Article (refereed) - postprint

Malik, Ashish A.; Roth, Vanessa-Nina; Hébert, Mathieu; Tremblay, Luc; Dittmar, Thorsten; Gleixner, Gerd. 2016. **Linking molecular size, composition and carbon turnover of extractable soil microbial compounds.** *Soil Biology and Biochemistry*, 100. 66-73. 10.1016/j.soilbio.2016.05.019

Linking molecular size, composition and carbon turnover of extractable soil microbial compounds

Ashish A. Malik[a,b,*], Vanessa-Nina Roth[a], Mathieu Hébert[c], Luc Tremblay[c], Thorsten Dittmar[d] and Gerd Gleixner[a]

[a] Department of Biogeochemical Processes, Max Planck Institute for Biogeochemistry, Jena, Germany

[b] Centre for Ecology & Hydrology, Wallingford, United Kingdom

[c] Department of Chemistry and Biochemistry, Université de Moncton, Moncton, New Brunswick, Canada

[d] Research Group for Marine Geochemistry (ICBM-MPI Bridging Group), Institute for Chemistry and Biology of the Marine Environment (ICBM), University of Oldenburg, Oldenburg, Germany

* Corresponding author:

Centre for Ecology & Hydrology, Maclean Building, Benson Lane, Crowmarsh Gifford, Wallingford, Oxfordshire, OX10 8BB, United Kingdom

Email: ashmalik@ceh.ac.uk

Phone: +44 (0) 1491 692227

Fax: +44 (0) 1491 692424

Abstract

Microbial contribution to the maintenance and turnover of soil organic matter is significant. Yet, we do not have a thorough understanding of how biochemical composition of soil microbial biomass is related to carbon turnover and persistence of different microbial components. Using a suite of state-of-the-art analytical techniques, we investigated the molecular characteristics of extractable microbial biomass and linked it to its carbon turnover time. A $^{13}CO_2$ plant pulse labelling experiment was used to trace plant carbon into rhizosphere soil microbial biomass, which was obtained by chloroform fumigation extraction (CFE). ^{13}C content in molecular size classes of extracted microbial compounds was analysed using size exclusion chromatography (SEC) coupled online to high performance liquid chromatography–isotope ratio mass spectrometry (SEC-HPLC-IRMS). Molecular characterization of microbial compounds was performed using complementary approaches, namely SEC-HPLC coupled to Fourier transform infrared spectroscopy (SEC-HPLC-FTIR) and electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI-FT-ICR-MS). SEC-HPLC-FTIR suggests that mid to high molecular weight (MW) microbial compounds were richer in aliphatic CH bonds, carbohydrate-like compounds and possibly P=O derivatives from phospholipids. On the contrary, the lower size range was characterised by more oxidised compounds with hydroxyl, carbonyl, ether and/or carboxyl groups. ESI-FT-ICR-MS suggests that microbial compounds were largely aliphatic and richer in N than the background detrital material. Both molecular characterization tools suggest that CFE derived microbial biomass was largely lipid, carbohydrate and protein derived. SEC-HPLC-IRMS analysis revealed that ^{13}C enrichment decreased with increasing MW of microbial compounds and the turnover time was deduced as 12.8 ±0.6, 18.5 ±0.6 and 22.9 ±0.6 days for low, mid and high MW size classes, respectively. We conclude that low MW compounds represent the rapidly turned-over metabolite fraction of extractable soil microbial biomass consisting of organic acids, alcohols, amino acids and sugars; whereas, larger structural compounds are part of the cell envelope (likely membrane lipids, proteins or polysaccharides) with a much lower renewal rate. This relation of microbial carbon turnover to its molecular size, structure and composition thus highlights the significance of cellular biochemistry in determining the microbial contribution to soil carbon cycling and specifically soil organic matter formation.

Keywords: Soil carbon, Microbial biomass, Chloroform fumigation extraction, HPLC-FTIR, ESI-FT-ICR-MS, HPLC-IRMS

1 Introduction

Microbial growth and activity largely control soil carbon cycling (Liang and Balser, 2011; Schimel and Schaeffer, 2012). It is readily accepted that the majority of plant organic carbon passes through the soil microorganisms, a fraction of which is used for cellular energy needs and the rest for biomass build-up; and that microbial biomass forms soil organic matter (SOM) mostly from cell fragments (Gleixner, 2013; Kögel-Knabner, 2002; Miltner et al., 2012). The contribution of microbial biomass to maintenance and accumulation of SOM is significant, some estimates suggest it could be as high as 80% of organic carbon in soil (Kindler et al., 2009; Liang et al., 2011; Simpson et al., 2007; Six et al., 2006). The residence time of microbial compounds in soil has been attributed to its molecular structure and biochemical composition as well as ecosystem specific effects (Simpson et al., 2007; Throckmorton et al., 2012; Tremblay and Benner, 2006). However, we do not have a thorough understanding of the dependencies of C turnover and persistence of microbial compounds on their molecular size and composition.

A widely used method to obtain microbial biomass from soil is biocidal fumigation using chloroform that lyses microbial cells and releases their contents which is followed by its extraction using K_2SO_4 solution (Tate et al., 1988; Vance et al., 1987). It is likely that chloroform fumigation does not lyse certain microbial groups with tougher cell envelopes and that the K_2SO_4 extraction selectively extracts only specific molecular compounds out of the lysed cellular products (Malik et al., 2013). Notwithstanding these shortcomings, chloroform fumigation extraction (CFE) has been extensively used to estimate soil microbial biomass carbon (Franzluebbers, 1999; Philippot et al., 2012) and its source and turnover time when coupled with stable isotope analysis (Dijkstra et al., 2006; Ryan and Aravena, 1994). However, in spite of its popularity in soil biology there is still no knowledge of the molecular structure and composition of the CFE microbial fraction. One of the aims of this report is to provide this understanding that is essential in making reliable ecological interpretations from CFE results. We applied two state-of-the-art molecular profiling tools to characterize the extracted microbial compounds: size exclusion chromatography (SEC) high performance liquid chromatography coupled with Fourier transform infrared spectroscopy (SEC-HPLC-

FTIR) and electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI-FT-ICR-MS). While SEC-HPLC-FTIR quantifies different functional groups in size classes of the dissolved natural organic matter (Landry and Tremblay, 2012); ESI-FT-ICR-MS is able to identify the elemental formulae of thousands of molecular ions over a wide mass range (Reemtsma, 2009; Sleighter and Hatcher, 2007). A combination of these complementary tools allows detailed molecular profiling investigations by revealing different information. Relating this information to stable carbon isotope ratios ($^{13}C/^{12}C$) of size separated compounds, when a ^{13}C tracer is applied in an experimental system, allows one to associate the molecular fingerprints to inherent carbon turnover rates. SEC-HPLC-IRMS involves size exclusion chromatography coupled online with an SEC-HPLC-IRMS interface that enables molecular size-dependent separation of organic compounds followed by direct online stable carbon isotope analysis of the eluted size fractions (Malik et al., 2012).

