The bio- and thermal lability of dissolved organic matter as revealed by high-resolution mass spectrometry and thermal chemical analyses

Kaijun Lu[1]*, Jianhong Xue[1], Laodong Guo[2], Zhanfei Liu[1]

[1] Marine Science Institute, The University of Texas at Austin, Port Aransas, Texas 78373
[2] School of Freshwater Sciences, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin 53204

*Corresponding author, kaijun.lu@utexas.edu

Abstract

As one of the largest reduced carbon pools on earth, it is critical to understand the chemical structures and lability of dissolved organic matter (DOM) in aquatic environments. To achieve a comprehensive view of its composition and reactivity, DOM were concurrently extracted via both ultrafiltration and solid-phase-extraction (SPE), and further compared through multi-dimensional analyses, including elemental and total hydrolyzable amino acid composition, high-resolution liquid chromatography mass spectrometry, and thermochemical analysis. With a very low percentage (< 5%) of shared molecules, ultrafiltered-DOM (UDOM) and SPE-extracted DOM (SPEDOM) showed distinct differences in their bio- and thermal labilities. SPEDOM not only contained more "degraded" biomarkers, such as glycine, β-alanine, and γ-aminobutyric acid, based on amino acid analysis, but was also dominated by a higher abundance of more bio-recalcitrant lignin-like structures. Furthermore, thermogravimetric analysis revealed higher activation energy for SPEDOM (138.8 ± 0.5 $kJ \cdot mol^{-1}$) than for UDOM (125.4 ± 2.8 $kJ \cdot mol^{-1}$), suggesting a more thermal recalcitrance nature of SPEDOM. Consistently, thermal slicing ramped pyrolysis showed a lower percentage of recalcitrant pyrolyzates (e.g., aromatic and phenol structures), but a higher contribution of labile compound classes (e.g., aliphatic structures) in UDOM, as well as a lower diversity of pyrolyzates from SPEDOM than UDOM. Taken together, our results demonstrated a highly heterogeneous nature in DOM composition and distinctly different bio- and thermal labilities between UDOM and SPEDOM, urging that characterization of different fractions of DOM is needed for a comprehensive evaluation of natural DOM.

1. Introduction

As one of the most active C pools on earth, dissolved organic matter (DOM), residing in a total amount of nearly 700×10^{15} g of C (e.g., Eglinton and Repeta, 2014), is constantly exchanging with atmospheric C through processes like photosynthesis and respiration. Given its important role in controlling global carbon cycling and climate, a thorough understanding of the chemical composition of DOM through molecular level analysis is necessary. Indeed, during the past several decades there has been an exponential increase in molecular level analyses of DOM. Along with this growing interest, the heterogeneity or diversity of DOM in terms of age, composition, and lability, as implied by many studies (e.g., Simjouw et al., 2005; Benner and Amon, 2015; Broek et al., 2017; Lin and Guo, 2020), has also received increasing attention.

Nevertheless, differences in DOM heterogeneity between different DOM fractions simultaneously extracted by different methods remain under studied. Also, little is known about potential DOM fractionation during extraction and resultant differences in DOM composition and reactivities or lability. With the great advancement in analytical approaches, acquiring purified DOM from natural waters, prior to any molecular analysis represents a pivotal initial step. Commonly applied extraction approaches include ultrafiltration and solid phase extraction (SPE). Ultrafiltration exploits the larger molecular size of DOM than those of most dissolved inorganic species, while SPE relies on the physicochemical sorption between acidified DOM and the stationary phase to extract DOM from a water sample. Both extraction approaches suffer from some degrees of selectivity, and neither ultrafiltered DOM (hereinafter as UDOM) nor SPE-isolated DOM (hereinafter as SPEDOM) can represent the total DOM (e.g., Broek et al., 2017). Previous studies mostly focused on a particular sub-fraction of DOM, either UDOM or SPEDOM, leading to a possible omission of the different compositional and functional components of DOM and therefore a potential underestimation of the diverse natural DOM. To the best of our knowledge, only a few studies have investigated different fractions of DOM, and demonstrated the preferences or biases of different sorbents and extraction methods albeit in an incomprehensive way (e.g., Simjouw et al., 2005; Chen et al., 2016; Broek et al., 2017, 2019).

In order to achieve a better understanding of the diversity of DOM, we systematically investigated the bulk and molecular properties of both UDOM and SPEDOM from multiple analytical angles, including elemental composition, total hydrolyzable amino acids (THAA), high-resolution mass spectrometry (MS), as well as thermochemical analyses (e.g., Hemingway et al., 2017, 2019; O'Connor et al., 2020; Rogers et al., 2021). Of all the characterizable chemical indicators, the compositional pattern of total hydrolyzable amino acids (THAA) has been widely employed to quantify the degradation status, or biolability, of natural organic matter (Dauwe and Middelburg, 1998; Dauwe et al., 1999; Amon et al., 2001; Li et al., 2018; Liu and Xue, 2020; O'Connor et al., 2020). Therefore, composition of THAA represents an effective tool of quantifying the biolability of DOM without involving bioassay experiments. In addition to specific compound analysis, high resolution mass spectrometry,

especially when in combination with chromatography separation, provides unparalleled insights into the formulas and potentially structures of DOM, and thus has been widely applied in DOM-related studies (e.g., Patriarca et al., 2018; Lu et al., 2018, 2021; Kim et al., 2019). The acquired DOM mass spectra, in the form of a large matrix of thousands of molecular formulas, and their corresponding chromatography retention time as well as signal intensities, greatly facilitate subsequent comparisons, including biolability, between different DOM samples. Finally, thermochemical analysis, such as pyrolysis coupled with gas chromatography, has long been applied to characterize environmental organic matter (e.g., Kimber and Searle, 1970; Boon and Haverkamp, 1979; Zsolnay, 1982; Saliot et al., 1984). With the recent introduction of programmed ramped pyrolysis oxidation (Rosenheim et al., 2008), and the combination with mass spectrometry (Seeley et al., 2018; O'Connor et al., 2020; Wang et al., 2020), this technique has been used to quantify the radiocarbon age and thermal stability, in terms of bonding strength, of natural organic matter (e.g., Hemingway et al., 2017, 2019; Rogers et al., 2021), and to decipher molecular composition of DOM based on compound class analysis (O'Connor et al., 2020).

Our goal was to acquire a broader picture of the DOM pool, with a focus on the compositional differences between two different DOM pools extracted simultaneously by ultrafiltration and SPE column. Through the comparison between UDOM and SPEDOM, we demonstrated that DOM contains fractions with distinct differences in thermal and biological lability based on molecular level specific compound and thermal chemical analyses.

2. Materials and Methods

2.1. Sample collection

Water samples were collected from three south Texas rivers, including the Guadalupe River (GR; N 28.48°, W 96.86°), Mission River (MR; N 28.29°, W 97.28°), and Nueces River (NR; N 27.88°, W 97.63°), and one coastal site from a ship channel (SC; N 27.84°, W 97.05°, Port Aransas, Texas) that is connected to the open Gulf of Mexico (Figure 1). The three river water samples all had a salinity of < 1‰, representing freshwater DOM, whereas the coastal water sample had a salinity of ca. 30 ‰, representing coastal marine DOM (Table S1). Water samples were collected with the pre-acid cleaned 10-L high-density polyethylene (HDPE) bottle or 20-L HDPE carboy and stored in a cooler with ice. Upon returning to lab within the same day, the samples were immediately filtered through in-series filter cartridges (Whatman™ Polycap HD 75 Capsule Filter, 5.0 μm, 2712 M; Whatman™ Polycap AS 75 Capsule filter, 0.2 μm, 2706 T) to remove suspended particles.

The prefiltered water (< 0.2 μm) was further processed to extract UDOM and SPEDOM using ultrafiltration and solid phase extraction, respectively. Ultrafiltration of the < 0.2 μm filtrate (without any treatment) was carried out with a spiral-wound Amicon S10N1 ultrafiltration cartridge (regenerated cellulose, Millipore), with a molecule-weight cutoff of 1 kDa. After ultrafiltration, the retentate (reduced from 20 L to 0.5 L) was further diafiltered with 20 L of ultrapure water to remove residual low molecular weight DOM and salts (Guo and Santschi, 1996). The purified UDOM was then freeze-dried to obtain powder DOM for further chemical and molecular characterization (Lin et al., 2021). Depending on the analysis, UDOM was either used in powder form, or re-dissolved with LC/MS grade water (Fisher).

Solid phase extraction followed the published procedure (Dittmar et al., 2008). Briefly, filtered water samples were first acidified with HCl to a pH of 2, and DOM was subsequently extracted using modified styrene divinyl benzene polymer (PPL) cartridges (Agilent Bond Elut PPL cartridge, 500 mg, P/N 12105006; a maximum of 2 L water was passed through one cartridge to avoid overloading). DOM extracts were eluted with LC/MS grade methanol (four cartridge volumes for each cartridge), re-concentrated to a final volume of 10 mL, and stored at -20 °C until further analysis. Depending on the analysis, the SPEDOM was used in either solution or powder form after being freeze-dried. The DOC concentrations at the four sampling sites varied from ca. 300 μmol $C \cdot L^{-1}$ in SC to ca. 1000 μmol $C \cdot L^{-1}$ in MR. The recovery rates varied from 18% to 49% for ultrafiltration (based on powder DOM after freeze drying), and from 46% to 52% for PPL-based solid phase extraction, respectively, with the lowest recovery rate found in the coastal saline water (SC) in both extraction approaches (Table S1). These recovery rates are consistent with previous studies using either ultrafiltration or SPE (e.g., Broek et al., 2017; Lu and Liu, 2019).

2.2. Organic carbon and nitrogen, and stable carbon isotope

The organic carbon, nitrogen, and $\delta^{13}C$ of UDOM and SPEDOM samples were measured using a Thermo FLASH 2000 CHN Elemental Analyzer coupled with a Thermo Delta V Plus isotope ratio mass spectrometer. The $\delta^{13}C$ values were expressed relative to

Vienna Pee Dee Belemnite standard. Precision for the C or N content was within 5% and for $\delta^{13}C$ within 0.2‰.

2.3. Total hydrolyzable amino acids analysis

Total hydrolyzable amino acids (THAAs) analysis followed previous works (Liu and Xue, 2020; O'Connor et al., 2020). For UDOM samples, ca. 0.5 mg was dissolved in 4 mL LC/MS grade H_2O, of which 2 mL subsamples were taken, mixed with 2 mL 12N HCl, sealed under N_2, and hydrolyzed at 110 °C for 20 h. For SPEDOM samples, 100 μL of SPEDOM solution was dried with N_2, re-dissolved in 2 mL H_2O, mixed with 2 mL 12N HCl, sealed under N_2, and hydrolyzed at 110 °C for 20 h. The solution after acid hydrolysis was then dried under N_2 in a vial and replaced with H_2O. THAAs were analyzed using high performance liquid chromatography (HPLC) after being derivatized by *o*-phthaldialdehyde (Lee et al., 2000; Kuznetsova and Lee, 2002; Liu and Xue, 2020).

Amino acid derivatives were eluted on a column (Phenomenex Luna C_{18}, 25 cm × 5 μm) with a flow rate of 0.8 $mL \cdot min^{-1}$ in a HPLC system (Shimadzu). A binary gradient of 0.05 M sodium acetate (pH 5.7) with 5% tetrahydrofuran (mobile phase A) and methanol (mobile phase B) was used, ramping from 20% B to 50% B in 40 min, then to 100% B in 20 min. The derivatized 16 amino acids, including aspartic acid (ASP), glutamic acid (GLU), histidine (HIS), serine (SER), arginine (ARG), glycine (GLY), threonine (THR), β-alanine (BALA), alanine (ALA), tyrosine (TYR), γ-aminobutyric acid (GABA), methionine (MET), valine (VAL), phenylalanine (PHE), isoleucine (ILE), and leucine (LEU), were detected by fluorescence and quantified through comparison with an authentic standard mixture (Sigma). Duplicate analyses of amino acids from the same extract generally agreed within 10%.

2.4. Molecular characterization of DOM through LC/MS

Molecular level information of DOM was acquired using an Ion Mobility Quadrupole Time of Flight Liquid Chromatography Mass Spectrometer (IM Q-TOF LC/MS, Agilent 6560) with an orthogonal electrospray ionization (ESI) source (Lu et al., 2018, 2021; Lu and Liu, 2019). Both ESI- and ESI+ modes were applied. Briefly, mobile phase A was H_2O with 0.1% (v/v) formic acid, and B was acetonitrile for ESI+ mode. Ten-μL sample was eluted through a StableBond C_{18} column (Poroshell 120 SB-C18; 2.7 μm, 2.1 × 100 mm; Agilent P/N 685775-902) at a flow rate of 0.5 $mL \cdot min^{-1}$. During the 21-min run, mobile phase B was increased from 3% to 90% in the first 15 min, held at 90% from 15 min to 20 min, and then dropped to 3% at 21 min. A post-run time of 4 min allowed the column to reach equilibrium before the next injection. For ESI- mode, mobile phase A was H_2O with 10 $mmol \cdot L^{-1}$ ammonium acetate, and B was acetonitrile. Ten-μL sample was eluted through a HILIC column (2.7 μm, 15 cm × 4. 6 mm SUPELCO, P/N 53981-U) at a flow rate of 0.5 $mL \cdot min^{-1}$. During the 10-min run, mobile phase B was held at 98% during the first 1 min, and dropped to 95% from 1 min to 10 min. A post-run period of 15 min allowed the column to reach equilibrium before the next injection. A total amount of ca. 0.1 – 0.2 μg C for both UDOM and SPEDOM samples was injected during the analysis, and each sample was analyzed in duplication.

The mass spectrum data were acquired with software MassHunter LC/MS Data Acquisition (Version B.07.00 and B.09.00) in both ESI- and ESI+ modes. For ESI+, the orthogonal electrospray ionization source (Dual Agilent Jet Stream ESI) was operated with N_2 sheath gas temperature of 350 °C at a flow rate of 12 $L \cdot min^{-1}$. N_2 drying gas applied at the source entrance was at temperature of 225 °C and maintained at a flow rate of 13 $L \cdot min^{-1}$ with a nebulizer pressure of 45 psig. The source operated in positive mode was set to a VCap voltage of 3500 V, and a nozzle voltage of 0 V. Q-TOF was in positive ion polarity in MS mode, with an MS mass range of 70–1200 m/z, and an acquisition rate of 1 spectrum/s. Reference masses of 121.050873 and 922.009798 (Agilent Tuning Mix) were used for mass calibration. For ESI-, the ion source was also the Dual AJS ESI. Sheath gas parameters were the same as in ESI+ mode. Drying gas temperature was also 225 °C, but the flow rate was adjusted to 5 $L \cdot min^{-1}$ and nebulizer pressure was maintained at 20 psig. The source operated in negative mode was maintained at 3500 V VCap voltage, but the nozzle voltage was raised to 2000 V. Q-TOF settings of ESI- MS resembled those of ESI+ mode, except that Q-TOF was operated under negative ion polarity, with reference masses of 112.985587 and 1033.988109 (Agilent Tuning Mix).

Data analysis employed MassHunter Qualitative Analysis (Version B.07.00, Service Pack 2; Lu et al., 2018, 2021; Lu and Liu, 2019). The built-in function "Find by Molecular Feature" in MassHunter Qualitative Analysis identified possible compounds detected. The mass inaccuracy tolerance was set at ≤ 1.5 ppm (< 0.9 mDa at a mass range of 150−600 Da). Possible CHNO formulas were computed with the function "Generate Formulas", which takes the exact mass, ^{13}C isotope abundance and spacing into account. Generated formulas were further screened with self-developed R codes following the basic criteria: (1) Double Bond Equivalent (DBE) = $1 + 1/2(2C - H + N) \geq 0$; (2) $O{:}C \leq 1$; (3) $N \leq 4$ and $N{:}C \leq 1$; (4) $0.333 \leq H/C \leq 2.25$; (5) Nitrogen Rule; and (6) isotopic spacing and abundance. Relative intensity of each detected feature/formula in each sample was extracted for further statistical analysis.

2.5. Thermogravimetric analysis

To evaluate thermal lability of DOM, thermogravimetric analysis (TGA) was carried out with a thermogravimetric analyzer (Shimadzu TGA-50). About 2 mg of UDOM or SPEDOM were used for the analysis. Samples were pyrolyzed under N_2 gas at a flow rate of 50 $mL \cdot min^{-1}$, with temperature ramping from ca. 25 °C to 700 °C at a ramping rate of 20 $°C \cdot min^{-1}$. The weight of the remaining sample was measured every second, and the thermograms were generated using the LabSolution software (LabSolutions TA: Acquisition Version 1.00 SP1).

Different from typical ramped pyrolysis oxidation (RPO) analyses (e.g., Rosenheim et al., 2008; Hemingway et al., 2017, 2019; Rogers et al., 2021), in which O_2 was introduced into the combustion process, the pyrolysis in this study, including both the TGA and the thermal slicing ramped pyrolysis GC-MS analysis (described in section 2.6), was conducted under anoxic environment. Therefore, instead of recording the amount of CO_2 generated

during the analysis, attributes of the volatile materials evolved during the pyrolysis process (e.g., weights and their GC-MS signatures) were monitored. Following the idea of RPO, the thermogram data were further converted into activation energy distribution based on the inverse method (Hemingway et al., 2017, 2019). Specifically, the weight loss of organic materials during thermal decomposition, which is analogous to the release of CO_2 in RPO, was treated as a function of the activation energy (*E*) required to volatilize a given carbon atom (Hemingway et al., 2019). The *E* in RPO studies represents the energy required to convert organic carbon into CO_2. In contrast, *E* in this work includes the energy needed to release the already-volatile compounds from the organic matter matrix, as well as the energy required to break covalent bonds to generate new volatile components (a process known as "cracking"). Therefore, the *E* calculated from anoxic pyrolysis, somewhat different from RPO *E* in definition but in a same principle, is used as a proxy for estimating the bond strength of organic matter matrix.

