Science of the Total Environment 703 (2020) 134764

Contents lists available at ScienceDirect

Science of the Total Environment

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

Linking the molecular composition of autochthonous dissolved organic matter to source identification for freshwater lake ecosystems by combination of optical spectroscopy and FT-ICR-MS analysis

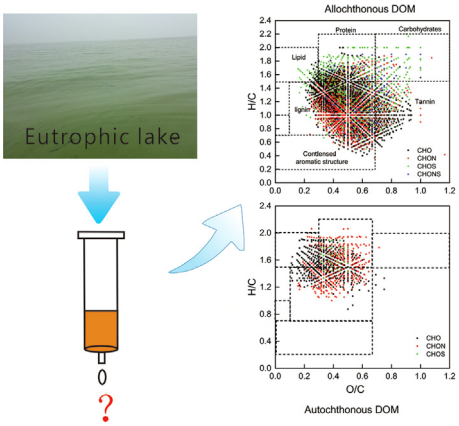
Shasha Liu [a], Zhongqi He [b], Zhi Tang [a,*], Leizhen Liu [c], Junwen Hou [a], Tingting Li [a], Yahe Zhang [d], Quan Shi [d], John P. Giesy [e], Fengchang Wu [a]

[a] State Key Laboratory of Environment Criteria and Risk Assessment, Chinese Research Academy of Environmental Sciences, Beijing 100012, China
[b] USDA-ARS Southern Regional Research Center, 1100 Robert E Lee Blvd, New Orleans, LA 70124, USA
[c] Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
[d] State Key Laboratory of Heavy Oil Processing, China University of Petroleum, 18 Fuxue Road, Changping, Beijing 102249, China
[e] Department of Biomedical and Veterinary Biosciences and Toxicology Centre, University of Saskatchewan, Saskatoon, Saskatchewan SK S7N 5B3, Canada

HIGHLIGHTS

- Autochthonous end-members were characterized by optical spectroscopy and FT-ICR-MS.
- Autochthonous DOM was more aliphatic and less oxidized than allochthonous DOM.
- Molecular DOM composition was directly associated with their optical properties.
- Multivariate analysis with FT-ICR-MS was useful for source analysis of freshwaters.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 7 August 2019
Received in revised form 25 September 2019
Accepted 29 September 2019
Available online 05 November 2019

Editor: Jay Gan

Keywords:
DOM
FT-ICR-MS
Optical spectroscopy
Multivariate analysis
Autochthonous sources
Lake

ABSTRACT

Autochthonous dissolved organic matter (DOM) is increasingly released in lakes due to eutrophication, and thus affects the composition and environmental behaviors of DOM in eutrophic lakes. However, there are only limited studies on the molecular characteristics of autochthonous DOM and its influencing mechanisms. Herein, end-member DOM samples of macrophytes, algae, sediments and freshwater DOM samples in eutrophic lakes (Ch:*Taihu and Dianchi*) were collected and characterized by optical spectroscopy and Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS). The results revealed the chemical structures of autochthonous DOM were more aliphatic and less oxidized, which was marked by increases in lipid compounds and decreases in the lignin components as compared to the allochthonous DOM-dominated freshwaters. More specially, algae-derived DOM contains more lipid compounds, while macrophyte-derived DOM was dominated by lignin and tannin compounds according to Van Krevelen plots. Sediment-derived DOM contained more N-containing compounds. The traditional optical indices indicated the relative aromaticity covaried with polyphenolic and polycyclic aromatics, whereas those reflecting autochthonous DOM covaried with more aliphatic compounds. Multivariate analysis of FT-ICR-MS data of end-members and freshwaters revealed the predominant terrestrial input

* Corresponding author.
E-mail address: tzwork@hotmail.com (Z. Tang).

https://doi.org/10.1016/j.scitotenv.2019.134764

to Lake Taihu and greater contribution of algae released DOM to Dianchi. This study provides critical information about the characteristics of autochthonous DOM at a molecular level and confirmed autochthonous DOM was compositionally distinct from allochthonous DOM. Overall autochthonous DOM should be gained more attention in the eutrophic lakes.

1. Introduction

Dissolved organic matter (DOM) is a compositional diverse assembly of molecules that is ubiquitous in natural waters (He and Wu, 2015; Patriarca et al., 2018). DOM plays a fundamental role in marine and lacustrine water. Increased knowledge of the composition of DOM will allow a better understanding on how it plays its biogeochemical functions (e.g., mobility, bioavailability and effects in the fate of organic pollutants and trace metals) (Cabaniss, 2011; Yan and Korshin, 2014). In addition, an improved understanding of its composition is also helpful in source identification of DOM in complex ecosystems (Antony et al., 2014; Kellerman et al., 2015; Ohno et al., 2010).

DOM is a pool of organic molecules that have various origins either autochthonous or allochthonous (Stedmon et al., 2007). Broadly, DOM is considered allochthonous when derived from degraded terrestrial material, and autochthonous when derived from aquatic macrophytes, phytoplankton, and sediments (McIntyre and Guéguen, 2013; Yang et al., 2016). In the last decades, autochthonous DOM increasingly released from biomass of algae or macrophyte due to the eutrophication and cyanobacterial bloom, and have a significant impact on DOM quality and reactivity in lakes (Zhou et al., 2018). However, limited information is known about the specific molecular characteristics of autochthonous end-member DOM and their possible influence to the biogeochemical processes of lake ecosystems (Gonsior et al., 2019).

Compositional and functional characterization of DOM is generally conducted with wet chemistry and traditional spectroscopic techniques, such as elemental analysis, ultraviolet–visible spectroscopy (Uv–Vis), excitation-emission matrix fluorescence spectroscopy (EEM), nuclear magnetic resonance spectroscopy (NMR) and bulk isotope compositions (Kellerman et al., 2018; Liu et al., 2018b, 2017; Song et al., 2019). However, advanced characterization needs more sophistic techniques on molecular levels. Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) is such an ultrahigh resolution instrumental analysis, which provides detailed information on the molecular constituents through detecting and calculating thousands of individual molecular formulas (Kellerman et al., 2014, 2015). The combination of optical spectroscopy with FT-ICR-MS and the exploration on their linkage would provide more complementary information for DOM characterization (Cortés-Francisco et al., 2014; Herzsprung et al., 2012).

Lake Taihu is one of the regions in recent decades with the most rapid economic growth in China. Its eutrophication has gradually extended from Meiliang Bay to the entire northwest zone of the lake, which has caused the drinking water crisis in late May 2007 (Qin et al., 2010). With a series of pollution control measures implemented since 2007 (Stone, 2011), the water quality has not further deteriorated, but the harmful cyanobacterial blooms still persist and the bloom induced anaerobic "black water" recur throughout summers in this lake (Zhang et al., 2016b). Dianchi Lake, another representative eutrophic freshwater lake in China, also explodes blue algae every year and received much attention in China and even throughout the world. The source identification and alleviation of nutrients are essential in blue algae management. With those concerns of cyanobacterial blooms, there are urgent needs to clarify the sources of organic matters and possible variations of DOM composition in these two lakes. Thus, in this study, DOM from the two lake waters as well as end-member source materials including algae, macrophyte and sediment was analyzed by UV–Vis, EEM and FT-ICR-MS. Multivariate statistical approaches were applied to explore the correlation of FT-ICR-MS spectral data and bulk properties of those DOM. Objectives of this study were: 1) characterizing molecular compositions of DOM in collected lake waters and end-member source materials; 2) linking their DOM molecular features to bulk properties and environmental factors; and 3) employing hierarchal cluster analysis (HCA) for their molecular formula assignments to derive insight on DOM source analysis for natural freshwaters.

2. Material and methods

2.1. Sample collection and preparation

For investigation of macrophyte-derived DOM, we studied *Myriophyllum verticillatum* L., the most abundant submerged species in Lake Taihu. The macrophytes were collected from the Xukou Bay in Lake Taihu and gently rinsed with distilled water, then cleaned using Milli-Q water. Its whole biomass was air-dried, ground to pass through a 1-mm sieve and then stored at −20 °C before use. It should be noted that the whole metabolome was investigated, not just the exo-metabolome, which is generally believed to affect the waters while the macrophytes lived.

For investigation of phytoplankton-derived DOM, 50 L water sample was collected from macrophyte free littoral region of Meiliang Bay in Lake Taihu. The sample was diluted and cleaned to remove CDOM attached on algal bloom using 10 L Milli-Q water. Then the sample was centrifuged through several 50 mL aliquots at 5000*g* for 30 min by using Thermo SORVALL® RC-6 Plus to separate mucilaginous cyanobacterial aggregations, and then freeze-dried, ground and stored at −20 °C before use.

Sediments were collected from two sites of Lake Taihu, macrophyte free littoral of Meiliang Bay and phytoplankton free littoral of Eastern Lake, respectively (Fig. S1), by use of a grab dredge. Sediments were transported to the laboratory on ice and immediately centrifuged at 8000*g* for 15 min. Filtrates were stored at −20 °C until use.

In order to verify the feasibility of source identification of freshwaters by using the chemical composition of end-members, several water samples were also collected from eutrophic lakes. Water samples were collected from eight sites throughout Lake Taihu and one site in Lake Dianchi for comparison (Supporting Information, SI, Fig. S1, Table S1) in 2-L glass bottles (acid-washed and precombusted) in May 2018 and transported to the laboratory in ice boxes within 4 h. After sampling, the water was immediately filtered with 0.47 μm GF/F glass fibers (Whatman, precombusted at 450 °C) and kept at 4 °C until analysis.

2.2. DOM extraction and bulk analysis

After the samples collected, DOM was extracted using a common method that allowed extraction of relatively fresh DOM because no further decomposition treatment was required (Bai

et al., 2017a). Details of the extraction procedure have been reported previously (Liu et al., 2016a,b). Briefly, the sample powder (1 g of macrophyte or 100 mg algae) were shaken in 30 mL Milli-Q water (1:30 w/v for macrophyte, 1:300 w/v for algae) for 18 h, respectively on a shaker at room temperature (25 °C). The DOM fractions in these supernatants, as well as in the supernatants of sediments and water samples were directly used for chemical analysis, UV–Vis and EEM measurement. For FT-ICR-MS, those samples were further enriched by solid phase extraction (SPE) as described below (Dittmar et al., 2008).

Dissolved organic carbon (DOC) and total dissolved nitrogen (TN) concentrations were determined by a multi N/C 3100 analyzer (Jena, Germany). Total phosphorous (TP) was determined using the molybdenum blue method after digestion at 120 °C for 40 min (Zhu et al., 2008). Chlorophyll-*a* (Chl-*a*) was quantified by use of spectrophotometry at wavelengths 665 nm and 750 nm, following extraction with hot (~80 °C) ethanol (90%) (Zhang et al., 2013).

UV absorbance spectra from 200 to 600 nm were measured with a UV–vis spectrophotometer (Agilent 8453, Wilmington, DE, USA), by use of a 1-cm path length cell. Specific ultraviolet absorbance ($SUVA_{254}$, L/mg·m) is the absorbance of ultraviolet light in a water sample at 254 nm that is normalized for dissolved organic carbon (DOC) concentration. The absorption ratio of E_2:E_3 is the ratio of absorbance at 250 to 365 nm and E_4:E_6 is the ratio of absorbance at 465 to 665 nm. Spectral slope of absorbance coefficients between 275 and 295, and 350 and 400 nm, were obtained by non-linear fitting of the exponential model: $a_\lambda = a_{\lambda 0}\, e^{S(\lambda_0 - \lambda)}$ where $\lambda_0 > \lambda$ and S is the spectral slope in the $\lambda_0 > \lambda$ spectral range. The spectral slope ratio (S_R) was defined as the ratio of the spectral slopes S of the shorter (275–295 nm) to the longer (350–400 nm) wavelength ranges (Helms et al., 2008b).