The objective of this study was to link the molecular size and structure of extractable microbial compounds to their carbon turnover time. CFE derived soil microbial biomass from a plant $^{13}CO_2$ pulse-labelling experiment over a time series was analysed to estimate the turnover time of its molecular size classes and to gain the molecular profiles of both bulk microbial fraction as well as its size classes. The combination of analytical tools used here allowed us to profile the CFE-derived microbial compounds and ascertain if different compound size classes have variable turnover time. The methods also provide valuable information on the molecular characteristics of compounds in soil organic matter, which has been discussed in relation to the microbial contribution to SOM formation.

2 Materials and methods

2.1 Soil sampling and experimental setup

Soil from an arable field at the Jena Biodiversity Experiment located in Jena, Germany was used in a greenhouse experiment. *Dysphania ambrosioides* (formerly *Chenopodium ambrosioides*), a temperate herb, was grown in soil mesocosms and after 3 months of plant growth; a $^{13}CO_2$ pulse labelling was performed for 10 h at a CO_2 concentration of 350-400 ppm in an airtight glass chamber. The plants were returned to the greenhouse at the end of the labelling period and destructive soil sampling was performed at 1, 3, 12 and 24 h, then at 2, 4, 7, 14, 21 and 28 days after the pulse labelling. Rhizosphere soil from 3 mesocosms was sampled at each time point. Soil was then sieved to <2mm, fine roots were extensively

117 removed (this excludes mesofauna and plant residues) and stored at -20 °C. Details about the
118 experimental design and the sampling strategy are given elsewhere (Malik et al., 2015).

119 2.2 Microbial biomass extraction

120 Microbial biomass from soil was obtained using a slightly modified version of the CFE
121 method (Vance et al., 1987; Malik et al., 2013). Soil (7 g wet weight) was fumigated in a
122 desiccator with chloroform gas for 24 h followed by repeated (8 times) evacuation. Organic
123 matter was extracted from fumigated and non-fumigated control soils with 0.05 M K_2SO_4
124 solution in a ratio of 1:4 (w/v). The mixture was homogenized on an orbital shaker (250 rev
125 per min, 30 min), centrifuged for 5 min at 12,000 g and then filtered using prewashed
126 Whatman filter paper. The resultant dissolved organic matter (DOM) was acidified and
127 purged with nitrogen gas in order to remove the dissolved inorganic carbon (Scheibe et al.,
128 2012). Fumigated and non-fumigated K_2SO_4 extracts from all time points and replicates (n
129 =30) were measured for stable carbon isotope ratios using SEC-HPLC-IRMS, whereas only
130 two composite samples from ten randomly pooled K_2SO_4 extracts were used for the other
131 detailed molecular analyses because as expected the microbial biomass content and
132 composition did not change in our steady state experimental system (Malik et al., 2015). The
133 K_2SO_4 extracts were directly measured on SEC-HPLC-IRMS without any further treatment
134 but for the other analyses a solid phase extraction (SPE) was performed in order to
135 concentrate and desalt the DOM. Fumigated and non-fumigated DOM was acidified to pH 2
136 and applied to activated SPE cartridges (Bond Elut PPL cartridge; 1g, Agilent Technologies;
137 Dittmar et al., 2008). CFE protocol blank (K_2SO_4 extract without any soil sample) and SPE
138 cartridge blank were also maintained throughout the molecular fingerprinting analyses. CFE
139 microbial biomass and background non-fumigated DOM represented approximately 1.2 and
140 0.3 % of total organic carbon, respectively, in the studied soil system.

141 2.3 SEC-HPLC-FTIR analysis

142 The SPE extracts were dissolved in methanol in order to obtain a DOC concentration of ~ 3
143 mg mL-1. 40 µL of this solution was injected into an Agilent 1200 HPLC system equipped
144 with the Polymer standards service (PSS) SUPREMA analytical Linear S (8 x 300 mm; 5 µm)
145 SEC column. Otherwise, the SEC-HPLC-FTIR technique and calibration used were the same
146 as previously described (Landry and Tremblay, 2012). Before deposition and FTIR analysis,
147 UV detection was carried out at a wavelength of 254 nm. SEC separated DOM was deposited
148 as tracks onto a rotating germanium disk where a background spectrum was taken on a

149 portion of the track where no DOM was eluting. Each band of every FTIR spectrum was
150 integrated, valley-to-valley. The absolute absorbance area of each band present in the protocol
151 blank was subtracted from the absolute area of the same band in the samples. Concentration
152 factors were considered for these subtractions. The corrected absorbance area of the band was
153 normalized by the corrected total absorbance area of the spectrum to obtain the relative
154 absorbance (%) for the band. These relative values enabled us to follow variations in the
155 proportion of each peak present in different spectra (i.e., in different samples and for different
156 MW).

157 2.4 ESI-FT-ICR-MS analysis

158 Mass spectral analyses were performed on a Bruker Solarix 15 Tesla FT-ICR-MS (Bruker
159 Daltonics, USA) in negative ionization mode. Prior to analysis the SPE DOM extracts were
160 diluted to achieve a dissolved organic carbon concentration of 20 mg/L in a 1:1 mixture of
161 methanol and water (v/v). The samples and blanks were injected into the ESI source with a
162 flow rate of 120 µL/h and an ESI needle voltage of -4 kV. The recorded m/z range was
163 between m/z 150 and 2000. 500 transients with an ion accumulation time of 0.2 s were added
164 up to one spectrum. A list of 67 compounds that covered the relevant m/z range was used for
165 linear internal mass calibration with a maximum mass error of 0.1 ppm. Molecular formula
166 assignment considered C, H, O, N, S and P using a self-written Matlab routine (Koch et al.,
167 2007; Stenson et al., 2003).

168 To remove detected masses or peaks that were not measured significantly, several criteria
169 were applied. First, m/z with a signal-to-noise ratio of the maximum of each m/z ($s/n_{Max,i}$) ≤ 5
170 were discarded (Pohlabeln and Dittmar, 2015 and references therein). The noise was defined
171 as the minimum intensity across all detected masses without blanks. $s/n_{Max,i}$ was defined for
172 each m/z_i individually by dividing the maximum intensity of each m/z_i by the noise. Second,
173 only detected masses that occurred more than one time in the set of measured samples were
174 kept. Third, all detected masses with a s/n_{Blanks} ratio ≥ 20 were removed. To determine
175 s/n_{Blanks} the average of signal intensity across all measured blanks was divided by noise. After
176 removing detected masses according to these criteria, the remaining detected masses were
177 normalized to the sum of intensities. For the following data analyses, only detected masses
178 with assigned molecular formulae were considered. For the list of non-fumigated samples
179 only formulae that were detected in both replicates were considered, the same was applied to
180 the list of peaks in the fumigated samples. Identified molecular formulae were assigned to

compound groups based on established molar ratios, aromaticity index and heteroatom contents (Seidel et al., 2015a, 2015b; Text S1). Formulae present in both fumigated and non-fumigated DOM extracts were considered as signatures of bulk SOM, whereas those unique to fumigated extracts were identified as microbial. In addition, a differential spectrum approach was used, wherein the relative intensities of non-fumigated samples were subtracted from those of the fumigated samples resulting in m/z with positive and negative values. Formulae that were present in higher abundance above a threshold (10% of the median intensity of the top 1 % of formulae ranked according to intensity) in either fumigated (positive values) or non-fumigated (negative values) were considered and identified as microbial or bulk SOM-related, respectively. More details about the data analysis are given in Supporting Information, Text S1.

