

Research Article

Molecular Signatures of Humic Acids from Different Sources as Revealed by Ultrahigh Resolution Mass Spectrometry

Shuai Qin,^{1,2} Chengbin Xu,² Yingzi Xu,¹ Yingchen Bai,¹ and Fei Guo ¹

¹State Key Laboratory of Environmental Criteria and Risk Assessment, Chinese Research Academy of Environmental Sciences, Beijing 100012, China

²School of Environmental Sciences, Liaoning University, Shenyang 110036, China

Correspondence should be addressed to Fei Guo; figoth@163.com

Received 27 December 2019; Accepted 5 February 2020; Published 1 April 2020

Guest Editor: Yihua Xiao

Copyright © 2020 Shuai Qin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Humic acid (HA) is extremely important for understanding the geochemical cycle of pollutants in different environments. Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) has performed molecular-level analysis of two standard HAs from the Suwannee River (SRHA) and leonardite (LEHA) and HA from Jiufeng forest in Beijing (JFHA), which is impossible for other conventional instruments. Regardless of the source of HA, compounds containing more heteroatoms (such as nitrogen and sulfur) have a higher degree of unsaturation and aromaticity. JFHA, SRHA, and LEHA from soil, river, and leonardite, respectively, are arranged in order from the lowest to highest degree of humification, according to molecular unsaturation and aromaticity of HAs. Soil HA is more labile and contains many large molecular weight compounds with low unsaturation. Regardless of unsaturation, molecules of River HA have a homogeneous molecular mass distribution and contain many plant-derived lignin- and tannin-like compounds, which are more stable than lipid and more labile than condensed aromatics. Leonardite HA with a high degree of humification contains a large number of compounds with high aromaticity and more heteroatoms and has low lability. Our results reveal the diversity of humic acid at molecular level because of different degree of humification and the lability. These conclusions are significant for understanding the role of humic acid from different sources in pollutant transformation and the geochemical cycle at the molecular level.

1. Introduction

Humic substances (HS) are the most widespread natural organic matter (NOM) in soil, water, and sediment. HS are derived mainly from plant, organism, and animal tissues [1], while the biomaterials lost their initial structures in chemical and biological degradation. Various turnover times of humic substances that affect global carbon balance are due to inherent structural resistance to biotic decomposition and binding to minerals [2–5]. They have complex chemical structures which are more stable than their precursors and have lost their chemical properties [6]. HS can be classified into humic acid (HA), fulvic acid (FA), and humin (HM) based on the solubility [7, 8]. HA generally represents the major fraction of HS [6] and is mainly derived from typical environments rich in biochemical reactions such as soil, rivers, and leonardite.

Soil, river, and leonardite are three typical sources of HAs. Soil HAs are mainly derived from the (bio)chemical degradation of plant and animal residues and synthetic activity from microorganisms and account for 20% of the total organic matter [6, 9, 10]. Rivers play a very important role in the transformation [11, 12] and global cycling of HAs [12, 13]. Elemental composition of molecules can cause the spontaneous generation of relatively labile HAs [14, 15] or change the HA reactivity and adsorption process. Lignite (low rank coal) is the second stage of coal formation [6]. Leonardite, the most oxidized variety of lignite, is the richest source of HS [16, 17]. Leonardite-derived HA accounts for 10–80% of leonardite organic matter, depending on its maturity level [18] and is further along the diagenetic path of humification. They are much older and contain more condensed aromatics than soil HAs [6]. There are a number

of studies concerning HA derived from different types of soil [19] and rivers worldwide [11], but contrast data on the molecular composition of HA from soil, river, and leonardite are still missing. The standard samples provided by International Humic Substances Society can more accurately reflect characteristics of HA from various sources.

It is difficult to analyse HA at the molecular level due to its extreme complexity, low concentration, and high polarity [20]. There are many conventional techniques to characterize HA from a variety of sources, such as liquid chromatography-mass spectrometry (LC-MS), pyrolysis gas chromatography-mass spectrometry (GC-MS), and nuclear magnetic resonance (NMR) [21]. They provide apparent or biased information about HA and fail to dig for complete characteristics [22]. Low-resolution GC-MS and LC-MS can only analyse a small part of HA due to volatility and solubility, respectively. NMR can give the functional group information of the bulk HA but fail to give molecular-level details [20]. Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) is a very suitable tool for analysing HA from different sources [23]. Due to its high resolution and accuracy, it can assign unique molecular formula to thousands of compounds [24]. Moreover, as a soft ionization technology, negative ion mode electrospray ionization (ESI) has high selectivity for polar molecules in complex mixtures [25], which can avoid complex separation procedures [26].

The purpose of this work is to investigate the molecular compositions of HAs from soil, river, and leonardite using ESI FT-ICR MS. The specific objectives are to (1) compare the diversity of molecular composition between HAs from different resources and (2) discuss the lability of diverse HAs.

2. Materials and Methods

2.1. Samples and Extraction Method. Humic acid (HA) collected from Suwannee River in South Georgia (SRHA) is a standard fulvic acid sample of International Humic Substances Society (IHSS) and its serial number is 2S101F. The IHSS standard leonardite-derived HA sample (LEHA) was obtained from the Gascoyne Mine in Bowman County, North Dakota, USA [27], and its serial number is 4S102F. Leonardite was produced by the natural oxidation of exposed lignite, which was a low-grade coal. The microbial action and other chemical processes of leonardite in the natural environment are longer and more complex, and therefore LEHA has a higher degree of humification. More additional details of the source material can be found at <http://humic-substances.org/>. HA collected from Beijing Jiufeng forest soils in China (JFHA) was extracted according to the standard method recommended by IHSS. The detailed extraction method is as follows.

Take certain amount of freeze-dried soil from Jiufeng forest and thoroughly mix it with 0.1 mol/L NaOH solution to obtain mixed solution with solid-liquid ratio of 1:10. Remove the residual soil after centrifugation. The 6.0 mol/L HCl solution was added to the mixed solution to make the pH value 1 and then left to stand for 24 hours to obtain

precipitate. 0.1 mol/L KOH solution was added to the precipitate, the pH was adjusted to 13.0, and then certain amount of potassium chloride was added to adjust the solution concentration to 0.4 mol/L. After half an hour of high-speed centrifugation (10000 r/min), the supernatant was separated, and 6.0 mol/L HCl was added to adjust the pH to 1.0, and concentrated HF was added to make the concentration to 0.3 mol/L. Impurities were removed by shaking and dialysis, and the solution was freeze-dried to obtain humic acid sample.

2.2. Parameters of FT-ICR MS. The spectra of JFHA, SRHA, and LEHA were obtained using a Bruker Apex-ultra Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) with a 9.4 T superconducting magnet and Apollo II electrospray ion source. The operating parameters of the FT-ICR MS analysis were mainly according to the previous studies [28]. Sample solutions were continuously injected into the electrospray ion source at a flow rate of 180 μ L/h with a syringe pump. The spray shield voltage, capillary column introducing voltage, and capillary column end voltage of the negative-ion ESI were set to 4.0 kV, 4.5 kV, and -320 V, respectively. For each spectrum multiple (200–300) transient scans (collected with a 4 M Word time domain) were coadded in the m/z 200–700 mass range. The spectra were externally calibrated from a homologous series of the Suwannee River Natural Organic Matter sample (SRNOM). The mass measurement accuracy for single charged molecular ions was better than ± 1 ppm. Mass peaks were exported if the signal-to-noise (S/N) ratio was greater than 6. Molecular formula assignments were limited by the following constraints for the three samples: 12C (0–100), 13C (0–2), 1H (0–100), 14N (0–3), 16O (0–30), and 32S (0–2) atoms, and only the possible element composition within ± 1 ppm mass measurement error was considered.