Three-dimensional excitation-emission matrices (EEMs) fluorescence spectroscopy were measured using a Hitachi F-7000 fluorescence spectrometer (Hitachi High-Technologies, Japan) with a 400-voltage xenon lamp at a scanning speed of 2400 nm/min and at room temperature (20 ± 2 °C). The scan ranges for excitation and emission were 200–450 nm (every 5 nm) and 250–600 nm (every 1 nm), respectively. The bandpass widths for excitation and emission were both set to 5 nm. Fluorescence index (FI) was calculated as the ratio of emission intensity at 450 nm to that of 500 nm for a fixed excitation wavelength of 370 nm (Cuss and Guéguen, 2015b).

2.3. FT-ICR-MS sample preparation and analysis

The DOM fractions in extracts of macrophyte and algae, as well as in the supernatants of sediments and water samples were further enriched with solid phase extraction cartridges for molecular characterization using Bruker Solarix 9.4 T FT-ICR-MS (Li et al., 2016). First, the SPE cartridges (Varian Bond Elute PPL, 1 g/6 mL) were rinsed with two cartridge volumes (6 mL) of methanol (MS grade). Each DOM sample was acidified with pure HCl to pH 2 and passed through the SPE cartridges by vacuum pump at a flow rate of about 2 mL/min. Then, the cartridges were rinsed with two cartridge volumes of 0.01 M HCl for the complete removal of salts. After the washing step, the cartridges were dried with nitrogen gas for 10 min and were then eluted with 10 mL of methanol (at the same ratio for up-scaling experiments). Aliquots leachates (1 mL) were N_2 dried and then diluted by 50 mL Milli-Q water to measure and calculate extraction efficiency by PPL. The remaining leachates were freeze-dried and re-dissolved in methanol (LC-MS Chromasolv grade) for the analysis of FT-ICR-MS.

For mass spectral analysis, the above leachates in methanol (i.e., DOM samples) were diluted with ultrapure water to DOC concentration of 20 mg/L for analysis on the 9.4 T FT-ICR-MS instrument. Samples were continuously infused into the electrospray ionization (ESI) unit at a flow rate of 120 μL/h in negative ion mode. The lower and upper mass limit was set to *m*/*z* 150 and 2000, respectively. Spray shield voltage was set to 3.0 kV. The capillary voltage was set to 3.5 kV, and the capillary column end voltage was −320 V. Mass spectra were collected over 500 scans, with an ion accumulation rate time of 0.02 s. Mathematically possible formulas were calculated with a signal-to-noise ratio ≥6. The detection error was limited to within 1 ppm absolute mass.

Molecular formulas containing the elements C, H, O, N, and S were assigned using a self-written software routine (Riedel et al., 2012) based on previously suggested criteria (Koch et al., 2007; Stenson et al., 2003). The modified aromaticity index (AI-mod) (Koch and Dittmar, 2006) and double bond equivalent (DBE) were calculated for each assigned molecular formula. These molecular descriptors are both indicative of the degree of aromaticity within a molecule. These and other molecular parameters (m/z, H/C, O/C, etc.), derived from peak formula assignments were expressed as intensity-weighted average (wa) values. Therefore, these values directly reflected the relative contribution of each *m*/*z* peak to the entire DOM mass spectrum.

Van Krevelen diagrams can help distinguish compound classes in samples on the basis of characteristic elemental ratios that each major biogeochemical compound class generally exhibits (Sleighter and Hatcher, 2007). The groups are delineated by the aromaticity index (AI) and H/C cutoffs (Kellerman et al., 2014; Koch and Dittmar, 2006): combustion-derived polycyclic aromatics (AI > 0.66), vascular plant-derived polyphenols (0.66 ≥ AI > 0.50), highly unsaturated and phenolic compounds (AI ≤ 0.50 and H/C < 1.5), and aliphatic compounds (2.0 ≥ H/C ≥ 1.5). The van Krevelen space for this study was also divided into seven discrete regions by elemental ratios (Hockaday et al., 2009): lipids (H:C = 1.5–2.0; O:C = 0–0.3); proteins (H:C = 1.5–2.2; O:C = 0.3–0.67; N/C ≥ 0.05); lignins (H:C = 0.7–1.5; O:C = 0.1–0.67); carbohydrates (H:C = 1.5–2.0; O:C = 0.67–1.2); unsaturated hydrocarbons (H:C = 0.7–1.0; O:C = 0–0.1); condensed aromatic structures (H:C = 0.2–0.7; O:C = 0–0.67); and tannin (H:C = 0.5–1.5; O:C = 0.67–1.2).

2.4. Statistical analysis

Linear correlation analyses (Pearson) of FT-ICR-MS data and optical parameters, water quality parameter were performed with IBM SPSS Statistics software (Version 22, IBM Corp.). Redundancy analysis (RDA) of FT-ICR-MS data and these parameters in R was conducted by package vegan. Hierarchal cluster analysis (HCA) was determined using between-groups linkage by applying squared Euclidean distance to measure the degree of similarity of FT-ICR-MS data of DOM between end-member and lake water samples. Deriving the squared Euclidean distance between two data points involves computing the sum of the squares of the differences between corresponding values. The similarities between the analyzed samples were presented in the dendrogram for each sample. The dendrogram similarity scales that are generated by the SPSS program range from zero (greater similarity) to 25 (lower similarity).

3. Results and discussion

3.1. Bulk spectral parameters analysis

Optical indices of UV–Vis absorbance and fluorescence spectroscopy of DOM from end-members and freshwaters were listed in Table S2. $SUVA_{254}$ values of DOM extracted from algae (0.09 L/mg·m), macrophyte (0.06 L/mg·m) and sediment (0.43–0.48 L/

mg·m) were lower than those in lake waters (0.83–2.33 L/mg·m) (Table S2). This observation revealed the low aromaticity and degree of humification of these end-member materials, apparently due to the fact that they were the precursors of DOM in water of lake (Cuss and Guéguen, 2015a; Zhang and He, 2015). Values of E_2/E_3, $S_{275-295}$, and S_R were lower in macrophyte-derived DOM than other collected DOMs, which might indicate the relatively larger percentage of high molecular weight in macrophyte-derived DOM than others. Inversely, algae- and sediment-derived DOM showed greater values, suggesting their lower molecular weight compared with macrophyte-derived DOM (Baker, 2001). FI determines source of DOM, which is either: microbial (high FI > 1.8, derived from extracellular release and leachate from bacteria and algae) or terrestrially derived (low FI < 1.2, terrestrial plant and soil organic matter) (Fellman et al., 2010). Sediment and algae-derived DOM had a higher FI values, which was consistent with previously reported (Fellman et al., 2010). However, macrophyte-derived DOM had lower FI value, which might indicate its more approximate property with terrestrial plant. In addition, values of S_R (6.81–7.16) and FI (1.81–1.90) for sediment-derived DOM are similar to those observed in microbial produced DOM previously reported (Derrien et al., 2017; McKnight et al., 2001; Yakimenko et al., 2018), which might indicate extensive microbial activities at interfaces between sediment and water. For the two freshwater lakes, the higher $SUVA_{254}$ and S_R values in Lake Taihu than Dianchi suggested the higher molecular weight and aromaticity in water of Lake Taihu than that in Dianchi, which could be attributed to the origin and the altered degree through biogeochemical process of DOM in lake ecosystems (Helms et al., 2008a). Additionally, the remarkable distinct of FI values between these two lakes also revealed their different DOM source, which would be discussed specially in section 3.4.

3.2. FT-ICR-MS spectral analysis of SPE-DOM

The extraction efficiency by PPL cartridges was about 8.2–37.9% possibly due to extensive loss of highly polar saccharides and small organic acids, which are too polar to be extracted (Gonsior et al., 2019). The presented results are only addressing the extractable and ionizable component of DOM. Spectra from negative-ion FT-ICR-MS analyses of DOM extracted by SPE exhibited thousands of peaks in the range of m/z from 200 to 750 Da (Figs. S2 and S3). The mean magnitude-weighted average (wa) composition varied among those DOM samples (Table 1). In comparison with water DOM samples in Lake Taihu, end-member autochthonous DOM samples extracted from macrophytes, algae, and sediment showed lower m/z_{wa}, C_{wa}, O_{wa}, AI-mod_{wa}, DBE_{wa}, and higher H/C_{wa}. There were least difference in the values of those parameters between Lake Taihu water and macrophyte DOM samples, but the differences between Lake Taihu water and other samples were apparent. Especially, algal DOM showed the greatest weight-averaged H/C_{wa} and least O/C_{wa}, AI-mod_{wa} and DBE_{wa} values. Compared to Lake Taihu, those mean molecular parameters of water in Lake Dianchi were more similar with algae-derived DOM.

Among samples examined, assigned molecular formulas consisted primarily of C, H and O (CHO) followed by formulas with additional N (CHON), S (CHOS) and NS (CHONS) (Fig. 1). The relative contribution of only CHO formulas was greater for DOM extracted from the macrophyte (79.9%), algae (86.1%) and Lake Dianchi water (87.1%) than that in Lake Taihu water (mean value: 59.5%) and sediments (mean value: 57.4%). However, greater proportions of formulas containing N and S were observed in sediment-derived DOM (N: 29.0–41.3%, S: 3.3–11.6%) and Lake Tai water (N: 24.2–27.2%, S: 11.6–14.9%) than those DOM extracted from macrophyte (N: 15.7%, 3.7%), algae (N: 12.9%, S: 1.0%) and Dianchi water (N: 9.4%, S: 3.6%). The extracted DOM in freshwater from Lake Taihu showed lesser proportion of exclusively CHO compounds than DOM in samples from the surface ocean, deep ocean and river samples (72.2–93.7%) (Kujawinski et al., 2009), while DOM samples from macrophytes, algae and Lake Dianchi water were intermediate to those values. This might be due to greater abundances of N-and S-containing molecular formulas from terrigenous and anthropogenic inputs to Lake Taihu (Wagner et al., 2015). The high abundance of CHON formulas in sediment samples might be due to greater activities of bacteria at the interface between water and sediments (Wagner et al., 2015), which was consistent with the UV–vis and fluorescence observations. It has also been reported that N-containing components of OM play a key role in its stabilization through N interactions with mineral surfaces (Hsu and Hatcher, 2005). Thus another explanation of greater percentage of CHON formulas in sediment might be the sequestration of N during humification (Hsu and Hatcher, 2005).