2.5 SEC-HPLC-IRMS analysis

Stable carbon isotope measurements were carried out using an HPLC system coupled to a Delta+ XP IRMS through an LC IsoLink interface (Thermo Fisher Scientific, Germany). SEC was performed on a mixed bed analytical column (TSK-GEL GMPWXL- 7.8 mm × 30 cm; Tosoh Bioscience, Germany). 100 µL aliquot of soil extracts was injected using an autosampler (Surveyor autosampler, Thermo Fisher Scientific) into the mobile phase that consisted of phosphate buffer 20 mM (pH 6.2) maintained at a constant flow rate of 500 µL min^{-1} using a Surveyor MS pump. Apparent MW was obtained using a calibration curve plotted with standards having known MW (Malik et al., 2012, 2013 for technical details). Empirical C turnover time (synonymously referred to as mean residence time) of microbial size classes was obtained by estimating the pulse ^{13}C dilution rate using an exponential function in SigmaPlot (Malik et al., 2015).

3 Results

3.1 SEC-HPLC-FTIR

The SEC chromatograms obtained from UV detection shows the distribution of fumigated and non-fumigated DOM with molecular weight ranging from 300 to 6500 Da (Supporting Information, Figure S1). The majority of fumigated and non-fumigated DOM was between 500 Da and 3500 Da with peak maxima around 1400 Da. Another peak was observed only in fumigated extracts between 300 Da and 480 Da with peak maxima around 400 Da. The same

two peaks were seen in SEC chromatograms obtained with FTIR detection (total spectrum absorbance) but their relative intensities differed when compared with UV absorbance peaks (not shown). FTIR absorption bands at differing wavenumbers were assigned to functional groups on the basis of published literature on natural organic matter and related complex molecules (Bellamy, 1975; Landry and Tremblay, 2012 and references therein). Unique presence of functional groups in fumigated extracts or their increased abundance in fumigated relative to non-fumigated DOM was the criteria used to identify microbial molecular structures (Figure 1). In general, fumigated DOM was richer in aliphatic CH bonds (3050-2830 cm^{-1}), carbohydrate-like compounds and possibly P=O derivatives (1080 cm^{-1}) (Figure 1-B, G; Table 1; Davis and Mauer, 2010). The distribution with MW suggests that these functional groups or structures were more abundant in mid to high MW CFE-derived microbial compounds. OH (3700-2700 cm^{-1}), C-O/C-O-C or C-OH (1400-1200 cm^{-1}) and C-N or C-O bonds (1250 cm^{-1}) were more abundant in low MW fumigated DOM (Figure 1-A, E, F; Table 1). The highest proportion of C-N or C-O groups in fumigated DOM was observed below 565 Da, which includes masses assigned to the low end of mid-MW and all the low MW.

3.2 ESI-FT-ICR-MS

Molecular formula assignment revealed that 59 formulae were unique to fumigated DOM and 112 formulae were more abundant in fumigated relative to non-fumigated extracts (Figure S2A-B). These 171 formulae were considered as signatures of microbial compounds. 659 formulae were present in both fumigated and non-fumigated extracts and therefore linked to the background SOM. Another 47 of them that were present in higher abundance in non-fumigated relative to fumigated DOM were also considered as SOM-derived (Figure S2C). Among the SOM related molecular formulae 80.6 % (569) were characterized with no nitrogen (N), 17 % (120) with 1 N and the rest 2.4 % (17) with 2-4 N. On the contrary, molecular formulae identified as microbial were rich in N; 53.2 % contained at least 1 N (Table 2). 17 % (29) of microbial molecular formulae were linked to peptides, compared to a tiny 2.5 % (18) of SOM-related formulae. Aromaticity index (AI) estimated for each formula identifies the molecules as non-aromatic ($AI \leq 0.5$), aromatic ($0.5 < AI < 0.67$) or condensed aromatic ($AI \geq 0.67$; (Koch and Dittmar, 2006). The percentage of non-aromatic formulae that were identified as microbial were higher (87.1 %) than those identified as SOM related (75.2 %; Table 2). Unique insights were obtained when the extracted formulae were plotted in the

H/C (hydrogen/carbon) versus O/C (oxygen/carbon) space called van Krevelen diagrams (Figure 2; (Kim et al., 2003; Sleighter and Hatcher, 2007). Comparing the van Krevelen analysis of microbial and SOM related molecular formulae with that of standard compound classes we could roughly infer that most microbial compounds extracted using CFE were lipid, amino sugar or protein derived. Assigned compound groups varied in relative abundance in microbial and SOM-related fraction; polyphenols and highly unsaturated compounds dominated the SOM-related fraction whereas unsaturated aliphatic and peptide like compounds were present in a higher fraction in the microbial extract (Figure 2E-F, Table S1). The mass:charge ratio (m/z) provides an indication of the MW of the microbial compounds because they were singly charged, which allowed us to investigate the molecular size to structure relationship. Most of the microbial formulae with higher m/z ratio were devoid of N and fell in the van Krevelen space for lipids, whereas those with relatively lower m/z ratio contained 1-2 N and fell in the protein and/or amino sugar van Krevelen space (Figure 2A-D). SOM-related formulae with higher molecular weight fell in the lignin and tannin van Krevelen space where those on the lower range were widely distributed across the van Krevelen diagram.

3.3 SEC-HPLC-IRMS

When used in combination with stable carbon isotope analysis the CFE fraction can be used to track the source of microbial carbon. Fumigated and non-fumigated DOM separated using SEC was assigned into three size classes (Figure S3) consistent with those used for SEC-HPLC-FTIR results: less than 408 Da (LMW/low MW), 408-2072 Da (MMW/mid MW) and 2072-10510 Da (HMW/high MW). ^{13}C enrichment in microbial compound size classes was estimated from the $\delta^{13}C$ values of fumigated and non-fumigated DOM size classes using a mass balance. ^{13}C enrichment in all size classes was highest immediately after the pulse labeling of plants and remained so for at least 12 h after the pulse (Figure 3). Among the size classes, highest enrichment was measured in the LMW fraction and it decreased with increasing MW. The mean $\Delta\delta^{13}C$ of LMW, MMW and HMW microbial compounds 1 h after pulse labeling was 151 ± 48.9, 107.2 ± 16.5 and 81 ± 29.9 ‰; it decreased to 28.7 ± 8.1, 29.3 ± 4.8 and 17.1 ± 6.7 ‰ by the final sampling point 4 weeks after pulse labelling. The ^{13}C enrichment was fitted into an exponential decay function and the fitted degradation rate constant (k) was used to calculate the turnover time of microbial compound size classes

(Figure 3). Empirical C turnover time was estimated at 12.8 ± 0.6, 18.5 ± 0.6 and 22.9 ± 0.7 days for LMW, MMW and HMW size classes, respectively (Table 3).