2.3. Molecular Indexes of Humic Acid. Normalized peak intensities (normalized to the sum of peak intensities of identified molecular formulas) were used to semiquantitatively evaluate differences in HA molecular composition [29]. The modified aromatic index (AI-mod) reflected the number of double C=C bonds, C-O bonds, and C=O bonds and was used to assess the aromaticity of compounds. AI-mod was calculated based on the assigned molecular formula (described as $C_cH_hO_oN_nS_s$) using the equation $AI\text{-mod} = (1 + c - 0.5o - s - 0.5h) / (c - 0.5o - s - n)$ [30]. The double bond equivalents (DBE) was used to assess the degree of unsaturation of compounds and calculated as follows: $DBE = c - h/2 + n/2 + 1$. Intensity weighted average values are displayed for number of carbons (C_{wa}), hydrogens (H_{wa}), oxygens (O_{wa}), nitrogens (N_{wa}), sulfurs (S_{wa}), molecular masses (m/z_{wa}), hydrogen to carbon ratios (H/C_{wa}), oxygen to carbon ratios (O/C_{wa}), double bond equivalents (DBE_{wa}), and modified aromatic indexes ($AI\text{-mod}_{wa}$). The molecular lability boundary index (MLB) is the bound when H/C is 1.5. $MLB_L\%$ which can be used to evaluate the lability of HA. When $H/C > 1.5$, the NOM constituents above MLB correspond to a more labile substance (MLB_L), and when

$H/C < 1.5$, the NOM constituents below MLB exhibit less labile and more recalcitrant characteristics (MLB_R). $MLB_L\%$ was calculated by dividing the total intensity of molecular formulas with $H/C \geq 1.5$ by the total intensity of molecular formulas [31]. Carboxyl-rich alicyclic molecules (CRAMs) which are often associated with refractory compounds are complex mixtures of carboxylated and fused alicyclic structures whose carboxyl-C: aliphatic-C ratio is 1:2 to 1:7 [32]. $CRAM\%$ was calculated by dividing the total intensity of molecular formulas with $DBE/C = 0.3-0.68$, $DBE/H = 0.2-0.95$, and $DBE/O = 0.77-1.75$ by the total intensity of molecular formulas [29]. The nominal oxidation state of C (NOSC) was used to assess the degradability of DOM [33, 34] and was calculated by using $NOSC = 4 - ((4c + h - 3n - 2o - 2s)/c)$ [34]. The assigned molecular formulas were examined using the van Krevelen diagram (v-K diagram) and modified aromatic index (AI-mod). The v-K diagram (a plot of atomic H/C and O/C) is an excellent method for graphical interpretation of large datasets and visualization of their chemical nature. Molecular formulas were further assigned to the following groups [30, 35, 36]. They are (i) polycyclic aromatics ($AI-mod > 0.66$), (ii) highly aromatic compounds ($0.66 \geq AI-mod > 0.50$), (iii) highly unsaturated compounds ($AI-mod \leq 0.50$ and $H/C < 1.5$), (iv) unsaturated aliphatic compounds ($2.0 > H/C \geq 1.5$ and $N = 0$), (v) unsaturated aliphatic compounds containing N ($2.0 > H/C \geq 1.5$ and $N > 0$), (vi) carbohydrates ($H/C \geq 2.0$ and $O/C \geq 0.9$), and (vii) fatty acids and sulfonic acids ($H/C \geq 2.0$ and $O/C < 0.9$) [37-39].

3. Results and Discussion

3.1. General Signatures of FT-ICR MS. The HAs from river, soil, and Leonardite are shown as representative spectra in Figures S1a-S1c. All HA samples have concentrated mass spectral peaks in the range of 200-700 Daltons (Da), and the overall peak shapes of SRHA and LEHA are approximately normal distribution (Figure S1). The number of assigned formulas of SRHA, JFHA, and LEHA is 3559, 2246, and 4398, respectively. The variation in the C, H, O, N, and S contents and molecular mass as well as indices based on these shows distinct differences (Table 1). There are the highest H content and lowest O and N content in JFHA, leading to the highest H/C and lowest O/C in JFHA. In addition, JFHA contains more saturated compounds with higher H/C and lower O/C (Figure 1), and its overall molecular mass is smaller (Figure S2b). This indicates that the degree of humification of JFHA is higher, because refractory humic acid fractions are marked by increases in the lipid components which have high H/C and low O/C [40]. The indexes of SRHA are relatively moderate, such as m/z_{wa} , H/C_{wa} , and DBE_{wa} . LEHA has the lowest H/C and the largest molecular mass and this may be because the condensed aromatic structure with low H/C is abundant in LEHA and has a large molecular weight. This is consistent with the previous conclusion that refractory humic acid fractions are marked by increases in the condensed aromatic components [40].

3.2. Van Krevelen Diagram Analysis. In the v-K diagram, the aromaticity of the compounds shows such a rule: the aromaticity of the compounds in lower layer is higher than that in upper layer [30]. The compound distribution of SRHA shows a rectangle in the v-K diagram, which indicates that the molecular distribution is relatively homogeneous (Figure 1). From the perspective of the number of molecular formulas and the peak magnitudes, there is no significant difference in the distribution proportion of the compounds (Table 2). The compounds are mainly concentrated in highly unsaturated compounds, accounting for 51% in terms of number of molecular formulas and 54% in terms of peak magnitudes. Highly aromatic compounds and polycyclic aromatics are ranked second and third, respectively, accounting for about 10% to 20% (Table 2).

The compounds of JFHA are mainly concentrated in the upper and lower layers, and sparse points are in the middle layer (Figure 1). In addition, from the perspective of the number of molecular formulas and peak magnitudes, the distribution of the compounds is quite different. From the number of molecule formulas, the compounds are mainly distributed in unsaturated aliphatic compounds, highly unsaturated compounds, and polycyclic aromatics, accounting for about 25%. From the peak magnitudes' point of view, the compounds are mainly distributed in unsaturated aliphatic compounds, accounting for 50%, followed by fatty acids, accounting for 25% (Table 2). By comparison, the high peak magnitudes of unsaturated aliphatic compounds and fatty acids and sulfonic acids are due to the high content of these compounds in JFHA firstly and secondly because these compounds are easily ionized in the negative ion mode of ESI FT-ICR MS [19].

The compound distribution of LEHA is a right-angled triangle in the v-K diagram; that is, the lower-layer compound is more abundant and has a wider distribution range than the upper-layer compound (Figure 1). There is no significant difference in the molecular formula number distribution and peak magnitude distribution of compounds. Obviously, the compounds are mainly concentrated in the lower layer, and the largest proportion is polycyclic aromatics, which accounts for about 50%, followed by highly aromatic compounds and highly unsaturated compounds, accounting for about 20-30% (Table 2).

DBE_{wa} for SRHA, JFHA, and LEHA are 10.02, 5.35, and 15.07, respectively, and $AI-mod_{wa}$ are 0.38, 0.17, and 0.59, respectively. These two molecular formula parameters consistently indicate that the aromaticity of LEHA is the highest among the three HAs samples, followed by SRHA, and the smallest is JFHA. This is concordant with previous research results. The DBE value increased in the progression of water-extractable organic matter from Adkins soil, labile humic acid, and refractory calcium humic acid, which indicates that as the degree of humification increases, the aromaticity of organic matter also increases [40]. The compounds that account for a large proportion of LEHA have higher degrees of unsaturation, and the most abundant compounds are polycyclic aromatics which have highest aromaticity. The presence of a large number of aromatic

TABLE 1: Molecular parameters of (a) humic acids from Jiufeng forest soil, (b) humic acids from Suwannee River, and (c) humic acids from leonardite derived from assigned molecular formulas^a.

