Molecular characterization of dissolved organic matter from subtropical wetlands: A comparative study through the analysis of optical properties, NMR and FTICR/MS.

N. Hertkorn[1], M. Harir[1], K. M. Cawley[2], P. Schmitt-Kopplin[1], R. Jaffé[2*]

[1] Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Research Unit Analytical Biogeochemistry (BGC), Ingolstaedter Landstrasse 1, D-85764 Neuherberg, Germany

[2] Southeast Environmental Research Center, and Department of Chemistry and Biochemistry, Florida International University, 11200 SW 8th Street, Miami, FL 33199, USA

* Corresponding author: R. Jaffé (jaffer@fiu.edu; 305.348.2456)

Abstract: Wetlands provide quintessential ecosystem services such as maintenance of water quality, water supply and biodiversity, among others; however, wetlands are also among the most threatened ecosystems worldwide. Natural dissolved organic matter (DOM) is an abundant and critical component in wetland biogeochemistry. This study describes the first detailed, comparative, molecular characterization of DOM in sub-tropical, pulsed, wetlands, namely the Everglades (USA), the Pantanal (Brazil) and the Okavango Delta (Botswana), using optical properties, high field nuclear magnetic resonance (NMR) and ultrahigh resolution mass spectrometry (FT-ICRMS), and compares compositional features to variations in organic matter sources and flooding characteristics (i.e. differences in hydroperiod). While optical properties showed a high degree of variability within and between the three wetlands, analogies in DOM fluorescence properties were such that an established excitation emission matrix fluorescence parallel factor analysis (EEM-PARAFAC) model for the Everglades was perfectly applicable to the other two wetlands. Area-normalized ^{1}H NMR spectra of selected samples revealed clear distinctions of samples while a pronounced congruence within the three pairs of wetland DOM readily suggested the presence of an individual wetland-specific molecular signature. Within sample pairs (long vs. short hydroperiod sites), internal differences mainly referred to intensity variations (denoting variable abundance) rather than to alterations of NMR resonances

positioning (denoting diversity of molecules). The relative disparity was largest between the Everglades long and short hydroperiod samples, whereas Pantanal and Okavango samples were more alike among themselves. Otherwise, molecular divergence was most obvious in the case of unsaturated protons (δ_H > 5 ppm). 2D NMR spectroscopy for a particular sample revealed a large richness of aliphatic and unsaturated substructures, likely derived from microbial sources such as periphyton in the Everglades. In contrast, the chemical diversity of aromatic wetland DOM likely originates from a combination of higher plant sources, progressive microbial and photochemical oxidation, and contributions from combustion-derived products (e.g. black carbon). FT-ICRMS spectra of both Okavango and Pantanal showed near 57 ± 2% CHO, 8 ± 2% CHOS, 33 ± 2 CHNO, and < 1% CHNOS molecules, whereas those of Everglades samples were markedly enriched in CHOS and CHNOS at the expense of CHO and CHNO compounds. In particular, the Everglades short hydroperiod site showed a large set of aromatic and oxygen-deficient "black sulphur" compounds whereas the long hydroperiods site containted oxygenated sulfur attached to fused-ring polyphenols. The elevated abundance of CHOS compounds for the Everglades samples likely results from higher inputs of agriculture-derived and sea spray derived sulphate. Although wetland DOM samples were found to share many molecular features, each sample was unique in its composition, which reflected specific environmental drivers and/or specific biogeochemical processes.

1) Introduction:

Natural dissolved organic matter (DOM) is a critical component of the global carbon cycle (Battin et al., 2009) and serves as an energy resource fuelling the microbial loop (Amon and Benner, 1996a), acts as a carrier facilitating the mobilization of trace metals and combustion derived products (Yamashita and Jaffé, 2008; Jaffé et al., 2013), and functions as a sun screen for aquatic organisms by limiting light penetration (Blough and Green, 1995; Foden et al., 2008) among other biogeochemical processes. In addition to comprising one of the largest organic matter pools in aquatic environments, DOM is one of the most complex mixtures of OM in natural systems containing millions of organic compounds (Koch et al., 2005; Hertkorn et al., 2008). While the molecular characterization of DOM (Hertkorn et al., 2006 and 2013; Jaffé et al., 2014; Woods et al., 2011 and 2012; Panagiotopolous et al., 2007; Lam et al., 2007;

Aluwihare and Repeta, 1999) has significantly advanced our understanding of its composition and ecological functions, a significant portion of this material remains uncharacterized at the molecular level. Although molecular similarities between bulk DOM from vastly different environments have been reported (Repeta et al., 2002; Perdue and Ritchie, 2003; Jaffé et al., 2012; Hertkorn et al., 2013), the variability in composition (quality) among samples can also be quite significant (Jaffé et al., 2008; Zhang et al., 2014) implying differences in the photo- and bio-reactivity of these materials (Amon and Benner, 1996b). Such compositional differences (or similarities) may have important implications with regards to carbon cycling and ecological functioning of DOM. While the characterization of DOM using targeted substrates such as amino acids (Yamashita and Tanoue, 2003), neutral sugars (Panagiotopolous et al., 2007), lignin phenols (Spencer et al., 2012) and others have actively been pursued, much of the bulk DOM remains uncharacterized (Hedges et al., 2000) and broader spectrum analyses are required. As such, multi-analytical approaches for the advanced molecular characterization of DOM are needed to advance this field (Hertkorn et al., 2013; Jaffé et al., 2012; Minor et al., 2014).

During synthesis of natural dissolved organic matter (DOM), common biosignatures characteristic of the respective sources are progressively attenuated by the combined action of biotic and abiotic reactions. While degradative analysis of DOM intentionally destroys the sample in the beginning to recover a suite of known decomposition products, such as amino acids, carbohydrates, lignin phenols and lipids (which typically account for about 5-30% of the organic carbon, depending on age, environment and diagenesis), non-target molecular-level analyses of DOM attempt to characterize the entire carbon present in DOM by means of information-rich detection methods, such as UV-Vis, fluorescence, NMR spectroscopy and FT-ICRMS. While optical properties have been widely applied for DOM bulk characterizations (Jaffé et al., 2014; Fellman et al., 2010) and allow large sample throughput at low cost, more advanced molecular level characterization techniques such as NMR provide unsurpassed insight into close-range molecular order (Hertkorn et al., 2006 and 2013; Lam et al., 2007; McCaul et al., 2011; Woods et al., 2011 and 2012; Zhang et al., 2014; Bell et al., 2015), while FT-ICRMS provides depiction of the compositional space with exceptional resolution (Kujawinski, 2002; Sleighter and Hatcher, 2007; D'Andrilli et al., 2010; Hertkorn et al., 2008; Hertkorn et al., 2013; Kaiser et al., 2003; Minor et al., 2014; Koch et al., 2005 and 2007). The combination of such techniques in the assessment of DOM dynamics has become more frequent (e.g. Tfaily et al.,

2015; Jaffé et al., 2012), and furthermore, the association between optical properties and the molecular characteristics of DOM have recently become an active research endeavour in an attempt to better link these parameters (Stubbins et al., 2014; Kellerman et al., 2015; Wagner et al., 2015). As such, this work should provide further advances in this field.

While significant efforts have been devoted to the detailed characterization of DOM in oceanic, lacustrine and riverine environments (Hertkorn et al., 2013; Kujawinski et al., 2009; Einsiedl et al., 2007; Minor et al., 2012; Jaffé et al., 2012), still little is known about its molecular features in large freshwater wetlands, environments that are critically threatened by anthropogenic influences such as pollution and drainage for flood control, agricultural and urban development. Organic matter dynamics in large wetlands are particularly complex (e.g. Chen et al., 2013; Yamashita et al., 2010; Cawley et al., 2012) due to a high variability in spatial and temporal organic matter sources, concentrations, and diagenetic transformations. These variations are to a large extent driven by interplay between complex hydrological and primary productivity patterns. In this study, DOM samples from three of the largest and most important sub-tropical, pulsed wetlands, the Everglades (USA), the Pantanal (Brazil) and the Okavango Delta (Botswana), were collected and analysed on a comparative basis using optical properties, including EEM-PARAFAC, high field ^{1}H NMR, and FT-ICRMS in order to assess similarities and differences in DOM composition and molecular structure in such vital ecosystems.

2) Experimental:

2a) Site descriptions, sample collection and analysis: The Everglades, Okavango Delta and Pantanal are three of the largest sub-tropical, pulsed, freshwater wetlands in the world and represent a wealth of biodiversity (Junk et al., 2006a). The Everglades ecosystem is a large (610,483 ha) subtropical wetland located in southern Florida, USA. Annually, the southern section of the system, namely Everglades National Park receives *ca.* 120 cm of precipitation (50-year averages from 1962-2012) with 21 cm falling during the dry season (December to April) and 99 cm falling during the wet season (May to November) (Southeast Regional Climate Center, *http://www.sercc.com*). The freshwater area of the Everglades consists primarily of grassy marshes dominated by sawgrass (*Cladium jamaciensis*) with some small stands of trees on higher ground. The freshwater marshes drain through two main slough areas, namely the peat soil dominated Shark River Slough and the less extensive, marl-soil based Taylor Slough, which

are characterized by longer and shorted hydroperiods (time and depth of inundation), respectively.

The Okavango Delta is a large wetland located in semi-arid NW Botswana and is subject to an annual flood event generated by water of the Okavango River flowing south from the highlands of Angola. During the flood event, the inundated area in the Delta expands in size from the annual minimum of 3,500-6,000 km^2 to the annual maximum of 9,000-13,000 km^2 (Gieske, 1997; McCarthy et al., 2003). About 88 % of inflowing water leaves the wetland through evaporation (Wolski et al., 2006). Flood water moves in the Okavango Delta as a combination of channel and floodplain flows. Several zones featuring differences in hydroperiod due to the seasonality of inundation are categorized as the panhandle, permanent swamp, seasonal floodplains, and occasional floodplains (Gumbricht et al., 2004; Cawley et al., 2012). The permanent swamp is characterized by extensive peat development and dominated by *Phragmites australis* and *C. papyrus* (Ellery et al., 2003; Mladenov et al., 2007). The seasonal floodplains are less peat rich and support mostly emergent sedges and aquatic macrophytes, while the occasional floodplains, characterized by the shortest hydroperiod are dominated by aquatic grasses.

The Pantanal is a large inland wetland of ca. 160,000 km^2, located mostly in SE Brazil, but also extends into Bolivia and Paraguay (Junk and Cunha, 2005; Junk et al., 2006b). The regional geology (depression) features natural levees along stream channels, and thus the wetland is comprised of a labyrinth of large river channels, small streams, canals, and lagoons. Climate conditions lead to clear wet and dry seasons creating a monomodal flood pulse system. The wetland discharges about 80% of its water to the Paraguay River in the southern section of the system. The climate is tropical to sub-tropical with a large number of habitats including savannas and dry forests, leading to broad species diversity. While still mostly pristine, the expansion of cattle ranching surrounding the protected national park and hydrological modifications in the greater watershed have been suggested as potential threats to this ecosystem (Junk and Cunha, 2005).

Surface water grab samples for the three above-described wetlands were collected in pre-cleaned, brown plastic bottles (60 ml for DOC and optical properties; 2 L for solid phase extracts ; SPE-DOM), placed on ice and filtered through GFF (0.7 μm nominal pore size), pre-combusted glass fiber filters within 6 hours after collection. For EEM-PARAFAC comparisons, multiple

156 samples collected monthly over several years for the FCE (n = 858; Chen et al., 2013), samples
157 collected along a trans-Okavango gradient (n = 38; Cawley et al., 2012), and samples collected in
158 different sub-environments of the Pantanal wetland (n = 22; rivers, lagoons, marshes;
159 unpublished) were used to assess differences and similarities in the fluorescence character of the
160 DOM. Sampling for SPE-DOM was performed during summer 2011 for the Florida coastal
161 Everglades (FCE) and during the summer 2010 for the Pantanal (PAN) and the Okavango Delta
162 (OKA) as part of on-going research programs. Only two SPE-DOM samples from each wetland
163 were selected for detailed NMR and FTICRMS analyses, and consisted of one sample
164 characteristic for long hydroperiod (-L) and one for short hydroperiod (-S) environments for each
165 wetland respectively. . For the Florida Coastal Everglades (FCE), samples were collected from
166 the freshwater marsh, peat-soil dominated Shark River Slough (FCE-L) and the marl-soil
167 dominated Taylor Slough (FCE-S), from the Okavango Delta (OKA) seasonal floodplain (OKA-
168 L) and occasional floodplain (OKA-S) along the Boro River (Cawley et al., 2012), and the
169 Paraguay River (PAN-L) and a wetland channel in Pantanal National Park (PAN-S; Chacra de
170 Solange) for the Pantanal (PAN). Representative sample selection for long and short hydroperiod
171 sites was based on previous reports for the FCE and OKA (Chen et al., 2013; Cawley et al.,
172 2012), and advise from local wetlands scientists for the PAN (C. Nunes da Cunha personal
173 communication). The filtered samples were subjected to SPE isolation (Dittmar et al., 2008).
174 Briefly, samples were acidified to a pH 2 using concentrated HCl. DOM in the acidified samples
175 was extracted using PPL (Varian Bond Elut) cartridges and eluted with methanol (Optima,
176 Fisher). The isolated SPE-DOM extracts (referred from here on as DOM for the NMR and FT-
177 ICRMS data) were stored in pre-combusted glass vials and kept in a freezer until analyzed. Milli-
178 Q water was used as a procedural blank and no contamination was observed. DOC
179 measurements were made within three weeks of sample collection at the Southeast
180 Environmental Research Center's water quality lab at Florida International University with a
181 Shimadzu TOC-V CSH TOC analyzer using a high temperature combustion method.

182 *2b) Optical properties analyses:* UV-Vis absorbance scans for filtered samples were collected on
183 a Varian Cary 50 Bio spectrophotometer and collected over a range of 200 nm to 800 nm in a 1-
184 cm quartz cuvette. The optical proxy for molecular weight (slope ratio; S_R) and the fluorescence
185 index (FI) were determined as described in the literature (Helms et al., 2008; McKnight et al.,

2001 respectively), where the S_R value is inversely proportional to the DOM molecular weight, and FI values determined by the ratio of 470/520 nm emission at 370 excitation (Jaffé et al., 2008) can range between 1.4 and 1.9 for soil/terrestrial higher plant and microbial DOM sources, respectively. A blank scan (Milli-Q water) was subtracted from each sample spectrum and spectra were baseline normalized using the average absorbance between 700-800 nm. The absorbance at 254 nm (A_{254}) was also determined and normalized to DOC to obtain standard UV absorbance values ($SUVA_{254}$; Weishaar et al., 2003). Samples were analyzed for fluorescence within two weeks of collection. Fluorescence EEMs were collected on a Horiba Jobin Yvon SPEX Fluoromax-3 spectrofluorometer using the methods of Maie et al. (2006) and Yamashita et al. (2010). Briefly, EEMs were collected over an excitation wavelength (λ_{ex}) range of 240 – 455 nm with an increment of 5 nm and an emission range of λ_{ex} + 10 nm to λ_{ex} + 250 nm with an increment of 2 nm in a 1 cm quartz cuvette. The excitation and emission slit widths were set to 5.7 nm and 2 nm, respectively. Fluorescence scans were collected in signal/reference ratio mode with an integration time of 0.25 s and reported in quinine sulfate units (QSU). EEMs were corrected for instruments optics and inner-filter effects according to Ohno (2002) and Raman normalized and blank subtracted using Matlab v2009a software. EEMs were modeled using Matlab v2009a and fit to an eight component PARAFAC model described in Chen et al. (2010) and Yamashita et al. (2010) that was comprised of FCE samples only.

2c) Nuclear magnetic resonance spectroscopy (NMR): ^{1}H NMR detected spectra of methanolic DOM extracts were acquired with a Bruker Avance NMR spectrometer at 500.13 (1D NMR only) / 800.13 MHz (B_0 = 11.7 / 18.7 T) at 283 K from a few mg of solid obtained by evaporation of original methanol-h_4 solution, dissolved in approx. 130 µL CD_3OD (Merck. 99.95% ^{2}H) solution with a 5 mm z-gradient ^{1}H / ^{13}C / ^{15}N / ^{31}P QCI cryogenic probe (90° excitation pulses: ^{13}C ~ ^{1}H ~ 10 µs) in sealed 2.5 mm Bruker MATCH tubes. ^{13}C NMR spectra were acquired with a Bruker Avance NMR spectrometer at 500.13 / 800.13 MHz (B_0 = 11.7 / 18.7 T) at 283 K from a few mg of solid obtained by evaporation of original methanol-h_4 solution (1 s acquisition time, 14 or 19 s relaxation delay; Table S3). 1D ^{1}H NMR spectra were recorded with a spin-echo sequence (10 µs delay) to allow for high-Q probe ringdown, and classical presaturation to attenuate residual water present "*noesypr1d"*, typically 512-2048 scans (5 s acquisition time, 5 s relaxation delay, 1 ms mixing time; 1 Hz exponential line broadening).

A phase sensitive, gradient enhanced TOCSY NMR spectrum with solvent suppression (*dipsi2etgpsi19*) was acquired for an acquisition time of 1 s, a mixing time of 70 ms, and a relaxation delay of 3 s. The one bond coupling constant ^{1}J(CH) used in 2D ^{1}H,^{13}C DEPT-HSQC spectra (*hsqcedetgpsisp2.2*) was set to 145 Hz; other conditions: ^{13}C 90 deg decoupling pulse, GARP (70 µs); 50 kHz WURST 180 degree ^{13}C inversion pulse (Wideband, Uniform, Rate, and Smooth Truncation; 1.2 ms); F2 (^{1}H): spectral width of 5981 Hz (11.96 ppm); 1.25 s relaxation delay; F1 (^{13}C): SW = 17607 Hz (140 ppm). HSQC-derived NMR spectra were computed to a 4096 × 512 matrix. Gradient (1 ms length, 450 µs recovery) and sensitivity enhanced sequences were used for all 2D NMR spectra. Absolute value JRES/COSY and phase sensitive echo-antiecho TOCSY spectra (with solvent suppression: *jresgpprqf, cosygpph19, dipsi2etgpsi19*) used a spectral width of 9615.4 Hz [JRES (F1) = 50 Hz] and were computed to a 16384 × 2048 matrix [JRES (F1) = 128]. Similarity of ^{1}H NMR spectra was computed from 0.01 ppm section integrals in the range δ_H = 0.5 – 9.5 ppm, with exclusion of methanol and residual water (Bruker AMIX software, version 3.9.4.) with Hierarchical Cluster Explorer (HCE); similarity versus distance metrics used Pearson correlation coefficients; minimum similarity values are provided in Fig. S2A. Other NMR acquisition conditions are given in Tab. S3.