2.6. Thermal slicing ramped pyrolysis GC-MS analysis

Protocols for thermal slicing ramped pyrolysis GC-MS (TSRP-GC-MS) have been described in previous works (Seeley et al., 2018; O'Connor et al., 2020). Briefly, ca. 2 mg of freeze-dried DOM were analyzed through a multi-shot pyrolyzer (EGA/PY-3030D, Frontier 96 Laboratories Ltd.) coupled with GC-MS (Shimadzu GCMS-TQ8040). The pyrolysis was conducted under helium, and the pyrolyzer was set to ramp from 50 °C to 650 °C at a constant heating rate of 20 °C·min^{-1}, consistent with the TGA settings and previous work (e.g., Collard and Blin, 2014). Six consecutive thermal slices, 50 – 90 °C, 90 – 170 °C, 170 – 290 °C, 290 – 370 °C, 370 – 530 °C, and 530 – 650 °C, were chosen. Based on previous literature on biopolymer pyrolysis and our own work (Goyal et al., 2008; Collard and Blin, 2014; Seeley et al., 2018; Wang et al., 2020), these six thermal slices can be further categorized into a dehydration zone of 50 – 170 °C, a dehydration and decarboxylation zone of 170 – 290 °C, a rapid depolymerization zone of 290 – 370 °C, and a fragmentation and cracking zone of 370 – 650 °C. Our previous work showed that this slicing procedure provided an ample separation of different compound classes and bond types for UDOM samples (O'Connor et al., 2020). The condensation and release of pyrolyzates from each slice was achieved through a cold trap, which was held at -190°C during the pyrolysis process by the N_2 gas connected to liquid N_2. The cold trap was turned off once the pyrolyzer reached the maximum temperature of a specific thermal slice, to allow the condensed pyrolyzates to pass through the GC column. The GC oven was set at 40 °C and held for 2 min before being ramped to 320 °C, by heating at 10 °C·min^{-1} to 150 °C and then at 20 °C·min^{-1} to 320 °C, and held for 3 min. The mass spectrometer was programed to detect ions in the range of 26 – 600 mass over charge (*m/z*) ratios.

Molecule identification was based upon several NIST open-source databases provided through the GC-MS software package. Only peaks with a match score of 70 or greater were integrated and taken into consideration for each sample (O'Connor et al., 2020). The resolution of the triple quadrupole mass spectrometer may not be high enough to identify a

compound with an accuracy of 100%. However, the database assignment provides key information on the core structure and/or the functional groups of a compound (O'Connor et al., 2020). Therefore, instead of the exact formula or structure, each detected compound was assigned to a specific "compound class". This assignment is based on the functional groups of the compounds, following a slightly modified order from previous work (O'Connor et al., 2020): nitrile → nitrogenous → phenol → aromatic → heterocyclic → ester → acetic → aldehyde → ketone → sterol → alcohol → ether → alkene → alkane. If one compound contains multiple functional groups, it will be assigned to the category with the highest rank. These same criteria were applied to all DOM samples.

2.7. Statistical Analyses

All statistical analyses were performed at a significance level of 0.05. Specifically, paired t test, conducted by Excel, was used to compare different properties between UDOM and SPEDOM. The principal component analysis (PCA; including amino acid composition data, and MS data) and hierarchical clustering analysis (HCA) were conducted using RStudio (2021.09.01 Build 372).

3. Results

3.1. Elemental and stable carbon isotope composition of UDOM and SPEDOM

Organic C content (OC%) in UDOM had a relatively large range, varying from 24.9% to 37.8% (Table 1), with lower OC% in the coastal saline SC sample. Compared with UDOM, OC% in SPEDOM samples was more unified, with the lowest value of 47.5% in GR and the highest of 49.1% in MR sample (Table 1). Overall, OC% of SPEDOM (48.6 ± 0.6%) was significantly higher than that of UDOM, 32.3 ± 5.5% ($p = 0.01$). A similar pattern was observed in organic N content (ON%; $p = 0.03$), as ON% of UDOM varied from 1.6% to 2.4%, and from 2.4% to 3.9% in SPEDOM (Table 1). However, there was no significant difference in the C/N ratios between the two types of DOM ($p = 0.50$), although the average C/N ratio was slightly higher in the SPEDOM samples (20.2 ± 4.4 vs. 18.8 ± 3.0). The lower OC content in UDOM samples likely resulted from the inclusion of inorganic materials such as colloidal metals and metalloids defined by membrane's cutoffs (e.g., Guo et al., 2000; Stolpe et al., 2013; Li et al., 2019). In contrast, SPE extraction is purely molecular property-based, thus only organic compounds with sufficient hydrophobicity can be retained on the resin, leading to a higher OC%.

$\delta^{13}C$ values of UDOM ranged from -23.1‰ to -25.9‰ (average -24.3 ± 1.1‰) and SEPDOM from -22.8‰ to -24.6‰ (average -23.9 ± 0.7‰; Table 1), showing no significant difference ($p = 0.69$). The $\delta^{13}C$ values of UDOM reported in this work were comparable to UDOM collected from the Mississippi River (Guo et al., 2009), In general, a similar trend of more depleted (negative) values in freshwater DOM samples was observed in UDOM (-24.5‰ in freshwater vs. -23.6‰ in saline water; Table 1). Overall, the C/N ratios and $\delta^{13}C$ values are indistinguishable between the UDOM and SPEDOM for either freshwater or saline water, indicating the same DOM sources.

3.2. Total hydrolyzable amino acids of UDOM and SPEDOM

Contribution of THAA carbon to total UDOM carbon varied from 2.9% in NR to 7.6% in GR (average 4.4 ± 1.9%), slightly higher than that for SPEDOM (2.0% to 2.8%, average 2.4 ± 0.4%), although the difference was not significant ($p = 0.10$; Table 1). Similarly, contribution of THAA nitrogen to total UDOM nitrogen varied from 16.7% to 29.1% (average 21.6 ± 4.6%), was also higher than that for SPEDOM (7.9% to 14.2%; average 13.8 ± 4.1%; $p = 0.10$; Table 1). The contributions of THAA-C and N were in the range of reported values for extracted DOM (e.g., Hubberten et al., 1995; Repeta, 2015).

THAA concentrations were further normalized to mole percentage (mol%) for each sample (Table 2). Glycine (GLY) was the most abundant amino acid in all samples, accounting for 20.3% of total THAA (by mol%) in UDOM samples and 22.3% in SPEDOM samples, consistent with the THAA composition of bulk DOM of the sampling region (Lu and Liu, 2019). Besides GLY, aspartic acid (ASP) and alanine (ALA) also contributed to a major fraction of the total THAA pool, representing 11.2% and 19.4% (ASP), and 13.5% and 12.8% (ALA) of UDOM and SPEDOM, respectively. Despite these similarities in the most abundant amino acids, the mol% of certain amino acids differed greatly between UDOM and

SPEDOM. For example, the mol% of serine (SER), arginine (ARG), threonine (THR), and phenylalanine (PHE) were significantly higher in UDOM samples than those in SPEDOM samples ($p < 0.05$). In contrast, amino acids, such as ASP, β-alanine (BALA), and γ-aminobutyric acid (GABA), were significantly more enriched in SPEDOM samples compared to UDOM samples ($p < 0.01$).

The compositional difference between UDOM and SPEDOM can be further visualized in a biplot derived from PCA, in which the samples are separated based on their THAA composition (Figure 2A). The first two principal components, namely PC1 and PC2, explained 56.6% and 16.5% of the data variance, respectively. The amino acids were clearly separated along the PC1 axis, with ASP, BALA, GABA, and GLY having the most negative loadings, while other amino acids (except for histidine, HIS) having positive loadings. UDOM samples were well separated from their SPE counterparts, with all UDOM samples in the positive region, while all SPEDOM samples in the negative region. This pattern is consistent with the observed higher concentrations of SER, ARG, THR, and PHE in UDOM, and higher concentrations of ASP, BALA, and GABA in SPEDOM.

To further evaluate the degradation status of UDOM and SPEDOM, degradation index (DI) values were calculated for each sample based on its THAA composition and the loading of each amino acid on the PC1 axis (Dauwe and Middelburg, 1998; Dauwe et al., 1999; Li et al., 2018; Liu and Xue, 2020). With higher DI values representing more labile organic materials, DI values ranged from -10.4 (SC) to -3.8 (MR) for SPEDOM samples, much lower than those of UDOM samples, varying from 5.6 (NR) to 13.7 (GR; Figure 2B).

3.3. Molecular level information of UDOM and SPEDOM

High resolution LC/MS can determine the molecular differences between UDOM and SPEDOM (Lu et al., 2018, 2021). Due to the predominant presence of carboxyl groups in DOM molecules (e.g., Zark and Dittmar, 2018), the negative mode of electrospray ionization (ESI-) has been commonly used in the molecular characterization of complex DOM mixtures (e.g., Kujawinski, 2002; Kim et al., 2003; Li and Minor, 2015; Wagner et al., 2019). However, the presence of high carboxyl-content aromatic compounds, especially in DOM samples enriched with terrestrial signals (e.g., riverine DOM samples; Kramer et al., 2012), can suppress the ionization of aliphatic and carbohydrate-like DOM (Ohno et al., 2016) during a negative ionization process. Considering that the latter structures are important fractions of UDOM (Benner et al., 1992; Whitehead, 2008), both negative (ESI-) and positive (ESI+) modes were applied in this work, aiming at a more comprehensive comparison between UDOM and SPEDOM.

The HPLC chromatograms, retention time vs. total ions, were distinctly different between UDOM and SPEDOM (Figure 3A&B). Specifically, the peak distributions at 3.6 min, and from 6.6 min to 8.6 min under ESI+ (Figure 3A), as well as from 3.7 min to 4.1 min under ESI- (Figure 3B), were different between UDOM and SPEDOM. But most strikingly, the total peak height and peak area of UDOM were much lower than those of SPEDOM under ESI+ mode, even though a similar amount of sample (ca. 0.1 – 0.2 μg of C) was

injected, and comparable chromatograms were acquired under ESI- mode. Such observation suggested a much lower ionization efficiency of UDOM compared with SPEDOM under ESI+ mode, hinting a difference in their molecular structures.

The number of features detected in each sample was consistent with the observed difference in ionization efficiency (lower in UDOM than SPEDOM, especially for ESI+). After blank subtraction, 167 – 204 features were detected in UDOM samples as compared to 1483 – 2953 for SPEDOM samples under ESI+ mode. Likewise, 181 – 751 features were detected in UDOM samples under ESI- mode, but a total number of 721 – 1245 was found in SPEDOM samples (Table 3). Roughly 50% of the detected features can be assigned with a CHNO formula for both UDOM and SPEDOM, and about 29% of the formulas assigned in UDOM samples can also be found in their corresponding SPEDOM samples under ESI+ mode, while the majority of formulas in SPEDOM (over 90%) was unique under both ESI modes. Such a fundamental difference in detected molecular formulas was also reflected by the different elemental stoichiometries between UDOM and SPEDOM. The averaged H/C ratios of UDOM under ESI+ ranged from 1.32 to 1.68, higher than the range of 1.26 to 1.66 for SPEDOM ($p = 0.13$, Table 3). Similarly, under ESI-, the averaged H/C ratios of UDOM (0.79–0.93) were higher than those of SPEDOM (0.75 – 0.86; $p = 0.08$; Table 3). The averaged O/C ratios varied from 0.25 to 0.29 for UDOM, and from 0.23 to 0.28 for SPEDOM under ESI+, while from 0.26 to 0.29 and from 0.29 to 0.37 for UDOM and SPEDOM under ESI-, respectively (Table 3). Contrary to H/C ratio, O/C was higher in UDOM under ESI+ ($p = 0.03$), but was lower in UDOM under ESI- ($p = 0.05$), consistent with a previous comparison between SPEDOM and dialysis-extracted DOM (Tfaily et al., 2012). The double bond equivalent (DBE) was higher in UDOM (6.3 – 12.1 under ESI+, 17.0 – 19.7 under ESI-) than in SPEDOM (4.9 – 12.7 under ESI+, 16.4 – 19.4 under ESI-), showing a higher degree of unsaturation for UDOM, even though the difference was not statistically significant (Table 3; $p = 0.10$ under ESI+, $p = 0.40$ under ESI-).

DOM molecules can be distinguished by their elemental composition, such as H/C, O/C, and N/C ratios, based on those of standard biomolecules (e.g., carbohydrates, lipids, etc.; Li and Minor, 2015). Seven discrete categories of DOM with different elemental stoichiometries were defined following previous studies (e.g., Kim et al., 2003; Ohno and Ohno, 2013; Mangal et al., 2016, 2017; Lu and Liu, 2019): (I) lipid-like structures with elemental ratios of: $0.01 \leq O/C \leq 0.1$; $1.5 \leq H/C \leq 2.0$; (II) unsaturated hydrocarbon-like structures: $0.01 \leq O/C \leq 0.1$; $0.75 \leq H/C < 1.5$; (III) condensed aromatic-like structures: $0.01 \leq O/C \leq 0.65$; $0.25 \leq H/C < 0.75$; (IV) protein-like structures: $0.1 < O/C \leq 0.65$; $1.5 \leq H/C \leq 2.25$; $N \geq 1$; (V) lignin-like structures: $0.1 < O/C \leq 0.65$; $0.75 \leq H/C < 1.5$; (VI) tannin-like structures: $0.65 < O/C \leq 0.85$; $0.75 \leq H/C \leq 1.5$; and (VII) carbohydrate-like structures: $0.65 < O/C \leq 1.0$; $1.5 < H/C \leq 2.25$. Note that without any actual standards analyzed, this broad categorization was an oversimplification of DOM, and did not necessarily guarantee all formulas within one category are truly sharing similar structures. However, such

simplification helps elucidate the chemical variations of DOM extracted via different methods.

Protein-like structures were a major component for both UDOM and SPEDOM under ESI+ mode, representing 11% – 33% and 22% – 36% of total assigned formulas, respectively. Following protein-like structures, 9% – 11% and 2% – 7% of formulas in UDOM and SPEDOM were categorized as lignin-like structures under ESI+, respectively (Figure 3C). On the other hand, formulas from ESI- were dominated by condensed aromatic-like structures with relatively low H/C ratios, accounting for 44% – 57% and 34% – 52% of total assigned formulas in UDOM and SPEDOM, respectively. Lignin-like structures were also important under ESI-, with 2% – 23% and 26% – 29% of assigned formulas, respectively (Figure 3D). Given that lignin-like structures also overlap with the well-recognized fraction of refractory DOM, namely carboxyl-rich alicyclic molecules (CRAM; Hertkorn et al., 2006), in the van Krevelen diagram, this observation of enriched lignin-like structures in river and coastal water DOM, especially under the ESI- mode, was consistent with the presence of CRAM in ultrahigh resolution MS studies (e.g., Hertkorn et al., 2006; Liu et al., 2011; Zigah et al., 2014; Li and Minor, 2015).

PCA was performed to further investigate the difference between detected UDOM and SPEDOM features (e.g., detected *m/z*s). UDOM samples were grouped more closely than SPEDOM samples under both ESI modes (Figures S1A&B), suggesting a rather similar DOM composition of all UDOM samples regardless of the sampling locations. On the other hand, SPEDOM samples were more diverse and also quite different from UDOM samples. The separation was more evident under ESI+ mode, consistent with the fact that many more features were detected for SPEDOM samples than for UDOM samples under ESI+ (Table 3).

PCA was further applied using the category results of UDOM and SPEDOM, where the differences were more evident (Figures S1C&D). PC1, mainly driven by positive scores of carbohydrate- and tannin-like structures, and negative scores of protein- and unsaturated hydrocarbon-like structures, explained 43.5% of the data variance under ESI+ mode, while PC2, explaining 28.7% of the variance, was driven by lignin- and lipid-like structures (Figure S1C). UDOM samples were mainly located in the upper right corner of the diagram, suggesting a higher abundance of lipid- and carbohydrate-like structures, consistent with previous studies (Simjouw et al., 2005; Tfaily et al., 2012). Carbohydrate- and tannin-like structures also contributed positively to PC1 (53.7%) under ESI- mode, while unsaturated hydrocarbon-like structures were the main driver for PC2 along the positive direction (24.1%; Figure S1D). Similarly, UDOM samples were in the right region of the PCA (higher abundance of carbohydrate- and lipid-like structures), and SPEDOM samples were in the left (enriched with lignin-like structures). In summary, UDOM samples were characterized by a higher fraction of carbohydrate- and lipid-like structures under both ESI modes, while SPEDOM samples were enriched with the more refractory lignin-like structures, consistent with the estimated higher fraction of humic substances in SPEDOM based on OC% analysis (Table 1).

In addition to PCA, HCA analysis, performed based on the log-transformed intensity of each detected feature in each sample, provided another way to visualize differences between UDOM and SPEDOM. The HCA results showed that certain compounds were more abundant with certain extraction method, with a difference of ≥ 2 in their normalized log-transformed relative intensity, reflected by a change of color from blue to yellow or red (Figures 4A&B). A zoom-in to these "unique" compounds revealed the high abundance of lignin-like structures in unique compounds in SPEDOM under both ESI modes (Figures 4D&F), as well as an enrichment of carbohydrate-like and unsaturated hydrocarbon-like features in UDOM under ESI- (Figure 4E). This pattern was consistent with the results from PCA based on compound category (Figure S1).