Sample	C_{wa}	H_{wa}	O_{wa}	N_{wa}	S_{wa}	m/z_{wa}	H/C_{wa}	O/C_{wa}	DBE_{wa}	$AI-mod_{wa}$	$MLBL\%^b$	$CRAM\%^c$	NOSC
JFHA	21.63	34.60	4.62	0.04	0	368.05	1.60	0.22	5.35	0.17	75.47	8.28	-0.16
SRHA	20.07	22.16	8.19	0.54	0.01	394.22	1.10	0.42	10.02	0.38	15.33	51.91	-0.25
LEHA	23.58	19.29	8.02	0.25	0.02	434.01	0.81	0.35	15.07	0.59	7.88	18.25	-007

^aIntensity weighted average values are displayed for number of carbons (C_{wa}), hydrogens (H_{wa}), oxygens (O_{wa}), nitrogens (N_{wa}), sulfurs (S_{wa}), molecular weights (m/z_{wa}), hydrogen to carbon ratios (H/C_{wa}), and oxygen to carbon ratios (O/C_{wa}). ^{b,c}Both $MLBL\%$ and $CRAM\%$ are obtained by dividing the total intensity of the corresponding molecular formula by the total intensity of the molecular formula.

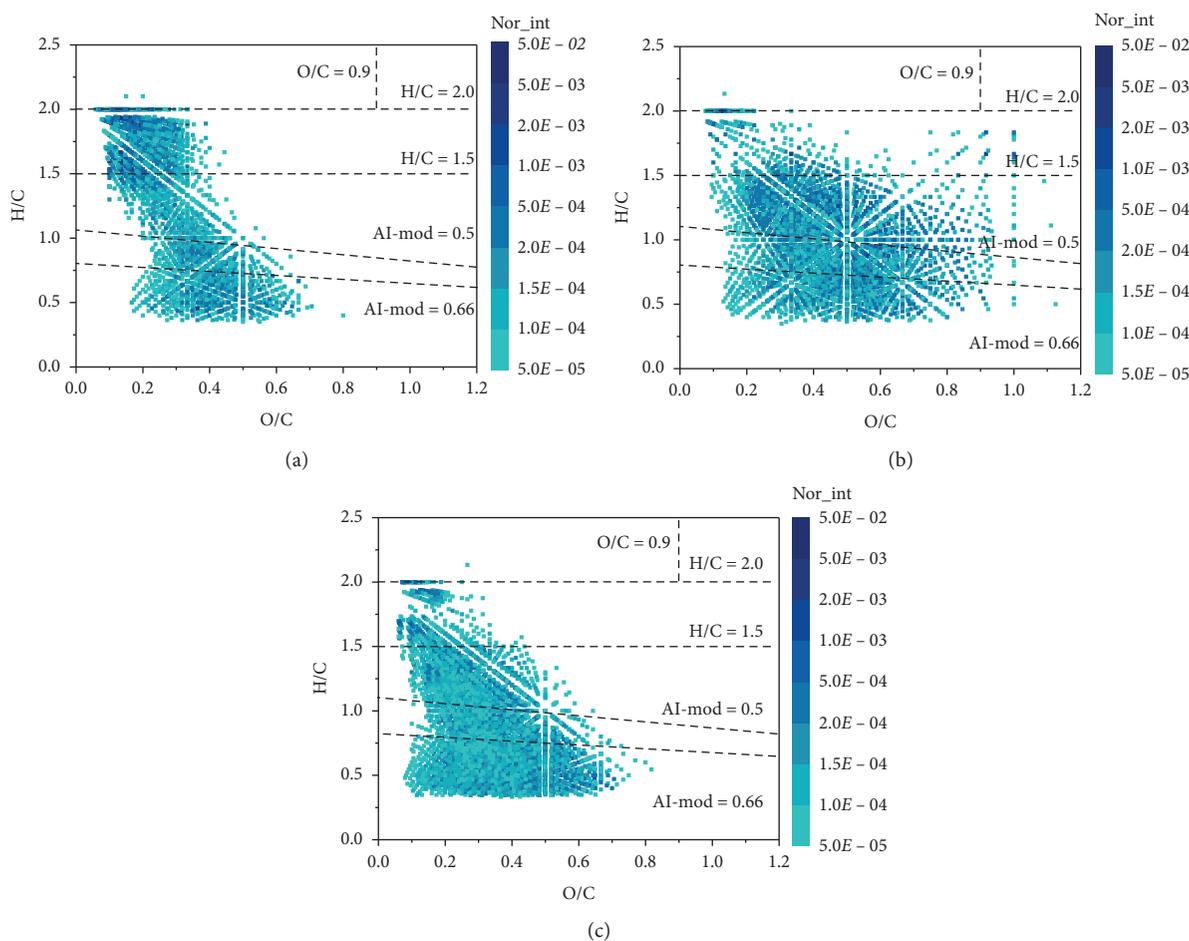


FIGURE 1: Van Krevelen diagrams from the mass spectra of (a) humic acids from Jiufeng forest soil, (b) humic acids from Suwannee River, and (c) humic acids from leonardite. The shades of color represent the peak magnitudes of the molecules. Broken lines divide the pattern into 7 areas which indicate major biomolecular compound classes: (1) polycyclic aromatics (PCAs, $AI-mod > 0.66$), (2) highly aromatic compounds ($0.66 \geq AI-mod > 0.50$), (3) highly unsaturated compounds ($AI-mod \leq 0.50$ and $H/C < 1.5$), (4) unsaturated aliphatic compounds ($2.0 > H/C \geq 1.5$ and $N = 0$), (5) unsaturated aliphatic compounds containing N ($2.0 > H/C \geq 1.5$ and $N > 0$), (6) carbohydrates ($H/C \geq 2.0$ and $O/C \geq 0.9$), and (7) fatty acids and sulfonic acids ($H/C \geq 2.0$ and $O/C < 0.9$). Nor_int is the normalization of the intensity of each molecule.

compounds indicates that LEHA is the most aromatic and has the highest degree of humification. The compounds of SRHA and JFHA are mainly concentrated in medially unsaturated compounds, which are highly unsaturated compounds and unsaturated aliphatic compounds, respectively. Highly unsaturated compounds are more aromatic than unsaturated aliphatic compounds. As a result, SRHA is more aromatic than JFHA.

DBE is roughly positively correlated with m/z among three HA samples. The compounds with DBE greater than 20 are distributed across m/z range of 450–700 DA. The number and peak magnitudes of these molecular formulas both increase in order of JFHA, SRHA, and LEHA (Figure 2). DBE is roughly negatively related to H/C among the three HA samples. The above compounds are polycyclic aromatics and highly aromatic compounds (Figure 3).

TABLE 2: Number of peaks assigned to molecular formulas having similar elemental composition to each compound group and the sum of their peak magnitudes.