2d) FTICR mass spectrometry: Ultrahigh-resolution Fourier transform ion cyclotron mass spectra were acquired using a 12 T Bruker Solarix mass spectrometer (Bruker Daltonics, Bremen, Germany) fitted with an electrospray ionization source in negative mode. Diluted SPE-DOM (5 µg/mL in methanol) were injected into the electrospray source using a micro-liter pump at a flow rate of 120 µL/h with a nebulizer gas pressure of 138 kPa and a drying gas pressure of 103 kPa. A source heater temperature of 200°C was maintained to ensure rapid desolvatation in the ionized droplets. Spectra were first externally calibrated on clusters of arginine in MeOH (0.57 µmol/L) and internal calibration was systematically done in the presence of natural organic matter reaching accuracy values lower than 500 ppb. The spectra were acquired with a time domain of 4 megawords and 1000 scans were accumulated for each spectrum. Calculation of elemental formulas for each peak was done in a batch mode by an in-house written software tool. The generated formulae were validated by setting sensible chemical constraints [N rule, O/C ratio ≤ 1, H/C ratio ≤ 2n + 2 (C_nH_{2n+2}). Element counts: C ≤ 100, H ≤ 200, O ≤ 80, N ≤ 3, S ≤ 2, P ≤ 1 and mass accuracy window (set at ± 0.5 ppm)]. Final formulae were generated and

categorized into groups containing CHO, CHNO, CHOS or CHNOS molecular compositions which were used to reconstruct the group-selective mass spectra (Schmitt-Kopplin et al., 2010). The computed average values for H, C, N, O and S (atom %) and the H/C and O/C ratios were based upon intensity-weighted averages of mass peaks with assigned molecular formulae, which comprised ~50% of observed mass peaks.

3) Results and Discussion:

3a) Optical properties: Optical properties consisting of A_{254}, $SUVA_{254}$, FI, S_R and EEM-PARAFAC for the six samples for SPE-DOM analysis are presented in Table 1. Although large differences in DOC and A_{254} were observed for the different samples (DOC range of 5.8 to 28.6 ppm; A_{254} from 0.202 to 0.844) some of the qualitative optical parameters such as the S_R values (range 0.91 to 0.98) and the FI values (range 1.30 to 1.44) all fell into a relatively narrow range. In contrast, the $SUVA_{254}$ values covered a larger range from 2.72 to 5.11. A linear correlation was observed between the DOC and the A_{254} (r^2=0.95).

S_R and FI values were quite similar for all sample pairs and among samples, suggesting that the molecular weight distribution and the soil/higher plant vs. microbial contributions were quite similar among these samples, or that the mineralization of wetland DOM leads to similar compositional features for systems with different organic matter sources. Detailed molecular characterizations of DOM in headwater streams from different climatic regions (biomes) have been reported to exhibit remarkably similar bulk characteristics, although site-specific features were also identified in each case (Jaffé et al., 2012). However, in the case of the $SUVA_{254}$, some clear compositional variations between different wetland DOM became apparent, where the samples from the more strongly soil-OM (or peat) influenced, long hydroperiod sites, featured higher $SUVA_{254}$ values compared to those with larger microbial and emergent/aquatic plant influence, short hydroperiod sites. Indeed, $SUVA_{254}$ showed higher values for the peat-based FCE-L compared to the marl-based FCE-S, the Paraguay River sample PAN-L compared to the wetland channel PAN-S, and in the seasonally flooded Boro River floodplain OKA-L compared to the occasional floodplain OKA-S. Although the differences in FI for the PAN samples were not significant, FI values for the other sites were expectedly inversely correlated to $SUVA_{254}$ (Jaffé et al., 2008).

The application of the FCE PARAFAC model to the OKA and PAN samples resulted in an excellent fit leaving no significant residues and was properly validated. The application of the FCE PARAFAC model to assess fluorescence characteristics of DOM from other wetlands was previously reported for the Okavango Delta (Cawley et al., 2012). In addition, the distribution of the EEM-PARAFAC components was also surprisingly similar among the six samples with C1 being dominant, followed by C3 > C5 > C4 > C6 and C7 and with C2 and C8 showing the lowest relative abundance. This trend is consistent with previous reports for the greater Everglades ecosystem (Yamashita et al., 2010; Chen et al., 2013). EEM-PARAFAC results for the three wetlands are shown in Figure 1. The data for FCE are presented as two separate sub-groups representing Everglades National Park (ENP) sites and the Water Conservation Area 2 (WCA2), an area located north of the ENP boundary where water resources are heavily managed and agricultural runoff is significant (Yamashita et al., 2010). In general terms, no significant differences were observed in the PARAFAC component distributions between FCE, OKA and PAN, and the only difference of significance was the relative abundance of the C2 PARAFAC component which was higher in the WCA2 compared to all other study regions. In agreement with the above, comparing the EEM-PARAFAC distributions between and among the six stations, and amongst the larger datasets of collected surface water samples (Table 1; Figure 1), no statistically significant differences were observed, although the range in values was large. Component C2 has been suggested to be photo-chemically stable or possibly a photo-degradation product (Chen et al., 2010; Cawley et al., 2012; Chen and Jaffé, 2014) and has also been identified as derived from the oxidation of soil OM, being exported from the Everglades Agricultural Area (EAA; located to the north of the WCA; Yamashita et al., 2010). As such it is not surprizing that the levels for C2 are enriched in waters from the WCA2, which receives significant canal inputs from the EAA. In the other wetlands and at freshwater marshes in more distant regions of the Everglades, C2 is only a relatively minor component of the DOM fluorescence signal. However, the most interesting aspect of this comparison is that the FCE-based PARAFAC model provided a perfect fit for both the OKA and PAN samples, suggesting that the overall fluorescent properties of the DOM in the three wetlands are quite similar.

3b) NMR study

NMR spectra of SPE-DOM: High field (800 MHz) NMR spectra with cryogenic detection performed on six samples (paired long and short hydroperiod sites from each wetland) revealed an exceptional coverage and chemical description of wetland organic proton and carbon chemical environments. The ^{1}H NMR spectra of wetland DOM acquired with solvent suppression showed the prevalence of rather smooth bulk signal envelopes reflecting intrinsic averaging from massive signal overlap with a considerable variance in abundance for all major chemical environments. In addition, rather minor superimposed sharp individual NMR resonances were indicative of biological signatures and occurred in the order PAN > OKA > FCE (Fig. 2; Fig. S1). From higher to lower field (from right to left), abundant (a) aliphatics, (b) "acetate-analogues", (c) carboxyl-rich alicyclic molecules (CRAM), (d) "carbohydrate-like" and methoxy, (e) olefinic, and (f) aromatic NMR resonances showed well visible and rather broad maxima (letters given according to Fig. S1).

Superimposed small NMR resonances indicative of comparatively abundant biological and biogeochemical molecules were most significant in the aromatic section (f), well noticeable in sections (e) and (a) and of continual lesser occurrence in the order c > b > d (Fig. S1). The area-normalized ^{1}H NMR spectra of the six DOM samples (Fig. 2) showed more variance than their respective ^{1}H NMR section integrals (Table 2), a plausible consequence of intrinsic averaging across sizable chemical shift windows (Hertkorn et al., 2007). One dimensional ^{1}H NMR spectra of wetland SPE-DOM revealed clear distinctions according to sample location, with pronounced congruence between the three pairs of samples (Fig. S1). Within sample pairs, internal differences mainly referred to intensity variations (denoting variable abundance) rather than to alterations of NMR resonances positioning (denoting molecular diversity). The relative disparity was largest between both FCE-L and FCE-S whereas PAN and OKA samples were more alike among themselves. Otherwise, molecular divergence was most obvious in the case of unsaturated protons (δ_H > 5 ppm). Subtle relative changes in composition between pairs of samples were readily visualized by superposition NMR spectra in which the relative NMR section integrals of each aromatic and aliphatic substructures had been normalized to 100% (Fig. S2E, S2F, S2G).

The larger discrimination observed between ^{1}H NMR spectra of DOM from different wetlands in comparison with the intrinsic variance among DOM within each wetland already suggested presence of an individual molecular signature, characteristic of each particular

wetland. Table 2 shows the respective ^{1}H NMR section integrals for the six samples under study. Generally, the OC**H**, XCC**H** and CCC**H** aliphatic chemical environments represented nearly equal contributions to make up approx. 90% of the spectrum with the CCC**H** units consistently exceeding 30%. Carboxyl-rich alicyclic molecules (CRAM) and functionalized and pure aliphatics followed the order FCE (L > S) > PAN ≈ OKA. Molecular divergence was most noticeable in the chemical environment of unsaturated protons, where the ratio of aromatic to olefinic protons declined in the order FCE > PAN > OKA. Here, $\mathbf{H_{ar}}$ (δ_H > 7 ppm) and C=C**H**, O_2C**H** (δ_H : 5.3 – 7 ppm) contributed less than 5% each to the overall spectra. Difference NMR spectra (L-S) obtained for FCE, OKA and PAN wetland SPE-DOM were computed from area-normalized NMR spectra (Fig. S2C, S2D) and indicated congruent behaviour for OKA and PAN SPE-DOM in the purely aliphatic section (δ_H < 3 ppm), with moderate increase of C_nC**H** groups (n > 1; δ_H < 1.6 ppm). The alterations in FCE-based aliphatics were governed by a marked increase of CRAM whereas the abundance of C_nC**H** decreased (Fig. S2D). Interestingly, rather concordant decline of methoxy groups (primarily methyl esters; Fig. 5) was observed for both FCE and PAN (Fig. S2D). Polycarboxylated and PAH-derived aromatics (δ_H > 8 ppm) were markedly increased in FCE-L as compared with FCE-S (cf. below).

For improved assessment of unsaturated protons, the respective chemical shift range was divided into several sections, comprising (f_1; letters according to Fig. S1) polycyclic and polycarboxylated aromatics as well as six-membered nitrogen heterocycles (δ_H > 8 ppm); (f_2) electron withdrawing substituents (COX; Perdue et al., 2007; δ_H ≈ 7.3 – 8.0 ppm); (f_3) electroneutral substituents (alkyl, H, R; δ_H ≈ 7.0 – 7.3 ppm); (f_4) electron-donating substituents (OR, OH, phenolics; δ_H ≈ 6.5 – 7.0 ppm); (e_1) polarized and conjugated olefins; (δ_H ≈ 5.5 – 6.5 ppm); (e_2) isolated olefins (δ_H ≈ 5.0 – 5.5 ppm), this section features however contributions from anomeric protons and certain ester groups (cf. discussion of 2D NMR spectra). The relative and absolute abundance of electroneutral substituted and phenolic aromatic compounds were maximal in OKA, and declined through PAN to FCE. The ratio of conjugated olefins and aromatics was similar in FCE and PAN; however, the abundance of these units was lower by ca. 30% in FCE. DOM from FCE-L showed higher proportions of isolated olefins and, possibly, anomeric positions within carbohydrates.

Within this, the FCE samples showed the lowest proportion of unsaturated protons, and among them, the short hydroperiod site FCE-S was marginally depleted in abundance of carboxylated aromatic protons compared to the longer hydroperiod site FCE-L, possibly due to higher light exposure at the short hydroperiod site. Such differences among samples from PAN and OKA were not significant. Ratios of aliphatic to aromatic signals (CCC**H**/**H**$_{ar}$; see data in Table 2) were also highest for the FCE samples, suggesting enrichment in microbial-derived DOC (periphyton sources) compared to the PAN and OKA samples, but also featuring differences between long and short hydroperiod sites, where preservation of aliphatics at long hydroperiod sites seemed to be favoured for all wetlands. These differences may at first conflict with previous reports where larger periphyton contributions to DOC at FCE-S compared to FCE-L (Chen et al., 2013) suggested to be related to drying and re-wetting of periphyton mats during the dry-to-wet transition at FCE-S and higher relative contributions of soil-derived DOM in FCE-L compared to FCE-S. Similarly, in the case of the long and short hydroperiod comparison, the higher CCC**H**/**H**$_{ar}$ ratios coincided with higher SUVA values for the DOM-L samples, suggesting a difference in the relative contribution of microbial vs. higher plant/soil derived DOM for CDOM compared to bulk DOM. CDOM, often used as a proxy for DOM only represents a small fraction of the bulk DOC and does not include aliphatic molecules as those determined here. As such, while being a convenient and useful proxy for DOC sources, CDOM-based measurements might be less sensitive for the evaluation ofcompositional differences between similar samples.

Methoxy NMR resonances for FCE-S compared to FCE-L were not only more abundant, but were also shifted to lower field, indicating increased fractions of aromatic methylethers and methylesters. FCE-S undergoes periodic drying and thus exposure of soil OM (SOM) to atmospheric conditions and intense sunlight exposure of DOM after high evaporation (drying) conditions. As such, much of the SOM can be aerobically oxidized to CO_2 creating marl soils. It is thus plausible that increased aerobic microbial oxidation and photo-exposure at this short hydroperiod site might enhance DOM oxidation compared to the long hydroperiod site (FCE-L). In addition, while OKA showed an appreciable shoulder at $\delta_H > 3.75$ ppm indicative of aromatic methyl esters and ethers at however, reduced relative abundance, this distinction was absent in both PAN and FCE (Fig. 2).

In addition to the characteristics described above, the FCE samples showed the largest proportion of aromatic compounds substituted with carbonyl derivatives (most likely carboxylic acids; δ_H > 7.3 ppm). This pattern is in accordance with the presence of dissolved black carbon (DBC) at these wetland sites, where the highest abundance was reported for the FCE samples (Ding et al., 2014a). The relatively large fraction of protons with very large downfield chemical shift (δ_H > 8 ppm) suggested the presence of six-membered nitrogen heterocycles as well as that of polycyclic aromatic hydrocarbons (PAH). These units followed the abundance order PAN > OKA > FCE and could be related in part to the presence of dissolved black nitrogen (DBN; Ding et al., 2014b). However, the ratio of olefinic protons (δ_H ~ 5.2 – 6.8 ppm) to aromatic protons (δ_H > 6.8 ppm; but see HSQC cross peaks; Fig. S3) followed the order FCE > PAN ≈ OKA. The distribution of aromatic protons in OKA indicated elevated abundance of electroneutral (alkyl, H; δ_H ≈ 7.0 – 7.3 ppm) and electron-donating substituents (OR, OH; δ_H < 7.0 ppm) in contrast to both FCE and PAN SPE-DOM which showed similar distribution of aromatic protons with larger proportions of electron-withdrawing substituents (COR; δ_H > 7.3 ppm) at however, different overall abundance (Fig. 2; Table 2). In contrast, the abundance of aromatics with electroneutral (R) or electron-donating substitution (OR) with δ_H ~ 7.3 - 6.6 ppm (Perdue et al., 2007) followed the order OKA > PAN >> FCE (Fig. 2), likely reflecting the enhanced relative contributions of higher plant derived DOM (in different degrees of oxidation) for the OKA and PAN compared to the FCE. In conclusion, one-dimensional ^{1}H NMR spectra show a considerable molecular divergence of aromatic molecules in the DOM of the three wetlands, where the compositional features seem driven by both source strengths and variations in biogeochemical processing.

Although some methoxy groups can be formed by reaction of hydroxyl groups in natural DOM and methanol during storage at ambient temperature (as SPE-DOM; Flerus et al., 2011), the **H**CO NMR section integral, which was found typically larger by ~ 2 % for the respective short hydroperiod samples (Table 2), might reflect larger abundance of native methyl esters at these sites or larger abundance of DOM methanolysis products.

^{13}C NMR spectra: ^{13}C NMR spectra of wetland DOM were not overly conspicuous, with limited variance of spectra appearance and ^{13}C NMR section integrals (Fig. S2; Table S1). The

abundance of non-functionalized aliphatics followed the order FCE-L > FCE-S > PAN > OKA, whereas aromaticity followed a near reverse order FCE-L ≈ FCE-S < OKA ≈ PAN. DOM from FCE-L showed depletion of carbohydrates and increase of lipid-like compounds (Table S1). The near invariant abundance of carbonyl derivatives (most likely carboxylic acids) for all DOM could imply that a sizable proportion of low field ^{1}H NMR resonances with chemical shift δ_H > 7.3 ppm, which were more abundant in PAN than in the others (Fig. 2; see also aromatic TOCSY cross peaks, Fig. 3), actually represented (substituted) PAH (with δ_C < 140 ppm; Hertkorn et al., 2013) rather than (poly)carboxylic aromatics (with δ_C ~ 167 – 187 ppm; Fig. 2; Tab. 2; Fig. S2; Table S1). Computed average H/C ratios from a basic reverse ^{13}C NMR based mixing model ranged in the order FCE-L > FCE-S > PAN-S ≈ OKA-L (^{13}C NMR spectra of PAN-L and OKA-S were not acquired) and primarily reflected variable content of aliphatic structures (δ_C ~ 0 – 47 ppm). The computed O/C ratio was near equal for the OKA, PAN and FCE-S samples, whereas that of FCE-L was lower by ~ 0.07 units. Here, a reduced abundance of oxidized aliphatic units (**H**C_{al}O) was primarily responsible, because phenolic and carboxylic content followed the order OKA-L ≈ PAN-S > FCE.

2D NMR spectra: The 2D NMR spectra provided remarkable richness in detail and refined preliminary assignment-proposals from the one-dimensional ^{1}H and ^{13}C NMR spectra. TOCSY NMR spectra (Fig. 3) revealed a wide range of methyl groups (**H**$_3$C-C**H**-X; X: C, O; Fig. 3A, section a); a contiguous, ill resolved cross peak reflected a large number of intra-aliphatic correlations (C-C**H**-C$_n$H-C**H**-C; n = 0 - 2; Fig. 3A, section b), and fewer cross peaks in-between oxygenated aliphatics (O-C**H**-C**H**-O; Fig. 3A, δ_H > 3.4 ppm). Protons bound to sp^2-hybridized carbon produced better resolved TOCSY cross peaks and were part of various α, β-unsaturated olefins (Fig. 3B, section c, d), oxygenated and carbonyl (COX) derivatives of benzenes with up to three COX substituents (Fig. 3C, section f, g, h) as well as six-membered nitrogen heterocycles and more extended aromatic systems with up to several aromatic rings (Fig. 3B, 3C, section e, 3D; Fig. 4). As mentioned earlier, such compounds might be related to the presence of combustion-derived compounds such as DBC and DBN (Ding et al., 2014a and b) and even black sulfur DBS (Hertkorn et al., 2013; see attendant discussion of FTICR mass spectra). In contrast to common five-membered heterocycles, (di)benzothiophene derivatives exhibit NMR

resonances ranging from δ_H: 7.4 – 8.1 ppm; corresponding HSQC cross peaks of DBS would appear in section g, Fig. 4).