3.4. TGA results and activation energy of UDOM and SPEDOM

TGA tracks the weight change of DOM during the anoxic pyrolysis process. Regardless of the sampling locations, thermograms for DOM weight loss showed similar patterns when the same extraction technique was applied (Figures 5A, B&S2). For UDOM samples, there was a notable weight loss (ca. 10% – 15% of total weight) at a rather low temperature range (ca. 20 – 100 °C), corresponding to a sharp peak in the weight loss rate (Figures 5A&S2). As temperature increased, the weight loss continued progressively, with ca. 35% – 40% of total weight lost in the temperature range of 200 – 500 °C. At the end of the pyrolysis, about 50% of DOM weight remained and the materials appeared to be black char residuals visually. In contrast, the SPEDOM samples did not show much immediate weight loss at low temperature ranges until ca. 150 °C. The major weigh loss in SPEDOM samples mainly occurred within the range of 150 – 600 °C (Figures 5B&S2). However, a similar proportion of DOM remained at the end of the pyrolysis (ca. 40% – 50% of total weight for SPEDOM, ca. 50% for UDOM), with the same appearance of black chars.

The weight loss patterns of UDOM and SPEDOM were used to compute the activation energy distribution [AE distribution or p(E); Figure 5C]. p(E) has been used successfully as a proxy for organic matter bond-strength diversity (Hemingway et al., 2019). Even though the calculated AE may not be equivalent to the activation energy of the actual enzymatic respiration reactions (Hemingway et al., 2019), AE and its distribution p(E) nonetheless provided important insights into the thermal-stability of tested DOM. As shown in Figure 5, UDOM samples were characterized by a two-peak distribution, with one thermo-labile component having an AE less than 80 $kJ \cdot mol^{-1}$ (the AE value with the highest probability), and one thermo-refractory component with an AE of ca. 135 $kJ \cdot mol^{-1}$. SPEDOM samples, on the contrary, followed a mono-peak distribution, with AE with the highest probability at ca. 127 $kJ \cdot mol^{-1}$. Due to the presence of the thermo-labile fraction in UDOM, the calculated AE was consistently higher in SPEDOM (138.1 – 139.5 $kJ \cdot mol^{-1}$) than in UDOM (122.0 – 128.5 $kJ \cdot mol^{-1}$; Table 4; $p = 0.002$).

3.5. Pyrolyzates of UDOM and SPEDOM from thermal slicing ramped pyrolysis GC-MS (TSRP-GC-MS)

As a typical example, the MR UDOM sample generally had the most productive thermal slice at the temperature between 290 and 370 °C (Figure S3C), accounting for 34% – 54% of total pyrolyzates in terms of peak intensity (average of 46%) during the TSRP-GC-MS run. The 170 – 290 °C and 370 – 530 °C slices also contributed significantly to the total amount of pyrolyzates, averaging 20% and 32% of total pyrolyzates, respectively (Figures 6&S3). Together, 98% of pyrolyzates from UDOM samples were released in thermal slices from 170 °C to 530 °C. Similarly, slices of 170 – 530 °C contributed 96% of pyrolyzates for SPEDOM samples, but with different allocations. The most productive thermal slice of the SPEDOM was in the "cracking zone" (Goyal et al., 2008; Collard and Blin, 2014; O'Connor et al., 2020) at 370 – 530 °C, with 41% – 58% of total pyrolyzates generated, significantly higher than the corresponding region in UDOM samples ($p = 0.03$). Slices of 170 – 290 °C and 290 – 370 °C contributed an average of 28% and 15% of pyrolyzates, respectively (Figure 6F). The amount of pyrolyzates in the slice of 290 – 370 °C was significantly lower in SPEDOM than the corresponding region in UDOM ($p = 0.004$). If the thermal slices were further categorized based on the reactions occurring, where temperature > 370 °C was often considered to induce significant fragmentation and secondary cracking (O'Connor et al., 2020), the amount of pyrolyzates was significantly higher for UDOM than for SPEDOM in the non-cracking zone (i.e., 50 – 370 °C; $p = 0.02$).

Each detected peak in the GC-MS chromatograms was further assigned to a compound class by comparing with the NIST databases (O'Connor et al., 2020; Tables S2 – S5). Nitrogenous structures were the predominant compound class, representing an average of 31% (9% – 48%) and 32% (24% – 41%) of all pyrolyzates across all thermal slices for UDOM and SPEDOM samples, respectively. In addition, pyrolyzates from UDOM samples were characterized by a considerable amount of O-containing heterocyclic (12%; 8% – 20%), alcohol (9%; 4% – 11%), aromatic (8%; 5% – 10%), ketone (7%; 4% – 9%) and acetic structures (7%; 3% – 16%; Table 5). In contrast, pyrolyzates from SPEDOM samples were more unified, represented by fewer classes, mostly aromatic structures (22%; 17% – 33%) and alkene structures (10%; 4% – 20%; Table 5).

Such pattern was also observed when examining the pyrolyzates on a per thermal slice basis (Figure 7). From a quantitative point of view, the molecular diversity scores (Shannon-Wiener index; S-W), which were calculated based on the relative intensity of each pyrolyzate class across all thermal slices (O'Connor et al., 2020), varied from 1.9 to 2.3 with an average of 2.1 for UDOM samples, and from 1.9 to 2.1 with an average of 2.0 for SPEDOM samples (Table 5; Figure 2). S-W scores of UDOM samples were higher than those of SPEDOM samples, though the difference was not statistically significant ($p = 0.2$). The DI values were positively correlated with S-W scores ($R^2 = 0.33$; $p = 0.14$), suggesting that less degraded, or more labile, DOM also had a higher pyrolyzate diversity, consistent with our previous work (O'Connor et al., 2020).

4. Discussion

4.1. Contrasting biolability between UDOM and SPEDOM

Understanding the biolability of DOM may represent the first yet crucial step to decipher the enigma of the long-term stability of DOM. It has been widely acknowledged that a minuscule fraction of annual DOM production could escape the decomposition and accumulate for a time scale of thousands of years (Dittmar, 2015). However, studies explicitly elucidating the molecular difference between the refractory and the labile fractions remain scarce (e.g., Simjouw et al., 2005). Aiming at isolating and differentiating sub-fractions of DOM, much previous research has compared the various attributes of DOM extracted by different approaches. For instance, results of ^{13}C solid state NMR revealed a strong signal of carbonyls at the frequency range of 175 – 180 ppm in UDOM (Mopper et al., 2007; Whitehead, 2008), but not in SPEDOM (e.g., Mopper et al., 2007; Hertkorn et al., 2013). In addition, peaks of anomeric protons (ca. 5.2 ppm), O-alkyl protons (3.5 – 4.5 ppm), and amide protons (ca. 8.0 ppm) in the ^{1}H NMR of UDOM are also absent in SPEDOM, while methyl and methylene protons (0.9 – 3.3 ppm) were much higher in SPEDOM (Kaiser et al., 2003; Repeta, 2015). Furthermore, the molecular compositions of DOM isolated via different SPE sorbents (e.g., C_{18}, XAD, PPL, diethylaminoethyl-cellulose, and St-DVB) have also been compared through UV-Visible spectrophotometry, and ultrahigh resolution mass spectrometry (Perminova et al., 2014; Li and Minor, 2015), showing the diversity of DOM, as well as a preferential extraction of CHON compounds by C_{18} sorbent. However, such molecular level information on sub-fractions of DOM has seldom been further connected to biolability or any indicator of biolability.

Even though THAAs only represent a small fraction of extracted DOM (2.0-7.6% C and 7.9-29.1% N, Table 1), they are sensitive enough to reflect the biolability or freshness of the bulk sample. For example, amino acids, such as GLU, ARG, LEU, and ILE, were often enriched in fresher DOM samples (Li et al., 2018), whereas more GLY, BALA, and GABA were enriched in highly altered or degraded samples (Dauwe and Middelburg, 1998; Chen et al., 2004; Davis et al., 2009; Peter et al., 2012). The contribution of THAA to DOM also can be used as an indicator for lability, as THAA-C/DOC and THAA-N/DON generally decreased with increasing recalcitrance (Repeta, 2015). The slightly higher contribution of THAA for UDOM already hinted a more labile nature of UDOM, even though the difference was not significant (Table 1). Our results further showed that GLY was more enriched in SPEDOM than in UDOM, consistent with previous work that compared amino acid composition between UDOM and SPEDOM extracted from ultrafiltration filtrate (e.g., over 25% in SPEDOM compared with ca. 20% in UDOM; Broek et al., 2019). The PCA results, based on amino acid composition, also suggested that UDOM was more labile or fresher than SPEDOM, as reflected by their highly different DI scores (Table 2; 7.9 ± 3.3 for UDOM and -7.9 ± 2.9 for SPEDOM samples). Such trend remained evident when a suite of organic materials with different diagenetic status were included in the data matrix (Figure 8; Dauwe and Middelburg, 1998). Both UDOM and SPEDOM were located far apart from the labile

organic materials (e.g., algae and phytoplankton, Phy-B and Phy-C; and sediment trap material, Saan-T; Dauwe et al., 1999), but were closer to the highly degraded samples (Dauwe et al., 1999), such as oxidized turbidites (Tur-ox), deep samples (SK), and deep sediments (Dab-S2 and Saan-S2). Moreover, SPEDOM samples had more negative loadings on PC1 (Figure 8), confirming a higher degradation status of SPEDOM than of UDOM.

A higher biolability of UDOM is also supported by the high-resolution LC/MS results. The H/C ratios of UDOM were higher than those of SPEDOM under both ESI modes (Table 3). H/C was often considered as an index indicating the lability of DOM in mass spectrometry studies (D'Andrilli et al., 2015; Spencer et al., 2016; Rogers et al., 2021), with higher H/C representing more labile fraction of DOM, thus a higher H/C detected in UDOM (1.51 vs. 1.49 under ESI+; 0.87 vs. 0.82 under ESI-) was indicative of higher lability in UDOM. Compound category analysis further demonstrated a higher abundance of lipid- (9.4% vs. 4.7% under ESI+; 1.5% vs. 1.2% under ESI-) and carbohydrate-like structures (2.4% vs. 1.9% under ESI+; 10.7% vs. 3.9% under ESI-) in UDOM, but a higher proportion of lignin-like structures (9.4% vs. 28.7% under ESI+; 1.5% vs. 26.9% under ESI-) in SPEDOM, consistent with results obtained via Fourier transform infrared spectroscopy (FTIR) and direct temperature-resolved mass spectrometry (DT-MS; Simjouw et al., 2005). Even though the actual molecular structures cannot be confirmed based on MS analysis alone, the difference in categories between UDOM and SPEDOM nonetheless hinted a more bio-refractory property (the recalcitrant lignin-like structures) of SPEDOM than UDOM (the bio-labile carbohydrate-like and lipid-like structures).

The different biolabilities between UDOM and SPEDOM further confirmed that natural DOM is highly heterogeneous in terms of its molecular composition and reactivity, and that different isolation methods/procedures only physically (i.e., ultrafiltration) or chemically (i.e., SPE) extract certain fractions of the bulk DOM. Indeed, the UDOM, taking the form of enriched high H/C ratio carbohydrate- and lipid-like compounds, represented the more biolabile fraction of DOM, as compared to SPEDOM with lower H/C ratio, and more hydrophobic compounds.

4.2. Thermal lability of UDOM vs. SPEDOM

Thermal lability of a DOM sample refers to the easiness of degradation upon pyrolysis, which can be monitored through the released pyrolyzates or the leftover weight of a sample with time. Thermal lability of natural organic matter has regained attention recently, and has been considered correlated to biolability albeit in a complicated manner (e.g., Leifeld and von Lützow, 2014; Chen et al., 2018; O'Connor et al., 2020), as the breakdown of chemical bonds in organic matter under heating conditions may not necessarily be equivalent to environmental enzymatic respiration reactions (Hemingway et al., 2019). Nevertheless, it has been shown that thermally labile fraction of organic matter was generally also ^{14}C-enriched or younger in ^{14}C ages, and that ^{14}C-depleted old DOM was predominantly derived from thermally refractory components (Zigah et al., 2017), suggesting that interpreting the thermal lability of DOM can offer clues to the biolability of a sample as an independent angle

(O'Connor et al., 2020). In addition to the bonding strength, the analysis of the volatile pyrolyzates released during the pyrolysis process provides molecular level information of DOM.

The TGA results provided direct evidence regarding the higher thermal lability of UDOM than SPEDOM. The thermochemical AE results demonstrated that higher activation energy is required to break down or release volatile components from SPEDOM (138.8 $kJ \cdot mol^{-1}$) than from UDOM (125.4 $kJ \cdot mol^{-1}$; Table 4). The AE of UDOM and SPEDOM calculated in this work was much lower than the reported values of SPEDOM, riverine particulate organic matter, sedimentary organic matter, and soil organic matter derived from RPO (Hemingway et al., 2019; Rogers et al., 2021), which generally had an averaged AE at around 160 $kJ \cdot mol^{-1}$. Such discrepancy was likely due to the sample matrix, as organic matter from sediments or soils is likely more structurally complex and thus more thermally stable than DOM collected from surface waters. The different pyrolysis methods used in different studies may have also played a factor. DOM was heated under anoxic conditions in this work, whereas an 8% O_2 was applied in RPO. With almost all organic carbon transferred to CO_2 in RPO (Rosenheim et al., 2008), the much higher/more complete extent of reaction would result in a higher AE. For comparison, ca. 50% of materials were left in anoxic pyrolysis, whereas the RPO may fully oxidize all the carbon (Rosenheim et al., 2008). Nevertheless, since the AE of SPEDOM and UDOM reported here was acquired using the same pyrolysis technique and procedures, our conclusion about higher thermal lability in UDOM than SPEDOM should be robust.

In addition to a higher amount of pyrolyzates released at higher temperature slices in SPEDOM than in UDOM (370 – 650 °C; Figure 6F; average of 33% for UDOM vs. average of 53% for SPEDOM; $p = 0.02$), the specific compound class results from the TSRP-GC-MS analysis also indicated a higher thermal lability of UDOM. Specifically, the different contributions of nitrogenous structures between UDOM and SPEDOM supported the conclusion, as a much higher fraction of nitrogenous structures in UDOM occurred at low temperature zone (50 – 170 °C; 71% of total pyrolyzates in UDOM vs. 28% of total pyrolyzates in SPEDOM; $p = 0.05$; Figure S4; Tables S2 – S5). Such enrichment in labile nitrogenous structures may reflect the presence of peptides and amide groups in UDOM based on NMR results (Kaiser et al., 2003). On the other hand, nitrogenous structures were consistently higher in SPEDOM in all high temperature thermal slices (170 – 650 °C; Figures 7&S4, Tables S2 – S5; p ranged from 0.13 to 0.43). As thermally stable nitrogenous structures (i.e., the nitrogenous structures eluted at high temperatures) were indicative of thermal stable and bio-recalcitrant components (O'Connor et al., 2020), higher contributions of nitrogenous components in SPEDOM confirmed its lower thermal lability as well as the lower biolability compared with UDOM.

Consistently, the relative amounts of aromatic and phenol pyrolyzates were higher in SPEDOM than in UDOM across all thermal slices (e.g., for aromatic structures, $p = 0.01$ for 50 – 170 °C; $p = 0.02$ for 170 – 290 °C; $p = 0.04$ for 290 – 370 °C; $p = 0.03$ for 370 –

530 °C; $p = 0.31$ for 530 – 650 °C; Tables S2 – S5). In contrast, the relative contribution of aliphatic structures (e.g., ketones, aldehyde, alcohol structures) was greater in UDOM than in SPEDOM, especially in the non-cracking zones (e.g., for ketone structures, $p = 0.07$ for 170 – 290 °C; $p = 0.04$ for 290 – 370 °C; $p = 0.04$ for 370 – 530 °C; Tables S2 – S5). As the thermally stable aromatic compounds with low H/C ratios (De la Rosa et al., 2008; Rogers et al., 2021) were considered as less biolabile than aliphatic compounds with higher H/C ratios (D'Andrilli et al., 2015; Spencer et al., 2016), the enrichment of aromatic pyrolyzates in SPEDOM thus not only confirmed its lower thermal lability, but also was in accordance with the less biolabile nature of SPEDOM.

In fact, this observed enrichment of aliphatic pyrolyzates in UDOM and the higher fraction of aromatic pyrolyzates in SPEDOM can be further explained by the molecular structure of DOM based on the MS results. The high abundance of refractory components in SPEDOM, namely the carboxyl-rich alicyclic molecules (CRAM) and the carotenoid-like linear terpenoids (MDLT; Arakawa et al., 2017), may facilitate the formation of aromatic structures during anoxic pyrolysis, due to their alicyclic structures. On the contrary, the carbohydrate-like and lipid-like structures in UDOM may results in the productions of ketone, aldehyde and alcohol structures (e.g., Zhou et al., 2017).

Finally, the overall distribution patterns of different groups of pyrolyzates offer insights into the labilities of UDOM and SPEDOM. Previous studies have confirmed the loss of molecular diversity with increasing degradation status of DOM (e.g., O'Connor et al., 2020), with the least diagenetically altered DOM having the highest molecular diversity of pyrolyzates. Taking all thermal slices into consideration, a slightly higher S-W was found in UDOM (2.1) than in SPEDOM (2.0) in almost all samples, except for the coastal SC samples (Figure 2B). Overall, together with a positive correlation between S-W scores and DI, the TSRP-GC-MS data consistently attested to a more labile nature of UDOM compared with SPEDOM.