Sample name	JFHA	SRHA	LEHA
Number of formulas	2246	3559	4398
Number of formulas similar to fatty acids and sulfonic acids	122 (5)	41 (1) ^a	34 (1)
Carbohydrates	0 (0)	0 (0)	0 (0)
Unsaturated aliphatic compounds containing N	0 (0)	1 (0)	0 (0)
Unsaturated aliphatic compounds	683 (30)	277 (8)	312 (7)
Highly unsaturated compounds	573 (26)	1828 (51)	1273 (29)
Highly aromatic compounds	331 (15)	788 (22)	937 (21)
Polycyclic aromatics	537 (24)	624 (18)	1842 (42)
Sum of peak magnitudes assigned to fatty acids and sulfonic acids ($\times 10^6$)	19052 (25)	2394 (4)	1120 (2)
Carbohydrates ($\times 10^6$)	0 (0)	0 (0)	0 (0)
Unsaturated aliphatic compounds containing N ($\times 10^6$)	0 (0)	3 (0)	0 (0)
Unsaturated aliphatic compounds ($\times 10^6$)	38187 (50)	6524 (11)	3681 (6)
Highly unsaturated compounds ($\times 10^6$)	6717 (9)	31135 (54)	13678 (22)
Highly aromatic compounds ($\times 10^6$)	4374 (6)	11535 (20)	12110 (20)
Polycyclic aromatics ($\times 10^6$)	7517 (10)	6600 (11)	30308 (50)
Sum ($\times 10^6$)	75847	58191	60897

^aValues in parentheses are percentages of either the corresponding number of assigned formulas or the corresponding summed peak magnitudes.

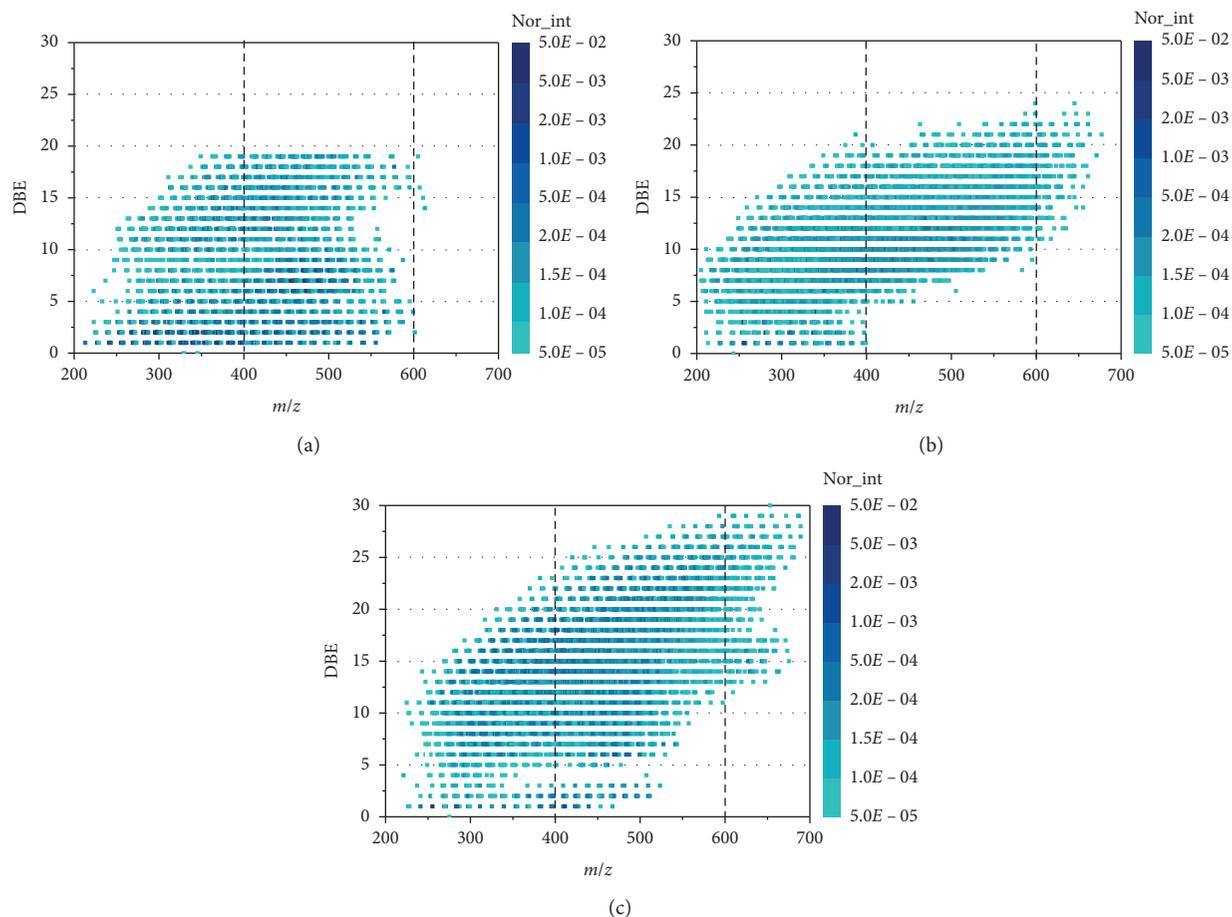


FIGURE 2: Respective DBE- m/z diagrams from (a) humic acids from Jiu Feng forest soil, (b) humic acids from Suwannee River, and (c) humic acids from Leonardite. Nor_int is the normalization of the intensity of each molecule.

Therefore, it can be inferred that the aromatic compounds with higher DBE and larger molecular mass have a significant effect on the aromaticity and humification of HA. Refractory humic acid fractions are marked by increases in

the condensed aromatic components with the highest DBE [40]. Humification processes involved in soil-derived organic matter formation are progressions where H/C ratios are steadily reduced [40] and the maximum molecular mass

