; chemical index of alteration (CIA) [23]: 53.5–98.4) (see Table S1 for details).

Soil organic carbon composition from different functional groups was assessed by solid-state <sup>13</sup>C nuclear magnetic resonance (<sup>13</sup>C NMR) spectroscopy. Briefly, 6 g soil was repeatedly washed with 7 treatments of 10% hydrofluoric acid. The residues were washed to a neutral pH and analyzed on a Bruker AVANCE III 600 spectrometer at a resonance frequency of 150.9 MHz. The 4 mm magic angle spinning probe packed with powdered samples spun at a rate of 12 kHz. The contact time was 4 ms, and the recycle delay was set at 2 s for measurement [24]. <sup>13</sup>C NMR spectra were divided into 7 regions: 0–45 ppm for alkyl C; 45–60 ppm for *N*-alkyl C and methoxyl groups; 60–95 ppm for *O*-alkyl C; 95–110 ppm for di-*O*-alkyl C; 110–145 ppm for aromatic C; 145–165 ppm for *O*-aromatic C; and 165–210 ppm for carboxyl and carbonyl groups [25–27]. Alkyl/*O*-alkyl C ratio and area ratio of 70–75 to 52–57 ppm was calculated to indicate the decomposition degree and bioavailability of soil organic matter [28].

### 2.2. DOM preparation and optical characterization

Soil DOM was extracted by shaking 10 g soil in 200 mL Milli-Q water at 200 rpm and 20 °C for 2 h. The suspension was filtered through 0.45

μm polyether sulfone membrane (Millipore) and stored at 4 °C prior to further analyses.

General water quality and optical characteristics were analyzed based on our previous report [29]. Specifically, pH was analyzed with a benchtop meter HQ440d and Intellical™ PHC201 pH probe. Dissolved organic carbon (DOC) and total dissolved nitrogen (TDN) concentrations were measured with a TOC-L CSH/CSN (Shimadzu, Japan), and the ratio of DOC to TDN (DOC/TDN) was calculated. Absorbance and three-dimension excitation-emission matrix (3D-EEM) fluorescence spectra were acquired by an Aqualog® spectrophotometer (Horiba, Japan).

We calculated the carbon-specific UV-absorbance at 254 nm (SUVA<sub>254</sub>) and the ratio of absorption at 250 and 365 nm (E2/E3 ratio) to indicate the DOM aromaticity and the apparent molecular weight, respectively [30,31]. Parallel factor (PARAFAC) analysis was used to decompose the fluorescence signal [32]. Two “humic-like” (C1 and C2) and one protein-like (C3) components explained 99.6% of the variance and were validated using split-half analysis with four splits. Fluorescence-based indices, including fluorescence index (FI), humification index (HIX), and biological index (BIX), were calculated to assess the DOM sources, humification degree, and the contribution of autochthonous DOM [33–35].

### 2.3. FT-ICR MS analysis

Soil DOM was solid-phase extracted (SPE) on a Bond Elut PPL (100 mg PPL in 3 mL cartridge; Agilent, USA) as described in Wang et al [36]. SPE extracts were analyzed on a 9.4 Tesla FT-ICR MS (Bruker, Germany) using standard methods described in He et al [37]. Specifically, the SPE DOM extracts with DOC concentration at about 50 mg/L were injected into ESI sources at a rate of 250 μL/h. The typical operating conditions for negative-ion ESI analysis were: 3.5 kV spray shield voltage, 4.0 kV capillary column introduced voltage, -320 V capillary column end voltage, and ions accumulated in the collision cell for 0.2 s then transferred into the ICR cell with a 1.2 ms time-of-flight. The ion transformation parameter for the quadrupole was optimized at *m/z* 300. The mass range was *m/z* 200–800. A total of 128 scans with 2 M word size were accumulated to enhance the signal-to-noise ratio. The spectra were calibrated with an internal calibration list that provided a mass accuracy of 0.2 ppm or higher throughout the mass range of interest. Peaks with a signal-to-noise ratio greater than 6 were assigned using in-house software. The mass accuracy window was set to 1.0 ppm. The elemental combination for each formula was set as: <sup>12</sup>C<sub>1–60</sub>, <sup>1</sup>H<sub>1–120</sub>, <sup>14</sup>N<sub>0–3</sub>, <sup>16</sup>O<sub>0–30</sub>, and <sup>32</sup>S<sub>0–1</sub>. In addition, the elemental formulae should meet the following criteria: (1) the number of H atoms must be at least 1/3 the number of C atoms and cannot exceed 2C + N + 2; (2) the sum of H and N atoms must be even (the nitrogen rule); and (3) the number of N or O atoms cannot exceed the number of C atoms. Formulae assigned with mass accuracy <1.0 ppm were categorized into four elemental groups (CHO, CHON, CHOS, and CHONS) and five biochemical groups into different groups based on elemental composition: lipid (O/C ≤ 0.6; H/C ≥ 1.32; N/C ≤ 0.126), protein (0.12 < O/C ≤ 0.6; 0.9 < H/C < 2.5; 0.126 ≤ N/C ≤ 0.7 or 0.6 < O/C ≤ 1.0; 1.2 < H/C < 2.5; 0.2 < N/C ≤ 0.7), amino sugar (O/C ≥ 0.61; H/C ≥ 1.45; 0.07 < N/C ≤ 0.2), carbohydrate (O/C ≥ 0.8; 1.65 ≤ H/C < 2.7; N = 0), and phytochemical (O/C ≤ 1.15; H/C < 1.32; N/C < 0.126) [38]. Modified aromaticity index (AI<sub>mod</sub>) was calculated to assess the degrees of aromaticity [39], and nominal oxidation state of carbon (NOSC) was calculated to indicate the relative nominal oxidation state of organic matter [18]. The percentage of labile compounds above the molecular lability boundary (%MLB<sub>L</sub>, H/C ≥ 1.5) was calculated to estimate the abundance of labile organic matter [40].

### 2.4. Data management and statistical analyses

We applied non-metric multidimensional scaling (NMDS) based on Bray-Curtis distance for permute fitting of molecular variables [15,18].