Van Krevelen (V-K) diagrams have been widely used to visualize complex DOM mixtures based on H/C and O/C ratios of molecular formulas assigned to individual mass spectral peaks (Kellerman et al., 2018; Kim et al., 2003). The V–K plots and relative abundances of chemical classes are shown in Figs. 2a–f and 3. Macrophyte-derived DOM was dominated by lignin materials (50.6%), followed by tannins (17.4%) and lipids (12.6%). This was similar with that observed in DOM derived from terrigenous plant previously reported using FT-ICR-MS (Ohno et al., 2010). The relative abundance of tannins were greatest in macrophyte-derived DOM than other collected DOM samples in our study, which might

Table 1
Intensity-weighted mean molecular parameters of different samples derived from assigned molecular formulae.[a]

Sample	m/z_{wa}	C_{wa}	H_{wa}	O_{wa}	N_{wa}	S_{wa}	H/C_{wa}	O/C_{wa}	AI-mod_{wa}	DBE_{wa}
TH-1	412.47	18.7	22.14	9.64	0.47	0.15	1.24	0.53	0.23	8.37
TH-2	408.4	18.47	21.82	9.56	0.47	0.16	1.24	0.53	0.23	8.29
TH-3	412.2	18.56	21.32	9.87	0.44	0.12	1.21	0.54	0.25	8.62
TH-4	412.51	18.74	22.26	9.63	0.46	0.15	1.23	0.52	0.23	8.34
TH-5	405.96	18.51	22.68	9.43	0.42	0.17	1.23	0.52	0.24	8.38
TH-6	409.55	18.67	22.13	9.53	0.45	0.14	1.24	0.52	0.24	8.32
TH-7	416.07	18.83	21.83	9.82	0.48	0.14	1.21	0.53	0.25	8.65
TH-8	413.75	18.71	22.09	9.75	0.48	0.13	1.24	0.53	0.23	8.41
Average	411.36	18.65	22.03	9.65	0.46	0.15	1.23	0.53	0.24	8.42
Macrophyte	380.6	17.18	21.05	9.28	0.24	0.04	1.29	0.55	0.2	7.27
Algae	337.9	18.02	29.13	5.48	0.31	0.01	1.67	0.31	0.08	4.11
Sediment-1	366.99	17.73	23.82	7.4	0.58	0.12	1.41	0.43	0.18	6.61
Sediment-7	343.76	16.37	23.13	6.81	1	0.03	1.48	0.42	0.14	5.81
Lake Dianchi	341.55	18.16	26.93	5.83	0.15	0.04	1.54	0.33	0.15	5.27

[a] Intensity weighted average values are displayed for r mass-to-charge ratio (m/z_{wa}), number of carbon (C_{wa}), hydrogen (H_{wa}), oxygen (O_{wa}), nitrogen (N_{wa}), and sulfur atoms (S_{wa}), hydrogen to carbon ratio (H/C_{wa}), oxygen to carbon ratio (O/C_{wa}), modified aromaticity index (AI-mod_{wa}) and double bond equivalent (DBE_{wa}).

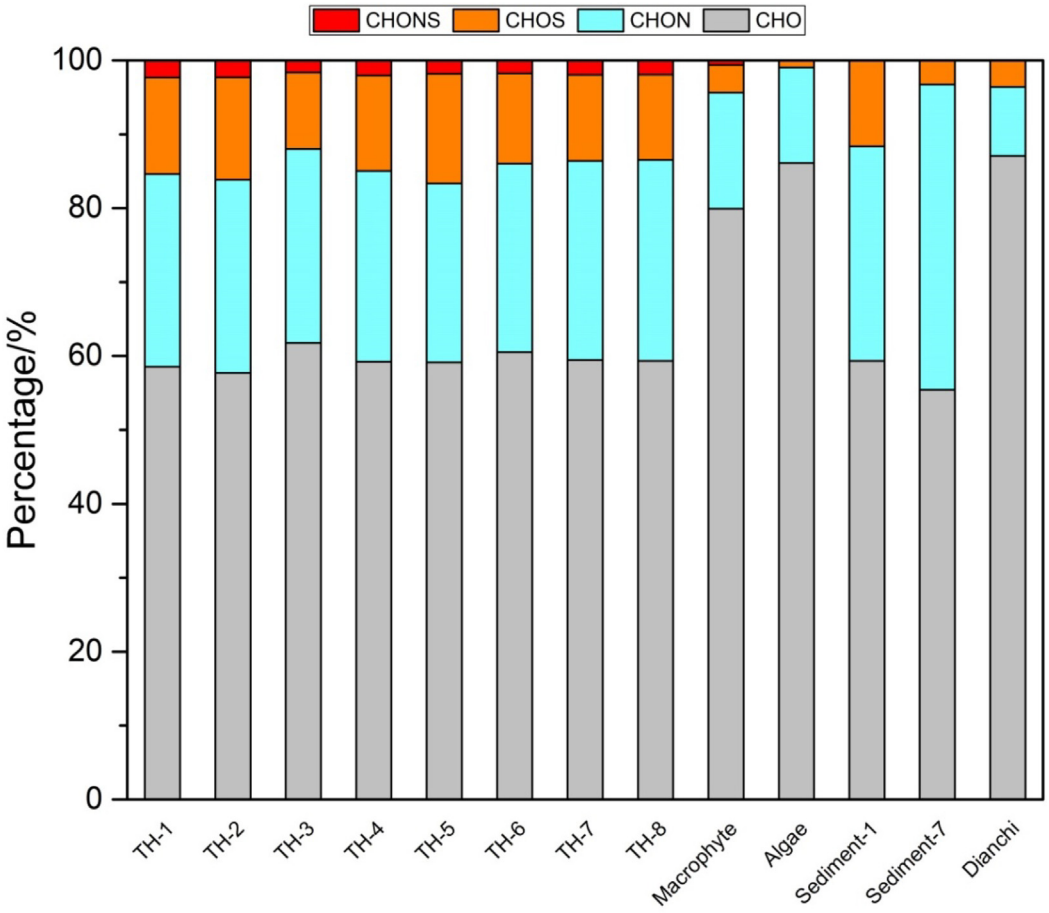


Fig. 1. Comparisons of compositions of DOM by major subcategories in various DOM samples.

be due to releases of phenolic acids richness in macrophytes (Liu et al., 2017). Algae-derived DOM contained greater proportions of lipids (56.8%), followed by lignin (15.0%) and protein (7.4%), but fewer tannins compared with macrophyte-derived DOM. The predominance of lipid compounds and absence of tannin formulas might therefore serve as an indicative feature of algae-derived DOM. Sediment-derived DOM was dominated by lignin (51.2–59.7%), followed by proteins (8.0–18.5%) and lipids (14.2%-16.7%). It was reported that lignin region potentially includes carboxyl rich alicyclic molecules (CRAMs) which would fall into the same space of the V-K diagrams (Hertkorn et al., 2006). The removal of labile DOM components by microbial processing has been suggested to result in the accumulation of CRAMs in sediment (Wagner et al., 2015). Thus, sediment-derived DOM here may reflect the preferential accumulation of recalcitrant compounds in the digenetic processing in sediment (Liu et al., 2016a; Wagner et al., 2015). It has been reported eutrophication promotes the accumulation of autochthonous DOM (especially algal sources) in lakes (Zhou et al., 2018). Thus, algae-derived compounds, including aliphatic substances and protein compounds would increase and thus affect the composition of DOM in eutrophic lakes. As a result, a fraction of autochthonous DOM explosively accumulated in lakes due to nutrient enrichment and algae bloom is likely to be highly biolabile and could be quickly utilized to support a growing community of heterotrophs (Kellerman et al., 2015). In addition, the binding capacity of DOM with metal or organic contaminants would also be altered and thus influence their bioavailability (Liu et al., 2018c; Yang et al., 2016).

Van Krevelen distributions of formulas identified in water samples of Lake Taihu were similar (Figs. S4 and 2e), with a dominance of lignin formulae (75.4–79.8%). The predominance of lignin components in Lake Taihu is similar to those frequently observed in global rivers and lakes (Kellerman et al., 2014; Wagner et al., 2015). These formulas were mainly heteroatomic and mostly were classified as highly unsaturated aliphatics. The CHOS and CHON formulas identified in Lake Taihu were more comparable to compounds found in abundance in wastewater effluents (Gonsior et al., 2011) and septic-impacted groundwaters (Arnold et al., 2014), which might suggest the wastewater inputs to Taihu from anthropogenic influence. Furthermore, the removal of labile DOM components by microbial processing or photo-degradation has been suggested to result in accumulation of refractory molecules such as CRAMs in lignin or protein region (Jaffé et al., 2012). DOM in water collected from Lake Dianchi enriched in lipid (43.8%) and lignin (40.9%), with limited amounts of tannin materials. The high percentage of lipid compounds in Dianchi might further indicate the contribution of DOM from algae. The DOM extraction selectivity of the SPE cartridges with respect to certain DOM compounds and also the use of electrospray ionization in both negative ion mode (ESI^-) and positive ion mode (ESI^+) for the characterization of DOM are important factors to distinguish compound classifications (Li et al., 2017; Ohno et al., 2016). PPL was reported to retain higher proportions of nitrogen-containing compounds from both fresh- and marine waters (Dittmar et al., 2008; Perminova et al., 2014). In general, there was a preferential increase in the number of assignments for the aliphatic and

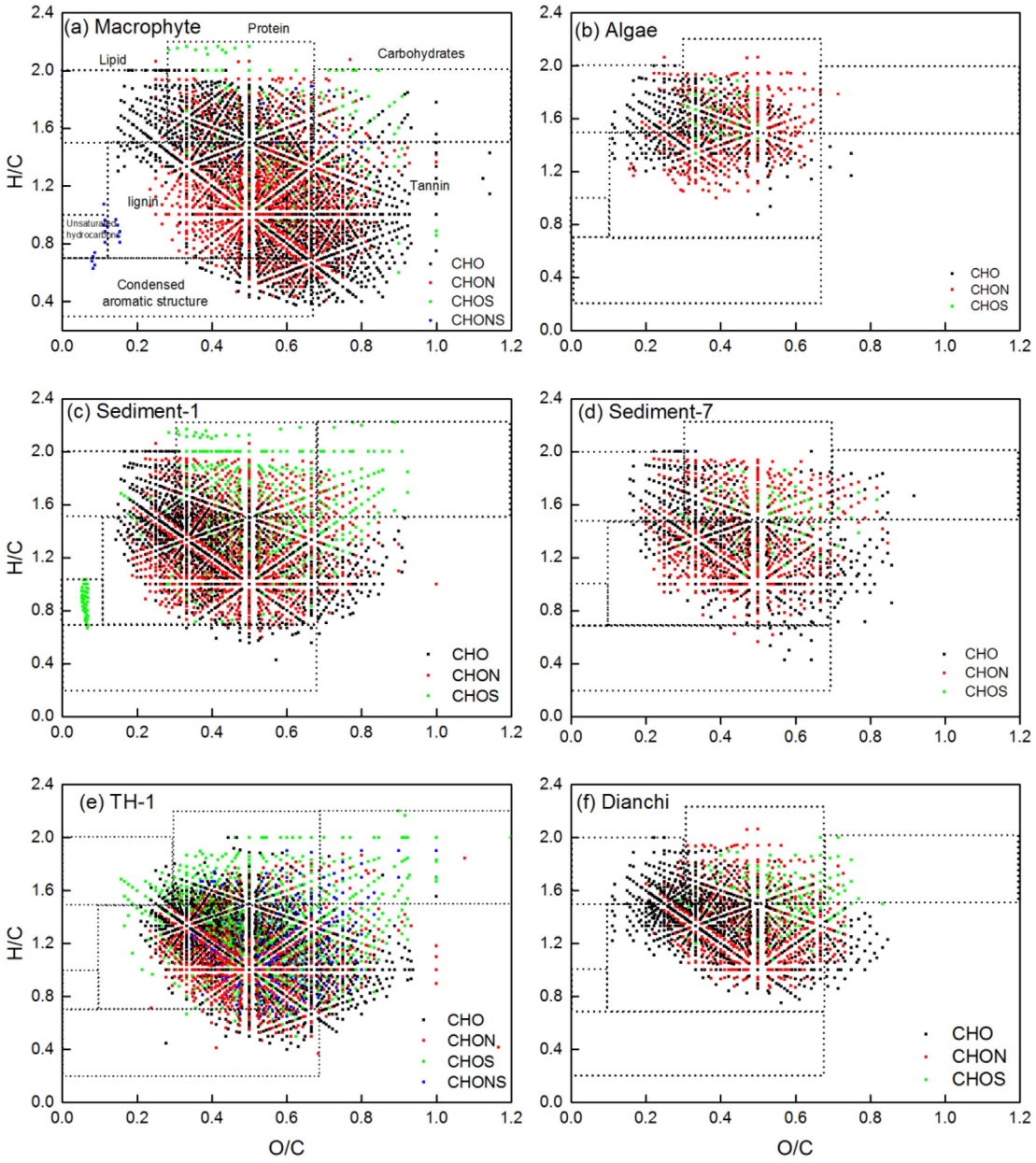


Fig. 2. Van Krevelen diagrams from the mass spectra of DOM derived from (a) macrophyte, (b) algae, (c and d) sediment, (e) Lake Taihu and (f) Lake Dianchi. Ovals overlain the plots indicate major compound classes: lipids (H:C = 1.5–2.0; O:C = 0–0.3); protein (H:C = 1.5–2.2; O:C = 0.3–0.67; N/C $\geq$ 0.05); lignins (H:C = 0.7–1.5; O:C = 0.1–0.67); carbohydrates (H:C = 1.5–2.0; O:C = 0.67–1.2); unsaturated hydrocarbons (H:C = 0.7–1.0; O:C = 0–0.1); condensed aromatic structures (H:C = 0.2–0.7; O:C = 0–0.67); and tannin (H:C = 0.5–1.5; O:C = 0.67–1.2).

carbohydrate-like DOM components in the ESI^+ mode (Ohno et al., 2016). Thus, It should be noted that, in our study, the relative abundance of aliphatic and carbohydrate components might be underestimated by PPL extraction and ESI^- mode.