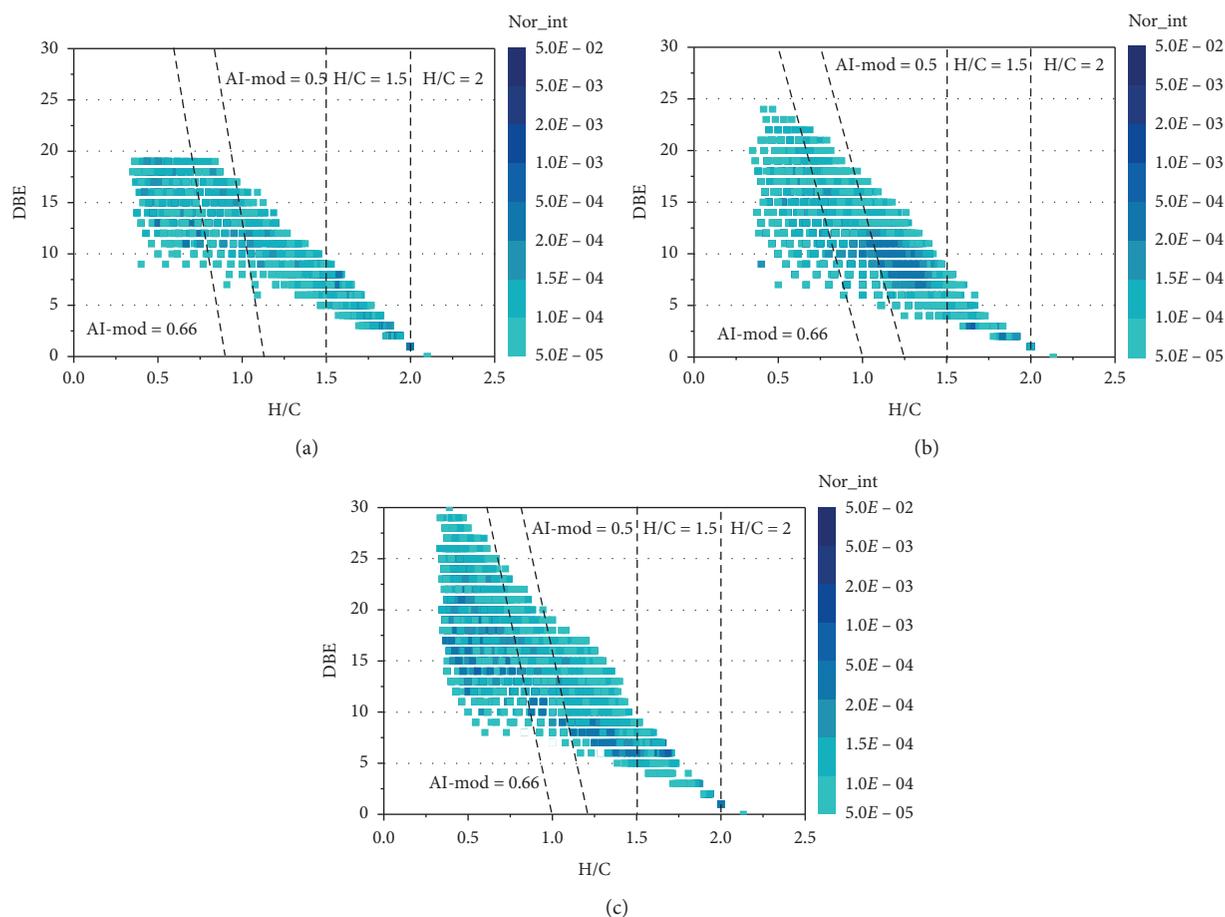


FIGURE 3: Respective DBE-H/C diagrams from (a) humic acids from Jiu Feng forest soil, (b) humic acids from Suwannee River, and (c) humic acids from Leonardite. Nor_int is the normalization of the intensity of each molecule.

of condensed aromatic components is increased [19]. There are more unsaturated aliphatic compounds in JFHA. The molecular formulas of these compounds are distributed across DBE range of 0–5 and m/z range of 400–600. That is, these compounds have low unsaturation and large molecular mass (Figure 2).

Most of molecular formulas for unsaturated aliphatic compounds ($2.0 > H/C \geq 1.5$) are distributed in the m/z range of 200–400 and a few are in the m/z range of 400–500 in SRHA. The molecular formulas for unsaturated aliphatic compounds are distributed densely in the m/z range of 200–600 in JFHA. The distribution of unsaturated aliphatic compounds in LEHA is similar to SRHA, which is sparser. It is inferred that the high content of unsaturated aliphatic compounds in JFHA has a large molecular mass. The distribution of highly unsaturated compounds and highly aromatic compounds in SRHA and LEHA is similar, while the distribution in JFHA is sparser and the distribution range is smaller. Polycyclic aromatics are distributed in the m/z range of 200–700 in LEHA, with the largest distribution range and the largest peak magnitudes. The content of polycyclic aromatics in LEHA is the highest and the molecular mass is the largest, indicating that LEHA has the strongest aromaticity [40] (Figure S2). The contribution of

aromatic compounds with larger molecular weights and higher unsaturation to the overall aromaticity of HAs is the most important.

Individual van Krevelen distribution of molecular formulas identified in each HA is included in the Supporting Information (Figure S3). Across all HAs, the assigned molecular formulas consisted primarily of C, H, and O (CHO) followed by formulas with additional N (CHON₁ and CHON₂) and S (CHOS) [11]. For these three HAs, the CHO molecular formula can make up all classes of compounds, while the CHON and CHOS molecular formulas only make up compounds with a high degree of unsaturation. CHOS molecular formula can form polycyclic aromatics and highly aromatic compounds in SRHA and LEHA. In addition, CHOS molecular formula can also form highly unsaturated compounds with relatively low unsaturation in SRHA. CHON₁ molecular formula is present in all three HAs, and the proportion is increasing according to the order of humification, namely, JFHA < SRHA < LEHA. The CHON₂ molecular formula with highest aromaticity exists in the polycyclic aromatics of JFHA and LEHA, and the aromaticity and unsaturation of the CHON₂ molecular formula are higher than the CHON₁ molecular formula, consistent with previous conclusions on FA [41]. The

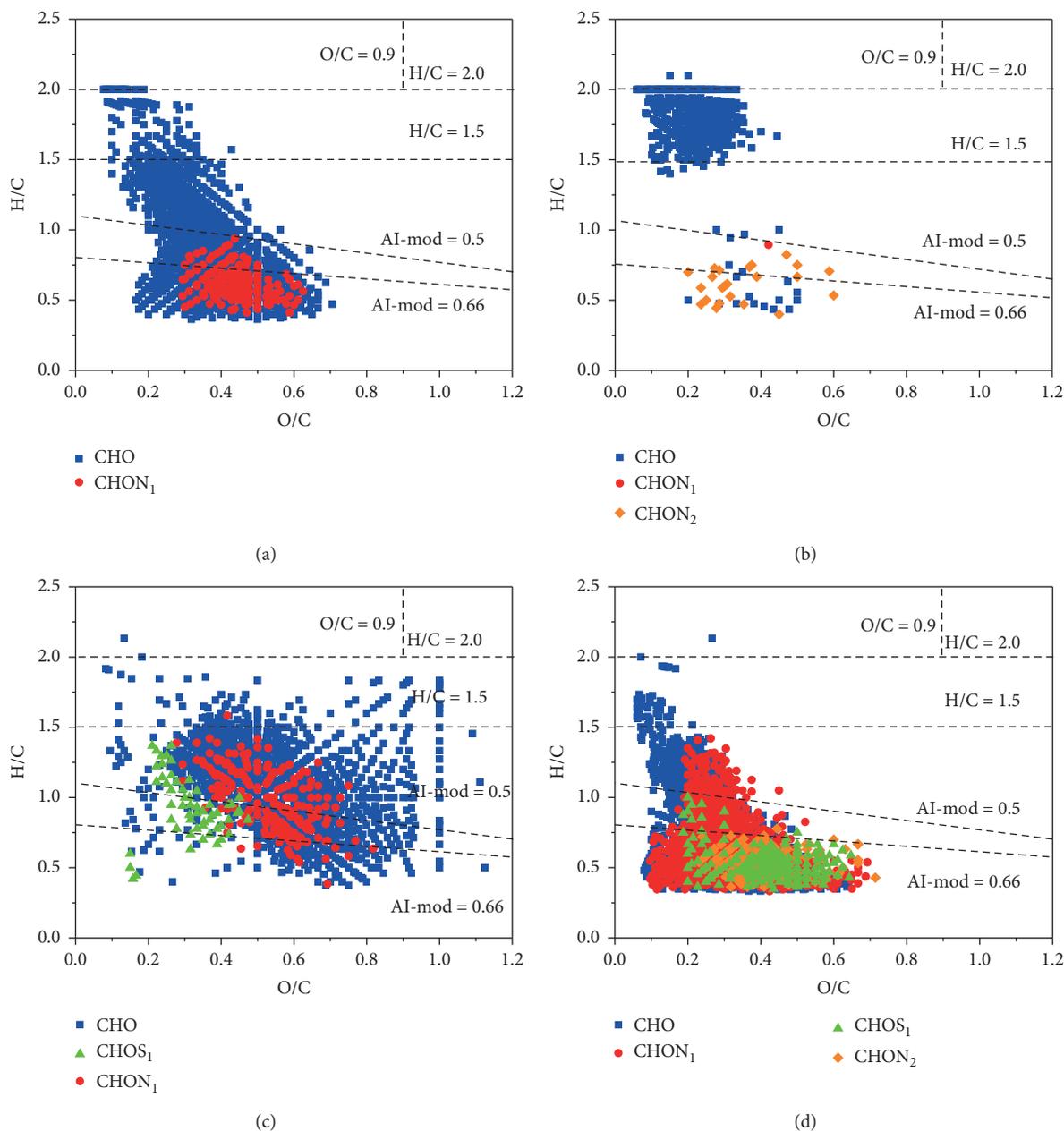


FIGURE 4: Van Krevelen diagrams of molecular formulas (a) common among all humic acid samples, (b) identified only in humic acids from Jiufeng forest soil, (c) identified only in humic acids from Suwannee River, and (d) identified only in humic acids from Leonardite. Formulas include CHO (blue), CHOS₁ (green), CHON₁ (red), and CHON₂ (orange).

aromaticity of compounds increased with the number of heteroatoms (such as nitrogen and sulfur) in the molecular formulas.