4.3. UDOM, SPEDOM, and total DOM

Our multifaceted DOM characterization demonstrated that UDOM and SPEDOM, as sub-fractions of total DOM, are fundamentally different in lability and reactivity, a conclusion consistent with previous studies in which C_{18} SPE was applied (Simjouw et al., 2005; Minor et al., 2014). Yet it is quite surprising to see such two molecularly distinct sub-fractions of DOM still possessed similar bulk properties (Table 1). For instance, as labile DOM generally has a lower C/N compared with refractory DOM (e.g., Hopkinson and Vallino, 2005), UDOM would be expected to have a lower C/N compared with SPEDOM. In fact, a previous study showed that the C/N ratio of UDOM (ca. 12 – 13) was significantly lower than that of SPEDOM (ca. 18 – 20) in open ocean (Broek et al., 2017). The similar C/N ratios between labile sub-fraction (i.e., UDOM) and refractory sub-fraction (i.e., SPEDOM) suggested that at least in this specific study area, N-containing compounds distributed uniformly in DOM, and were not preferentially enriched in the labile sub-fraction in comparison with open ocean DOM.

The low overlap of common molecular formulas between UDOM and SPEDOM, ca. 30% for UDOM and 10% for SPEDOM (Figure 4), showed an extremely diverse nature of bulk DOM, even though the calculated overlap may only represent a lower estimation of the actual commonality, given the limitation of MS analysis. It further suggested that the commonality between DOMs recovered by size- and chemistry-based approaches was quite limited, at least in the investigated systems (i.e., south Texas rivers and more oligotrophic coastal waters). A detailed molecular comparison between UDOM and SPEDOM has rarely been done thus far, but when compared with previous studies using other extraction techniques, the reported overlap of molecular formulas in this work was much lower than that between SPEDOM extracted using different sorbents (ca. 70%; Li and Minor, 2015), and somewhat lower than that between SPEDOM and dialysis-extracted DOM (ca. 40%; Tfaily et al., 2012). Even though direct comparison has not been documented previously, an overlap between UDOM and SPEDOM can be roughly estimated based on the recovery rates in the coupled ultrafiltration-SPE work (e.g., Simjouw et al., 2005; Broek et al., 2017). With a recovery rate of ca. 10 – 30% for ultrafiltration, ca. 45% for SPE on bulk seawater, and ca. 45% for SPE on ultrafiltration permeate (Broek et al., 2017), the overlap between UDOM and SPEDOM was estimated to be 10% for SPEDOM, and 45% for UDOM, not quite different from the values in this work.

Such large difference in molecular formulas between UDOM and SPEDOM was not unexpected, as previous NMR work has already demonstrated their large compositional difference (Mopper et al., 2007; Whitehead, 2008; Hertkorn et al., 2013; Repeta, 2015). Furthermore, the distinct difference between SPEDOM and UDOM was supported by the published ^{14}C age data, where PPL SPEDOM generally had a depleted ^{14}C signal ranging from ca. -400 ‰ to lower than -500 ‰ (Bercovici et al., 2018), substantially lower than that of for UDOM (e.g., Guo et al., 2009; Walker et al., 2011; Broek et al., 2017). In this work the SPEDOM and UDOM together may explain over 70% of the total DOM pool, not quite different from previously published values in the coupled ultrafiltration-SPE approach (e.g., 70% in Simjouw et al., 2005; 40% in Broek et al., 2017). Compared with single extraction approach, combination of multifaceted characterization techniques and comparisons in molecular composition between concurrently isolated UDOM and SPEDOM samples clearly provide a more complete picture of bulk DOM, reveal inherent fractionation during DOM isolation, and point to an urgent need in better deciphering the composition, reactivity and heterogeneity of natural DOM.

5. Conclusion

In this work DOM was concurrently isolated using two most commonly used isolation methods, ultrafiltration (UDOM) and PPL-SPE (SPEDOM), and their bio- and thermal lability were systematically evaluated using a series of chemical/molecular analyses. High resolution LC/MS demonstrated that UDOM and SPEDOM represented fundamentally different fractions of the total DOM pool, as suggested by the low percentage of overlap for the assigned formulas. In addition, amino acid composition analysis, elemental analysis, as well as thermochemical analyses including activation energy calculation and specific pyrolyzates, further showed that UDOM contained more labile, both biologically and thermally, DOM fractions, while SPEDOM contained more recalcitrant components. A higher activation energy (139 $kJ \cdot mol^{-1}$) was observed for SPEDOM samples compared to UDOM (125 $kJ \cdot mol^{-1}$). Furthermore, this work demonstrated that specific analyses (e.g., the nitrogenous fraction, aliphatic and aromatic fractions) through multiple approaches (e.g., combination of LC/MS and TSRP-GC-MS) should provide new insights into the formation and preservation of DOM in aquatic systems. Finally, in light of the diverse nature of DOM, this work further confirmed the occurrence of DOM fractionation during isolation and the bias in DOM molecular composition/structures based on currently available isolation approaches, and further suggested that current extraction techniques isolate fundamentally different sub-fractions of DOM. Thus, caution is needed when interpreting DOM data using isolates from specific extraction techniques.

Acknowledgement

We thank Dr. Ryan Hladyniuk from UTMSI Core Facility Lab for helping assemble and maintain the equipment. Special thanks to Drs. Robert Dickey and Ed Buskey for making the instrumentation available at UTMSI. We also would like to thank John O'Connor for his help in sampling trips and Hui Lin for his help during sample ultrafiltration. This study was supported by American Chemical Society (ACS) Petroleum Research Fund (60531-ND2 to ZL). Publication supported in part by an Institutional Grant (NA18OAR4170088) to the Texas Sea Grant College Program from the National Sea Grant Office, National Oceanic and Atmospheric Administration (NOAA), U.S. Department of Commerce.

The data that support the findings of this study are available on request from the corresponding author, KL.

Tables

Table 1. C and N contents, C/N ratio, $\delta^{13}C$ value, and contributions of total hydrolyzable amino acid (THAA) C and N to UDOM and SPEDOM.

Samples	UDOM						SPEDOM					
	OC%	ON%	C/N	$\delta^{13}C$ (‰)	THAA-C/DOC (%)	THAA-N/DON (%)	OC%	ON%	C/N	$\delta^{13}C$ (‰)	THAA-C/DOC (%)	THAA-N/DON (%)
GR	29.2	2.4	14.2	-25.9	7.6	29.1	47.5	3.9	14.2	-24.0	2.0	7.9
MR	37.8	2.0	22.1	-23.1	3.3	20.5	49.1	2.3	24.9	-24.6	2.8	19.4
NR	37.3	2.1	20.7	-24.6	2.9	16.7	48.9	3.2	17.8	-22.8	2.8	14.2
SC	24.9	1.6	18.2	-23.6	3.9	20.0	48.9	2.4	23.8	-24.3	2.1	14.0
Avg ± SD	32.3 ± 5.5	2.0 ± 0.3	18.8 ± 3.0	-24.3 ± 1.1	4.4 ± 1.9	21.6 ± 4.6	48.6 ± 0.6	3.0 ± 0.7	20.2 ± 4.4	-23.9 ± 0.7	2.4 ± 0.4	13.8 ± 4.1

Table 2. THAA mol% and Degradation Index (DI) for each sample.

Sample	ASP	GLU	HIS	SER	ARG	GLY	THR	BALA	ALA	TYR	GABA	MET	VAL	PHE	ILE	LEU	DI
UDOM-MR	11.3	9.7	1.1	9.1	2.5	21.9	11.5	2.3	11.9	1.4	0.7	0.4	6.3	2.4	3.0	4.4	5.9
UDOM-NR	10.9	8.7	0.7	8.8	2.6	23.2	9.5	2.9	13.3	1.7	0.7	0.8	6.1	2.5	2.9	4.9	5.6
UDOM-SC	9.7	8.8	0.4	10.3	2.8	20.4	13.4	2.6	14.6	1.6	0.5	0.9	5.2	2.1	2.5	4.1	6.6
UDOM-GR	13.0	10.2	0.7	7.9	3.9	15.4	9.5	2.0	14.2	2.4	0.8	0.4	7.4	3.2	2.9	6.0	13.7
SPEDOM-MR	19.2	8.9	2.1	5.4	1.8	20.7	7.5	5.3	11.8	1.9	1.9	0.5	5.4	2.0	2.3	3.3	-3.8
SPEDOM-NR	17.2	8.9	0.8	3.8	1.5	25.0	6.5	7.4	12.0	1.1	2.3	0.3	5.6	1.9	2.4	3.4	-8.0
SPEDOM-SC	18.4	8.7	0.2	3.9	0.8	23.4	8.2	7.0	15.1	0.5	2.4	0.0	6.1	1.6	0.0	3.7	-10.4
SPEDOM-GR	22.6	8.2	0.8	3.7	1.1	20.3	6.3	8.0	12.2	1.3	2.4	0.3	5.4	2.0	2.1	3.5	-9.5

Table 3. Molecular information of UDOM and SPEDOM based on LC/MS analysis

Samples	Number of features detected	Number of formulas assigned	Averaged H/C	Averaged O/C	Averaged DBE
UDOM GR (ESI+)	167	36	1.68	0.25	6.3
UDOM MR (ESI+)	188	67	1.50	0.25	10.8
UDOM NR (ESI+)	168	33	1.56	0.29	8.0
UDOM SC (ESI+)	204	47	1.32	0.28	12.1
SPEDOM GR (ESI+)	2953	1642	1.66	0.23	4.9
SPEDOM MR (ESI+)	2826	1505	1.52	0.25	6.6
SPEDOM NR (ESI+)	2235	1164	1.52	0.28	6.6
SPEDOM SC (ESI+)	1483	744	1.26	0.27	12.7
UDOM GR (ESI-)	730	318	0.90	0.28	17.0
UDOM MR (ESI-)	707	394	0.79	0.27	19.7
UDOM NR (ESI-)	751	367	0.84	0.26	18.7
UDOM SC (ESI-)	181	60	0.93	0.29	18.9
SPEDOM GR (ESI-)	1245	565	0.86	0.29	17.7
SPEDOM MR (ESI-)	926	476	0.82	0.29	18.5
SPEDOM NR (ESI-)	721	329	0.75	0.30	20.9
SPEDOM SC (ESI-)	1110	391	0.86	0.37	17.1

Table 4. Calculated activation energy ($kJ \cdot mol^{-1}$) for each sample

Samples	UDOM ($kJ \cdot mol^{-1}$)	SPEDOM ($kJ \cdot mol^{-1}$)
GR	123.2	138.1
MR	127.8	138.8
NR	128.5	138.8
SC	122.0	139.5
Avg ± SD	125.4 ± 2.8	138.8 ± 0.5

Table 5. Molecular information of UDOM and SPEDOM based on TSRP-GC-MS analysis

Samples	Alkene	Alkane	Ketone	Alcohol	Aromatic	Phenol	Heterocyclic	Acetic	Aldehyde	Sterol	Ester	Ether	Nitrogenous	Nitrile	S-W
UDOM-MR	7.4	4.3	7.3	4.3	9.8	4.1	9.3	3.4	6.5	0.5	4.0	0.5	35.1	3.5	2.2
UDOM-NR	7.1	0.6	8.7	14.9	9.1	1.7	10.9	3.0	1.0	0.0	4.0	2.2	29.6	7.1	2.2
UDOM-SC	2.8	6.3	4.4	7.0	5.0	0.5	8.3	4.4	3.0	0.0	1.3	0.7	48.3	7.9	1.9
UDOM-GR	6.0	4.8	7.1	11.3	6.9	0.9	20.1	15.6	8.8	0.0	3.7	0.8	9.2	4.5	2.3
SPEDOM-MR	11.4	4.7	2.3	6.1	33.0	5.8	2.5	4.4	0.3	0.2	3.7	0.4	24.5	0.7	2.0
SPEDOM-NR	4.9	5.5	3.7	9.1	20.9	1.7	3.7	4.6	0.8	0.0	2.6	0.8	41.1	0.6	1.9
SPEDOM-SC	20.5	3.6	2.6	1.9	17.2	3.1	5.6	13.8	4.0	0.0	2.5	0.0	22.5	2.6	2.1
SPEDOM-GR	4.4	7.6	2.8	8.0	16.9	2.8	5.8	6.6	0.4	0.1	2.9	0.2	40.3	1.4	1.9

Figures

Figure 1. Sampling locations

Figure 2. (A) PCA based on THAA composition; (B) Degradation Index (DI) values and Shannon-Wiener index (S-W) scores of DOM samples extracted by ultrafiltration (red) and solid phase extraction (blue). The DI values were derived from the loadings of THAA-PCA. The S-W scores were derived from the relative intensity of each pyrolyzate class.

Figure 3. (A) HPLC chromatograms of UDOM (red) and SPEDOM (blue) acquired under ESI+ mode; (B) HPLC chromatograms of UDOM (red) and SPEDOM (blue) acquired under ESI- mode; (C) van Krevelen diagram of UDOM (red) and SPEDOM (blue) acquired under ESI+ mode; (D) van Krevelen diagram of UDOM (red) and SPEDOM (blue) acquired under ESI- mode. The van Krevelen diagrams are divided into seven discrete regions (indicated by Roman numbers): (I) lipid-like structures, (II) unsaturated hydrocarbon-like structures, (III) condensed aromatic-like structures, (IV) protein-like structures, (V) lignin-like structures, (VI) tannin-like structures, and (VII) carbohydrate-like structures.

Figure 4. HCA based on the log-transformed relative intensity under (A) ESI+ and (B) ESI-, with a brighter color (yellow and red) indicating a higher relative intensity, while a colder color (blue) representing a lower relative intensity; (C) DOM molecules with elevated intensities in UDOM samples under ESI+; (D) DOM molecules with elevated intensities in SPE samples under ESI+; (E) DOM molecules with elevated intensities in UDOM samples under ESI-; (F) DOM molecules with elevated intensities in SPE samples under ESI-.

Figure 5. (A) Thermogram of UDOM collected at NR, with the remaining weight% shown in green, and the weight loss rate ($\%\cdot s^{-1}$) shown in red; (B) Thermogram of SEPDOM collected at NR, with the remaining weight% shown in green, and the weight loss rate ($\%\cdot s^{-1}$) shown in blue; (C) Calculated averaged activation energy distributions for UDOM (red) and SPEDOM (blue) based on the inverse model; the shaded area representing the standard error for either extraction method.

Figure 6. Percentage of pyrolyzates (in terms of intensities) at each thermal slice/temperature zone for UDOM samples (red) and SPEDOM samples (blue).

Figure 7. Averaged percentage of each compound class for every thermal slice for (A) UDOM samples and (B) SPEDOM samples.

Figure 8. PCA based on THAA composition with THAA data from Dauwe and Middelburg (Dauwe and Middelburg, 1998) included.

References

Amon R. M. W., Fitznar H.-P. and Benner R. (2001) Linkages among the bioreactivity, chemical composition, and diagenetic state of marine dissolved organic matter. *Limnol. Oceanogr.* **46**, 287–297.

Arakawa N., Aluwihare L. I., Simpson A. J., Soong R., Stephens B. M. and Lane-Coplen D. (2017) Carotenoids are the likely precursor of a significant fraction of marine dissolved organic matter. *Sci. Adv.* **3**, e1602976.

Benner R. and Amon R. M. W. (2015) The Size-Reactivity Continuum of Major Bioelements in the Ocean. *Annu. Rev. Mar. Sci.* **7**, 185–205.

Benner R., Pakulski J. D., McCarthy M., Hedges J. I. and Hatcher P. G. (1992) Bulk chemical characteristics of dissolved organic matter in the ocean. *Science*.

Bercovici S. K., Koch B. P., Lechtenfeld O. J., McCallister S. L., Schmitt-Kopplin P. and Hansell D. A. (2018) Aging and molecular changes of dissolved organic matter between two deep oceanic end-members. *Glob. Biogeochem. Cycles* **32**, 1449–1456.

Boon J. J. and Haverkamp J. (1979) Pyrolysis mass spectrometry of a benthic marine ecosystem—the influence of Arenicola marina on the organic matter cycle. *Neth. J. Sea Res.* **13**, 457–478.

Broek T. A. B., Bour A. L., Ianiri H. L., Guilderson T. P. and McCarthy M. D. (2019) Amino acid enantiomers in old and young dissolved organic matter: Implications for a microbial nitrogen pump. *Geochim. Cosmochim. Acta* **247**, 207–219.

Broek T. A. B., Walker B. D., Guilderson T. P. and McCarthy M. D. (2017) Coupled ultrafiltration and solid phase extraction approach for the targeted study of semi-labile high molecular weight and refractory low molecular weight dissolved organic matter. *Mar. Chem.* **194**, 146–157.

Chen J., Li Y., Yin K. and Jin H. (2004) Amino acids in the Pearl River Estuary and adjacent waters: origins, transformation and degradation. *Cont. Shelf Res.* **24**, 1877–1894.

Chen M., Kim S., Park J.-E., Kim H. S. and Hur J. (2016) Effects of dissolved organic matter (DOM) sources and nature of solid extraction sorbent on recoverable DOM composition: Implication into potential lability of different compound groups. *Anal. Bioanal. Chem.* **408**, 4809–4819.

Chen W.-H., Chu Y.-S., Liu J.-L. and Chang J.-S. (2018) Thermal degradation of carbohydrates, proteins and lipids in microalgae analyzed by evolutionary computation. *Energy Convers. Manag.* **160**, 209–219.

Collard F.-X. and Blin J. (2014) A review on pyrolysis of biomass constituents: Mechanisms and composition of the products obtained from the conversion of cellulose, hemicelluloses and lignin. *Renew. Sustain. Energy Rev.* **38**, 594–608.