Optical indices were fit to the ordination to reveal molecular characteristics that covaried with those properties. Significance was assessed through a Monte Carlo test (999 permutations), and variables with a significance level of  $P < 0.1$  were shown. NMDS was calculated using the vegan package in R software. Spearman correlation was calculated between the sum-normalized intensity of peaks that were present in at least 90% of the samples and optical indices, SOM composition and elemental constituents [4,14,20]. Formulae with significant correlations ( $P < 0.1$ ) to other indexes were shown in van Krevelen plots. Linear correlations between multiple proxies were assessed using SPSS. Structural equation models (SEM) evaluated complex network relationships among soil properties and the mobilization of soil DOM. Before SEM analysis, we conducted principal components analysis based on  $^{13}\text{C}$  NMR results to generate a multivariate index with the loading on the first principal component axis representing the SOM composition. The first principal component associated with high alkyl C abundance and the low area ratio of 70–75/52–57 ppm explained 77% of the total variance. SEM analyses were performed by AMOS 21.0, and model fit was evaluated by model  $\chi^2$  test and the root-mean-square error of approximation. A database of measured parameters is available from the corresponding author upon request. All other data are included in Supplementary Information files.

### 3. Results and discussion

#### 3.1. DOC yield from soils driven by SOC content

DOC yield ranged from 0.1 to 1.3 mg-C/g-soil and was linearly correlated with SOC content ( $r = 0.87$ ,  $P < 0.01$ ), but not with the relative abundances of SOC fractions or soil elemental composition ( $P > 0.05$ ; Fig. S2). Based on the linear relationship, every 1% increase in SOC content resulted in an increase of approximately  $0.68 \pm 0.08$  mg DOC yield from 1 g soil. Although DOM polarity (e.g., abundance of carboxyl C in SOC) and pH should theoretically affect the solubility of SOC [2], DOC yield from soils was not correlated with either of these factors, but rather appeared to be strongly controlled by SOC content across our wide range of soil types.

#### 3.2. Linkages between molecular and optical characters of soil-derived DOM

FT-ICR MS analysis resolved 9418 molecular formulae across all samples. We used NMDS to explore linkages between molecular characteristics of DOM and widely used optical indices, such as SUVA<sub>254</sub>, E2/E3 ratio, FI, HIX, BIX and the fractions of fluorescent DOM components (%C1, %C2 and %C3; details in Fig. S3). Results showed that DOM from the less weathered alkaline soils (pH =  $7.8 \pm 0.4$ ; CIA =  $59.6 \pm 3.8$ ; Si/Al =  $8.6 \pm 1.1$ ) distributed at the negative end of the first dimension and had high SUVA<sub>254</sub>, HIX and %C1, as well as high aromaticity ( $\text{AI}_{\text{mod}}$ ) and high portion of condensed aromatic and polyphenolic compounds (Fig. 1). These findings indicate that the most abundant “terrestrial humic-like” fluorescent DOM (C1) in soils is highly aromatic, oxygen-rich and has a relatively high molecular weight, similar to findings from natural waters, such as boreal lakes [18], rivers [19] and seawater [20]. Conversely, DOM from the highly weathered acidic soils (pH =  $6.0 \pm 0.4$ ; CIA =  $91.2 \pm 5.9$ ; Si/Al =  $4.1 \pm 1.5$ ) was found at the positive end of the first dimension and had high %MLB<sub>L</sub>, as well as the high portion of peptide-like and aliphatic components. Those characteristics are consistent with high BIX that suggests a high relative contribution of microbial-derived DOM [4,40]. Therefore, DOM from the highly weathered acidic soils is more likely to be biolabile.

To better resolve the possible relationships of DOM molecules to the various fluorescent components, we investigated the relative abundances of fluorescent components with the normalized mass spectral peak intensities of 2087 formulae that occurred

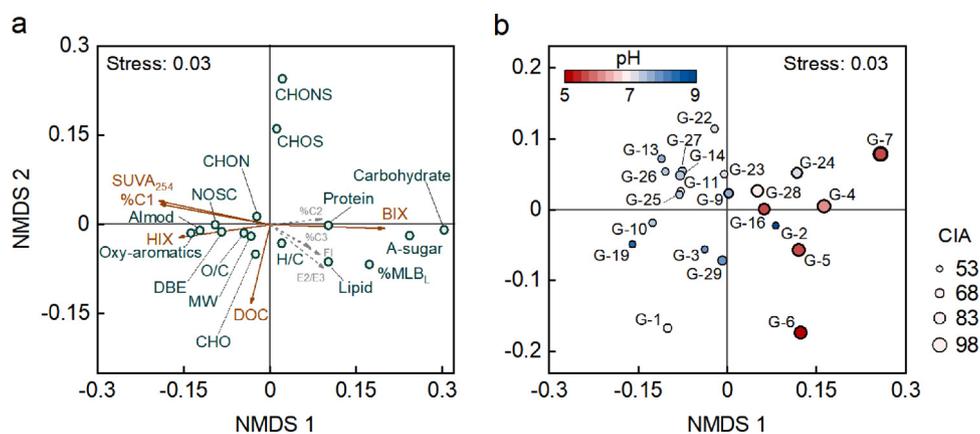
in 90% of soil samples (accounting for 64–75% of the total intensity for each sample). Based on the projection of Spearman correlation coefficients on van Krevelen plots (Fig. 2), the distribution of associated formulae in soils did not suggest a direct connection with the elemental or biochemical groups. Instead, the fluorescent-component-related formulae clustered in four regions: (i) “terrestrial-humic-like” C1 related formulae in Region I ( $\text{H/O} \leq 3$  and  $\text{AI}_{\text{mod}} \geq 0.1$ ), which contains large amounts of O-rich highly unsaturated compounds, polyphenols and condensed aromatics; (ii) “microbial-humic-like” C2 related formulae in Region II ( $\text{H/O} > 3$  and  $\text{AI}_{\text{mod}} \geq 0.1$ ), which contains large amounts of O-poor highly unsaturated aliphatics and aromatics; (iii) “microbial-humic-like” C2 related formulae in Region III ( $\text{H/O} \leq 3$  and  $\text{AI}_{\text{mod}} < 0.1$ ), which contains large amounts of amino sugars and carbohydrates; and (iv) “protein-like” C3 related formulae in Region IV ( $\text{H/O} > 3$  and  $\text{AI}_{\text{mod}} < 0.1$ ), which contains large amounts of saturated aliphatic and peptide-like formulae. After dividing the 2087 formulae into these four regions, we calculated the arithmetic mean numbers of the C, H, O, N, and S for each region, and the average formulae belonging to Regions I–IV were  $\text{C}_{16.71}\text{H}_{17.13}\text{O}_{8.12}\text{N}_{0.77}\text{S}_{0.00}$ ,  $\text{C}_{19.29}\text{H}_{25.73}\text{O}_{6.18}\text{N}_{0.34}\text{S}_{0.00}$ ,  $\text{C}_{14.48}\text{H}_{22.04}\text{O}_{9.33}\text{N}_{1.01}\text{S}_{0.01}$ , and  $\text{C}_{16.19}\text{H}_{27.18}\text{O}_{6.66}\text{N}_{0.80}\text{S}_{0.03}$ , respectively (Table S2). The relative abundances of the common formulae in the four regions strongly covaried with the relative abundances of total formulae clustered in the corresponding regions (Fig. S4). These four clusters showed distinct patterns contrary to fluorescent-component-related formulae clusters identified in lakes [18]. Importantly, this analysis highlights that similar fluorescent signals may have completely distinct molecular identities in DOM from terrestrial and aquatic environments.

