



**HAL**  
open science

## Comparison of sample preparation techniques for the (-)ESI-FT-ICR-MS analysis of humic and fulvic acids

Aleksandar Goranov, Amanda Tadini, Ladislau Martin-Neto, Alberto Bernardi, Patricia Oliveira, José Pezzopane, Débora Milori, Stéphane Mounier, Patrick Hatcher

### ► To cite this version:

Aleksandar Goranov, Amanda Tadini, Ladislau Martin-Neto, Alberto Bernardi, Patricia Oliveira, et al.. Comparison of sample preparation techniques for the (-)ESI-FT-ICR-MS analysis of humic and fulvic acids. *Environmental Science and Technology*, In press, 10.1021/acs.est.2c01125. hal-03752314

**HAL Id: hal-03752314**

**<https://hal.science/hal-03752314>**

Submitted on 16 Aug 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Comparison of sample preparation techniques for the (-)ESI-FT-ICR-MS analysis of humic and fulvic acids

Aleksandar I. Goranov<sup>a,#</sup>, Amanda M. Tadini<sup>a,b,#</sup>, Ladislau Martin-Neto<sup>b</sup>, Alberto C. C. Bernardi<sup>c</sup>, Patricia P. A. Oliveira<sup>c</sup>, José R. M. Pezzopane<sup>c</sup>, Débora M.B.P.Milori<sup>b</sup>, Stéphane Mounier<sup>d</sup>, Patrick G. Hatcher<sup>a,\*</sup>

a) Department of Chemistry and Biochemistry, Old Dominion University, Norfolk, VA, USA 23529

b) Embrapa Instrumentação, Brazilian Agricultural Research Corporation - Embrapa, São Carlos, SP, Brazil 13560-970

c) Embrapa Pecuária Sudeste, Brazilian Agricultural Research Corporation - Embrapa, São Carlos, SP, Brazil 13560-970

d) University of Toulon, Aix Marseille University, CNRS/INSU, IRD, MIO UM 110, Mediterranean Institute of Oceanography, CS 60584, 83041 – Toulon, France

#Authors contributed equally to this work.

\*Corresponding author: Dr. Patrick G. Hatcher ([phatcher@odu.edu](mailto:phatcher@odu.edu))

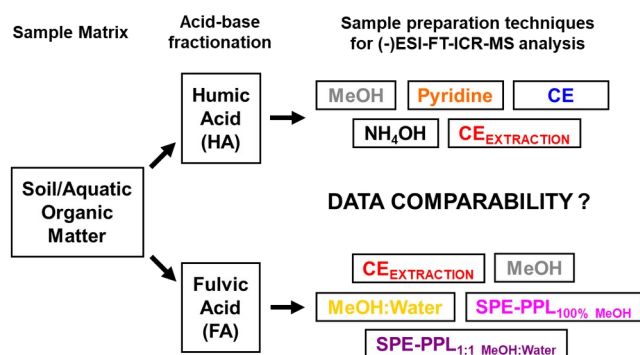
## ABSTRACT

Soil organic matter (SOM) plays a key role in the global carbon and nitrogen cycles. Soil biogeochemistry is commonly studied by extracting the base-soluble fractions of SOM: the acid-insoluble humic acid (HA) and acid-soluble fulvic acid (FA). Electrospray ionization – Fourier transform – ion cyclotron resonance – mass spectrometry (ESI-FT-ICR-MS) is commonly utilized for molecularly characterizing these fractions. Different sample preparation techniques exist for the analysis of HA and FA though questions remain regarding data comparability following different preparations. Comparisons of different sample preparation techniques revealed that the negative-mode ESI-FT-ICR-MS analytical window can be skewed to detect different groups of molecules, with primary differences in oxygenation, aromaticity, and molecular weight. It was also observed that HA and FA from soils versus an aquatic matrix behaved very differently and thus, we concluded that sample preparation techniques determined to be “most optimal” in our study are in no way universal, and we recommend that future studies of HA and FA involve such comparative studies for determining the most suitable sample preparation technique for their particular type of HA or FA matrices. This will enhance data comparability among different studies and environmental systems and ultimately, allow us to better understand the complex composition of environmental matrices.

**Subjects:** Soil organic matter, electrospray ionization (ESI), FT-ICR-MS, data comparability.

**Synopsis:** Minimal research exists on the comparability of mass spectral data of humic and fulvic acid extracts. This study compares different sample preparation techniques and discusses potential analytical biases to characterized molecular composition.

## Graphical Abstract:



## 1. INTRODUCTION

Humic substances (HS) are the biogeochemical products of the natural decomposition of plant litter by microorganisms<sup>1, 2</sup> and abiotic processes<sup>3</sup> contributing to the production of soil organic matter (SOM). HS account for 60-75 % of SOM<sup>4</sup> and have a unique structure dissimilar to other compound

51 groups present in SOM<sup>5-9</sup>. HS play an essential role in soil fertility, agricultural productivity, stabilization of  
52 aggregates, carbon sequestration, and as a whole are fundamental players in the biogeochemical cycling  
53 of SOM<sup>1, 2, 4, 10-16</sup>. HS have been a significant focus of soil science research and are still objects of intense  
54 discussion, research, and even controversy<sup>4, 11, 17-27</sup>. Therefore, understanding the molecular composition  
55 of HS is important for deciphering their enigmatic cycling as well as their involvement in carbon  
56 sequestration processes and mitigating climate change.

57 Ultrahigh resolution mass spectrometry, namely Fourier transform - ion cyclotron resonance -  
58 mass spectrometry (FT-ICR-MS), is a powerful technique providing a detailed view of the molecular  
59 composition of natural organic matter<sup>28-32</sup>, including HS<sup>3, 32-42</sup>. Nowadays, electrospray ionization (ESI),  
60 particularly in negative ionization mode (-), has become the norm in the environmental sciences<sup>43, 44</sup>. As  
61 SOM cannot be directly solubilized (a requirement for ESI analysis), HS are extracted using an alkali  
62 extraction (0.1 M NaOH). The unextracted remainder of SOM, the humin, is insoluble in most solvents  
63 making it difficult to analyze and thus, the humin fraction is rarely employed in SOM research. The NaOH-  
64 extracted HS are further fractionated into a fulvic acid (FA) fraction and a humic acid (HA) fraction based  
65 on the adjusting the solution to pH = 1 by the addition of HCl: the FA fraction remains fully solubilized  
66 while the HA fraction precipitates out. These two HS are then usually freeze-dried into a powder for  
67 sequential analyses<sup>45</sup>.

68 HA and FA fractions have been studied for decades and it has been found that they are  
69 analytically challenging samples for ESI methods: FA samples are difficult to purify and thus, are often rich  
70 in inorganic species – a severe issue, which can bias, and even prevent the ionization of molecules in the  
71 ESI source. The HA fraction, in contrast, has high aromatic content and tends to aggregate<sup>46</sup>, and  
72 sometimes it can also contain inorganic (e.g., Ca, Fe) constituents<sup>47, 48</sup> additionally skewing its solubility  
73 and ionizability. Once acid-precipitated and dried, the HA fraction is also difficult to re-solubilize back into  
74 solution.

75 Different sample preparation techniques have been developed to overcome the challenging  
76 matrices of HA and FA. HA fractions have been previously dissolved in neutral MeOH:Water mixtures<sup>49, 50</sup>,  
77 basified MeOH: Water mixtures with NaOH<sup>51</sup>, basified MeOH: Water mixtures with NH<sub>4</sub>OH<sup>35-37, 51</sup>, acidified  
78 MeOH: Water mixtures with acetic acid<sup>51</sup>, basified MeOH: Water mixtures with NaOH followed with cation-  
79 exchange<sup>52</sup>, neutral MeOH<sup>44</sup>, basified water with NH<sub>4</sub>OH<sup>44, 53-55</sup>, as well as in basified water with NaOH  
80 followed by cation-exchange and solvent extraction of the resin<sup>3</sup>. Commonly, NH<sub>4</sub>OH is added after HA  
81 had been already solubilized and thus, NH<sub>4</sub>OH only serves as an ionization enhancer<sup>35-37, 49, 50</sup>. The addition  
82 of NH<sub>4</sub>OH is necessary as HA fractions are often rich in poorly ionizable aromatic and condensed aromatic  
83 molecules<sup>56, 57</sup>. In other studies, however, NH<sub>4</sub>OH was simultaneously added when the powdered HA was  
84 mixed with the solubilization solution and thus, NH<sub>4</sub>OH served both as a basification agent and an  
85 ionization enhancer<sup>44, 51, 53-55</sup>. Recent work has shown that the use of NH<sub>4</sub>OH can induce nitrogen-  
86 incorporation and possibly bias the mass spectrometric analytical window<sup>58</sup>, but this has not been  
87 assessed in HA matrices.

88 FA fractions have been previously solubilized in neutral MeOH: Water mixtures<sup>43, 51</sup>, basified  
89 isopropanol:water mixtures with NH<sub>4</sub>OH<sup>51, 59</sup>, basified MeOH: Water mixtures with NH<sub>4</sub>OH<sup>38</sup>, acidified  
90 MeOH: Water mixtures with acetic acid<sup>59</sup>, basified water with NH<sub>4</sub>OH<sup>55</sup>. Solid-phase extraction using Bond  
91 Elut Priority Pollutant cartridges (SPE-PPL), a popular approach for obtaining high purity organic matter  
92 extracts from aquatic samples<sup>60</sup>, has also been employed<sup>61</sup>.

93 Clearly, there is a large diversity in the possible sample preparation techniques to solubilize HA  
94 and FA fractions in solution and prepare samples suitable for ESI-FT-ICR-MS analysis. Different sample  
95 preparation techniques have been previously discussed in the literature<sup>62</sup> though questions remain about  
96 the comparability of data from these different sample preparation techniques. The composition of solvent  
97 (e.g., 100% MeOH, MeOH:Water or isopropanol:water mixtures) and addition of modifiers (NH<sub>4</sub>OH, acetic  
98 acid, etc.) are controlling factors in the ionizability of molecules in the ESI source<sup>63, 64</sup>. It has been shown  
99 that when HA is analyzed in MeOH: Water mixtures of different ratios (e.g., 25:75, 50:50, 70:30), the  
100 highest quality spectra are obtained when MeOH is more than 50%<sup>53, 59, 65</sup>. Solvents and modifiers affect  
101 the molecular conformation, charge distribution, the stability of the formed ions, and consistency of  
102 ionization currents<sup>66-69</sup>. Additional factors include the pH of the solution, ratio of organic solvent:water, as  
103 well the quantity of ions (i.e., the ionic strength)<sup>59, 70, 71</sup>. Thus, variability in the sample preparation will affect  
104 the ionizability of molecules and thus likely affect the observed molecular composition of HA and FA using  
105 ESI-FT-ICR-MS.

106 Previous studies have shown that different solvents fractionate organic matter and further bias the  
107 analytical window of ESI-FT-ICR-MS in the context of sediments<sup>72</sup>, aerosols<sup>73</sup>, soils<sup>74</sup>, and petroleum and  
108 its fractions<sup>67, 69</sup>. However, in these studies the solvent was used both for obtaining solvent-extractable  
109 organic matter and for a carrier solvent in the ESI analysis. By contrast in HA and FA analysis, the humic  
110 substances are already extracted and only need to be solubilized.

111 In this manuscript, we assess how data for freeze-dried HA and FA compares when different  
112 sample preparation techniques are used. This is the first comprehensive study of sample preparation  
113 techniques for HA and FA fractions. Though these fractions are key SOM constituents and are commonly  
114 employed in soil science studies, this humic fractionation technique has started to appear in aquatic  
115 dissolved organic matter (DOM) studies too. This is likely because DOM partly originates from SOM  
116 leaching into aquatic environments establishing the connection between terrestrial (land) and aquatic  
117 interfaces. Different sample preparation techniques were tested on HA and FA fractions from several soils  
118 and one river. We investigate compositional differences using (-)ESI-FT-ICR-MS and assess how sample  
119 preparation and/or solvent composition affect the molecular results. Results from this study will enhance  
120 our understanding of HA and FA and create a framework for researchers to perform similar sample  
121 preparation comparative studies in order to obtain more representative molecular information about HA,  
122 FA, and other environmental samples.

123

## 124 2. EXPERIMENTAL SECTION

125

### 126 2.1 Soil samples and extraction of humic substances

127

128 Two haplorthox (red-yellow latosol) soil samples, representing agricultural and forest soils, were  
129 collected in the Embrapa Southeast Livestock Research Center (Embrapa Pecuária Sudeste, São Carlos,  
130 São Paulo State, Brazil). A spodosol soil profile, characteristic of Amazonian soils, was obtained from  
131 nearby Barcelos city, Amazon State, Brazil. Obtained were samples of its Horizon A (0 – 15 cm, sample  
132 labeled Amazonian-A) and its Horizon C (350 cm, sample labeled Amazonian-C).