3.3. Linking DOM molecular compositions to optical properties and environmental factors

In order to further clarify the unique features of various sources, multivariate analysis of molecular composition using redundancy analysis (RDA) were conducted (Fig. 4). The input parameters for the ordination included bulk properties (i.e. TN, DOC, $SUVA_{254}$, E_2:E_3, E_4:E_6, S, S_R, and FI) and FT-ICR-MS-derived molecular composition data [i.e. weight-averaged parameters H/C_{wa}, O/C_{wa}, DBE_{wa}, $AI\text{-}mod_{wa}$ and relative abundances (%) of classified compounds]. The compound categories based on V-K diagrams included aliphatic compounds (aliphatics), polyphenolic compounds (polyphenols), highly unsaturated and phenolic compounds (HUPs), polycyclic aromatics (PolcArom), lipid, protein, carbohydrate, lignin, tannin and condensed aromatic structure. The results showed a clear separation between categories of greater aromaticity in DOM in water samples from Lake Taihu on the negative end of the first axis (RDA1) with aliphatic DOM from algae, macrophyte and sediments on the positive end (Fig. 4). The RDA1 explained 89.7% of the variance in the sample and is closely related to the nucleo-cytoplasmic ratio, aromaticity, and contents of oxygen. Samples that have greater m/z, more aromaticity and greater oxygen content (i.e. great m/z_{wa}, DBE_{wa}, $AI\text{-}mod_{wa}$, O_{wa}, O/C_{wa}) exhibited negative RDA1, while samples that were more aliphatic (i.e. great H_{wa} and H/C_{wa} values) exhibited more positive RDA1. All

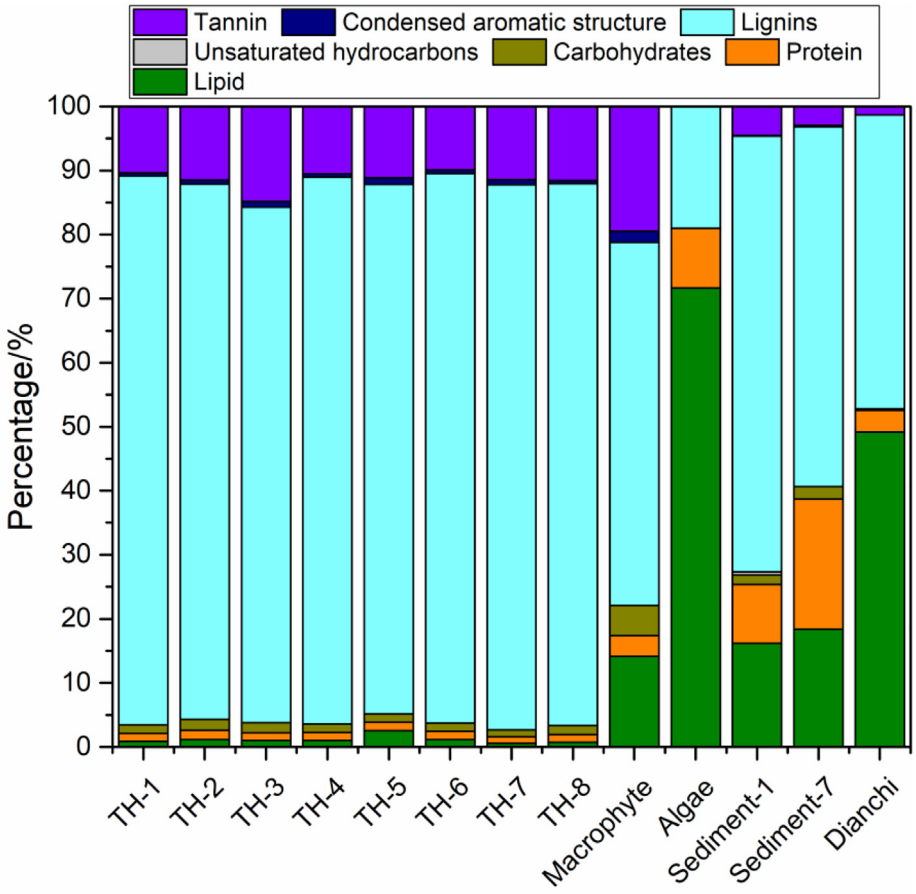


Fig. 3. Relative abundance (%) of van-Krevelen diagram-derived classification classes from the FT-ICR-MS analysis of SPE-DOM components extracted from various sources (macrophyte, algae, sediment) and natural freshwaters (Lake Taihu and Lake Dianchi). The major ionizable compound classes are divided into lipids, proteins, carbohydrates, unsaturated hydrocarbons, lignins, tannins, and condensed aromatic structures.

samples of water from Lake Taihu exhibited negative RDA1 values, which indicated their more aromaticity and greater oxygen contents than other samples collected in the study. Conversely, autochthonous DOM derived from macrophytes, algae and sediment as well as water sample in Dianchi showed more aliphatic feature. The result of autochthonous DOM distributed in positive RDA1 was in agreement with previous studies, which found autochthonous DOM was characterized with molecular formulas depleted in O and enriched in H with lower AI-mod and generally include biolabile compound classes including proteins and lipids (Koch and Dittmar, 2006; Sleighter and Hatcher, 2007). Molecular formulas with these features in negative RDA1 might reflect terrestrial inputs from plant-derived compounds such as lignins and tannins. $SUVA_{254}$ covaried with Lake Taihu water samples with a large proportion of polyphenolic compounds, polycyclic aromatics, unsaturated and phenolic compounds, and condensed aromatic structures, such as lignins and tannins (Fig. 4). Conversely, DOC, $S_{275-295}$, FI, TN were positively related with RDA1 and samples that were less aromatic and more aliphatic and dominated by DOM from phytoplankton and sediment (Fig. 4). This was consistent with lipids, aliphatic and proteins that were categorized into the positive end of the RDA1. In brief, variation explained by RDA1 was closely related to the trend in DOM from aromatic to aliphatic, which corresponds to the distinguishable molecular characteristics shift from the more aromatic, allochthonous DOM to the more aliphatic, autochthonous DOM. Segregation of aromatic contributions from polyphenolic compounds versus aliphatic DOM from lipid compounds is consistent with past findings from previous studies in which sources of DOM were investigated by using of FT-ICR-MS (Kellerman et al., 2018).

RDA2 explained only 4.9% of the variation and is related to nitrogen content (N_{wa}). Samples with greater contents of N_{wa} exhibited a positive RDA2 score, while samples with lesser contents of N_{wa} exhibited a negative RDA2 score. S_R and E_4/E_6, indices for biological and humification parameters, positively related with RDA2. This suggests that the second axis might represent microbial production, which is also supported by the scores on the RDA2 of the weighted average of N_{wa}. Samples separated along the axis such as sediment-derived DOM contains more aliphatic compounds with nitrogen atoms. The higher contents of nitrogen might suggest enhanced microbial activities in sediment samples, which is consistent with the results deduced from optical indices in Section 3.1.

3.4. Implications for sources identification

Could these unique molecular characteristics of end-member DOM be used to identify potential sources of DOM in freshwater lakes? In this work, hierarchal cluster analysis (HCA) based on FT-ICR-MS data was carried out between freshwater samples and known source materials to identify potential sources of DOM in Lake Taihu and Lake Dianchi. Because thousands of formulas can be observed in a sample of DOM, there are 57,630 formulas in the data base for the 13 samples. Apparently, the majority of the formulas were repeatedly appeared in the 13 samples. As the results, there were 8860 exclusive/unique formulas in these sam-

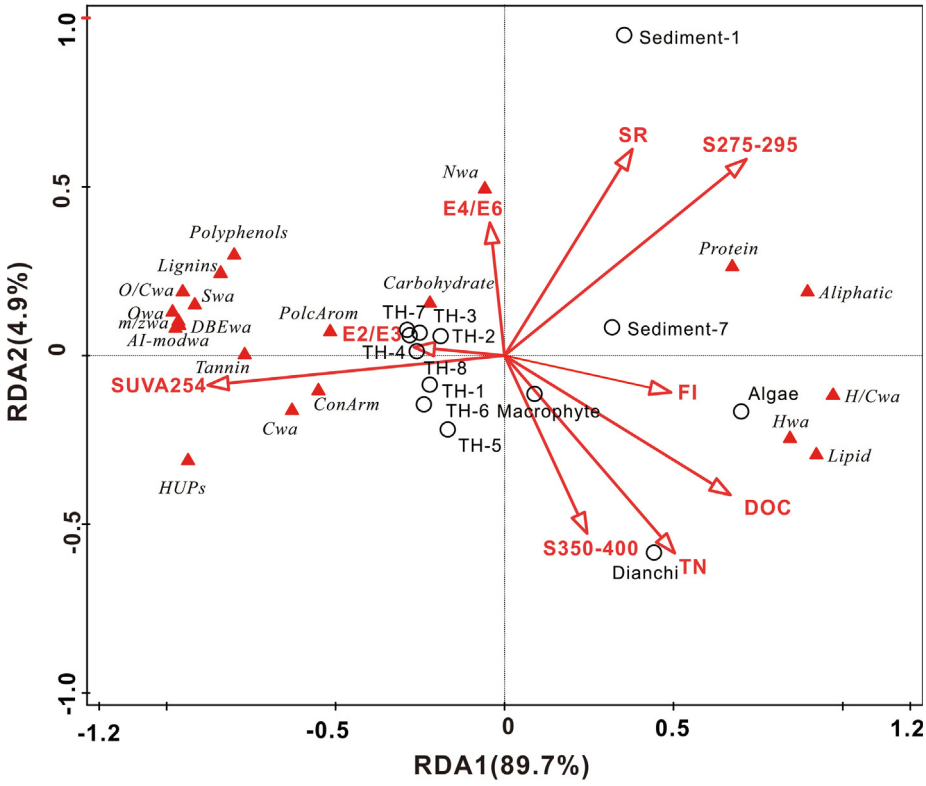


Fig. 4. Multivariate analysis of compounds and drivers using RDA. Ordinations are based on Bray-Curtis, which utilizes relative abundance of compounds. Environmental variables were fit to the ordination.