#### 3.3. Linking DOM signatures to SOM and mineral elemental constraints

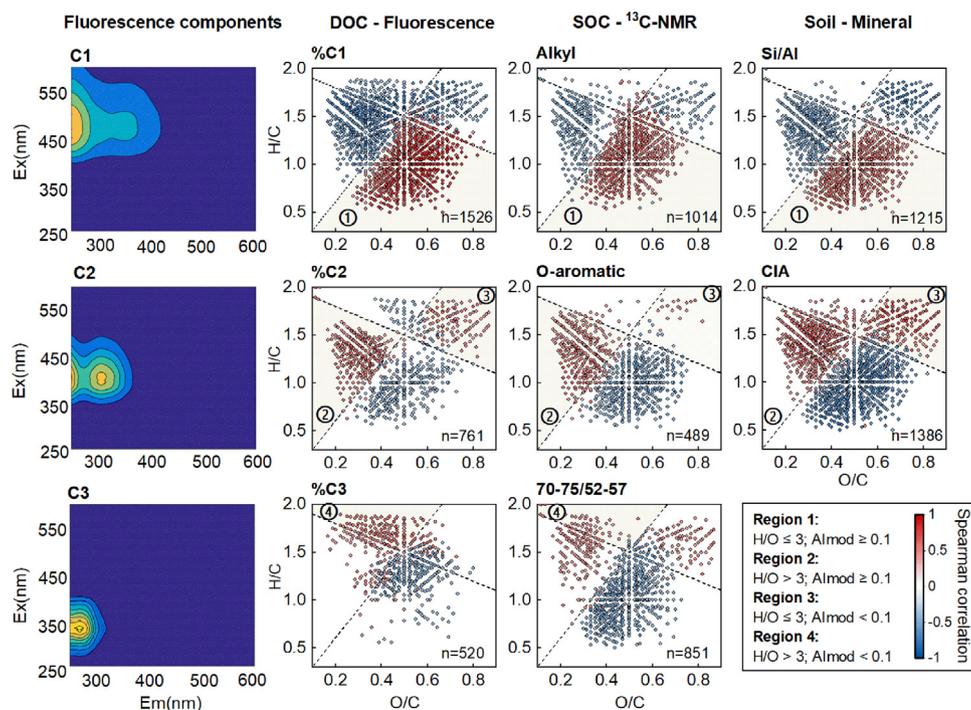
To explore potential SOM and mineral elemental constraints on soil DOM characteristics, we applied Spearman correlation analyses between individual formulae and soil properties (examples in Fig. 2 and full analysis in Figs. S5–S7) and linear correlation analyses (Fig. 3) of the relative abundances of DOM formulae in the four regions (%R1–%R4) versus soil/mineralogical properties.

The “terrestrial-humic-like” C1-related Region I contains formulae clusters that positively correlated with SOC content, pH, alkyl C, N-alkyl C, alkyl/O-alkyl ratio,  $\text{SiO}_2$ , MgO, CaO,  $\text{Na}_2\text{O}$ , and FeO contents, and FeO/ $\text{Fe}_2\text{O}_3$  and Si/Al ratios, and negatively correlated with O-aromatic C, carboxyl C, an area ratio of 70–75/52–57 ppm,  $\text{Al}_2\text{O}_3$  and  $\text{Fe}_2\text{O}_3$  contents and CIA (Figs. S5–S7). The relative abundance of formulae in Region I (%R1) showed similar correlations with these indices (Fig. 3). Therefore, the more highly weathered soils having a higher CIA and lower Si/Al ratio and pH values were found to contain less O-rich DOM molecular formulae in Region I. This phenomenon can be explained by ligand-exchange complexation of DOM molecules with high aromaticity and acidic functional groups at surfaces of Al/Fe (hydr)oxides prevalent in highly weathered and acidic soils [41–43]. This strong complexation can reduce the possible dissolution of the terrestrial humic-like DOC components and lead to the long-term accumulation of O-aromatic C in the SOC. Additionally, in the strongly weathered soils with relatively low pH, the highly oxygenated DOM molecular formulae in Region I tend to protonate and cause lower solubility and DOM yields [44,45].

Contrasting with Region I, the “microbial-humic-like” C2-related Regions II and III contained formulae clusters that showed negative correlations with the degree of mineral weathering (e.g., pH, elemental composition, Si/Al ratio and CIA; Figs. S5–S7). Relative abundances of DOM molecular formulae in Regions II and III (%R2 and %R3) also showed significant correlations with these parameters (Fig. 3). Clearly, the molecular formulae in these two regions were also mainly constrained by the weathering-related dimension. In highly weathered soils, these constituents tend to be weakly sorbed by Al/Fe (hydr)oxides [41–43].