133 Additional information about the four soils can be found in Section 1 of the supporting information  
134 (SI). Powdered HA and FA samples were prepared from the four soils following the procedure  
135 recommended by the International Humic Substances Society (IHSS)<sup>45</sup>. The full protocol is described in  
136 Section 2 of the SI. Additional standardized and internationally recognized HA and FA samples were  
137 purchased from the IHSS (<https://humic-substances.org/>): Elliott Soil HA (sample label ESHA, cat. no.  
138 5S102H, lot V) and FA (sample label ESFA, cat. no. 5S102F, lot V) as well as Suwannee River HA  
139 (sample label SRHA, cat. no. 2S101H, lot II) and FA (sample label SRFA, cat. no. 2S101F, lot II).

140

### 141 2.2 Sample preparation techniques

142

143 Calculated sample amounts (0.5 – 0.7 mg), depending on carbon content, were weighed into glass  
144 vials to result in 0.2 mg-carbon equivalents. The solid HA or FA powders were then solubilized in 2000  $\mu$ L  
145 solvent (as described below) to yield a HA or FA solution at a concentration of 100 mg sample-carbon per  
146 L solution ( $\text{mgC}\cdot\text{L}^{-1}$ ). This solution was then diluted 2-fold to yield a final sample concentration of 50  
147  $\text{mgC}\cdot\text{L}^{-1}$  that was to be infused in the ESI-FT-ICR-MS per the standard operating procedure for the  
148 instrument. For all sample preparations, MilliQ laboratory-grade water (18.1 M $\Omega$ ) and methanol (MeOH,  
149 Fisher Scientific, Optima LC-MS grade) were used, and all glassware was acid-cleaned and combusted (8  
150 hours at 450 °C) prior to usage.

151

152 Five different solubilization procedures were employed to solubilize HA fractions. Procedural blanks  
153 corresponding to the five sample preparation techniques were also prepared.

154

- 155 • **Solubilization in methanol (MeOH approach)** – tested only on the agricultural HA sample. The  
156 powdered HA sample was mixed with 2000  $\mu$ L MeOH and then sonicated for 10 minutes. Due to the  
157 incomplete visual dissolution of the sample, the suspension was filtered using a 0.45  $\mu$ m acid-washed  
158 filter (ThermoFisher Scientific Target2 F2513-3, PTFE membrane). The sample was then diluted with  
159 additional 2000  $\mu$ L MeOH. Solubilized HA carbon was not quantified. A 30 min sonication was also  
160 employed as a separate test but complete HA solubilization was not achieved. A (-)ESI-FT-ICR mass  
161 spectrum was not acquired for the 30 min sonication test.
- 162 • **Solubilization in pyridine (Pyridine approach)** – tested only on the agricultural HA sample. The  
163 powdered HA sample was mixed with 2000  $\mu$ L pyridine (Sigma-Aldrich,  $\geq$  99.9%, HPLC grade) and  
164 then sonicated for 10 minutes. Due to the incomplete visual dissolution of the sample, the suspension  
165 was filtered using a 0.45  $\mu$ m acid-washed filter (Fisher Target2 F2513-3, PTFE membrane). The  
166 sample was then diluted with additional 2000  $\mu$ L pyridine. Solubilized HA carbon was not quantified. A  
167 30 min sonication was also employed as a separate test but complete HA solubilization was not  
168 achieved. A (-)ESI-FT-ICR mass spectrum was not acquired for the 30 min sonication test.
- 169 • **Solubilization with ammonium hydroxide (NH<sub>4</sub>OH approach)** – The powdered HA sample was  
170 mixed with 2000  $\mu$ L basified MilliQ water (pH 9, NH<sub>4</sub>OH, Fluka,  $\geq$  25 wt.%, LC-MS grade) and then  
171 sonicated for 10 minutes. Complete dissolution was achieved. A 500  $\mu$ L aliquot of the aqueous

172 supernatant was mixed with 500  $\mu\text{L}$  MeOH to result in a 1:1 MeOH: Water solution at a final sample-  
173 carbon concentration of 50  $\text{mgC}\cdot\text{L}^{-1}$ .

174 • **Base-solubilization with cation-exchange (CE approach)** – The powdered HA sample was mixed  
175 with 1800  $\mu\text{L}$  basified MilliQ water (pH 9, NaOH, Fisher Chemical SS254-4, 50% w/w Certified) and  
176 then sonicated for 10 minutes. Complete dissolution was achieved. To this solution, 200  $\mu\text{L}$  of a  
177 suspension of cation-exchange resin (Dowex 50WX8 100-200 mesh, Acros Organics) in water was  
178 added yielding a total water volume of 2000  $\mu\text{L}$ . The resin suspension was vortexed for 30 seconds,  
179 and let sit for 10 minutes. A 500  $\mu\text{L}$  aliquot of the aqueous supernatant was mixed with 500  $\mu\text{L}$  MeOH  
180 to result in a 1:1 MeOH: Water solution at a final sample-carbon concentration of 50  $\text{mgC}\cdot\text{L}^{-1}$ .

181 • **Base-solubilization with cation-exchange and solvent-extraction (CE<sub>EXTRACTION</sub> approach)** – The  
182 powdered HA sample was mixed with 1800  $\mu\text{L}$  basified MilliQ water (pH 9, NaOH, Fisher Chemical  
183 SS254-4, 50% w/w Certified) and then sonicated for 10 minutes. Complete dissolution was achieved.  
184 To this solution, 200  $\mu\text{L}$  of a suspension of cation-exchange resin (Dowex 50WX8 100-200 mesh,  
185 Acros Organics) in water was added yielding a total water volume of 2000  $\mu\text{L}$ . The resin suspension  
186 was vortexed for 30 seconds, and let sit for 10 minutes. Then, the supernatant was aliquoted out of the  
187 resin, and 2000  $\mu\text{L}$  of MeOH were added to the resin to solvent-extract any organic molecules that  
188 were retained onto it. The vial with the resin solvent-extraction was vortexed for 30 seconds and let sit  
189 for 10 min. Exactly 500  $\mu\text{L}$  of the methanolic supernatant were mixed with 500  $\mu\text{L}$  of the cation-  
190 exchanged aqueous solution to result in a 1:1 MeOH: Water solution at a final sample-carbon  
191 concentration of 50  $\text{mgC}\cdot\text{L}^{-1}$ . The CE<sub>EXTRACTION</sub> approach is graphically described on Figure S2.

192

193 Four different solubilization procedures were employed to solubilize FA fractions. Procedural blanks  
194 corresponding to the four approaches were also prepared.

195

196 • **Base-solubilization with cation-exchange and solvent-extraction (CE<sub>EXTRACTION</sub> approach)** – see  
197 above.

198 • **Solubilization in methanol (MeOH approach)** – The powdered FA sample was mixed with 2000  $\mu\text{L}$   
199 MeOH and then sonicated for 10 minutes. Complete dissolution was achieved. The sample was then  
200 diluted with additional 2000  $\mu\text{L}$  MeOH.

201 • **Solubilization in methanol-water mixture (MeOH:Water approach)** – The powdered FA sample  
202 was mixed with 2000  $\mu\text{L}$  MeOH and then sonicated for 10 minutes. Complete dissolution was  
203 achieved. Then, 500  $\mu\text{L}$  of the methanolic solution was mixed with 500  $\mu\text{L}$  of MQ water resulting in a  
204 1:1 MeOH: Water solution.

205 • **Solid-phase extraction using PPL cartridges (SPE-PPL approach)** – The powdered FA sample  
206 was solubilized in 5000  $\mu\text{L}$  of neutral (pH = 7) MilliQ water and then sonicated for 10 minutes.  
207 Complete dissolution was achieved. Then, the solution was acidified to pH = 2 using dropwise addition  
208 of HCl (Fisher, Certified ACS Plus). Solid-phase extraction was performed using PPL cartridges  
209 (Agilent Technologies, Bond Elut Priority Pollutant (PPL), 100 mg styrene divinyl copolymer). The  
210 extraction was performed following previously published procedures<sup>60, 75</sup> with slight modifications  
211 including more extensive cartridge washing, as described below, to minimize the number of  
212 contaminant species introduced into the samples from the PPL resin. The PPL cartridge is firstly  
213 conditioned with three cartridge volumes of MeOH and then washed using six cartridge volumes of  
214 acidified (pH = 2) MilliQ water. Then, the acidified sample solution is passed through the column.  
215 Then, the cartridge is washed with six cartridge volumes of acidified (pH = 2) MilliQ water. Volumes  
216 passing through the PPL cartridges were allowed to flow through by gravity, i.e., the flow rate was not  
217 enhanced using a peristaltic pump setup. The cartridges are then dried under a flow of ultrapure N<sub>2</sub>  
218 gas (AirGas, NI UHP300). Cartridges are eluted with 2000  $\mu\text{L}$  of MeOH to result in a methanolic  
219 extract at a final sample-carbon concentration of 50  $\text{mgC}\cdot\text{L}^{-1}$  (assuming 50% recovery).

220

221 A complete summary of samples and sample preparation techniques is presented in Table SX.

222

### 223 **2.3 Ultrahigh resolution mass spectrometry: (-)ESI-FT-ICR-MS**

224

225 Prepared HA and FA samples with different sample preparation techniques were directly injected into an  
226 Apollo II ESI source of a Bruker Daltonics 10-Tesla Apex Qe FT-ICR-MS at the College of Sciences Major  
227 Instrumentation Cluster (COSMIC) at Old Dominion University (Norfolk, VA). Samples were analyzed in  
228 (-)ESI mode with the same tuning parameters instrumental variability from influencing the spectral  
229 comparability. The samples of this project were analyzed in three separate days and data comparability  
230 was ensured by tune validation using SRFA and the recommendations by Hawkes and D'Andrilli<sup>43</sup>. All  
231 spectra are internally calibrated using naturally present fatty acids, dicarboxylic acids, and compounds  
232 belonging to CH<sub>2</sub>-homologous series<sup>76</sup>. Peaks with signal-to-noise ratio  $\geq 3$  were used for further

233 processing. Using an in-house MATLAB script, peaks found in the procedural blank, peaks of inorganic  
234 origin ("salt" peaks), and isotopologue (<sup>13</sup>C, <sup>32</sup>S) peaks are removed. Salt peaks are defined as following:  
235

236 **Table S1.** Definition of salt peaks

m/z range	Mass defects of salt peaks
< 300	0.4 – 0.97
300-500	0.5 – 0.96
500-800	0.6 – 0.95
> 800	0.7 – 0.94

237  
238 Mass spectra are assessed based on total number of formulas and number of salt peaks. Number of  
239 assignable peaks is also calculated as total peaks – salt peaks – blank peaks – isotopologue peaks. Salt  
240 and assignable peaks are assessed based on number-based percentages (i.e., salt % = number of salt  
241 peaks/number of total peaks x 100).

242 To assign molecular formulas and assess molecular composition, refined mass lists of assignable  
243 peaks were combined in a dataset of 81 samples in total. This dataset included SRFA quality control  
244 spectra as well as mass spectra of various agricultural and Amazonian samples from previous studies of  
245 Tadini<sup>77, 78</sup>. The mass spectra of all samples have been published in the *Mendelay Data* repository<sup>79-81</sup>.  
246 Spectra were then aligned to minimize formula assignment errors<sup>82</sup> using a previously published MATLAB  
247 script<sup>83</sup>. Aligned spectra were then assigned molecular formulas using a previously published MATLAB  
248 script<sup>84</sup>. Assignment was limited to <sup>12</sup>C<sub>5-∞</sub>, <sup>1</sup>H<sub>5-100</sub>, <sup>16</sup>O<sub>1-50</sub>, <sup>14</sup>N<sub>0-5</sub>, <sup>32</sup>S<sub>0-4</sub>, and <sup>31</sup>P<sub>0-2</sub> with maximum assignment  
249 error of 1 ppm. Formulas that did not adhere to previously established molecular rules for natural organic  
250 matter<sup>85-87</sup> were eliminated. Mass peaks with any residual ambiguous assignments (multiple possible  
251 molecular formulas per mass peak) were further refined using inclusion within homologous series (CH<sub>2</sub>,  
252 H<sub>2</sub>, COO, CH<sub>2</sub>O, C<sub>2</sub>H<sub>4</sub>O, O, H<sub>2</sub>O, NH<sub>3</sub>)<sup>85, 86</sup>. Any further ambiguous assignments were refined based on  
253 assignment error and the molecule with lower assignment error was selected. In the final list of molecular  
254 formulas, there was only one molecular formula per mass peak. For all formulas, the modified aromaticity  
255 index<sup>88, 89</sup> is calculated as shown below:  
256