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

Dauwe B. and Middelburg J. J. (1998) Amino acids and hexosamines as indicators of organic matter degradation state in North Sea sediments. *Limnol. Oceanogr.* **43**, 782–798.

Dauwe B., Middelburg J. J., Herman P. M. J. and Heip C. H. R. (1999) Linking diagenetic alteration of amino acids and bulk organic matter reactivity. *Limnol. Oceanogr.* **44**, 1809–1814.

Davis J., Kaiser K. and Benner R. (2009) Amino acid and amino sugar yields and compositions as indicators of dissolved organic matter diagenesis. *Org. Geochem.* **40**, 343–352.

De la Rosa J. M., González-Pérez J. A., González-Vázquez R., Knicker H., López-Capel E., Manning D. A. C. and González-Vila F. J. (2008) Use of pyrolysis/GC–MS combined with thermal analysis to monitor C and N changes in soil organic matter from a Mediterranean fire affected forest. *CATENA* **74**, 296–303.

Dittmar T. (2015) Reasons behind the long-term stability of dissolved organic matter. In *Biogeochemistry of Marine Dissolved Organic Matter (Second Edition)* (ed. C. A. Carlson). Academic Press, Boston. pp. 369–388.

Dittmar T., Koch B., Hertkorn N. and Kattner G. (2008) A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnol Ocean. Methods* **6**, 230–235.

Eglinton T. I. and Repeta D. J. (2014) Organic matter in the contemporary ocean. In *Treatise on Geochemistry (Second Edition)* (eds. H. D. Holland and K. K. Turekian). Elsevier, Oxford. pp. 151–189.

Goyal H. B., Seal D. and Saxena R. C. (2008) Bio-fuels from thermochemical conversion of renewable resources: A review. *Renew. Sustain. Energy Rev.* **12**, 504–517.

Guo L. and Santschi P. H. (1996) A critical evaluation of the cross-flow ultrafiltration technique for sampling colloidal organic carbon in seawater. *Mar. Chem.* **55**, 113–127.

Guo L., Santschi P. H. and Warnken K. W. (2000) Trace metal composition of colloidal organic material in marine environments. *Mar. Chem.* **70**, 257–275.

Guo L., White D. M., Xu C. and Santschi P. H. (2009) Chemical and isotopic composition of high-molecular-weight dissolved organic matter from the Mississippi River plume. *Mar. Chem.* **114**, 63–71.

Hemingway J. D., Rothman D. H., Grant K. E., Rosengard S. Z., Eglinton T. I., Derry L. A. and Galy V. V. (2019) Mineral protection regulates long-term global preservation of natural organic carbon. *Nature* **570**, 228.

Hemingway J. D., Rothman D. H., Rosengard S. Z. and Galy V. V. (2017) An inverse method to relate organic carbon reactivity to isotope composition from serial oxidation. *Copernic. Publ.* **14**, 5099–5114.

Hertkorn N., Benner R., Frommberger M., Schmitt-Kopplin P., Witt M., Kaiser K., Kettrup A. and Hedges J. I. (2006) Characterization of a major refractory component of marine dissolved organic matter. *Geochim. Cosmochim. Acta* **70**, 2990–3010.

Hertkorn N., Harir M., Koch B. P., Michalke B. and Schmitt-Kopplin P. (2013) High-field NMR spectroscopy and FTICR mass spectrometry: powerful discovery tools for the molecular level characterization of marine dissolved organic matter. *Biogeosciences* **10**, 1583–1624.

Hopkinson C. S. and Vallino J. J. (2005) Efficient export of carbon to the deep ocean through dissolved organic matter. *Nature* **433**, 142–145.

Hubberten U., Lara R. J. and Kattner G. (1995) Refractory organic compounds in polar waters: Relationship between humic substances and amino acids in the Arctic and Antarctic. *J. Mar. Res.* **53**, 137–149.

Kaiser E., Simpson A. J., Dria K. J., Sulzberger B. and Hatcher P. G. (2003) Solid-state and multidimensional solution-state NMR of solid phase extracted and ultrafiltered riverine dissolved organic matter. *Environ. Sci. Technol.* **37**, 2929–2935.

Kim D., Kim Sungjune, Son S., Jung M.-J. and Kim Sunghwan (2019) Application of online liquid chromatography 7 T FT-ICR mass spectrometer equipped with quadrupolar detection for analysis of natural organic matter. *Anal. Chem.* **91**, 7690–7697.

Kim S., Kramer R. W. and Hatcher P. G. (2003) Graphical Method for Analysis of Ultrahigh-Resolution Broadband Mass Spectra of Natural Organic Matter, the Van Krevelen Diagram. *Anal. Chem.* **75**, 5336–5344.

Kimber R. W. L. and Searle P. L. (1970) Pyrolysis gas chromatography of soil organic matter. 1. Introduction and methodology. *Geoderma* **4**, 47–55.

Kramer M. G., Sanderman J., Chadwick O. A., Chorover J. and Vitousek P. M. (2012) Long-term carbon storage through retention of dissolved aromatic acids by reactive particles in soil. *Glob. Change Biol.* **18**, 2594–2605.

Kujawinski E. B. (2002) Electrospray ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (ESI FT-ICR MS): Characterization of complex environmental mixtures. *Environ. Forensics* **3**, 207–216.

Kuznetsova M. and Lee C. (2002) Dissolved free and combined amino acids in nearshore seawater, sea surface microlayers and foams: Influence of extracellular hydrolysis. *Aquat. Sci.* **64**, 252–268.

Lee C., Wakeham S. G. and I. Hedges J. (2000) Composition and flux of particulate amino acids and chloropigments in equatorial Pacific seawater and sediments. *Deep Sea Res. Part Oceanogr. Res. Pap.* **47**, 1535–1568.

Leifeld J. and von Lützow M. (2014) Chemical and microbial activation energies of soil organic matter decomposition. *Biol. Fertil. Soils* **50**, 147–153.

Li H. and Minor E. C. (2015) Dissolved organic matter in Lake Superior: insights into the effects of extraction methods on chemical composition. *Environ. Sci. Process. Impacts* **17**, 1829–1840.

Li X., Guo H., Zheng H., Xiu W., He W. and Ding Q. (2019) Roles of different molecular weights of dissolved organic matter in arsenic enrichment in groundwater: Evidences from ultrafiltration and EEM-PARAFAC. *Appl. Geochem.* **104**, 124–134.

Li X., Liu Z., Chen W., Wang L., He B., Wu K., Gu S., Jiang P., Huang B. and Dai M. (2018) Production and transformation of dissolved and particulate organic matter as indicated by amino acids in the Pearl River estuary, China. *J. Geophys. Res. Biogeosciences* **123**, 3523–3537.

Lin H. and Guo L. (2020) Variations in Colloidal DOM Composition with Molecular Weight within Individual Water Samples as Characterized by Flow Field-Flow Fractionation and EEM-PARAFAC Analysis. *Environ. Sci. Technol.* **54**, 1657–1667.

Lin H., Xu H., Cai Y., Belzile C., Macdonald R. W. and Guo L. (2021) Dynamic changes in size-fractionated dissolved organic matter composition in a seasonally ice-covered Arctic River. *Limnol. Oceanogr.* **66**, 3085–3099.

Liu Z., Sleighter R. L., Zhong J. and Hatcher P. G. (2011) The chemical changes of DOM from black waters to coastal marine waters by HPLC combined with ultrahigh resolution mass spectrometry. *Estuar. Coast. Shelf Sci.* **92**, 205–216.

Liu Z. and Xue J. (2020) The lability and source of particulate organic matter in the northern Gulf of Mexico hypoxic zone. *J. Geophys. Res. Biogeosciences* **125**, e2020JG005653.

Lu K., Gardner W. S. and Liu Z. (2018) Molecular structure characterization of riverine and coastal dissolved organic matter with ion mobility quadrupole time-of-flight LCMS (IM Q-TOF LCMS). *Environ. Sci. Technol.* **52**, 7182–7191.

Lu K., Li X., Chen H. and Liu Z. (2021) Constraints on isomers of dissolved organic matter in aquatic environments: Insights from ion mobility mass spectrometry. *Geochim. Cosmochim. Acta*.

Lu K. and Liu Z. (2019) Molecular level analysis reveals changes in chemical composition of dissolved organic matter from south Texas rivers after high flow events. *Front. Mar. Sci.* **6**.

Mangal V., Shi Y. X. and Guéguen C. (2017) Compositional changes and molecular transformations of dissolved organic matter during the arctic spring floods in the lower Churchill watershed (Northern Manitoba, Canada). *Biogeochemistry* **136**, 151–165.

Mangal V., Stock N. L. and Guéguen C. (2016) Molecular characterization of phytoplankton dissolved organic matter (DOM) and sulfur components using high resolution Orbitrap mass spectrometry. *Anal. Bioanal. Chem.* **408**, 1891–1900.

Minor E. C., Swenson M. M., Mattson B. M. and Oyler A. R. (2014) Structural characterization of dissolved organic matter: a review of current techniques for isolation and analysis. *Environ. Sci. Process. Impacts* **16**, 2064–2079.

Mopper K., Stubbins A., Ritchie J. D., Bialk H. M. and Hatcher P. G. (2007) Advanced instrumental approaches for characterization of marine dissolved organic matter: extraction techniques, mass spectrometry, and nuclear magnetic resonance spectroscopy. *Chem. Rev.* **107**, 419–442.

O'Connor J. A., Lu K., Guo L., Rosenheim B. E. and Liu Z. (2020) Composition and lability of riverine dissolved organic matter: insights from thermal slicing ramped pyrolysis GC–MS, amino acid, and stable isotope analyses. *Org. Geochem.* **149**, 104100.

Ohno T. and Ohno P. E. (2013) Influence of heteroatom pre-selection on the molecular formula assignment of soil organic matter components determined by ultrahigh resolution mass spectrometry. *Anal. Bioanal. Chem.* **405**, 3299–3306.

Ohno T., Sleighter R. L. and 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**, 2497–2504.

Patriarca C., Bergquist J., Sjöberg P. J. R., Tranvik L. and Hawkes J. A. (2018) Online HPLC-ESI-HRMS method for the analysis and comparison of different dissolved organic matter samples. *Environ. Sci. Technol.* **52**, 2091–2099.

Perminova I. V., Dubinenkov I. V., Kononikhin A. S., Konstantinov A. I., Zherebker A. Ya., Andzhushev M. A., Lebedev V. A., Bulygina E., Holmes R. M., Kostyukevich Y. I., Popov I. A. and 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**, 7461–7468.

Peter S., Shen Y., Kaiser K., Benner R. and Durisch-Kaiser E. (2012) Bioavailability and diagenetic state of dissolved organic matter in riparian groundwater. *J. Geophys. Res. Biogeosciences* **117**.

Repeta D. J. (2015) Chemical characterization and cycling of dissolved organic matter. In *Biogeochemistry of Marine Dissolved Organic Matter (Second Edition)* (ed. D. A. H. A. Carlson). Academic Press, Boston. pp. 21–63.

Rogers J. A., Galy V., Kellerman A. M., Chanton J. P., Zimov N. and Spencer R. G. M. (2021) Limited presence of permafrost dissolved organic matter in the Kolyma River, Siberia revealed by ramped oxidation. *J. Geophys. Res. Biogeosciences* **126**, e2020JG005977.

Rosenheim B. E., Day M. B., Domack E., Schrum H., Benthien A. and Hayes J. M. (2008) Antarctic sediment chronology by programmed-temperature pyrolysis: Methodology and data treatment. *Geochem. Geophys. Geosystems* **9**, Q04005.

Saliot A., Ulloa-Guevara A., Viets T. C., de Leeuw J. W., Schenck P. A. and Boon J. J. (1984) The application of pyrolysis-mass spectrometry and pyrolysis-gas chromatography-mass spectrometry to the chemical characterization of suspended matter in the ocean. *Org. Geochem.* **6**, 295–304.

Seeley M. E., Wang Q., Bacosa H., Rosenheim B. E. and Liu Z. (2018) Environmental petroleum pollution analysis using ramped pyrolysis-gas chromatography–mass spectrometry. *Org. Geochem.* **124**, 180–189.

Simjouw J.-P., Minor E. C. and Mopper K. (2005) Isolation and characterization of estuarine dissolved organic matter: Comparison of ultrafiltration and C18 solid-phase extraction techniques. *Mar. Chem.* **96**, 219–235.

Spencer R. G. M., Mann P. J., Dittmar T., Eglinton T. I., McIntyre C., Holmes R. M., Zimov N. and Stubbins A. (2016) Detecting the signature of permafrost thaw in Arctic rivers. *J. Geophys. Res. Oceans*, 2830–2835.

Stolpe B., Guo L., Shiller A. M. and Aiken G. R. (2013) Abundance, size distributions and trace-element binding of organic and iron-rich nanocolloids in Alaskan rivers, as revealed by field-flow fractionation and ICP-MS. *Geochim. Cosmochim. Acta* **105**, 221–239.

Tfaily M. M., Hodgkins S., Podgorski D. C., Chanton J. P. and Cooper W. T. (2012) Comparison of dialysis and solid-phase extraction for isolation and concentration of dissolved organic matter prior to Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Bioanal. Chem.* **404**, 447–457.

Wagner S., Fair J. H., Matt S., Hosen J. D., Raymond P., Saiers J., Shanley J. B., Dittmar T. and Stubbins A. (2019) Molecular hysteresis: Hydrologically driven changes in riverine dissolved organic matter chemistry during a storm event. *J. Geophys. Res. Biogeosciences* **0**.

Walker B. D., Beaupré S. R., Guilderson T. P., Druffel E. R. M. and McCarthy M. D. (2011) Large-volume ultrafiltration for the study of radiocarbon signatures and size vs. age relationships in marine dissolved organic matter. *Geochim. Cosmochim. Acta* **75**, 5187–5202.

Wang Q., Leonce B., Seeley M. E., Adegboyega N. F., Lu K., Hockaday W. C. and Liu Z. (2020) Elucidating the formation pathway of photo-generated asphaltenes from light Louisiana sweet crude oil after exposure to natural sunlight in the Gulf of Mexico. *Org. Geochem.* **150**, 104126.

Whitehead K. (2008) Marine organic geochemistry. In *Chemical Oceanography and the Marine Carbon Cycle* Cambridge University Press. pp. 261–302.

Zark M. and Dittmar T. (2018) Universal molecular structures in natural dissolved organic matter. *Nat. Commun.* **9**, 3178.

Zhou X., Li W., Mabon R. and Broadbelt L. J. (2017) A critical review on hemicellulose pyrolysis. *Energy Technol.* **5**, 52–79.

Zigah P. K., Minor E. C., Abdulla H. A. N., Werne J. P. and Hatcher P. G. (2014) An investigation of size-fractionated organic matter from Lake Superior and a tributary stream using radiocarbon, stable isotopes and NMR. *Geochim. Cosmochim. Acta* **127**, 264–284.

Zigah P. K., Minor E. C., McNichol A. P., Xu L. and Werne J. P. (2017) Constraining the sources and cycling of dissolved organic carbon in a large oligotrophic lake using radiocarbon analyses. *Geochim. Cosmochim. Acta* **208**, 102–118.

Zsolnay A. (1982) A new pyrolysis—mass spectrometry approach to organic marine chemistry using chemical ionization. *J. Anal. Appl. Pyrolysis* **4**, 47–58.