4 Discussion

Chloroform fumigation extraction (CFE) is a widely used technique to quantify soil microbial biomass and although the extraction is incomplete and selective it is the easiest and fastest method of characterising soil microbial biomass. CFE gives lysed cellular components (living microbial biomass) in addition to K_2SO_4-extractable SOM that includes microbial residues and necromass. Therefore, the non-fumigated extract has to be subtracted from the fumigated one in order to quantify and profile microbial organic matter. The high-resolution analytical tools reported here allowed us to characterise and decipher microbial molecular fingerprints from that of the background SOM. We observed some clear relationships between molecular size and structure of microbial compounds. However, SEC and mass spectrometry measure molecular size using different mechanisms and therefore the size, structure and carbon turnover relationships obtained from these techniques may not be entirely comparable. In addition, the starting material used in the molecular characterization techniques was DOM concentrated through SPE with approximately 60 % recovery (Roth et al., 2015), while that used for isotope analyses was unprocessed fumigated and non-fumigated DOM. This means we are missing a fraction of DOM in the molecular characterization analyses.

SEC-HPLC-FTIR analysis revealed that low to mid MW compounds were more oxidised and rich in hydroxyl (OH and C-OH), carbonyl (C=O), amine (C-N), ether or carbohydrates (C-O/C-O-C) and/or carboxyl (COOH) groups suggesting that this is the 'metabolite fraction' of the microbial biomass (Baldock et al., 1990). ESI-FT-ICR-MS suggests that lower MW microbial compounds were enriched in N and fell in the peptide and highly unsaturated compounds' space of van Krevelen analysis and ^{13}C analysis revealed that lower MW compounds have faster turnover. Thus we could conclude that the low MW microbial metabolite fraction consists of fast turnover components often monomers like amino acids, sugars, organic acids and alcohols (Ma et al., 2012). Towards the higher MW range, SEC-HPLC-FTIR observations suggest higher abundance of aliphatic compounds, carbohydrates and phosphate groups indicating the presence of larger cellular biosynthetic components (Simpson et al., 2007). ESI-FT-ICR-MS indicates that mid MW microbial compounds were unsaturated aliphatic compounds, devoid of N and possibly lipid-derived. It is clear from

307 synthesis of data from all three techniques that the majority of microbial compounds were in
308 the mid to high MW range and were polymeric or oligomeric biosynthetic cellular
309 components of lipid, polysaccharide or protein origin (Hart et al., 2013; Spence et al., 2011).
310 SEC-HPLC-IRMS analysis based turnover time estimates suggest that these cellular structural
311 compounds have slower turnover in comparison to the metabolite fraction. These compounds
312 can be degraded into smaller and more oxidised fragments, in agreement with the observation
313 that O-rich functional groups were more abundant in low MW microbial extracts and in bulk
314 SOM. Microorganisms appear to be important contributors of bulk low to mid MW SOM
315 extractable fraction, while bulk high MW SOM seems mostly from plant-derived
316 macromolecules such as lignin and tannin. Through this greenhouse based mesocosm
317 experiment we clearly demonstrate the molecular size and structure relationship to turnover
318 time of microbial compounds. A similar correlation of size and turnover time of microbial
319 compounds has been previously demonstrated in a C3-C4 vegetation change field experiment
320 on two different soil types and this report validates the earlier results (Malik et al., 2013).
321 However, the turnover time of high MW compounds was 22 days that appears to be rather
322 small in the SOM context (Schmidt et al., 2011). This could mean that CFE microbial
323 biomass is largely labile and/or many high turnover time compounds are not extracted (Malik
324 et al., 2015).

325 The complementary molecular characterization tools reported here are neither exhaustive nor
326 specific, but they allowed us for the first time to profile the molecular fingerprints of living
327 microbial biomass as well as the SOM that is largely derived from microbial necromass. CFE
328 is a popular technique in soil biology employed to measure microbial biomass, its source,
329 turnover and fate in soil systems. It is often referred to as “quick and dirty” because there is
330 not much information about the molecular composition of the fumigated and non-fumigated
331 extracts. Assuming a strong dominance of microbial-derived material in the fumigated
332 extracts and by constructing differential spectra, the molecular techniques used in this report
333 suggest that CFE derived microbial biomass is largely lipid, carbohydrate and protein derived.
334 This study sheds light on the hitherto unknown molecular characteristics of the CFE
335 extractable soil microbial biomass.

336 Unique molecular signatures of the extracted SOM were also obtained by focussing on the
337 non-fumigated DOM fraction. SEC-HPLC-FTIR observations suggest that hydroxyl, carboxyl
338 and ester groups were abundant across the MW range in SOM derived extracts. Interestingly,

aliphatic compounds, carbohydrates and phosphate groups were less abundant compared to the fumigated fraction. Complementary ESI-FT-ICR-MS analysis supports the lower abundance of aliphatic compounds and furthermore suggests the lack of nitrogen in most molecules detected. It also identified that SOM-related high MW compounds were likely lignin and tannin compounds, indicators of vascular plants. Those on the lower MW range were also lignin and tannin-related but in addition were also identified as lipid, carbohydrate and protein derived, which are all molecular signatures of microbial biomass (Hart et al., 2013; Kelleher et al., 2006; Malik and Gleixner, 2013; Mann et al., 2015; Spence et al., 2011). This overlap suggests contribution of microbial necromass or microbially processed or altered compounds to SOM formation. It is further substantiated by reports about marine sequestration of carbon in bacterial metabolites where a similar overlap between bacterial and surrounding DOM molecular formulae is observed (Lechtenfeld et al., 2015).

The turnover time relationship with molecular size and structure of microbial compounds highlights the importance of cellular biochemistry in determining the microbial contribution to soil C formation. This dependency could be altered in conditions of environmental change as a function of physiological adaptation of soil microorganisms. The latter could lead to change in the balance of microbial anabolic and catabolic processes that is often linked to the capacity of soil microorganisms to regulate the flow of C in soil systems. A more efficient metabolic pathway with higher anabolic fluxes could lead to increased microbial biomass and subsequently higher SOC storage. In this context, the molecular size-turnover time relationship could be useful in understanding the microbial physiological adaptations to climate change and help improving the mechanistic understanding of soil C cycling processes.