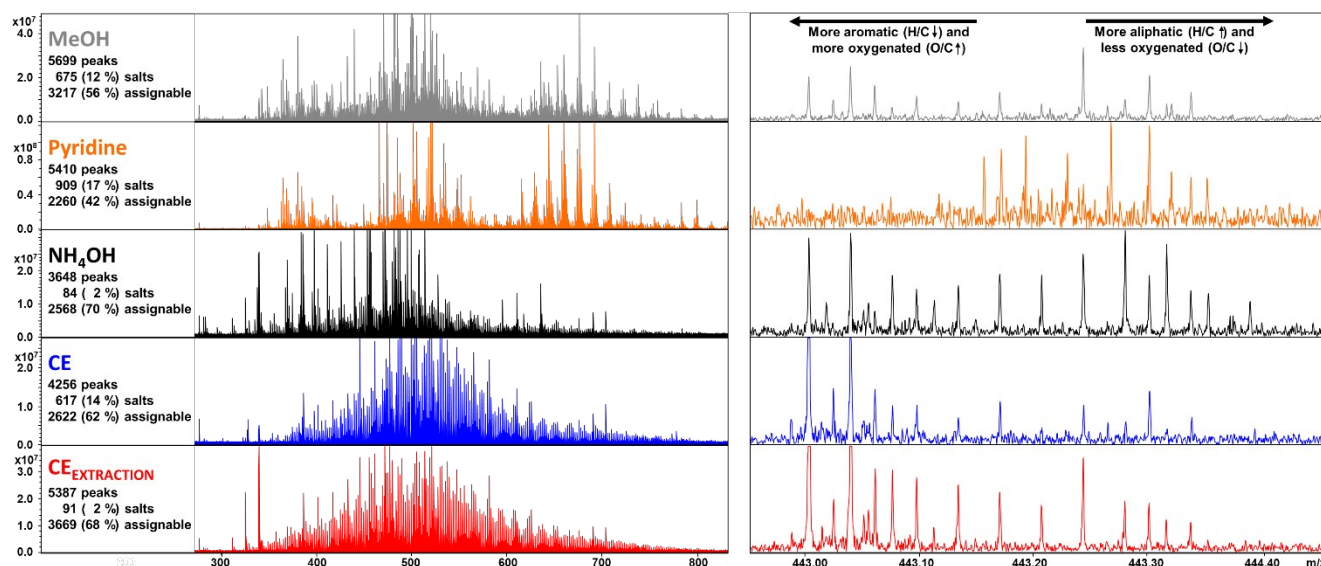
$$257 \quad AI_{MOD} = \frac{1 + C - \frac{1}{2}O - S - \frac{1}{2}(N + P + H + Cl)}{C - \frac{1}{2}O - N - S - P}$$

258  
259 Formulas were classified as following: Condensed aromatic ( $AI_{MOD} \geq 0.67$ ), aromatic ( $0.67 < AI_{MOD} \leq 0.50$ ),  
260 olefinic/alicyclic ( $0 < AI_{MOD} < 0.50$ ), and aliphatic ( $AI_{MOD} = 0$ ).  
261

### 262 3. RESULTS AND DISCUSSION

#### 263 264 3.1 Sample preparation techniques for humic acids (HA)

265  
266 The agricultural soil HA was prepared using five solubilization approaches and was analyzed on the  
267 same day with identical tuning parameters (Figure 1). Corresponding procedural blank samples are shown  
268 in Figure S4.



270

271 **Figure 1:** Whole (-)ESI-FT-ICR mass spectra (left) and expanded windows at nominal mass of 443 (right)  
 272 for an agricultural soil humic acid prepared using the MeOH, Pyridine, NH<sub>4</sub>OH, CE, and CE<sub>EXTRACTION</sub>  
 273 sample preparation techniques.

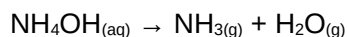
274

275 The mass spectra had variable shapes indicative of different ionization distribution when different  
 276 preparation techniques were employed (Figure 1). The MeOH and Pyridine approaches yielded trimodal  
 277 spectra whereas the NH<sub>4</sub>OH, CE, and CE<sub>EXTRACTION</sub> approaches yielded unimodal Gaussian-like spectra.  
 278 Unimodal spectra are commonly observed for natural organic matter, including for HA and FA fractions<sup>30</sup>.  
 279 <sup>31, 35-37, 51, 74, 90, 91</sup>. The observed spectral multimodality following the MeOH and Pyridine approaches was  
 280 likely caused by aggregation, selective ionization of different molecular classes, or preferential extraction  
 281 of different molecular classes because the HA samples were not entirely solubilized in MeOH or pyridine.  
 282 A detailed discussion of the observed spectral multimodality in the MeOH and Pyridine treatments can be  
 283 found in Section 7 of the SI.

284 The peak distributions at single nominal mass ranges are inspected to evaluate differences among  
 285 the detected ions in each sample (right panels of Figure 1). As spectral shape and magnitude are highly  
 286 liable to bias by instrumental parameters<sup>92, 93</sup>, per-m/z distributions are more reliable for evaluating how the  
 287 sample preparation techniques affected the observed HA molecular composition. These expanded ranges  
 288 reflect the chemical stability, prevalence, and ionizability of molecules of different compositions present  
 289 in the analyzed sample<sup>93</sup>. The distribution of ionized molecules with the MeOH approach appeared to contain  
 290 a variety of different molecules covering mass defects of 0.00 – 0.35. The distribution of ionized molecules  
 291 with the Pyridine approach appears to be missing ions with mass defects 0.95 – 0.15 indicative of  
 292 discrimination against more oxygenated aromatic molecules and preferential ionization of poorly  
 293 oxygenated aliphatic molecules. Additionally, molecular formulas were assessed based on the abundance  
 294 non-oxygen heteroelements (i.e., abundance of CHO, CHON, CHOS, and CHOP formulas). NSP  
 295 heteroelements are actively involved in the global biogeochemical cycles and in soil systems are  
 296 responsible for modulating plant growth as well as carbon decomposition and humification reactions<sup>94-96</sup>.  
 297 Highly elevated contents of CHON formulas were observed following the MeOH and Pyridine approaches  
 298 and these are discussed in detail in Section 8 of the SI. Because these two approaches (MeOH and  
 299 Pyridine) did not fully solubilize the samples and appeared to significantly bias the molecular composition,  
 300 they were not further considered for testing on other HA samples.

301 To further assess this, characteristic H/C and O/C clustering patterns of molecular formulas were  
 302 assessed using van Krevelen diagrams<sup>90, 97</sup> (Figure S6). The modified aromaticity index (AI<sub>MOD</sub>) metric  
 303 allowed for distinguishing aliphatic, olefinic, aromatic, and condensed aromatic molecules<sup>88, 89</sup>. The MeOH  
 304 approach enhanced the ionization of aliphatic/olefinic molecules (AI<sub>MOD</sub> < 0.50) but still allowed for the  
 305 observation of aromatic (0.67 < AI<sub>MOD</sub> ≤ 0.50) and condensed aromatic compounds (AI<sub>MOD</sub> ≥ 0.67). When  
 306 the agricultural HA was solubilized in pyridine (Figure 1), the ionization of aliphatic and olefinic molecules  
 307 (AI<sub>MOD</sub> < 0.50) was enhanced, and aromatic molecules (AI<sub>MOD</sub> > 0.50) were almost entirely undetected.  
 308 Aromatic molecules could have been suppressed in the ESI source (outcompeted by the aliphatic  
 309 molecules), but aromatic molecules could also have been not solubilized by the pyridine at all. Previous  
 310 studies have shown that pyridine extractions of organic matter are selective towards molecules of less  
 311 polar aliphatic character<sup>73</sup> agreeing with our findings.

312 Using basified aqueous solutions (either with NaOH or NH<sub>4</sub>OH) achieved complete solubilization of  
313 HA samples. Therefore, all sample molecules could be introduced into the ESI source. Using NH<sub>4</sub>OH is  
314 very practical and is commonly employed in HA analysis: In addition to serving as a basification agent,  
315 NH<sub>4</sub>OH is an ionization enhancer in (-)ESI<sup>98</sup>. Upon nebulization, NH<sub>4</sub>OH transforms into gaseous species,  
316 which are easily evacuated while promoting molecular ionization:  
317



318 In contrast, when powdered HA is solubilized by basified solutions with NaOH, sodium cations  
319 negatively affect the ionization and also cause the formation of peaks of inorganic origin, such as solvent-  
320 salt or HA-salt adducts<sup>51, 59</sup> also known as “salt” peaks. Thus, sodium cations must be removed using  
321 cation-exchange (CE) resin prior to analysis. When the spectra following NH<sub>4</sub>OH and CE approaches are  
322 compared (Figure 1), it is evident that peaks with mass defect > 0.20 are lost after the CE approach. Van  
323 Krevelen diagrams of the assigned formulas (Figure S6) indicate that higher molecular weight and poorly  
324 oxygenated molecules of both aromatic and aliphatic character sorbed onto the resin<sup>3</sup>. To recover any  
325 possible CE-entrapped molecules, the resin was further solvent-extracted with MeOH. This MeOH fraction  
326 was mixed with the CE-exchanged aqueous solution (CE<sub>EXTRACTION</sub> approach, Figure S2). Thus, all  
327 molecules were combined to provide a more representative solution of the original HA sample. The peaks  
328 that were not observed when only CE was employed have now been detected following the CE<sub>EXTRACTION</sub>  
329 approach (Figure 1). These ions may also have been lost due to unoptimized instrument parameters, but  
330 this is highly unlikely as all agricultural HA samples were analyzed on the same day with the same  
331 parameters and all samples exhibited stable ionization currents.

332 The spectra following the NH<sub>4</sub>OH and CE<sub>EXTRACTION</sub> appeared most to contain the highest  
333 diversity of HA ions (Figure 1), but it is evident that the per-m/z magnitude distributions were different.  
334 When NH<sub>4</sub>OH was used, ions of higher mass defect were detected indicative that less oxygenated  
335 molecules (primarily aliphatic but some aromatic too) are preferentially ionized at the expense of more  
336 oxygenated of various aromatic character. Not all aliphatic molecules that were observed following the  
337 NH<sub>4</sub>OH approach were detected when the CE<sub>EXTRACTION</sub> approach was employed. This indicates that while  
338 there is a significant overlap, the NH<sub>4</sub>OH and CE<sub>EXTRACTION</sub> approaches both yield a bias towards different  
339 molecular classes. Notably, the CE<sub>EXTRACTION</sub> approach allowed for detecting 48% more assignable peaks  
340 (5387) than the NH<sub>4</sub>OH approach (3648) making the CE<sub>EXTRACTION</sub> approach superior for this particular  
341 sample matrix. We hypothesize that NH<sub>4</sub>OH sucks up all of the charge which would be otherwise  
342 available for more molecules to be ionized (i.e., ionization poorly transfers to molecules and they do not  
343 become ions). Alternatively, NH<sub>4</sub>OH may be overionizing the oxygen-rich molecules cheating  
344 supercharged ions. Such ions would likely have too many charged groups (e.g., COO<sup>-</sup>) in close proximity  
345 that would cause the ions to collapse and degrade by the time they reach the ICR cell. These two  
346 hypotheses explain why NH<sub>4</sub>OH reduced the number of total and assignable peaks, as well as why less  
347 oxygenated compounds are preferentially detected with the NH<sub>4</sub>OH approach.

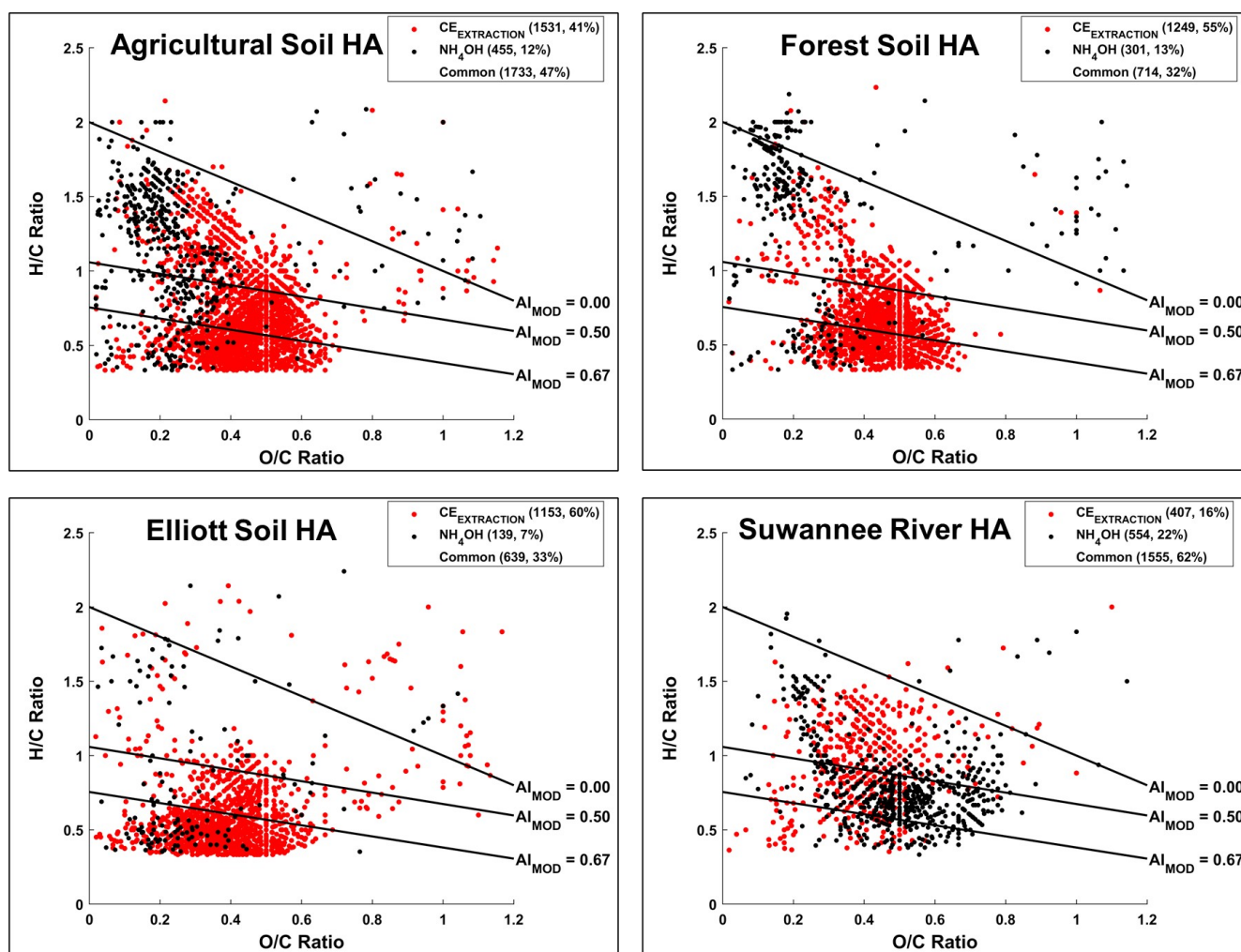
348 Only the three solubilization approaches which allowed for full solubilization (NH<sub>4</sub>OH, CE,  
349 CE<sub>EXTRACTION</sub>) were employed on the HA from the forest soil and similar trends were observed (Figure SX),  
350 as described in Section X of the SI. The NH<sub>4</sub>OH and CE<sub>EXTRACTION</sub> approaches were additionally tested on  
351 HA from the Elliott Soil (ESHA) and Suwannee River (SRHA). The spectral trends for ESHA (Figure SX)  
352 were similar to those observed for the agricultural and forest HA: CE<sub>EXTRACTION</sub> allowed for the detection of  
353 higher number of peaks over a large mass defect range. By contrast, SRHA performed very differently:  
354 the CE<sub>EXTRACTION</sub> spectrum had a bimodal distribution and there was an evident loss of ions with higher  
355 mass defect (> 0.23). The NH<sub>4</sub>OH preparation allowed for detecting a higher number of total and  
356 assignable peaks, and the per per-m/z distributions appeared to contain most diverse types of ions.