Figure 1

Click here to access/download;Figure;Figrue 1.pdf

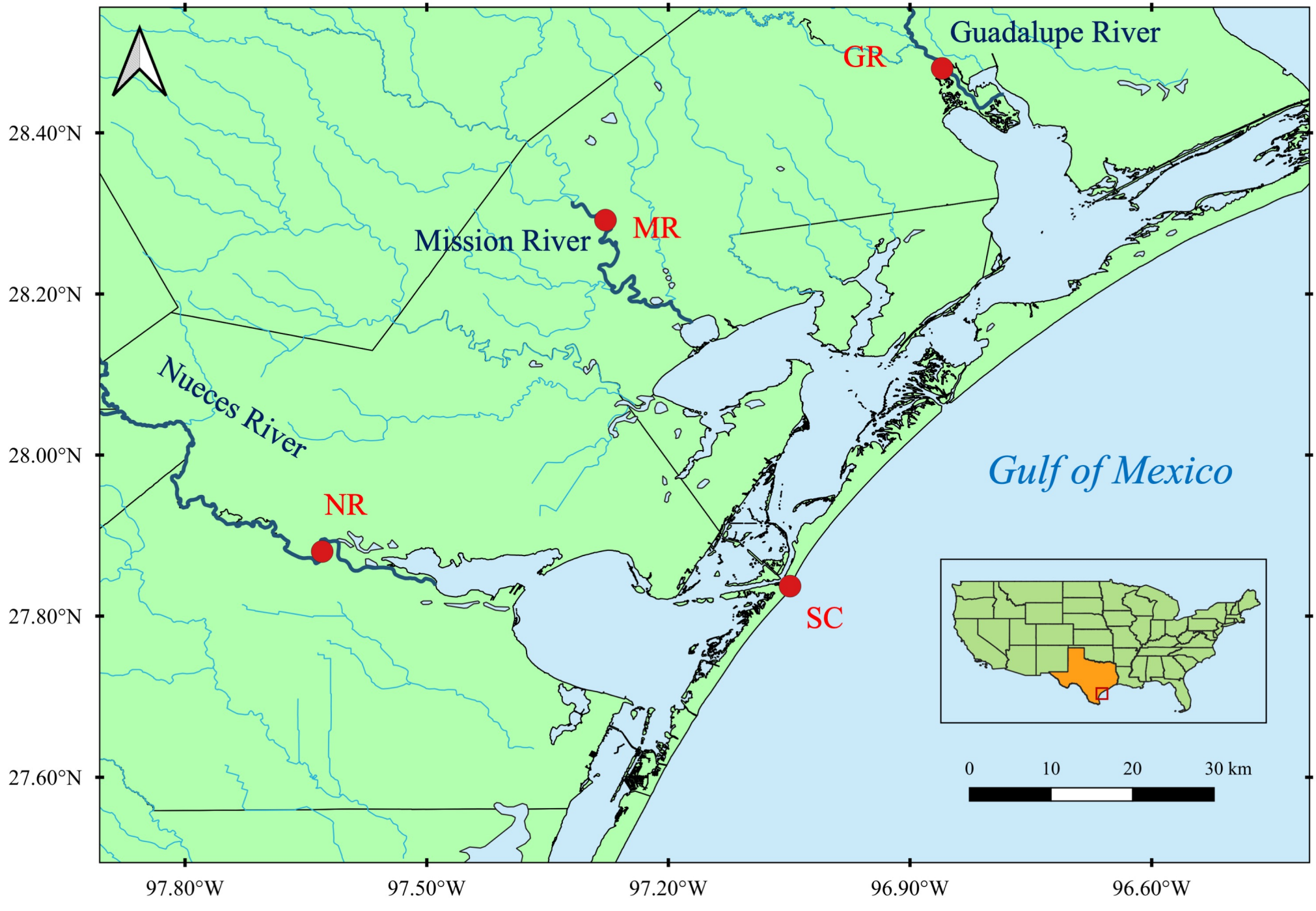


Figure 2

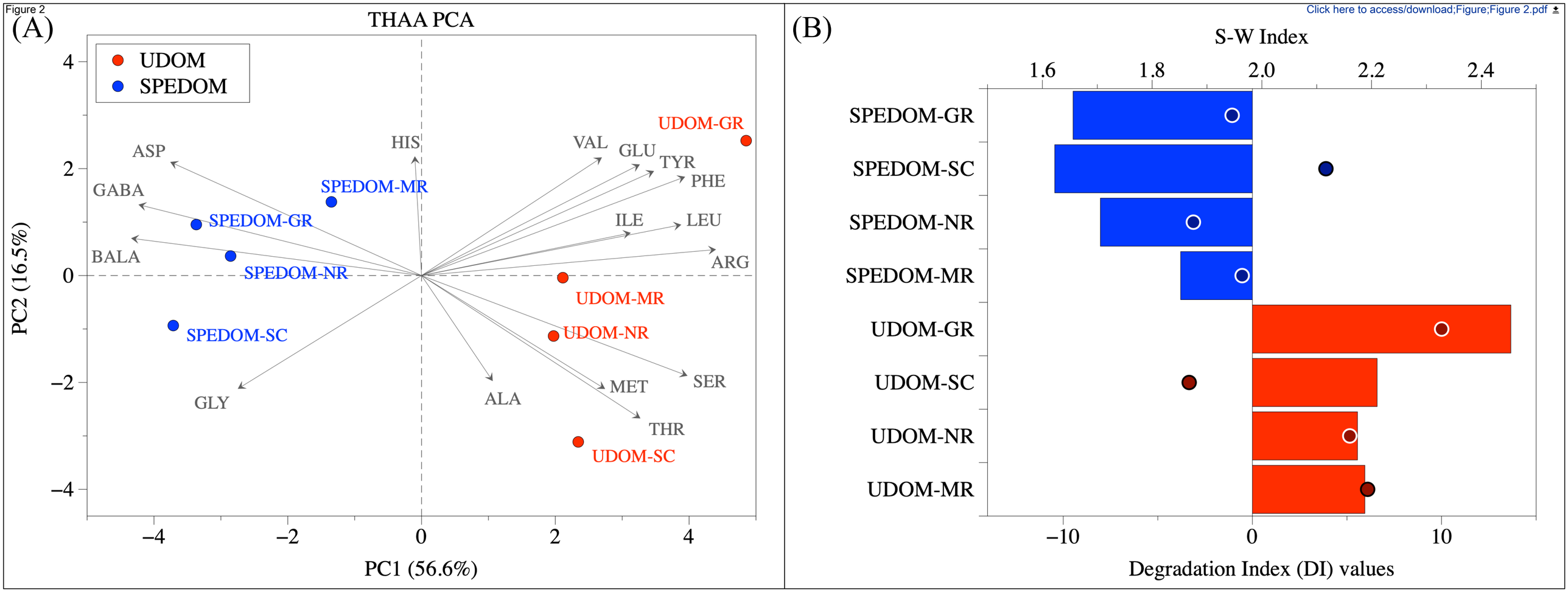


Figure 3

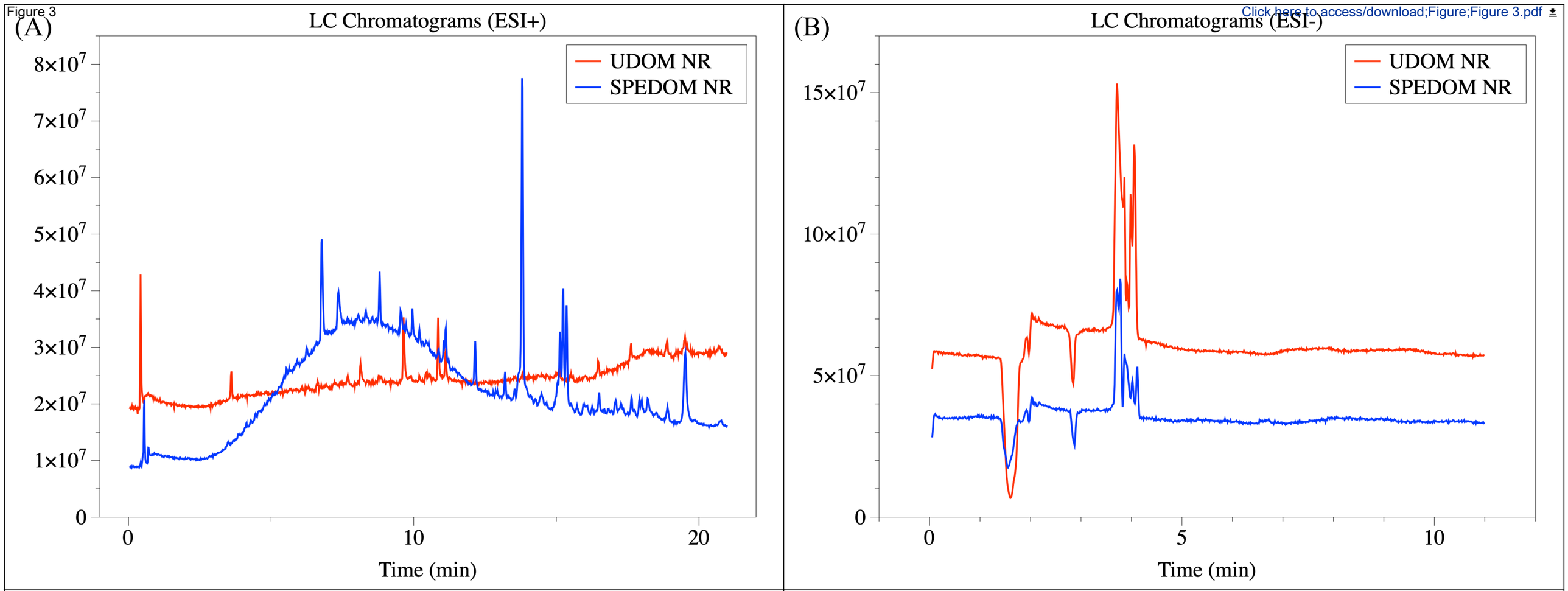
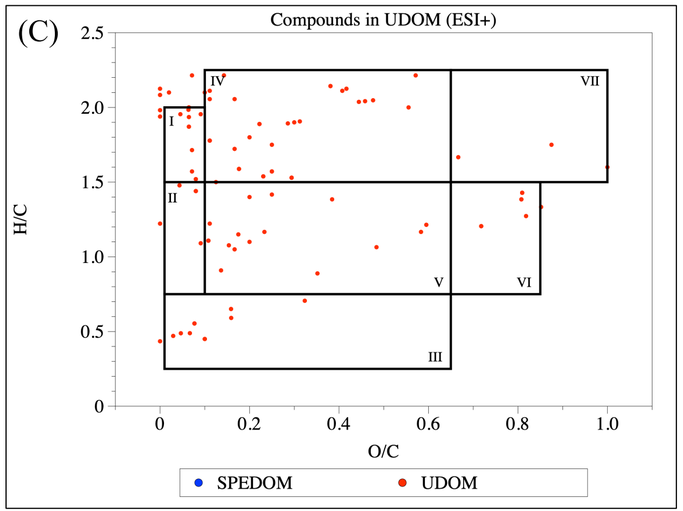
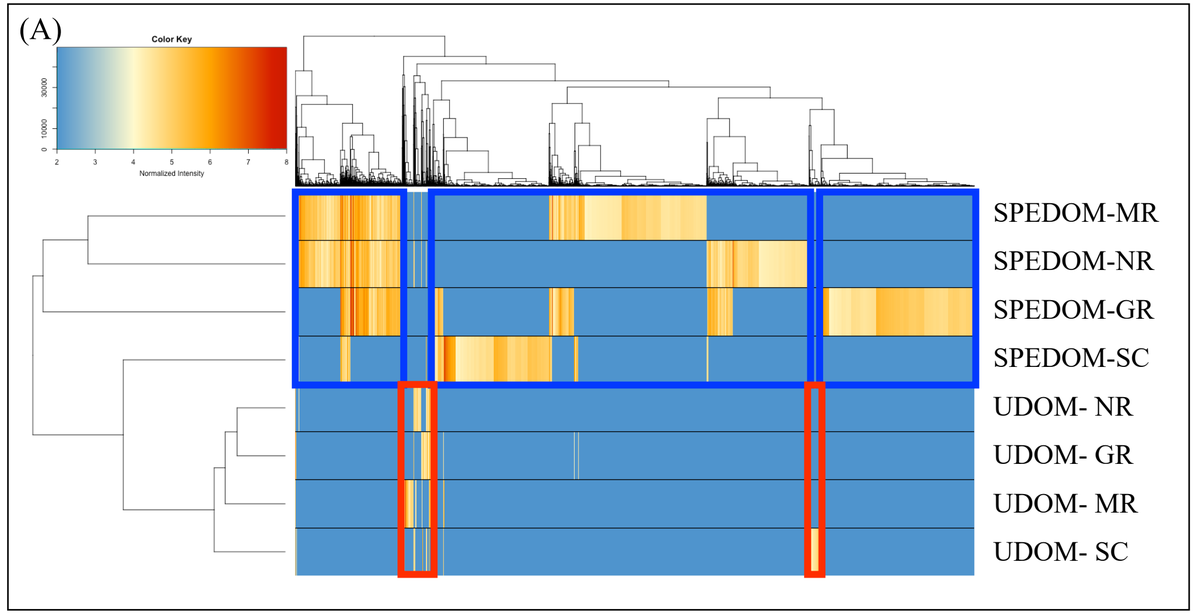
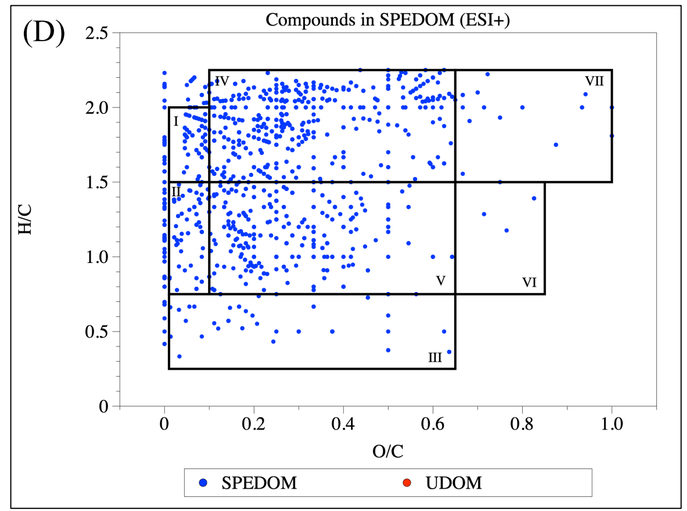


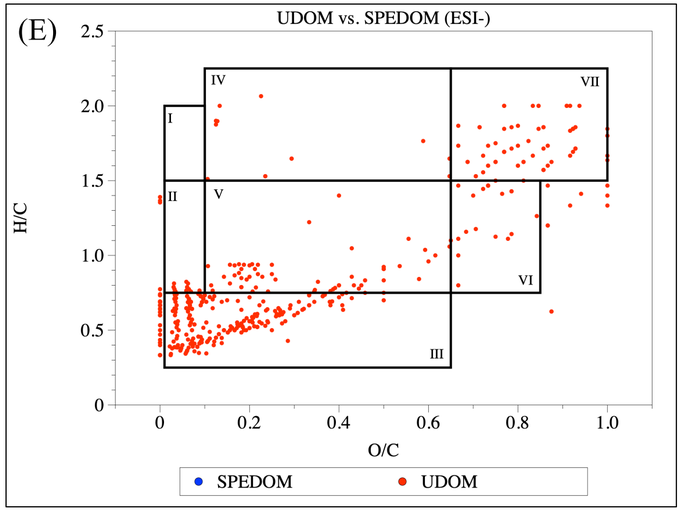
Figure 4

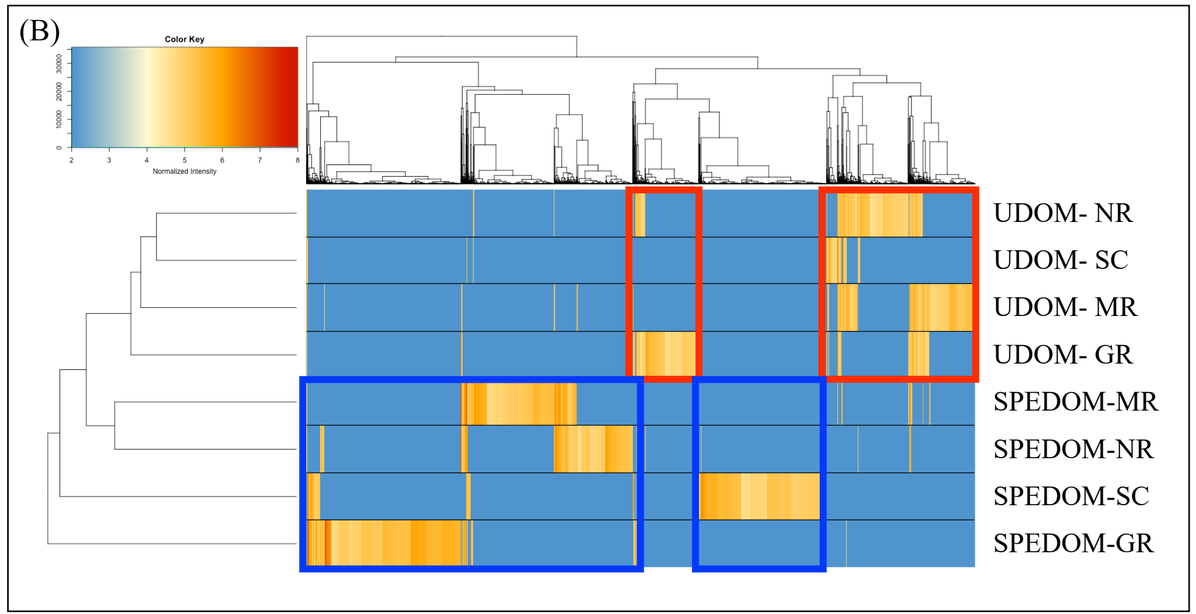
Click here to access/download;Figure;Figure 4.pdf