HSQC NMR spectra of PAN and OKA did not show peculiar features which were not observable in those of both FCE samples and therefore will be not discussed here. The HSQC NMR spectra of both FCE-S and FCE-L were remarkably similar and produced near identical overlay NMR spectra with some discernible variance in HSQC cross peak amplitude rather than positioning (data not shown). This behavior is expected from comparison of the one-dimensional ^{1}H NMR spectra. These display differences in relative amplitude rather than positioning of NMR resonances which is indicative of variance in abundance of certain molecules rather than variance in molecular diversity (see, however, discussion of CHOS compounds present in FCE DOM as derived from FTICR mass spectrometry). About 90 % of overall HSQC cross peak integral resided in a contiguous expansive superimposed assembly of HSQC cross peaks originating from protons bound to sp^3-hybridized carbon (Fig. S3).

The resolution of these expansive aliphatic HSQC cross-peaks of FCE-S (Fig. S3) could be remarkably improved by spectral editing according to carbon multiplicity (Fig. 5). The combination of methyl- and methylene-selective DEPT-HSQC NMR spectra revealed well discriminated cross peaks for all three types of protonated carbon; i.e. methyl, methylene and methine (Fig. 5). The chemical diversity of X-CH_3 groups as indicated by DEPT HSQC cross peaks (section a, Fig. 5) was noteworthy, and the near Gaussian distribution of C-CH_3 cross peak amplitude in ^{1}H and ^{13}C NMR frequencies indicated near maximum diversity of aliphatic chemical environments associated with these methyl groups. However, classical methyl groups terminating extensive, purely aliphatic units (δ_H < 1.0 ppm; CCC**$\underline{CH_3}$** units) contributed less than 20% to the total C**$\underline{CH_3}$** HSQC cross peak integral. The large majority of C-**$\underline{CH_3}$** units was sufficiently proximate to carbonyl derivatives (i.e., most likely carboxylic acids) to let those experience downfield chemical shift anisotropy from these nearby carbonyl groups, resulting in chemical shifts ranging from δ_H ~ 1.0 - 1.7 ppm, respectively (cross peak a; Fig. 5). Alicyclic structures (e.g. CRAM; Hertkorn et al., 2006) facilitate clustering of chemical environments as shorter paths of chemical bonds between different substituents are realized in rings rather than in open chains. Another ~20% of C**$\underline{CH_3}$** in FCE was bound to olefins [C=C-**$\underline{CH_3}$**], with a possible contribution of S-**$\underline{CH_3}$** groups (section b; Fig. 5).

The carbon bound methylene (C-**CH$_2$**-C) cross peak occupied an impressively large area down to $\delta_H \sim 3.5$ ppm, well into the proton chemical shift range commonly attributed to OC**H** units. The two major chemical environments discriminated were methylene more distant to COX (C-**CH$_2$**-C$_n$-COX, with $n \geq 1$, and $\delta_H < 2.1$ ppm cross peak d; Fig. 5), and methylene groups directly proximate to carboxylic groups (in α-position; i.e. C-**CH$_2$**-COX, with $\delta_H > 2.1$ ppm cross peak e; Fig. 5). The former shows a wider range of remote carbon substitution as indicated by the substantial spread of respective carbon chemical shifts ($\Delta\delta_C$: 24 / 16 ppm, respectively for section d / e HSQC cross peaks; Fig. 5; see also Fig. 8b in Hertkorn et al., 2013). A wide variety of aliphatic and aromatic methylesters and methylethers were also found, the latter being virtually absent in marine SPE-DOM. Here, aliphatic methyl esters were most abundant (section g_2 in Fig. 5), aromatic methyl esters (section g_3 in Fig. 5) and methyl ethers (section g_4 in Fig. 5); were of similar abundance, and clearly recognizable aliphatic methyl ethers were also present (section g_5 in Fig. 5). Oxomethylene (O**CH$_2$**) occurred in the form of carbohydrate side chains (section h; Fig. 5), and a remarkable set of aliphatic oxomethylene (OCH_2) HSQC cross peaks ($\delta_{H/C} \sim 3.4 - 4.0/58 - 72$ ppm; section j in Fig. 5) was present in SPE-DOM FCE-S, which does not correspond to common lignin β-aryl ether units, which resonate in this ^{1}H and ^{13}C NMR chemical shift range, but commonly comprise C$_{ar}$-**CH**-O-, i.e. methine substructures. Analogous oxomethine substructures are also found in phenylcoumaran, resinol and dibenzodioxocin units as well, whereas oxomethylene units with $\delta_H > 4.5$ are rare in common lignins (Ralph et al., 1998; Yelle et al., 2008; Martinez et al., 2008; Wen et al., 2013; Yuan et al., 2011). This peculiar HSQC cross peak was discovered in FCE-S wetland SPE-DOM (section j HSQC cross peak in Fig. 5) but since then has also been observed (in retrospect) with lesser distinction in other SPE-DOM including those from marine sources. The singular positioning of a methylene group in the ^{1}H and ^{13}C NMR chemical shift space strongly restrains the potential diversity of its chemical environments: it has to represent a OCH_2 group (methylene as defined by the phase in ^{1}H, ^{13}C DEPT HSQC NMR spectra; single oxygen because of δ_C: any O-CH_2-O environment would resonate at $\delta_C > 90$ ppm). Similarly, common O-CH_2-N chemical environments would resonate at higher field than observed in both $\delta_{H/C}$, but cannot be excluded entirely in case of peculiar remote substitution. The most plausible substructure is OCH_2C; then, δ_H from 5.3 – 5.7 ppm warrants presence of an ester group: this implies a -C-(C=O)-O-CH_2-C substructure. However,

alkylation alone will not produce the necessary low field δ_H observed. This leaves -C-(C=O)-O-CH_2-C=O as a plausible group; possibly confined with a carboxylic group such as -C-(C=O)-O-CH_2-COOH or as an ester –C-(C=O)-O-CH_2-COOR. Both these substructures have a decent propensity to form enols with variable double bond character -C-(C=O)-O-CH=$CH(OH)_2$. A partial double bond character, which might be possibly controlled by mutual interactions in the complex DOM mixture of molecules, would also explain the observed spread of chemical shift in ^{1}H and ^{13}C NMR frequencies in this HSQC cross peak even if the methylene group itself in -C-(C=O)-O-CH_2-COOH is four (carbon) or five (proton) bonds away from the most proximate atom position where substitution may affect its chemical shift.

Several thousands of acid and ester derivatives of acetoacetic acid [H_3C-(C=O)-O-CH_2-COOH] are known in literature. Here, many of the common esters comprise lipid substructures such as n-alkanes, sterane and other polyalicyclic hydrocarbons, trimethylammonium salts, among others, suggesting a natural origin of these compounds also in wetland SPE-DOM. While substructures with –O-CH_2-COOZ (Z: H, R) will produce distinct "oxomethylene (O**$\underline{CH_2}$**C)" cross peaks in ^{1}H, ^{13}C DEPT HSQC NMR spectra (section j cross peak; Fig. 5), the derivatives with –O-$CHCH_n$-COOZ substructures (Z: H, R) will contribute to the ^{1}H NMR downfield section of the expansive "oxomethine (O**$\underline{CH}$**C_2)" ^{1}H, ^{13}C DEPT HSQC cross peak (section i cross peak; Fig. 5) and will not be readily discerned owing to a larger variance in remote substitution. In addition, oxomethylene units without geminal and vicinal adjacent protons will very likely produce intense singlet NMR resonances, contributing to the enhanced visibility of HSQC cross peaks even at rather limited relative abundance. Further evaluation of aliphatic spin systems in FCE-L provided evidence for massive aliphatic branching in CCC**$\underline{H}$** units and of large chemical diversity of remote carboxylic substitution (Fig. S5).

TOCSY and HSQC NMR spectra demonstrated presence of olefinic and aromatic unsaturation in all wetland SPE-DOM (Fig. 3 and Fig. 4). The FCE-S showed the most informative detail of HSQC cross peaks arising from unsaturated C_{sp2}H groups (Fig. 4). In comparison with marine SPE-DOM (cf. Fig. S4 and attendant discussion), wetland SPE-DOM displayed a more restricted chemical diversity of conjugated olefins (Fig. 4) whereas all kinds of oxygenated aromatics, i.e. those substituted with electron-withdrawing (e.g. COOH) and electron-donating (e.g. OR) substituents were much more abundant and chemically diverse in (all) wetland DOM (not all data shown). The latter finding is indicative of polyphenol input from

vascular plants (e.g. lignin-derivatives) into wetland DOM whereas aromatics in marine SPE-DOM mainly reflect marine natural products (Fig. S4). In general, aromatic unsaturation (as deduced from proton NMR integrals; Table 2) followed the order PAN > OKA > FCE (Fig. S1), whereas olefinic unsaturation followed the order OKA ~ PAN > FCE (Fig. S1). Aliphatic to aromatic ratios changed across the different samples with an order of FCE > PAN > OKA, suggesting higher relative contributions from periphyton in the FCE, whilst the PAN and OKA were more influenced by higher plant-derived organic matter including lignins. The olefinic to aromatic ratios (FCE-L: 0.44; FCE-S: 0.39; ; PAN: 0.38; OKA: 0.41) were computed from adapted ^{1}H NMR section integrals [δ_H: 10 – 6.5 ppm (aromatics) / 6.5 – 5.0 ppm (olefins); Table 2]´owing to HSQC cross peak positioning wich indicated major contribution of oxygenated aromatics $\underline{\mathbf{C_{ar}}}$O at δ_H: 7.0 – 6.5 ppm; Fig. 4] and showed lower values than oceanic DOC (Hertkorn et al., 2013), who reported olefinic to aromatic ratios in the range of 1.2 to 3.0. It is likely that this significant difference is due to the contributions of higher plant, lignin-rich carbon in the wetlands compared to marine DOM. The slightly elevated olefin content found in FCE-L may result from the contribution of periphyton-derived DOM in the Everglades (Maie et al., 2005; Chen et al., 2013), (see also above). In addition, all three sites are known for frequent and seasonal fires and have been reported to contain dissolved black carbon (DBC) (Ding et al., 2014a) in abundances close to 10% of their DOC on a global average (Jaffé et al., 2013). However, the DBC content (as %DOC) in the FCE was higher (as high as 20% of DOC) than for the PAN samples (13% and 14% for PAN-L and PAN-S respectively), and these higher than the OKA samples (9.4% and 6.3% for OKA-L and OKA-S respectively) studied here (Jaffé et al., 2013). In addition, the presence of six-membered N-containing heterocycles in these samples might be indicative of the presence of dissolved black nitrogen (DBN), which has previously been reported in the FCE (Maie et al., 2006) and proposed to consist of polyaromatic molecules containing pyrrolic-N, and multiple carboxylic substituents (Wagner et al., 2015); (see also attendant FTMS-based discussion of dissolved black sulfur compounds (DBS) in FCE-S; section 3e). With regards to the degree of oxidation of the aromatic signal, the OKA showed the highest proportion of electron donating groups (Fig. 2; Fig. S1 and S2), such as phenols and ethers, possibly related to lignin oxidation products, while the FCE featured the highest shares of electron withdrawing substituents (e.g. carboxyl groups) possibly associated with DBC. Although all three ecosystems have climates leading to high light exposure, the high levels of

DOC in the FCE suggest some degree of self-shading, while DOC in the OKA is generally lower and the system is known for its capacity to photo-degrade DOM (Cawley et al., 2012). Thus, the degree of photo-exposure of the DOM combined with combustion by-products such as DBC, may be the driver controlling the oxidation state of the aromatic fraction. The photo-reactivity of DBC in marine environments has recently been shown (Stubbins et al., 2012) and may play a role in the lower DBC levels observed in the OKA samples.

3c) FTICR mass spectrometry:

Ultrahigh resolution Fourier transform ion cyclotron mass spectra (FTICR/MS) of SPE-DOM may provide several thousands of mass peaks for individual samples (Koch et al., 2005; Kujawinski et al., 2009), of which many hundreds were assigned here to extended CHO, CHNO, CHOS and CHNOS molecular series (Schmitt-Kopplin et al., 2010) based on the technique's excellent mass accuracy and mass resolution (Fig. 6 and Table 3). Although detailed FTICR/MS data are derived only from a few paired SPE-DOM samples (Long and Short hydroperiod) for each wetland, a slightly higher number of mass peaks (relative difference < 6%) and of assigned molecular formulas (relative difference < 1%) was observed for the FCE-L compared to the FCE-S, whereas elevated counts of mass peaks and assigned molecular compositions were found in case of the PAN-S and OKA-S samples (relative difference < 2%; Table 3). Molecular weights ranged in the order FCE-S > FCE-L ~ PAN > OKA (Table 3). This admittedly minor molecular weight difference was not reflected in the S_R values of these samples (Table 1) which were quite similar. However, S_R only represents a molecular weight proxy for CDOM and might not be sensitive enough to reflect minor differences accurately. In general, while SPE-DOM of both OKA and PAN showed near 57 ± 2% CHO, 8 ± 2% CHOS, 33 ± 2 CHNO, and < 1% CHNOS molecules, the mass spectra of FCE samples were fundamentally different compared with respect to both OKA and PAN as well as among themselves (Fig. 6; Table 3; see also Fig. 7). Sample FCE-S appeared most distinct from all other samples both with respect to total count of ions, overall mass peak distribution and with respect to molecular diversity within nominal mass ranges (Fig. 6). Here, FTICR mass spectra of both FCE samples showed the conspicuous doublets of CHO/CHOS pairs visible at high resolution (Δm ($C_{-3}H_4S$) = 2.4 mDa) indicating a nominal exchange of H_4S against C_3 (Schmitt-Kopplin et al., 2010), whereas all other samples showed both lower abundances and diversity of CHOS compounds (Fig. 6 and Fig. 7). In case of

612 the FCE samples, CHOS and CHNOS compounds were markedly enriched at the expense of
613 CHO and CHNO compounds. While the proportion of CHNO (21 ± 1%) and CHNOS (9 ± 1%)
614 molecules were similar for both FCE samples, the abundance of CHOS molecules in FCE-S was
615 elevated by more than 10%, predominantly at the expense of CHO molecules. The overall
616 abundance of sulphur in the FCE was nearly four-fold when compared with that of the OKA and
617 PAN samples (Table 3).

618 CHOS compounds observed in all wetland samples already showed a remarkable
619 chemical diversity (Fig. 8B). However, the chemical dissimilarity of CHOS compounds common
620 to both FCE samples remarkably exceeded that found in OKA and PAN, covering a substantial
621 share of the CHOS chemical space from O/C ratio: 0.3 - 0.8 and H/C ratio 0.6 - 1.7, respectively
622 (Fig. 8C). Here, four groups of CHOS molecules were differentiated based on their positioning
623 in H/C against O/C van Krevelen diagrams (Fig. 7F): (a) saturated sulfolipids with H/C ratio > 2
624 and intermediate O/C ratio, suggesting the presence of sulphur in elevated oxidation states; (b)
625 unsaturated sulfolipids with a rather restricted H/C and O/C ratio; (c) a very large and expansive
626 set of molecularly diverse CHOS molecules with a bandwidth of O/C ratios similar to CHO
627 compounds but reaching out to higher saturation (larger H/C ratio) than the latter (Fig. 7E, Fig.
628 7F); (d) unique to FCE-S (with traces in OKA-L) was a large set of aromatic and oxygen-
629 deficient "black sulphur" compounds (DBS; section d; Fig. 7F, similarly positioned like CHOS
630 compounds in Atlantic open Ocean abyssopelagic SPE-DOM at 5446 m depth (Fig. S8, in
631 Hertkorn et al., 2013), but covering a larger mass range (Fig. S7F). Section (b) and (c) CHOS
632 compounds were also observed in PAN and OKA, whereas black sulphur compounds were rare
633 in OKA-L and virtually absent in the other samples except FCE-S. DOM-type CHOS
634 compounds common to all six wetland samples were on average more saturated and oxygenated
635 than their respective CHO and CHNO counterparts, suggesting also here presence of sulphur in
636 elevated oxidation states (Fig. 8B).

637 The CHOS compounds of both FCE samples not only differed fundamentally from those
638 found in OKA and PAN, but were also remarkably diverse in both FCE-L and FCE-S samples
639 itself. Figure 9 indicates CHOS compounds present with elevated abundance in either FCE-S
640 (Fig. 9A) or FCE-L (Fig. 9B). The most peculiar feature of FCE-S was a hydrogen-deficient pool
641 of (poly)aromatic CHOS compounds (section a mass peaks; Fig. 9A) in extended molecular
642 series with limited degree of oxidation (O/C ratio < 0.22), ranging from m/z ~ 300 - 600. The

positioning in both van Krevelen and mass-edited H/C ratio diagrams (Fig. 9) was in accordance with that of ´black sulphur´ in abyssopelagic South Atlantic SPE-DOM (Fig. S8; Fig. 16 in Hertkorn et al., 2013), but its signature was more conspicuous and showed larger richness of diverse CHOS compounds in FCE-S. While sulphur can be readily inserted into any C-C and C-H bond, analogous to oxygen, organic sulphur can also occupy oxidation states ranging from -2 to +6, an option not available to oxygen. Nevertheless, the manifest oxygen-deficiency of the proposed highly unsaturated CHOS molecules (section d; Fig. 7F) suggests the presence of reduced sulphur in the form of sulphides. Aromatic CHOS molecules will then most likely occur as benzothiophene derivatives, a chemical environment of sulphur largely favoured in mineral oils (Purcell et al., 2007; Liu et al., 2010; Muller et al., 2012). While both black carbon as well as black nitrogen (Wagner et al., 2015) have been reported in the FCE (Ding et al., 2014 a and b; Maie et al., 2006) the presence of this 'black sulphur' was not previously observed at FCE. The environmental factors driving the high abundance of these compounds at FCE-S remain unclear but may be related to the higher fire frequency at short hydroperiod sites and possibly soil charring. A small set of CHOS compounds with more average H/C and O/C ratios (section b mass peaks; Fig. 9A) was accompanied by a rather minor set of highly oxygenated CHNOS compounds, with an O/C ratio > 0.75 (Fig. 9A).

In contrast to the FCE-S, the FCE-L sample displayed an oxygenated set (O/C ratio > 0.6) of a few dozen hydrogen-deficient (H/C ratio < 1.1) CHOS molecules in truncated molecular series and at rather low mass (m/z < 400; Fig. 9B section d mass peaks). These molecules were most likely composed of oxygenated aromatics connected by (some) ether bridges, which rather likely originate ultimately from plant and/or algal polyphenols. Apart from PAH derived compounds, which are commonly rather oxygen-deficient (O/C ratio < 0.3), these structures represent one of the most plausible motifs of very hydrogen-deficient molecules found in DOM (H/C ratio < 1). The large extent of average oxygenation makes sulphur functional groups in elevated oxidation states, e.g. sulfones, sulfonates or sulfates, likely candidates for this group of CHOS compounds.