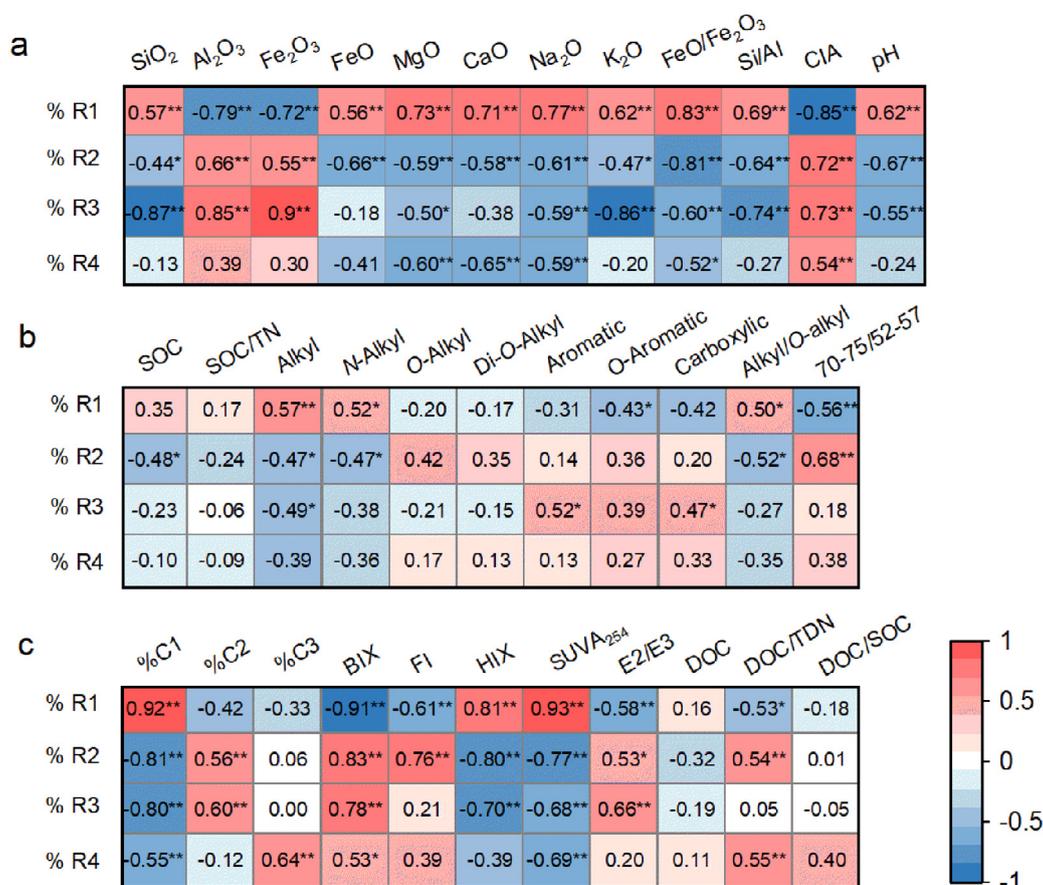
**Fig. 1. Multivariate analysis of molecular and optical signatures of DOM using non-metric multidimensional scaling.** (a) Ordination was based on Bray-Curtis dissimilarities of variables derived from FT-ICR MS molecular composition ( $k=3$ , stress=0.03). Optical indices were fit to the same ordination over 999 permutations. Variables (% C2, %C3, and FI) with significance levels  $>0.1$  are shown in light gray, and others with  $P<0.1$  are shown in brown. (b) Samples are color-coded based on the pH gradient, and sized by the weathering degree indicated by the chemical index of alteration (CIA). Abbreviations:  $SUVA_{254}$ , specific ultraviolet absorbance at 254 nm; HIX, humification index; E2/E3, absorbance at 254 nm divided by the absorbance at 365 nm; BIX, biological index; FI, fluorescence index; %C1–%C3, relative abundances of fluorescence components C1–C3.



**Fig. 2. Molecular associations between individual molecules and three fluorescence components, soil organic matter composition and soil mineral characteristics.** FT-ICR MS compounds with significant Spearman correlations ( $P<0.1$ ) with other indexes are plotted in van-Krevelen diagrams. Red color indicates a positive correlation, and blue color indicates a negative correlation. Regional lines are combination of  $AI_{mod} = 0.1$  and  $H/O = 3$ .

Thus, these O-poor highly unsaturated compounds and aromatics (Region II) and the amino sugars and carbohydrates (Region III) account for higher proportions of the DOM pool. The O-poor molecular formulae in Region II that were significantly correlated with BIX, FI and DOC/TDN may have higher contributions from freshly produced microbial byproducts [46]. The molecular formulae in Region III contained a significant amount of amino sugars and carbohydrates that should not contribute to the fluorescent signals of the “microbial-humic-like” C2 fraction due to the lack of a conjugation structure [47]. Nevertheless, amino sugars are specifically sourced from microbes and their high correlation with “microbial-humic-like” C2 support the feasibility of using this fluorescence signal to indicate microbial contributions to SOC.

The “protein-like” C3 related Region 4 contained substantial amounts of saturated aliphatic and peptide-like molecular formulae, which supports the possibility of using this fluorescence signal as an indicator of protein contributions. Formulae in this region negatively correlated with CaO and MgO contents and  $SUVA_{254}$ , and positively correlated with BIX and DOC/TDN (Figs. S5–S7). These relationships suggest that they are important contributor to the non-aromatic fraction and are influenced by localized nutrient conditions. Nevertheless, this region contains only a few molecular formulae that correlated with weathering-related parameters, such as Si/Al ratio and CIA, suggesting that this region is less affected by mineral weathering than other biochemical regions.



**Fig. 3.** Heat maps showing the relationship between the relative abundance of all soil dissolved organic matter formulae clustered in four regions (see region definition in Fig. 2 legend) and soil parameters. Correlations with categorical variables of (a) soil mineral properties, (b) soil organic carbon characters, and (c) soil dissolved organic matter characters. Correlations are scaled from red for a positive correlation to blue for a negative correlation. Significance levels: \*  $P < 0.05$ , \*\*  $P < 0.01$ .

Overall, the four clusters that linked to the optical parameters of DOM showed distinct tendencies for being released from SOM, interacting with Al/Fe (hydr)oxides, and being microbiologically processed based on their relations with various soil properties. Importantly, the mineral weathering-related dimension (e.g., pH, elemental composition and mineral weathering indices) was identified as the most important constraint on the major molecular signatures of soil-derived DOM, especially those in Regions I–III, but not those in Region IV.