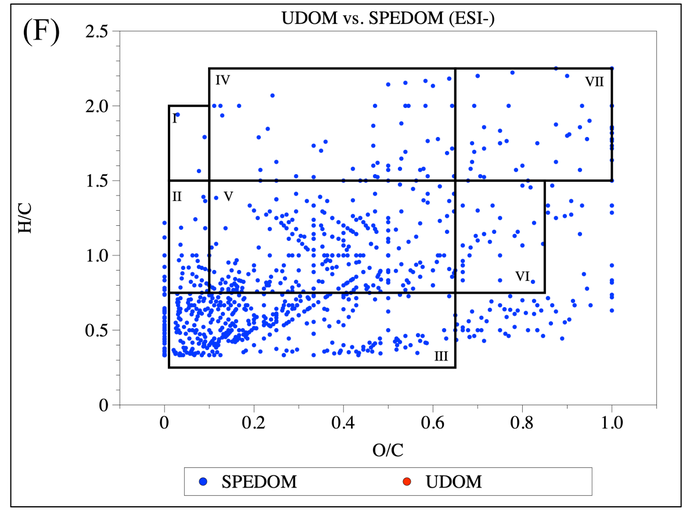












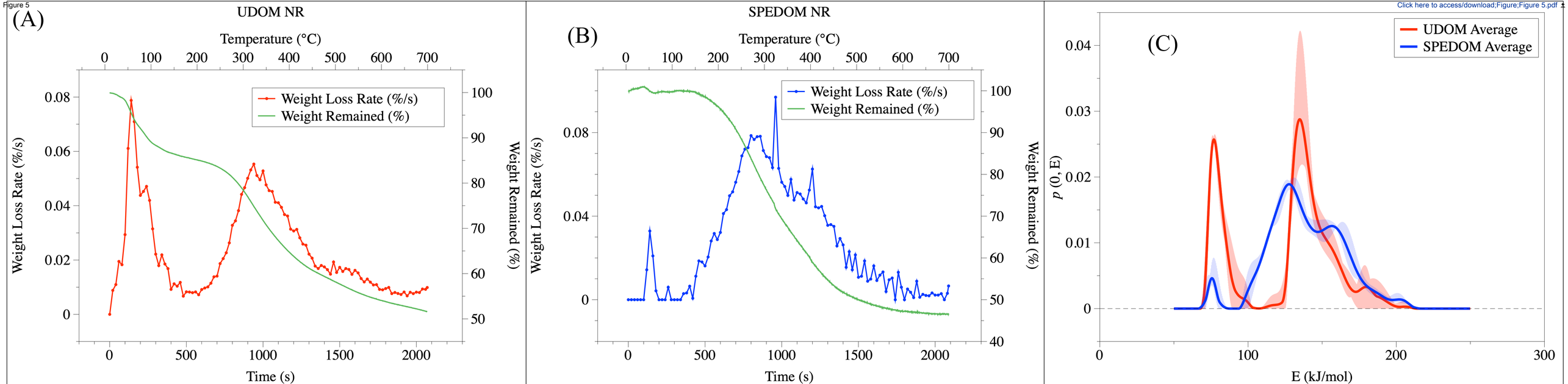


Figure 6

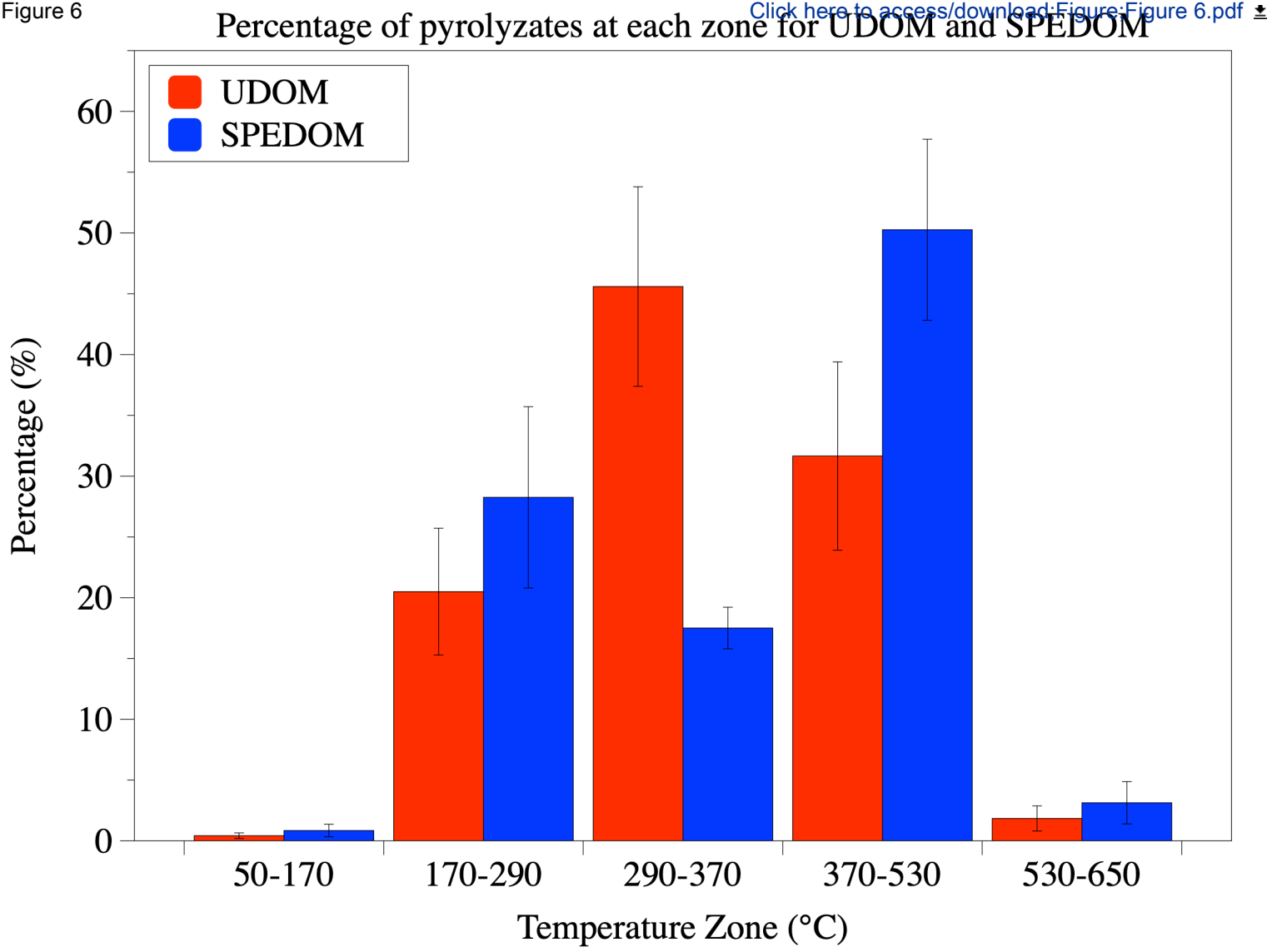


Figure 7

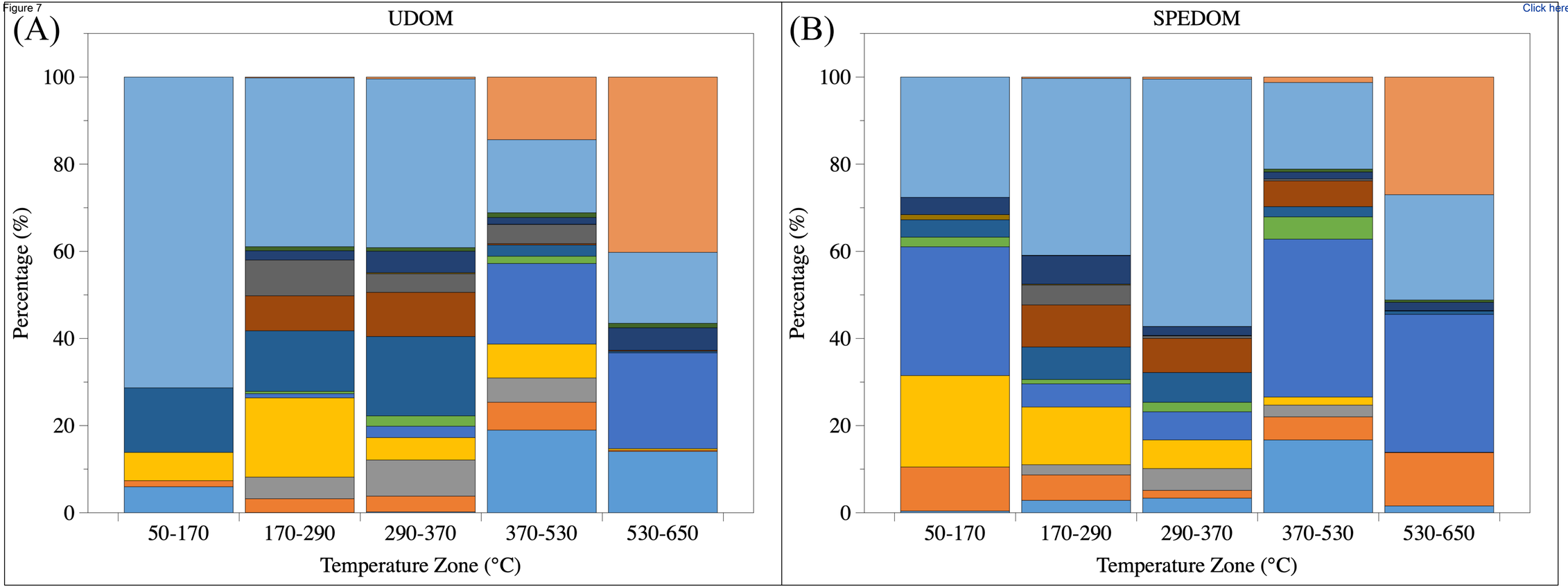
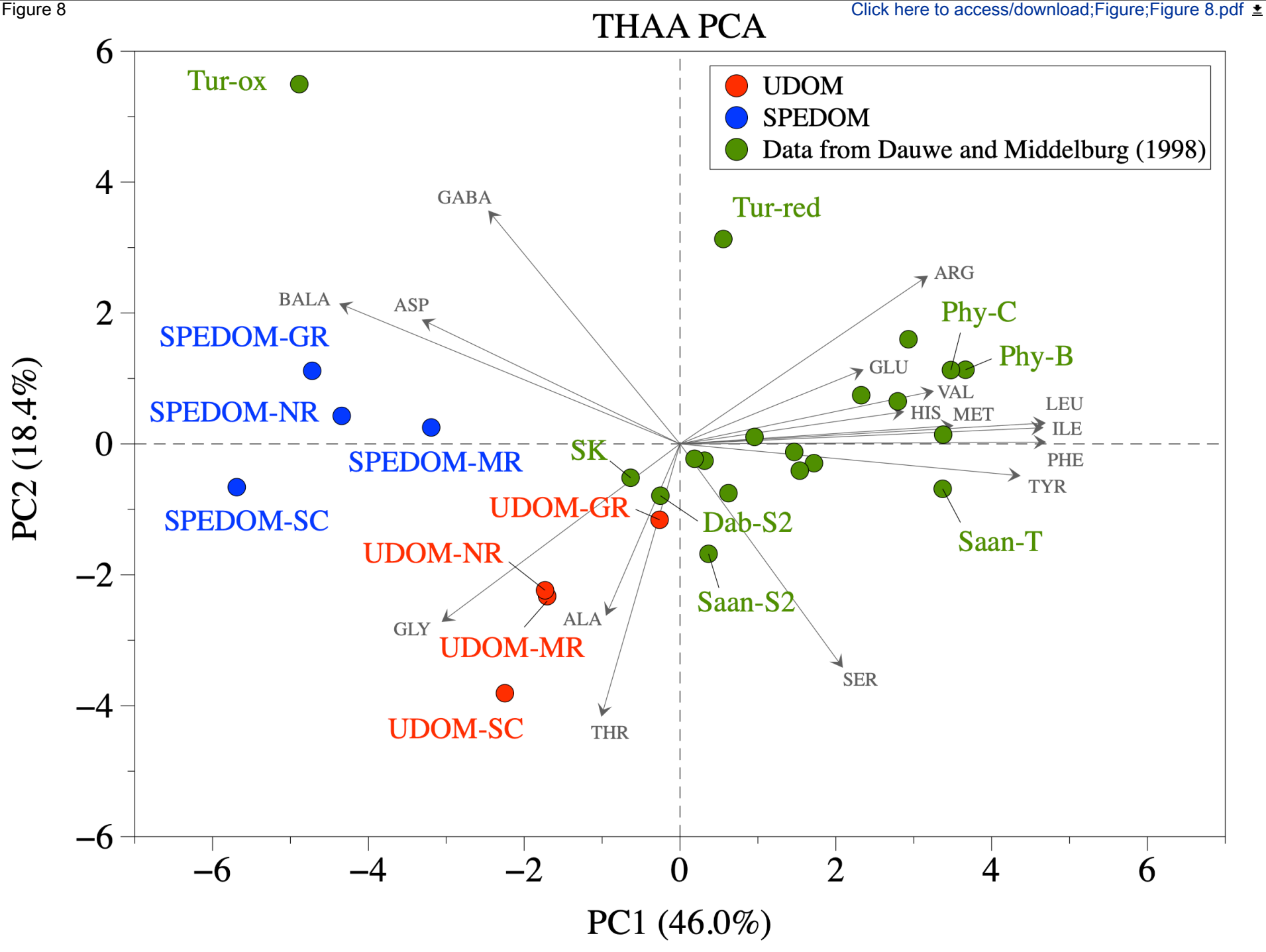


Figure 8



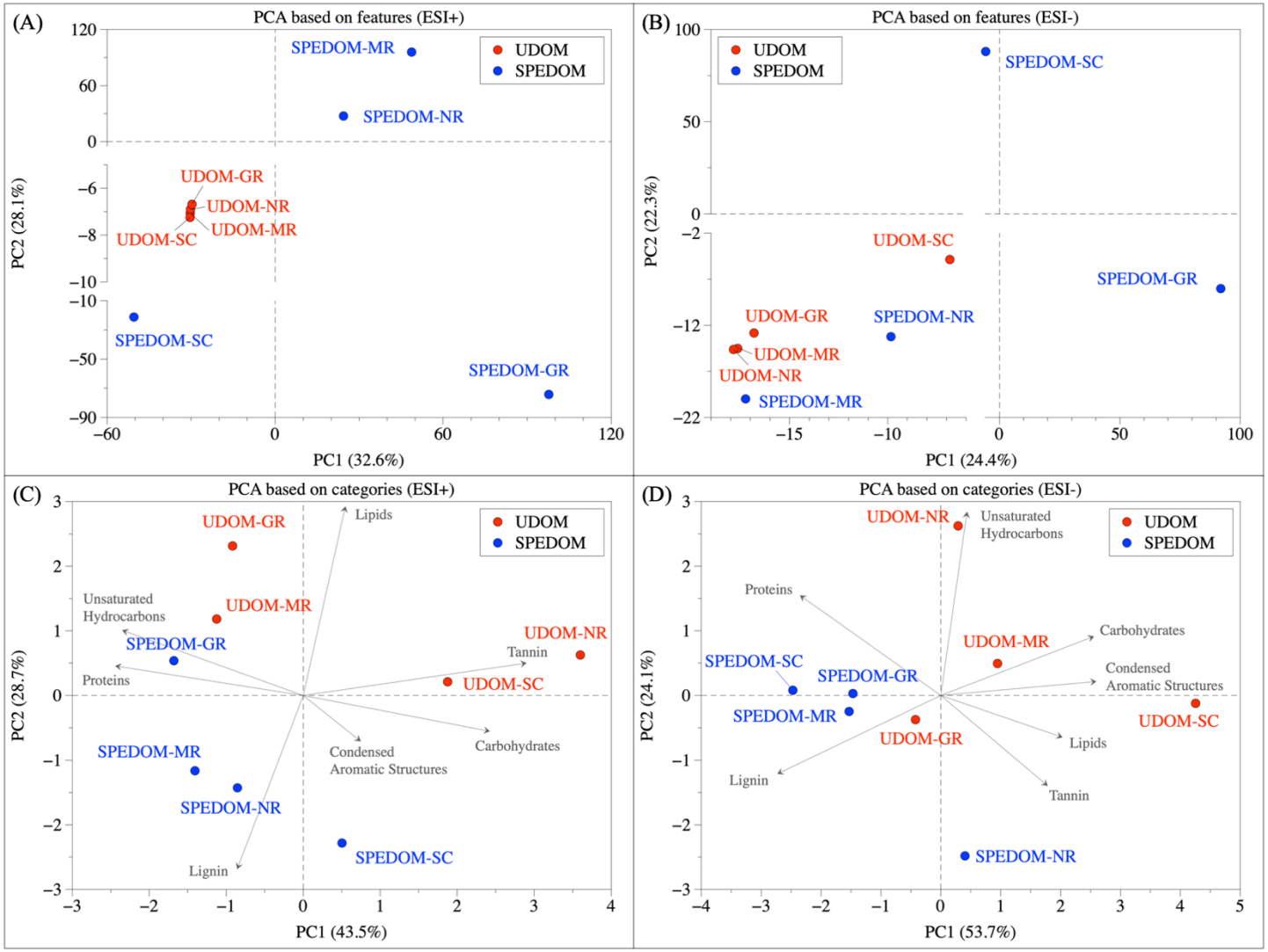


Figure S1. (A) PCA based on the log-transformed relative intensity under ESI+; (B) PCA based on the normalized log-transformed relative intensity under ESI-; (C) PCA based on the category under ESI+; (D) PCA based on the category under ESI-.