In addition, a rather expansive cloud of abundant CHOS and less common CHNOS compounds at mass range m/z: 200 – 550, with large and variable extent of oxygenation (O/C ratio: 0.4 – 0.95) was prominent in FCE-L and near absent in FCE-S. The sizable expansion of this cloud with a huge range of H/C ratios testified for a rather large overall diversity of these

unique CHOS molecules found solely in FCE-L; CHNOS compounds seemed to follow suit but with a lesser overall diversity: highly oxygenated (O/C ratio > 0.8) and hydrogen-rich CHNOS molecules (H/C > 1.6) were missing even if every added nitrogen carried one intrinsic hydrogen into analogous CHO molecular formulas. This higher molecular diversity for the FCE-L site may be driven by higher soil-derived (peat soils) DOM contributions at this site compared to FCE-S (marl soils) (Chen et al., 2013) and a higher degree of DOM preservation at this deeper, less photo-exposed site.

It has to be mentioned that this cloud encircled the common molecular series of several hundreds of CHOS and CHNOS compounds found in both FCE samples (Fig. 7E, 7F, 8C). The significantly higher presence of sulphur-containing molecular formulas for the FCE samples is likely the result of higher inputs of sulphate to the Everglades compared to the Pantanal and Okavango. Firstly, Everglades is a coastal wetland where sea-spray may be an important contributor to sulphate. In addition, it is the most anthropogenically impacted wetland of the three being compared, where runoff from agricultural lands within the Everglades watershed is likely the most important contributing factor to the sulphur load of the system as it is an ingredient of fertilizer applications. The CHO and CHNO components specific for the OKA and PAN samples are shown in Figure 8F and suggest, in agreement with the NMR data, a higher degree of oxidized, H-deficient materials at these sites compared to the FCE. This is particularly true for the PAN which show unique molecular formulas for oxidized, H-deficient CHO and CHNO components (Fig. 8D),whereas molecular formulas unique for the OKA are relatively few (Fig. 8E).

Comparative analyses of van Krevelen diagrams between the six sites as shown in Fig. 8A clearly cluster the FCE samples separately from the OKA and PAN. Cluster analysis showed a clear distinction between the FCE on one hand and the OKA and PAN samples on the other hand, with less pronounced but significant differences between the paired, long and short hydroperiod samples at each site (Fig. 8A). Among pairs of DOM samples, similarity according to FTICR/MS-based cluster analysis was in the order PAN > OKA > FCE (Fig. 8A), whereas one-dimensional ^{1}H NMR spectra clustered according to increasing dissimilarity in the order OKA < PAN < FCE (Fig. S2). This discrepancy is readily explained by the different recognition of aliphatic groups in FTMS (insensible) and NMR spectra (quantitative depiction). CHO and CHNO molecules ionized by negative ESI occupied rather similar expansive regions with near

705 average H/C and O/C elemental ratios (Fig. 8B). This is a common feature of DOM molecular
706 distribution as derived from FTICR mass spectra. Here, the largest number of feasible and
707 chemically reasonable isomeric molecules will project on single mass peaks at average H/C and
708 O/C elemental ratios, contributing to larger overall mass peak amplitude – this applying even
709 more specifically to van Krevelen diagrams, in which different molecular compositions with
710 identical elemental ratios contribute to the same data points (Hertkorn et al., 2007; Lechtenfeld et
711 al., 2014). Analogously, the distribution of CHO, CHOS, CHNO and CHNOS molecular series
712 roughly coincided, with some displacement of CHOS molecules in both FCE samples, towards
713 higher H/C ratio (i.e. higher aliphatic character).

714 At first glance, the H/C vs. O/C (Fig. 7) as well as the H/C vs. m/z (Fig. S7) plots,
715 showed near uniform fingerprints for OKA and PAN, covering larger areas in the van Krevelen
716 diagrams in case of CHO compared with CHNO compounds (Fig. 8B), suggesting an increased
717 overall chemical diversity of CHO compounds. In addition, the paired wetland samples clustered
718 separately for the high and low hydroperiods respectively, suggesting that molecular
719 compositions differ among sites with different hydrology. The weighted average O/C and H/C
720 values were remarkably similar for PAN and OKA showing rather marginal variance between
721 different sites or between high and low hydroperiod (Fig. 7A-7D; Table 3). In comparison, FCE-
722 S showed a considerably decreased O/C ratio. While computed O/C ratios of wetland DOM
723 exceeded those found in oceanic DOM by about 0.2 units (Table 3; Table 4 in Hertkorn et al.,
724 2013), the H/C ratio of wetland DOM was approximately 0.15 units higher in comparison. Even
725 if ionization selectivity in negative ESI FTICR mass spectra applied, the ^{1}H NMR section
726 integrals indicate analogous trends of relative saturation, or alternatively, hydrogen deficiency
727 between wetland and marine DOM. In comparison with average wetland SPE-DOM, average
728 open ocean SPE-DOM showed lesser abundance of aromatics (by 2-3%), lower proportions of
729 OC**H** chemical environments (by 8%), and, especially, higher abundance of pure aliphatics (i.e.
730 CCC**H** units; by 12%). This implies that marine DOM shows lower abundance of hydrogen-
731 deficient (unsaturated) and higher abundance of hydrogen-rich (purely aliphatic) molecules than
732 wetland DOM, in line with the elevated H/C ratio as derived from FTICR mass spectra.
733 Similarly, the higher abundance of oxygen-rich OC**H** chemical environments in wetland DOM
734 as seen by ^{1}H NMR section integral (8% relative increase) was in accordance with the increased
735 O/C ratio found in their FTICR mass spectra (increase by 0.2 units).

Comparative analysis of van Krevelen diagrams (Fig. S8) obtained solely from the four PAN and OKA samples confirmed the previously observed higher similarity between PAN-L and PAN-S compared to OKA-L and OKA-S sample pairs (Fig. 7; Fig. 8D, 8E and 8F). Molecular compositions with unique high abundance when derived from all six wetland samples were sparse and non-significant in case of OKA (Fig. 8E; see also Fig. S8B), whereas molecular compositions with unique high abundance in all four OKA and PAN samples occupied a rather dense, contiguous section of hydrogen-deficient (H/C ratio < 1) and oxygenated (O/C ratio ~ 0.3 – 0.7) CHO and CHNO molecules (Fig. S8F). In agreement with its high degree of photo-oxidation, the OKA contained higher proportions of highly oxygenated CHO (O/C ratio > 0.7) and CHNO (O/C ratio > 0.5) molecules, and a few rather abundant (and easily ionizable) sulfolipids (Fig. S8B), whereas PAN SPE-DOM displayed larger proportions of hydrogen-deficient CHO molecules of considerable chemical diversity and extent of oxygenation (O/C ratio ~ 0.2–0.9), and several dozens of CHNO molecules similarly positioned but with more limited range of oxygenation and hence, overall chemical diversity (Fig. 8D).

Remarkably, with the exception of a tiny section of CHO molecules (H/C ~ 1.1; O/C ~ 0.4), both CHO and CHNO molecular series for OKA and PAN nearly perfectly superimpose in the H/C against O/C van Krevelen diagram (Fig. 8F). It is very likely that these CHO molecules jointly present in PAN + OKA mainly represent oxygenated aromatic molecules, possibly connected by ether linkages. This is one of the most comprehensive ways to envision such hydrogen-deficient molecules of conceivable natural product origin, and in agreement with the NMR data suggesting a higher degree of oxidized, H-deficient materials at these sites compared to the FCE. This is particularly true for the PAN which shows unique molecular formulas for oxidized, H-deficient CHO and CHNO components, whereas molecular formulas unique for the OKA are relatively few (Fig. 8E and Fig. 8F; Figs. S8B and S8C).

4) Conclusions and biogeochemical significance: Very detailed molecular analyses of DOM samples from three different sub-tropical freshwater wetlands suggest in agreement with previous reports on riverine and marine DOM characterizations, that many of the bulk molecular characteristics in freshwater DOM are shared by ecosystems despite being very different in their environmental settings (Repeta et al., 2002; Jaffé et al., 2012). Nevertheless, organic structural spectroscopy provided evidence for wetland-specific molecular assemblies. NMR and FTMS

analysis provided exceptional coverage of wetland SPE-DOM composition and structure, confirming individual wetland organic matter molecular characteristics. Those were directly revealed in ^{1}H NMR spectra, while extensive mathematical analysis was mandatory to discern analogous distinction in FTICR mass spectra, which show extensive projection of structural variance on the primary measured variable. These detailed analyses revealed significant variations in the molecular composition that can, in some cases, be controlled by site-specific environmental conditions. Among those are hydrological drivers such as hydroperiod (lengths and depths of inundation), resulting in variations in light penetration and associated photochemical processes along with seasonal drying of surface soils and associated aerobic oxidation processes. Other drivers include (i) external sources of sulphur, such as agricultural activities and sea spray, resulting in the formation of a variety of sulphur compounds in DOM, (ii) fire regime, possibly causing soil OM charring during wildfires, and (iii) natural DOM source variations and source strength in the contribution from vascular plants, grasses, and aquatic vegetation including microbial contributions from periphyton. As such, while not all molecular differences could be explained through one or more of these drivers, this study illustrates for the first time the extensive molecular diversity and compositional complexity of DOM in wetlands, and as such should serve as a database for future characterization efforts. Further detailed molecular-level characterizations of wetland DOM are encouraged as a means to better understand spatial and seasonal variability in sources, transformations and reactivity, which can be ultimately used to aid in constraining carbon cycling models.

Author contributions: N.H. performed NMR study, contributed to the study design, data interpretation and actively participated in the writing of the manuscript; M.H. performed the FT-ICR/MS analyses and data manipulation and participated in data interpretation; K.C. collected samples from Okavango Delta and Everglades and performed optical properties study; P.S-K. Provided support for the FT-ICR/MS analyses and general data interpretations; R.J. collected Pantanal and Everglades samples, participated in all data interpretations, took the lead in coordinating this study and writing this manuscript, and generated funding in support of this research. All authors provided significant input on the final manuscript.

Acknowledgments: The authors acknowledge the financial support from the George Barley Chair and the Florida Coastal Everglades LTER (DEB – 1237517; to RJ) for funding this research. We are grateful to C. Nunes, N. Rodriguez, N. Mladenov and P. Wolski who assisted in field logistics and to S. Thaller (HMGU) who expertly assisted in NMR sample preparation and assembly of figures. K.C. thanks Everglades National Park for a Post-doctoral fellowship. This is SERC contribution number ###.

References:

Aluwihare, L. I. and Repeta, D. J.: A comparison of the chemical characteristics of oceanic DOM and extracellular DOM produced by marine algae, *Mar. Ecol. Prog. Ser.* **186**, 105-117, 1999.

Amon, R. M. W. and Benner, R.: Bacterial utilization of different size classes of dissolved organic matter, *Limnol. Oceanogr*. **41**, 41-51, 1996a.

Amon, R. M. W. and Benner, R.: Photochemical and microbial consumption of dissolved organic carbon and dissolved oxygen in the Amazon River system, *Geochimica et Cosmochimica Acta*. **60**, 1783-1792, 1996b.

Battin, T. J., Luyssaert, S., Kaplan, L. A., Aufdenkampe, A. K., Richter, A., Tranvik, L. J.: The boundless carbon cycle. *Nature Geoscience.* **2**, 598-600, 2009.

Bell, N. G. A., Michalchuk, A. A. L., Blackburn, J. W. T., Graham, M. C., and Uhrin, D.: Isotope-filtered 4D NMR spectroscopy for structure determination of humic substances, *Angewandte Chemie* International Edition, **54,** 1-5, 2015.

Blough, N. V. and Green, S. A.: Spectroscopic characterization and remote sensing on non-living organic matter. p. 23-45. *In*: R. G. Zepp and C. Sonntag (Eds.), The role of non-living organic matter on the earth's carbon cycle. John Wiley and Sons. 2009.

Cawley, K., Wolski, P., Mladenov, N. and Jaffé, R.: Dissolved organic matter biogeochemistry along a transect of the Okavango Delta, Botswana. *Wetlands* **32**, 475-486**,** 2012.

Chen, M. and Jaffé, R.: Photo- and bio-reactivity patterns of dissolved organic matter from biomass and soil leachates and surface waters in a subtropical wetland, *Water Research* **61**, 181-190, 2014.

Chen, M., Price, R. M., Yamashita, Y. and Jaffé, R.: Comparative study of dissolved organic matter from groundwater and surface water in the Florida coastal Everglades using multi-dimensional spectrofluorometry combined with multivariate statistics, *Applied Geochemistry* **25**, 872-880, 2010.

Chen, M., Maie, N., Parish, K. and Jaffé, R.: Spatial and temporal variability of dissolved organic matter quantity and composition in an oilgotrophic subtropical coastal wetland, *Biogeochemistry* **115**, 167-183, 2013.

D'Andrilli, J., Dittmar, T., Koch, B. P., Purcell, J. M., Marshall, A. G., Cooper, W. T.: Comprehensive Characterization of Marine Dissolved Organic Matter by FT-ICR Mass Spectrometry with Electrospray and Atmospheric Pressure Photoionization, *Rap. Comm. Mass Spectrom*. **24**, 643-650, 2010.

Ding, Y., Cawley, K., Nunes, C., and Jaffé, R., Environmental dynamics of dissolved black carbon in wetlands, *Biogeochemistry* **119**, 259-273, 2014a.

Ding, Y., Watanabe, A. and Jaffé, R.: Dissolved black nitrogen (DBN) in freshwater environments: Source and land to ocean flux assessment, *Organic Geochemistry* **68**, 1-4, 2014b.

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

Einsiedl, F., Hertkorn, N., Wolf, M., Frommberger, M., Schmitt-Kopplin, P. and Koch, B. P.: Rapid biotic molecular transformation of fulvic acids in a karst aquifer, *Geochimica et Cosmochimica Acta* **71**, 5474-5482, 2007.

Ellery, W. N., McCarthy, T. S. and Smith, N. D.: Vegetation, hydrology and sedimentation patterns on the major distributary system of the Okavango Delta, Botswana, control of a river system by vegetation, *Wetlands* **23**, 357-375, 2003.

Fellman, J. B., Hood, E. and Spencer, R. G. M.: Fluorescence spectroscopy opens new windows into dissolved organic matter dynamics in freshwater ecosystems: a review. *Limnology Oceanography* **55**, 2452–2462. 2010.

Flerus, R., Koch, B. P., Schmitt-Kopplin, P., Witt, M. and Kattner, G.: Molecular level investigation of reactions between dissolved organic matter and extraction solvents using FT-ICR MS, *Marine Chemistry* **124**, 100-107, 2011.

Foden, T., Sivyer, D. B., Mills, D. K. and Devlin, M. J.: Spatial and temporal distribution of chromophoric dissolved organic matter (CDOM) fluorescence and its contribution to light attenuation in UK waterbodies, *Est. Coastal Shelf Sci.* **79**, 707-717, 2008.

Gieske, A.: Modelling outflow from the Jao/Boro river system in the Okavango delta, Botswana, *Journal of Hydrology* **193**, 214–239, 1997.

Gumbricht, T., McCarthy, J. and McCarthy, T. S.: Channels, wetlands and islands in the Okavango Delta, Botswana, and their relation to hydrological and sedimentological processes, *Earth Surface Processes and Landforms* **29**,15–29, 2004.

Hedges, J. I., Eglinton, G., Hatcher, P. G., Kirchman, D. L., Arnosti, C., Derenne, S., Evershed, R. P., Kogel-Knabner, I., de Leeuw, J.W., Littke, R., Michaelis, W. and Rullkotter, J.: The molecularly-uncharacterized component of nonliving organic matter in natural environments. *Organic Geochemistry* **31**, 945–958, 2000.

Helms, J. R., Stubbins, A., Ritchie, J. D. and Minor, E. C.: Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photobleaching of chromophoric dissolved organic matter, *Limnology & Oceanography* **53**, 955–969, 2008.

Hertkorn, N., Harir, M., Koch, B. P., Michalke, B. and Schmitt-Kopplin, P.: 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, 2013.

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

Hertkorn, N., Frommberger, M., Witt, M., Koch, B. P., Schmitt-Kopplin, P. and Perdue, E. M.: Natural organic matter and the event horizon of mass spectrometry, *Analytical Chemistry* **80**, 8908-8919, 2008.

Hertkorn, N., Ruecker, C., Meringer, M., Gugisch, R., Frommberger, M., Perdue, E. M., Witt, M. and Schmitt-Kopplin, P.: High-precision frequency measurements: indispensable tools at the core of the molecular-level analysis of complex systems, *Analytical Bioanalytical Chemistry* **389**, 1311-1327, 2007.

Jaffé, R., Ding, Y., Niggemann, J., Vähätalo, A., Stubbins, A., Spencer, R., Campbell, J. L. and Dittmar T.: Global mobilization of charcoal from soils via dissolution and subsequent riverine transport to the oceans, *Science* **340**, 345-347, 2013.

Jaffé, R., Cawley, K. M. and Yamashita, Y.: Applications of excitation emission matrix fluorescence with parallel factor analysis (EEM-PARAFAC) in assessing environmental dynamics of natural dissolved organic matter (DOM) in aquatic environments: A review. *Advances in the Physicochemical Characterization of Organic Matter*, ACS Series, Rosario F. - Editor. 27-73, 2014.

Jaffé, R., Yamashita, Y., Maie, N., Cooper, W. T., Dittmar, T., Dodds, W. K., Jones, J. B., Myoshi, T., Ortiz-Zayas, J. R., Podgorski, D. C. and Watanabe, A.: Dissolved Organic Matter in Headwater Streams: Compositional Variability Across Climatic Regions, *Geochimica et Cosmochimica Acta* **94**, 95-108, 2012.

Jaffé, R., McKnight, D. M., Maie, N., Cory, R. M., McDowell, W. H. and Campbell, J. L.: Spatial and temporal variations in DOM composition in ecosystems: The importance of long-term monitoring of optical properties. *Journal of Geophysical Research - Biogeosciences*. doi:10.1029/2008JG000683, 2008.

Junk, W. J., Brown, M., Campbell, I. C., Finlayson, M., Gopal, B., Ramberg, L. and Warner, B. G.: The comparative biodiversity of seven globally important wetlands: a synthesis, *Aquatic Sciences* **68**, 400–414, 2006a.

Junk, W. J., Nunes da Cunha, C., Wantzen, K. M., Petermann, P., Strüssmann, C., Marques, M. I. and Adis, J.: Biodiversity and its conservation in the Pantanal of Mato Grosso, Brazil. *Aquatic Sciences* **68**, 278–309, 2006b.

Junk, W. T. and Nunes da Cunha, C.: Pantanal: a large South American wetland at a crossroads, *Ecological Engineering* **24**, 391-401, 2005.

Kaiser, E., Simpson, A. J., Dria, K. J., Sulzberger, B. and Hatcher, P. G.: Solid-state and multidimensional solution-state NMR of solid phase extracted and ultrafiltered riverine dissolved organic matter, *Environmental Science & Technology* **37**, 2929-2935, 2003.