### 3.4. Geochemical, environmental and ecological significance

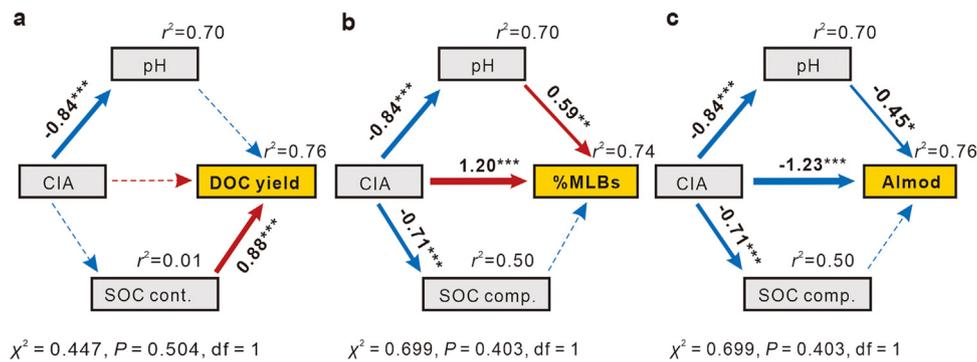
Quantity and quality of soil DOM (e.g., bioavailability and aromaticity) are critical factors regulating several soil functions, such as fueling soil microbes for carbon/nutrient cycling with labile substrates and transporting pollutants via cation- $\pi$  and  $\pi$ - $\pi$  interactions [3,4,48]. We used structural equation models (SEM) based on the relationships between DOC with SOC and soil elemental characteristics (Figs. S2, S8, and S9) to assess the critical factors controlling DOC yield, %MLB<sub>L</sub> and Al<sub>mod</sub>. The models explained a high degree of variance in the quantity and quality of soil DOM ( $r^2 > 0.7$ ) across the diversity of soils. The models confirmed the important role of SOC content as the dominant driver of DOC yield (Fig. 4a). While mineral weathering (indicated by CIA) can affect pH and SOC composition, mineral weathering rather than pH or SOC composition had the strongest direct effect on both %MLB<sub>L</sub> and Al<sub>mod</sub> of DOM (Fig. 4b and c). Accordingly, the soils with

a greater degree of weathering yielded more biolabile DOM components with lower DOM aromaticity.

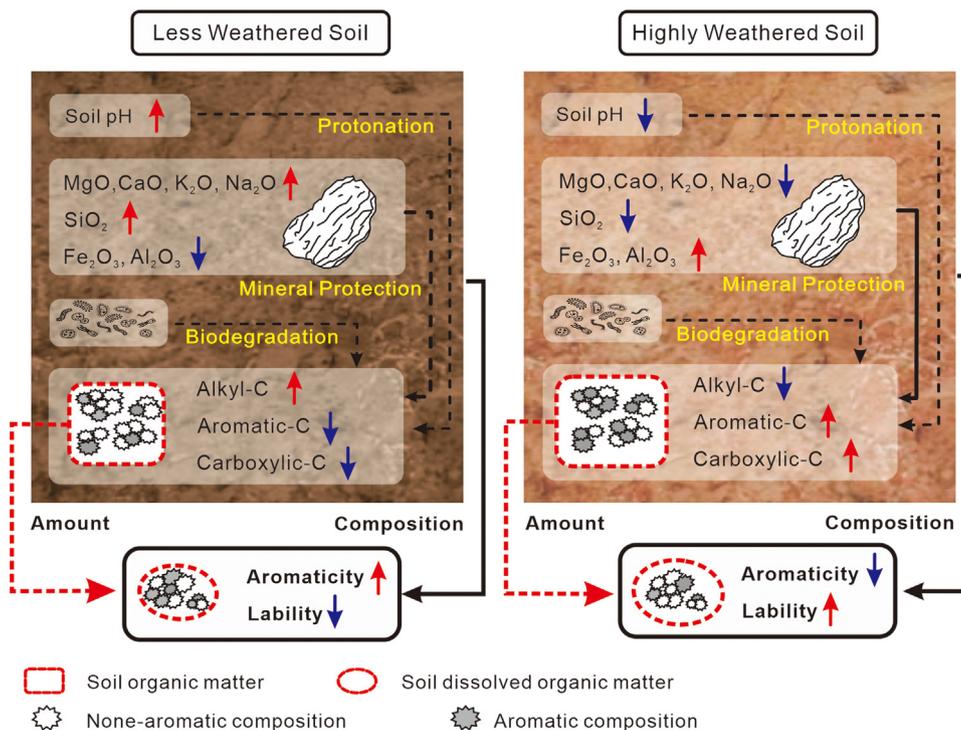
Our study highlights the dominant role of SOM content and soil elemental/mineralogical characteristics in regulating the yield and molecular signatures of soil-derived DOM, respectively (Fig. 5). Within highly weathered soils, aromatic carbon constituents tend to be adsorbed and chemically protected by sorption to Al/Fe (hydr)oxides, thereby accumulating in soils. Additionally, a lower aromatic content of soil DOM can make it more available to microbes and thus more biodegradable. The molecular signatures of soil-derived DOM and their variations with soil properties provide a fundamental dataset for understanding, modeling and optimizing specific DOM functions across various geographical areas using soil-mapping inventories. While the present findings are derived from a limited series of reference soils (22 diverse soils), this study provides a conceptual framework for future investigations to expand the geographical extent of soil characteristics to further unify the interpretation and modeling of molecular signatures and functions of soil-derived DOM.

### Data availability

The data supporting the findings of this study is available within supplementary information. The detailed molecular data is archived in open research data repository under Science Data Bank (<http://www.doi.org/10.11922/sciencedb.01064>).



**Fig. 4.** Structure equation models examining the multivariate effects of soil elemental and soil organic composition on (a) soil DOC yield and (b, c) soil DOM composition. Arrows indicate the hypothesized path of causation, while dotted arrows indicate insignificant pathways (\*\* $P < 0.001$ ; \* $P < 0.01$ ;  $P < 0.05$ ). Arrow width is proportional to the strength of the relationship, while red and blue arrows indicate positive and negative relationships, respectively. The numbers adjacent to arrows are standardized path coefficients, which reflect the effect size of the relationship. SOM is represented by the first component from the principal component analysis based on  $^{13}\text{C}$  NMR.



