DOI: 10.1016/j.envres.2023.116315

Environmental Research

Molecular Composition and Transformations of Labile Soil Organic Matter Fractions in Mediterranean arable soils: Agronomic and Environmental Implications

--Manuscript Draft--

Manuscript Number:	ER-22-11152
Article Type:	VSI: SoilSci-EnvRes
Section/Category:	Environmental Chemistry and Toxicology
Keywords:	Soil organic carbon; Particulate organic carbon; Molecular composition; Land use; Climatic changes; Environmental and agronomic relevance
Corresponding Author:	Hamada Abdelrahman, PhD Cairo University Giza, EGYPT
First Author:	Hamada Abdelrahman, PhD
Order of Authors:	Hamada Abdelrahman, PhD
	Diana Hofmann, PhD
	Rachel L. Sleighter, PhD
	Dan C. Olk, PhD
	Anne E. Berns, PhD
	Teodoro Miano, PhD
	Sabry M. Shaheen, PhD
	Claudio Cocozza, PhD
Abstract:	With the increased global interest in sequestering carbon in soil and reducing its emission, it is necessary to understand the composition and transformations of different pools of soil organic matter (SOM). To explore in detail the chemical composition of agroecologically relevant yet distinct fractions of SOM, the light fraction of SOM (LFOM), the 53-μm particulate organic matter (POM), and the mobile humic acid (MHA) fractions were sequentially extracted from soil collected for agricultural soils and characterized using both 13C cross polarization magic angle spinning nuclear magnetic resonance (CPMAS NMR) spectroscopy and also electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI-FT-ICR-MS). The 13C CPMAS NMR results showed a decrease in the O-alkyl C region assigned to carbohydrates (51–110 ppm) and an increase in the aromatic region (111–161 ppm) proceeding from the LFOM to the POM and then to the MHA fraction. Similarly, based on the thousands of molecular formulae assigned to the peaks detected by FT-ICR-MS, condensed hydrocarbons were dominant only in the MHA, while aliphatic formulae were abundant in the POM and LFOM fractions. The molecular formulae of the LFOM and POM were mainly grouped in the high H/C lipid-like and aliphatic space, as the expected carbohydrate signatures were likely not detected due to low ionization efficiencies. For a portion of the MHA compounds, the double bond equivalent (DBE) values were extremely high (17–33, average of 25), corresponding to low H/C values of 0.3–0.6 that are representative of condensed hydrocarbons. The molecular formulae of lignin-like compounds were essentially negligible in all fractions. The chemical similarities and differences between the LFOM, POM, and especially the MHA may explain the process of organic C stabilization in soil. Their suites of chemical traits, as captured here, illustrates the environmental and agronomic relevance of these SOM fraction and their roles in depicting short- and mid-term changes responses of SOM to management- and climate induced changes.
Suggested Reviewers:	Giancarlo Renella, PhD Prof. of Soil and Environmental Sciences, University of Padua giancarlo.renella@unipd.it Expert on the topic
	Vasileios Antoniadis, PhD

	Associate Professor, University of Thessaly vasilisrev@yahoo.com Expert on soil organic matter cycling and environmental monitoring
	Jifu Ma, PhD Tianjin University majf@ieecas.cn Expert on soil quality and organic matter
	Claudio Zaccone, PhD Associate Professor, University of Verona claudio.zaccone@univr.it Expert on soil organic matter.. President, Soil System Sciences division, European Geosciences Union (EGU)
	Samir Haddad, PhD Associate Professor, Minia University samir.mohamed@mu.edu.eg Expert on soil organic matter cycling in response to soil (micro)biology

December 21, 2022

Cover Letter

To the Editor in Chief
Environmental Research

Dear Editor-in-Chief, dear Managing Guest Editor Prof. Avelino Núñez-Delgado,

many thanks for the kind invitation to contribute to the Virtual Special Issue **"Soil Science and Environmental Research: VSI: SoilSci-EnvRes**" of Environmental Research.

Please find enclosed of our manuscript entitled:

Molecular Composition and Transformations of Labile Soil Organic Matter Fractions in Mediterranean arable soils: Agronomic and Environmental Implications,

written by
Hamada Abdelrahman, Diana Hofmann, Rachel L. Sleighter, Dan C. Olk, Anne E. Berns, Teodoro Miano, Sabry M. Shaheen, Claudio Cocozza,

to be submitted as an original research paper to the VSI: SoilSci-EnvRes of Environmental Research.

With the increasing interest in soil organic carbon sequestration and the race toward reducing emission, it is important to understand the nature and transformation of different SOM fractions. We explored the chemical composition of SOM fractions in arable soils and discussed its agricultural and environmental relevance. Specifically, the light fraction of SOM (LFOM), the 53-µm particulate organic matter (POM), and the mobile humic acid (MHA) fractions were extracted sequentially then characterized using ^{13}C NMR and FT-ICR-MS. The analytical techniques allowed describing the molecular compositio and distinguish LF, POM and MHA: i) aliphatic substances were common in the three fractions but were most abundant in the LFOM and POM; ii) molecular formulae assigned to the lignin range were practically negligible; and iii) molecular formulae assigned to condensed aromatic hydrocarbons were characteristic of only the MHA. Hence, the MHA contained both labile fragments with H/C >1.5, as well as more condensed aromatic compounds with m/z >400 and DBE >17 (>4 rings), which suggest both labile and recalcitrant nature of MHA.
The composition and transformations of SOM fractions, reported in this work, are be agronomically and ecologically relevant as it helps predicting the sensitivity and response of individual SOM fraction to climatic and land use changes, and thus, can help defining conservative agronomic management and environmental monitoring.
The topic is novel and highly relevant for soil and environmental scientists, and thus fits well with the scope of the targeted journal and special issue.
The manuscript is an original work that has not been published previously and it is not under consideration for publication elsewhere.

The publication has been approved by all co-authors.
I will serve as corresponding author during the review process.

Kind regards
Corresponding author
Hamada Abdelrahman, PhD
Associate Professor, Cairo University
Corresponding author
hamada@cu.edu.eg

Highlights

- The molecular composition and transformations of SOM fractions are explored
- SOM labile fractions, LFOM, POM & MHA, were sequentially extracted
- LFOM and POM are mainly composed of aliphatic compounds (DBE <10, H/C 1.5–2)
- MHA contained labile fragments (H/C >1.5) & condensed aromatics (DBE >17; >4 rings)
- Molecular changes in SOM can be used to monitor agronomic & environmental impacts

Molecular Composition and Transformations of Labile Soil Organic Matter Fractions in Mediterranean arable soils: Agronomic and Environmental Implications

Hamada Abdelrahman[a *], Diana Hofmann[b], Rachel L. Sleighter[c], Dan C. Olk[d], Anne E. Berns[b], Teodoro Miano[e], Sabry M. Shaheen[f], Claudio Cocozza[e]

[a] Soil Science Department, Faculty of Agriculture, Cairo University, Giza 12613, Egypt

[b] Forschungszentrum Jülich GmbH, Institute of Bio- and Geosciences, IBG-3: Agrosphere, Jülich, Germany

[c] FBSciences, Inc., Research and Development, Norfolk, VA, USA

[d] USDA–ARS, National Laboratory for Agriculture and the Environment, Ames, IA, USA

[e] DiSSPA–Università degli Studi di Bari "Aldo Moro", Bari, Italy

[f] University of Wuppertal, Laboratory of Soil- and Groundwater-Management, Pauluskirchstraße 7, 42285 Wuppertal, Germany

[*] Corresponding Author:

Hamada Abdelrahman, hamada@cu.edu.eg, ORCID: 0000-0002-6069-7239

Abstract

With the increased global interest in sequestering carbon in soil and reducing its emission, it is necessary to understand the composition and transformations of different pools of soil organic matter (SOM). To explore in detail the chemical composition of agroecologically relevant yet distinct fractions of SOM, the light fraction of SOM (LFOM), the 53-µm particulate organic matter (POM), and the mobile humic acid (MHA) fractions were sequentially extracted from soil collected for agricultural soils and characterized using both ^{13}C cross polarization magic angle spinning nuclear magnetic resonance (CPMAS NMR) spectroscopy and also electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI-FT-ICR-MS). The ^{13}C CPMAS NMR results showed a decrease in the O-alkyl C region assigned to carbohydrates (51–110 ppm) and an increase in the aromatic region (111–161 ppm) proceeding from the LFOM to the POM and then to the MHA fraction. Similarly, based on the thousands of molecular formulae assigned to the peaks detected by FT-ICR-MS, condensed hydrocarbons were dominant only in the MHA, while aliphatic formulae were abundant in the POM and LFOM fractions. The molecular formulae of the LFOM and POM were mainly grouped in the high H/C lipid-like and aliphatic space, as the expected carbohydrate signatures were likely not detected due to low ionization efficiencies. For a portion of the MHA compounds, the double bond equivalent (DBE) values were extremely high (17–33, average of 25), corresponding to low H/C values of 0.3–0.6 that are representative of condensed hydrocarbons. The molecular formulae of lignin-like compounds were essentially negligible in all fractions. The chemical similarities and differences between the LFOM, POM, and especially the MHA may explain the process of organic C stabilization in soil. Their suites of chemical traits, as captured here, illustrates the environmental and agronomic relevance of these SOM fraction and their roles in depicting short- and mid-term changes responses of SOM to management- and climate induced changes.

Keywords: Soil organic carbon; Particulate organic carbon; Molecular composition; Land use; Climatic changes; Environmental and agronomic relevance

1. Introduction

Soil organic matter (SOM) content is one of the main factors of soil chemical, physical, and biological properties. Furthermore, SOM is a major pool of global C; it contains more than three times as much C as either the atmosphere or terrestrial vegetation (Prăvălie et al., 2021). However, the mechanisms responsible for its formation and decomposition, and its agroecological relevance are still a matter of debate (Kleber and Lehmann, 2019; Olk et al., 2019). The composition and roles of SOM have been investigated for decades, but information on its transformation and response to changes remain controversial (Berns and Knicker, 2014). To understand the role of SOM in the soil ecosystem and its response to land use and climatic changes, it is necessary to identify the composition of SOM and/or its fractions (Mustafa et al., 2023). The obstacle is that only a portion of SOM components can be assigned to known chemical groups, while other components have lost their original structure. Microbial processing and possibly chemical alterations result in a gradual loss of relatively labile compounds, such as amino acids and carbohydrates, leading to enrichment of more recalcitrant compounds, such as unsubstituted aromatic rings and carboxyl functionalities (Ikeya et al., 2015).