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

Koch, B. P., Dittmar, T., Witt, M. and Kattner, G.: Fundamentals of molecular formula assignment to ultrahigh resolution mass data of natural organic matter, *Analytical Chemistry* **79**, 1758-1763, 2007.

Koch, B. P., Witt, M. R., Engbrodt, R., Dittmar, T. and Kattner, G.: Molecular formulae of marine and terrigenous dissolved organic matter detected by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry, *Geochimica et Cosmochimica Acta* **69**, 3299-3308, 2005.

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

Kujawinski, E.B.: Electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI FT-ICR MS): Characterization of complex environmental mixtures. *Environ. Forensics*, **3**, 207-216, 2002.

Kujawinski, E. B., Longnecker, K., Blough, N. V., Del Vecchio, R., Finlay, L., Kitner, J. B. and Giovannoni, S. J.: Identification of possible source markers in marine dissolved organic matter using ultrahigh resolution mass spectrometry. *Geochimica et Cosmochimica Acta* **73**, 4384-4399, 2009.

Lam, B., Alaee, M., Lefebvre, B., Moser, A., Williams, A. and Simpson, A. J.: Major Structural Components in Freshwater Dissolved Organic Matter, *Environmental Science and Technology*, **41**, 2840-2847, 2007.

Lechtenfeld, O. J., Kattner, G., Flerus, R., McCallister, S. L., Schmitt-Kopplin, P., Koch, B. P.: Molecular transformation and degradation of refractory dissolved organic matter in the Atlantic and Southern Ocean, *Geochimica et Cosmochimica Acta*, **126**, 321-337, 2014.

Liu, P., Xu, C., Shi, Q., Pan, N., Zhang, Y., Zhao, S., Chung, K. H.: Characterization of Sulfide Compounds in Petroleum: Selective Oxidation Followed by Positive-Ion Electrospray Fourier Transform Ion Cyclotron Resonance Mass Spectrometry, *Analytical Chemistry* **82**, 6601-6606, 2010.

Maie, N., Yang, C.-Y., Miyoshi, T., Parish, K. and Jaffé, R.: Chemical characteristics of dissolved organic matter in an oligotrophic subtropical wetland/estuarine ecosystem. *Limnology & Oceanography* **50**, 23-35, 2005.

Maie, N., Parish, K., Watanabe, A., Knicker, H., Benner, R., Abe, T., Kaiser, K. and Jaffé, R.: Chemical characteristics of dissolved organic nitrogen in an oligotrophic subtropical coastal ecosystem, *Geochimica et Cosmochimica Acta* **70**, 4491-4506. 2006.

Martinez, A. T., Rencoret, J., Marques, G., Gutierrez, A., Ibarra, D., Jimenez-Barbero, J., del Rio, J. C.: Monolignol acylation and lignin structure in some nonwoody plants: A 2D NMR study. *Phytochemistry* **69**, 2831-2843, 2008.

McCarthy, J. M., Gumbricht, T., McCarthy, T., Frost, P., Wessels, K. and Seidel, F.: Flooding patterns of the Okavango wetland in Botswana between 1972 and 2000, *Ambio* **32**, 453–457, 2003.

McCaul, M. V., Sutton, D., Simpson, A. J., Spence, A., McNally, D. J., Moran, B.W., Goel, A., O'Connor, B., Hart, K. and Kelleher, B. P.: Composition of dissolved organic matter within a lacustrine environment, *Environmental Chemistry* **8**, 146–154, 2011.

McKnight, D. M., Boyer, E. W., Westerhoff, P. K., Doran, P. T., Kulbe, T. and Andersen, D. T.: Spectorflourometric characterization of dissolved organic matter for indication of precursor organic material and aromaticity. *Limnology & Oceanography* **46,** 38–48. 2001.

Minor, E. C., Swenson, M., Mattson, B. M. and Oyler, A. R.: Structural characterization of dissolved organic matter: A review of current techniques for isolation and analysis. *Environmental Science: Processes and Impacts*, DOI: 10.1039/c4em00062e, 2014.

Minor, E. C., Steinbring, C. J., Longnecker, K. and Kujawinski, E.B.: Characterization of dissolved organic matter in Lake Superior and its watershed using ultrahigh resolution mass spectrometry, *Organic Geochemistry*, **43**, 1-11, 2012.

Mladenov, N., McKnight, D. M., Macko, S. A., Norris, M., Cory, R. M. and Ramberg, L.: Chemical characterization of DOM in channels of a seasonal wetland. *Aquatic Sciences* **69**, 456–471. 2007.

Muller, H., Adam, F. M., Panda, S. K., Al-Jawad, H.-H., Al-Hajji, A. A.: Evaluation of Quantitative Sulphur Speciation in Gas Oils by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: Validation by Comprehensive Two-Dimensional Gas Chromatography. *J. Am. Mass Spectrom.* **23**, 806-815, 2012.

Ohno, T.: Fluorescence inner-filtering correction for determining the humification index of dissolved organic matter. *Environmental Science and Technology* **36**, 742–746, 2002.

Panagiotopoulos, C., Repeta, D. J. and Johnson, C. G.: Characterization of methyl sugars, 3-deoxysugars and methyl deoxysugars in marine high molecular weight dissolved organic matter, *Organic Geochemistry* **38**, 884-896. 2007.

Perdue, E.M., Hertkorn, N. and Kettrup, A.: Substitution patterns in aromatic rings by increment analysis. Model development and application to natural organic matter. *Analytical Chemistry* **79**, 1010-1021, 2007.

Perdue, E.M. and Ritchie, J. D.: Dissolved Organic Matter in Fresh Waters. In *Treatise on Geochemistry* (eds. H.D. Holland and K.K. Tuerekian); Volume **5**: *Surface and Ground Water; Weathering, Erosion and Soils* (ed. J.I. Drever), 273-318, 2003.

Purcell, J. M., Jayal, P., Kim, D.-G., Rodgers, R. P., Hendrickson, C. L., Marshall, A. G.: Sulphur Speciation in Petroleum: Atmospheric Pressure Photoionization or Chemical Derivatization and Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry, *Energy & Fuels* **21**, 2869-2874, 2007.

Ralph, J., Hatfield, R. D., Piquemal, J., Yahiaou,i N., Pean, M., Lapierre, C. and Boudet, A. M.: NMR characterization of altered lignins extracted from tobacco plants down-regulated for lignification enzymes cinnamyl-alcohol dehydrogenase and cinnamoyl-CoA reductase, *P. Natl. Acad. Sci. USA* **95**, 12803-12808, 1998.

Repeta, D. J., Quan, T. M., Aluwihare, L. I. and Accardi, A.: Chemical characterization of high molecular weight dissolved organic matter in fresh and marine waters. *Geochimica et Cosmochimica Acta* **66**, 955–962, 2002.

Schmitt-Kopplin, Ph., Gelencsér, A., Dabek-Zlotorzynska, E., Kiss, G., Hertkorn, N., Harir, M., Hong, Y. and Gebefügi, I.: Analysis of the unresolved organic fraction in atmospheric aerosols with ultrahigh resolution mass spectrometry and nuclear magnetic resonance spectroscopy: Organosulfates as photochemical smog constituents, *Analytical Chemistry* **82**, 8017–8026, 2010.

Sleighter, R. L. and Hatcher, P. G.: The application of electrospray ionization coupled to ultrahigh resolution mass spectrometry for the molecular characterization of natural organic matter, *J. Mass Spectrom.* **42**, 559-574, 2007.

Spencer, R. G. M., Hernes, P. J., Aufdenkampe, A. K., Baker, A., Gulliver, P., Stubbins, A., Aiken, G., Dyda, R. Y., Butler, K. D., Mwambai, V. L., Mangangu, A. M., Wabakanghanzi, J. N. and Six, J.: An initial investigation into the organic matter biogeochemistry of the Congo River. *Geochimica et Cosmochimica Acta* **84**, 614-626, 2012.

Stubbins, A., Lapierre, J.-F., Berggren, M., Prairie, Y. T., Dittmar, T. and del Giorgio P.A.: What's in an EEM? Molecular signatures associated with dissolved organic fluorescence in boreal Canada, *Environ. Sci. Technol.* **48**, 10598-10606, 2014.

Stubbins, A., Niggemann, J. and Dittmar, T.: Photo-lability of deep ocean dissolved black carbon, *Biogeosciences Discussion* **9**, 485–505, 2012.

Tfaily, M. M., Corbett, J. E., Lin, X., Chanton, J. P., Glaser, P. H., Cawley, K. M., Jaffé, R. and Cooper, W. T.: Variations in the composition of porewater dissolved organic matter across ecological and hydrological gradients in a northern peatland, *Photochem. and Photobiol.* **91**, 684-695, 2015.

Wagner, S., Dittmar, T. and Jaffé, R.: Molecular characterization of dissolved black nitrogen by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry, *Organic Geochemistry* **79**, 21-30, 2015.

Wagner, S., Jaffé, R., Cawley, K., Dittmar, T. and Stubbins, A.: Associations between the molecular and optical properties of dissolved organic matter in the Florida Everglades: a model coastal wetland system. *Frontiers in Marine Biogeochemistry.* DOI:10.3389/fchem.2015.00066, 2015.

Weishaar, J. L, Aiken, G. R, Bergamaschi, B. A., Fram, M. S, Fujii, R. and Mopper, K.: Evaluation of specific ultraviolet absorbance as an indicator of the chemical composition and reactivity of dissolved organic carbon, *Environmental Science & Technology* **37**, 4702–4708, 2003.

Wen, J.-L., Sun, S.-L., Xue, B.-L., Sun, R.-C.: Recent Advances in Characterization of Lignin Polymer by Solution-State Nuclear Magnetic Resonance (NMR) Methodology, *Materials* **6**, 359-391, 2013.

Wolski, P., Savenije, H. H. G., Murray-Hudson, M. and Gumbricht, T.: Modelling of the flooding in the Okavango Delta, Botswana, using a hybrid reservoir-GIS model, *Journal of Hydrology* **331**, 58–72, 2006.

Woods, G. C., Simpson, M. J., Koerner, P. J., Napoli, A., and Simpson, A. J.: HILIC-NMR: towards the identification of individual molecular components in dissolved organic matter. *Environmental Science & Technology*, **45**, 3880-3886. 2011.

Woods, G. C., Simpson, M. J., Simpson, A. J.: Oxidized sterols as a significant component of dissolved organic matter: Evidence from 2D HPLC in combination with 2D and 3D NMR spectroscopy *Water Research* **46**, 3398-3408, 2012.

Yamashita, Y. and Jaffé, R.: Characterizing the Interactions between Trace Metals and Dissolved Organic Matter Using Excitation-Emission Matrix and Parallel Factor Analysis. *Environmental Science & Technology* **42**, 7374-7379. 2008.

Yamashita, Y. and Tanoue, E.: Chemical characterization of protein-like fluorophores in DOM in relation to aromatic amino acids, *Marine Chemistry* **82**, 255–271, 2003.

Yamashita, Y., Scinto, L., Maie, N. and Jaffé, R.: Assessing the environmental dynamics of dissolved organic matter in an oligotrophic subtropical wetland by optical means, *Ecosystems* **13**, 1006-1019, 2010.

Yuan, T.-Q., Sun, S.-N., Xu, F., Sun, R.-C.: Characterization of Lignin Structures and Lignin-Carbohydrate Complex (LCC) Linkages by Quantitative ^{13}C and 2D HSQC NMR Spectroscopy, *Journal of Agricultural and Food Chemistry*, **59**, 10604-10614. 2011.

Yelle, D. J., Ralph, J. and Frihart, C. R.: Characterization of nonderivatized plant cell walls using high-resolution solution-state NMR spectroscopy, *Magnetic Resonance in Chemistry* **46**, 508-517, 2008.

Zhang, F., Harir, M., Moritz, F., Zhang, J., Witting, M., Wu, Y., Schmitt-Kopplin, P., Fekete, A., Gaspar, A. and Hertkorn, N.: Molecular and structural characterization of dissolved organic matter during and post cyanobacterial bloom in Taihu by combination of NMR spectroscopy and FTICR mass spectrometry, *Water Research* **57**, 280-294, 2014.

Figure Captions:

Fig. 1. Box plots of the relative abundance (%) of PARAFAC components for the FCE, OKA, and PAN regions. Top of the blue box represents the 75th percentile, middle of the blue box is the median, lower edge of blue box is the 25th percentile, top of the black bar is the maximum value and bottom of the black bar is the minimum value. ENP = Everglades National Park; WCA2 = Water Conservation Area 2.

Fig. 2. ^{1}H NMR spectra of six wetland SPE-DOM (500 MHz; CD_3OD); overlay: intensities are normalized to total NMR resonance area in the entire chemical shift range shown (δ_H = 0-10 ppm), with exclusion of residual water and methanol NMR resonances (Fig. S1). (A) entire NMR spectrum (δ_H = 0-10 ppm), with section of unsaturated protons (δ_H = 5 - 10 ppm), highlighted in orange; (B) section of unsaturated protons (δ_H = 5 - 10 ppm); (C) section of aliphatic protons (δ_H = 0 - 5 ppm); highlighted in purple colour: vertical expansion of functionalized aliphatic compounds, associated also with CRAM (carboxyl-rich alicyclic compounds).

Fig. 3. TOCSY NMR spectra (800 MHz, CD_3OD) of wetland SPE-DOM, with (A) FCE-S, (B) SPE-DOM, and (C) PAN-L and (D) OKA-S SPE-DOM. Panel (A) depicts TOCSY cross peaks between aliphatic protons (X-C_{sp3}**H**-C_{sp3}**H**-X; X: C, O), whereas panels (B-D) depict TOCSY cross peaks between unsaturated protons (X-C_{sp2}**H**-C_{sp2}**H**-X; X: C, O). Section a: **H$_3$**C-C_n-X cross peaks, with n = 1 (δ_H > 3) and n > 1 (δ_H < 3); where X is any heteroatom, likely oxygen; section b: -C-C**H**-C**H**-C_n-X-, intra-aliphatic cross peaks; section c: α,β-unsaturated and conjugated double bonds: **H**C_{olefin}=C_{olefin}**H**-(C=O)-X; section d: polarized α,β-unsaturated double bonds: **H**C_{olefin}=C_{olefin}**H**-(C=O)-X; section e: congested fjord region in polycyclic aromatics; section f: aromatics **H**$C_{aromatic}$-$C_{aromatic}$**H** with ortho or/and para oxygenated substituents (classic aromatic substitution of DOM); section g: condensed and strongly electron withdrawing aromatics **H**$C_{aromatic}$-$C_{aromatic}$**H** (multiply carboxylated, N-heterocycles); section h: (more extended) polycyclic aromatics, polycarboxylated aromatics, N-heterocycles. Panel D: Sections of chemical shift for substituted aromatics as proposed by SPARIA model (**s**ubstitution

patterns in **a**romatic **r**ings by **i**ncrement **a**nalysis): COR: electron withdrawing substituents; R: electroneutral substituents; OR: electron-donating substituents (Perdue et al., 2007).

Fig. 4. ^{1}H, ^{13}C HSQC NMR cross peaks of FCE-S; section of unsaturated (olefinic and aromatic) protons δ_H = 4…10.5 ppm. Assignment in analogy to South Atlantic SPE-DOM FMAX (Hertkorn et al., 2013) with key substructures denoted as follows: section a: anomeric CH in carbohydrates (sp^3-hybridized); section b: isolated olefins; section c: C-conjugated olefins, certain five membered N-, O- and S-heterocycles (δ_H < 6.5 ppm); section d: multiply oxygenated aromatics including oxygen heterocycles, lignin derivatives, syringyl units (S2/6); section e: phenols, classical oxygenated DOM aromatics, lignin derivatives, guaiacyl units (G2), certain admixture of carbonyl derivatives (likely carboxylic units), causing downfield ^{1}H NMR chemical shift (δ_H > 7.3 ppm); section f: classical DOM aromatic, lignin derivatives, guaiacyl units (G5/6), para-coumarate (C3/5); section g: classical DOM aromatics with high proportion of carboxylated units; at δ_H > 8 ppm: multiply carboxylated aromatics, classical PAH and certain six-membered nitrogen heterocycles; sterically uncongested PAH; section h: α,β-unsaturated double bonds for δ_C > 140 ppm, including double bonds adjacent to aromatics: C-**HC**$_{olefin}$=**C**$_{olefin}$**H**-(C=O), C_{ar}-X; section i: nitrogen heterocycles, heteroatom substituted polycyclic aromatics; section j: certain six-membered nitrogen heterocycles, very likely with more than one nitrogen. The green area highlights the HSQC cross peak region accessible to single benzene rings substituted by common electron withdrawing, neutral and electron-donating common substituents of natural organic matter; SPARIA: **s**ubstitution **p**atterns in **a**romatic **r**ings by **i**ncrement **a**nalysis (Perdue et al., 2007).

Fig. 5. Methylene (CH_2) selective ^{1}H, ^{13}C DEPT – HSQC NMR spectrum of SPE-DOM FCE-S with assignment of major substructures; general colours: CH_3: red; CH_2 green, and CH: gray; section a: C-**CH**$_3$ cross peaks; section b: C=C-**CH**$_3$ and -S**CH**$_3$ cross peaks; section c: acetate **H**$_3$**C**-C(=O)-O-C-; section d: C_2**CH**$_2$ cross peaks; section e: -C-**CH**$_2$-COOH cross peaks; section f: C_3**CH** cross peaks; section g: only methoxy (O**CH**$_3$) cross peaks are shown here; see insert: section g_1: **H**$_3$**C**OH (HD_2COD shows methine carbon); sections g_2 and g_3: aliphatic (g_2) and aromatic (g_3) methyl esters **H**$_3$**C**O-C(=O)-C-; section g_4 and g_5: aromatic (g_4) and aliphatic (g_5) methyl ethers **H**$_3$**C**O-C-C; section f: C_3**CH** cross peaks; section h: oxomethylene (O**CH**$_2$) cross

peaks, likely from carbohydrates; section i: OC_2**CH** cross peaks; section j: methylene bound to esters –C-$\mathbf{H_2C}$O-C(=O)-Z- (cf. main text).

Fig 6. Left panel: negative electrospray 12T FTICR mass spectra of Wetlands SPE-DOM (insert in figure 7F show an enlarged mass view of a mass range of 6.0 Da. Right panel: expansion of the mass segment m/z = 465.00-465.20 (Δm = 0.2 Da; asterisk in insert figure 7F), with assignment according to CHO, CHNO, CHOS and CHNOS molecular series. (A) OKA-L; (B) OKA-S; (C) PAN-L; (D) PAN-S; (E) FCE-L; (F) FCE-S.