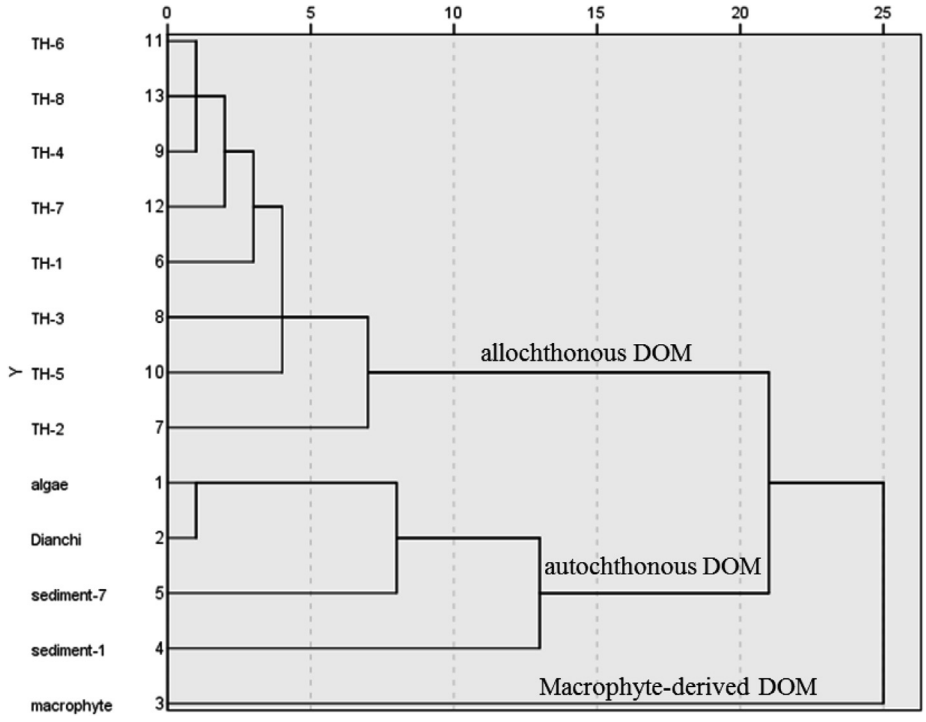


Fig. 5. Dendrogram from the cluster analysis using relative peak intensities of the total formulas in 13 DOM samples.

ples. A data matrix was then constructed from the relative magnitudes of the 8860 formulas in each of the 13 samples. Relative magnitudes of formulas absent in samples were set to 0. A cluster analysis was then applied to the data matrix (Fig. 5). Sampling locations TH-4 (Gonghu Bay), TH-6 (Center Lake) and TH-8 (Xukou Bay) exhibited the least degree of similarity with other locations. This result might be explained by their similar conditions since these three regions are less affected by terrestrial and anthro-

pogenic inputs, and water was clearer than that at other locations (Qin et al., 2018). Then locations TH-7 (Eastern Lake), TH-1 (Meiliang Bay), TH-3 (Southern Coast), TH-5 (Zhushan Bay), and TH-2 (Western Coast) were also grouped together, which reflected their generally poorer water quality and the degree of anthropogenic influence, such as inputs of wastewater and runoff (Liu et al., 2018a). Similarly, DOM derived from phytoplankton exhibited the greatest similarity with Lake

Dianchi water. This result further demonstrated the significant contribution of DOM produced by phytoplankton to Lake Dianchi. During the second clustering, DOM in extracts of sediments and phytoplankton were clustered together, which can be explained by the similar molecular characteristics of DOM derived from bacteria and that derived from phytoplankton, which could also include cyanobacteria. Particularly, DOM from macrophyte was classified as a separate class and not grouped with DOM in water from Lake Taihu. This result further clarified the speculation that terrigenous DOM might be the primary contributor of DOM in water of Lake Taihu, but not from macrophyte-derived DOM. Molecular characteristics of macrophyte-derived DOM might be intermediate between allochthonous DOM and autochthonous DOM, and play a limited role in the whole Lake Taihu since the low level of biomass input in recent years (Zhang et al., 2016a). In the study, statistical analysis using relative magnitudes effectively explored large, complex data sets to obtain more detailed molecular structure and source information. The molecular characteristics of diverse end-members should serve as a database for future study of source identification. In addition, the multivariate statistics based on FT-ICR-MS data provide a potential methodology to conduct source analysis in the natural freshwaters.

3.5. Geochemical and environmental implications

In the past decades, the nutrient enrichment induced by anthropogenic activity has caused excessive growth of emergent and submerged macrophyte in shallow freshwater lakes. Also, the occurrence of algae bloom is spreading frequently and globally in freshwater lakes due to climate change and nutrient enrichment (Paul, 2008; Wang and Jiang, 2016). Consequently, the considerable autochthonous dissolved organic matters in biomass of algae and macrophyte are released through cell death/lysis, and significantly contribute to the organic matter pool in lakes (Zhou et al., 2018). In this study, we revealed that the special chemical structures of autochthonous DOM were more aliphatic and less oxidized compared with allochthonous DOM. Therefore, it is to be expected that the algae bloom would lead to changes in chemical components of DOM in eutrophic lakes, which will further impact the biogeochemical processes of organic matters (such as degradation and sedimentation) (Liu et al., 2018c) and behaviors of contaminants (such as sorption, transportation and transformation) (Bai et al., 2017a,b; Bravo et al., 2017) in the environment. It has been reported phytoplankton-derived organic compounds could enhance mercury methylation rates in boreal lake sediments (Bravo et al., 2017). Thus autochthonous DOM was an important DOM pool and should gained more attention in eutrophic environment. Future studies are also necessary to investigate the detailed mechanisms of how autochthonous DOM in eutrophic lakes affects the large-scale ecosystem and also the behaviors of contaminants in molecular-scale in the environment.

4. Conclusions

Molecular-level characterization of DOM extracted from autochthonous end-members was conducted by combination of optical spectroscopy and FT-ICR-MS. Our results revealed that autochthonous DOM was more aliphatic and less oxidized compared with allochthonous DOM, which was marked by increases in lipid compounds and decreases in lignin components as compared to the allochthonous DOM-dominated freshwaters. Also, the optical indicators showed consistent results with that deduced from FT-ICR-MS.

The combination of multivariate analysis with FT-ICR-MS data is an invaluable method to further explore the potential information from FT-ICR-MS. This study has shown how to unravel the source identification of DOM in natural freshwaters by the combination of multivariate analysis with FT-ICR-MS. Further studies can now be designed to further investigate the different biogeochemical processes and environmental fates of diverse sources DOM in molecular level in lakes.

Declaration of Competing Interest

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

Acknowledgements

We are grateful to the financial support from the National Natural Science Foundation of China (Nos. 41807372, 41630645, and 41521003), and China Postdoctoral Science Foundation (2017M622280). We also thank Xiang Wan for help with sampling collection.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2019.134764.

References

Antony, R., Grannas, A.M., Willoughby, A.S., Sleighter, R.L., Thamban, M., Hatcher, P. G., 2014. Origin and sources of dissolved organic matter in snow on the east antarctic ice sheet. Environ. Sci. Technol. 48 (11), 6151–6159.
Arnold, W.A., Longnecker, K., Kroeger, K.D., Kujawinski, E.B., 2014. Molecular signature of organic nitrogen in septic-impacted groundwater. Environ. Sci. Proc. Impacts. 16 (10), 2400–2407.
Bai, L., Cao, C., Wang, C., Wang, C., Zhang, H., Jiang, H., 2017a. Roles of phytoplankton- and macrophyte-derived dissolved organic matter in sulfamethazine adsorption on goethite. Environ. Pollut. 230, 87–95.
Bai, L., Zhao, Z., Wang, C., Wang, C., Liu, X., Jiang, H., 2017b. Multi-spectroscopic investigation on the complexation of tetracycline with dissolved organic matter derived from algae and macrophyte. Chemosphere 187, 421–429.
Baker, A., 2001. Fluorescence excitation–emission matrix characterization of some sewage-impacted rivers. Environ. Sci. Technol. 35 (5), 948–953.
Bravo, A.G., Bouchet, S., Tolu, J., Björn, E., Mateos-Rivera, A., Bertilsson, S., 2017. Molecular composition of organic matter controls methylmercury formation in boreal lakes. Nat. Commun. 8, 14255.
Cabaniss, S.E., 2011. Forward modeling of metal complexation by NOM: II. Prediction of binding site properties. Environ. Sci. Technol. 45 (8), 3202–3209.
Cortés-Francisco, N., Harir, M., Lucio, M., Ribera, G., Martínez-Lladó, X., Rovira, M., Schmitt-Kopplin, P., Hertkorn, N., Caixach, J., 2014. High-field FT-ICR mass spectrometry and NMR spectroscopy to characterize DOM removal through a nanofiltration pilot plant. Water Res. 67, 154–165.
Cuss, C.W., Guéguen, C., 2015a. Labile Organic Matter—Chemical Compositions, Function, and Significance in Soil and the Environment. Soil Sci. Soc. Am. Inc., Madison, WI, pp. 237–274.
Cuss, C.W., Guéguen, C., 2015b. Relationships between molecular weight and fluorescence properties for size-fractionated dissolved organic matter from fresh and aged sources. Water Res. 68, 487–497.
Derrien, M., Yang, L., Hur, J., 2017. Lipid biomarkers and spectroscopic indices for identifying organic matter sources in aquatic environments: a review. Water Res. 112, 58–71.
Dittmar, T., Koch, B., Hertkorn, N., Kattner, G., 2008. A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. Limnol. Oceanogr. Methods 6 (6), 230–235.
Fellman, J.B., Hood, E., Spencer, R.G.M., 2010. Fluorescence spectroscopy opens new windows into dissolved organic matter dynamics in freshwater ecosystems: a review. Limnol. Oceanogr. 55 (6), 2452–2462.