As SOM is made of different pools, isolation of specific sub-fractions of SOM that have different yet well-defined chemical characteristics and turnover rates, can allow depicting various stages of organic matter development in soil, as affected by land management, can be depicted (Mustafa et al., 2023; Schnitzler et al., 2007). (Cao et al., 2011) introduced an integrated physical-chemical fractionation procedure to sequentially separate the light fraction of soil organic matter (LFOM), the 500- and 53-µm particulate organic matter (POM), and the mobile humic acid (MHA) fraction.

The LFOM and POM fractions consist of partially decomposed plant residues and are physically separated from soil by size and density (Gregorich et al., 2006), while the MHA is composed of more stabilized materials (Abdelrahman et al., 2016; Olk, 2006) and is isolated through NaOH extraction and subsequent purification steps. Results have shown that the MHA and the Ca-bound humic fraction contributed disproportionately to the aromatic portion of the ^{13}C NMR spectrum for whole SOM, while the LFOM and POM contributed to the aliphatic portion (Cao et al., 2011), suggesting that the joint use of these fractions can better describe whole SOM than can either a stabilized SOM approach or the LFOM/POM alone.

The masses of the LFOM and the POM fractions have been found to respond to changes in land management and land use over the short-term (Abdelrahman et al., 2020; Cao et al., 2011) and thus can serve as indicators for such short-term management-induced changes in soil C. Published evidence indicates that the MHA has ecological meaning and can contribute to a better understanding of SOM and nutrients cycling and accumulation across different climatic and environmental settings (corn and cotton in USA, rice in Philippines, Wheat in Italy; see case studies in Olk et al., 2019). Further studies investigated the occurrence of specific chemical structures, such as carbohydrates and amino compounds in these fractions (e.g., Abdelrahman et al., 2017, 2016). Yet, more detailed characterization of their compositions and transformation may enhance our understanding of the ecological role and relevance these fractions play in response to management.

The ^{13}C NMR spectroscopy has been widely applied as a non-destructive technique to characterize SOM composition in solid phases (Berns and Knicker, 2014; Mustafa et al., 2022) and has been applied to LFOM (Wang et al., 2012), POM (Yeasmin et al., 2020), and MHA fractions (Cao et al., 2011). These studies showed that LFOM and POM were strongly aliphatic, while in contrast the MHA contained more carboxyl and aromatic functional groups. Another advanced technique, electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI-FT-ICR-MS), has attracted attention for SOM characterization due to its ultrahigh mass resolution and accuracy at the molecular level (Fox et al., 2017; Kaplan et al., 2016). FT-ICR-MS is a sophisticated characterization tool for assigning molecular formulae to thousands of individual peaks that can be detected in the mass spectrum of complex mixtures (Sleighter and Hatcher, 2011). Its application has led to new insights into the molecular composition of different natural organic matter sources, including water, soil, and aerosols (Koch et al., 2007). FT-ICR-MS has been mainly used for dissolved organic matter from aqueous environments, however, it is increasingly used to characterize soil extracts (Fernández et al., 2008; Ohno et al., 2010; Seifert et al., 2016) and humic substances (Ikeya et al., 2015, 2013) where soil humic acids (HAs) showed molecular formulae with H/C and O/C ratios similar to condensed hydrocarbons. However, only one study (Ohno et al., 2010) used FT-ICR-MS to characterize the MHA, comparing its chemical nature to those of water-extractable organic C from plants and soils. In that study, the MHA were enriched in lipid and condensed

aromatic components but depleted in lignin-like compounds, whereas the aqueous plant and soil extracts contained more diverse mixtures of lipid-, protein-, carbohydrate-, and lignin-like compounds.

With the global increasing interest ins sequestering carbon in soil, reducing emissions, and the development of carbon market with its projects and methodologies, it is significantly important to understand the composition and transformations of different SOM fractions in response to management and climate induced changes. In particular, the Mediterranean, basin and agroecosystem worldwide, is characterized by climatic, pedogenic and agricultural management that seem to favor the loss of soil organic matter, which results in significant environmental and financial impacts (Ferreira et al., 2022; Jones et al., 2012). Although of available practices and measure to conserve SOM, soils around the Mediterranean continue to be depleted from SOM and nutrients. Therefore, an understanding of SOM pools composition and transformation is seriously necessary.

Previous studies have characterized some of SOM fractions (e.g., Mustafa et al., 2023; NDZELU et al., 2022; Savarese et al., 2021) but not to extent we present here. To our knowledge, no previous work has presented molecular-level characterization of a suit of SOM fraction that represent SOM continuum. Our work utilized FT-ICR-MS to elucidate the molecular composition and transformations of three different SOM pools sequentially extracted from the same soil sample. We hypothesized that the LFOM and POM might share some compositional similarities, whereas the MHA will be compositionally different. Consequently, the objective of this work was to describe the molecular compositions of the LFOM, POM, and MHA, using the combination of ^{13}C CPMAS NMR and FT-ICR-MS to gain better insights into their chemical characteristics and transformations. The findings can allow for a better understanding of the nature of these SOM fractions and their cycling in different agricultural and ecological systems. Comparing the progression in chemical characteristics from the physically isolated LFOM and POM to the mineral bound MHA fraction allows a better understanding of organic C stabilization in soils, which is very relevant for C retention in soil.

2. Materials and Methods

2.1. Site description and sampling

Soil samples were collected from two experimental field trials (Foggia and Metaponto) at different locations in the South of Italy (41°27′35″ N and 15°30′18″ E; 40°24′25″ N and 16°48′24″ E). The soils are Vertisols, specifically Typic Calcixererts (Soil Survey Staff 2014) derived from alluvium with silt and clay texture.

The cultivation history of both soils is a rotation of lentil (*Lens culinaris*) and wheat (*Triticum* sps.). In 2009, the experimental fields were converted to organic management, having fertilization treatments for different types of organic fertilizers, with three field replicates. Annual fertilization was applied to provide 100 kg N ha^{-1} to each wheat crop, or 13.1 kg P_2O_5 ha^{-1} to each lentil crop. This study focuses on compost and organo-mineral fertilizer treatments, as the common practice in Mediterranean management. In the organo-mineral fertilizer, the N component was collagen-based while the P component was collagen- and ground rock phosphate-based, and both were applied to provide the required N or P to the designated crop.

Soils were sampled (0–30 cm depth) after harvest where each soil sample was as a composite of three cores each with two additional field replicates (collected around the center of the plot), air-dried, ground to pass through a 2-mm sieve, and then stored for subsequent analyses. More information about soils, field trials, fertilizers, and sampling are provided elsewhere (Abdelrahman et al., 2016).

2.2. SOM extraction and characterization

Soils were sequentially extracted for the LFOM, 500–53 μm POM, and MHA fractions. Extraction of the LFOM and POM fractions was performed with modification of the procedure by Cao et al. (2011), as described by Abdelrahman et al. (2016). The LFOM was separated by density (1.60 g mL^{-1}) using Na polytungstate followed by dispersion in Na metaphosphate and wet sieving separation of the 500–53 μm POM. The remaining silt + clay size particles were dried overnight and then extracted for the MHA fraction through overnight shaking in 0.25 M NaOH under an N_2 atmosphere, followed by decantation and acidification of the supernatant to pH 1.0 using 2.0 M HCl. The humic fraction precipitates were separated through centrifugation and then cleansed of soil contaminants through re-

solubilization in KOH and re-precipitation by HCl, followed by dialysis for three days in successively weaker HCl solutions and then deionized water (Mao et al., 2008). Finally, the MHA samples were lyophilized. Each field replicate was extracted separately for the SOM fractions. Averaged across both sites and both fertilizer treatments, the LFOM, POM, and MHA fractions contained 331, 307, and 442 g C kg^{-1} fraction and 18, 24, and 36 g N kg^{-1} fraction, respectively (Abdelrahman et al., 2017, 2016).

2.3. ^{13}C Cross-Polarization Magic-Angle Spinning (CPMAS) NMR

Due to the reported similarities (Abdelrahman et al., 2017, 2016) in the same fraction among different sites/treatments, only four samples of each fraction representing sites/treatments were characterized using ^{13}C CPMAS NMR. Spectra were obtained at a ^{13}C resonance frequency of 75.4 MHz on a 7.05 T Varian INOVA™ Unity (Varian Inc., Palo Alto, CA, USA). An HX Apex probe was used with a stator holding 6 mm diameter cylindrical zirconia Pencil® rotors with Vespel® drive tips. Samples were spun at 8000 ± 3 Hz at 22°C. The spectra were collected with a sweep width of 25 kHz and an acquisition time of 20 ms. The optimal contact time and recycle delay for the cross-polarization experiment were determined in preliminary experiments to be 1 ms and 5 s, respectively. During cross-polarization, the ^{1}H radio frequency (RF) field strength was set to 47.0 kHz and the ^{13}C RF field strength to 41.1 kHz. To compensate for inhomogeneities of the Hartmann-Hahn condition, an ascending ramp of 15.3 kHz on the ^{1}H-RF field was used (Berns and Conte, 2011). Proton decoupling was done using a SPINAL sequence with a ^{1}H field strength of 50.4 kHz, a phase of 4.5°, and a pulse length of 12 µs.