Fig. 7. Van Krevelen diagrams of six wetlands SPE-DOM; (A) OKA-L; (B) OKA-S; (C) PAN-L; (D) PAN-S; (E) FCE-L; (F) FCE-S, obtained from negative electrospray 12T FTICR mass spectra. Only molecular assignments bearing combinations of C,-H,-O,-N, and –S atoms are shown; color coded according to molecular series as follows: CHO-blue, CHOS-green, CHNO-orange, CHNOS-red. Bubble areas reflect the relative intensities of respective mass peaks. Panel F: labels for CHOS compounds correspond to key molecules, section a: saturated sulfolipids; section b: unsaturated sulfolipids; section c: common CHOS compounds in DOM, possibly sulfonated carboxylic-rich alicyclic compounds (CRAM); d: aromatic black sulphur.

Fig. 8. Comparative analysis of van Krevelen diagrams derived from negative electrospray 12T FT-ICR mass spectra of all six wetlands SPE-DOM. (A) Clustering diagram based on the similarity values between the spectra of six wetlands SPE-DOM using Pearson correlation coefficient; (B) molecular compositions common to all six wetlands SPE-DOM, (C) unique molecular compositions common in FCE samples (FCE-L and FCE-S); (D) unique molecular compositions with high abundance in both PAN samples; (E) unique molecular compositions with high abundance in both OKA samples; (F) unique molecular compositions common in all four PAN and OKA. The aromaticity index AI (Koch and Dittmar, 2006) provided denotes single aromatic compounds for AI > 0.5 (bright blue triangle).

Fig. 9. Comparative analysis of (left) H/C vs. m/z and (right) H/C vs. O/C van Krevelen diagrams derived from negative electrospray 12T FTICR mass spectra of the two Florida Coastal Everglades SPE-DOM FCE-S and FCE-L (see also Fig. 8). (A) Molecular compositions with high abundance in Florida Coastal Everglades SPE-DOM FCE-S; section a: oxygen-deficient

(poly)aromatic black sulphur; CHNOS: suite of highly oxygenated CHNOS molecules; section b: common CHOS molecules in DOM; section c: saturated sulfolipids. The aromaticity index AI (Koch and Dittmar, 2006) provided in the upper right van Krevelen diagram denotes single aromatic compounds for AI > 0.5 (bright blue triangle) and polyaromatic compounds for AI > 0.67 (bright purple triangle); (B) Molecular compositions with high abundance in Florida Coastal Everglades SPE-DOM FCE-L; section d: a distinct set of oxygen-rich aromatic CHOS compounds, likely associated with ether-linked aromatic units; cf. text.

Supporting Online Information

Fig. S1. ^{1}H NMR spectra of six wetland SPE-DOM (CD_3OD; 500 MHz), acquired with solvent suppression and exclusion regions used in the computation of NMR section integrals and overlay NMR spectra (Fig. 2 and this figure) which denote HD_2COD and residual HDO, with section of unsaturated protons (δ_H > 5 ppm) vertically expanded. Intensities are normalized to 100% total integral in the entire chemical shift range shown (δ_H = 0…10 ppm). Fundamental substructures are indicated from higher to lower field (from right to left), (a) aliphatics, **H**CCC; (b) "acetate-analogue", **H**$_3$CC(=O)-O-; (c) carboxyl-rich alicyclic materials (CRAM), **H**C(C)-COX; (d) "carbohydrate-like" and methoxy, **H**CO; (e) olefinic, **H**C=C; and (f) aromatic NMR resonances **H**C_{ar} (cf. text). Further division of unsaturated protons provided (f_1) polycyclic and polycarboxylated aromatics as well as six-membered nitrogen heterocycles (δ_H > 8 ppm); (f_2) electron withdrawing substituents (COX; Perdue et al., 2007; $\delta_H \approx$ 7.3 – 8.0 ppm); (f_3) electroneutral substituents (alkyl, H, R; $\delta_H \approx$ 7.0 – 7.3 ppm); (f_4) electron-donating substituents (OR, OH, phenolics; $\delta_H \approx$ 6.5 – 7.0 ppm); (e_1) polarized and conjugated olefins; ($\delta_H \approx$ 5.5 – 6.5 ppm); (e_2) isolated olefins ($\delta_H \approx$ 5.0 – 5.5 ppm), with conceivable contributions from anomeric protons and ester groups (cf. discussion of 2D NMR spectra).

Fig. S2. ^{1}H NMR spectra of wetland SPE-DOM (CD_3OD; 500 MHz). Ssimilarity assessment by means of (panel A) cluster analysis (Pearson) and (panel B) PCA as well as (panels C, D) computed difference ^{1}H NMR spectra of 3 wetland SPE-DOM pairs (L-S: long minus short hydroperiod) as derived from 0.001 ppm buckets in area-normalized ^{1}H NMR spectra; used chemical shift range :δ_H = 9.5 – 0.5 ppm, with exclusion of residual water and methanol NMR resonances. Panels E, F, G: Manual overlay according to identical ^{1}H NMR section integral in the respective regions of ^{1}H NMR chemical shift shown: (left column) entire NMR spectrum (δ_H = 0 - 10 ppm); (center column) section of unsaturated protons (δ_H = 5 - 10 ppm); (right column) section of aliphatic protons (δ_H = 0 - 5 ppm). Panel E: OKA; panel F: PAN, and panel G: FCE SPE-DOM. Sections f_n of unsaturated protons are denoted as provided in Fig. S1.

Fig. S3. ^{13}C NMR spectra of selected wetland SPE-DOM; full spectra computed with 35 Hz exponential line broadening; insert: section of methoxy peaks (δ_C = 51-59 ppm; computed with 2 Hz line broadening); OKA-L and PAN-S: in $^{12}CD_3OD$ at B_0 = 11.7 T; FCE in CD_3OD at B_0 = 18.8 T.

Fig. S4. ^{1}H, ^{13}C HSQC NMR spectrum of SPE-DOM FCE-S, with regions shown in figures: (A) chemical environments of sp^3-hybridized carbon (aliphatic CH_n units; Fig. 6); (B) chemical environments of sp^2-hybridized carbon (unsaturated, i. e. olefinic and aromatic CH units; Fig. 5). Sensitivity enhanced apodization is used to emphasize less abundant sp^2-hybridized carbon (overall HSQC cross peak integral <4% of aliphatic units) environments at the cost of resolution in case of aliphatic CH_n units (n = 1 – 3).

Fig. S5. Overlay of ^{1}H, ^{13}C HSQC NMR spectra of SPE-DOM FCE-S (dark blue) and South Atlantic SPE-DOM at fluorescence maximum (48 mg, FMAX; orange: Hertkorn et al., 2013), together with region of HSQC NMR cross peaks accessible for single aromatic rings with full range of electron-withdrawing (COX), electroneutral (R, H) and electron donating substitution (OH, OR), shown in green color (SPARIA: Perdue et al., 2007). Wetland SPE-DOM shows more exhaustive coverage of single aromatic rings from contributions of multiply oxygenated aromatics (δ_H < 7 ppm; δ_C < 120 ppm), likely originating from plant phenolics but also from polycarboxylated aromatics and PAH derivatives (δ_H > 8.5 ppm). In contrast, open ocean SPE-DOM FMAX exhibits a larger abundance as well as overall chemical diversity of α,β unsaturated and C-conjugated olefins, likely originating from marine natural products; for assignment of HSQC cross peaks, see Figs. 4 and 5, and Hertkorn et al., 2013.

Fig. S6. Further evaluation of aliphatic spin systems of wetland SPE-DOM FCE-L. Panel A: overall ^{1}H, ^{1}H JRES NMR spectrum with sections a_1, a_2, a_3, denoting the area of panels B, C, D, which display ^{1}H NMR projections along JRES and ^{1}H, ^{13}C DEPT HSQC NMR spectra (copied from Fig. 6); panel B: section of O**C**<u>**H**</u> aliphatic units, demonstrating (section b_1) presence of intense JRES cross peaks from O<u>**CH$_3$**</u> groups, with absence of J_{HH} splittings; panel C: section of aliphatic CC<u>**H**</u> units, with dominance of HOOC-C<u>**H$_n$**</u>-CH_2- units (triplet J_{HH} splitting; n = 1, 2) over HOOC-C<u>**H$_n$**</u>-$CH_,$- units (doublet J_{HH} splitting; n = 1, 2) shown in section c_1; section c_2

indicates panel D; panel D: section of aliphatic CCC**H** units, showing a remarkable clustering of **H**$_3$C-CH- units at $\delta_{\mathbf{H}}$: 1.0 – 1.4 ppm, which indicate pronounced aliphatic branching in section d_1 (doublet splitting from J_{HH}), whereas ethyl groups **H**$_3$C-CH_2- dominate the low field section $\delta_H < 1$ ppm (section d_2).

Fig. S7. Mass edited H/C ratios from negative electrospray 12T FTICR mass spectra of Wetlands SPE-DOM: (A) OKA-L; (B) OKA-S; (C) PAN-L; (D) PAN-S; (E) FCE-L; (F) FCE-S. Insert histograms show the number of assigned molecular compositions. Colour code for elemental compositions bearing combinations of C, H, O, N, and S atoms are defined as follows: blue (CHO), orange (CHNO), green (CHOS) and red (CHNOS). Bubble areas reflect the relative intensities of each mass peak.

Fig. S8. Comparative analysis of van Krevelen diagrams derived from negative electrospray 12T FTICR mass spectra derived from four Pantanal and Okavango SPE-OM only. (A) Clustering diagram based on the similarity values between the FTICR mass spectra of these four SPE-DOM; (B) Molecular compositions with rather high abundance in both Okavango SPE-DOM; (C) Molecular compositions with rather high abundance in both Pantanal SPE-DOM, with color code according to molecular series (cf. text). The bright blue triangle denotes aromatic compounds, with aromaticity index AI > 0.5 (Koch and Dittmar, 2006); see Fig. 9 and attendant discussion.

Table 1: DOC and optical properties of the six bulk water samples collected for SPE-DOM

	Sample	DOC (ppm)	$SUVA_{254}$	Abs_{254}	S_R	FI	TFI (QSU)	%C1	%C2	%C3	%C4	%C5	%C6	%C7
1	FCE-L	28.57	2.95	0.844	0.95	1.34	968.34	28%	9%	23%	12%	14%	3%	6%
2	FCE-S	9.67	2.72	0.263	0.98	1.44	360.11	34%	3%	15%	13%	15%	10%	5%
3	OKA-L	6.33	3.19	0.202	0.91	1.36	180.99	35%	1%	19%	11%	17%	7%	7%
4	OKA-S	9.87	2.98	0.294	0.97	1.33	158.16	31%	3%	20%	11%	16%	5%	10%
5	PAN-L	5.82	5.11	0.297	0.92	1.41	267.59	34%	2%	16%	10%	15%	8%	10%
6	PAN-S	6.60	4.49	0.296	0.91	1.39	270.05	37%	0%	20%	11%	16%	7%	6%

TFl = total fluorescence (QSU)
%CX = Relative abundance of PARAFAC component X
L and S indicate Long or Short hydroperiod

δ (^{1}H) [ppm]	**10.0 – 6.50**	**6.5 - 5.3**	**4.9 - 3.1**	**3.1 - 1.9**	**1.9 - 0.0**
key substructures	**H**$_{ar}$	C=C**H**, O_2C**H**	OC**H**	XCC**H**	CCC**H**
OKA-L	7.2	3.0	29.6	26.9	33.4
OKA-S	7.2	3.0	31.0	26.5	32.3
PAN-L	7.5	2.9	28.1	28.5	33.1
PAN-S	7.2	2.7	29.7	27.9	32.4
FCE-L	5.5	2.4	27.9	30.5	33.7
FCE-S	5.3	2.1	29.3	28.9	34.4

Table 2. ^{1}H NMR section integral for key substructures of natural organic matter (SPE-DOM) as computed from 0.001 ^{1}H NMR bucket NMR integrals (cf. Fig. S2); owing to distribution of HSQC cross peaks, the distinction between aromatic and olefinic molecules was placed at δ_H = 6.5 ppm (cf. Fig. 4).

Members of Molecular series	OKA-L	OKA-S	PAN-L	PAN-S	FCE-L	FCE-S
CHO compounds	1581 (57.6 %)	1772 (60.0 %)	1711 (58.8 %)	1844 (56.5 %)	1400 (37.2 %)	1201 (32.2 %)
CHOS compounds	266 (9.7 %)	207 (6.8 %)	211 (7.3 %)	260 (8.0 %)	1127 (29.9 %)	1400 (37.5 %)
CHNO compounds	893 (32.5 %)	1075 (35.2 %)	984 (33.8 %)	1151 (35.3 %)	864 (22.9 %)	761 (20.4 %)
CHNOS compounds	5 (0.2 %)	3 (0.1 %)	5 (0.2%)	8 (0.3 %)	375 (10.0 %)	372 (10.0 %)
total number of assigned mass peaks	2745	3057	2911	3263	3766	3734
total number of mass peaks	9830	10315	10588	10818	11692	10989
percent of mass peaks attributed to CHO, CHOS, CHNO and CHNOS compositions	28%	30%	27.5%	30%	32%	34%
average H [%]	40.66	41.41	39.34	39.54	40.82	41.14
average C [%]	38.12	37.75	38.77	38.73	37.14	39.39
average O [%]	20.69	20.33	21.45	21.27	21.06	18.36
average N [%]	0.38	0.38	0.40	0.42	0.37	0.27
average S [%]	0.16	0.13	0.04	0.04	0.61	0.83
computed average H/C ratio	1.06	1.09	1.01	1.02	1.09	1.04
computed average O/C ratio	0.54	0.53	0.55	0.54	0.56	0.46
computed average C/N ratio	101.2	100.6	97.4	93.2	99.7	145.5
computed average C/S ratio	246	285	877	885	60.5	47.3
mass weighted average	378.2	375.2	388.8	386.0	386.7	402.2

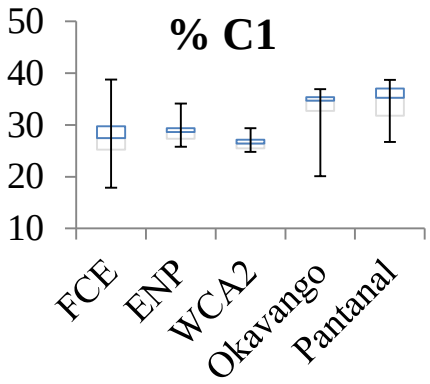
Table 3. Counts of mass peaks in wetland SPE-DOM as computed from negative electrospray (ESI) 12 T FTICR mass spectra for singly charged ions.

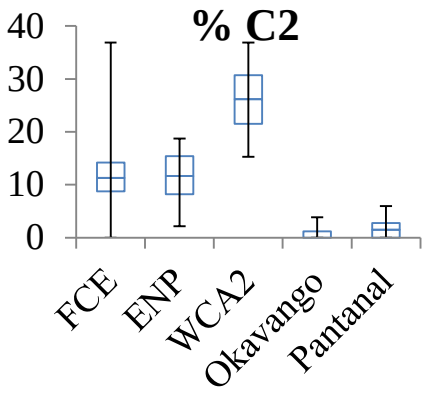
δ(^{13}C) ppm	**220-187**	**187-167**	**167-145**	**145-108**	**108-90**	**90-59**	**59-51**	**47-0**	**H/C ratio**	**O/C ratio**
key substructures	**C**=O	**C**OX	**C**$_{ar}$-O	**C**$_{ar}$-C,H	O_2**C**H	O**C**H	O**C**H_3	C**C**H		
FCE-S	2.5	13.8	2.5	10.3	2.4	14.2	12.6	41.7	**1.62**	**0.64**
FCE-L	1.6	13.8	2.2	9.5	0.9	11.9	11.6	48.5	**1.70**	**0.57**
OKA-L	2.2	14.8	5.2	17.2	2.4	14.7	7.9	35.6	**1.44**	**0.64**
PAN-S	1.8	14.0	5.0	17.2	2.7	14.5	6.9	37.9	**1.45**	**0.62**
NMR mixing model	**C=O**	**COOH**	**C_{ar}-O**	**C_{ar}-H**	**O_2CH**	**OCH**	**OCH_3**	**CH_2**		
H/C ratio	**0**	**1**	**0**	**1**	**1**	**1**	**3**	**2**		
O/C ratio	**1**	**2**	**1**	**0**	**2**	**1**	**1**	**0**		

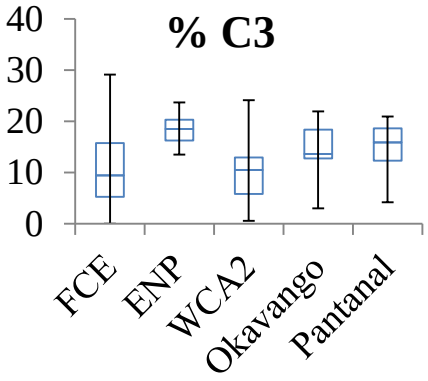
Table S1. (Top): ^{13}C NMR section integrals (percent of total carbon) and key substructures of wetland SPE-DOM. Bottom: Substructures used for basic NMR-derived reverse mixing model with nominal H/C and O/C ratios given (Hertkorn et al., 2013).