Acknowledgements

This truly collaborative project was funded by the Max-Planck-Gesellschaft (MPG)/Max Planck Society, the Natural Science and Engineering Research Council of Canada (NSERC), and the New Brunswick Innovation Foundation (NBIF). We acknowledge Deutsche Forschungsgemeinschaft (DFG)/German Research Foundation for the PhD fellowship to A.A.M. in the research training group 1257 ‘Alteration and element mobility at microbe-mineral interface’. V.-N. R. received financial support from the foundation “Zwillenberg-Tietz Stiftung” and Deutsche Forschungsgemeinschaft (DFG)/German Research Foundation as part of the collaborative research centre (CRC) 1076 “AquaDiva”. A.A.M. has also

371 received funding from the European Union's Horizon 2020 research and innovation
372 programme under the Marie Sklodowska-Curie grant agreement No 655240. We thank
373 Steffen Ruehlow for technical support with stable isotope analyses; Agnes Fastnacht, Karl
374 Kuebler and Iris Kuhlmann for support in establishing the experimental setup and K.
375 Klapproth for technical support with FT-ICR-MS measurements. We also thank the reviewers
376 and the editor for constructive comments that helped improve the manuscript.

377

378

Figure legends:

Figure 1: FTIR relative absorbance of seven peaks (panels A to G) of size separated DOM in fumigated and non-fumigated samples. A greater proportion of a particular functional group in the microbial fraction is indicated by its higher relative absorbance in the fumigated fraction in comparison to the non-fumigated.

Figure 2: van Krevelen plots of ESI-FT-ICR-MS derived formulae over molecular weight (A,B), number of nitrogen (C,D) and compound classes (E,F) that are microbial- (unique to or in higher abundance in fumigated extracts) or soil organic matter (SOM) related (unique to or in higher abundance in non-fumigated extracts). van Krevelen diagrams plot the H/C (hydrogen/carbon) of each assigned molecular formula against its O/C (oxygen/carbon).

Figure 3: ^{13}C isotope incorporation in different microbial size classes (n=3) over the experimental period following $^{13}CO_2$ pulse labeling of plants (LMW: low molecular weight, MMW: mid molecular weight and HMW: high molecular weight).

References

Baldock, J.A., Oades, J.M., Vassallo, A.M., Wilson, M.A., 1990. Solid-state CP/MAS 13C N.M.R. analysis of bacterial and fungal cultures isolated from a soil incubated with glucose. Australian Journal of Soil Research 28, 213–25.

Davis, R., Mauer, L., 2010. Fourier transform infrared (FT-IR) spectroscopy: a rapid tool for detection and analysis of foodborne pathogenic bacteria, in: Méndez-Vilas, A. (Ed.), Current Research, Technology and Education Topics in Applied Microbiology and Microbial Biotechnology. FORMATEX, Spain, pp. 1582–1594.

Dijkstra, P., Ishizu, A., Doucett, R., Hart, S.C., Schwartz, E., Menyailo, O. V., Hungate, B.A., 2006. 13C and 15N natural abundance of the soil microbial biomass. Soil Biology and Biochemistry 38, 3257–3266. doi:10.1016/j.soilbio.2006.04.005

Dittmar, T., Koch, B., Hertkorn, N., Kattner, G., 2008. A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. Limnol Oceanogr. Methods 6, 230–235. doi:10.4319/lom.2008.6.230

Franzluebbers, A.J., 1999. Introduction to Symposium — Microbial Biomass : Measurement and Role in Soil Quality 30677, 505–506.

Gleixner, G., 2013. Soil organic matter dynamics: a biological perspective derived from the use of compound-specific isotopes studies. Ecological Research 28, 683–695. doi:10.1007/s11284-012-1022-9

Hart, K.M., Kulakova, A.N., Allen, C.C.R., Simpson, A.J., Oppenheimer, S.F., Masoom, H., Courtier-Murias, D., Soong, R., Kulakov, L. a, Flanagan, P. V, Murphy, B.T., Kelleher, B.P., 2013. Tracking the fate of microbially sequestered carbon dioxide in soil organic matter. Environmental Science & Technology 47, 5128–37. doi:10.1021/es3050696

Kelleher, B.P., Simpson, M.J., Simpson, A.J., 2006. Assessing the fate and transformation of plant residues in the terrestrial environment using HR-MAS NMR spectroscopy. Geochimica et Cosmochimica Acta 70, 4080–4094. doi:10.1016/j.gca.2006.06.012

Kim, S., Kramer, R.W., Hatcher, P.G., 2003. Graphical method for analysis of ultrahigh-resolution broadband mass spectra of natural organic matter, the van Krevelen diagram. Analytical Chemistry 75, 5336–44. doi:10.1021/ac034415p

Kindler, R., Miltner, A., Thullner, M., Richnow, H.-H., Kästner, M., 2009. Fate of bacterial biomass derived fatty acids in soil and their contribution to soil organic matter. Organic Geochemistry 40, 29–37. doi:10.1016/j.orggeochem.2008.09.005

Koch, B.P., Dittmar, T., 2006. From mass to structure: An aromaticity index for high-resolution mass data of natural organic matter. Rapid Communications in Mass Spectrometry 20, 926–932. doi: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. Analytical Chemistry 79, 1758–1763. doi:10.1021/ac061949s

Kögel-Knabner, I., 2002. The macromolecular organic composition of Plant and microbial residues as inputs to soil organic matter. Soil Biology and Biochemistry 34, 139–162. doi:10.1016/S0038-0717(01)00158-4

Landry, C., Tremblay, L., 2012. Compositional differences between size classes of dissolved

organic matter from freshwater and seawater revealed by an HPLC-FTIR system. Environmental Science and Technology 46, 1700–1707. doi:10.1021/es203711v

Lechtenfeld, O.J., Hertkorn, N., Shen, Y., Witt, M., Benner, R., 2015. Marine sequestration of carbon in bacterial metabolites. Nature Communications 6, 6711. doi:10.1038/ncomms7711

Liang, C., Balser, T.C., 2011. Microbial production of recalcitrant organic matter in global soils: implications for productivity and climate policy. Nature Reviews Microbiology 9, 75–75. doi:10.1038/nrmicro2386-c1

Liang, C., Cheng, G., Wixon, D.L., Balser, T.C., 2011. An Absorbing Markov Chain approach to understanding the microbial role in soil carbon stabilization. Biogeochemistry 106, 303–309. doi:10.1007/s10533-010-9525-3

Ma, J., Ibekwe, a. M., Wang, H., Xu, J., Leddy, M., Yang, C.-H., Crowley, D.E., 2012. Assimilable Organic Carbon (AOC) in Soil Water Extracts Using Vibrio harveyi BB721 and Its Implication for Microbial Biomass. PLoS ONE 7, e28519. doi:10.1371/journal.pone.0028519

Malik, A., Blagodatskaya, E., Gleixner, G., 2013. Soil microbial carbon turnover decreases with increasing molecular size. Soil Biology and Biochemistry 62, 115–118. doi:10.1016/j.soilbio.2013.02.022

Malik, A., Gleixner, G., 2013. Importance of microbial soil organic matter processing in dissolved organic carbon production. FEMS Microbiology Ecology 86, 139–148. doi:10.1111/1574-6941.12182