The free induction decays were recorded by VnmrJ (Version 1.1 RevisionD, Varian Inc., Palo Alto, CA, USA) and processed by Mestre-C (Version 4.9.9.9, Mestrelab Research, Santiago de Compostela, Spain). Fourier transformation was done with an exponential filter function with a line broadening (LB) of 25–50 Hz. Baseline correction was done using the manual baseline correction function of Mestre-C. The ^{13}C chemical shifts are reported relative to tetramethylsilane (=0 ppm) using adamantane as an external reference. The relative intensities of the regions were determined using the integration routine of the MestReC software. When spinning sidebands (ssb) were present,

188 the relative intensities were corrected for the ssb as described in Berns and Conte (2011). The NMR
189 spectra were divided into seven regions: aliphatic (45–0 ppm), N-alkyl and methoxy (64–45 ppm), O-
190 alkyl (90–64 ppm), di-O-alkyl (108–90 ppm), aromatics (161–108 ppm), COO/N–C=O (190–161
191 ppm), and ketones and aldehydes (215–190 ppm). Aliphaticity [Aliphatic C peak area (0–110
192 ppm)]100/[Total peak area (0–160 ppm)] and aromaticity [Aromatic C peak area (110–160
193 ppm)]100/[Total peak area (0–160 ppm)] of samples were calculated according to González Pérez et
194 al. (2004).
195
196 **2.4. Molecular characterization using FT-ICR-MS**
197 Due to the reported similarities (Abdelrahman et al., 2017, 2016) in the same fraction among
198 different sites/treatments, only four samples of each fraction representing sites/treatments were
199 dissolved in an adequate amount of methanol in an ultrasonic bath under heat without further cleanup
200 steps to prevent further sample loss (especially of more polar compounds) that commonly occurs
201 during solid phase extraction. These methanol solutions were introduced by flow injection with a
202 syringe pump at 8 µl/min into the FT-ICR-MS (LTQ-FT Ultra, ThermoFisher Scientific), equipped
203 with an electrospray ionization (ESI) source and a 7 T superconducting magnet. Negative ion mode
204 was utilized to prevent the dual detection of protonated and sodiated [$(M+H)^+$ and $(M+Na)^+$] ions that
205 occur in positive ion mode.
206 All samples were analyzed consecutively during one day (and repeated the next day with
207 comparable results) under the following conditions: spray voltage/capillary voltage/tube lens 2.9 kV/-
208 50 V/-130 V, respectively; sheath-gas 3 arb, without aux-gas or sweep gas; transfer capillary 275°C;
209 mass range 200–1000 Da. To prevent sample carry-over, flushing of the spray capillary was done first
210 with an isopropanol-methylene chloride-chloroform mixture, followed by 10–20 fold with pure
211 methanol. Prior to and between analyses, blanks were analyzed to exclude any memory effect in the
212 form of a DOM pattern.
213 Detailed information on external and internal calibration can be found in the Supplementary
214 Information. Peaks at several randomly selected consecutive recalibrated masses (odd and
215 corresponding even nominal masses) were first characterized manually in order to confirm the actual

mass accuracy and to determine the proper constraints in the formula assignment program (in house-developed Scilab routines). This test showed that the most intense peaks were mainly CHO peaks. Therefore, the number of heteroatoms was permitted as follows: carbon, hydrogen, and oxygen were unlimited, sulfur ≤ 1, nitrogen ≤ 2, and P=0. The molecular formula rules outlined by Stubbins et al. (2010) eliminated the vast majority of duplicate formulae for a single peak. In the cases where duplicate formulae existed (only 0.1–1.3% of the peaks), the formula extension approach (Kujawinski and Behn, 2006) was utilized, where the formula chosen was that which fell into a CH_2 homologous series.

Once molecular formulae had been assigned, various visualization diagrams were plotted. The van Krevelen diagram plots the elemental ratios of H/C and O/C of the formulae, where this alignment can be representative of a biochemical compound class. Kendrick mass defect (KMD) analysis separates components that have similar skeletal formulae yet different quantities of a specific functional group (Ikeya et al., 2015). CH_2 and COO are two common repeating units conventionally used during KMD analysis. To calculate the KMD for the CH_2 group, its Kendrick mass (KM) is first calculated, leading to calculation of its KMD value through the following equations (Ikeya et al., 2015):

$$KM(CH_2) = (\text{calculated mass}) \times [(\text{nominal mass of } CH_2) / (\text{exact mass of } CH_2)] \quad (1)$$

$$KMD(CH_2) = KM(CH_2) - [\text{nominal } KM(CH_2)] \quad (2)$$

Formulae of the same family, for example the members of an alkylation series, have the same Kendrick mass defect but different nominal Kendrick masses and therefore are positioned along a horizontal line in the Kendrick (CH_2) plot. Horizontal lines of different Kendrick mass defects correspond to formulae of different composition, such as degrees of saturation or heteroatom content.

Double bond equivalent (DBE) is calculated as the sum of the number of ring structures plus the number of double bonds per molecular structure, described as $C_cH_hN_nO_oS_sP_p$ (Ikeya et al., 2015), by the following equation:

$$DBE = c - h/2 + n/2 + p/2 + 1 \quad (3)$$

As described by Koch and Dittmar (2006), aliphatic compounds have DBE/C<0.3 and H/C>1, and a modified aromaticity index, AI_{mod}, can be used to distinguish aromatic (AI_{mod} 0.5–0.67) from condensed aromatic ($AI_{mod} \geq 0.67$) entities.

3. Results and Discussion

3.1. ^{13}C CPMAS NMR of SOM fractions

The ^{13}C CPMAS NMR spectra of the LFOM fractions (Fig. S1A) were dominated by the signal at 72 ppm (O-alkyl) and the N, O-substituted alkyl structures in general (108-90, 90-64 and 64-45 ppm). The largest signal around 72 ppm corresponds to the chemical shift of C-OH groups, mostly found in carbohydrates, but also in lignin side chains. The sharp signal at 105 ppm originates from di-O-alkyl bridges in polymeric carbohydrate chains. These chemical structures are found in cellulose, hemicellulose, starch, pectin, and lignin, though the latter usually contributes the least. The NMR spectrum of the LFOM hence confirmed the reported nature of the LFOM as partially decomposed plant material (Gregorich et al., 2006). The signal at 72 ppm was also the strongest signal in the spectra of the POM fractions (Fig. S1B), but the overall N,O-substituted alkyl regions (108-90, 90-64 and 64-45 ppm) were greatly reduced. This relative reduction of the carbohydrate regions indicates a slightly more advanced state of decomposition of the POM fraction compared to the LFOM fraction.

Fig. 1

The MHA spectra (Fig. S1C) were characterized by a very large and broad signal centered at 130 ppm indicating aromatic and/or unsaturated C. The aromatic region (161-108 ppm) encompassed 46–49% of the total spectral area in the MHA fraction from the Foggia plots and the unsubstituted aliphatic compounds (45–0 ppm) accounted for around 15% (Table S1). Although it is not the aim of the work to report variations between the same fraction from different sites, significant observations are discussed. In fact, the MHA from the Metaponto plots also displayed significantly higher amounts of aromatic compounds (31-32%) compared to the LFOM and POM fractions, while the relative amounts of the unsubstituted aliphatic region (45-0 ppm) were around 26-27%. Aromatic and unsubstituted aliphatic compounds require larger amounts of oxygen and energy to be metabolized and degraded by microorganisms and hence tend to accumulate in soil. In combination with the

strongly reduced carbohydrate regions, the MHA spectra were consistent with stabilized C fractions (Sonsri et al., 2022). As the three fractions were isolated sequentially from the same sample, we postulate that these fractions represent a continuum of organic matter stabilization in soil.

The chemical compositions of each fraction at each site were not remarkably different between the compost and fertilizer treatments. Similarly, the spectra of LFOM and POM fractions from each fertilizer treatment did not differ notably between the two field sites. Only the MHA fraction differed visibly between the two sites. Comparison of the relative signal intensities (Table S1, Fig. S2) showed that in both treatments the MHA fraction were significantly enriched in aromatic compounds (161-108 ppm) at the Foggia site, while at the Metaponto site it showed significantly higher contents of substituted and unsubstituted aliphatics (45-0, 64-45, and 90-64 ppm). The lack of fertilizer treatment effects on the chemical compositions of the fractions was surprising, as especially the LFOM and POM fractions of the compost treatments were expected to display significantly higher signals in the carbohydrate region stemming from degraded plant material. However, the collagen-base (i.e., protein-base) of the fertilizer probably generated similar signals, as signals from amino acids also fall within this region. Another explanation likely lies in the fractionation procedure itself. As these fractions are extracted through flotation and wet sieving, compounds soluble in aqueous solutions (i.e., containing hydrophilic polar groups) are removed and hence part of the SOM is not recovered.

3.2. General characteristics of the FT-ICR-MS spectra

Similar to the observations for the CPMAS NMR spectra, the FT-ICR mass spectra (data not shown) were quite similar within each fraction for the different sites and compost/fertilizer treatments. For all LFOM samples, most of the peaks were detected in the *m/z* range of 500–650 Da. Some signals with high magnitude [>40% of the base (i.e., largest) peak in each spectrum] were observed at nominal masses 507, 509, 511, 535, and 537. These peaks corresponded to long-chain dicarboxylic acid formulae with predominantly aliphatic character (e.g., $C_{32}H_{64}O_4$, H/C=2.0, O/C=0.125). The POM spectra were mainly in the *m/z* range of 400–800 Da, with some signals having higher magnitudes at 250–300 Da. Peaks in the MHA spectra were detected mostly in the *m/z* ranges of 370–440 and 480–700 Da. For the POM and MHA samples, high magnitude peak clusters were detected at

nominal masses 704, 819, and 935. These peaks possibly originated from triply charged Na polytungstate that had been used to separate the POM but remained in the POM and subsequently extracted MHA fraction, despite multiple washing steps. These peaks, as well as other high mass defect peaks of inorganic origin due to possible incomplete de-salting (Sleighter and Hatcher, 2011), were manually deleted from the peak lists.

Due to the presence of salt peaks and the inherent similarities among the spectra for each fraction type, only peaks that were present in all samples of the same fractional type were analyzed further. By doing so, comparisons between the LFOM, POM, and MHA fractions were more reliable, since sample preparation and instrumental reproducibility were ensured by removing peaks that were only detected in 1–3 samples of the same fractional type.

3.3. van Krevelen plots of SOM fractions

From the molecular formulae assigned to the common peaks detected in all samples of the same fraction for sites/treatments, van Krevelen diagrams were constructed (Fig. 2a), which are colored according to their heteroatom content. Table 1 summarizes the averaged mass spectral characteristics (O/C, H/C, DBE, modified aromaticity index: AI_{mod}), as well as the percentages of formulae that fall into various categories (%CHO, %CHON, %CHOS, aliphatic vs. aromatic). A conceptual van Krevelen diagram depicting the associated regions is given in Fig. S3. The CHO compounds (black points in Figs. 2-4) were by far the most abundant components, accounting for 90–92% of all formulae in the LFOM and POM and about 83% of the formulae in the MHA (Table 1). Formulae containing S and N (blue and red points, respectively, in Figs. 2-4) were less abundant and most fell into the aliphatic category with high H/C.