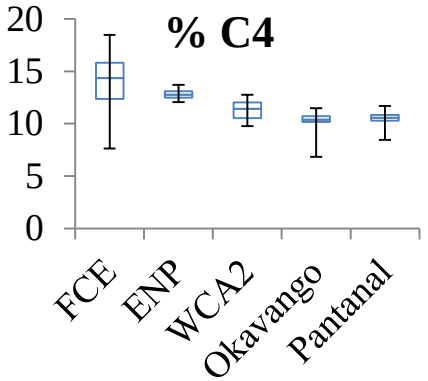
spectrum	Figure	PK	NS	AQ [ms]	D1 [ms]	NE	WDW1	WDW2	PR1	PR2	SPE-DOM [mg]
^{1}H NMR	2, S1	5TXI	512-1024	5000	10000	-	-	EM	-	1	3.7 – 9.5 mg
^{13}C NMR	S3	5D	74496	1000	14000	-	-	EM	-	35	OKA-L 4.7 mg
^{13}C NMR	S3	5D	44224	1000	14000	-	-	EM	-	35	PAN-S 4.2 mg
^{13}C NMR	S3	8QCO	23420	1000	19000	-	-	EM	-	35	FCE-L 9.5 mg
^{13}C NMR	S3	8QCO	3728	1000	19000	-	-	EM	-	35	FCE-S 9.1 mg
^{1}H, ^{1}H TOCSY	3	5TXI	24	1000	2500	1024	QS	EM	2.5	2.5	see caption
^{1}H, ^{1}H TOCSY	3	8QCI	12	1000	2500	1794	QS	EM	2.5	2.5	FCE-S 9.1 mg
^{1}H,^{13}C DEPT HSQC	4	8QCI	320	250	1250	256	QS	EM	2.5	2.5	FCE-S
^{1}H, ^{1}H JRES	S6	8QCI	3072	1000	500	49	QS	QS	0	0	FCE-S 9.1 mg
^{1}H, ^{13}C HSQC	5, S4, S5	8QCI	1600	250	1250	167	QS	EM	4	7.5	FCE-S

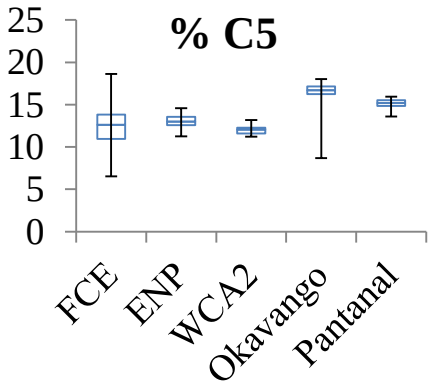
Table S2. Acquisition parameters of NMR spectra, shown according to figures. PK: probeheads used for acquisition of NMR spectra, 8QCI: cryogenic inverse geometry 5 mm z-gradient ^{1}H/^{13}C/^{15}N/^{31}P QCI probe (B_0 = 18.8 T); 8QCO: cryogenic classical geometry 3 mm z-gradient ^{1}H/^{13}C/^{15}N/^{31}P probe (B_0 = 18.8 T); 5TXI: cryogenic inverse geometry 5 mm z-gradient ^{1}H, ^{13}C, ^{15}N probe (B_0 = 11.7 T); 5D: cryogenic classical geometry 5 mm z-gradient ^{13}C, ^{1}H probe (B_0 = 11.7 T); NS: number of scans (for 2D NMR: F2); AQ: acquisition time [ms]; D1: relaxation delay [ms]; NE: number of F1 increments in 2D NMR spectra; WDW1, WDW2: apodization functions in F1/ F2 (EM/GM: line broadening factor [Hz]; QS: shifted square sine bell; SI: sine bell); PR1, PR2: coefficients used for windowing functions WDW1, WDW2, EM/GM are given in [Hz], SI/QS derived functions indicate shift by π/n.