357 To better understand how the NH<sub>4</sub>OH and CE<sub>EXTRACTION</sub> approaches affect the molecular  
358 composition of HA fractions, molecular formulas of the CE<sub>EXTRACTION</sub> and NH<sub>4</sub>OH samples of the four HA  
359 samples were assessed as using presence/absence analysis<sup>99</sup> (Figure 2). The formulas of soil HA  
360 samples preferentially observed for following the NH<sub>4</sub>OH approach were less oxygenated than those  
361 observed following the CE<sub>EXTRACTION</sub> approach. The common formulas of these samples (Figures X) had  
362 H/C and O/C clustering patterns characteristic of soil HA fractions indicating that both NH<sub>4</sub>OH and  
363 CE<sub>EXTRACTION</sub> approaches manage to representatively characterize soil HA fractions<sup>3</sup>. Furthermore, the  
364 CE<sub>EXTRACTION</sub> also enhanced the detection of CHON and CHOS species for the soil HA matrices (Figures  
365 SX) but suppressed the detection of CHON species in SRHA (Figure SX).

366 In summary: 1) the unique formulas observed following the CE<sub>EXTRACTION</sub> approach generally had  
367 the same H/C and O/C clustering pattern as the common formulas (Figures SX and SX); and 2):  
368 CE<sub>EXTRACTION</sub> increased the number of total and assignable peaks; and 3) CE<sub>EXTRACTION</sub> allowed for observing  
369 the most diverse types of molecules (Figure S6) including more CHON species as well as condensed  
370 aromatic compounds of different degrees of oxygenation characteristic for the heterogenous nature of HA



371 fractions<sup>3</sup>. For these reasons we conclude that  $CE_{EXTRACTION}$  is the most optimal approach for the analysis of  
 372 the soil HA matrices assessed here. Thus, the  $CE_{EXTRACTION}$  sample preparation technique should be  
 373 considered more often for the analysis of HA samples. There was a notable difference in how the aquatic  
 374 HA sample behaved relative to the HA samples. The  $CE_{EXTRACTION}$  approach enhanced the detection of less  
 375 oxygenated compounds of uniform aromaticity and higher molecular weight (> 600 Da) whereas the  
 376  $NH_4OH$  approach enhanced the detection of more aromatic and more oxygenated compounds across the  
 377 entire MW range (300-800 Da). The  $NH_4OH$  approach also allows for the detection of more total and  
 378 assignable peaks. For these reasons, and because the uniquely detected peaks following the  $NH_4OH$   
 379 approach have a similar H/C and O/C clustering pattern to the common peaks among the  $NH_4OH$  and  
 380  $CE_{EXTRACTION}$  approaches, we conclude that the  $NH_4OH$  sample preparation technique is more optimal for  
 381 aquatic HA samples.  
 382  
 383

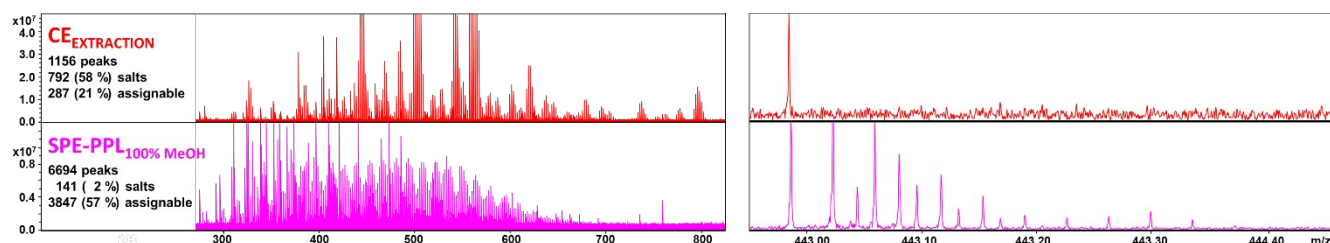


384  
 385 **Figure 2.** Van Krevelen diagrams (H/C versus O/C molar ratios) of assigned molecular formulas to humic  
 386 acids prepared using the  $NH_4OH$  and  $CE_{EXTRACTION}$  sample preparation techniques. Formulas are assessed  
 387 using presence/absence analysis<sup>99</sup> and only unique formulas to each sample are shown. The number of  
 388 formulas found in each of these pools is listed in the legends along with corresponding percentages  
 389 (relative to total number of formulas in the two samples being compared). The black lines indicate modified  
 390 aromaticity index cutoffs ( $AI_{MOD}$ )<sup>88, 89</sup>. Common formulas and assessments relative to molecular weight are  
 391 shown in Figures S7 and S8.  
 392

### 393 3.2 Sample preparation techniques for fulvic acids (FA)

394  
 395 The FA fraction is generally less problematic than the HA one due to the solubility of FA in both acid and  
 396 base. However, FA can have high mineral content even after clean-up with resins. To analyze the FA  
 397 samples, Amazon-C FA was first prepared using the  $CE_{EXTRACTION}$  approach. This approach was chosen as  
 398 we assumed that the sample is white due to its high content of Al complexes or minerals<sup>100</sup> that had  
 399 survived cation-exchange and DAX-8 resin purifications. Employing the same solubilization approach (i.e.,  
 400  $CE_{EXTRACTION}$ ) for both HA and FA fractions would have been ideal for achieving fully comparable data if HA  
 401 and FA from the same soil are compared. The resultant (-)ESI-FT-ICR mass spectrum was of poor quality

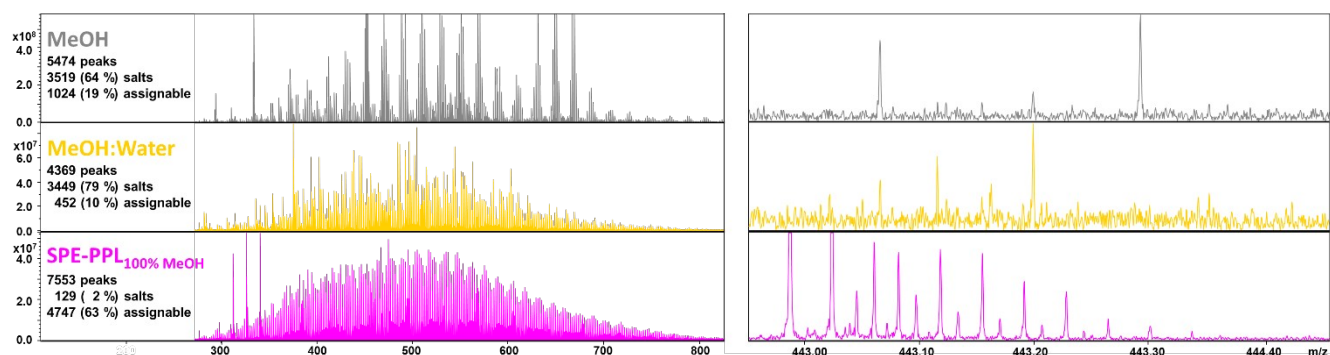
402 and only 1156 peaks were detected, 792 (58%) of which were salts (Figure 3). The sample also exhibited  
403 unstable ionization currents which provided early evidence that inorganic species which were not  
404 successfully removed by the CE resin. The molecular formulas corresponding to the 287 assignable peaks  
405 did not exhibit characteristic H/C and O/C clustering patterns for FA (Figure SX) but formulas within  
406 homologous series (CH<sub>2</sub>, H<sub>2</sub>O, CO, etc.) were found validating their existence in the Amazon-C sample.  
407 Thus, solid-phase extraction using Bond Elut Priority Pollutant (PPL) cartridges<sup>60</sup> was employed (SPE-PPL  
408 approach) yielding a spectrum of very high quality. The observed stable ionization currents during the  
409 analysis and the detected 6694 peaks indicated that the SPE-PPL preparation was successful in removing  
410 the inorganic impurities.  
411



412  
413 **Figure 3:** Whole (-)ESI-FT-ICR mass spectra (left) and expanded windows at nominal mass of 443 (right)  
414 for Amazon-C fulvic acid prepared using the **CE EXTRACTION** and **SPE-PPL<sub>100% MeOH</sub>** sample preparation  
415 techniques.  
416

417 Though SPE-PPL is widely used in the environmental sciences, it is well known that it only  
418 recovers 40-60% of the sample-carbon which fractionates the sample and biases the composition  
419 towards. Thus, we compared SPE-PPL with two more sample preparation techniques: MeOH and  
420 MeOH:Water approaches (Figure 4). For these additional tests we used Amazonian-A FA as the  
421 Amazonian-C FA was in limited amounts.  
422