Gonsior, M., Powers, L.C., Williams, E., Place, A., Chen, F., Ruf, A., Hertkorn, N., Schmitt-Kopplin, P., 2019. The chemodiversity of algal dissolved organic matter from lysed Microcystis aeruginosa cells and its ability to form disinfection by-products during chlorination. Water Res. 155, 300–309.
Gonsior, M., Zwartjes, M., Cooper, W.J., Song, W., Ishida, K.P., Tseng, L.Y., Jeung, M.K., Rosso, D., Hertkorn, N., Schmitt-Kopplin, P., 2011. Molecular characterization of effluent organic matter identified by ultrahigh resolution mass spectrometry. Water Res. 45 (9), 2943–2953.
He, Z., Wu, F. (Eds.), 2015. Labile Organic Matter – Chemical Compositions, Function, and Significance in Soil and the Environment, SSSA Special Publication. Soil Sci. Soc. Am, Madison, Wisconsin.
Helms, J.R., Stubbins, A., Ritchie, J.D., Minor, E.C., Kieber, D.J., Mopper, K., 2008. Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photobleaching of chromophoric dissolved organic matter. Limnol. Oceanogr. 53 (3), 955–969.
Hertkorn, N., Benner, R., Frommberger, M., Schmitt-Kopplin, P., Witt, M., Kaiser, K., Kettrup, A., Hedges, J.I., 2006. Characterization of a major refractory component of marine dissolved organic matter. Geochim. Cosmochim. Acta 70 (12), 2990–3010.
Herzsprung, P., von Tümpling, W., Hertkorn, N., Harir, M., Büttner, O., Bravidor, J., Friese, K., Schmitt-Kopplin, P., 2012. Variations of DOM quality in inflows of a drinking water reservoir: linking of van Krevelen diagrams with EEMF spectra by rank correlation. Environ. Sci. Technol. 46 (10), 5511–5518.
Hockaday, W.C., Purcell, J.M., Marshall, A.G., Baldock, J.A., Hatcher, P.G., 2009. Electrospray and photoionization mass spectrometry for the characterization of organic matter in natural waters: a qualitative assessment. Limnol. Oceanogr. Methods 7 (1), 81–95.
Hsu, P.-H., Hatcher, P.G., 2005. New evidence for covalent coupling of peptides to humic acids based on 2D NMR spectroscopy: a means for preservation. Geochim. Cosmochim. Acta 69 (18), 4521–4533.
Jaffé, R., Yamashita, Y., Maie, N., Cooper, W.T., Dittmar, T., Dodds, W.K., Jones, J.B., Myoshi, T., Ortiz-Zayas, J.R., Podgorski, D.C., Watanabe, A., 2012. Dissolved organic matter in headwater streams: compositional variability across climatic regions of North America. Geochim. Cosmochim. Acta 94, 95–108.
Kellerman, A.M., Dittmar, T., Kothawala, D.N., Tranvik, L.J., 2014. Chemodiversity of dissolved organic matter in lakes driven by climate and hydrology. Nat. Commun. 5, 3804.
Kellerman, A.M., Guillemette, F., Podgorski, D.C., Aiken, G.R., Butler, K.D., Spencer, R. G.M., 2018. Unifying concepts linking dissolved organic matter composition to persistence in aquatic ecosystems. Environ. Sci. Technol. 52 (5), 2538–2548.
Kellerman, A.M., Kothawala, D.N., Dittmar, T., Tranvik, L.J., 2015. Persistence of dissolved organic matter in lakes related to its molecular characteristics. Nat. Geosci. 8, 454.
Kim, S., Kramer, R.W., Hatcher, P.G., 2003. Graphical method for analysis of ultrahigh-resolution broadband mass spectra of natural organic matter, the Van Krevelen diagram. Anal. Chem. 75 (20), 5336–5344.
Koch, B.P., Dittmar, T., 2006. From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter. Rapid Commun. Mass Spectrosc. 20 (5), 926–932.
Koch, B.P., Dittmar, T., Witt, M., Kattner, G., 2007. Fundamentals of molecular formula assignment to ultrahigh resolution mass data of natural organic matter. Anal. Chem. 79 (4), 1758–1763.
Kujawinski, E.B., Longnecker, K., Blough, N.V., Vecchio, R.D., Finlay, L., Kitner, J.B., Giovannoni, S.J., 2009. Identification of possible source markers in marine dissolved organic matter using ultrahigh resolution mass spectrometry. Geochim. Cosmochim. Acta 73 (15), 4384–4399.
Li, Y., Harir, M., Lucio, M., Kanawati, B., Smirnov, K., Flerus, R., Koch, B.P., Schmitt-Kopplin, P., Hertkorn, N., 2016. Proposed guidelines for solid phase extraction of Suwannee river dissolved organic matter. Anal. Chem. 88 (13), 6680–6688.
Li, Y., Harir, M., Uhl, J., Kanawati, B., Lucio, M., Smirnov, K.S., Koch, B.P., Schmitt-Kopplin, P., Hertkorn, N., 2017. How representative are dissolved organic matter (DOM) extracts? A comprehensive study of sorbent selectivity for DOM isolation. Water Res. 116, 316–323.
Liu, S., Wu, F., Feng, W., Guo, W., Song, F., Wang, H., Wang, Y., He, Z., Giesy, J.P., Zhu, P., Tang, Z., 2018a. Using dual isotopes and a Bayesian isotope mixing model to evaluate sources of nitrate of Tai Lake, China. Environ. Sci. Pollut. Res. 25 (32), 32631–32639.
Liu, S., Zhao, T., Zhu, Y., Qu, X., He, Z., Giesy, J.P., Meng, W., 2018b. Molecular characterization of macrophyte-derived dissolved organic matters and their implications for lakes. Sci. Total Environ. 616, 602–613.
Liu, S., Zhu, Y., Liu, L., He, Z., Giesy, J.P., Bai, Y., Sun, F., Wu, F., 2018c. Cation-induced coagulation of aquatic plant-derived dissolved organic matter: Investigation by EEM-PARAFAC and FT-IR spectroscopy. Environ. Pollut. 234, 726–734.
Liu, S., Zhu, Y., Meng, W., He, Z., Feng, W., Zhang, C., Giesy, J.P., 2016a. Characteristics and degradation of carbon and phosphorus from aquatic macrophytes in lakes: insights from solid-state 13C NMR and solution 31P NMR spectroscopy. Sci. Total Environ. 543, 746–756.
Liu, S., Zhu, Y., Wu, F., Meng, W., He, Z., Giesy, J.P., 2016b. Characterization of plant-derived carbon and phosphorus in lakes by sequential fractionation and NMR spectroscopy. Sci. Total Environ. 566–567, 1398–1409.
Liu, S., Zhu, Y., Wu, F., Meng, W., Wang, H., He, Z., Guo, W., Song, F., Giesy, J.P., 2017. Using solid ^{13}C NMR coupled with solution ^{31}P NMR spectroscopy to investigate molecular species and lability of organic carbon and phosphorus from aquatic plants in Tai Lake, China. Environ. Sci. Pollut. Res. 24 (2), 1880–1889.
McIntyre, A.M., Guéguen, C., 2013. Binding interactions of algal-derived dissolved organic matter with metal ions. Chemosphere 90 (2), 620–626.
McKnight, D.M., Boyer, E.W., Westerhoff, P.K., Doran, P.T., Kulbe, T., Andersen, D.T., 2001. Spectrofluorometric characterization of dissolved organic matter for indication of precursor organic material and aromaticity. Limnol. Oceanogr. 46 (1), 38–48.
Ohno, T., He, Z., Sleighter, R.L., Honeycutt, C.W., Hatcher, P.G., 2010. Ultrahigh resolution mass spectrometry and indicator species analysis to identify marker components of soil- and plant biomass-derived organic matter fractions. Environ. Sci. Technol. 44 (22), 8594–8600.
Ohno, T., Sleighter, R.L., Hatcher, P.G., 2016. Comparative study of organic matter chemical characterization using negative and positive mode electrospray ionization ultrahigh-resolution mass spectrometry. Anal. Bioanal. Chem. 408 (10), 2497–2504.
Patriarca, C., Bergquist, J., Sjöberg, P.J.R., Tranvik, L., Hawkes, J.A., 2018. Online HPLC-ESI-HRMS method for the analysis and comparison of different dissolved organic matter samples. Environ. Sci. Technol. 52 (4), 2091–2099.
Paul, V.J., 2008. In: Hudnell, H.K. (Ed.), Cyanobacterial Harmful Algal Blooms: State of the Science and Research Needs. Springer New York, New York, NY, pp. 239–257.
Perminova, I.V., Dubinenkov, I.V., Kononikhin, A.S., Konstantinov, A.I., Zherebker, A. Y., Andzhushev, M.A., Lebedev, V.A., Bulygina, E., Holmes, R.M., Kostyukevich, Y. I., Popov, I.A., Nikolaev, E.N., 2014. Molecular mapping of sorbent selectivities with respect to isolation of arctic dissolved organic matter as measured by Fourier transform mass spectrometry. Environ. Sci. Technol. 48 (13), 7461–7468.
Qin, B., Yang, G., Ma, J., Wu, T., Li, W., Liu, L., Deng, J., Zhou, J., 2018. Spatiotemporal changes of cyanobacterial bloom in large shallow eutrophic Lake Taihu, China. Front. Microbiol. 9, 451.
Qin, B., Zhu, G., Gao, G., Zhang, Y., Li, W., Paerl, H.W., Carmichael, W.W., 2010. A drinking water crisis in Lake Taihu, China: linkage to climatic variability and lake management. Environ. Manage. 45 (1), 105–112.
Riedel, T., Biester, H., Dittmar, T., 2012. Molecular fractionation of dissolved organic matter with metal salts. Environ. Sci. Technol. 46 (8), 4419–4426.
Sleighter, R.L., Hatcher, P.G., 2007. The application of electrospray ionization coupled to ultrahigh resolution mass spectrometry for the molecular characterization of natural organic matter. J. Mass Spectrom. 42 (5), 559–574.
Song, F., Wu, F., Feng, W., Liu, S., He, J., Li, T., Zhang, J., Wu, A., Amarasiriwardena, D., Xing, B., Bai, Y., 2019. Depth-dependent variations of dissolved organic matter composition and humification in a plateau lake using fluorescence spectroscopy. Chemosphere 225, 507–516.
Stedmon, C.A., Thomas, D.N., Granskog, M., Kaartokallio, H., Papadimitriou, S., Kuosa, H., 2007. Characteristics of dissolved organic matter in Baltic Coastal Sea Ice: allochthonous or autochthonous origins?. Environ. Sci. Technol. 41 (21), 7273–7279.
Stenson, A.C., Marshall, A.G., Cooper, W.T., 2003. Exact masses and chemical formulas of individual Suwannee River Fulvic acids from ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectra. Anal. Chem. 75 (6), 1275–1284.
Stone, R., 2011. China aims to turn tide against toxic lake pollution. Science 333 (6047), 1210.
Wagner, S., Riedel, T., Niggemann, J., Vähätalo, A.V., Dittmar, T., Jaffé, R., 2015. Linking the molecular signature of heteroatomic dissolved organic matter to watershed characteristics in world rivers. Environ. Sci. Technol. 49 (23), 13798–13806.
Wang, C., Jiang, H.-L., 2016. Chemicals used for in situ immobilization to reduce the internal phosphorus loading from lake sediments for eutrophication control. Crit. Rev. Environ. Sci. Technol. 46 (10), 947–997.
Yakimenko, O., Khundzhua, D., Izosimov, A., Yuzhakov, V., Patsaeva, S., 2018. Source indicator of commercial humic products: UV-Vis and fluorescence proxies. J. Soil Sediment 18 (4), 1279–1291.
Yan, M., Korshin, G.V., 2014. Comparative examination of effects of binding of different metals on chromophores of dissolved organic matter. Environ. Sci. Technol. 48 (6), 3177–3185.
Yang, C., Liu, Y., Zhu, Y., Zhang, Y., 2016. Insights into the binding interactions of autochthonous dissolved organic matter released from Microcystis aeruginosa with pyrene using spectroscopy. Mar. Pollut. Bull. 104 (1), 113–120.
Zhang, M. and He, Z., 2015. Labile Organic Matter—Chemical Compositions, Function, and Significance in Soil and the Environment. Soil Sci. Soc. Am. Inc., Madison, WI, pp. 1–22.
Zhang, Y., Liu, X., Qin, B., Shi, K., Deng, J., Zhou, Y., 2016a. Aquatic vegetation in response to increased eutrophication and degraded light climate in Eastern Lake Taihu: implications for lake ecological restoration. Sci. Rep. 6, 23867.
Zhang, Y., Liu, X., Wang, M., Qin, B., 2013. Compositional differences of chromophoric dissolved organic matter derived from phytoplankton and macrophytes. Org. Geochem. 55 (1), 26–37.
Zhang, Y., Shi, K., Liu, J., Deng, J., Qin, B., Zhu, G., Zhou, Y., 2016b. Meteorological and hydrological conditions driving the formation and disappearance of black blooms, an ecological disaster phenomena of eutrophication and algal blooms. Sci. Total Environ. 569–570, 1517–1529.
Zhou, Y., Davidson, T.A., Yao, X., Zhang, Y., Jeppesen, E., de Souza, J.G., Wu, H., Shi, K., Qin, B., 2018. How autochthonous dissolved organic matter responds to eutrophication and climate warming: evidence from a cross-continental data analysis and experiments. Earth-Sci. Rev. 185, 928–937.
Zhu, G., Wang, F., Gao, G., Zhang, Y., 2008. Variability of phosphorus concentration in large, shallow and eutrophic Lake Taihu, China. Water Environ. Res. 80 (9), 832–839.