Malik, A., Scheibe, A., LokaBharathi, P. a, Gleixner, G., 2012. Online stable isotope analysis of dissolved organic carbon size classes using size exclusion chromatography coupled to an isotope ratio mass spectrometer. Environmental Science & Technology 46, 10123–9. doi:10.1021/es302467y

Malik, A.A., Dannert, H., Griffiths, R.I., Thomson, B.C., Gleixner, G., 2015. Rhizosphere bacterial carbon turnover is higher in nucleic acids than membrane lipids: implications for understanding soil carbon cycling. Frontiers in Microbiology 6, 268. doi:10.3389/fmicb.2015.00268

Mann, B.F., Chen, H., Herndon, E.M., Chu, R.K., Tolic, N., Portier, E.F., Roy Chowdhury, T., Robinson, E.W., Callister, S.J., Wullschleger, S.D., Graham, D.E., Liang, L., Gu, B., 2015. Indexing Permafrost Soil Organic Matter Degradation Using High-Resolution Mass Spectrometry. PloS One 10, e0130557. doi:10.1371/journal.pone.0130557

Miltner, A., Bombach, P., Schmidt-Brücken, B., Kästner, M., 2012. SOM genesis: microbial biomass as a significant source. Biogeochemistry 111, 41–55. doi:10.1007/s10533-011-9658-z

Philippot, L., Ritz, K., Pandard, P., Hallin, S., Martin-Laurent, F., 2012. Standardisation of methods in soil microbiology: progress and challenges. FEMS Microbiology Ecology 82, 1–10. doi:10.1111/j.1574-6941.2012.01436.x

Pohlabeln, A.M., Dittmar, T., 2015. Novel insights into the molecular structure of non-volatile marine dissolved organic sulfur. Marine Chemistry 168, 86–94. doi:http://dx.doi.org/10.1016/j.marchem.2014.10.018

Reemtsma, T., 2009. Determination of molecular formulas of natural organic matter

478 molecules by (ultra-) high-resolution mass spectrometry. Status and needs. Journal of
479 Chromatography A 1216, 3687–3701. doi:10.1016/j.chroma.2009.02.033

480 Roth, V.-N., Dittmar, T., Gaupp, R., Gleixner, G., 2015. The Molecular Composition of
481 Dissolved Organic Matter in Forest Soils as a Function of pH and Temperature. PLoS
482 ONE 10, e0119188.

483 Ryan, M.C., Aravena, R., 1994. Combining C-13 Natural-Abundance and Fumigation
484 Extraction Methods to Investigate Soil Microbial Biomass Turnover. Soil Biology &
485 Biochemistry 26, 1583–1585. doi:10.1016/0038-0717(94)90101-5

486 Scheibe, A., Krantz, L., Gleixner, G., 2012. Simultaneous determination of the quantity and
487 isotopic signature of dissolved organic matter from soil water using high-performance
488 liquid chromatography/isotope ratio mass spectrometry. Rapid Communications in Mass
489 Spectrometry 26, 173–180. doi:10.1002/rcm.5311

490 Schimel, J.P., Schaeffer, S.M., 2012. Microbial control over carbon cycling in soil. Frontiers
491 in Microbiology 3, 1–11. doi:10.3389/fmicb.2012.00348

492 Schmidt, M.W.I., Torn, M.S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I. a.,
493 Kleber, M., Kögel-Knabner, I., Lehmann, J., Manning, D. a. C., Nannipieri, P., Rasse,
494 D.P., Weiner, S., Trumbore, S.E., 2011. Persistence of soil organic matter as an
495 ecosystem property. Nature 478, 49–56. doi:10.1038/nature10386

496 Seidel, M., Beck, M., Greskowiak, J., Riedel, T., Waska, H., Suryaputra, I.G.N.A., Schnetger,
497 B., Niggemann, J., Simon, M., Dittmar, T., 2015a. Biogeochemistry of dissolved organic
498 matter in an intertidal sandy beach. Marine Chemistry 176, 150–163.
499 doi:10.1016/j.marchem.2015.08.011

500 Seidel, M., Yager, P.L., Ward, N.D., Carpenter, E.J., Gomes, H.R., Krusche, A. V., Richey,
501 J.E., Dittmar, T., Medeiros, P.M., 2015b. Molecular-level changes of dissolved organic
502 matter along the Amazon River-to-ocean continuum. Marine Chemistry 177, 218–231.
503 doi:10.1016/j.marchem.2015.06.019

504 Simpson, A.J., Simpson, M.J., Smith, E., Kelleher, B.P., 2007. Microbially derived inputs to
505 soil organic matter: Are current estimates too low? Environmental Science and
506 Technology 41, 8070–8076. doi:10.1021/es071217x

507 Six, J., Frey, S.D., Thiet, R.K., Batten, K.M., 2006. Bacterial and Fungal Contributions to
508 Carbon Sequestration in Agroecosystems. Soil Science Society of America Journal 70,
509 555–569. doi:10.2136/sssaj2004.0347

510 Sleighter, R.L., Hatcher, P.G., 2007. The application of electrospray ionization coupled to
511 ultrahigh resolution mass spectrometry for the molecular characterization of natural
512 organic matter. Journal of Mass Spectrometry 42, 559–574. doi:10.1002/jms.1221

513 Spence, A., Simpson, A.J., Mcnally, D.J., Moran, B.W., McCaul, M. V., Hart, K., Paull, B.,
514 Kelleher, B.P., 2011. The degradation characteristics of microbial biomass in soil.
515 Geochimica et Cosmochimica Acta 75, 2571–2581. doi:10.1016/j.gca.2011.03.012

516 Stenson, A.C., Marshall, A.G., Cooper, W.T., 2003. Exact masses and chemical formulas of
517 individual Suwannee River fulvic acids from ultrahigh resolution electrospray ionization
518 Fourier transform ion cyclotron resonance mass spectra. Analytical Chemistry 75, 1275–
519 1284. doi:10.1021/ac026106p

520 Tate, K.R., Ross, D.J., Feltham, C.W., 1988. A direct extraction method to estimate soil

521 microbial c: effects of experimental variables and some different calibration procedures.
522 Soil Biology and Biochemistry 20, 329–335. doi:10.1016/0038-0717(88)90013-2

523 Throckmorton, H.M., Bird, J.A., Dane, L., Firestone, M.K., Horwath, W.R., 2012. The source
524 of microbial C has little impact on soil organic matter stabilisation in forest ecosystems.
525 Ecology Letters 15, 1257–1265. doi:10.1111/j.1461-0248.2012.01848.x

526 Tremblay, L., Benner, R., 2006. Microbial contributions to N-immobilization and organic
527 matter preservation in decaying plant detritus. Geochimica et Cosmochimica Acta 70,
528 133–146. doi:10.1016/j.gca.2005.08.024

529 Vance, E.D., Brookes, P.C., Jenkinson, D.S., 1987. An extraction method for measuring soil
530 microbial biomass C. Soil Biology and Biochemistry 19, 703–707. doi:10.1016/0038-
531 0717(87)90052-6

532

Table

Table 1: Proportion of functional groups in different molecular weight (MW) size classes of microbial compounds as obtained by SEC-LC-FTIR analyses. Number of pluses signifies relative abundance in fumigated soil extracts in comparison to non-fumigated ones.