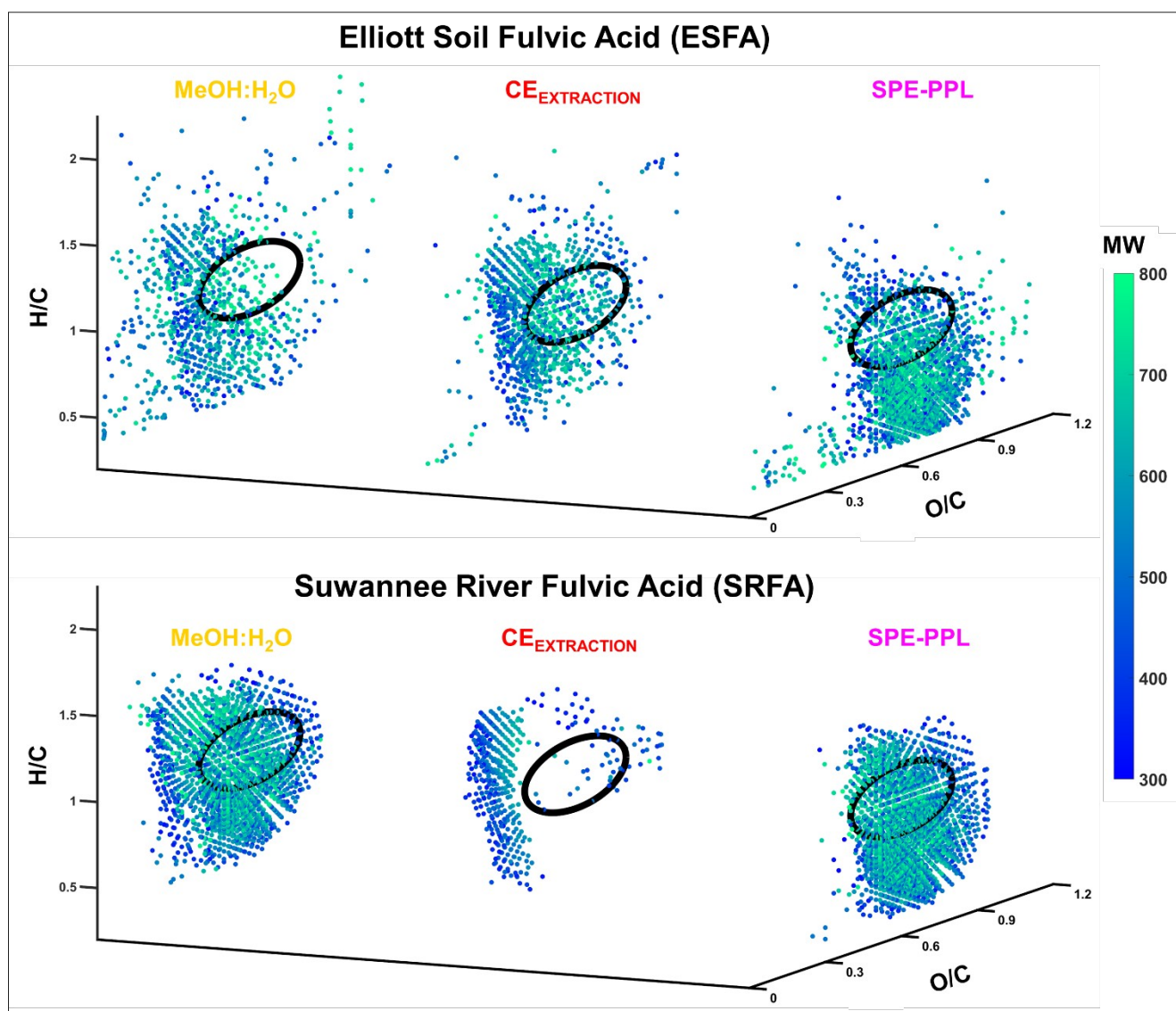
423 We first solubilized the powdered Amazonian-A FA in MeOH (MeOH approach). We assumed that  
424 MeOH will only solubilize the organic matter and avoid most inorganic substances due to the poor  
425 solubility of inorganic salts in MeOH<sup>101</sup>. The resultant spectrum was again populated with salt peaks. Then,  
426 we mixed the MeOH solution with water, which we have seen sometimes to improve the spectral quality  
427 (MeOH:Water approach)<sup>57</sup>. Unfortunately, this approach did not produce a usable spectrum for the  
428 Amazonian-A FA either. By contrast, the resultant mass spectrum of following SPE-PPL contained 7553  
429 peaks, which included 2% salt peaks and 4747 assignable peaks. The assigned molecular formulas to  
430 Amazonian-A and Amazonian-C spectra after SPE-PPL showed H/C and O/C clustering patterns typical  
431 for fulvic acids providing additional support that that the SPE-PPL approach was most appropriate for the  
432 soil FA samples of this study.



433  
434 **Figure 4:** Whole (-)ESI-FT-ICR mass spectra (left) and expanded windows at nominal mass of 443 (right)  
435 for Amazon-A fulvic acid prepared using the **MeOH**, **MeOH:Water**, and **SPE-PPL<sub>100% MeOH</sub>** sample  
436 preparation techniques.  
437

438 However, because the MeOH: Water approach was ineffective for our Amazonian FA samples,  
439 and this is the recommended sample preparation technique for standardized FA samples from IHSS<sup>43</sup>, we  
440 suspected that the “most optimal” sample preparation technique will differ for FA of different levels of purity  
441 or types of environments (e.g., soil versus aquatic) or matrices (e.g., peat versus loam soil). To investigate  
442 this, the Elliott Soil FA (ESFA) and Suwannee River FA (SRFA) samples from IHSS were solubilized using  
443 three different approaches: MeOH:Water (recommended for IHSS FA standards<sup>43</sup>), CE EXTRACTION (most  
444 optimal for the soil HA of our study), and SPE-PPL (most optimal for the Amazon soil FA samples).

445 For both ESHA and SRFA, the MeOH:Water approach yielded high number of peaks (4723 and  
446 5681 peaks, respectively) with spectra (Figures SX) comparable to those published by the interlaboratory  
447 comparison study<sup>43</sup>. Assigned molecular formulas also had typical H/C and O/C clustering patterns for  
448 these two samples (Figure 5)<sup>43</sup>. For ESFA, the molecular composition following ESFA and CE<sub>EXTRACTION</sub>  
449 approaches was of high similarity. This was probably because the molecules lost during the SPE-PPL  
450 extraction were likely small and undetectable by (-)ESI-FT-ICR-MS. Similar comparisons of CE and SPE-  
451 PPL preparations on ESFA were published previously paralleling our findings<sup>61</sup>. By contrast, the  
452 CE<sub>EXTRACTION</sub> approach caused a significant loss of lignin and tannin-like compounds of O/C > 0.4 in the  
453 SRFA sample indicating that CE<sub>EXTRACTION</sub> is not suitable for this matrix. The SPE-PPL approach produced  
454 similar spectra to those after the MeOH:Water approach but a closer inspection revealed that SPE-PPL  
455 skewed the molecular compositions towards a more oxygenated composition causing a loss of molecules  
456 with O/C > 0.5. Thus, for both Elliott soil and Suwannee River FA, the most appropriate sample  
457 preparation technique appears to be MeOH:Water concluding that the “most optimal” sample preparation  
458 technique will vary with the type of FA sample (e.g., aquatic versus soil) and its matrix constituents (e.g.,  
459 inorganic contents). These findings validated that the MeOH:Water sample preparation technique is the  
460 most suitable approach for working with SRFA and ESFA, which are key reference materials used for  
461 mass spectrometric tuning, tune validation and general quality control, and for ensuring inter- and intra-  
462 laboratory data comparability<sup>43</sup>. It must be noted that SPE-PPL greatly enhanced the detection of CHON  
463 species which might make the SPE-PPL technique more optimal for cases where soil nitrogen dynamics  
464 are researched.  
465



466  
467 **Figure 5.** Presence/absence analysis<sup>99</sup> on Elliott soil fulvic acid (ESFA, top) and Suwannee River fulvic  
468 acid (SRFA, bottom) molecular compositions obtained following MeOH: Water, CE<sub>EXTRACTION</sub>, and SPE-PPL  
469 solubilization approaches. Shown are van Krevelen diagrams (H/C versus O/C molar ratio plots) of unique

470 formulas to two or three samples. Formulas are color-coded based on molecular weight (MW). The black  
471 ovals serve as visual references for comparisons. Common formulas are shown in Figures SX and SX.  
472

473 The SPE-PPL technique is highly utilized in the community of aquatic sciences, but little knowledge exists  
474 on how data compares if the methanolic eluent is analyzed with 2-fold dilution with MeOH (i.e., as 100%  
475 MeOH) or with Water (i.e., as 1:1 MeOH:Water). Different research groups analyze SPE-PPL extracts  
476 differently without published rationale for the chosen solvent system. We have found that the addition of  
477 water can sometimes improve the spectral quality, but in other cases can deteriorate it. We have seen that  
478 this can vary even within the same type of sample matrix. For example, biochar extracts analyzed for the  
479 study by Wozniak<sup>56</sup> gave the highest quality spectrum following SPE-PPL<sub>100% MeOH</sub> whereas biochar extracts  
480 analyzed for the studies by Goranov<sup>57, 102</sup> gave the highest quality spectrum following SPE-PPL<sub>1:1 MeOH:Water</sub>.  
481 Samples were analyzed by the same analyst and sample preparation technique comparisons were made  
482 for each dataset. To test this for a soil FA, we analyzed ESFA using both SPE-PPL<sub>100% MeOH</sub> and SPE-  
483 PPL<sub>1:1 MeOH:Water</sub>. The addition of water reduced the total number of peaks from 6209 to 5006 and caused  
484 molecules with O/C > 0.5 to be suppressed whereas molecules with O/C < to be preferentially ionized  
485 (Figure SX).  
486

### 487 **3.3 Implications and recommendations for future studies**

488

489 The employment of sample preparation techniques for for HA and FA fractions can significantly  
490 alter the analytical window of the (-)ESI-FT-ICR-MS technique. The utilization of MeOH or pyridine as  
491 solvents did not provide representative spectra of the samples of this study likely due to the incomplete HA  
492 solubilization and the enhanced ionization of aliphatic compounds at the expense of aromatic ones (Figure  
493 1). However, it must be noted that the observed trends in this study are possibly specific to our sample  
494 matrices: the agricultural and forest HA samples used here were not fully soluble in MeOH but MeOH had  
495 been previously successful in solubilizing the aquatic SRHA without the addition of NH<sub>4</sub>OH or NaOH<sup>44</sup>.  
496 Thus, MeOH and Pyridine approaches should not be disregarded and be tested in future studies for HA  
497 matrices other than the ones employed here. The two approaches that were shown most prominent for the  
498 HA of our study, NH<sub>4</sub>OH and CE<sub>EXTRACTION</sub>, produced high-quality spectra but introduced biases in the  
499 observed molecular composition caused from different solvent composition affecting the (-)ESI. The  
500 CE<sub>EXTRACTION</sub> approach has been rarely employed<sup>3</sup> while using NH<sub>4</sub>OH has been the most common  
501 approach in HA analysis<sup>35-37, 44, 51, 53-55</sup>. Furthermore, while we did not observe this in our study (see section  
502 X of the SI), NH<sub>4</sub>OH has been shown to react with molecules and produce N-adducts<sup>58</sup>. Thus, it is likely  
503 that previously published HA data has been biased towards a more aliphatic (H-rich), N-rich, and less  
504 oxygenated (O-poor) composition, which may lead to false interpretations that the refractory HA fraction of  
505 SOM is biologically reactive and possibly labile towards microbial consumption.

506 The Amazonian FA samples of our study had challenging salt-rich matrices and the organic  
507 matter could not be purified using CE or DAX-8 resins. The SPE-PPL approach is rarely used for  
508 analyzing FA samples<sup>61</sup> but here we show that it has a significant merit for preparing FA samples of  
509 challenging composition (e.g., with high inorganic content). This agrees broad application of SPE-PPL in  
510 desalinating marine and other inorganic-rich aquatic matrices<sup>60, 75</sup> but questions remain regarding the  
511 molecular fractionation following SPE-PPL. Our comparisons reveal that SPE-PPL fractionates the sample  
512 causing a loss of less oxygenated molecules, but this may be unavoidable for salty FA matrices. For  
513 cleaner FA matrices, such as ESFA, employing SPE-PPL is not necessary as both MeOH:Water and  
514 CE<sub>EXTRACTION</sub> produced comparable data. Similar comparisons by Han<sup>61</sup> agree with our results for ESFA and  
515 disagree with those for SRFA, likely because of the significant differences among soil and aquatic  
516 molecules and how their chemistries.

517 While we make conclusions for the “most optimal” approach in our study, we must emphasize that  
518 these approaches are in no way universally optimal. Our work shows that the “most optimal” sample  
519 preparation technique is dependent on the type of sample and its matrix. Thus, we recommend that the  
520 most effective sample preparation technique is determined on a per-dataset basis. For future studies  
521 researchers should identify 1-3 key samples that are representative of the particular sample matrix of a  
522 dataset and test several different sample preparation techniques to determine the most optimal sample  
523 preparation technique. Performing such methodical comparisons will enhance (-)ESI-FT-ICR-MS data  
524 quality allowing us to better understand the molecular-level role of HS in the global biogeochemical cycles.

525 This study fills a key research gap in our knowledge on the (-)ESI-FT-ICR-MS analytical window as  
526 well as on data comparability. Ideally, all samples should be prepared using the same technique, but this  
527 is incredibly challenging for HS as the HA and FA fractions are incredibly different in their composition. We  
528 showed the utility of CE<sub>EXTRACTION</sub> for HA analysis and our comparisons on ESFA revealed that CE<sub>EXTRACTION</sub>  
529 is useful for FA analysis too. Thus, depending on the sample matrix, it may be possible that both HA and  
530 FA fractions of certain datasets are prepared using the same techniques. The large differences in trends

531 of aquatic versus soil HS indicate that comparing terrestrial and aquatic data is extremely challenging and  
532 more universal sample preparation techniques should be developed for such comparisons.

533 Any future studies combining datasets should carefully assess what sample preparation techniques  
534 were used and consider possible ways to trim formulas that are not reproducible among methods. For  
535 example, one possible way for this would be to align all formulas throughout the dataset and trim formulas  
536 found in only one sample while keeping formulas found in at least two (or more) samples. Another  
537 possible option would be to optimize for higher signal-to-noise peak picking thresholds removing  
538 irreproducible peaks of poor signal-to-noise from the spectra<sup>99</sup>. Such refinements are incredibly important  
539 for combining datasets and assessing them using multivariate statistical methods<sup>103</sup>. The usage of multiple  
540 sample preparation techniques, on the other hand, can be useful in expanding the ESI-FT-ICR-MS  
541 analytical window and by combining the data of the multiple preparations one can obtain a more  
542 comprehensive catalog of molecular formulas for a particular sample. Though only (-)ESI was employed  
543 here, it is well known that other ionization techniques can provide additional aspects to the molecular  
544 composition of environmental matrices. Thus, future studies should consider other ionization methods,  
545 such as positive-mode ESI or atmospheric pressure photoionization (APPI) techniques, and perform  
546 similar sample preparation technique comparisons to assess data comparability. This would advance the  
547 knowledge of the complex composition and heterogeneity of HS and more broadly enhance our  
548 understanding of the global biogeochemical cycles.