The van Krevelen diagrams for the compiled formulas of the LFOM and POM fractions (Fig. 2a) were similar, having average O/C and H/C values of 0.24–0.26 and 1.69–1.72, respectively. Most of these formulae fell into the lipid-like (H/C 1.5–2.0, O/C 0–0.3) aliphatic region, as previously reported (Hockaday et al., 2009; Ohno et al., 2010; Sleighter and Hatcher, 2008).

In contrast, the MHA molecular formulae were grouped in two clearly separated ranges: the lipid-like region and the condensed hydrocarbon region (H/C 0.2–0.7, O/C 0.0–0.67) often categorized as

black carbon. Carbohydrates (especially those without carboxylic groups), which have high O/ C and high H/C, generally cannot adequately compete for a charge during the ESI process of complex mixtures and are therefore not detected in the mass spectra here, or those reported elsewhere (Hockaday et al., 2009; Ohno et al., 2010). The van Krevelen plot of the MHA fraction reported by Ohno et al. (2010) also showed molecular formulae that clustered mostly in the two separate regions of lipid-like and condensed hydrocarbon spaces, with some also aligning in the lignin-like central region (H/C 0.7–1.5, O/C 0.1–0.67). Consistently, (Olk et al., 2006) used ^{13}C CPMAS NMR to find the MHA of paddy rice soils was enriched in phenolic lignin residues, suggesting incomplete decomposition as phenolic lignin compounds are normally oxidized and degraded during the humification processes (DiDonato et al., 2016). The present MHA fraction showed scarce formula assignments in the lignin space. This lack of lignin assignments might be related to the fully aerobic soil conditions, which would promote lignin oxidation (DiDonato et al., 2016). Having essentially no molecular formulae assigned to the lignin space in the MHA plot (Fig 2a) indicated that the MHA had progressed beyond initial humification status, which is in line with Ohno et al. (2010) who concluded that formation of the MHA is marked by a decrease in lignin components and an increase in condensed aromatic components.

D'Andrilli et al. (2015) introduced the molecular lability boundary (MLB) to differentiate between molecular constituents of natural organic matter that vary in intrinsic lability. The MLB is set at H/C=1.5, so formulae having H/C>1.5 are defined as more labile compounds (lipid-like and aliphatic region) while formulae having H/C<1.5 are defined as more recalcitrant materials in the aromatic and condensed aromatic regions. The proportion of formulae in the labile region appear most pronounced in the POM (93% of formulae have H/C$\geq$1.5) similar to LFOM (89% of formulae have H/C$\geq$1.5) but in contrast to MHA (74% of formulae have H/C$\geq$1.5), as shown in Table 1 and Fig. 2a.

Intrinsic biochemical lability of compounds can differ from their degradability in soils, likely depending on their location within the soil matrix, their strength of binding to the soil mineral surfaces or within SOM, and the presence of specific microbial enzymes. Yet D'Andrilli et al. (2015) provided some validation for the MLB by associating relative labilities estimated through this process for dissolved organic carbon samples with their degradation rates during weekly incubations.

Fig. 2b shows the relationship between H/C ratios and *m*/*z* for the common formulas of the SOM fractions. The molecular formulae with H/C>1.5 were distributed across the whole range of 250–800 Da for all fractions. The molecular formulae in the MHA with low H/C (0.3–0.6), which are typical for condensed aromatics, were observed only within the *m*/*z* range of 400–800 Da.

Table 1

Fig. 2

3.4. The Kendrick mass analyses

The LFOM and POM fractions clearly differed from the MHA fraction in their relationships between the KMD (CH_2) and the nominal Kendrick mass (Fig. 3a). The LFOM and POM compounds were situated in one wide band extending to higher KMD values, which corresponds to the high H/C aliphatic compounds shown in the van Krevelen diagrams (Fig. 2a). This wide band was also present in the MHA fraction, but a second, narrower band having considerably lower KMD (0.1–0.5) was distinctly separate, correlating to the condensed aromatics at low H/C in the van Krevelen diagram. Formulae that aligned at the same KMD value across the mass range are indicative of an alkylation series. Formulae with high KMD (i.e., high H/C) have longer CH_2 homologous series covering a wider mass range. The lower KMD (i.e., low H/C) formulae have fewer components in each homologous series, because condensed aromatic ring systems have less opportunity for CH_2 expansions.

Fig. 3

3.5. The distribution of double bond equivalent (DBE) values

As shown in Fig. 3b, DBE values increased progressing from the LFOM and POM (average DBE of approximately 6, with values as high as 25) to the MHA (average DBE of 9, with values up to 33). The vast majority of molecular formulae in the LFOM and POM had DBE $\leq$11, which held true across the entire mass range. The MHA fraction showed two clearly different DBE regions of molecular formulae, whether graphed against *m*/*z*, H/C, or O/C (Figs. 3-4), similar to the Inogashira humic acid as reported by Ikeya et al. (2012). The MHA region with lower DBE values (mostly $\leq$12 across the *m*/*z* range of 300–800 Da with higher H/C of 1.4–2.1, Fig. 4a) is characteristic of aliphatic

lipid-like components, consistent with those found in the LFOM and POM (Ikeya et al., 2012). The second region – found exclusively in the MHA – exhibited extremely high DBE values (from 17 at *m*/*z* 400 Da up to 33 at *m*/*z* 800 Da), low H/C values (0.33–0.60), and higher O/C values (0.35–0.65, Fig. 4b). This set of traits indicates components having yet higher degrees of microbially driven oxidation and condensation (corresponding to a higher number of double bonds) than the results reported by Ikeya et al. (2015)

Fig. 4

Exemplary structures of specific formulae with exceptionally high DBE values in the MHA fraction are provided in Fig. 5, in order to illustrate the types of structures that could give rise to such DBE values. It should be noted that many other structural isomers are also possible. In general, the low H content dictates that these structures be condensed polycyclic aromatic hydrocarbons (PAHs) with abundant double bonds and rings. Simultaneously, abundant oxygen – often as much as hydrogen – is also incorporated, accounting for the high O/C ratios. These presumptions are not incongruent with the NMR results, as we expect these condensed PAHs to represent only a subset of all inherent organic molecules in the MHA fraction.

Fig. 5

Presuming that in general carbonyl compounds (as in Fig. 5c) as a reactive species should be rather unstable in soil, and therefore seldom found, and above all, their ionization efficiencies in negative ion mode are lower than carboxylic groups, we hypothesize that these high DBE formulae are oxygenated condensed PAHs bearing mostly hydroxyl- and carboxylic side groups as postulated in Fig. 5a–b. These condense PAHs shall be consistent with the significant presence of carboxyl C in the NMR spectrum (Fig. 1) and of aromatic C in both the NMR spectrum and the van Krevelen diagram (Fig. 2a) of the MHA. Such C contents between 59 and 65% in the examples given in Fig. 5 are similar to those of humic substances, where the humic acids commonly contain >55% C.

The DBE–H/C plots (Fig. 4a) suggest the sequential development of SOM fractions from the plant-derived and physically non-sequestered fractions (LFOM and POM) to the bound MHA fraction. The molecular formulae of the LFOM and POM were mainly in the aliphatic region (DBE <10, H/C 1.5–2). The MHA had a greater proportion of its molecular formulae with distinctly higher

DBE and lower H/C than the POM or LFOM fractions, which are characteristic of condensed PAHs. The various DBE plots taken together with the NMR spectra, the van Krevelen plots, and their derived calculations, illustrate a continuum toward more degraded/stabilized composition of the MHA than of the POM and LFOM. This continuum is consistent within the hypothesized increasing degree of decomposition from the LFOM/POM to the MHA, which was also shown by the carbohydrate compositions of these fractions (Abdelrahman et al. 2016), which indicated the MHA is composed of both plant and microbially derived components.

The literature suggests three possible sources of condensed aromatic C in the MHA and SOM. First, previous works (e.g., Ikeya et al., 2015) hypothesized that smaller condensed aromatic structures (2–5 rings) such as polynuclear quinones that are produced by fungi could be precursors of condensed aromatic components in SOM. A second potential source is charred-like plant materials that were subjected to degradation and oxidation processes within the soils (Abdelrahman et al., 2018). A third possible precursor is lignin; the condensed aromatic structures are produced possibly from the transformation of lignin in the presence of hydroxyl radicals (Ikeya et al., 2015; Waggoner et al., 2015). The availability of lignin in soil and the absence of lignin signals in the van Krevelen plots of MHA underscores the potential of this mechanism for creating condensed aromatic C, but our data are not suited to distinguish between the three potential sources. Previous studies provided ample evidence that condensed aromatic C structures in humic fractions such as the MHA are not created artificially during the alkaline extraction and are instead naturally occurring species (Olk et al., 2019).