Functional group	Microbial compounds				
	Low MW	Mid MW		High MW	Bulk
	343-398 Da	453-620 Da	1593-2107 Da	2788-5844 Da	
Aliphatic CH	nd	++	+	++	++
Carbohydrate, P=O	+[#]	++	+	+	+
OH	++	+			
C-N, C-O	+[#]	+		[#]	
C-O-C, C-OH	++				
C=O (COOH)	+[#]				

= Relative absorbance values between replicates were variable, thus comparison was less reliable

nd = not determined because the signal was too low over the entire FTIR spectra for this part of the track

Table 2: Absolute and relative abundance of ESI-FT-ICR-MS derived formulae identified as microbial and SOM related. Note: AI < 0.5: non-aromatic, AI > 0.5: aromatic and AI > 0.67: condensed aromatic.

	Microbial in %	SOM-related in %	Microbial absolute	SOM-related absolute
Total formulae			171	706
0 N	46.8	80.6	80	569
1 N	41.5	17	71	120
2-4 N	11.7	2.4	20	17
AI ≤ 0.5	87.1	75.2	149	531
0.5 < AI < 0.67	5.8	14.2	10	100
AI ≥ 0.67	7.1	10.6	12	75

Table 3: Estimated turnover time for the different molecular size fractions of soil microbial biomass. Also given is the empirical molecular size range and relative abundance of the size fractions.

Molecular size class	Turnover time (d)	Relative abundance (%)
Low MW (< 408 Da)	12.8 ± 0.6	8
Mid MW (408-2072 Da)	18.5 ± 0.6	80.4
High MW (2072-10510 Da)	22.9 ± 0.7	11.6

Figure 1

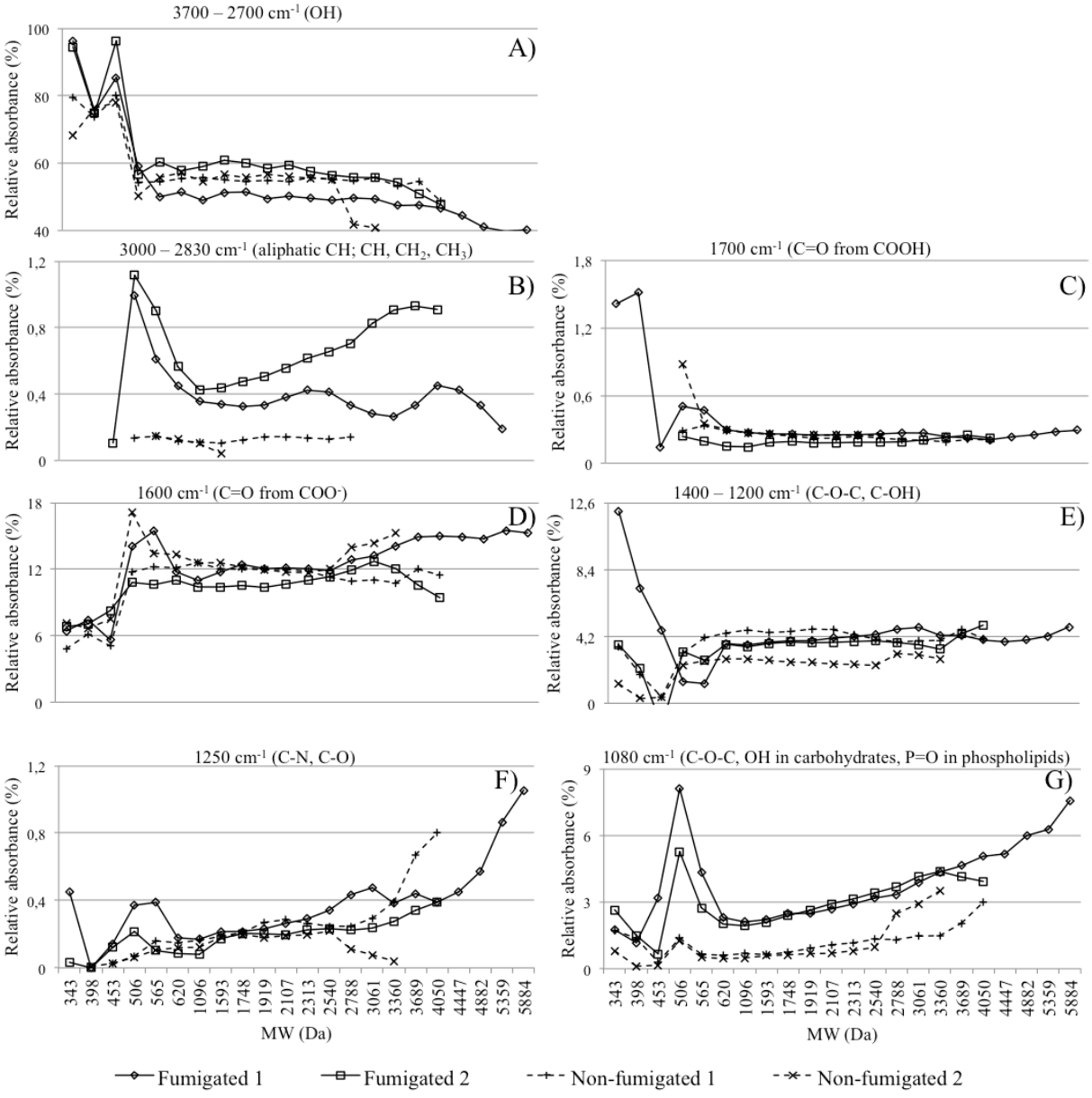

3700 – 2700 cm-1 (OH)
A)
Relative absorbance (%)
3000 – 2830 cm-1 (aliphatic CH; CH, CH2, CH3)
B)
1700 cm-1 (C=O from COOH)
C)
1600 cm-1 (C=O from COO-)
D)
1400 – 1200 cm-1 (C-O-C, C-OH)
E)
1250 cm-1 (C-N, C-O)
F)
1080 cm-1 (C-O-C, OH in carbohydrates, P=O in phospholipids)
G)
MW (Da)
Fumigated 1
Fumigated 2
Non-fumigated 1
Non-fumigated 2