Of all formulas assigned to peaks detected by ultrahigh resolution mass spectral analysis ($n = 5329$), approximately 24% ($n = 1282$) are shared among all HA samples. The common molecular formulas consisted of the CHO molecular formulas and the CHON₁ molecular formulas. And the CHON₁ molecular formulas constituted highly aromatic compounds and polycyclic aromatics (Figure 4(a)). There is similarity in these common molecular formulas regardless of climatic regions, land use characteristics, diagenetic process

[42, 43], and other factors. The content of compounds with high aromaticity and unsaturated degree is higher than other compounds and heteroatom-containing compounds are more aromatic and unsaturated. These compounds are accumulation of a similar pool of refractory molecules [44] resulting from the removal of labile DOM components by microbial processing [45] or photodegradation [46]. There are 492 molecular formulas identified only in JFHA, and most of them are unsaturated aliphatic compounds with $1.5 < H/C < 2$ and $0 < O/C < 0.4$. Molecular formulas depleted in O and enriched in H with lower AI-mod may be derived from autochthonous sources and are generally

categorized into biolabile compound classes including carbohydrates, proteins, and lipids [20, 30]. Therefore, JFHA is mainly derived from autochthonous sources [47, 48]. These lipid-like compounds may be the products of secondary soil microbes [49]. There are 1483 molecular formulas detected only in SRHA. These compounds have higher O/C and contain more heteroatom-containing formulas (CHON_1 and CHOS_1). This indicates that SRHA has been anthropogenically impacted, such as urban region and cropland. Anthropogenic inputs to the rivers have altered in-stream DOM composition and increased the content of heteroatoms [11]. 39% of the molecular formulas are detected only in LEHA, including all classes of molecular formulas containing heteroatom, and the overall shape is a right-angled triangle. According to the order of CHO, CHON_1 , and CHON_2 (CHOS_1), the area of each class of molecular formulas becomes smaller and smaller and gradually concentrates on the lower layer of the v-K diagram, that is, the area of aromatic compounds. Therefore, the aromaticity and unsaturation of the molecular formulas containing heteroatoms are significantly higher. Molecular formulas enriched in O and depleted in H with increased AI-mod reflect terrestrial inputs from plant-derived biomacromolecules such as lignins, tannins, and carboxyl-rich acyclic molecules [20, 30, 32, 50]. Compounds containing heteroatoms (nitrogen and sulfur) are more aromatic [16], and the higher the number of heteroatoms, the stronger the aromaticity. At the same time, there is higher content for compounds containing heteroatoms in HAs with higher degree of humification.

3.3. Molecular Lability of Humic Acids. The quality of HA is determined by the mixture of different types of chemical species, including labile and recalcitrant compounds. Microbial-derived bioavailable fractions are classified as more labile organic matter, which has been correlated with high H/C and low O/C ratios in the lipid, protein, and amino sugar-like regions of the v-K diagram [48, 51–54]. Classification of less labile, more recalcitrant components from each carbon source corresponds to molecular species that cluster in more lignin-, tannin-, and black carbon-like regions of the v-K diagram [31]. The MLB_1 molecular richness ranked HAs from most to least labile as JFHA > SRHA > LEHA, whose $\text{MLB}_1\%$ is 75.47%, 15.33%, and 7.88%. JFHA represented the most purely autochthonous, microbially derived HA and has the highest contributions of more hydrogen-saturated species and labile fractions. The CRAM formulas are commonly linked with refractory compounds widely detected in diverse environments including deep ocean [32]. The structural diversity found within CRAM and its substantial content of alicyclic rings and branching contribute to its resistance to biodegradation and refractory nature [32]. The content of CRAM is SRHA, LEHA, and JFHA in order from high to low, which are 51.91%, 18.25%, and 8.28%, respectively. CRAM seems to consist mainly of the decomposition products of biomolecules, as indicated by the prevalence of its carboxyl groups and the manner of oxidation that increases with decreasing molecular size [32]. Therefore, SRHA contains plant-derived lignin- and tannic-like compounds with the relatively small molecular

mass. SRHA may have the most significance for pollutants because CRAM affected the reactivity of organic matter and the bioavailability of associated nutrients and trace metals. The nominal oxidation state of carbon (NOSC) is a formal parameter that gives information about the average oxidation state of all carbons in each chemical formula, regardless of chemical structure [33]. The NOSC ranked HAs from highest to lowest as LEHA > JFHA > SRHA, whose NOSC is -0.07 , -1.16 , and -0.25 . DOM degrades along the gradient from aromatics to aliphatics and high to low nominal oxidation states of carbon [34]. JFHA is the most labile HA because it contains a large amount of unsaturated aliphatic compounds. SRHA is easy to combine with nutrients and trace metals to affect its bioavailability. LEHA shows the highest nominal oxidation states of carbon due to higher degree of humification.

4. Conclusions

Humic acids from different sources of soil, river, and Leonardite were analyzed at the molecular level by ultrahigh resolution mass spectrometry, and thus multiple similarities and differences were found. Despite sources of HA, the compounds containing more heteroatoms (such as nitrogen and sulfur) are more unsaturated and aromatic. AI-mod and DBE consistently indicated that LEHA, SRHA, and JFHA are in order from the highest to lowest degree of humification. JFHA contained amounts of compounds with large molecular mass and low degree of unsaturation. Different classes of compounds uniformly distributed in SRHA, especially lignin- and tannin-like compounds with relatively high degree of unsaturation. LEHA contains an amount of aromatic compounds involving more heteroatoms with high degree of unsaturation and large molecular weight due to high degree of humification. Different molecular composition can cause the diversity of HA lability. JFHA is relatively labile and is prone to interact with microorganisms. The lignin- and tannin-like compounds of SRHA stem from plant-derived organic matter, which are easily combined with nutrients and trace metals to reduce the bioavailability of SRHA. LEHA is relatively less labile due to amounts of aromatic compounds. Our results reveal the diversity of HA due to different degrees of humification at the molecular level and the ways of combining with other substances because of different lability. Ultrahigh resolution FT-ICR MS is a powerful tool to address these novel hypotheses and is of great significance for studying the role of HA from different sources in the pollutant transformation and geochemical cycle at the molecular level.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Authors' Contributions

Shuai Qin and Chengbin Xu contributed equally to this work and Shuai Qin and Chengbin Xu are the co-first authors.

Acknowledgments

This study was sponsored jointly by the National Natural Science Foundation of China (Grant No. 331772057) and the Major Science and Technology Projects in Guangxi Province (Grant No. AA17204078).