3.6. Combination of analytical techniques

FT-ICR-MS analyses showed differences in the molecular compositions of the successively isolated SOM fractions. According to the molecular formulae signatures, aliphatic substances were common in all fraction but more abundant in the LFOM and POM, whereas the condensed aromatic hydrocarbons were characteristic of only the MHA

composition of the LFOM, POM, and the MHA can be characterized as follows: i) aliphatic substances were common in all three fractions and were most abundant in the LFOM and POM; ii)

438 molecular formulae assigned to the lignin range were practically negligible; and iii) molecular
439 formulae assigned to condensed aromatic hydrocarbons were characteristic of only the MHA. Hence,
440 the MHA contained both labile fragments with H/C >1.5, as well as more condensed aromatic
441 compounds with *m*/*z* >400 and DBE >17 (>4 rings), which suggest recalcitrant SOM substances.
442 In contrast to the characterization of the soluble portion by FT-ICR-MS of the LFOM as largely
443 lipid and aliphatic, solid state NMR found a largely carbohydrate nature, consistent with earlier NMR
444 analyses of the LFOM (Cao et al. 2011). This is a common discrepancy, as large carbohydrate
445 polymers have limited solubility and, moreover, have limited ionization efficiencies in negative ion
446 mode ESI, because they are less ionic than components containing carboxylic acid groups. A second
447 potential difference between the methods is that unsubstituted aromatic C is more visible in NMR
448 than in ESI-FT-ICR-MS, due to its nonpolar nature. Advanced forms of ^{13}C NMR are considered to
449 be a more quantitative measure of all C forms within SOM (Mao et al., 2017), while the qualitative
450 ESI prefers ionic polar compounds and is only semi-quantitative at best. Yet the FT-ICR-MS more
451 precisely characterized the components that it could detect, providing individual formulae instead of
452 the broad classes of functionality that NMR provides. The absence of a lignin signature in the MHA
453 fraction is consistent with the decomposition/humification continuum of soil C, as lignin may have
454 been further degraded/synthesized to the condensed aromatic hydrocarbons observed in the MHA.
455 Thus, the intrinsic chemical nature of inputted materials may indeed affect their respective cycling
456 rates, contrasting with the belief that SOM cycling is dictated solely by microbial access.
457

458 **4. Conclusions and environmental implications**

459 This work described the molecular composition of different SOM pools, and distinguish them
460 according to the molecular formulae signatures: i) aliphatic substances were common in the three
461 fractions but were most abundant in the LFOM and POM; ii) molecular formulae assigned to the
462 lignin range were practically negligible; and iii) molecular formulae assigned to condensed aromatic
463 hydrocarbons were characteristic of only the MHA. Hence, the MHA contained both labile fragments
464 with H/C >1.5, as well as more condensed aromatic compounds with *m*/*z* >400 and DBE >17 (>4

rings), which suggest recalcitrant SOM substances.

These insights demonstrate the usefulness of studying a diversity of labile SOM fractions by using both chemical and physical extractions, in combination with different analytical techniques that provide complementary insights. The Mediterranean agroecosystems, extended worldwide beyond the Mediterranean basin, are defined by their characteristic climate and soils that together with agriculture favor the rapid decrease of SOM, which has counteract strategies and interventions toward reducing emissions and negatively impact agricultural productivity. Understanding the composition and transformations of labile SOM fractions, as reported in this work, can be agronomically and ecologically relevant as it helps revealing the impact of certain environmental/agronomic conditions on soil carbon stabilization in more recalcitrant SOM forms. Further characterization of different SOM pools from different agroecosystems and different environmental settings are required in order to evaluate whether the reported fractional characteristics are case-specific or are extendable to beyond the sampled locations.

Acknowledgement: Authors thank Roland Bol, IBG-3, Forschungszentrum Jülich, for his comments on the first draft and Bernhard Steffen (retired) for his support in the development and expansion of the evaluation program for FT-ICR-MS data.

References

Abdelrahman, H., Cocozza, C., Olk, D.C., Ventrella, D., Miano, T., 2017. Carbohydrates and Amino Compounds as Short-Term Indicators of Soil Management. Clean - Soil, Air, Water 45. https://doi.org/10.1002/clen.201600076

Abdelrahman, H., Cocozza, C., Olk, D.C., Ventrella, D., Montemurro, F., Miano, T., 2020. Changes in Labile Fractions of Soil Organic Matter During the Conversion to Organic Farming. J. Soil Sci. Plant Nutr. 20, 1019–1028. https://doi.org/10.1007/s42729-020-00189-y

Abdelrahman, H., Hofmann, D., Berns, A.E.A.E., Meyer, N., Bol, R., Borchard, N., 2018. Historical charcoal additions alter water extractable, particulate and bulk soil C composition and stabilization. J. Plant Nutr. Soil Sci. 181, 809–817. https://doi.org/10.1002/jpln.201800261

Abdelrahman, H.M., Olk, D.C., Dinnes, D., Ventrella, D., Miano, T., Cocozza, C., 2016. Occurrence and abundance of carbohydrates and amino compounds in sequentially extracted labile soil organic matter fractions. J. Soils Sediments 16, 2375–2384. https://doi.org/10.1007/s11368-016-1437-y

Berns, A.E., Conte, P., 2011. Effect of ramp size and sample spinning speed on CPMAS 13C NMR spectra of soil organic matter. Org. Geochem. 42, 926–935.

Berns, A.E., Knicker, H., 2014. Soil Organic Matter. eMagRes 3, 43–54. https://doi.org/10.1002/9780470034590.EMRSTM1345

Cao, X., Olk, D.C., Chappell, M., Cambardella, C.A., Miller, L.F., Mao, J., 2011. Solid-State NMR Analysis of Soil Organic Matter Fractions from Integrated Physical–Chemical Extraction. Soil Sci. Soc. Am. J. 75, 1374–1384. https://doi.org/10.2136/SSSAJ2010.0382

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

DiDonato, N., Chen, H., Waggoner, D., Hatcher, P.G., 2016. Potential origin and formation for molecular components of humic acids in soils. Geochim. Cosmochim. Acta 178, 210–222. https://doi.org/10.1016/J.GCA.2016.01.013

Fernández, J.M., Hockaday, W.C., Plaza, C., Polo, A., Hatcher, P.G., 2008. Effects of long-term soil amendment with sewage sludges on soil humic acid thermal and molecular properties. Chemosphere 73, 1838–1844.

Ferreira, C.S.S., Seifollahi-Aghmiuni, S., Destouni, G., Ghajarnia, N., Kalantari, Z., 2022. Soil degradation in the European Mediterranean region: Processes, status and consequences. Sci. Total Environ. 805, 150106. https://doi.org/10.1016/J.SCITOTENV.2021.150106

Fox, P.M., Nico, P.S., Tfaily, M.M., Heckman, K., Davis, J.A., 2017. Characterization of natural organic matter in low-carbon sediments: Extraction and analytical approaches. Org. Geochem. 114, 12–22.

González Pérez, M., Martin-Neto, L., Saab, S.C., Novotny, E.H., Milori, D.M.B.P., Bagnato, V.S., Colnago, L.A., Melo, W.J., Knicker, H., 2004. Characterization of humic acids from a Brazilian Oxisol under different tillage systems by EPR, 13C NMR, FTIR and fluorescence spectroscopy. Geoderma 118, 181–190. https://doi.org/10.1016/S0016-7061(03)00192-7

Gregorich, E.G., Beare, M.H., McKim, U.F., Skjemstad, J.O., 2006. Chemical and Biological Characteristics of Physically Uncomplexed Organic Matter. Soil Sci. Soc. Am. J. 70, 975–985.

Hockaday, W.C., Purcell, J.M., Marshall, A.G., Baldock, J.A., Hatcher, P.G., 2009. Electrospray and photoionization mass spectrometry for the characterization of organic matter in natural waters: A qualitative assessment. Limnol. Oceanogr. Methods 7. https://doi.org/10.4319/lom.2009.7.81

Ikeya, K., Sleighter, R.L., Hatcher, P.G., Watanabe, A., 2015. Characterization of the chemical composition of soil humic acids using Fourier transform ion cyclotron resonance mass spectrometry. Geochim. Cosmochim. Acta 153, 169–182. https://doi.org/10.1016/J.GCA.2015.01.002

Ikeya, K., Sleighter, R.L., Hatcher, P.G., Watanabe, A., 2013. Fourier transform ion cyclotron resonance mass spectrometric analysis of the green fraction of soil humic acids. Rapid Commun. Mass Spectrom. 27, 2559–2568.

Ikeya, K., Sleighter, R.L., Hatcher, P.G., Watanabe, A., 2012. Compositional features of Japanese Humic Substances Society standard soil humic and fulvic acids by Fourier transform ion

cyclotron resonance mass spectrometry and X-ray diffraction profile analysis. Humic Subst. Res. 9.
Jones, A., Panagos, P., Barcelo, S., Bouraoui, F., Bosco, C., Dewitte, O., Gardi, C., Hervás, J., Hiederer, R., Jeffery, S., Montanarella, L., Penizek, V., Toth, G., Van Den Eeckhaut, M., Van Liedekerke, M., Verheijen, F.G.A., Yigini, Y., Erhard, M., Lukewille, A., Petersen, J., Marmo, L., Olazabal, C., Strassburger, T., Viestova, E., 2012. State of Soil in Europe. https://doi.org/10.2788/77361
Kaplan, D.I., Xu, C., Huang, S., Lin, Y., Tolić, N., Roscioli-Johnson, K.M., Santschi, P.H., Jaffé, P.R., 2016. Unique Organic Matter and Microbial Properties in the Rhizosphere of a Wetland Soil. Environ. Sci. Technol. 50, 4169–4177.
Kleber, M., Lehmann, J., 2019. Humic Substances Extracted by Alkali Are Invalid Proxies for the Dynamics and Functions of Organic Matter in Terrestrial and Aquatic Ecosystems. J. Environ. Qual. 48, 207–216.
Koch, B.P., Dittmar, T., 2006. From mass to structure: An aromaticity index for high-resolution mass data of natural organic matter. Rapid Commun. Mass Spectrom. 20. https://doi.org/10.1002/rcm.2386
Koch, B.P., Dittmar, T., Witt, M., Kattner, G., 2007. Fundamentals of molecular formula assignment to ultrahigh resolution mass data of natural organic matter. Anal. Chem. 79. https://doi.org/10.1021/ac061949s
Kujawinski, E.B., Behn, M.D., 2006. Automated analysis of electrospray ionization fourier transform ion cyclotron resonance mass spectra of natural organic matter. Anal. Chem. 78. https://doi.org/10.1021/ac0600306
Lattová, E., Snovida, S., Perreault, H., Krokhin, O., 2005. Influence of the labeling group on ionization and fragmentation of carbohydrates in mass spectrometry. J. Am. Soc. Mass Spectrom. 16. https://doi.org/10.1016/j.jasms.2005.01.021
Mao, J., Cao, X., Olk, D.C., Chu, W., Schmidt-Rohr, K., 2017. Advanced solid-state NMR spectroscopy of natural organic matter, Progress in Nuclear Magnetic Resonance Spectroscopy. Pergamon.
Mao, J., Olk, D.C., Fang, X., He, Z., Schmidt-Rohr, K., 2008. Influence of animal manure application on the chemical structures of soil organic matter as investigated by advanced solid-state NMR and FT-IR spectroscopy. Geoderma 146, 353–362.
Mustafa, A., Frouz, J., Naveed, M., Ping, Z., Nan, S., Minggang, X., Núñez-Delgado, A., 2022. Stability of soil organic carbon under long-term fertilization: Results from 13C NMR analysis and laboratory incubation. Environ. Res. 205, 112476. https://doi.org/10.1016/J.ENVRES.2021.112476
Mustafa, A., Saeed, Q., Karimi Nezhad, M.T., Nan, S., Hongjun, G., Ping, Z., Naveed, M., Minggang, X., Nú nez-Delgado, A., 2023. Physically separated soil organic matter pools as indicators of carbon and nitrogen change under long-term fertilization in a Chinese Mollisol. Environ. Res. 216, 114626. https://doi.org/10.1016/J.ENVRES.2022.114626
Ndzelu, B.S., Dou, S., Zhang, X., ZHANG, Y., 2022. Molecular composition and structure of organic matter in density fractions of soils amended with corn straw for five years. Pedosphere. https://doi.org/10.1016/J.PEDSPH.2022.06.057
Ohno, T., He, Z., Sleighter, R.L., Honeycutt, C.W., Hatcher, P.G., 2010. Ultrahigh Resolution Mass Spectrometry and Indicator Species Analysis to Identify Marker Components of Soil- and Plant Biomass-Derived Organic Matter Fractions. Environ. Sci. Technol. 44, 8594–8600. https://doi.org/10.1021/ES101089T
Olk, D.C., 2006. A Chemical Fractionation for Structure-Function Relations of Soil Organic Matter in Nutrient Cycling. Soil Sci. Soc. Am. J. 70. https://doi.org/10.2136/sssaj2005.0108
Olk, D.C., Bloom, P.R., Perdue, E.M., McKnight, D.M., Chen, Y., Farenhorst, A., Senesi, N., Chin, Y.-P., Schmitt-Kopplin, P., Hertkorn, N., Harir, M., 2019. Environmental and Agricultural Relevance of Humic Fractions Extracted by Alkali from Soils and Natural Waters. J. Environ. Qual. 48, 217–232. https://doi.org/10.2134/JEQ2019.02.0041
Olk, D.C., Cassman, K.G., Schmidt-Rohr, K., Anders, M.M., Mao, J.D., Deenik, J.L., 2006. Chemical stabilization of soil organic nitrogen by phenolic lignin residues in anaerobic agroecosystems. Soil Biol. Biochem. 38, 3303–3312.