Supporting Information for

Linking the molecular composition of autochthonous dissolved organic matter to source identification for freshwater lake ecosystems by combination of optical spectroscopy and FT-ICR-MS analysis

Shasha Liu [a], Zhongqi He [b], Zhi Tang [a,*], Leizhen Liu [c], Junwen Hou [a], Tingting Li [a], Yahe Zhang [d], Quan Shi [d], John P. Giesy [e], Fengchang Wu [a]

[a] State Key Laboratory of Environment Criteria and Risk Assessment, Chinese Research Academy of Environmental Sciences, Beijing 100012, China

[b] USDA-ARS Southern Regional Research Center, 1100 Robert E Lee Blvd, New Orleans, LA 70124, USA

[c] Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China

[d] State Key Laboratory of Heavy Oil Processing, China University of Petroleum, 18 Fuxue Road, Changping, Beijing 102249, China

[e] Department of Biomedical and Veterinary Biosciences and Toxicology Centre, University of Saskatchewan, Saskatoon, Saskatchewan SK S7N 5B3, Canada

*Corresponding author: Phone/fax: +86 10 84915312; Email: tzwork@hotmail.com.

Contents:

Number of pages:10

Number of tables: 2

Number of figures: 4

Study area

Lake Taihu (30°55′40″–31°32′58″ N and 119°52′32″–120°36′10″ E)，located in the southern Changjiang (Yangtze) River Delta, is one of the most densely populated regions in China. Lake Taihu has an area of 2338 km^2, with a maximum length and width of 68.5 and 56 km, respectively. The average depth is about 1.9 m. The average annual air temperature in Lake Taihu is 16.0–18.0℃. The average annual precipitation is 1100–1150 mm. In the current study, there were a total of 8 sampling sites distributed in 8 regions of Lake Taihu (Fig. S1, Table S1). There are TH-1 in Meiliang Bay, TH-2 in western coast, TH-3 in southern coast, TH-4 in Gonghu Bay, TH-5 in Zhushan Bay, TH-6 in Center Bay, TH-7 in Eastern Lake and TH-8 in Xukou Bay. The water quality and phytoplankton biomass of Lake Taihu exhibit spatial heterogeneity. According to the monitoring station data in Lake Taihu, the northern bays (Gonghu Bay, Meiliang Bay, and Zhushan Bay) and western bays are under poor water quality and frequent cyanobacterial bloom, whereas the eastern bays and Xukou Bay is dominated by macrophyte with well water quality.

Lake Dianchi Basin (24°28 to 25°28 N, 102°30 to 103°00 E) is located in the middle of Yunnan Guizhou Plateau in Southwest China. With an area of 308.6 km^2 and an average depth of only 4.4 m, Lake Dianchi is the sixth largest body of fresh water in China. The annual mean temperature is 14.7℃, with an average precipitation of 797–1,007 mm. The lake is divided in two by an artificial dam in the northeast. The smaller half (10.7 km^2) is called Caohai. The larger half (297.9 km^2), is called Waihai. Eutrophication has been an intractable problem since the 1980s in Lake Dianchi Basin.

In this study, in order to compare the source identification with Lake Taihu, one water sampling site in Haigeng Park in eutrophic region of Dianchi was also collected for molecular characterization and source analysis (Table S1).

Table S1. The location of sampling sites collected in the study.

Sample	Label	longitude	latitude	District
Taihu water	TH-1	120°11′39.59″	31°28′34.79″	Meiliang Bay
	TH-2	119°56′42.00″	31°18′52.20″	Western Coast
	TH-3	120°07′07.18″	30°57′49.21″	Southern Coast
	TH-4	120°22′36.88″	31°26′06.40″	Gonghu Bay
	TH-5	120°01′41.41″	31°27′00.04″	Zhushan Bay
	TH-6	120°08′37.90″	31°06′59.44″	Center Lake
	TH-7	120°30′47.84″	31°05′21.88″	Eastern Lake
	TH-8	120°24′21.46″	31°10′36.59″	Xukou Bay
Dianchi water	Dianchi	102°39′30.84″	24°57′42.06″	Haigeng Park
End-member	Macrophyte	120°24′21.46″	31°10′36.59″	Xukou Bay
	Algae	120°11′39.59″	31°28′34.79″	Meiliang Bay
	Sediment-1	120°11′39.59″	31°28′34.79″	Meiliang Bay
	Sediment-7	120°30′47.84″	31°05′21.88″	Eastern Lake

Table S2. Physicochemical characterization and optical properties.

Sample	TN ($mg \cdot L^{-1}$)	TP ($mg \cdot L^{-1}$)	Chl-a ($ug \cdot L^{-1}$)	DOC ($mg \cdot L^{-1}$)	$SUVA_{254}$ $L \cdot (mg \cdot m)^{-1}$	E_2/E_3	$S_{275-295}$	$S_{350-400}$	S_R	FI
TH-1	1.72	0.08	55.35	3.08	2.25	5.92	0.02	0.004	4.902	1.69
TH-2	2.03	0.09	43.75	3.34	1.86	7.92	0.022	0.016	1.375	1.78
TH-3	9.32	0.4	367.11	2.37	1.87	6.7	0.021	0.011	1.841	1.69
TH-4	1.11	0.03	5.68	2.78	2.29	5.64	0.02	0.009	2.202	1.67
TH-5	4.57	0.26	174.24	2.59	1.9	5.07	0.018	0.01	1.74	1.77
TH-6	2.56	0.14	79.16	3.14	1.55	6.83	0.023	0.009	2.404	1.7
TH-7	1.69	0.06	13.39	2.54	2.33	8.36	0.022	0.011	2.004	1.67
TH-8	1.33	0.03	7.09	2.57	1.98	10.56	0.024	0.013	1.782	1.65
average	3.04	0.14	93.22	2.80	2.00	7.13	0.02	0.01	2.28	1.70
Lake Dianchi	13.59	1.04	-	51.61	0.83	7.67	0.023	0.023	1.017	2.09
Macrophyte	7.58	-	-	66.04	0.06	3.77	0.016	0.014	1.104	1.03
Algae	11.87	-	-	53.97	0.09	6.3	0.043	0.013	3.3	2.14
Sediment-1	0.32	-	-	6.84	0.43	6.71	0.048	0.007	6.814	1.81
Sediment-7	0	-	-	4.55	0.48	6.51	0.032	0.005	7.156	1.90

$SUVA_{254}$ is the ratio of absorbance (m^{-1}) at 254 nm to DOC ($mg \cdot L^{-1}$);

E_2/E_3 is an absorption ratio at 250 to 365 nm, and E_4/E_6 is the absorption ratio of 465/665;

$S_{275-295}$ is the absorption slope from 275nm to 295nm, and $S_{350-400}$ is the absorption slope from 350 nm to 400 nm;

S_R is the ratio of the spectral slopes S of the shorter (275 - 295 nm) to the longer (350 - 400 nm) wavelength ranges;

FI was calculated as the ratio of emission intensity at 450 nm to that of 500 nm for a fixed excitation wavelength of 370 nm.

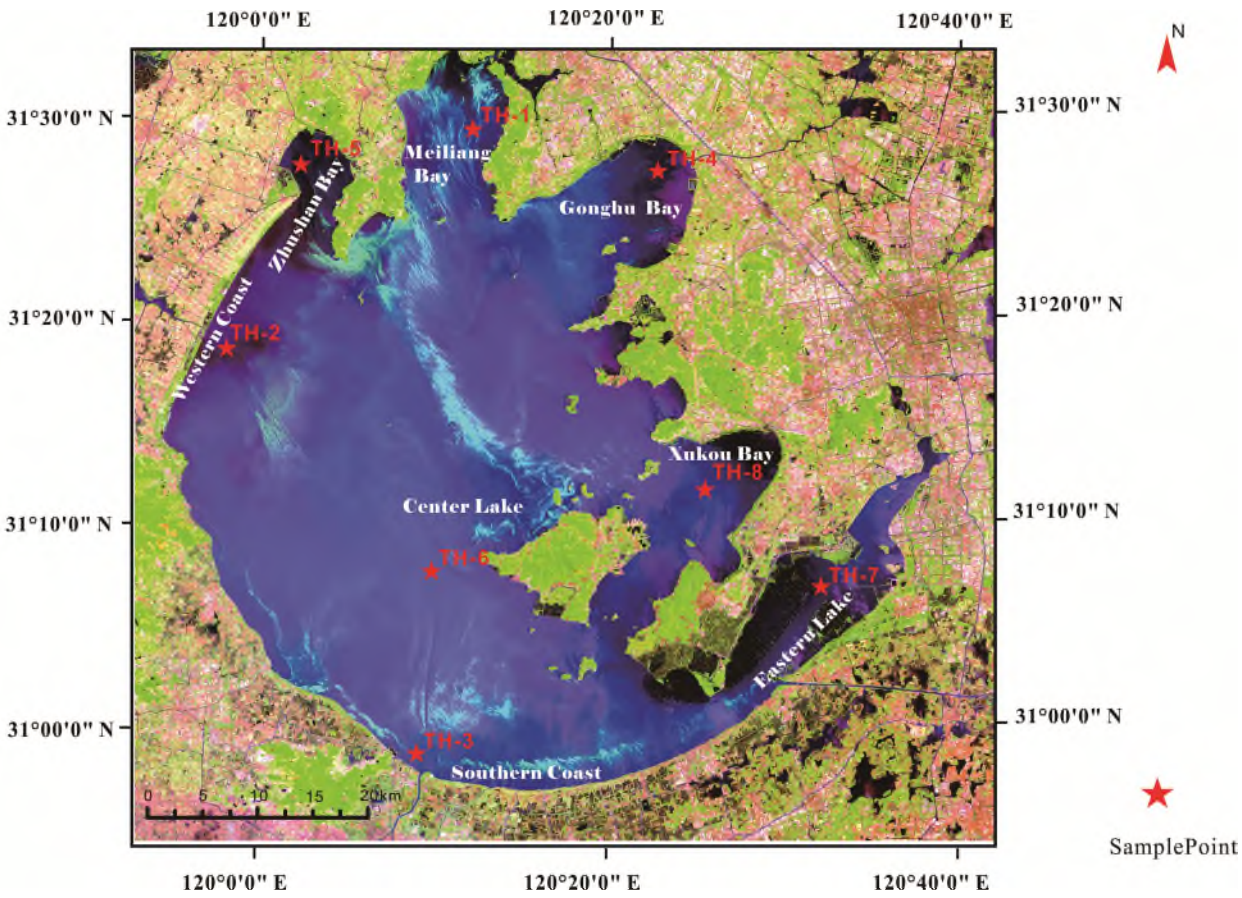


Fig. S1. Sampling sites in various regions of Lake Taihu. Samples of water were collected from all regions and named identified as: TH-1, TH-2, TH-3, TH-4, TH-5, TH-6, TH-7, respectively. Macrophyte was collected from Xukou Bay, algae were collected from Meiliang Bay, and sediments were collected from Meiliang Bay (Sediment-1) and Eastern Lake (Sediment-7).