**Fig. 5.** Conceptual model showing the impacts of soil weathering and soil organic matter on the quantity and quality of soil dissolved organic matter. The characters of less weathered and highly weathered soils were extracted based on redundancy analysis and principal component analysis (Figs. S10 and S11). Positive and negative responses are indicated by red and blue arrows, respectively. With increasing weathering degree, the loss of base cations ( $\text{Mg}^{2+}$ ,  $\text{Ca}^{2+}$ ,  $\text{Na}^+$ , and  $\text{K}^+$ ) increased, and silica (Si) preferentially dissolves/leaches versus Al and Fe, leaving Al/Fe (hydr)oxides to adsorb and preserve aromatic carbon. In addition, a lower pH caused by intense weathering/leaching may increase the protonation of phenolic DOM and decreased the solubility of soil aromatic carbon. With strong processing by microbes, DOM from less weathered soil is featured with less aromaticity and a higher amount of microbial processed components, and therefore has a lower molecular weight and higher lability. Conversely, the DOM concentration is primarily determined by the amount of soil organic matter.

#### 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.

#### CRedit authorship contribution statement

**Ying-Hui Wang:** Data curation, Formal analysis, Writing – original draft, Writing – review & editing. **Peng Zhang:** Data curation, Formal

analysis. **Chen He:** Data curation, Formal analysis. **Jian-Chun Yu:** Data curation, Formal analysis. **Quan Shi:** Data curation, Formal analysis. **Randy A. Dahlgren:** Data curation, Writing – original draft, Writing – review & editing. **Robert G.M. Spencer:** Data curation, Writing – original draft, Writing – review & editing. **Zhi-Bing Yang:** Data curation, Writing – original draft, Writing – review & editing. **Jun-Jian Wang:** Conceptualization, Investigation, Funding acquisition, Data curation, Writing – original draft, Writing – review & editing.

## Acknowledgements

This work was financially supported by the National Natural Science Foundation of China (42122054, 42192513, [41807360](#)), Guangdong Basic and Applied Basic Research Foundation ([2021B1515020082](#)), Key Platform and Scientific Research Projects of Guangdong Provincial Education Department ([2019KZDXM028](#), and [2020KCXTD006](#)), and Science and Technology Development Fund Project of Shenzhen ([JCYJ20190809142611503](#) and [JCYJ20190809162205531](#)).

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.fmre.2022.01.032](https://doi.org/10.1016/j.fmre.2022.01.032).