549

## 550 SUPPORTING INFORMATION

551

552 Graphical description of sampling sites; Sample preparation flowchart for the CE<sub>EXTRACTION</sub> approach;  
553 Description of ESI-FT-ICR-MS tuning and quality control; Mass spectra of procedural blanks; Assigned  
554 molecular formulas and van Krevelen diagrams; Discussion of mass spectral multimodality; Discussion of  
555 potential nitrogen-incorporation.

556

557 Mass spectral data (formula assignments) has been published in the *Mendeley Data* repository<sup>79</sup> (doi:  
558 10.17632/bwrgwx8b4s.1).

559

## 560 ACKNOWLEDGMENTS

561

562 The authors are grateful for the financial support provided by the São Paulo State Research Foundation  
563 (FAPESP, projects numbers: 2011/03250-2, 2012/51469-6, 2013/07276-1, 2013/13013-3, 2017/22950-1,  
564 2017/20084-5, and 2019/02939-9) and the Frank Batten Endowment fund to Dr. Patrick G. Hatcher by Old  
565 Dominion University (Norfolk, VA). We also thank Dr. Célia R. Montes (Centro de Energia Nuclear na  
566 Agricultura, Universidade de São Paulo, Brazil) for providing the Amazon soil samples.

567

## 568 REFERENCES

569

570 (1) Cotrufo, M. F.; Soong, J. L.; Horton, A. J.; Campbell, E. E.; Haddix, Michelle L.; Wall, D.  
571 H.; Parton, W. J. Formation of soil organic matter via biochemical and physical pathways of  
572 litter mass loss. *Nature Geoscience* **2015**, 8 (10), 776-779. DOI: 10.1038/ngeo2520.

573 (2) Liang, C.; Schimel, J. P.; Jastrow, J. D. The importance of anabolism in microbial control  
574 over soil carbon storage. *Nature Microbiology* **2017**, 2 (8), 1-6. DOI:  
575 10.1038/nmicrobiol.2017.105.

576 (3) DiDonato, N.; Chen, H.; Waggoner, D.; Hatcher, P. G. Potential origin and formation for  
577 molecular components of humic acids in soils. *Geochimica et Cosmochimica Acta* **2016**, 178,  
578 210-222. DOI: <https://doi.org/10.1016/j.gca.2016.01.013>.

579 (4) De Nobili, M.; Bravo, C.; Chen, Y. The spontaneous secondary synthesis of soil organic  
580 matter components: A critical examination of the soil continuum model theory. *Applied Soil*  
581 *Ecology* **2020**, 154, 1-13. DOI: <https://doi.org/10.1016/j.apsoil.2020.103655>.

582 (5) Hedges, J. I. The formation and clay mineral reactions of melanoidins. *Geochimica et*  
583 *Cosmochimica Acta* **1978**, 42 (1), 69-76. DOI: [https://doi.org/10.1016/0016-7037\(78\)90218-](https://doi.org/10.1016/0016-7037(78)90218-1)  
584 [1](https://doi.org/10.1016/0016-7037(78)90218-1).

- 585 (6) Stevenson, F. J. *Humus chemistry: Genesis, composition, reactions*; John Wiley & Sons,  
586 1994.
- 587 (7) Jokic, A.; Wang, M. C.; Liu, C.; Frenkel, A. I.; Huang, P. M. Integration of the polyphenol  
588 and Maillard reactions into a unified abiotic pathway for humification in nature: the role of  $\delta$ -  
589 MnO<sub>2</sub>. *Organic Geochemistry* **2004**, 35 (6), 747-762. DOI:  
590 <https://doi.org/10.1016/j.orggeochem.2004.01.021>.
- 591 (8) Schaeffer, A.; Nannipieri, P.; Kästner, M.; Schmidt, B.; Botterweck, J. From humic  
592 substances to soil organic matter—microbial contributions. In honour of Konrad Haider and  
593 James P. Martin for their outstanding research contribution to soil science. *Journal of Soils  
594 and Sediments* **2015**, 15 (9), 1865-1881. DOI: 10.1007/s11368-015-1177-4.
- 595 (9) Hayes, M. H. B.; Swift, R. S. An appreciation of the contribution of Frank Stevenson to the  
596 advancement of studies of soil organic matter and humic substances. *Journal of Soils and  
597 Sediments* **2018**, 18 (4), 1212-1231. DOI: 10.1007/s11368-016-1636-6.
- 598 (10) García, A. C.; van Tol de Castro, T. A.; Santos, L. A.; Tavares, O. C. H.; Castro, R. N.;  
599 Berbara, R. L. L.; García-Mina, J. M. Structure-property-function relationship of humic  
600 substances in modulating the root growth of plants: A review. *Journal of Environmental  
601 Quality* **2019**, 48 (6), 1622-1632. DOI: <https://doi.org/10.2134/jeq2019.01.0027>.
- 602 (11) Olk, D. C.; Bloom, P. R.; Perdue, E. M.; McKnight, D. M.; Chen, Y.; Farenhorst, A.;  
603 Senesi, N.; Chin, Y.-P.; Schmitt-Kopplin, P.; Hertkorn, N.; et al. Environmental and  
604 agricultural relevance of humic fractions extracted by alkali from soils and natural waters.  
605 *Journal of Environmental Quality* **2019**, 48 (2), 217-232. DOI:  
606 <https://doi.org/10.2134/jeq2019.02.0041>.
- 607 (12) Islam, M. A.; Morton, D. W.; Johnson, B. B.; Angove, M. J. Adsorption of humic and  
608 fulvic acids onto a range of adsorbents in aqueous systems, and their effect on the  
609 adsorption of other species: A review. *Separation and Purification Technology* **2020**, 247, 1-  
610 19. DOI: <https://doi.org/10.1016/j.seppur.2020.116949>.
- 611 (13) Lipczynska-Kochany, E. Humic substances, their microbial interactions and effects on  
612 biological transformations of organic pollutants in water and soil: A review. *Chemosphere*  
613 **2018**, 202, 420-437. DOI: <https://doi.org/10.1016/j.chemosphere.2018.03.104>.
- 614 (14) Valenzuela, E. I.; Cervantes, F. J. The role of humic substances in mitigating  
615 greenhouse gases emissions: Current knowledge and research gaps. *Science of The Total  
616 Environment* **2021**, 750, 1-14. DOI: <https://doi.org/10.1016/j.scitotenv.2020.141677>.
- 617 (15) Bayer, C.; Martin-Neto, L.; Mielniczuk, J.; Pavinato, A.; Dieckow, J. Carbon  
618 sequestration in two Brazilian Cerrado soils under no-till. *Soil and Tillage Research* **2006**, 86  
619 (2), 237-245. DOI: <https://doi.org/10.1016/j.still.2005.02.023>.
- 620 (16) Sá, J. C. d. M.; Lal, R.; Cerri, C. C.; Lorenz, K.; Hungria, M.; de Faccio Carvalho, P. C.  
621 Low-carbon agriculture in South America to mitigate global climate change and advance food  
622 security. *Environ Int* **2017**, 98, 102-112. DOI: <https://doi.org/10.1016/j.envint.2016.10.020>.
- 623 (17) Hatcher, P. G.; Waggoner, D. C.; Chen, H. Evidence for the existence of humic acids in  
624 peat soils based on solid-state <sup>13</sup>C NMR. *Journal of Environmental Quality* **2019**, 48 (6),  
625 1571-1577. DOI: 10.2134/jeq2019.02.0083.
- 626 (18) Gerke, J. Concepts and misconceptions of humic substances as the stable part of soil  
627 organic matter: A review. *Agronomy* **2018**, 8 (76), 1-16. DOI: 10.3390/agronomy8050076.

- 628 (19) Kleber, M.; Lehmann, J. Humic substances extracted by alkali are invalid proxies for the  
629 dynamics and functions of organic matter in terrestrial and aquatic ecosystems. *Journal of*  
630 *Environmental Quality* **2019**, *48* (2), 207-216, <https://doi.org/10.2134/jeq2019.01.0036>. DOI:  
631 <https://doi.org/10.2134/jeq2019.01.0036>.
- 632 (20) Lehmann, J.; Kleber, M. The contentious nature of soil organic matter. *Nature* **2015**, *528*  
633 (7580), 60-68. DOI: 10.1038/nature16069.
- 634 (21) Myneni, S. C. B. Chemistry of natural organic matter - the next step: Commentary on a  
635 humic substances debate. *Journal of Environmental Quality* **2019**, *48* (2), 233-235,  
636 <https://doi.org/10.2134/jeq2019.02.0002c>. DOI: <https://doi.org/10.2134/jeq2019.02.0002c>.
- 637 (22) Janzen, H. The future of humic substances research: Preface to a debate. *Journal of*  
638 *Environmental Quality* **2019**, *48* (2), 205-206, <https://doi.org/10.2134/jeq2019.01.0001c>. DOI:  
639 <https://doi.org/10.2134/jeq2019.01.0001c>.
- 640 (23) Kelleher, B. P.; Simpson, A. J. Humic substances in soils: Are they really chemically  
641 distinct? *Environmental Science & Technology* **2006**, *40* (15), 4605-4611. DOI:  
642 10.1021/es0608085.
- 643 (24) Schmidt, M. W. I.; Torn, M. S.; Abiven, S.; Dittmar, T.; Guggenberger, G.; Janssens, I.  
644 A.; Kleber, M.; Kögel-Knabner, I.; Lehmann, J.; Manning, D. A. C.; et al. Persistence of soil  
645 organic matter as an ecosystem property. *Nature* **2011**, *478* (7367), 49-56. DOI:  
646 10.1038/nature10386.
- 647 (25) Ohno, T.; Hess, N. J.; Qafoku, N. P. Current understanding of the use of alkaline  
648 extractions of soils to investigate soil organic matter and environmental processes. *Journal of*  
649 *Environmental Quality* **2019**, *48* (6), 1561-1564. DOI: 10.2134/jeq2019.08.0292.
- 650 (26) Olk, D. C.; Bloom, P. R.; De Nobili, M.; Chen, Y.; McKnight, D. M.; Wells, M. J. M.;  
651 Weber, J. Using humic fractions to understand natural organic matter processes in soil and  
652 water: Selected studies and applications. *Journal of Environmental Quality* **2019**, *48* (6),  
653 1633-1643. DOI: <https://doi.org/10.2134/jeq2019.03.0100>.
- 654 (27) Hayes, M. H. B.; Swift, R. S. Chapter One - Vindication of humic substances as a key  
655 component of organic matter in soil and water. In *Advances in Agronomy*, Sparks, D. L. Ed.;  
656 Vol. 163; Academic Press, 2020; pp 1-37.
- 657 (28) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. Fourier transform ion cyclotron  
658 resonance mass spectrometry: A primer. *Mass Spectrom Rev* **1998**, *17* (1), 1-35. DOI: [10.1002/\(Sici\)1098-2787\(1998\)17:1<1::Aid-Mas1>3.0.Co;2-K](https://doi.org/10.1002/(Sici)1098-2787(1998)17:1<1::Aid-Mas1>3.0.Co;2-K).
- 660 (29) Kujawinski, E. B. Electrospray ionization Fourier transform ion cyclotron resonance  
661 mass spectrometry (ESI FT-ICR MS): Characterization of complex environmental mixtures.  
662 *Environmental Forensics* **2002**, *3* (3), 207-216.
- 663 (30) Sleighter, R. L.; Hatcher, P. G. The application of electrospray ionization coupled to  
664 ultrahigh resolution mass spectrometry for the molecular characterization of natural organic  
665 matter. *Journal of Mass Spectrometry* **2007**, *42* (5), 559-574. DOI: [doi:10.1002/jms.1221](https://doi.org/10.1002/jms.1221).
- 666 (31) Sleighter, R. L.; Hatcher, P. G. Fourier transform mass spectrometry for the molecular  
667 level characterization of natural organic matter: Instrument capabilities, applications, and  
668 limitations In *Fourier Transforms - Approach to Scientific Principles*, Nikolic, G. Ed.; InTech,  
669 2011; pp 295-320.