Supplementary Materials

Figure S1 includes respective ESI FT-ICR mass spectra from (a) humic acids from Jiufeng forest soil, (b) humic acids from Suwannee River, and (c) humic acids from leonardite. Figure S2 includes respective H/C-*m/z* diagrams from (a) humic acids from Jiufeng forest soil, (b) humic acids from Suwannee River, and (c) humic acids from leonardite. Figure S3 includes van Krevelen diagrams from the mass spectra of (a) humic acids from Jiufeng forest soil, (b) humic acids from Suwannee River, and (c) humic acids from leonardite. Figure S4 includes van Krevelen diagrams of molecular formulas (a) common among in humic acids from Jiufeng forest soil and Suwannee River, (b) common among in humic acids from Jiufeng forest soil and leonardite, and (c) common among in humic acids from Suwannee River and leonardite. (*Supplementary Materials*)

References

- [1] F. J. Stevenson, *Humus Chemistry: Genesis, Composition, Reactions*, John Wiley & Sons, Hoboken, NJ, USA, 1994.
- [2] J. A. Baldock and J. O. Skjemstad, "Role of the soil matrix and minerals in protecting natural organic materials against biological attack," *Organic Geochemistry*, vol. 31, no. 7-8, pp. 697-710, 2000.
- [3] E. S. Krull, J. A. Baldock, and J. O. Skjemstad, "Importance of mechanisms and processes of the stabilisation of soil organic matter for modelling carbon turnover," *Functional Plant Biology*, vol. 30, no. 2, pp. 207-222, 2003.
- [4] E. J. W. Wattel-Koekkoek, P. Buurman, J. Van Der Plicht, E. Wattel, and N. Van Breemen, "Mean residence time of soil organic matter associated with kaolinite and smectite," *European Journal of Soil Science*, vol. 54, no. 2, pp. 269-278, 2003.
- [5] R. Mikutta, M. Kleber, M. S. Torn, and R. Jahn, "Stabilization of soil organic matter: association with minerals or chemical recalcitrance?" *Biogeochemistry*, vol. 77, no. 1, pp. 25-56, 2006.
- [6] B. Allard, "A comparative study on the chemical composition of humic acids from forest soil, agricultural soil and lignite deposit," *Geoderma*, vol. 130, no. 1-2, pp. 77-96, 2006.
- [7] J. Xu, B. Zhao, W. Chu, J. Mao, and J. Zhang, "Chemical nature of humic substances in two typical Chinese soils (upland vs paddy soil): a comparative advanced solid state NMR study," *Science of the Total Environment*, vol. 576, pp. 444-452, 2017.
- [8] J. Dai, W. Ran, B. Xing, M. Gu, and L. Wang, "Characterization of fulvic acid fractions obtained by sequential extractions with pH buffers, water, and ethanol from paddy soils," *Geoderma*, vol. 135, pp. 284-295, 2006.
- [9] L. Grasset and A. Amblès, "Structure of humin and humic acid from an acid soil as revealed by phase transfer catalyzed hydrolysis," *Organic Geochemistry*, vol. 29, no. 4, pp. 881-891, 1998.
- [10] L. Grasset, A. Amblès, and A. Pyrolysis, "Structural study of soil humic acids and humin using a new preparative thermochemolysis technique," *Journal of Analytical and Applied Pyrolysis*, vol. 47, no. 1, pp. 1-12, 1998.
- [11] S. Wagner, T. Riedel, J. Niggemann, A. V. Vähätalo, T. Dittmar, and R. Jaffé, "Linking the molecular signature of heteroatomic dissolved organic matter to watershed characteristics in world rivers," *Environmental Science & Technology*, vol. 49, no. 23, pp. 13798-13806, 2015.
- [12] J. J. Cole, Y. T. Prairie, N. F. Caraco et al., "Plumbing the global carbon cycle: integrating inland waters into the terrestrial carbon budget," *Ecosystems*, vol. 10, no. 1, pp. 172-185, 2007.
- [13] T. J. Battin, L. A. Kaplan, S. Findlay et al., "Erratum: biophysical controls on organic carbon fluxes in fluvial networks," *Nature Geoscience*, vol. 2, no. 8, p. 595, 2009.
- [14] J. Hilton, M. O'hare, M. J. Bowes, and J. I. Jones, "How green is my river? A new paradigm of eutrophication in rivers," *Science of the Total Environment*, vol. 365, no. 1-3, pp. 66-83, 2006.
- [15] S. Bertilsson and B. Jones Jr., "Supply of dissolved organic matter to aquatic ecosystems: autochthonous sources," in *Aquatic Ecosystems*, pp. 3-24, Elsevier, Amsterdam, Netherlands, 2003.
- [16] A. Y. Zhrebek, Y. I. Kostyukevich, A. S. Kononikhin, E. N. Nikolaev, and I. V. Perminova, "Molecular compositions of humic acids extracted from leonardite and lignite as determined by Fourier transform ion cyclotron resonance mass spectrometry," *Mendeleev Communications*, vol. 26, no. 5, pp. 446-448, 2016.
- [17] O. S. Yakimenko and V. A. Terekhova, "Humic preparations and the assessment of their biological activity for certification purposes," *Eurasian Soil Science*, vol. 44, no. 11, pp. 1222-1230, 2011.
- [18] L. Cavani, C. Ciavatta, and C. J. B. T. Gessa, "Identification of organic matter from peat, leonardite and lignite fertilisers using humification parameters and electrofocusing," *Bioresource Technology*, vol. 86, no. 1, pp. 45-52, 2003.
- [19] K. Ikeya, R. L. Sleighter, P. G. Hatcher, and A. Watanabe, "Characterization of the chemical composition of soil humic acids using Fourier transform ion cyclotron resonance mass spectrometry," *Geochimica et Cosmochimica Acta*, vol. 153, pp. 169-182, 2015.
- [20] R. L. Sleighter and P. G. Hatcher, "The application of electrospray ionization coupled to ultrahigh resolution mass spectrometry for the molecular characterization of natural organic matter," *Journal of Mass Spectrometry*, vol. 42, no. 5, pp. 559-574, 2007.
- [21] P. G. Hatcher, K. J. Dria, S. Kim, and S. W. Frazier, "Modern analytical studies of humic substances," *Soil Science*, vol. 166, no. 11, pp. 770-794, 2001.
- [22] J. A. Leenheer and J.-P. Croué, "Peer reviewed: characterizing aquatic dissolved organic matter," *Environmental Science & Technology*, vol. 37, no. 1, pp. 18A-26A, 2003.
- [23] S. McDonald, J. M. Pringle, P. D. Prenzler, A. G. Bishop, and K. Robards, "Bioavailability of dissolved organic carbon and fulvic acid from an Australian floodplain river and billabong,"