Prăvălie, R., Nita, I.A., Patriche, C., Niculiță, M., Birsan, M.V., Roșca, B., Bandoc, G., 2021. Global changes in soil organic carbon and implications for land degradation neutrality and climate stability. Environ. Res. 201, 111580. https://doi.org/10.1016/J.ENVRES.2021.111580

Savarese, C., Drosos, M., Spaccini, R., Cozzolino, V., Piccolo, A., 2021. Molecular characterization of soil organic matter and its extractable humic fraction from long-term field experiments under different cropping systems. Geoderma 383, 114700. https://doi.org/10.1016/J.GEODERMA.2020.114700

Schnitzler, F., Lavorenti, A., Berns, A.E., Drewes, N., Vereecken, H., Burauel, P., 2007. The influence of maize residues on the mobility and binding of benazolin: Investigating physically extracted soil fractions. Environ. Pollut. 147, 4–13.

Seifert, A.G., Roth, V.N., Dittmar, T., Gleixner, G., Breuer, L., Houska, T., Marxsen, J., 2016. Comparing molecular composition of dissolved organic matter in soil and stream water: Influence of land use and chemical characteristics 571, 142–152.

Shen, X., Perreault, H., 1998. Characterization of carbohydrates using a combination of derivatization, high-performance liquid chromatography and mass spectrometry. J. Chromatogr. A 811, 47–59. https://doi.org/10.1016/S0021-9673(98)00238-6

Sleighter, R.L., Hatcher, P.G., 2011. Fourier Transform Mass Spectrometry for the Molecular Level Characterization of Natural Organic Matter: Instrument Capabilities, Applications, and Limitations. Fourier Transform. - Approach to Sci. Princ. https://doi.org/10.5772/15959

Sleighter, R.L., Hatcher, P.G., 2008. Molecular characterization of dissolved organic matter (DOM) along a river to ocean transect of the lower Chesapeake Bay by ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. Mar. Chem. 110, 140–152. https://doi.org/10.1016/J.MARCHEM.2008.04.008

Soil Survey Staff (2014) Keys to Soil Taxonomy, 12th ed. USDA-Natural Resources Conservation Service, Washington, DC.

Sonsri, K., Naruse, H., Watanabe, A., 2022. Mechanisms controlling the stabilization of soil organic matter in agricultural soils as amended with contrasting organic amendments: Insights based on physical fractionation coupled with 13C NMR spectroscopy. Sci. Total Environ. 825, 153853. https://doi.org/10.1016/J.SCITOTENV.2022.153853

Stubbins, A., Spencer, R.G.M.M., Chen, H., Hatcher, P.G., Mopper, K., Hernes, P.J., Mwamba, V.L., Mangangu, A.M., Wabakanghanzi, J.N., Six, J., 2010. Illuminated darkness: Molecular signatures of Congo River dissolved organic matter and its photochemical alteration as revealed by ultrahigh precision mass spectrometry. Limnol. Oceanogr. 55, 1467–1477.

Waggoner, D.C., Chen, H., Willoughby, A.S., Hatcher, P.G., 2015. Formation of black carbon-like and alicyclic aliphatic compounds by hydroxyl radical initiated degradation of lignin. Org. Geochem. 82, 69–76.

Wang, Q., Zhang, L., Zhang, J., Shen, Q., Ran, W., Huang, Q., 2012. Effects of compost on the chemical composition of SOM in density and aggregate fractions from rice-wheat cropping systems as shown by solid-state 13C-NMR spectroscopy. J. Plant Nutr. Soil Sci. 175. https://doi.org/10.1002/jpln.201100350

Yeasmin, S., Singh, B., Smernik, R.J., Johnston, C.T., 2020. Effect of land use on organic matter composition in density fractions of contrasting soils: A comparative study using 13C NMR and DRIFT spectroscopy. Sci. Total Environ. 726, 138395. https://doi.org/10.1016/J.SCITOTENV.2020.138395

Figure Captions

Fig. 1. ^{13}C CPMAS NMR spectra of a) the light fraction of soil organic matter (LFOM), b) the 500–53 µm particulate organic matter (POM) fraction, and c) the mobile humic acid (MHA) fraction from the compost treated plots at the Metaponto site (arrow = center glitch).

Fig. 2. van Krevelen diagrams of the FT-ICR-MS molecular formulae that were common to all 4 site/treatment combinations for each fraction (a) and the relationship between atomic H/C and *m/z* for the common molecular formulae (b) for the light fraction of soil organic matter (LFOM), the 500–53 µm particulate organic matter (POM), and the mobile humic acid (MHA) fractions. Formulas are colored according to heteroatom content (CHO, CHON, and CHOS formulas are black, red, and blue, respectively).

Fig. 3. Kendrick mass defect analysis plots as a function of CH_2 (a) and the relationship between double bond equivalents (DBE) and *m/z* for the molecular formulae that were common to all 4 site/treatment combinations for each fraction (b) for the light fraction of soil organic matter (LFOM), the 500–53 µm particulate organic matter (POM), and the mobile humic acid (MHA) fractions. Formulas are colored according to heteroatom content (CHO, CHON, and CHOS formulas are black, red, and blue, respectively).

Fig. 4. Relationships between double bond equivalents (DBE) and H/C (a) or O/C (b) for the molecular formulae that were common to all 4 site/treatment combinations for the light fraction of soil organic matter (LFOM), the 500–53 µm particulate organic matter (POM), and the mobile humic acid (MHA) fraction. Formulas are colored according to heteroatom content (CHO, CHON, and CHOS formulas are black, red, and blue, respectively).

Fig. 5. Structure proposals to illustrate possible skeletal structures and functional groups for randomly selected formulae with varying high DBE values in the mobile humic acid (MHA) fraction.

Table 1. The number-averaged O/C, H/C, DBE, and AI_{mod} values for the molecular formulas that were common to all four site/treatment combinations for each fraction (LFOM, POM, and MHA), as well as the percentage of formulas that fall into categories based on aromaticity (aliphatic, aromatic, condensed aroamtic), heteroatom content (CHO, CHON, CHOS), and lability (calculated from the molecular lability boundary, D'Andrilli et al., 2015).

Fraction	O/C	H/C	DBE	AI_{mod}	% Aliphatic	% Aromatic	% Condensed Aromatic	% CHO	% CHON	% CHOS	% Labile
LFOM	0.24	1.69	6.2	0.09	90.7%	0.8%	0.0%	90.2%	5.8%	4.0%	88.5%
POM	0.26	1.72	5.4	0.06	94.9%	0.0%	0.0%	92.2%	2.8%	5.0%	92.9%
MHA	0.30	1.51	8.9	0.18	76.3%	0.4%	17.1%	82.8%	4.4%	12.9%	73.6%

LFOM; light fraction of soil organic matter; POM: 500–53 µm particulate organic matter; MHA: mobile humic acid; DBE: double bond equivalent; AI_{mod}: modified aromaticity index