- 670 (32) Kujawinski, E. B.; Freitas, M. A.; Zang, X.; Hatcher, P. G.; Green-Church, K. B.; Jones,  
671 R. B. The application of electrospray ionization mass spectrometry (ESI MS) to the structural  
672 characterization of natural organic matter. *Organic Geochemistry* **2002**, 33 (3), 171-180.
- 673 (33) Stenson, A. C.; Marshall, A. G.; Cooper, W. T. Exact masses and chemical formulas of  
674 individual Suwannee River fulvic acids from ultrahigh resolution electrospray ionization  
675 Fourier transform ion cyclotron resonance mass spectra. *Analytical Chemistry* **2003**, 75 (6),  
676 1275-1284. DOI: 10.1021/ac026106p.
- 677 (34) Witt, M.; Fuchser, J.; Koch, B. P. Fragmentation studies of fulvic acids using collision  
678 induced dissociation Fourier transform ion cyclotron resonance mass spectrometry.  
679 *Analytical Chemistry* **2009**, 81 (7), 2688-2694.
- 680 (35) Ikeya, K.; Sleighter, R. L.; Hatcher, P.; Watanabe, A. Compositional features of  
681 Japanese Humic Substances Society standard soil humic and fulvic acids by Fourier  
682 transform ion cyclotron resonance mass spectrometry and X-ray diffraction profile analysis.  
683 *Humic Subst. Res.* **2012**, 9, 25-33.
- 684 (36) Ikeya, K.; Sleighter, R. L.; Hatcher, P. G.; Watanabe, A. Fourier transform ion cyclotron  
685 resonance mass spectrometric analysis of the green fraction of soil humic acids. *Rapid*  
686 *Communications in Mass Spectrometry* **2013**, 27 (22), 2559-2568. DOI: 10.1002/rcm.6718.
- 687 (37) Ikeya, K.; Sleighter, R. L.; Hatcher, P. G.; Watanabe, A. Characterization of the chemical  
688 composition of soil humic acids using Fourier transform ion cyclotron resonance mass  
689 spectrometry. *Geochimica et Cosmochimica Acta* **2015**, 153, 169-182. DOI:  
690 10.1016/j.gca.2015.01.002.
- 691 (38) Ikeya, K.; Sleighter, R. L.; Hatcher, P. G.; Watanabe, A. Chemical compositional  
692 analysis of soil fulvic acids using Fourier transform ion cyclotron resonance mass  
693 spectrometry. *Rapid Communications in Mass Spectrometry* **2020**, 34 (15), 1-11. DOI:  
694 10.1002/rcm.8801.
- 695 (39) Qin, S.; Xu, C.; Xu, Y.; Bai, Y.; Guo, F. Molecular signatures of humic acids from  
696 different sources as revealed by ultrahigh resolution mass spectrometry. *J. Chem.* **2020**,  
697 *2020*, 1-11.
- 698 (40) Qin, S.; Xu, C.; Guo, F.; Qi, J.; Xu, L.; Xu, Y.; Song, F.; Bai, Y. Molecular signatures of  
699 three fulvic acid standard samples as revealed by electrospray ionization Fourier transform  
700 ion cyclotron resonance mass spectrometry. *ChemistrySelect* **2019**, 4 (47), 13940-13946.  
701 DOI: 10.1002/slct.201903293.
- 702 (41) Reemtsma, T.; These, A.; Springer, A.; Linscheid, M. Fulvic acids as transition state of  
703 organic matter: Indications from high resolution mass spectrometry. *Environmental Science*  
704 *& Technology* **2006**, 40 (19), 5839-5845. DOI: 10.1021/es060318c.
- 705 (42) Kramer, R. W.; Kujawinski, E. B.; Hatcher, P. G. Identification of black carbon derived  
706 structures in a volcanic ash soil humic acid by Fourier transform ion cyclotron resonance  
707 mass spectrometry. *Environmental Science & Technology* **2004**, 38 (12), 3387-3395.
- 708 (43) Hawkes, J. A.; D'Andrilli, J.; Agar, J. N.; Barrow, M. P.; Berg, S. M.; Catalán, N.; Chen,  
709 H.; Chu, R. K.; Cole, R. B.; Dittmar, T.; et al. An international laboratory comparison of  
710 dissolved organic matter composition by high resolution mass spectrometry: Are we getting  
711 the same answer? *Limnology and Oceanography: Methods* **2020**, 18 (6), 235-258. DOI:  
712 10.1002/lom3.10364.



- 713 (44) Zhrebker, A.; Kim, S.; Schmitt-Kopplin, P.; Spencer, R. G. M.; Lechtenfeld, O.;  
714 Podgorski, D. C.; Hertkorn, N.; Harir, M.; Nurfajin, N.; Koch, B.; et al. Interlaboratory  
715 comparison of humic substances compositional space as measured by Fourier transform ion  
716 cyclotron resonance mass spectrometry (IUPAC Technical Report). *Pure and Applied*  
717 *Chemistry* **2020**, 92 (9), 1447-1467. DOI: doi:10.1515/pac-2019-0809.
- 718 (45) Swift, R. S. Organic matter characterization. In *Methods of Soil Analysis*, 1996; pp 1011-  
719 1069.
- 720 (46) Wershaw, R. L. Molecular aggregation of humic substances. *Soil Science* **1999**, 164  
721 (11), 803-813.
- 722 (47) Zanin, L.; Tomasi, N.; Cesco, S.; Varanini, Z.; Pinton, R. Humic Substances Contribute  
723 to Plant Iron Nutrition Acting as Chelators and Biostimulants. *Frontiers in Plant Science*  
724 **2019**, 10 (675), 1-10, Mini Review. DOI: 10.3389/fpls.2019.00675.
- 725 (48) Hakim, A.; Suzuki, T.; Kobayashi, M. Strength of humic acid aggregates: Effects of  
726 divalent cations and solution pH. *ACS Omega* **2019**, 4 (5), 8559-8567. DOI:  
727 10.1021/acsomega.9b00124.
- 728 (49) Ohno, T.; Ohno, P. E. Influence of heteroatom pre-selection on the molecular formula  
729 assignment of soil organic matter components determined by ultrahigh resolution mass  
730 spectrometry. *Analytical and Bioanalytical Chemistry* **2013**, 405 (10), 3299-3306. DOI:  
731 10.1007/s00216-013-6734-3.
- 732 (50) Ohno, T.; He, Z.; Sleighter, R. L.; Honeycutt, C. W.; Hatcher, P. G. Ultrahigh resolution  
733 mass spectrometry and indicator species analysis to identify marker components of soil- and  
734 plant biomass-derived organic matter fractions. *Environmental Science & Technology* **2010**,  
735 44 (22), 8594-8600. DOI: 10.1021/es101089t.
- 736 (51) Stenson, A. C.; Landing, W. M.; Marshall, A. G.; Cooper, W. T. Ionization and  
737 fragmentation of humic substances in electrospray ionization Fourier transform-ion cyclotron  
738 resonance mass spectrometry. *Analytical Chemistry* **2002**, 74 (17), 4397-4409. DOI:  
739 10.1021/ac020019f.
- 740 (52) Jiménez-Morillo, N. T.; González-Pérez, J. A.; Almendros, G.; De la Rosa, J. M.;  
741 Waggoner, D. C.; Jordán, A.; Zavala, L. M.; González-Vila, F. J.; Hatcher, P. G. Ultra-high  
742 resolution mass spectrometry of physical speciation patterns of organic matter in fire-affected  
743 soils. *Journal of Environmental Management* **2018**, 225, 139-147. DOI:  
744 <https://doi.org/10.1016/j.jenvman.2018.07.069>.
- 745 (53) Kujawinski, E. B.; Hatcher, P. G.; Freitas, M. A. High-resolution Fourier transform ion  
746 cyclotron resonance mass spectrometry of humic and fulvic acids: Improvements and  
747 comparisons. *Analytical Chemistry* **2002**, 74 (2), 413-419. DOI: 10.1021/ac0108313.
- 748 (54) Jiménez-González, M. A.; Almendros, G.; Waggoner, D. C.; Álvarez, A. M.; Hatcher, P.  
749 G. Assessment of the molecular composition of humic acid as an indicator of soil carbon  
750 levels by ultra-high-resolution mass spectrometric analysis. *Organic Geochemistry* **2020**,  
751 143, 1-10. DOI: <https://doi.org/10.1016/j.orggeochem.2020.104012>.
- 752 (55) Panettieri, M.; Jiménez-González, M. A.; De Sosa, L. L.; Almendros, G.; Madejón, E.  
753 Chemical diversity and molecular signature of soil humic fractions used as proxies of soil  
754 quality under contrasted tillage management. *Spanish Journal of Soil Science* **2021**, 11 (1),  
755 39-54. DOI: <https://doi.org/10.3232/SJSS.2021.V11.N1.04>.

- 756 (56) Wozniak, A. S.; Goranov, A. I.; Mitra, S.; Bostick, K. W.; Zimmerman, A. R.; Schlesinger,  
757 D. R.; Myneni, S.; Hatcher, P. G. Molecular heterogeneity in pyrogenic dissolved organic  
758 matter from a thermal series of oak and grass chars. *Organic Geochemistry* **2020**, *148* (1), 1-  
759 18. DOI: <https://doi.org/10.1016/j.orggeochem.2020.104065>.
- 760 (57) Goranov, A. I.; Wozniak, A. S.; Bostick, K. W.; Zimmerman, A. R.; Mitra, S.; Hatcher, P.  
761 G. Photochemistry after fire: Structural transformations of pyrogenic dissolved organic matter  
762 elucidated by advanced analytical techniques. *Geochimica et Cosmochimica Acta* **2020**, *290*,  
763 271-292. DOI: <https://doi.org/10.1016/j.gca.2020.08.030>.
- 764 (58) Leverton, T. Method considerations for compound identification in complex mixtures  
765 using electrospray ionization ultrahigh resolution mass spectrometry. Master's Thesis,  
766 Michigan Technological University, Houghton, Michigan, 2019.
- 767 (59) Brown, T. L.; Rice, J. A. Effect of Experimental Parameters on the ESI FT-ICR Mass  
768 Spectrum of Fulvic Acid. *Analytical Chemistry* **2000**, *72* (2), 384-390. DOI:  
769 10.1021/ac9902087.
- 770 (60) Dittmar, T.; Koch, B.; Hertkorn, N.; Kattner, G. A simple and efficient method for the  
771 solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnology*  
772 *and Oceanography: Methods* **2008**, *6*, 230-235. DOI: <https://doi.org/10.4319/lom.2008.6.230>.
- 773 (61) Han, R.; Lv, J.; Luo, L.; Wen, B.; Zhang, S. Molecular-scale investigation of soil fulvic  
774 acid and water-extractable organic matter by high-resolution mass spectrometry and <sup>1</sup>H NMR  
775 spectroscopy. *Environmental Chemistry* **2019**, *16* (2), 92-100. DOI:  
776 <https://doi.org/10.1071/EN18124>.
- 777 (62) Bahureksa, W.; Tfaily, M. M.; Boiteau, R. M.; Young, R. B.; Logan, M. N.; McKenna, A.  
778 M.; Borch, T. Soil Organic Matter Characterization by Fourier Transform Ion Cyclotron  
779 Resonance Mass Spectrometry (FTICR MS): A Critical Review of Sample Preparation,  
780 Analysis, and Data Interpretation. *Environmental Science & Technology* **2021**, *55* (14), 9637-  
781 9656. DOI: 10.1021/acs.est.1c01135.
- 782 (63) Liigand, J.; Kruve, A.; Leito, I.; Girod, M.; Antoine, R. Effect of mobile phase on  
783 electrospray ionization efficiency. *Journal of The American Society for Mass Spectrometry*  
784 **2014**, *25* (11), 1853-1861. DOI: 10.1007/s13361-014-0969-x.
- 785 (64) Kostianen, R.; Kauppila, T. J. Effect of eluent on the ionization process in liquid  
786 chromatography - mass spectrometry. *Journal of Chromatography A* **2009**, *1216* (4), 685-  
787 699. DOI: <https://doi.org/10.1016/j.chroma.2008.08.095>.
- 788 (65) Soule, M. C. K.; Longnecker, K.; Giovannoni, S. J.; Kujawinski, E. B. Impact of  
789 instrument and experiment parameters on reproducibility of ultrahigh resolution ESI FT-ICR  
790 mass spectra of natural organic matter. *Organic Geochemistry* **2010**, *41* (8), 725-733. DOI:  
791 10.1016/j.orggeochem.2010.05.017.
- 792 (66) Van Berkel, G. J.; McLuckey, S. A.; Glish, G. L. Electrochemical origin of radical cations  
793 observed in electrospray ionization mass spectra. *Analytical Chemistry* **1992**, *64* (14), 1586-  
794 1593. DOI: 10.1021/ac00038a015.
- 795 (67) Miyabayashi, K.; Naito, Y.; Tsujimoto, K.; Miyake, M. Structure characterization of  
796 polyaromatic hydrocarbons in Arabian mix vacuum residue by electrospray ionization Fourier  
797 transform ion cyclotron resonance mass spectrometry. *International Journal of Mass*  
798 *Spectrometry* **2004**, *235* (1), 49-57. DOI: <https://doi.org/10.1016/j.ijms.2004.04.001>.