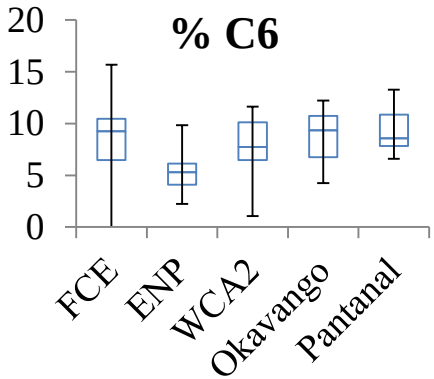

% C1
50
40
30
20
10
FCE
ENP
WCA2
Okavango
Pantanal

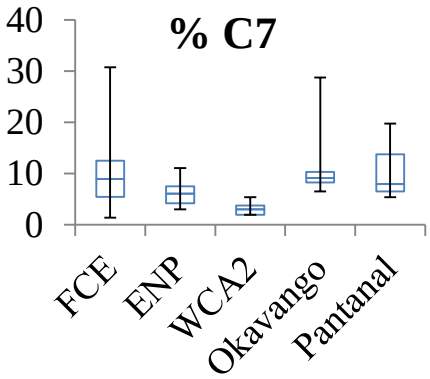

% C2
40
30
20
10
0
FCE
ENP
WCA2
Okavango
Pantanal

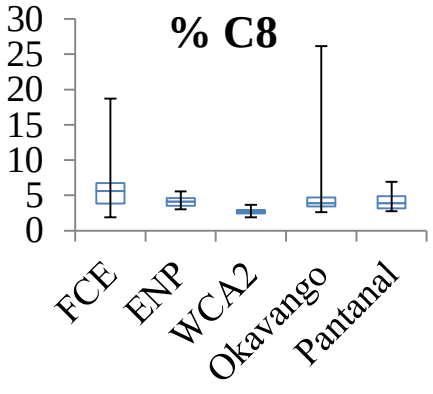

% C3
40
30
20
10
0
FCE
ENP
WCA2
Okavango
Pantanal

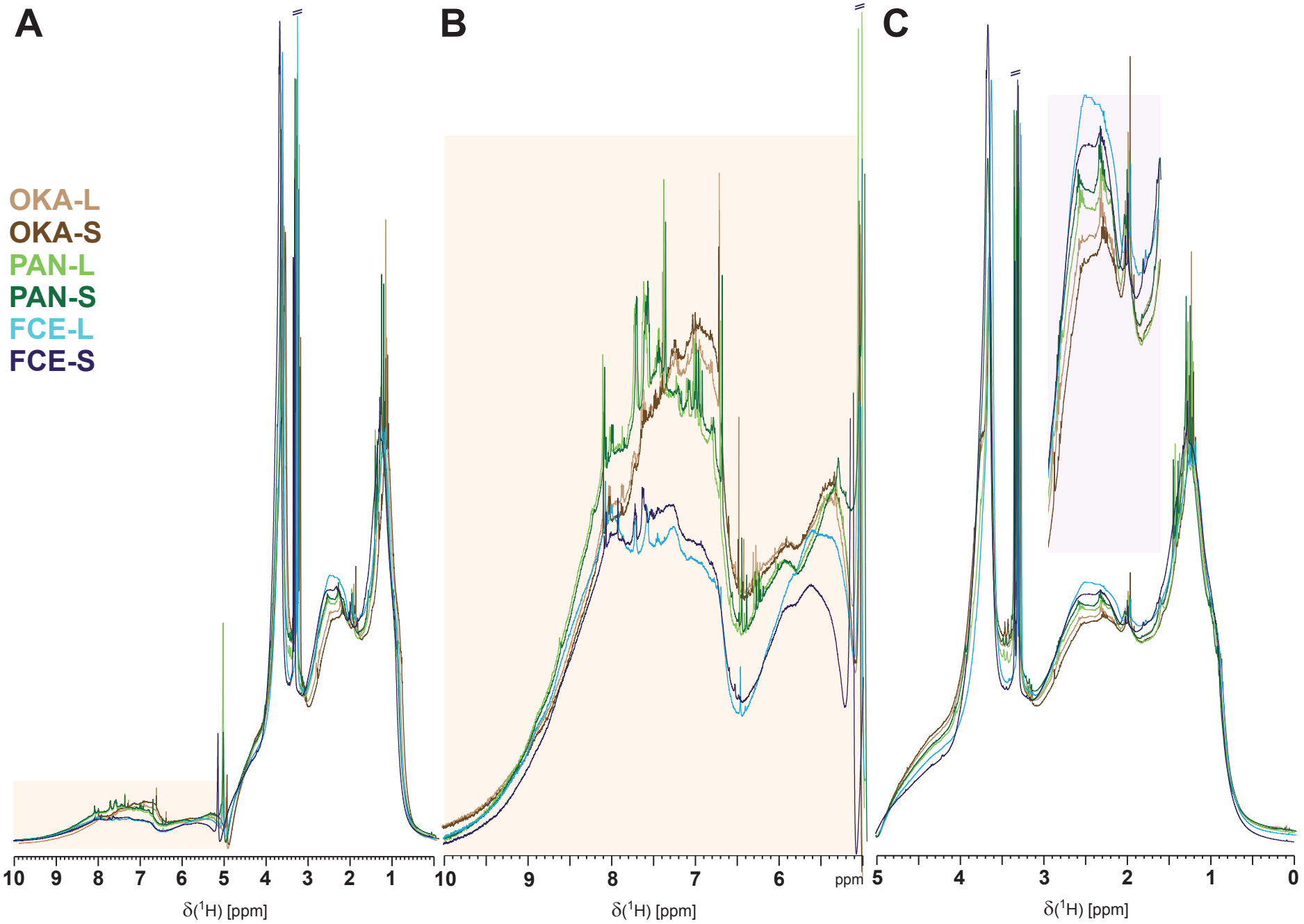

% C4
20
15
10
5
0
FCE
ENP
WCA2
Okavango
Pantanal

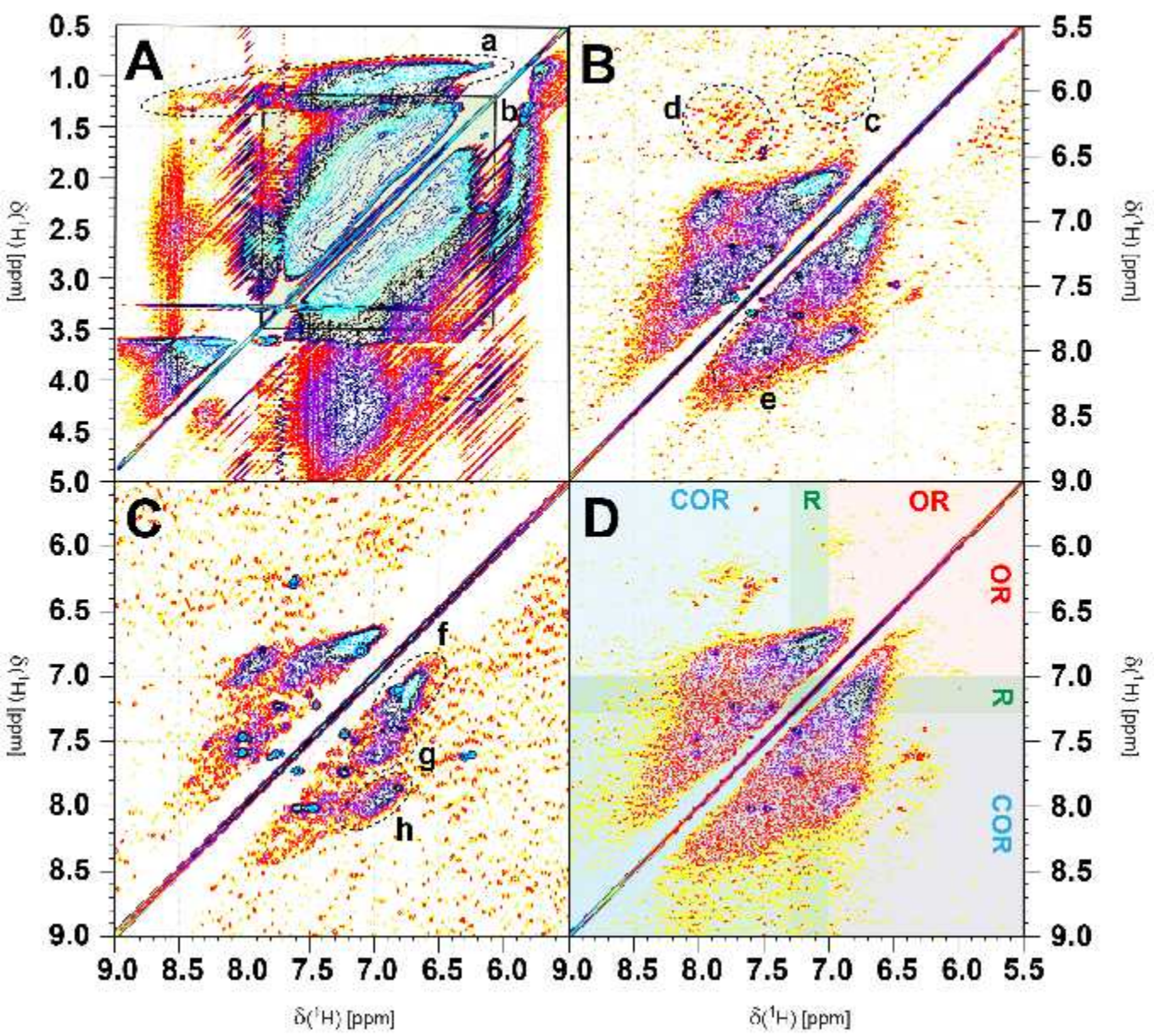

% C5
25
20
15
10
5
0
FCE
ENP
WCA2
Okavango
Pantanal

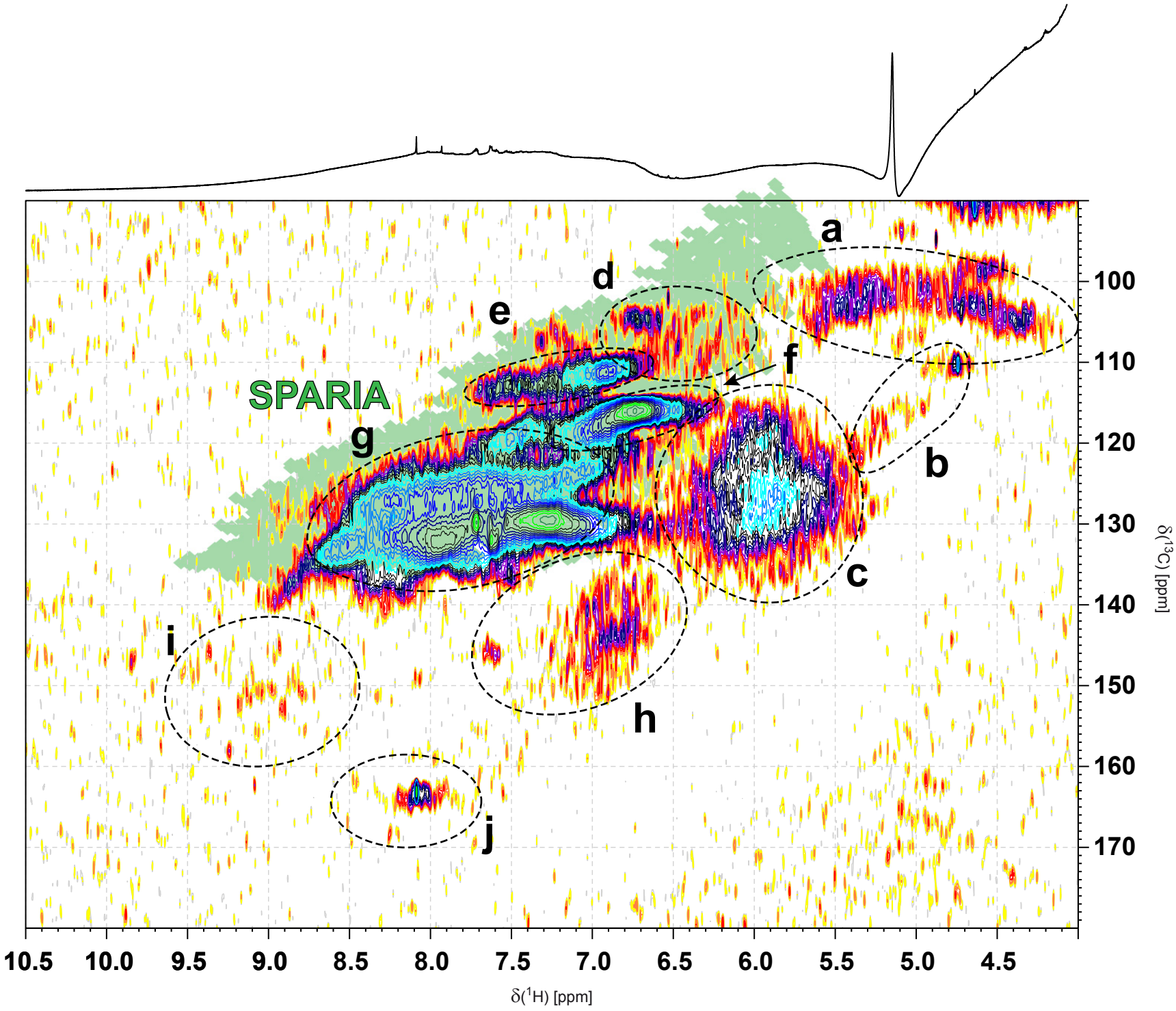

% C6
20
15
10
5
0
FCE
ENP
WCA2
Okavango
Pantanal

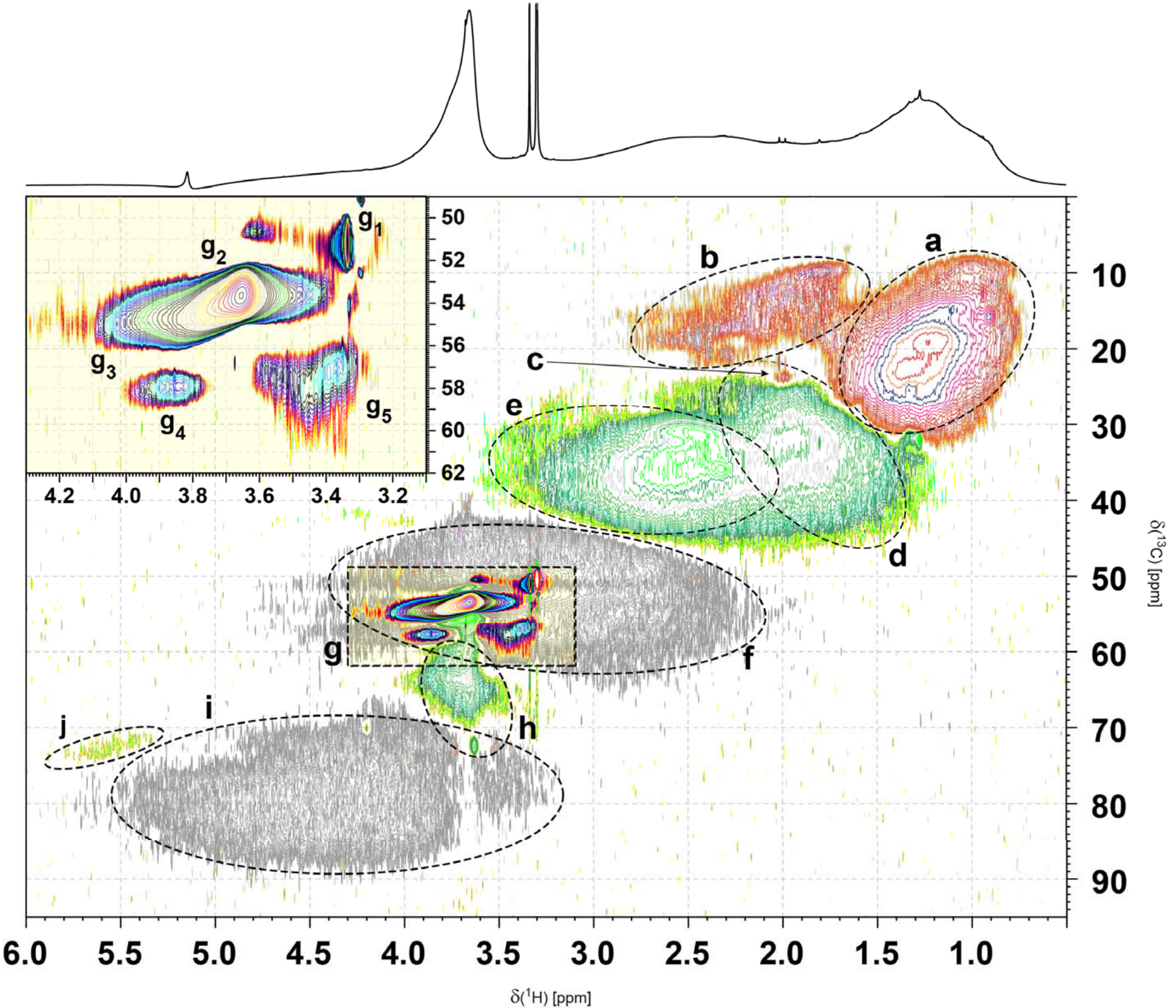

% C7
40
30
20
10
0
FCE
ENP
WCA2
Okavango
Pantanal

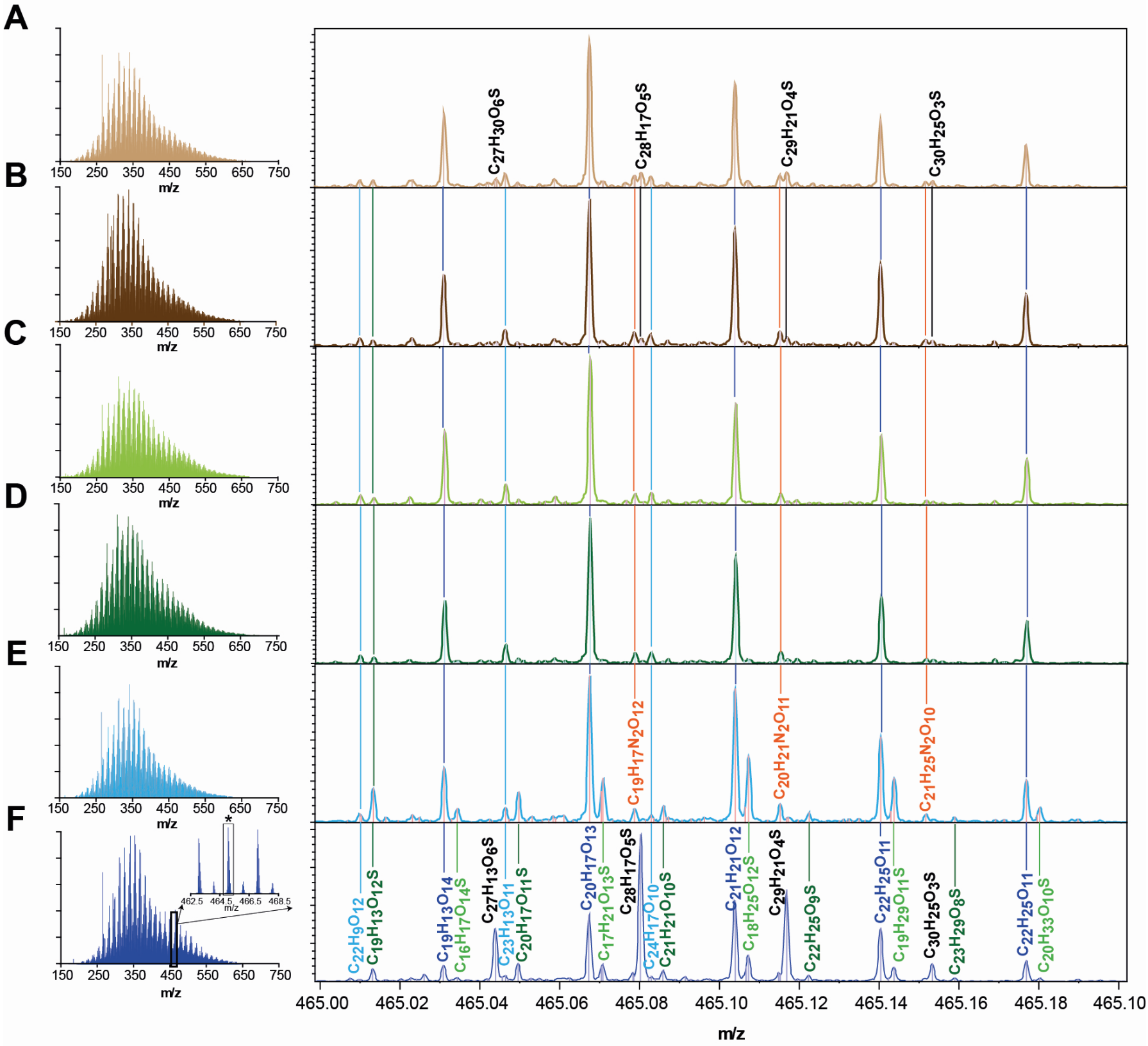

% C8
30
25
20
15
10
5
0
FCE
ENP
WCA2
Okavango
Pantanal

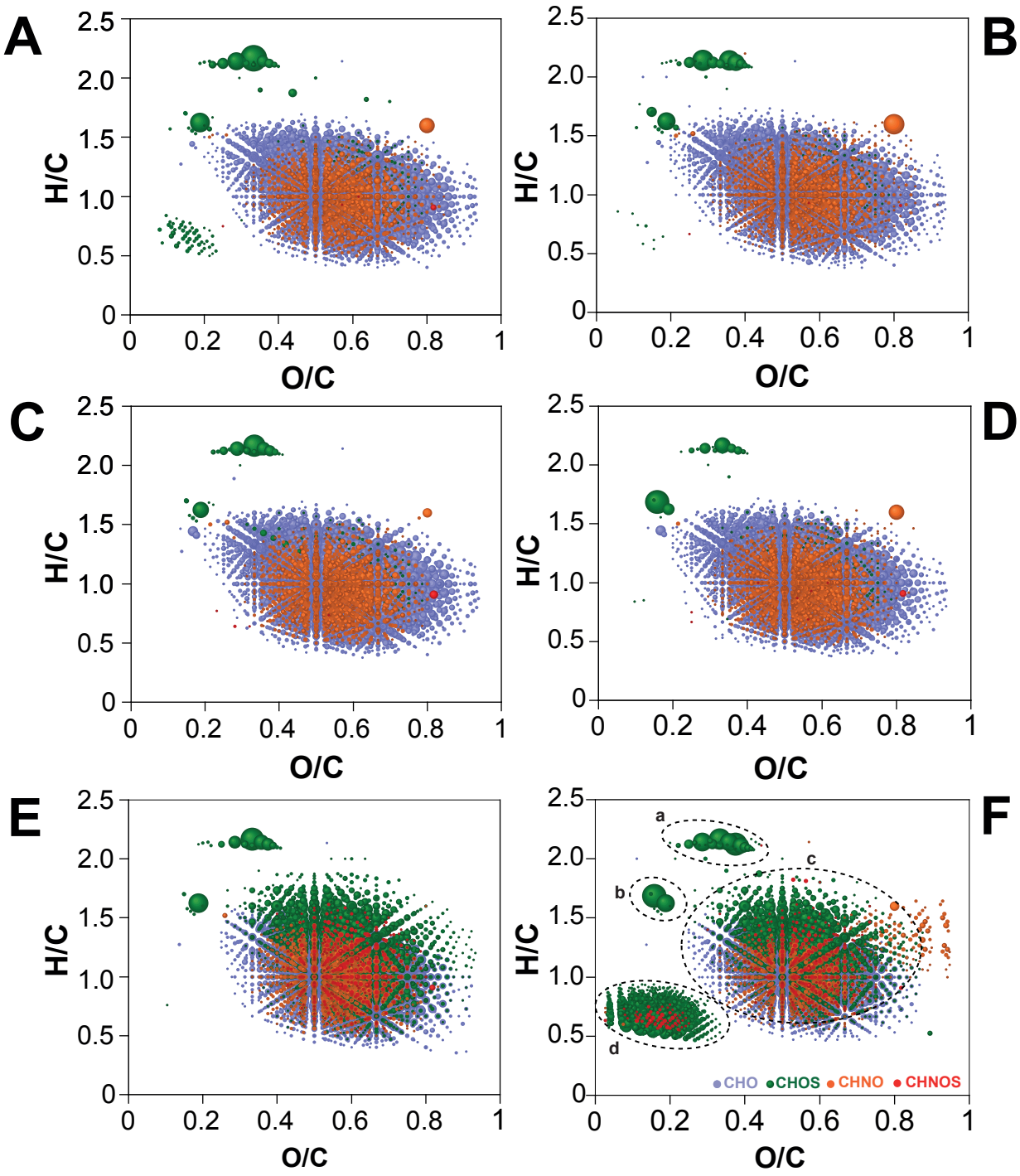

A
B
C
OKA-L
OKA-S
PAN-L
PAN-S
FCE-L
FCE-S
10 9 8 7 6 5 4 3 2 1
10 9 8 7 6
ppm
5 4 3 2 1 0
δ(¹H) [ppm]
δ(¹H) [ppm]
δ(¹H) [ppm]

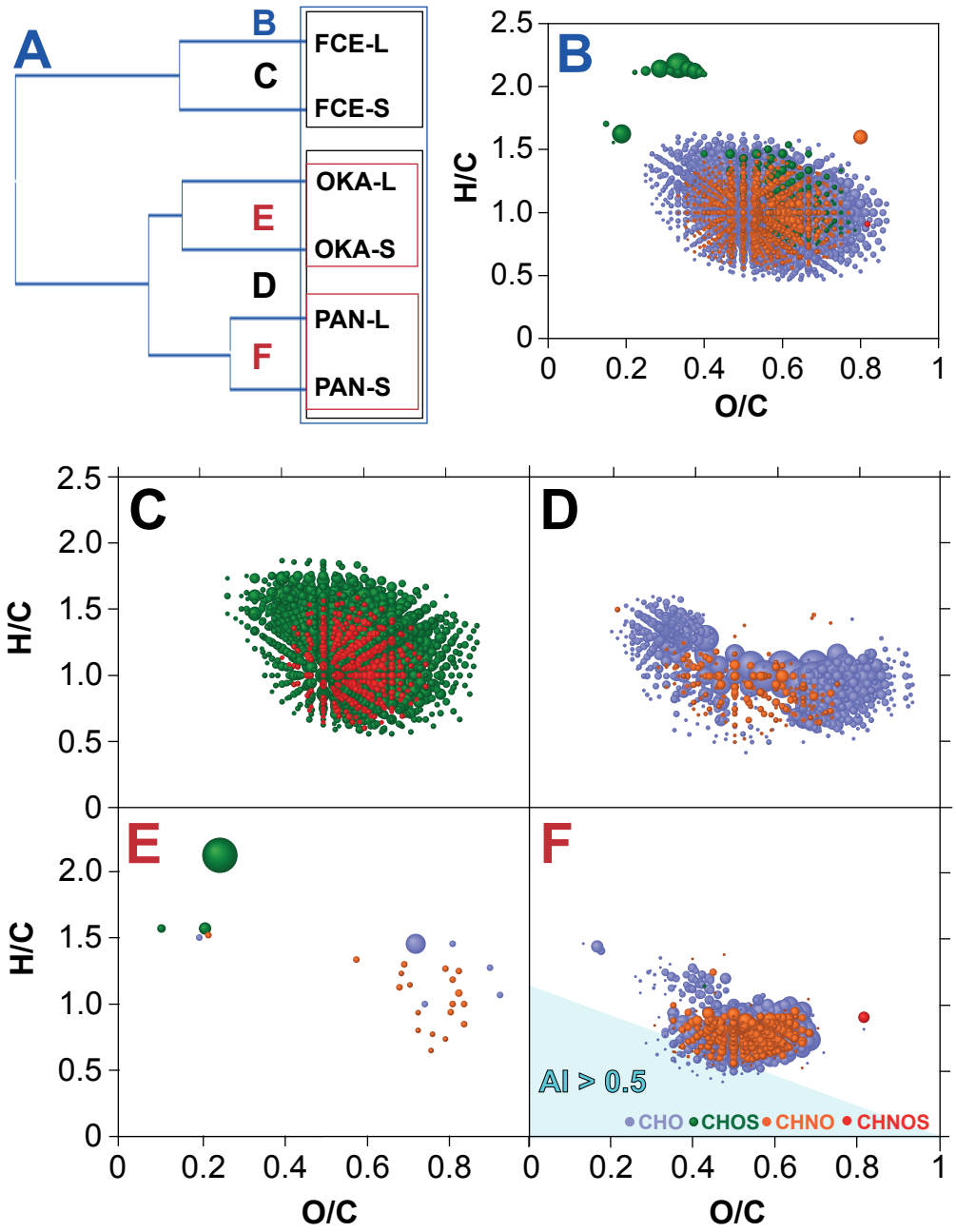

A
B
C
D
a
b
c
d
e
f
g
h
COR
R
OR
δ(¹H) [ppm]
δ(¹H) [ppm]
δ(¹H) [ppm]
δ(¹H) [ppm]

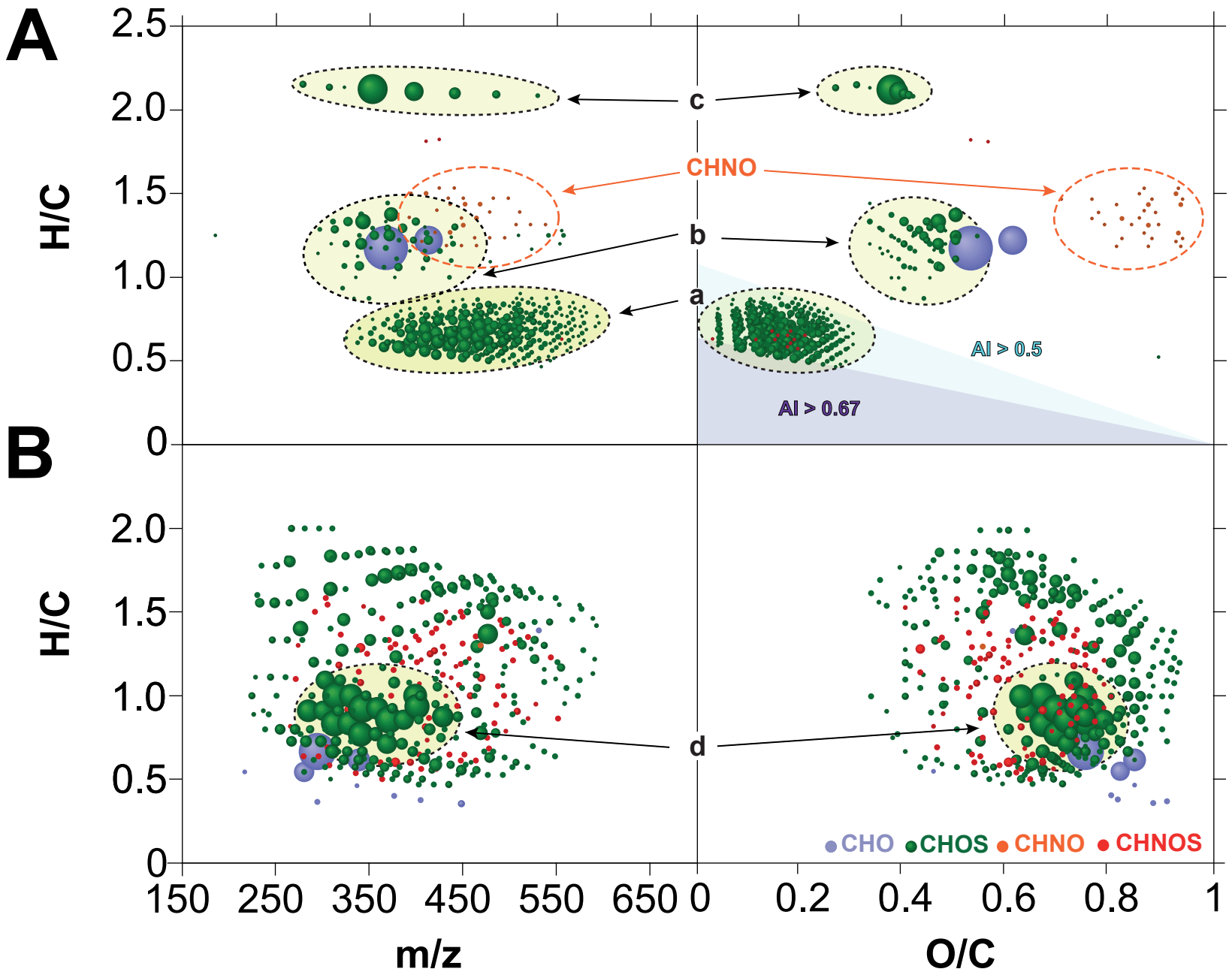

a
b
c
d
e
f
g
h
i
j
SPARIA
100
110
120
130
140
150
160
170
δ(13C) [ppm]
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
δ(1H) [ppm]


a
b
c
d
e
f
g
h
i
j
g1
g2
g3
g4
g5
δ(1H) [ppm]
δ(13C) [ppm]


A
B
C
D
E
F
150 250 350 450 550 650 750
m/z
462.5 464.5 466.5 468.5
m/z
$C_{27}H_{30}O_6S$
$C_{28}H_{17}O_5S$
$C_{29}H_{21}O_4S$
$C_{30}H_{25}O_3S$
$C_{19}H_{17}N_2O_{12}$
$C_{20}H_{21}N_2O_{11}$
$C_{21}H_{25}N_2O_{10}$
$C_{22}H_9O_{12}$
$C_{19}H_{13}O_{12}S$
$C_{19}H_{13}O_{14}$
$C_{16}H_{17}O_{14}S$
$C_{27}H_{13}O_6S$
$C_{23}H_{13}O_{11}$
$C_{20}H_{17}O_{11}S$
$C_{20}H_{17}O_{13}$
$C_{17}H_{21}O_{13}S$
$C_{28}H_{17}O_5S$
$C_{24}H_{17}O_{10}$
$C_{21}H_{21}O_{10}S$
$C_{21}H_{21}O_{12}$
$C_{18}H_{25}O_{12}S$
$C_{29}H_{21}O_4S$
$C_{22}H_{25}O_9S$
$C_{22}H_{25}O_{11}$
$C_{19}H_{29}O_{11}S$
$C_{30}H_{25}O_3S$
$C_{23}H_{29}O_8S$
$C_{22}H_{25}O_{11}$
$C_{20}H_{33}O_{10}S$
465.00 465.02 465.04 465.06 465.08 465.10 465.12 465.14 465.16 465.18 465.10
m/z


A
B
C
D
E
F
H/C
O/C
0
0.5
1.0
1.5
2.0
2.5
0.2
0.4
0.6
0.8
1
a
b
c
d
CHO
CHOS
CHNO
CHNOS


A
B
C
FCE-L
FCE-S
E
OKA-L
OKA-S
D
F
PAN-L
PAN-S
B
H/C
O/C
C
D
E
F
Al > 0.5
CHO
CHOS
CHNO
CHNOS


A
B
H/C
m/z
O/C
c
CHNO
b
a
d
AI > 0.5
AI > 0.67
CHO
CHOS
CHNO
CHNOS