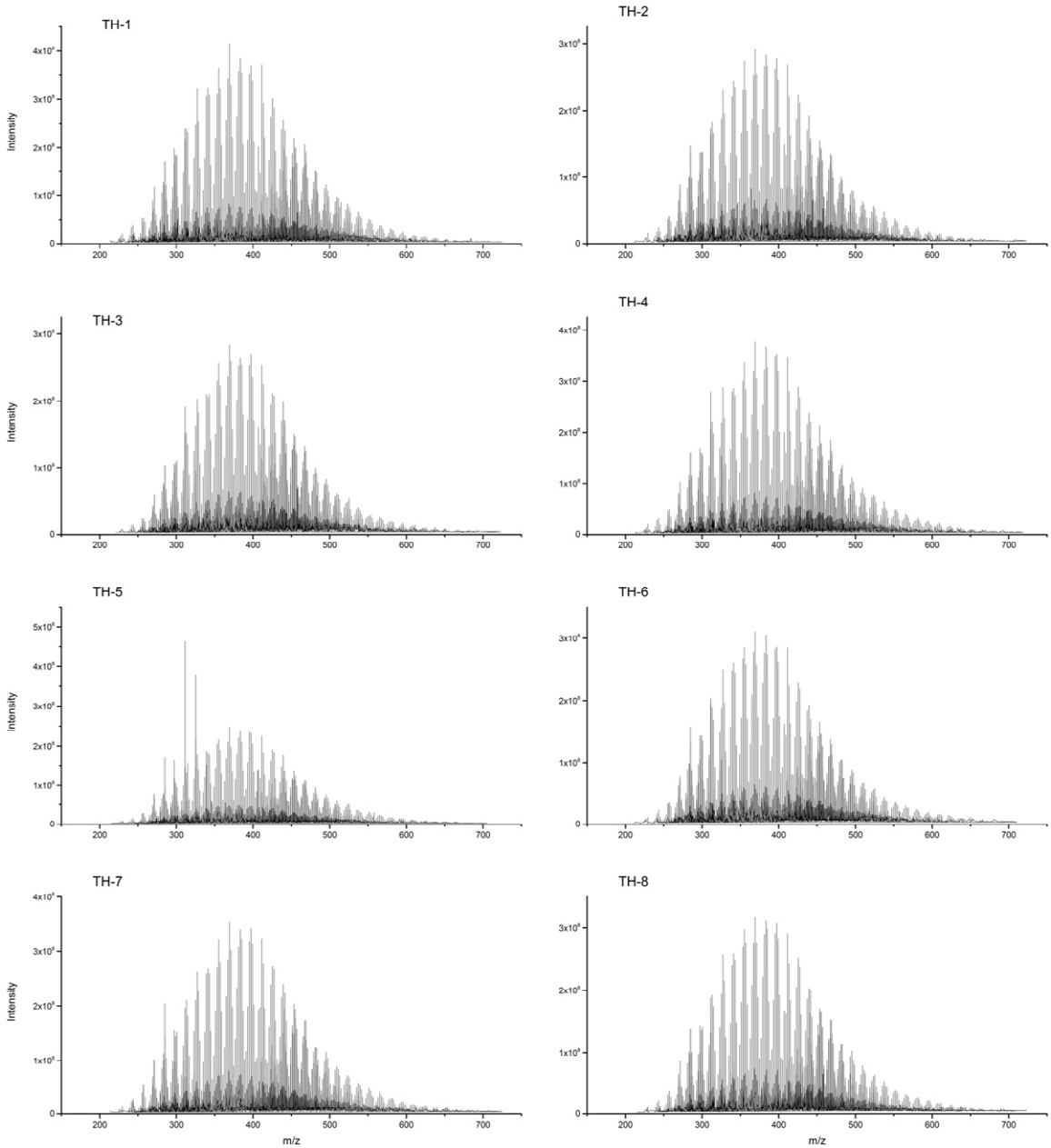


Fig. S2. Negative ion mass spectra of SPE-DOM from water samples in different regions of Lake Taihu isolated by PPL.

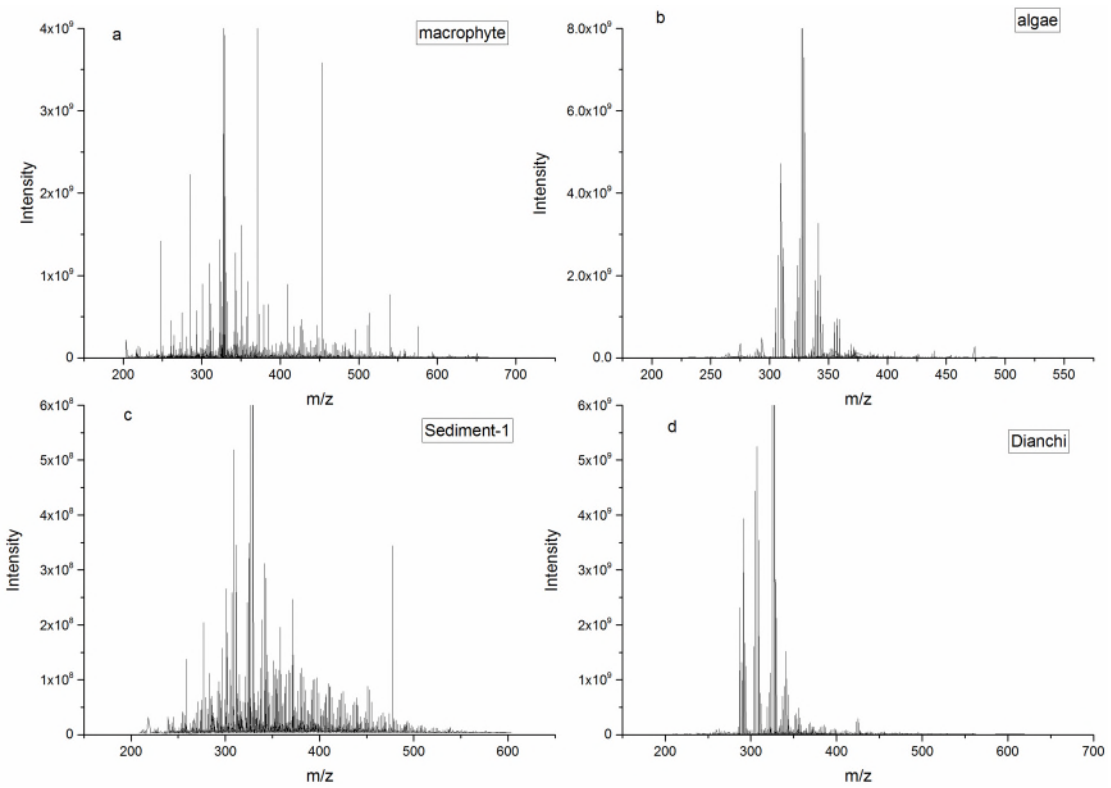


Fig. S3. Negative ion mass spectra of SPE-DOM from different source samples. The different sources include a) macrophyte-derived DOM; b) algae-derived DOM; c) sediment-derived DOM; and d) SPE extracted DOM from Dianchi.

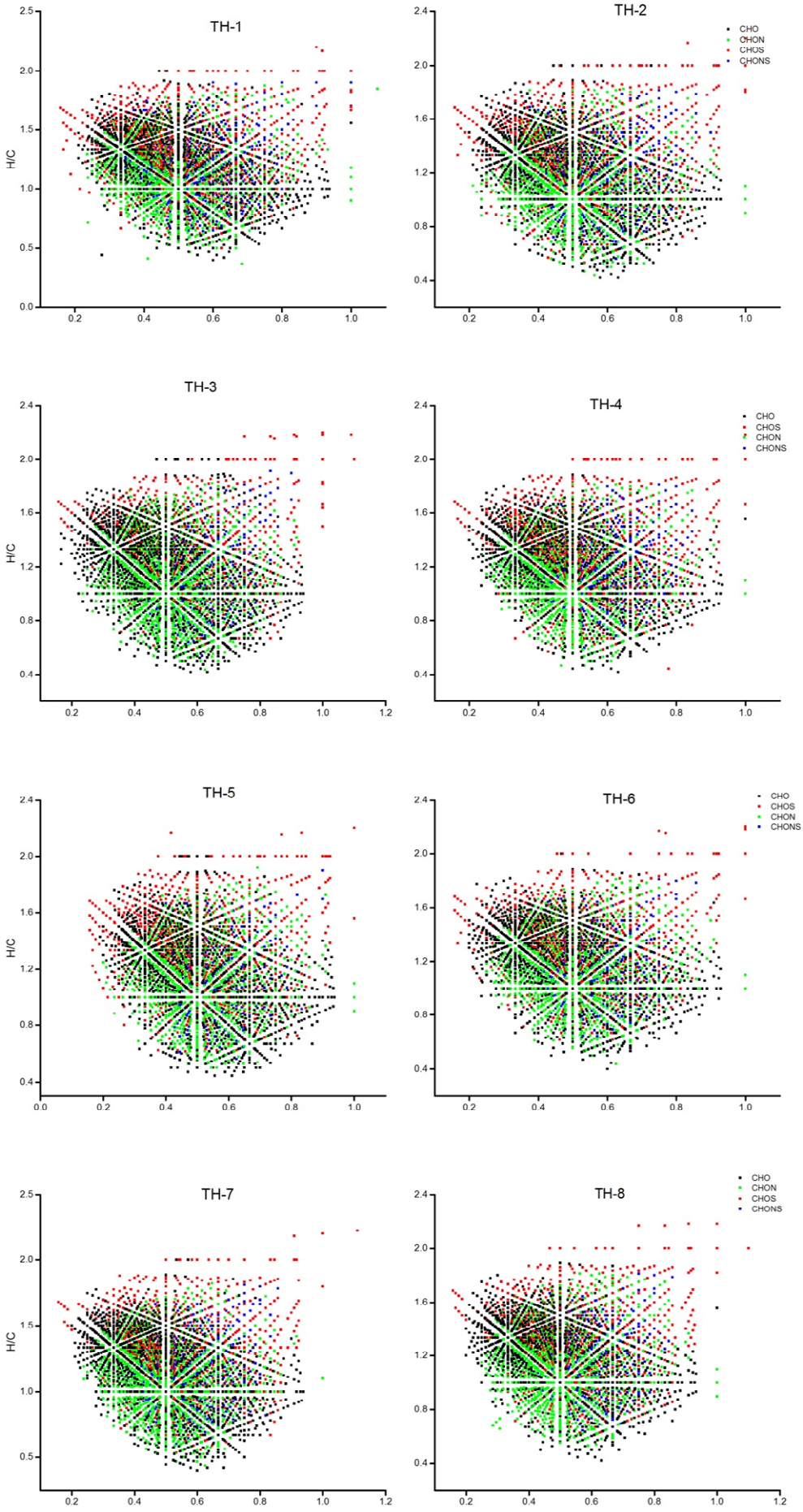

TH-1
TH-2
TH-3
TH-4
TH-5
TH-6
TH-7
TH-8
H/C
CHO
CHON
CHOS
CHONS

Fig. S4. Van Krevelen diagrams from the mass spectra of SPE-DOM of water from different region in Lake Taihu. Ovals overlain the plots indicate major compound classes: lipids (H:C=1.5-2.0; O:C=0-0.3); protein (H:C=1.5-2.2; O:C=0.3-0.67; N/C $\geqslant$ 0.05); lignins (H:C=0.7-1.5; O:C=0.1-0.67); carbohydrates (H:C=1.5-2.0; O:C=0.67-1.2); unsaturated hydrocarbons (H:C=0.7-1.0; O:C=0-0.1); condensed aromatic structures (H:C=0.2-0.7; O:C=0-0.67); and tannin (H:C=0.5-1.5; O:C=0.67-1.2).