Figure 2

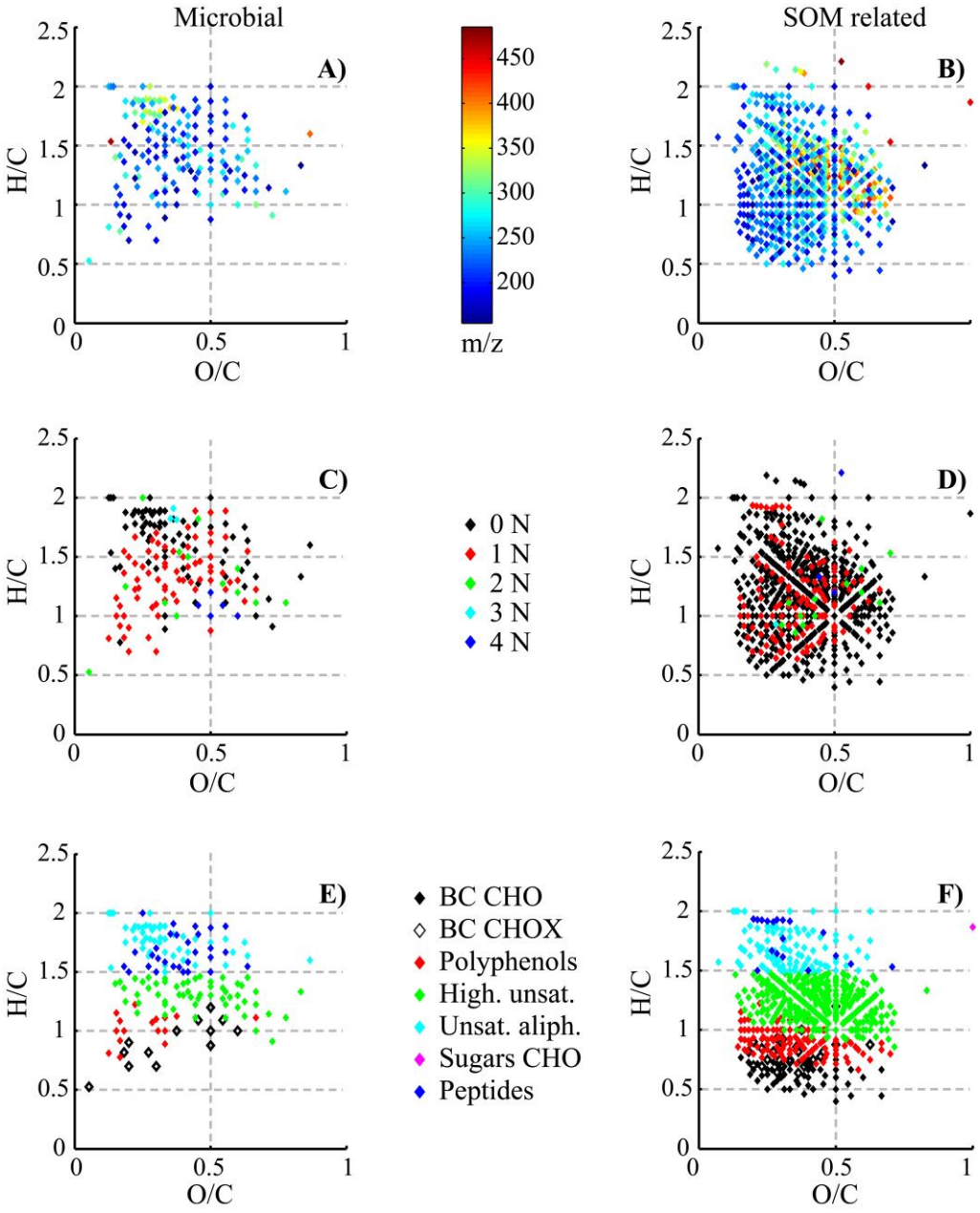
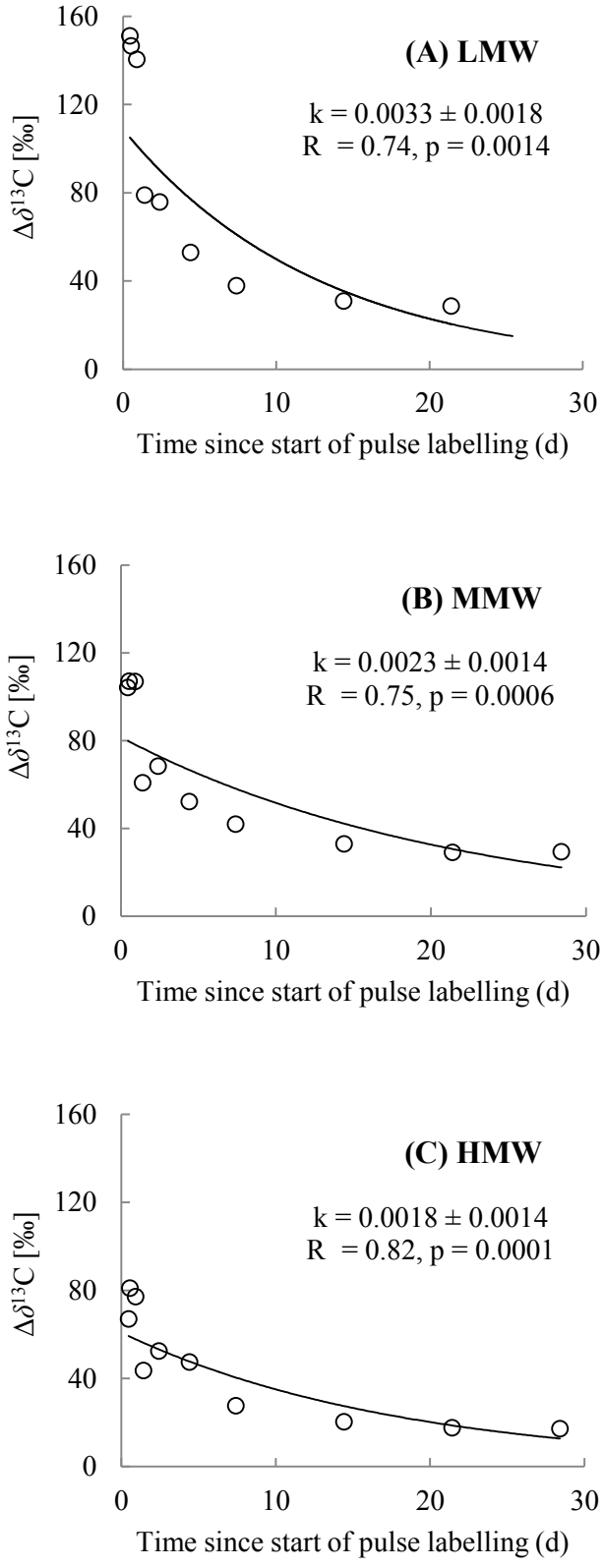

Microbial
SOM related
A)
B)
C)
D)
E)
F)
H/C
O/C
m/z
450
400
350
300
250
200
0 N
1 N
2 N
3 N
4 N
BC CHO
BC CHOX
Polyphenols
High. unsat.
Unsat. aliph.
Sugars CHO
Peptides

Figure 3


(A) LMW
k = 0.0033 ± 0.0018
R = 0.74, p = 0.0014
(B) MMW
k = 0.0023 ± 0.0014
R = 0.75, p = 0.0006
(C) HMW
k = 0.0018 ± 0.0014
R = 0.82, p = 0.0001
Δδ13C [‰]
Time since start of pulse labelling (d)
0
40
80
120
160
0
10
20
30