- 799 (68) Van Berkel, G. J.; Asano, K. G. Chemical derivatization for electrospray ionization mass  
800 spectrometry. 2. Aromatic and highly conjugated molecules. *Analytical Chemistry* **1994**, 66  
801 (13), 2096-2102. DOI: 10.1021/ac00085a027.
- 802 (69) Molnárné Guricza, L.; Schrader, W. Electrospray ionization for determination of non-  
803 polar polyaromatic hydrocarbons and polyaromatic heterocycles in heavy crude oil  
804 asphaltenes. *Journal of Mass Spectrometry* **2015**, 50 (3), 549-557. DOI:  
805 <https://doi.org/10.1002/jms.3561>.
- 806 (70) Gaspar, A.; Kunenkov, E. V.; Lock, R.; Desor, M.; Perminova, I.; Schmitt-Kopplin, P.  
807 Combined utilization of ion mobility and ultra-high-resolution mass spectrometry to identify  
808 multiply charged constituents in natural organic matter. *Rapid Communications in Mass*  
809 *Spectrometry* **2009**, 23 (5), 683-688. DOI: <https://doi.org/10.1002/rcm.3924>.
- 810 (71) Krueve, A.; Kaupmees, K.; Liigand, J.; Leito, I. Negative electrospray ionization via  
811 deprotonation: Predicting the ionization efficiency. *Analytical Chemistry* **2014**, 86 (10), 4822-  
812 4830. DOI: 10.1021/ac404066v.
- 813 (72) McKee, G. A.; Hatcher, P. G. A new approach for molecular characterisation of  
814 sediments with Fourier transform ion cyclotron resonance mass spectrometry: Extraction  
815 optimisation. *Organic Geochemistry* **2015**, 85, 22-31. DOI:  
816 <https://doi.org/10.1016/j.orggeochem.2015.04.007>.
- 817 (73) Willoughby, A. S.; Wozniak, A. S.; Hatcher, P. G. A molecular-level approach for  
818 characterizing water-insoluble components of ambient organic aerosol particulates using  
819 ultrahigh-resolution mass spectrometry. *Atmospheric Chemistry and Physics* **2014**, 14 (18),  
820 10299-10314. DOI: 10.5194/acp-14-10299-2014.
- 821 (74) Tfaily, M. M.; Chu, R. K.; Tolić, N.; Roscioli, K. M.; Anderton, C. R.; Paša-Tolić, L.;  
822 Robinson, E. W.; Hess, N. J. Advanced solvent based methods for molecular  
823 characterization of soil organic matter by high-resolution mass spectrometry. *Analytical*  
824 *Chemistry* **2015**, 87 (10), 5206-5215. DOI: 10.1021/acs.analchem.5b00116.
- 825 (75) Li, Y.; Harir, M.; Lucio, M.; Kanawati, B.; Smirnov, K.; Flerus, R.; Koch, B. P.; Schmitt-  
826 Kopplin, P.; Hertkorn, N. Proposed guidelines for solid phase extraction of Suwannee River  
827 dissolved organic matter. *Analytical Chemistry* **2016**, 88 (13), 6680-6688. DOI:  
828 10.1021/acs.analchem.5b04501.
- 829 (76) Sleighter, R. L.; McKee, G. A.; Liu, Z.; Hatcher, P. G. Naturally present fatty acids as  
830 internal calibrants for Fourier transform mass spectra of dissolved organic matter. *Limnology*  
831 *and Oceanography: Methods* **2008**, 6, 246-253. DOI: <https://doi.org/10.4319/lom.2008.6.246>.
- 832 (77) Tadini, A. M.; Nicolodelli, G.; Senesi, G. S.; Ishida, D. A.; Montes, C. R.; Lucas, Y.;  
833 Mounier, S.; Guimarães, F. E. G.; Milori, D. M. B. P. Soil organic matter in podzol horizons of  
834 the Amazon region: Humification, recalcitrance, and dating. *Science of The Total*  
835 *Environment* **2018**, 613-614, 160-167. DOI: <https://doi.org/10.1016/j.scitotenv.2017.09.068>.
- 836 (78) Tadini, A. M.; Bernardi, A. C. C.; Milori, D. M. B. P.; Oliveira, P. P. A.; Pezzopane, J. R.  
837 M.; Martin-Neto, L. Spectroscopic characteristics of humic acids extracted from soils under  
838 different integrated agricultural production systems in tropical regions. *Geoderma Regional*  
839 **2022**, 28, e00476. DOI: <https://doi.org/10.1016/j.geodrs.2021.e00476>.
- 840 (79) Goranov, A.; Tadini, A. M.; Martin-Neto, L.; Bernardi, A. C. C.; Oliveira, P. P. A.;  
841 Pezzopane, J. R. M.; Milori, D. M. B. P.; Mounier, S.; Hatcher, P. G. DATASET: Comparison

842 of sample preparation techniques for the ESI-FT-ICR-MS analysis of humic and fulvic acids.  
843 Mendeley Data, 2022.

844 (80) Tadini, A. M.; Goranov, A.; Martin-Neto, L.; Bernardi, A. C. C.; Oliveira, P. P. A.;  
845 Pezzopane, J. R. M.; Hatcher, P. G. DATASET: Characterization of humic acid fractions from  
846 crop-livestock-forest systems using ultrahigh resolution mass spectrometry. Mendeley Data,  
847 2022.

848 (81) Tadini, A. M.; Goranov, A.; Mounier, S.; Milori, D. M. B. P.; Martin-Neto, L.; Hatcher, P.  
849 G. DATASET: Characterization of humic and fulvic acid fractions from Amazonian Spodosol  
850 soils using ultrahigh resolution mass spectrometry. Mendeley Data, 2022.

851 (82) Merder, J.; Freund, J. A.; Feudel, U.; Niggemann, J.; Singer, G.; Dittmar, T. Improved  
852 mass accuracy and isotope confirmation through alignment of ultrahigh-resolution mass  
853 spectra of complex natural mixtures. *Analytical Chemistry* **2020**, 92 (3), 2558-2565. DOI:  
854 10.1021/acs.analchem.9b04234.

855 (83) Mantini, D.; Petrucci, F.; Pieragostino, D.; Del Boccio, P.; Di Nicola, M.; Di Ilio, C.;  
856 Federici, G.; Sacchetta, P.; Comani, S.; Urbani, A. LIMPIC: a computational method for the  
857 separation of protein MALDI-TOF-MS signals from noise. *BMC Bioinformatics* **2007**, 8 (1),  
858 101. DOI: 10.1186/1471-2105-8-101.

859 (84) Obeid, W. A. Investigation of the potential for algae to produce hydrocarbon based  
860 fuels from algae by hydrous pyrolysis. Doctoral Dissertation, Old Dominion University, 2015.  
861 [https://digitalcommons.odu.edu/cgi/viewcontent.cgi?article=1040&context=chemistry\\_etds](https://digitalcommons.odu.edu/cgi/viewcontent.cgi?article=1040&context=chemistry_etds).

862 (85) Kujawinski, E. B.; Behn, M. D. Automated analysis of electrospray ionization Fourier  
863 transform ion cyclotron resonance mass spectra of natural organic matter. *Analytical*  
864 *Chemistry* **2006**, 78 (13), 4363-4373. DOI: <https://doi.org/10.1021/ac0600306>.

865 (86) Koch, B. P.; Dittmar, T.; Witt, M.; Kattner, G. Fundamentals of molecular formula  
866 assignment to ultrahigh resolution mass data of natural organic matter. *Analytical Chemistry*  
867 **2007**, 79 (4), 1758-1763. DOI: <https://doi.org/10.1021/ac061949s>.

868 (87) Stubbins, A.; Spencer, R. G. M.; Chen, H.; Hatcher, P. G.; Mopper, K.; Hernes, P. J.;  
869 Mwamba, V. L.; Mangangu, A. M.; Wabakanghanzi, J. N.; Six, J. Illuminated darkness:  
870 Molecular signatures of Congo River dissolved organic matter and its photochemical  
871 alteration as revealed by ultrahigh precision mass spectrometry. *Limnology and*  
872 *Oceanography* **2010**, 55 (4), 1467-1477. DOI: 10.4319/lo.2010.55.4.1467.

873 (88) Koch, B. P.; Dittmar, T. From mass to structure: An aromaticity index for high-resolution  
874 mass data of natural organic matter. *Rapid Communications in Mass Spectrometry* **2006**, 20  
875 (5), 926-932. DOI: <https://doi.org/10.1002/rcm.2386>.

876 (89) Koch, B. P.; Dittmar, T. From mass to structure: An aromaticity index for high-resolution  
877 mass data of natural organic matter (Erratum). *Rapid Communications in Mass Spectrometry*  
878 **2016**, 30 (1), 1. DOI: <https://doi.org/10.1002/rcm.7433>.

879 (90) Kim, S.; Kramer, R. W.; Hatcher, P. G. Graphical method for analysis of ultrahigh-  
880 resolution broadband mass spectra of natural organic matter, the van Krevelen diagram.  
881 *Analytical Chemistry* **2003**, 75 (20), 5336-5344. DOI: <https://doi.org/10.1021/ac034415p>.

882 (91) Kim, S.; Simpson, A. J.; Kujawinski, E. B.; Freitas, M. A.; Hatcher, P. G. High resolution  
883 electrospray ionization mass spectrometry and 2D solution NMR for the analysis of DOM

884 extracted by C<sub>18</sub> solid phase disk. *Organic Geochemistry* **2003**, 34 (9), 1325-1335. DOI:  
885 10.1016/s0146-6380(03)00101-3.

886 (92) Cao, D.; Lv, J.; Geng, F.; Rao, Z.; Niu, H.; Shi, Y.; Cai, Y.; Kang, Y. Ion accumulation  
887 time dependent molecular characterization of natural organic matter using electrospray  
888 ionization - Fourier transform ion cyclotron resonance mass spectrometry. *Analytical*  
889 *Chemistry* **2016**, 88 (24), 12210-12218. DOI: 10.1021/acs.analchem.6b03198.

890 (93) Kew, W.; Blackburn, J. W. T.; Uhrin, D. Response to Comment on "Laser  
891 desorption/ionization coupled to FTICR mass spectrometry for studies of natural organic  
892 matter". *Analytical Chemistry* **2018**, 90 (9), 5968-5971. DOI:  
893 10.1021/acs.analchem.8b00332.

894 (94) Zhao, Q.; Thompson, A. M.; Callister, S. J.; Tfaily, M. M.; Bell, S. L.; Hobbie, S. E.;  
895 Hofmockel, K. S. Dynamics of organic matter molecular composition under aerobic  
896 decomposition and their response to the nitrogen addition in grassland soils. *Science of The*  
897 *Total Environment* **2022**, 806, 150514. DOI: <https://doi.org/10.1016/j.scitotenv.2021.150514>.

898 (95) Lu, M.; Zhou, X.; Luo, Y.; Yang, Y.; Fang, C.; Chen, J.; Li, B. Minor stimulation of soil  
899 carbon storage by nitrogen addition: A meta-analysis. *Agriculture, Ecosystems &*  
900 *Environment* **2011**, 140 (1), 234-244. DOI: <https://doi.org/10.1016/j.agee.2010.12.010>.

901 (96) Huang, X.; Terrer, C.; Dijkstra, F. A.; Hungate, B. A.; Zhang, W.; van Groenigen, K. J.  
902 New soil carbon sequestration with nitrogen enrichment: a meta-analysis. *Plant and Soil*  
903 **2020**, 454 (1), 299-310. DOI: 10.1007/s11104-020-04617-x.

904 (97) Van Krevelen, D. W. Graphical-statistical method for the study of structure and reaction  
905 processes of coal. *Fuel Processing Technology* **1950**, 29, 269-228.

906 (98) Mallet, C. R.; Lu, Z.; Mazzeo, J. R. A study of ion suppression effects in electrospray  
907 ionization from mobile phase additives and solid-phase extracts. *Rapid Communications in*  
908 *Mass Spectrometry* **2004**, 18 (1), 49-58. DOI: <https://doi.org/10.1002/rcm.1276>.

909 (99) Sleighter, R. L.; Chen, H.; Wozniak, A. S.; Willoughby, A. S.; Caricasole, P.; Hatcher, P.  
910 G. Establishing a measure of reproducibility of ultrahigh-resolution mass spectra for complex  
911 mixtures of natural organic matter. *Analytical Chemistry* **2012**, 84 (21), 9184-9191. DOI:  
912 <https://doi.org/10.1021/ac3018026>.

913 (100) Tadini, A. M.; Nicolodelli, G.; Marangoni, B. S.; Mounier, S.; Montes, C. R.; Milori, D. M.  
914 B. P. Evaluation of the roles of metals and humic fractions in the podzolization of soils from  
915 the Amazon region using two analytical spectroscopy techniques. *Microchemical Journal*  
916 **2019**, 144, 454-460. DOI: <https://doi.org/10.1016/j.microc.2018.10.009>.

917 (101) Pinho, S. P.; Macedo, E. A. Solubility of NaCl, NaBr, and KCl in water, methanol,  
918 ethanol, and their mixed solvents. *Journal of Chemical & Engineering Data* **2005**, 50 (1), 29-  
919 32. DOI: 10.1021/je049922y.

920 (102) Goranov, A. I.; Wozniak, A. S.; Bostick, K. W.; Zimmerman, A. R.; Mitra, S.; Hatcher, P.  
921 G. Microbial labilization and diversification of pyrogenic dissolved organic matter.  
922 *Biogeosciences* **2022**, 19 (5), 1491-1514. DOI: 10.5194/bg-19-1491-2022.

923 (103) Sleighter, R. L.; Liu, Z.; Xue, J.; Hatcher, P. G. Multivariate statistical approaches for  
924 the characterization of dissolved organic matter analyzed by ultrahigh resolution mass  
925 spectrometry. *Environmental Science & Technology* **2010**, 44, 7576-7582.