## References

- [1] B. Marschner, K. Kalbitz, Controls of bioavailability and biodegradability of dissolved organic matter in soils, *Geoderma* 113 (2003) 211–235, [doi:10.1016/S0016-7061\(02\)00362-2](https://doi.org/10.1016/S0016-7061(02)00362-2).
- [2] K. Kalbitz, S. Solinger, J.H. Park, et al., Controls on the dynamics of dissolved organic matter in soils: a review, *Soil Sci.* 165 (2000) 277–304, [doi:10.1097/00010694-200004000-00001](https://doi.org/10.1097/00010694-200004000-00001).
- [3] G.R. Aiken, H. Hsu-Kim, J.N. Ryan, Influence of dissolved organic matter on the environmental fate of metals, nanoparticles, and colloids, *Environ. Sci. Technol.* 45 (2011) 3196–3201, [doi:10.1021/es103992s](https://doi.org/10.1021/es103992s).
- [4] A.M. Kellerman, F. Guillemette, D.C. Podgorski, et al., Unifying concepts linking dissolved organic matter composition to persistence in aquatic ecosystems, *Environ. Sci. Technol.* 52 (2018) 2538–2548, [doi:10.1021/acs.est.7b05513](https://doi.org/10.1021/acs.est.7b05513).
- [5] J. Lehmann, M. Kleber, The contentious nature of soil organic matter, *Nature* 528 (2015) 60–68, [doi:10.1038/nature16069](https://doi.org/10.1038/nature16069).
- [6] D. Markewitz, E.A. Davidson, R.D.O. Figueiredo, et al., Control of cation concentrations in stream waters by surface soil processes in an Amazonian watershed, *Nature* 410 (2001) 802–805, [doi:10.1038/35071052](https://doi.org/10.1038/35071052).
- [7] T.S. Bianchi, The role of terrestrially derived organic carbon in the coastal ocean: a changing paradigm and the priming effect, *Proc. Natl. Acad. Sci. U.S.A.* 108 (2011) 19473–19481, [doi:10.1073/pnas.1017982108](https://doi.org/10.1073/pnas.1017982108).
- [8] S. Fontaine, S. Barot, P. Barré, et al., Stability of organic carbon in deep soil layers controlled by fresh carbon supply, *Nature* 450 (2007) 277–280, [doi:10.1038/nature06275](https://doi.org/10.1038/nature06275).
- [9] G.D. Joannis, R.L. Bradley, C.M. Preston, et al., Soil enzyme inhibition by condensed litter tannins may drive ecosystem structure and processes: the case of *Kalmia angustifolia*, *New Phytol.* 175 (2007) 535–546, [doi:10.1111/j.1469-8137.2007.02113.x](https://doi.org/10.1111/j.1469-8137.2007.02113.x).
- [10] J.A. Keuskamp, I.C. Feller, H.J. Laanbroek, et al., Short- and long-term effects of nutrient enrichment on microbial exoenzyme activity in mangrove peat, *Soil Biol. Biochem.* 81 (2015) 38–47, [doi:10.1016/j.soilbio.2014.11.003](https://doi.org/10.1016/j.soilbio.2014.11.003).
- [11] V.N. Roth, M. Lange, C. Simon, et al., Persistence of dissolved organic matter explained by molecular changes during its passage through soil, *Nat. Geosci.* 12 (2019) 755–761, [doi:10.1038/s41561-019-0417-4](https://doi.org/10.1038/s41561-019-0417-4).
- [12] Q.H. Ye, Y.H. Wang, Z.T. Zhang, et al., Dissolved organic matter characteristics in soils of tropical legume and non-legume tree plantations, *Soil Biol. Biochem.* 148 (2020) 107880, [doi:10.1016/j.soilbio.2020.107880](https://doi.org/10.1016/j.soilbio.2020.107880).
- [13] Y. Zhang, K.V. Heal, M. Shi, et al., Decreasing molecular diversity of soil dissolved organic matter related to microbial community along an alpine elevation gradient, *Sci. Total Environ.* (2021) 151823, [doi:10.1016/j.scitotenv.2021.151823](https://doi.org/10.1016/j.scitotenv.2021.151823).
- [14] Y. Ding, Z. Shi, Q. Ye, et al., Chemodiversity of soil dissolved organic matter, *Environ. Sci. Technol.* 54 (2020) 6174–6184, [doi:10.1021/acs.est.0c01136](https://doi.org/10.1021/acs.est.0c01136).
- [15] H.Y. Li, H. Wang, H.T. Wang, et al., The chemodiversity of paddy soil dissolved organic matter correlates with microbial community at continental scales, *Microbiome* 6 (2018), [doi:10.1186/s40168-018-0561-x](https://doi.org/10.1186/s40168-018-0561-x).
- [16] J.L. Dixon, A.M. Heimsath, R. Amundson, The critical role of climate and saprolite weathering in landscape evolution, *Earth Surf. Process. Landf.* 34 (2009) 1507–1521, [doi:10.1002/esp.1836](https://doi.org/10.1002/esp.1836).
- [17] M. Egli, A. Mirabella, G. Sartori, The role of climate and vegetation in weathering and clay mineral formation in late quaternary soils of the Swiss and Italian Alps, *Geomorphology* 102 (2008) 307–324, [doi:10.1016/j.geomorph.2008.04.001](https://doi.org/10.1016/j.geomorph.2008.04.001).
- [18] A.M. Kellerman, D.N. Kothawala, T. Dittmar, et al., Persistence of dissolved organic matter in lakes related to its molecular characteristics, *Nat. Geosci.* 8 (2015) 454–457, [doi:10.1038/Ngeo2440](https://doi.org/10.1038/Ngeo2440).
- [19] A. Stubbins, J.F. Lapierre, M. Berggren, et al., What's in an EEM? Molecular signatures associated with dissolved organic fluorescence in boreal Canada, *Environ. Sci. Technol.* 48 (2014) 10598–10606, [doi:10.1021/es502086e](https://doi.org/10.1021/es502086e).
- [20] S. Wagner, R. Jaffé, K. Cawley, et al., Associations between the molecular and optical properties of dissolved organic matter in the Florida everglades, a model coastal wetland system, *Front. Chem.* 3 (2015) 66, [doi:10.3389/fchem.2015.00066](https://doi.org/10.3389/fchem.2015.00066).
- [21] *Soil Survey Staff, Keys to Soil Taxonomy, 12th Edition, USDA, Natural Resources Conservation Service, Washington DC, 2014.*
- [22] B.P. Ruxton, Measures of the degree of chemical weathering of rocks, *J. Geol.* 76 (1968) 518–527, [doi:10.1086/627357](https://doi.org/10.1086/627357).
- [23] H.W. Nesbitt, G.M. Young, Early proterozoic climates and plate motions inferred from major element chemistry of lutites, *Nature* 299 (1982) 715–717, [doi:10.1038/299715a0](https://doi.org/10.1038/299715a0).
- [24] M.W.I. Schmidt, H. Knicker, P.G. Hatcher, et al., Improvement of C-13 and N-15 CPMAS NMR spectra of bulk soils, particle size fractions and organic material by treatment with 10% hydrofluoric acid, *Eur. J. Soil Sci.* 48 (1997) 319–328, [doi:10.1111/j.1365-2389.1997.tb00552.x](https://doi.org/10.1111/j.1365-2389.1997.tb00552.x).
- [25] J.A. Baldock, R.J. Smernik, Chemical composition and bioavailability of thermally altered *Pinus resinosa* (Red Pine) wood, *Org. Geochem.* 33 (2002) 1093–1109, [doi:10.1016/S0146-6380\(02\)00062-1](https://doi.org/10.1016/S0146-6380(02)00062-1).
- [26] G. Bonanomi, G. Incerti, E. Barile, et al., Phytotoxicity, not nitrogen immobilization, explains plant litter inhibitory effects: evidence from solid-state <sup>13</sup>C NMR spectroscopy, *New Phytol.* 191 (2011) 1018–1030, [doi:10.1111/j.1469-8137.2011.03765.x](https://doi.org/10.1111/j.1469-8137.2011.03765.x).