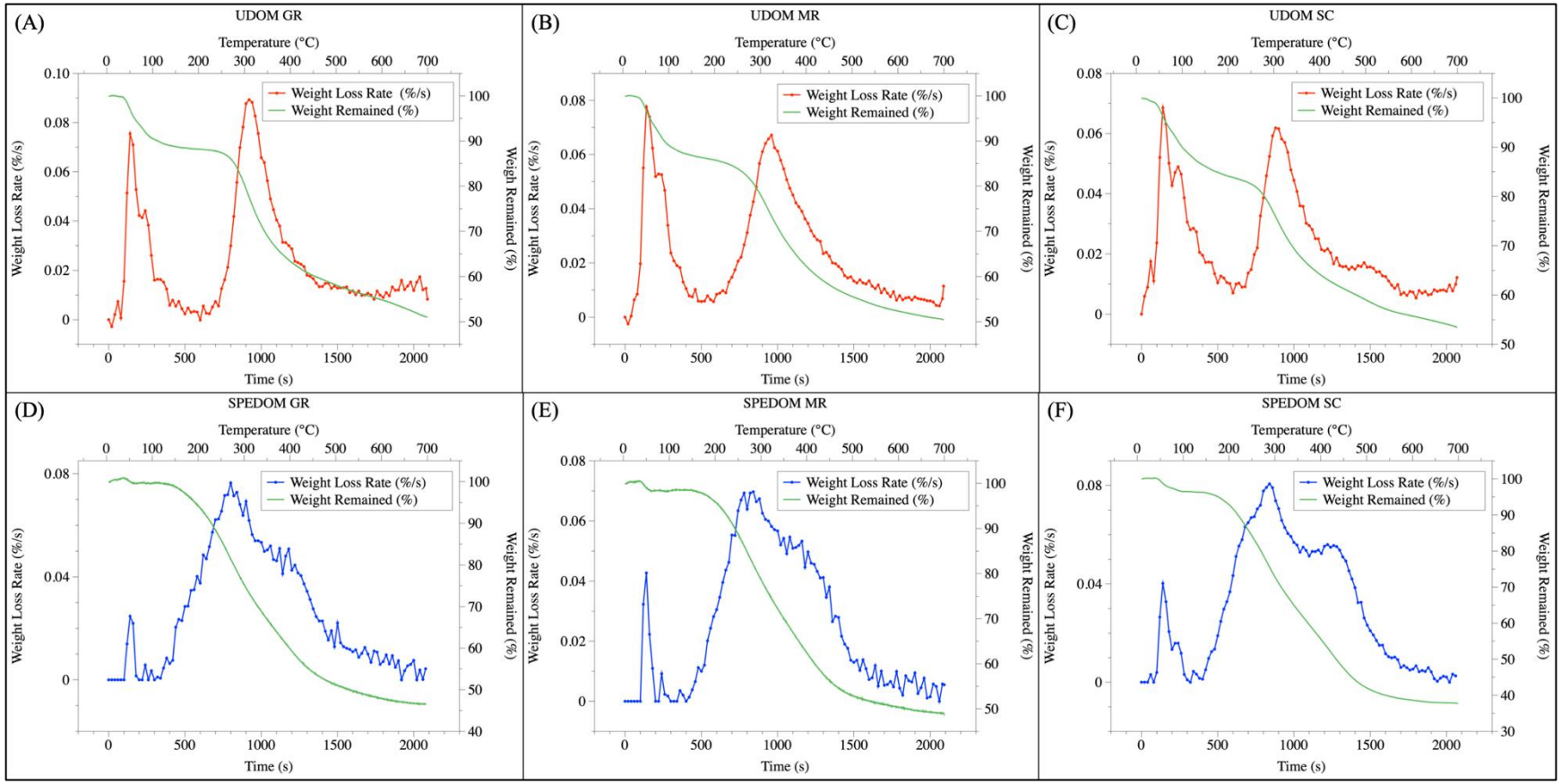


Figure S2. (A) – (C) Thermogram of UDOM collected at GR, MR, and SC, with the remaining weight% shown in green, and the weight loss rate (%·s^{-1}) shown in red; (B) Thermogram of SEPDOM collected at GR, MR, and SC, with the remaining weight% shown in green, and the weight loss rate (%·s^{-1}) shown in blue.

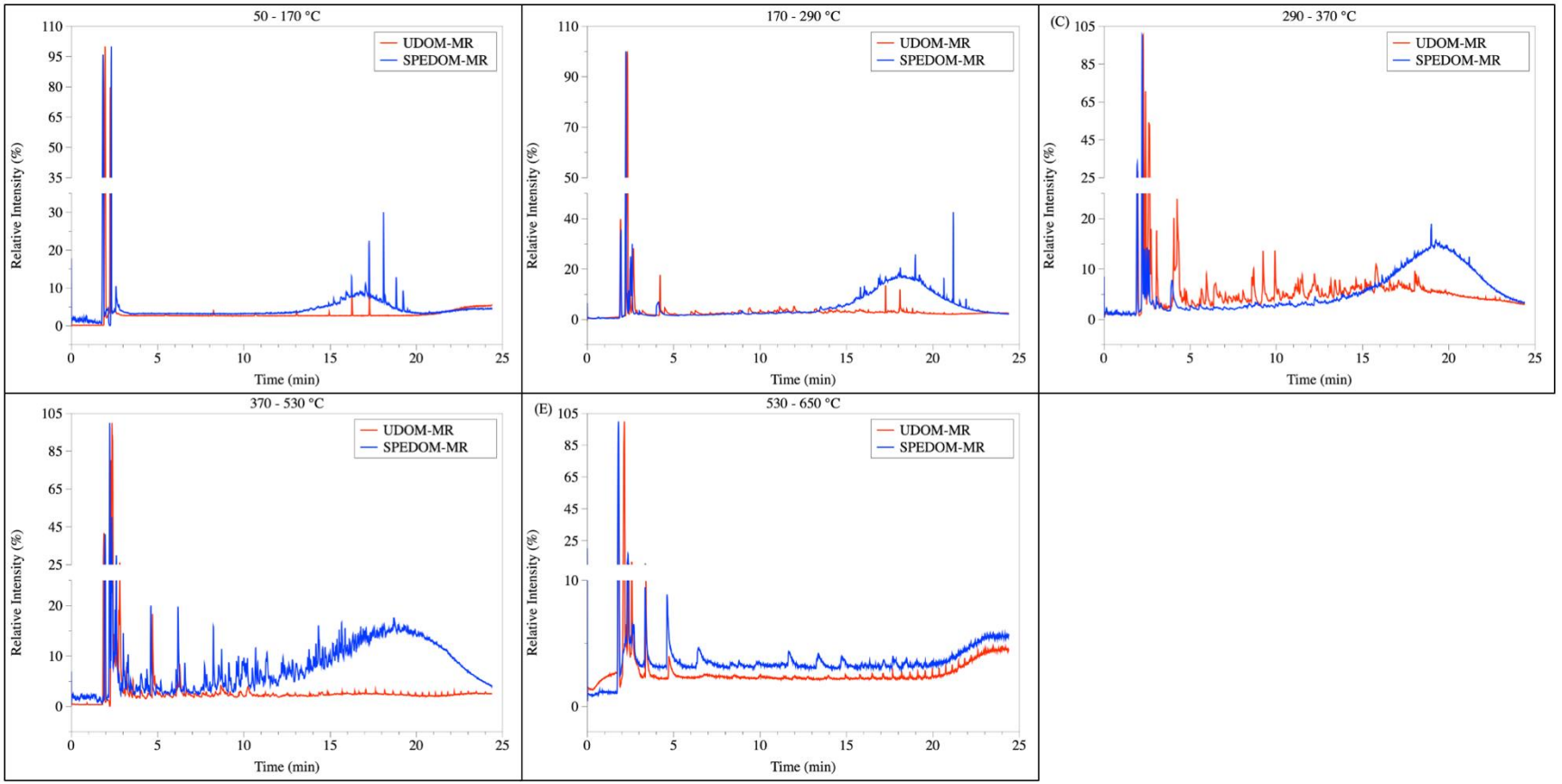


Figure S3. TSRP-GC-MS chromatograms at temperature zone of 50 – 170 °C (A), 170 – 290 °C (B), 290 – 370 °C (C), 370 – 530 °C (D), and 530 – 650 °C (E) for UDOM and SPEDOM samples collected from MR.

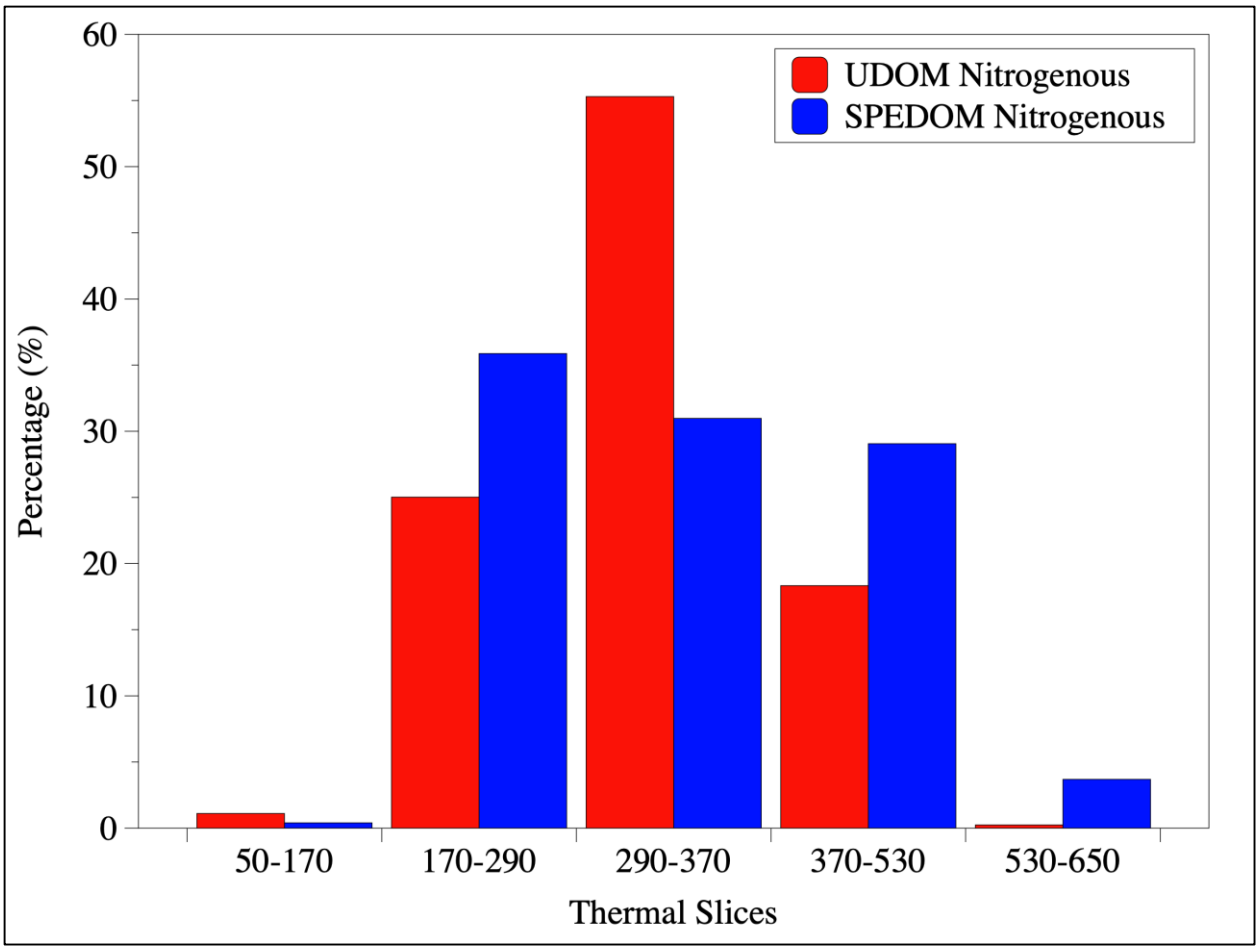


Figure S4. The distribution pattern of nitrogenous structures across different thermal slices between UDOM and SPEDOM.

Table S1. Salinity, DOC concentrations, and C-based recovery rates for each sample.

Samples	Salinity (‰)	DOC (μmol C·L^{-1})	Ultrafiltration Recovery	SPE Recovery
GR	0.4	428.9	24.0%	50.9%
MR	0.1	1008.3	48.7%	55.9%
NR	0.3	300.7	28.9%	48.5%
SC	30.2	600.3	17.6%	46.0%

Table S2. Compound classes of pyrolyzates for MR samples.

Compound Class	MR UDOM					MR SPEDOM				
	50-170	170-290	290-370	370-530	530-650	50-170	170-290	290-370	370-530	530-650
Alkene	1.7	0.1	0.0	28.4	21.6	0.0	4.4	2.6	17.5	0.0
Alkane	2.2	0.0	0.3	16.9	1.2	0.0	0.2	5.7	6.4	0.0
Ketone	0.0	3.7	10.8	3.1	0.0	0.0	2.5	5.4	1.4	0.0
Methanol	26.0	17.7	1.0	0.5	1.4	31.4	15.6	6.7	2.0	0.0
Aromatic	0.0	0.7	8.4	19.1	21.6	32.6	3.6	10.3	50.1	72.8
Phenol	0.0	0.6	6.6	1.8	0.0	8.8	1.6	6.7	7.3	0.0
Heterocyclic (o)	55.8	9.5	13.7	0.0	0.0	1.6	4.8	3.0	1.6	0.0
Acetic	0.0	6.3	4.0	0.0	0.8	0.0	8.0	11.2	1.1	0.0
Aldehyde	0.0	9.4	0.8	17.5	0.0	0.0	0.0	0.8	0.3	0.0
Sterol	0.0	0.0	0.9	0.1	0.0	4.7	0.6	0.1	0.0	0.0
Ester	0.0	3.3	6.3	0.0	0.4	7.2	14.4	2.1	0.3	0.0
Ether	0.0	0.8	0.4	0.2	2.0	0.0	0.2	0.0	0.6	0.0
Nitrogenous	14.3	47.8	46.6	3.2	0.0	13.6	44.1	44.2	11.6	3.3
Nitrile	0.0	0.0	0.0	9.1	50.9	0.0	0.0	1.2	0.0	24.0

Table S3. Compound classes of pyrolyzates for NR samples.

Compound Class	NR UDOM					NR SPEDOM				
	50-170	170-290	290-370	370-530	530-650	50-170	170-290	290-370	370-530	530-650
Alkene	0.0	0.1	0.5	15.8	3.1	1.5	0.2	3.7	9.1	0.0
Alkane	0.0	0.0	0.4	1.0	0.0	33.7	6.3	1.4	3.2	30.1
Ketone	0.0	4.5	9.4	9.9	0.0	0.0	2.1	8.1	3.2	0.0
Methanol	0.0	16.7	2.1	27.4	0.0	40.2	20.5	7.4	2.4	0.1
Aromatic	0.0	2.0	1.2	17.5	32.0	11.7	4.3	5.8	40.1	6.9
Phenol	0.0	1.1	1.5	2.3	0.0	0.0	0.4	0.7	3.2	0.0
Heterocyclic (o)	3.4	4.8	24.4	0.8	1.5	0.0	6.3	2.4	2.9	3.0
Acetic	0.0	4.2	6.0	0.0	0.0	0.0	5.5	7.9	3.2	0.4
Aldehyde	0.0	6.4	0.3	0.1	0.0	0.0	0.5	0.7	1.0	0.0
Sterol	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0
Ester	0.0	2.8	7.6	1.1	4.1	4.8	3.7	2.4	2.0	2.6
Ether	0.0	0.6	0.8	4.0	2.1	0.0	0.0	0.0	1.7	0.4
Nitrogenous	96.6	56.9	45.2	6.5	6.5	8.0	49.9	59.0	27.1	54.7
Nitrile	0.0	0.0	0.5	13.4	50.7	0.0	0.2	0.5	0.8	1.8

Table S4. Compound classes of pyrolyzates for SC samples.

Compound Class	SC UDOM					SC SPEDOM				
	50-170	170-290	290-370	370-530	530-650	50-170	170-290	290-370	370-530	530-650
Alkene	0.0	0.0	0.3	7.9	0.6	0.0	6.0	3.3	31.9	2.0
Alkane	3.3	10.0	8.9	1.1	0.0	0.0	8.4	0.0	1.7	16.9
Ketone	0.0	3.0	4.8	5.5	0.0	0.0	2.6	3.8	2.6	0.0
Methanol	0.0	17.4	5.8	0.3	0.0	0.0	2.3	4.4	1.2	0.2
Aromatic	0.0	1.1	0.3	11.0	34.4	31.7	5.4	8.1	23.6	22.6
Phenol	0.0	0.4	0.5	0.7	0.0	0.0	0.1	1.0	5.1	0.0
Heterocyclic (o)	0.0	13.8	12.2	0.6	0.0	6.3	10.6	15.6	1.6	0.0
Acetic	0.0	5.2	8.3	0.1	0.0	0.0	14.0	6.0	16.9	0.0
Aldehyde	0.0	5.4	4.2	0.0	0.0	0.0	17.1	0.6	0.3	0.0
Sterol	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Ester	0.0	1.9	0.9	1.3	0.6	0.0	3.9	0.7	2.6	0.0
Ether	0.0	1.1	0.9	0.1	0.0	0.0	0.1	0.2	0.0	0.0
Nitrogenous	96.7	40.7	51.9	53.4	5.1	62.0	28.7	56.5	10.1	37.6
Nitrile	0.0	0.0	1.1	17.8	59.3	0.0	0.8	0.0	2.5	20.7

Table S5. Compound classes of pyrolyzates for GR samples.

Compound Class	GR UDOM					GR SPEDOM				
	50-170	170-290	290-370	370-530	530-650	50-170	170-290	290-370	370-530	530-650
Alkene	22.2	0.0	0.0	23.8	30.9	0.0	0.7	3.8	8.3	4.2
Alkane	0.0	2.7	4.8	6.6	0.0	6.7	8.5	0.0	9.9	2.0
Ketone	0.0	8.6	8.2	3.7	0.0	0.0	2.1	2.8	3.7	0.0
Methanol	0.0	20.9	11.7	2.9	0.0	12.3	14.5	7.7	1.8	0.0
Aromatic	0.0	0.0	0.5	26.4	0.0	42.2	7.9	1.6	31.2	24.4
Phenol	0.0	0.0	0.8	1.9	0.0	0.0	1.9	0.3	4.8	0.0
Heterocyclic (o)	0.0	27.6	22.6	9.0	0.0	8.1	8.2	6.4	3.4	0.0
Acetic	0.0	16.4	22.1	1.0	0.0	0.0	11.2	6.3	2.4	0.0
Aldehyde	0.0	11.7	11.8	0.0	0.0	0.0	0.6	0.2	0.1	0.0
Sterol	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.3	0.0	0.0
Ester	0.0	0.5	5.0	3.7	15.6	3.8	4.2	2.8	1.5	5.0
Ether	0.0	1.2	1.1	0.0	0.0	0.0	0.0	0.0	0.4	1.8
Nitrogenous	77.8	9.6	11.2	3.9	53.5	26.8	39.6	67.6	30.7	1.1
Nitrile	0.0	0.8	0.1	17.1	0.0	0.0	0.3	0.1	1.8	61.5