- Marine and Freshwater Research*, vol. 58, no. 2, pp. 222–231, 2007.
- [24] J. W. T. Blackburn, W. Kew, M. C. Graham, and D. Uhrin, “Laser desorption/ionization coupled to FTICR mass spectrometry for studies of natural organic matter,” *Analytical Chemistry*, vol. 89, no. 8, pp. 4382–4386, 2017.
- [25] Z. Fang, C. He, Y. Li, K. H. Chung, C. Xu, and Q. Shi, “Fractionation and characterization of dissolved organic matter (DOM) in refinery wastewater by revised phase retention and ion-exchange adsorption solid phase extraction followed by ESI FT-ICR MS,” *Talanta*, vol. 162, pp. 466–473, 2017.
- [26] J. M. Jarvis, A. M. Mckenna, R. N. Hilten, K. C. Das, R. P. Rodgers, and A. G. Marshall, “Characterization of pine pellet and peanut hull pyrolysis bio-oils by negative-ion electrospray ionization fourier transform ion cyclotron resonance mass spectrometry,” *Energy & Fuels*, vol. 26, no. 6, pp. 3810–3815, 2012.
- [27] Introduction to the International Humic Substances Society, <http://humic-substances.org/source-materials-for-ihss-samples/>.
- [28] K. Wang, Y. Pang, C. He et al., “Optical and molecular signatures of dissolved organic matter in Xiangxi Bay and mainstream of Three Gorges Reservoir, China: spatial variations and environmental implications,” *Science of The Total Environment*, vol. 657, pp. 1274–1284, 2019.
- [29] D. He, C. He, P. Li et al., “Optical and molecular signatures of dissolved organic matter reflect anthropogenic influence in a coastal river, northeast China,” *Journal of Environment Quality*, vol. 48, no. 3, 2019.
- [30] B. P. Koch and T. Dittmar, “From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter,” *Rapid Communications in Mass Spectrometry*, vol. 20, no. 5, pp. 926–932, 2006.
- [31] J. D’andrilli, W. T. Cooper, C. M. Foreman, and A. G. Marshall, “An ultrahigh-resolution mass spectrometry index to estimate natural organic matter lability,” *Rapid Communications in Mass Spectrometry*, vol. 29, no. 24, pp. 2385–2401, 2015.
- [32] N. Hertkorn, R. Benner, M. Frommberger et al., “Characterization of a major refractory component of marine dissolved organic matter,” *Geochimica et Cosmochimica Acta*, vol. 70, no. 12, pp. 2990–3010, 2006.
- [33] T. Riedel, H. Biester, and T. Dittmar, “Molecular fractionation of dissolved organic matter with metal salts,” *Environmental Science & Technology*, vol. 46, no. 8, pp. 4419–4426, 2012.
- [34] A. M. Kellerman, D. N. Kothawala, T. Dittmar, and L. J. Tranvik, “Persistence of dissolved organic matter in lakes related to its molecular characteristics,” *Nature Geoscience*, vol. 8, no. 6, pp. 454–457, 2015.
- [35] P. M. Medeiros, M. Seidel, L. C. Powers, T. Dittmar, D. A. Hansell, and W. L. Miller, “Dissolved organic matter composition and photochemical transformations in the northern North Pacific Ocean,” *Geophysical Research Letters*, vol. 42, no. 3, pp. 863–870, 2015.
- [36] T. Šantl-Temkiv, K. Finster, T. Dittmar et al., “Hailstones: a window into the microbial and chemical inventory of a storm cloud,” *PLoS One*, vol. 8, no. 1, Article ID e53550, 2013.
- [37] M. Seidel, M. Beck, T. Riedel et al., “Biogeochemistry of dissolved organic matter in an anoxic intertidal creek bank,” *Geochimica et Cosmochimica Acta*, vol. 140, pp. 418–434, 2014.
- [38] M. Seidel, P. L. Yager, N. D. Ward et al., “Molecular-level changes of dissolved organic matter along the Amazon River-to-Ocean Continuum,” *Marine Chemistry*, vol. 177, pp. 218–231, 2015.
- [39] R. L. Sleighter, G. A. Mckee, and P. G. Hatcher, “Direct Fourier transform mass spectral analysis of natural waters with low dissolved organic matter,” *Organic Geochemistry*, vol. 40, no. 1, pp. 119–125, 2009.
- [40] T. Ohno, Z. He, R. L. Sleighter, C. W. Honeycutt, and P. G. Hatcher, “Ultrahigh resolution mass spectrometry and indicator species analysis to identify marker components of soil- and plant biomass-derived organic matter fractions,” *Environmental Science & Technology*, vol. 44, no. 22, pp. 8594–8600, 2010.
- [41] S. Qin, C. Xu, F. Guo et al., “Molecular signatures of three fulvic acid standard samples as revealed by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry,” *ChemistrySelect*, vol. 4, no. 47, pp. 13940–13946, 2019.
- [42] J. I. Hedges, E. Mayorga, E. Tsamakis et al., “Organic matter in Bolivian tributaries of the Amazon River: a comparison to the lower mainstream,” *Limnology and Oceanography*, vol. 45, no. 7, pp. 1449–1466, 2000.
- [43] J. E. Richey, R. L. Victoria, E. Mayorga, L. A. Martinelli, and R. H. Meade, “Case study 1: integrated analysis of a humid tropical region—the Amazon Basin,” in *Vegetation, Water, Humans and the Climate*, pp. 415–428, Springer, Berlin, Germany, 2004.
- [44] R. Jaffé, Y. Yamashita, N. Maie et al., “Dissolved organic matter in headwater streams: compositional variability across climatic regions of North America,” *Geochimica et Cosmochimica Acta*, vol. 94, pp. 95–108, 2012.
- [45] R. L. Sleighter, R. M. Cory, L. A. Kaplan, H. a. N. Abdulla, and P. G. Hatcher, “A coupled geochemical and biogeochemical approach to characterize the bioreactivity of dissolved organic matter from a headwater stream,” *Journal of Geophysical Research: Biogeosciences*, vol. 119, no. 8, pp. 1520–1537, 2014.
- [46] A. Stubbins, R. G. M. Spencer, H. Chen et al., “Illuminated darkness: molecular signatures of Congo River dissolved organic matter and its photochemical alteration as revealed by ultrahigh precision mass spectrometry,” *Limnology and Oceanography*, vol. 55, no. 4, pp. 1467–1477, 2010.
- [47] R. L. Sleighter and P. G. Hatcher, “Molecular characterization of dissolved organic matter (DOM) along a river to ocean transect of the lower Chesapeake Bay by ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry,” *Marine Chemistry*, vol. 110, no. 3–4, pp. 140–152, 2008.
- [48] A. C. Stenson, A. G. Marshall, and W. T. Cooper, “Exact masses and chemical formulas of individual Suwannee River fulvic acids from ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectra,” *Analytical Chemistry*, vol. 75, no. 6, pp. 1275–1284, 2003.
- [49] T. Riedel, S. Iden, J. Geilich, K. Wiedner, W. Durner, and H. Biester, “Changes in the molecular composition of organic matter leached from an agricultural topsoil following addition of biomass-derived black carbon (biochar),” *Organic Geochemistry*, vol. 69, pp. 52–60, 2014.
- [50] N. Hertkorn, M. Frommberger, M. Witt, B. P. Koch, P. Schmitt-Kopplin, and E. M. Perdue, “Natural organic matter and the event horizon of mass spectrometry,” *Analytical Chemistry*, vol. 80, no. 23, pp. 8908–8919, 2008.
- [51] L. Chipman, D. Podgorski, S. Green, J. Kostka, W. Cooper, and M. Huettel, “Decomposition of plankton-derived dissolved organic matter in permeable coastal sediments,”

- Limnology and Oceanography*, vol. 55, no. 2, pp. 857–871, 2010.
- [52] J. D’Andrilli, C. M. Foreman, A. G. Marshall, and D. M. Mcknight, “Characterization of IHSS Pony Lake fulvic acid dissolved organic matter by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry and fluorescence spectroscopy,” *Organic Geochemistry*, vol. 65, pp. 19–28, 2013.
- [53] A. M. Grannas, W. C. Hockaday, P. G. Hatcher, L. G. Thompson, and E. Mosley-Thompson, “New revelations on the nature of organic matter in ice cores,” *Journal of Geophysical Research*, vol. 111, no. D4, 2006.
- [54] M. P. Bhatia, S. B. Das, K. Longnecker, M. A. Charette, and E. B. Kujawinski, “Molecular characterization of dissolved organic matter associated with the Greenland ice sheet,” *Geochimica et Cosmochimica Acta*, vol. 74, no. 13, pp. 3768–3784, 2010.