- [27] S. Sjögersten, S. Caul, T.J. Daniell, et al., Organic matter chemistry controls greenhouse gas emissions from permafrost peatlands, *Soil Biol. Biochem.* 98 (2016) 42–53, [doi:10.1016/j.soilbio.2016.03.016](https://doi.org/10.1016/j.soilbio.2016.03.016).
- [28] G. Bonanomi, G. Incerti, F. Giannino, et al., Litter quality assessed by solid state <sup>13</sup>C NMR spectroscopy predicts decay rate better than C/N and Lignin/N ratios, *Soil Biol. Biochem.* 56 (2013) 40–48, [doi:10.1016/j.soilbio.2012.03.003](https://doi.org/10.1016/j.soilbio.2012.03.003).
- [29] W. Cai, Z.L. Du, A.P. Zhang, et al., Long-term biochar addition alters the characteristics but not the chlorine reactivity of soil-derived dissolved organic matter, *Water Res.* 185 (2020) 116260, [doi:10.1016/j.watres.2020.116260](https://doi.org/10.1016/j.watres.2020.116260).
- [30] J. Peuravuori, K. Pihlaja, Molecular size distribution and spectroscopic properties of aquatic humic substances, *Anal. Chim. Acta* 337 (1997) 133–149, [doi:10.1016/S0003-2670\(96\)00412-6](https://doi.org/10.1016/S0003-2670(96)00412-6).
- [31] J.L. Weishaar, G.R. Aiken, B.A. Bergamaschi, et al., Evaluation of specific ultraviolet absorbance as an indicator of the chemical composition and reactivity of dissolved organic carbon, *Environ. Sci. Technol.* 37 (2003) 4702–4708, [doi:10.1021/es030360x](https://doi.org/10.1021/es030360x).
- [32] K.R. Murphy, C.A. Stedmon, D. Graeber, et al., Fluorescence spectroscopy and multi-way techniques, PARAFAC, *Anal. Methods-UK* 5 (2013) 6557–6566, [doi:10.1039/c3ay41160e](https://doi.org/10.1039/c3ay41160e).
- [33] D.M. McKnight, E.W. Boyer, P.K. Westerhoff, et al., Spectrofluorometric characterization of dissolved organic matter for indication of precursor organic material and aromaticity, *Limnol. Oceanogr.* 46 (2001) 38–48, [doi:10.4319/lo.2001.46.1.0038](https://doi.org/10.4319/lo.2001.46.1.0038).
- [34] T. Ohno, Fluorescence inner-filtering correction for determining the humification index of dissolved organic matter, *Environ. Sci. Technol.* 36 (2002) 742–746, [doi:10.1021/es0155276](https://doi.org/10.1021/es0155276).
- [35] H.F. Wilson, M.A. Xenopoulos, Effects of agricultural land use on the composition of fluvial dissolved organic matter, *Nat. Geosci.* 2 (2009) 37–41, [doi:10.1038/ngeo0391](https://doi.org/10.1038/ngeo0391).
- [36] Y.H. Wang, R.G.M. Spencer, D.C. Podgorski, et al., Spatiotemporal transformation of dissolved organic matter along an alpine stream flow path on the Qinghai-Tibet plateau: importance of source and permafrost degradation, *Biogeosciences* 15 (2018) 6637–6648, [doi:10.5194/bg-15-6637-2018](https://doi.org/10.5194/bg-15-6637-2018).
- [37] C. He, Y.H. Zhang, Y.Y. Li, et al., In-house standard method for molecular characterization of dissolved organic matter by FT-ICR mass spectrometry, *ACS Omega* 5 (2020) 11730–11736, [doi:10.1021/acsomega.0c01055](https://doi.org/10.1021/acsomega.0c01055).
- [38] A. Rivas-Ubach, Y. Liu, T.S. Bianchi, et al., Moving beyond the van krevelen diagram: A new stoichiometric approach for compound classification in organisms, *Anal. Chem.* 90 (2018) 6152–6160, [doi:10.1021/acs.analchem.8b00529](https://doi.org/10.1021/acs.analchem.8b00529).
- [39] B.P. Koch, T. Dittmar, From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter, *Rapid Commun. Mass Spectrom.* 30 (2016) 250–250, [doi:10.1002/rcm.7433](https://doi.org/10.1002/rcm.7433).
- [40] J. D'Andrilli, W.T. Cooper, C.M. Foreman, et al., An ultrahigh-resolution mass spectrometry index to estimate natural organic matter lability, *Rapid Commun. Mass Spectrom.* 29 (2015) 2385–2401, [doi:10.1002/rcm.7400](https://doi.org/10.1002/rcm.7400).
- [41] J. Lv, S. Zhang, S. Wang, et al., Molecular-scale investigation with ESI-FT-ICR-MS on fractionation of dissolved organic matter induced by adsorption on iron oxyhydroxides, *Environ. Sci. Technol.* 50 (2016) 2328–2336, [doi:10.1021/acs.est.5b04996](https://doi.org/10.1021/acs.est.5b04996).
- [42] A. Oren, B. Chefetz, Successive sorption-desorption cycles of dissolved organic matter in mineral soil matrices, *Geoderma* 189–190 (2012) 108–115, [doi:10.1016/j.geoderma.2012.05.004](https://doi.org/10.1016/j.geoderma.2012.05.004).
- [43] Y. Wang, H. Wang, J.S. He, et al., Iron-mediated soil carbon response to water-table decline in an alpine wetland, *Nat. Commun.* 8 (2017) 15972, [doi:10.1038/ncomms15972](https://doi.org/10.1038/ncomms15972).
- [44] J.B. Christensen, T.H. Christensen, The effect of pH on the complexation of Cd, Ni and Zn by dissolved organic carbon from leachate-polluted groundwater, *Water Res.* 34 (2000) 3743–3754, [doi:10.1016/S0043-1354\(00\)00127-5](https://doi.org/10.1016/S0043-1354(00)00127-5).
- [45] N.R. Urban, S.E. Bayley, S.J. Eisenreich, Export of dissolved organic-carbon and acidity from peatlands, *Water Resour. Res.* 25 (1989) 1619–1628, [doi:10.1029/WR025i007p01619](https://doi.org/10.1029/WR025i007p01619).
- [46] P.A. Meyers, R. Ishiwatari, Lacustrine organic geochemistry - an overview of indicators of organic matter sources and diagenesis in lake sediments, *Org. Geochem.* 20 (1993) 867–900, [doi:10.1016/0146-6380\(93\)90100-P](https://doi.org/10.1016/0146-6380(93)90100-P).
- [47] D.M. Reynolds, *The principles of fluorescence*, in: P.G. Coble, J. Lead, A. Baker, D.M. Reynolds, R.G. Spencer (Eds.), *Aquatic Organic Matter Fluorescence*, Cambridge University Press, New York, 2014, pp. 3–34.
- [48] M. Ravichandran, Interactions between mercury and dissolved organic matter - a review, *Chemosphere* 55 (2004) 319–331, [doi:10.1016/j.chemosphere.2003.11.011](https://doi.org/10.1016/j.chemosphere.2003.11.011).



**Ying-Hui Wang** is a postdoc researcher fellow under the supervision of Dr. Jun-Jian Wang at the School of Environmental Science and Engineering, Southern University of Science and Technology. She received her Ph.D. Degree from Peking University in 2019. Her current research is focused on the chemodiversity and dynamics of dissolved organic matter in soil systems.



**Jun-Jian Wang**, associate professor at Southern University of Science and Technology. His research focuses on the organic carbon biogeochemistry and environmental quality in the plant–soil–water continuum. He received the National Science Fund for Excellent Young Scholars of China and Young Scientist Award by China Environmental Science Association. He published more than 70 papers in top-tier journals including *Nature Communications*, *Environmental Science & Technology*, and *Water Research*. He is serving as an associate editor of the Journal of Environmental Quality and on the young editorial boards of the Critical Reviews in Environmental Science and Technology and Chinese Chemical Letters.