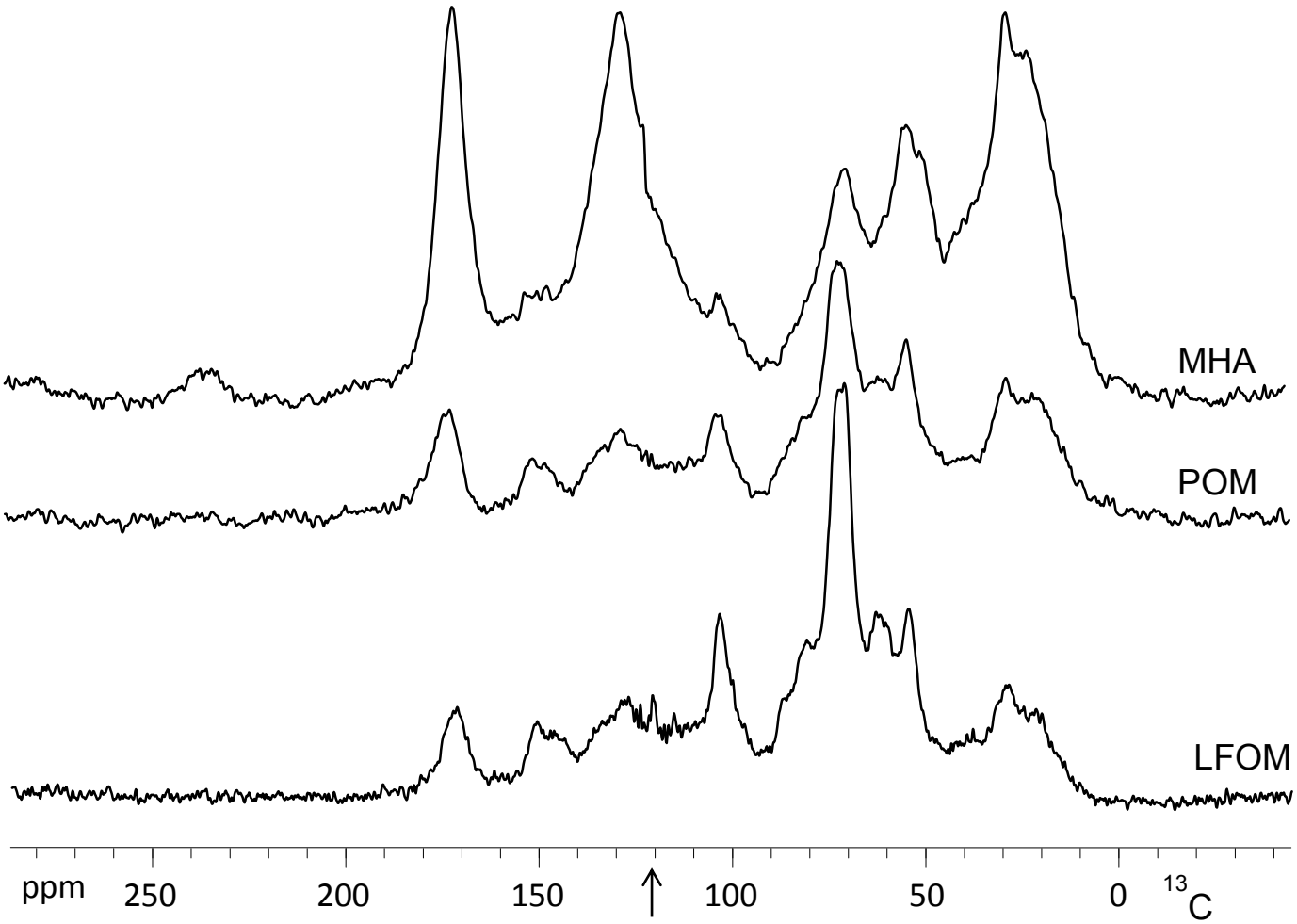


Figure 1

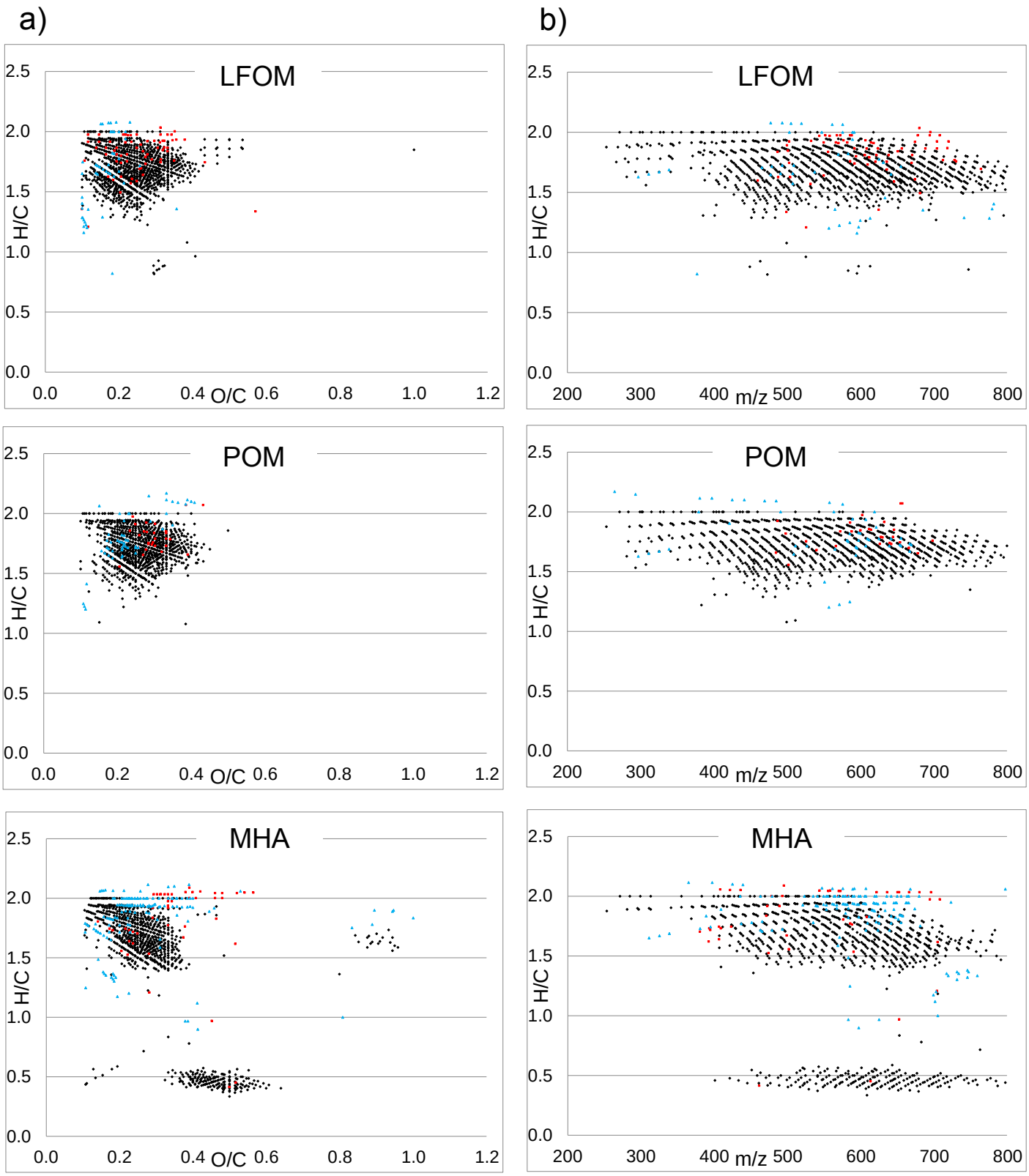

a)
b)
LFOM
POM
MHA
H/C
O/C
m/z
0.0
0.5
1.0
1.5
2.0
2.5
0.2
0.4
0.6
0.8
1.2
200
300
400
500
600
700
800
CHO
CHON
CHOS

Figure 2

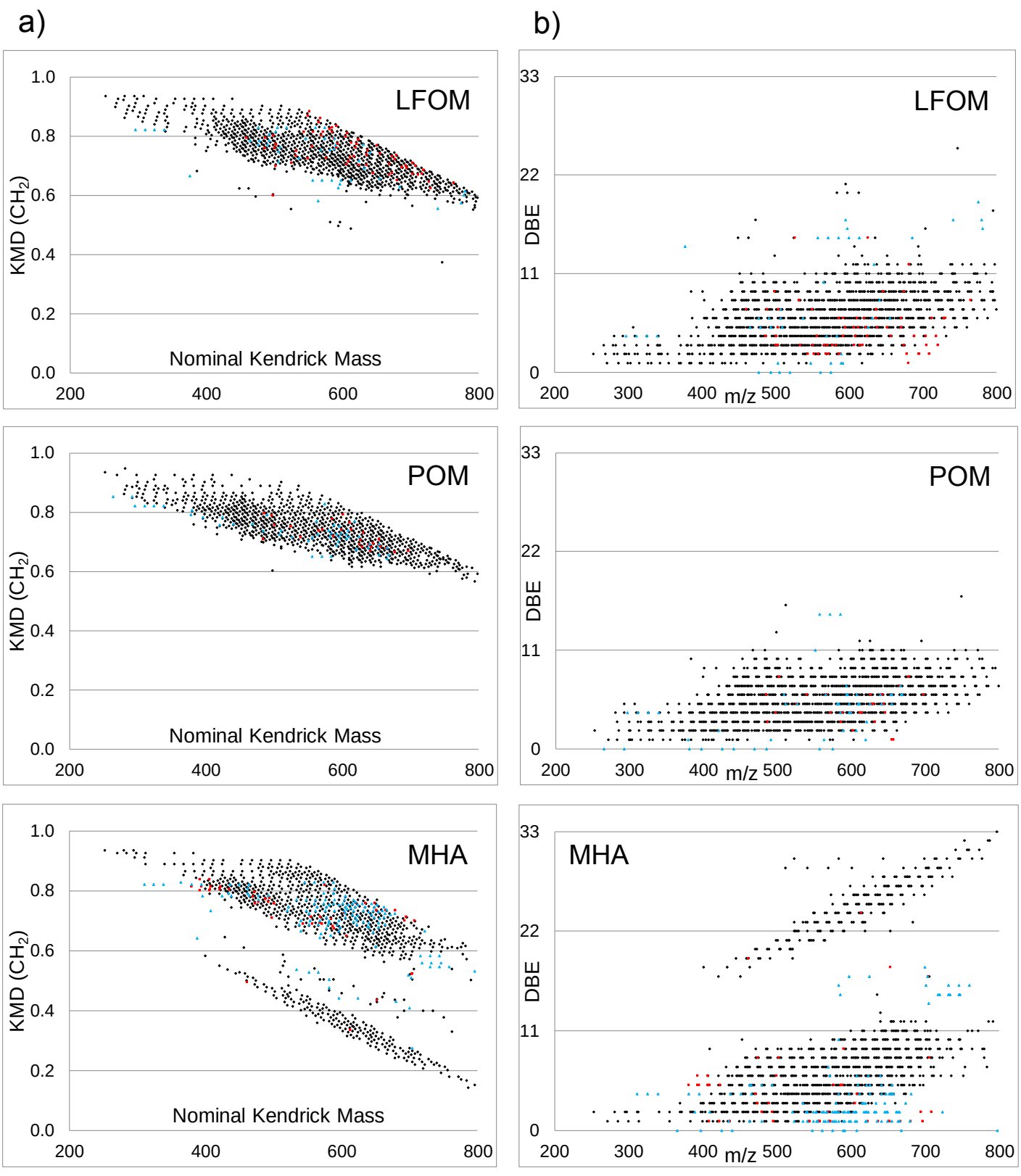

a)
b)
LFOM
POM
MHA
KMD (CH2)
Nominal Kendrick Mass
DBE
m/z
CHO CHON CHOS

Figure 3

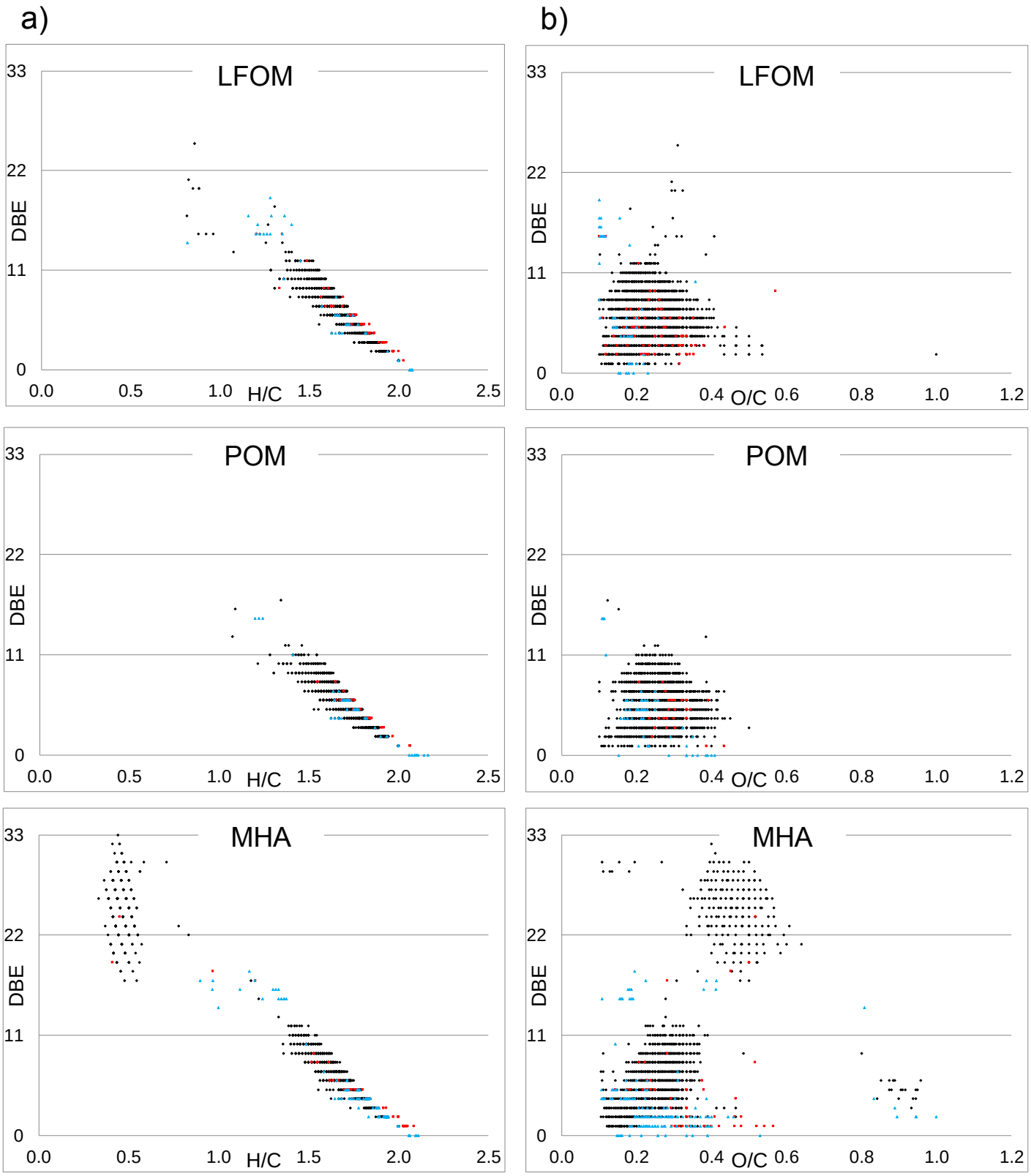
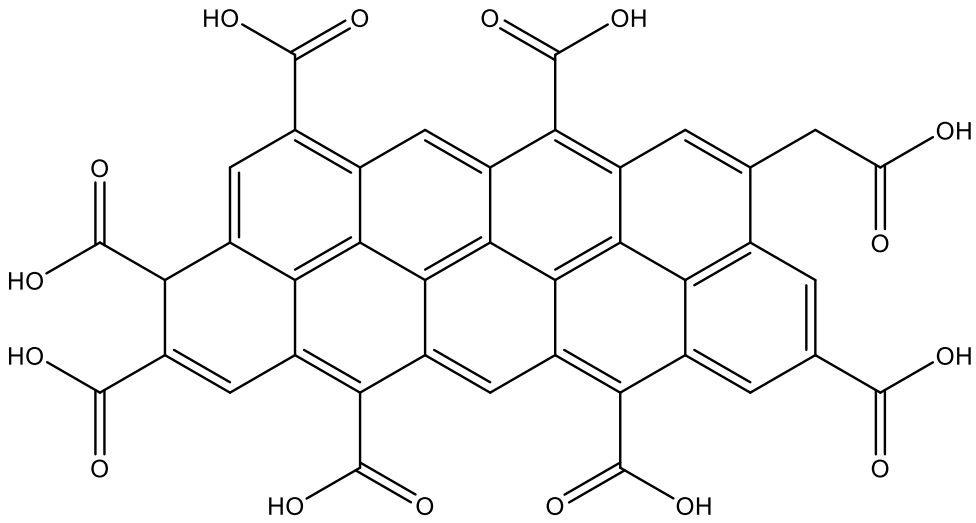
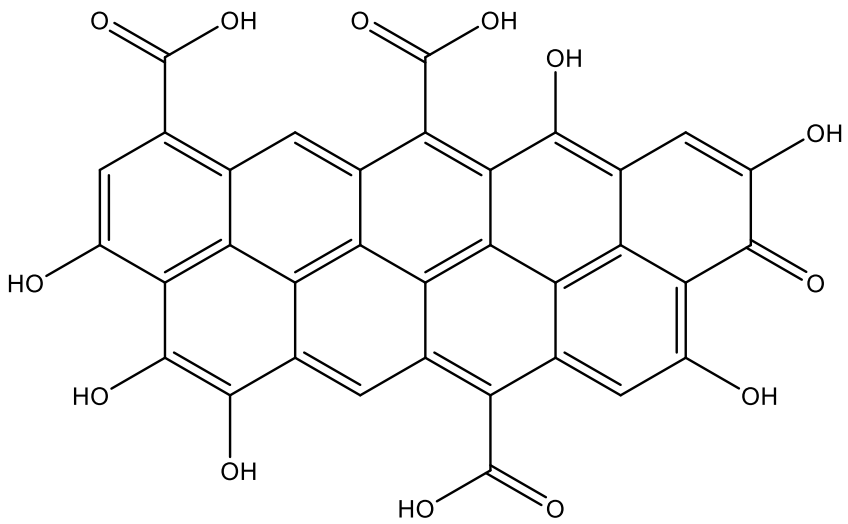

a)
b)
LFOM
POM
MHA
DBE
H/C
O/C
0
11
22
33
0.0
0.5
1.0
1.5
2.0
2.5
0.2
0.4
0.6
0.8
1.2
CHO CHON CHOS

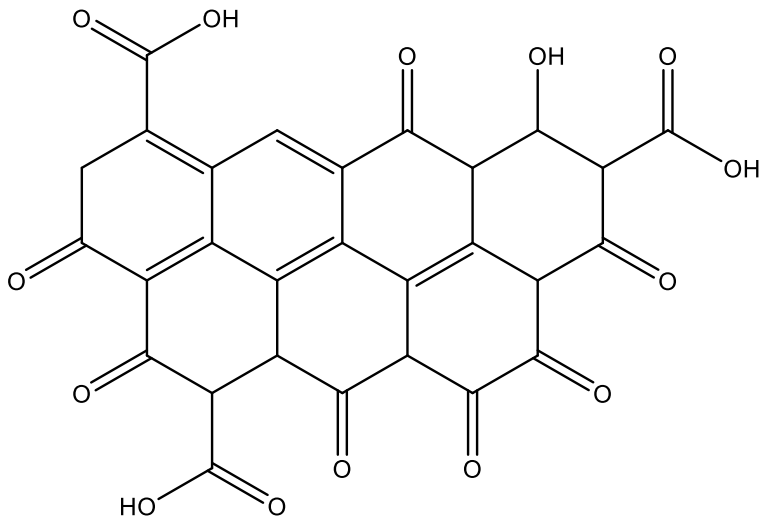
Figure 4



a)
$C_{40}H_{18}O_{16}$
64% C
DBE=32
H/C=0.40
O/C=0.45
753.052214 (1-)



b)
$C_{34}H_{14}O_{13}$
65% C
DBE=28
H/C=0.38
O/C=0.41
629.036169 (1-)



c)
$C_{28}H_{14}O_{14}$
59% C
DBE=22
H/C=0.50
O/C=0.50
573.031084 (1-)

Figure 5

Supp Mat

CRediT authorship contribution statement:

Hamada Abdelrahman: Development of the original idea, Investigation, Methodology, Writing - original draft, review and editing. **Diana Hofmann** Methodology - FT-ICR-MS analysis, writing and review **Rachel L. Sleighter**: Data analysis - FT-ICR-MS data, writing, review and editing. **Dan C. Olk**: writing original draft, review, and editing. **Anne E. Berns**: Methodology - NMR analysis, writing and review. **Teodoro Miano**: reviewing and editing **Sabry M. Shaheen:** writing, review, and editing **Claudio Cocozza**: writing, review, and editing.

Declaration of interests

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

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Declaration of interests

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

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: