

OPEN

# Molecular level study of hot water extracted green tea buried in soils - a proxy for labile soil organic matter

Nicholle G. A. Bell <sup>\*</sup>, Alan J. Smith, Yufan Zhu, William H. Beishuizen, Kangwei Chen, Dan Forster, Yiran Ji & Elizabeth A. Knox

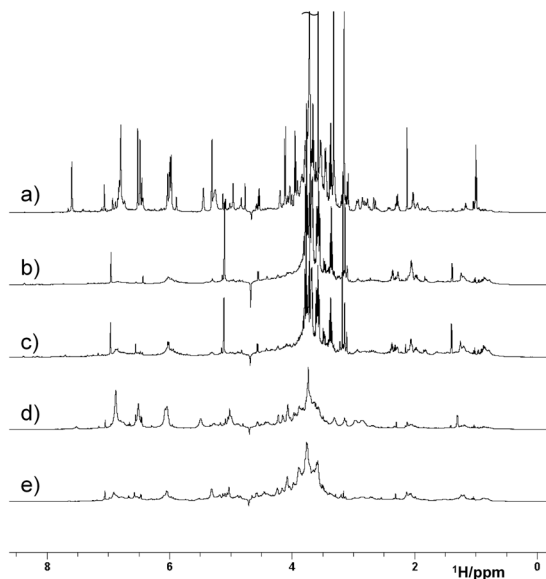
Understanding the composition of soil organic matter (SOM) is vital to our understanding of how soils form, evolve and respond to external stimuli. The sheer complexity of SOM, an inseparable mixture of thousands of compounds hinders the determination of structure-function relationships required to explore these processes on a molecular level. Litter bags and soil hot water extracts (HWE) have frequently been used to study the transformation of labile SOM, however these are still too complex to examine beyond compound classes. In this work, a much simpler mixture, HWE buried green tea, was investigated by Nuclear Magnetic Resonance (NMR) spectroscopy and Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR-MS), as a proxy for labile SOM. Changes induced by the burial over 90 days in a grassland, woodland and two peatland sites, one damaged by drainage and one undergoing restoration by drain-blocking, were analysed. Major differences between the extracts were observed on the level of compound classes, molecular formulae and specific molecules. The causes of these differences are discussed with reference to abiotic and biotic processes. Despite the vastly different detection limits of NMR and MS, chemometric analysis of the data yielded identical separation of the samples. These findings provide a basis for the molecular level interrogation of labile SOM and C-cycling processes in soils.

Earth's soils store a greater amount of C than both the vegetation and atmosphere combined and thus play a crucial role in the global C cycle<sup>1</sup>. The accumulation of C is largely the result of restricted decomposition rates, which are ecosystem dependent. Three major causes for these restrictions include climate (temperature and moisture), substrate quality (chemical and physical characteristics), and the composition, abundance and activity of the soil biota<sup>2,3</sup>.

Soil decomposition has historically been studied using litter bags, which consist of plant material of known mass enclosed in a screened container<sup>4,5</sup>. A variety of different litter bag compositions and protocols have been used over the years, which makes a cross-study comparison practically impossible<sup>6,7</sup>. More recently, a standardized method using green and rooibos Lipton tea bags has emerged<sup>8</sup>, which involves burying pairs of pre-weighed green and rooibos tea bags 15 cm apart to a depth of 8 cm. After 90 days, the tea bags are collected, dried and weighed. Using the fact that the rooibos tea decomposes slower than the green tea permits the so-called stabilization factor,  $S$  and decomposition rate,  $k$ , to be calculated from the determined weight losses. These two variables constitute the Tea Bag Index (TBI) which has been measured for different soils worldwide in order to define the relationships between soil types, decomposition rates and environmental factors<sup>6,8-10</sup>. The tea bag decomposition rates,  $k$ , have been shown to positively correlate with those of local litter and respond in similar ways to biotic factors<sup>7,9,11</sup>.

On a molecular level soil litter decomposition can be divided into, at least, two stages<sup>12</sup>. The first stage involves the non-lignin and soluble compounds, e.g. carbohydrates, small phenolics, amino acids, which are easily leached, hydrolysed or oxidised, while the second stage involves the transformation of recalcitrant molecules<sup>6</sup>, e.g. lignin-like compounds. Considering a burial period of 90 days and the typical turnover rates of different litter compound classes<sup>13</sup>, the latter stage will not occur to a significant extent<sup>8</sup> in soils of temperate climates, therefore largely the labile SOM is affected on this time scale. Labile SOM is considered to be the most sensitive indicator of SOM quality and decomposability<sup>14,15</sup>, which justifies the study of this soil pool.

EaStChem, School of Chemistry, University of Edinburgh, David Brewster Rd, Edinburgh, EH9 3FJ, UK. \*email: [Nicholle.Bell@ed.ac.uk](mailto:Nicholle.Bell@ed.ac.uk)



**Figure 1.** 600 MHz  $^1\text{H}$  NMR spectra of the HWE (a) unburied, (b) grassland, (c) woodland, (d) RMII and (e) RMI buried green tea. The assignment of the unburied tea signals is given in Supplementary Fig. S2.

Green tea is produced from the same tea leaves as black tea (*Camellia sinensis*) with the only difference being the lack of an oxidation step in the manufacturing process. Rooibos tea is produced from the leaves and stems of the *Aspalathus linearis* bush, which are oxidized<sup>16</sup>. The addition of stems as well as the oxidation process make rooibos tea less labile (less soluble)<sup>7,8,17</sup>, and thus less suitable for monitoring the initial stages of decomposition over the 90 day TBI burial period.

The composition of green tea has been studied extensively and it is known to consist of a mixture of phenolics, terpenoids, xanthines, alkaloids, flavones, flavonoids, fatty acids, amino acids and sugars<sup>18–21</sup>, with each compound class containing many molecules. For example Hooft *et al.* produced the most comprehensive list to date of tea phenolics identifying a total of 177 compounds<sup>22</sup>. Of these, polyphenolics called catechins exist in the highest concentration dominated by epigallocatechin gallate (EGCG) followed by epigallocatechin (EGC)<sup>23</sup>.

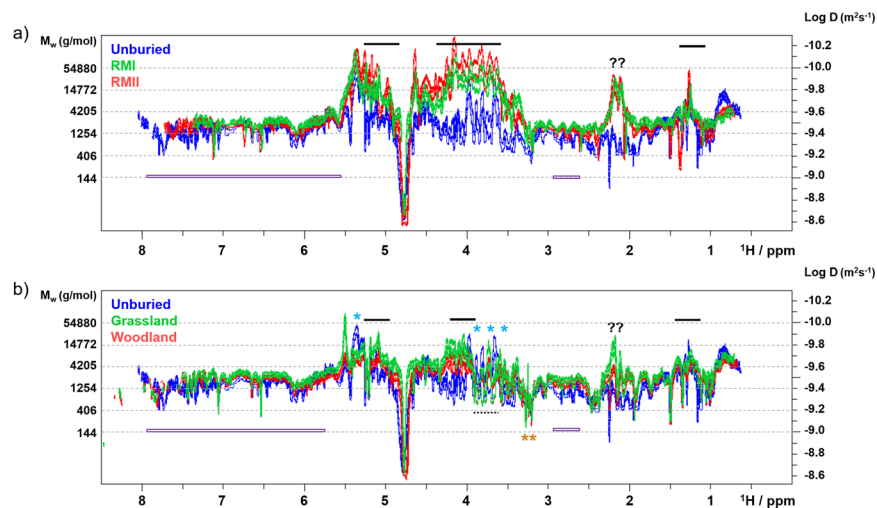
Undoubtedly simpler than the heterogeneous plant litter or SOM itself, monitoring the changes of buried tea on the molecular level is therefore feasible and could advance our understanding of decomposition processes that occur in soils. To study these processes, some form of OM isolation is required. Considering that hot water extracted (HWE) SOM is widely regarded as representative of labile SOM<sup>15,24,25</sup>, buried tea was subjected to HWE in this study. Unlike cold water extraction, HWE yields several times the amount of C and also extracts microbial metabolites and rhizodeposits<sup>15,26</sup>. It therefore represents both the labile OM and microbial biomass.

Solubilised tea OM is a complex mixture that requires powerful analytical techniques, liquid-state NMR and FT-ICR-MS<sup>27,28</sup>, to be studied. This study presents molecular level compositional differences of HWE buried green tea as a proxy for labile SOM decomposition of three different soil types: species rich grassland, woodland and peatland.

## Results

**NMR analysis of Hot Water Extracted (HWE) green tea samples.** An overlay of representative 1D  $^1\text{H}$  spectra from HWE buried and unburied tea is shown in Fig. 1.

The spectra can be divided into three main regions: 8.5–6 ppm – aromatics (mostly catechins), caffeine (in the unburied tea); 5.5–3 ppm – mostly carbohydrates, caffeine, amino acids, catechins and organic acids; and 3–0 ppm, aliphatic molecules and moieties, such as amino acids, fatty acids and catechins. Assignment of the major signals present in the  $^1\text{H}$  NMR spectrum of the HWE unburied tea based on literature data<sup>22,29,30</sup> and 2D NMR analysis<sup>31</sup>, is given in Supplementary Fig. S2. The spectrum of unburied green tea shows a number of sharp resonances, as one would expect for a mixture of small to medium size molecules. Signals of a polysaccharide, amylopectin<sup>32</sup>, were also identified (Supplementary Fig. S3). There are clear differences between the HWE spectra of the buried and unburied tea, which indicate that burial has changed the molecular composition. Starting with the aromatic region, the signals of catechins dominate the spectra obtained from the unburied tea and the Red Moss of Balerno raised bog under restoration (herein referred to as RMII), as broad signals in the latter sample, while these are greatly reduced in the other samples (Fig. 1 and Supplementary S4). The carbohydrate region separates spectra into two groups: peatland and grassland/woodland samples. For both groups, most of the unburied tea carbohydrates have disappeared (Supplementary Figs. S5 and 6), including the polysaccharide, amylopectin. New polysaccharide signals appeared in the spectra of peatland samples that belong to arabinogalactans - a known component of the green tea<sup>33</sup>. While in the woodland/grassland samples the small carbohydrates, trehalose and mannitol were identified, which are known stress molecules of microbial origin (Supplementary Fig. S7)<sup>34,35</sup>. Finally, in the unburied tea, the region below 3 ppm contains signals of typical tea compounds ( $\text{CH}_2$  groups of



**Figure 2.** The 800 MHz 2D DOSY spectra of the HWE tea samples. Overlay of the (a) unburied tea spectra with RMI/RMII and (b) grassland/woodland spectra. Selected areas are labelled and discussed in the text. Blue stars above the spectra indicate signals of amylopectin, arabinogalactans are labelled by horizontal bars, dashed bars indicate signals of trehalose and mannitol. Orange stars indicate OCH<sub>3</sub> groups. Question marks label unassigned signals from large molecules. Catechin signals are marked by purple bars. The shift of the signals towards larger molecular weight seen in the grassland spectra could be due to presence of oxidation products of catechins, e.g. Theasinensins<sup>42</sup>.

catechins, theanine - a tea specific amino acid -, other amino acids, carboxylic acids and fatty acids). In the peatland samples these CH<sub>2</sub> signals are broad with additional broad humps appearing at around 2.1, 1.3 and 0.9 ppm in all four samples, again with distinct differences between the peatland samples and the grassland/woodland samples. The signals of theanine (and also caffeine) disappeared in all samples and compounds such as alanine (1.47, 3.76 ppm) became more pronounced in the spectra of the latter two samples, while lactic acid (1.32, 4.1 ppm) appeared strongly in the grassland samples. Several grassland samples also contained elevated levels of fatty acids. Some of the spectra obtained from the damaged Red Moss of Balerno raised bog site (herein referred to as RMI) were similar to the RMII spectra, while the others appeared similar to the woodland spectra. The characteristics described above are typical for the majority of the spectra obtained from an individual site (Supplementary Figs. S8–S11).

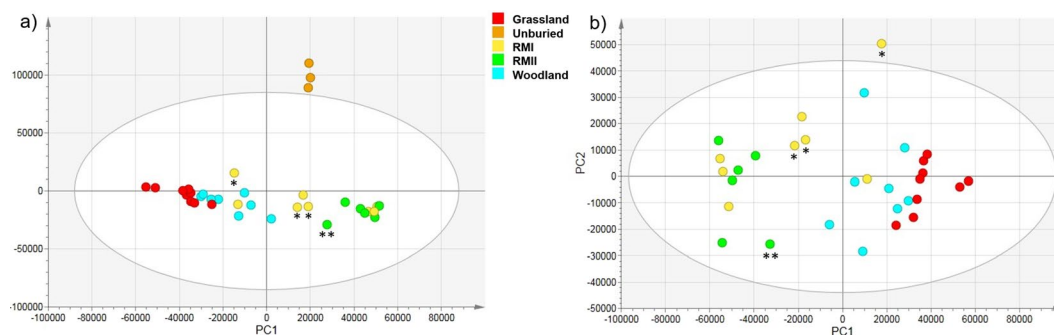
**Relaxation and Diffusion Order Spectroscopy (DOSY) analysis of HWE green tea samples.** To investigate the nature of the changes in the <sup>1</sup>H spectra discussed above, *T*<sub>1</sub> and *T*<sub>2</sub> relaxation and DOSY experiments were recorded on the representative samples presented in Fig. 1. Generally, the *T*<sub>1</sub> and *T*<sub>2</sub> relaxation times of buried samples shortened significantly (see Supplementary Tables SIII–SV and Figs. S12–S21) relative to the unburied tea sample. The RMI sample had the shortest relaxation times, while the woodland sample had values close to the unburied tea sample. These experiments indicated that, within the NMR detection limits, both large and small molecules were generated by the burial process, unexpectedly across the compound class range.

Analysis of the DOSY spectra (Fig. 2) confirmed the results of the relaxation experiments showing a spread of molecular sizes between 90 g/mol to 55,000 g/mol (Supplementary Fig. S22). The DOSY peaks of the unburied tea sample appeared in the expected molecular range, reflecting the presence of small (amino acids such as theanine, caffeine, theobromine, gallic acid, glucose), medium (sucrose, catechins) and large (mainly amylopectin) molecules.

The DOSY spectrum of the RMII sample confirmed that the broad signals in the aromatic region belong to catechins and not to large molecules. The broadness thus being a consequence of heterogeneity due to their partial modifications or dynamic aggregation. The dominant and unique feature of the DOSY spectra of the RM samples (Fig. 2, Supplementary Figs. S24 and 25) is the appearance of a large arabinogalactan (Supplementary Figs. S5 and S6)<sup>33</sup>. On the other hand, sharp and intense signals resonating between 3.5 and 4 ppm in the spectra of the woodland and grassland samples (Supplementary Figs. S26 and 27) belong to trehalose and mannitol<sup>34,35</sup>.

**Chemometric analysis of <sup>1</sup>H NMR spectra of HWE green tea samples.** Principal Component Analysis (PCA) of <sup>1</sup>H spectra of 34 HWE tea samples was performed to characterise the sample set. Two principle components were extracted (Fig. 3a) accounting for 76% of the total variance (39 and 37% by PC1 and PC2, respectively). The percentage of the total predicted variance was 85% (28 and 57% by PC1 and PC2, respectively). As expected, the unburied tea samples clearly separated from the rest, falling outwith the 95% Hotelling's *T*<sup>2</sup> ellipse.

Focusing on the buried samples, PCA was repeated excluding spectra of the unburied tea (Fig. 3b). Two principle components were again extracted accounting for 72% of the total variance (59 and 12% by PC1 and PC2, respectively). The percentage of the total predicted variance was 75% (56 and 20% by PC1 and PC2, respectively). The samples separated along PC1 in a similar manner as seen in the relaxation and DOSY experiments, into



**Figure 3.** PCA score plot based on the 1D  $^1\text{H}$  600 MHz NMR spectra of (a) all 34 green tea samples: unburied, tea buried in grassland, woodland, RMI and RMII, (b) only buried tea samples. The legend indicates the colour coding for each sample site. The Hotelling's  $T^2$  ellipse represents the 95% confidence interval. Outlying RMI or RMII samples are marked by a single or double asterisks, respectively.

RMI/RMII and grassland/woodland groups. The largest separation was seen between the RMII and grassland samples, while the woodland samples partially overlapped with the grassland samples. The RMI samples were the most heterogeneous group, with some overlapping with the RMII and others with the woodland samples. Visual inspection of the RMI  $^1\text{H}$  NMR spectra confirmed that the respective samples share a similar appearance to either the RMII or woodland spectra (Supplementary Figs. S10 and S11). Overall, the PCA analysis of the  $^1\text{H}$  NMR spectra of the HWE buried tea indicated that the within-group variation was significantly less than the between-group variation for the RMII and grassland samples. The PC1 loadings (Supplementary Fig. S28) identified spectral regions responsible for this separation reflecting the disappearance of catechins and the appearance of trehalose, mannitol and aliphatic compounds (unassigned signals at  $\sim 3.15$  ppm) in the woodland, grassland and some RMI samples, as well as the appearance of arabinogalactans in most RMII and some RMI samples. The amylopectin anomeric signals visible in the PC2 loadings at  $\sim 5.4$  ppm were most abundant in some of the RMI spectra.

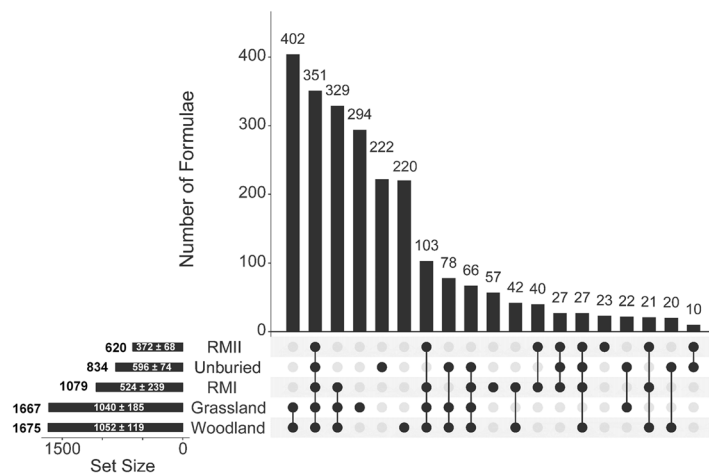
The subsequent PLS-DA analysis (Supplementary Figs. S29 and 30), in addition to demonstrating a clear separation of the RMII and grassland samples seen by PCA, also separated the woodland samples from the rest. Some RMI samples continued to overlap with RMII samples, while some moved towards the woodland samples. The regions of the NMR spectra that contributed most to the PLS clustering are indicated in the form of biplots (Supplementary Fig. S29); the compounds resonating in these areas were identical to those identified by PCA analysis presented above. The hierarchical clustering analysis (HCA) dendrogram based on PLS-DA (Supplementary Fig. S30) mirrors the clustering of the PLS-DA score plots with labelling of individual samples linking them to the original  $^1\text{H}$  NMR spectra (Supplementary Figs. S8–11).

Despite significant overlap of  $^1\text{H}$  signals, chemometric analysis of the NMR data identified differences between the composition of the HWE buried tea from different soil types and provided structures of discriminatory molecules.

**FT-ICR-MS analysis of the HWE green tea samples.** The sensitivity and resolution of FT-ICR-MS provides a complementary view to the analysis of tea offered by NMR. High quality MS spectra were obtained for unburied and buried green tea samples (Table SVIII) with between 489 and 2023 peaks observed with  $S/N > 5$  in the  $m/z$  range between 150 and 1000 g/mol. An assignment rate of 81–86% (including isotopologues) with a standard deviation  $\leq 2\%$  was achieved for all samples. The average number of formulae assigned per site had a low standard deviation ( $\pm 11$ –18%) with the exception of the RMI samples ( $\pm 46\%$ ) (Table SIX).

Altogether, 2385 unique monoisotopic molecular formulae were identified in the complete sample set (24%  $\text{C}_a\text{H}_b\text{O}_c\text{N}_d$ , and 76%  $\text{C}_a\text{H}_b\text{O}_c$ ). Woodland and grassland samples yielded the largest number of formulae (1675 and 1667, respectively), approximately double the number seen in the unburied samples (834). This is in stark contrast with the peatland samples that yielded numbers closer to those observed for the unburied tea, nevertheless, with distinct differences. For RMI samples, 29% more formulae (1079) were identified, while the trend was the opposite for the RMII samples, with 26% less formulae (620) identified. A closer inspection of the six RMI samples indicated that half of the samples had a larger average number of molecular formulae assigned (721), while the other samples had less than half (327) – an amount comparable to the RMII samples. The number of nitrogen compounds identified in the buried samples was in the order: grassland > woodland > RMI > RMII. This order follows the elemental analysis of the non-HWE buried samples (Table SII).

The assigned molecular formulae were analysed with the help of an UpSet plot (Fig. 4, Table SX)<sup>36</sup>. These plots help visualize the distribution of compounds in a sample set, revealing a wealth of information. Molecular formulae for samples from  $n = 5$  different environments can be represented by a maximum of  $2^5 - 1 = 31$  intersections ( $I_1$  to  $I_{31}$ ) containing compounds common to  $k = 1$ –5 sample sets. Symbols  $I_i \binom{5}{k}$  are used here to refer to individual intersections; these were ordered in the descending order of the number of formulae they contain. Formulae were only found up to  $I_{28} \binom{5}{k}$  and the first 19 intersections shown in Fig. 4 represent more than 99% of the compounds.

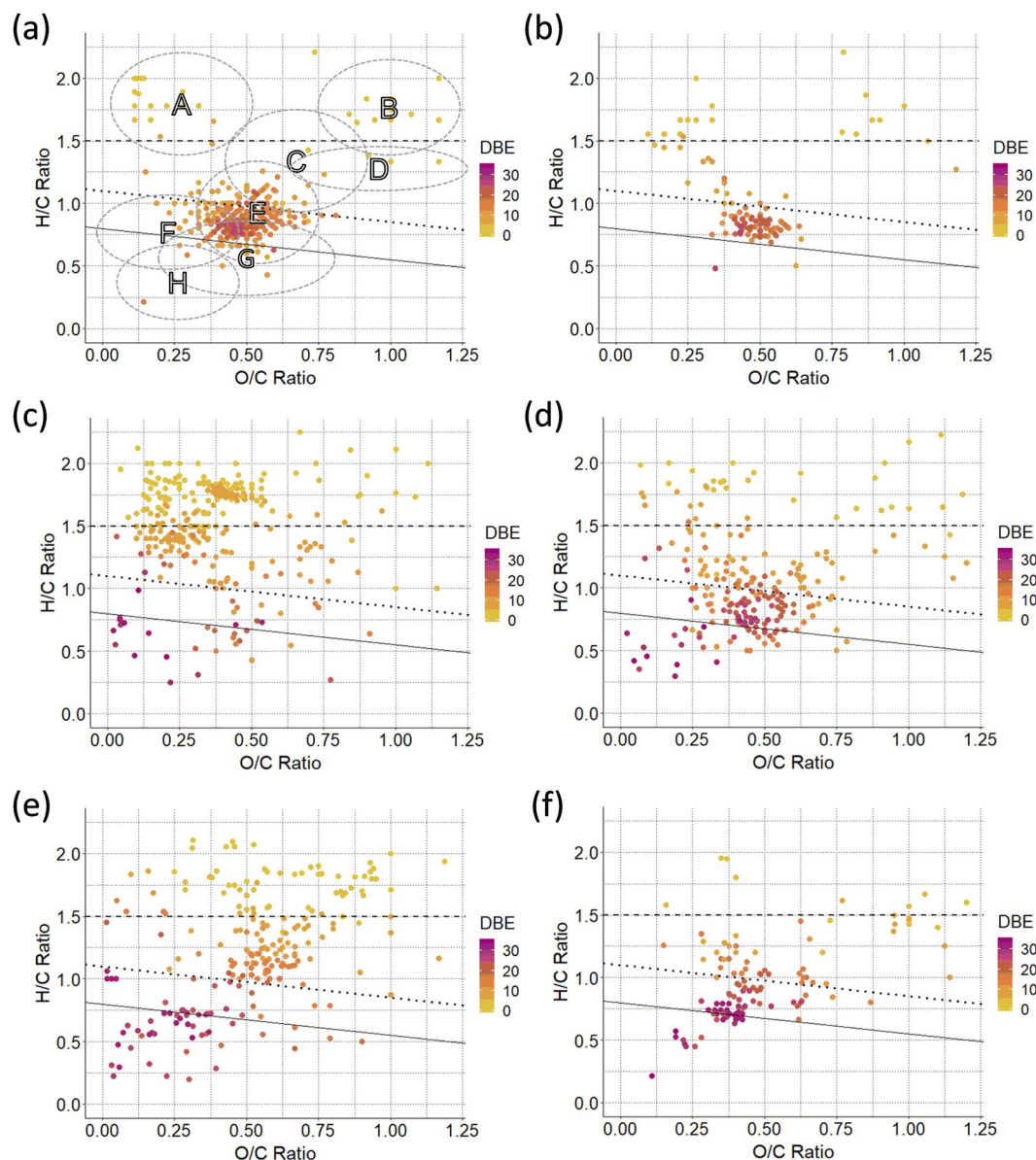


**Figure 4.** UpSet plot of the (-) ESI FT-ICR-MS data of the green tea samples. The plot shows for each site: the total number of unique total formulae (on the left); the average + standard deviation of formulae assigned (inside the horizontal bar). The vertical bars indicate the number of common assigned formulae for individual intersections defined by black circles and ordered in the descending population. Intersections  $I_{1-3}$  covered 46% of the assigned formulae, while the  $I_{4-6}$  and  $I_{7-9}$  intersection cumulatively extended the coverage to 77% and 88%, respectively.

A closer inspection of the intersection plots gives further insight into the molecular make-up of the tea buried in different sites. As  $I_1$  to  $I_9$  contained 88% of unique formulae, a detailed analysis presented here is limited to these intersections. The largest number of formulae (402) belongs to the  $I_1 \binom{5}{2}$  group containing formulae found in woodland and grassland samples only. Similarity between the grassland and woodland samples is reinforced by the fact that the  $I_2 \binom{5}{5}$  intersection, containing 351 formulae common to all sites, was the second most populated. It was interesting to see that the next most abundant intersection was  $I_3 \binom{5}{3}$ , containing 329 formulae from the grassland, woodland and RMI samples. In contrast, the  $I_{22} \binom{5}{3}$  intersection, containing formulae from the grassland, woodland and RMII, showed only 5 formulae (data not shown). This points towards similarities between the green tea buried at RMI and grassland/woodland sites and indicates differences between the green tea buried at the two RM sites. The  $I_4$  to  $I_6$  intersections are all of the  $\binom{5}{1}$  type, i.e. contain unique formulae found in one site only. The 222 formulae identified in the unburied tea only,  $I_5$ , represent compounds that were degraded in all environments. The numbers in the grassland ( $I_4$ ) and woodland ( $I_6$ ) intersections are significant (half of the  $I_1$  compounds), indicating the different effect these environments have on the tea samples. The next three intersections,  $I_{7-9}$ , extend the coverage to 88% of compounds. The  $I_7 \binom{5}{4}$  intersection, encompassing all but the unburied tea samples, showed 103 formulae. These represent compounds that were removed from samples in all environments. The next two intersections  $I_8$  and  $I_9$  are of  $\binom{5}{3}$  and  $\binom{5}{4}$  type, respectively. Neither includes the RMII site samples, both include unburied, woodland and grassland samples, and in the case of the  $I_9 \binom{5}{4}$  also the RMI samples. These intersections therefore identify which compounds were removed from the RMI and RMII samples (78 formulae) and exclusively from the RMII samples (66 formulae), respectively. It is interesting to note that the  $I_{24} \binom{5}{4}$  intersection that contained formulae exclusively removed from the RMI samples had only three entries. These observations further emphasize the unique makeup of the tea buried in the RMII site. From the point of view of new formulae in the RM samples, these appear in  $I_{10} \binom{5}{1}$ ,  $I_{15} \binom{5}{1}$  and  $I_{12} \binom{5}{2}$  intersections, corresponding exclusively to RMI (57 formulae), RMII (23 formulae), and both RM samples (40 formulae), respectively. This is approximately 10 times less than the unique formulae detected in the woodland and grassland samples.

**Molecular level analysis of the FT-ICR-MS data.** The H/C and O/C ratios were determined for the individual molecular formulae obtained from the FT-ICR-MS data and these were inspected using van Krevelen diagrams. These are presented for the representative samples in Supplementary Fig. S31. At the same time, the molecular formulae were examined by calculating their double bond equivalent (DBE) and a modified aromaticity index ( $AI_{\text{mod}}$ )<sup>37</sup>. The results presented here focus on van Krevelen diagrams of selected intersections (Fig. 5) as identified by the UpSet plot; these intersections were also characterized by the oxygen class and mass distribution of their compounds (Supplementary Figs. S32 and S33).

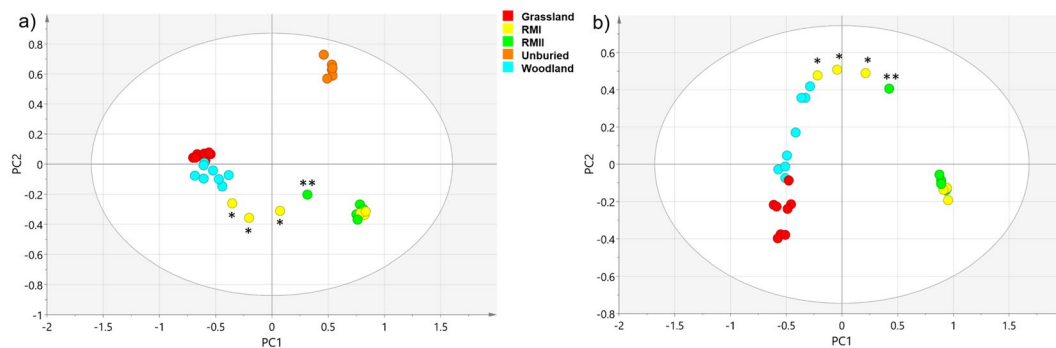
*Preserved in all environments ( $I_2$ , 351 compounds).* Using the  $AI_{\text{mod}}$  index, the majority of these compounds (189,  $M_w$  174–900 g/mol) can be classified as non-aromatic, followed by aromatic compounds (130,  $M_w$  168–900 g/mol).



**Figure 5.** Van Krevelen diagrams for selected intersections identified by the UpSet plot. (a) preserved in all,  $I_2$ ; (b) produced in all  $I_7$ ; (c) produced in grassland only,  $I_{41}$ ; (d) produced in woodland only,  $I_6$ ; (e) removed in all environments,  $I_5$ ; (f) produced in RMI or RMII ( $I_{10} + I_{12} + I_{15}$ ). Colour coding denotes the DBE. The lines represent aliphatics (above the dashed line), aromatics (between the dotted and the full line) and condensed aromatics (below the full line) based on the  $AI_{mod}$  index<sup>76</sup>. Letters denote the following compound classes, A: fatty acids/lipids, B: carbohydrates, C: glycosides, D: nucleic acids, E: polyphenols, F: lignin-like compounds, G: oxidised polyphenols, H: condensed aromatics.

Their mass distribution is centred between 200 and 600 g/mol, while their oxygen classes show a broad, skewed Gaussian distribution between  $O_2$  and  $O_{17}$  with a maximum at  $O_{12}$ . The majority of molecular formulae identified in green tea by others are present in this intersection<sup>22,38–40</sup>. These include fatty acids, namely, oleic acid, palmitic acid, linolenic acid, linoleic acid, stearic acid, myristic acid<sup>38,39</sup>, while others are tentatively assigned to pentadecanoic acid and palmitoleic acid. The glycoside region contains molecules, such as rutin and kaempferol glycosides<sup>22,41</sup>. The molecular formula  $C_{12}H_{22}O_{11}$  represents many disaccharides, including sucrose and trehalose that were identified by NMR. All major catechins were identified, the most abundant being ECGC, however dimers and oxidative products, such as theaflavins, were also found. Removal of catechins in some samples, as detected by NMR, was therefore not complete. There is very little of condensed aromatics, all of which are low molecular weight compounds (12,  $M_w$  162–302 g/mol).

*Produced in all environments ( $I_7$ , 103 compounds).* Only 103 (61 non-aromatic, 39 aromatic, 3 condensed aromatic) new compounds were detected in all soil environments. Their mass distribution was bimodal with maxima



**Figure 6.** PCA score plot from the (-) ESI FT-ICR-MS data. Each circle represents an individual HWE tea sample mass spectrum. (a) analysis of unburied tea, tea buried in grassland, woodland, RMI and RMII; (b) analysis of only buried tea samples. The legend indicates the colour coding for each sample site. The Hotelling's  $T^2$  ellipse represents 95% confidence interval. Outlying RMI and RMII samples are marked with a single and double asterisks, respectively.

at 300 and 600 g/mol, respectively, while their oxygen class distribution peaked at  $O_{15}$ . The non-aromatic molecules were a mix of fatty acids and glycosides with a few carbohydrates. Molecular formulae of theaflavins and theanaphoquinones<sup>42,43</sup>, oxidative products of the green tea typically found in the black tea, were identified.

*Removed from all environments ( $I_5$ , 222 compounds).* The majority of these compounds are non-aromatic (166,  $M_w$  150–800 g/mol) falling into the glycoside area, followed by fatty acids, carbohydrates and condensed aromatics. Their mass distribution was Gaussian, centred at around 600 g/mol, showing a fairly even range of masses, while higher oxygen classes were more populated. Few tentative assignments were made, namely  $C_5H_{10}O_5$  which is likely arabinose and  $C_{15}H_{18}O_9$ , which has been identified as caffeoyl-D-glucose. Many of the  $I_5$  molecular formulae are related to  $I_2$  formulae by a specific molecular unit such as oxygen, or isoprene units. Comparing the produced in all and removed from all sites, more glycosides were removed than were produced.

*Produced only in grassland ( $I_4$ , 294), only in woodland ( $I_6$ , 220), and only in peatland ( $I_{10} + I_{12} + I_{15}$ , 120 compounds).* The van Krevelen plots show a clear difference between molecules produced in grassland only ( $I_4$ ) and woodland only ( $I_6$ ), with grassland samples having more fatty acid compounds, while woodland samples contain more flavonoids/glycosides. In addition, more aromatic molecules were produced in the woodland (57 aromatic, 27 condensed) compared to the grassland (17 aromatic, 12 condensed) samples. The molecular weight distribution in the grassland samples shows mainly low  $M_w$  compounds (centred on 400 g/mol), while the woodland produced larger molecules (centred on 600 g/mol). The grassland compounds belong to low oxygen classes (centred on  $O_6$ ), while the woodland compounds have a broad distribution up to  $O_{20}$ . A few amino acids (histidine and tryptophan) could be identified in the grassland samples only, while xanthines such as theophylline as well as certain theaflavin mono/digallates were found only in the woodland samples. The abundance of fatty acids in the grassland samples (also seen by NMR) is a striking feature.

Compared to the grassland and woodland, far fewer compounds were produced in the peatland, over 70% of which were found in the oxidised polyphenols and glycoside regions of the van Krevelen with a bimodal distribution in the oxygen classes centred on  $O_7$  and  $O_{16}$ . A number of their formulae belong to large compounds, e.g. theasinensin digallate and its homologues.

The stress induced molecule, mannitol, was not found in these intersections; it only appeared in the  $I_3$  intersection, in agreement with NMR data.

**Chemometric analysis of the FT-ICR-MS spectra of the HWE green tea samples.** The PCA analysis of the FT-ICR-MS spectra using two principle components accounted for 68% of the total variance (53 and 16% by PC1 and PC2, respectively). The percentage of the total predicted variance was 62% (47 and 15% by PC1 and PC2, respectively). The MS PCA score plots (Fig. 6) indicate significant between-group variation with a tight, well-separated cluster of the unburied tea (two samples measured in triplicate confirming the reproducibility of the method).

Grassland and woodland samples appeared close together in one area of the score plot, with more variation in the woodland sample set. Three RMI samples were located closer to the woodland samples, while the rest overlaid with the RMII samples. Only one RMII sample trended towards the RMI/woodland samples, while the rest clustered tightly in a separate area of the score plot.

The PCA analysis of the FT-ICR-MS spectra excluding the unburied tea samples using two principle components accounted for 72% of the total variance (60 and 12% by PC1 and PC2, respectively). The percentage of the total predicted variance was 65% (56 and 10% by PC1 and PC2, respectively). Removing the unburied tea samples from the analysis (Fig. 6b) preserved the features discussed above and accentuated within-group differences in the grassland and woodland samples, as seen by their spread along the PC2. This was not the case for the RM samples, which did not change their clustering pattern. This PCA score plot shows a significant “arch effect”, which may occur with multivariate methods where the data is projected onto low dimensions. This may reflect a

single dominant gradient in the data and that the samples are highly structured according to PC1. Nevertheless, as discussed<sup>44</sup> the arch effect is not necessary artefactual. Since the MS data show a similar ordination of samples as the NMR data, this is also likely the case here.

PCA loadings for the analysis with and without the unburied tea are shown in Supplementary Figs. S34–S36, respectively. The majority of the discriminating molecular formulae were assigned to the most likely compounds, considering those previously found in green tea<sup>18–20,22,30,40,42,45,46</sup>.

Based on well-separated PCA scores, PLS-DA was performed next. Similar to the analysis of NMR data, the PLS-DA of the MS data accentuated the within-group differences in the grassland and woodland samples, while it did not change the clustering pattern of the RM samples (Supplementary Fig. S37). The corresponding HCA dendrogram (Supplementary Fig. S38) mirrors the clustering of the PLS-DA score plots. The overlaid biplots singled out the molecular formulae/compounds that contributed most to the PLS-DA score plot definition. All were already identified as significant by the PCA and are highlighted in the figure legends of Supplementary Figs. S35 and S36.

The biplots (Supplementary Fig. S37) reflect higher concentrations of EGCG, ECG/CG, gallic acid and  $C_{24}H_{20}O_{15}$  in the RM samples, while trehalose was particularly elevated in the grassland samples. EGC/GC,  $C_{16}H_{14}O_9$ , flavonols (such as myricetin),  $C_{30}H_{26}O_{13}$  (catechin dimers), epicatechin trimer procyanidin C1, as well as, polyphenol flavonoids (such as theasinsin B and C) were elevated in woodland and some RMI samples.

## Discussion

Both high resolution NMR spectroscopy and MS provided a wealth of information regarding changes to the composition of the HWE buried green tea. Analysis of the identified molecular formulae yielded insights into the compound class distribution in individual intersections as defined by the UpSet plot analysis together with their attributes such as molecular mass, aromaticity index and oxygen class. Furthermore, the simplicity of the green tea has allowed the assignment of several individual molecules from the MS and NMR data. Such level of detail has not been achieved to date for buried plant litter or labile SOM due to the complexity of these mixtures, limiting their analysis to compound classes only<sup>47–49</sup>.

Treated chemometrically, both NMR and MS data of the HWE buried green tea revealed significant differences between samples. As PCA showed inter-group variations above the level of intra-group differences, a supervised dimensionality reduction method, PLS-DA, was employed to classify the samples. Reflecting unique transformations of the green tea buried in different soils, this approach has successfully provided molecular signatures of green tea decomposition, including compounds responsible for the ordering of the samples. It is remarkable that despite the several orders of magnitude difference in sensitivity between NMR and FT-ICR-MS, both techniques classified samples in a very similar way. Not only were the woodland and grassland samples clearly separated from the peatland samples, but also some damaged peatland samples (RMI) separated from the samples from the peatland site under restoration (RMII). Labile SOM is sensitive to land-use changes or management<sup>50–52</sup> and therefore analysis of the HWE buried tea has the potential to report on peatland ecosystem health. These observations will form basis of a more extensive study focusing on peatlands (raised and blanket bogs) in different condition.

**Relating the molecular changes to biotic and abiotic factors.** What changes are expected to occur upon burial of tea bags and what drives them? Lipton tea bags are made from a non-woven material that will stop macrofauna from getting inside<sup>8</sup>. Therefore, only physical, microbial and chemical degradation of the buried tea is possible. The observed molecular transformations are the result of a combination of processes which resulted in the (i) disappearance or reduction in the amount of compounds, either by their biodegradation, chemical modification or physical removal via dissolution in water and water flux; (ii) appearance of new, smaller molecules as a consequence of the above processes; (iii) partial degradation of green tea polymers that are not soluble, producing smaller, water extractable poly/oligomers; (iv) oligo/polymerization of compounds; and finally (v) the appearance of small or large molecules of biological origin that are either excreted by, or extracted from, microorganisms after they have been imbedded in the tea samples.

Temperature is one of the major factors affecting decomposition of litter and SOM<sup>53,54</sup>; burial of tea in winter months with relatively stable but low temperature will slow down biotic decomposition<sup>6</sup>, allowing any differences due to other factors to be accentuated. Stable rainfall/snowfall at the sites of the buried tea during winter meant that moisture was not the limiting factor across soils in same region<sup>55</sup>. Another abiotic factor, pH, is an important factor that affects the enzymatic and chemical activity and therefore the level of SOM transformation<sup>56–58</sup>. Access to oxygen is yet another inherent feature that, e.g. in waterlogged peatlands, is a limiting factor for some enzymatic activities<sup>56,59</sup>.

Soil exoenzymes produced by microorganisms are proximate agents of OM decomposition<sup>58</sup> and can largely be classed as hydrolytic (breakdown of labile compounds, e.g. carbohydrates) and oxidative (responsible for transformation of less labile compounds, e.g. aromatics)<sup>58,60,61</sup>. Hydrolytic enzyme activities are dependent on labile carbon content, temperature, soil C/N ratio and pH, while oxidative enzymes are limited mainly by pH and the availability of oxygen<sup>58,62</sup>. While the C/N ratio was not measured for the soils examined in this study, it was for the unburied and buried tea, which can be compared to studies examining the C/N ratio of buried litter. The C/N ratio decreased after burial, which is in an agreement with the incubation studies of litter that reported a decrease in C/N believed to be associated with either the increase in microbial N-containing metabolites or non-decomposable N-containing compounds<sup>15</sup>. Sinsabaugh *et al.* examined the activity of the seven-most widely measured soil enzymes, using data from 40 ecosystems, and found that the oxidase enzyme activity and decomposition increased with pH, while hydrolase enzyme activity showed a weak or strong negative correlation with pH<sup>58</sup>. For the soil types in this study, the C/N ratio of the buried tea decreased in order



of peatland > woodland > grassland, while the pH of the soils decreased in the order of grassland > woodland > peatland. Thus, both factors suggest that biotic decomposition, and specifically oxidative enzyme activity, was higher in grassland soils. This is supported by the fact that the greatest mass loss and greatest number of compounds produced (microbial N-containing plus tea decomposition products) was for tea bags buried in the grassland. The decrease in decomposition from grassland to peatland matches exactly with the results of the initial Tea Bag Index study<sup>8</sup>.

**Molecular level transformations.** The molecular transformations that occur when litter decomposes have been summarised generally as a decrease in carbohydrates and increase in aromatics and aliphatics<sup>47,63</sup>, through hydrolysis or oxidative processes. As tea is a simpler mixture compared to both plant litter and SOM, decomposition can be studied to the highest level of specific molecular changes. For example, considering that black tea is produced by oxidation of the same tea leaves as green tea (*Camellia sinensis*), the major products of which are oxidised catechins such as thearubins, it is possible that oxidative processes in soils could produce the same compounds that are found in black tea<sup>42,64</sup>. Inspection of HWE buried green tea using high resolution MS provided evidence for this. Molecules such as catechin dimers or trimers, as well as thearubins such as theasinensin B and C, were found in woodland and peat samples. However, other molecules such as theaflavin-3-gallate, associated with black tea<sup>42</sup>, were found in all samples, including the unburied tea. It is possible that these compounds were already present in unburied tea at very low concentrations and were detected due to the high sensitivity of FT-ICR-MS.

Appearance of cold stress molecules, trehalose and mannitol, in the grassland, woodland and some RM samples is intriguing and has not yet been reported for buried litter. Trehalose is excreted by fungi and its release accompanies root nodule senescence<sup>65</sup>. Together with mannitol, trehalose is known to protect cells from frost damage<sup>35,65,66</sup>. The appearance of these molecules is likely a consequence of the winter burial period; additional experiments performed in summer/autumn are required to prove this hypothesis.

MS analysis also uncovered an increased presence of lipids in the grassland samples. While their exact structures could not be determined, the majority form homologous series differing by the number of oxygens or CH<sub>2</sub> groups. The abundance of unique hydroxyl and methyl branched fatty acids in grassland compared to woodland soils in colder months has been reported previously due to the greater presence of gram negative bacteria or actinobacteria than gram positive bacteria<sup>67,68</sup>.

Partially hydrolysed arabinogalactans were detected in both the damaged peatland (RMI) and the peatland undergoing restoration (RMII) (but not grassland and woodland samples). Large insoluble arabinogalactans, which have been reported in green tea<sup>20</sup>, are unobservable by liquid state NMR, therefore these must have been produced by hydrolysis which could have occurred abiotically (due the acidic pH of peatlands) and/or biotically (by peat-inhabiting hydrolytic planctomycetes that are active even at low temperatures)<sup>69,70</sup>.

Major differences were observed in aromatic regions of the NMR spectra. Catechins – phenolic molecules – were mostly degraded in the RMI samples (the damaged peatland site). This finding can be explained by the enzyme latch theory, which states that phenolics are elevated in a well-functioning peatland with persistent anoxia and low pH, which hinder the activity of oxidative enzymes like phenol oxidase<sup>56</sup>. Functioning peatlands have an anoxic zone below the predominant water table. Following the instructions of the TBI project, the samples were buried at a depth of 8 cm. This is above the typical water table in both RMI and RMII sites (~15 cm). However, it is possible that during the wet winter months (snow also laid on these sites for two months of the burial period), these could have been submerged under water and hence starved from oxygen. It is also possible that winter temperatures limited the enzymatic activity<sup>71</sup>. One cannot rule out that different microbial communities were present in the damaged and restored sites at the start of the study, which could have led to this difference. Concomitant examination of the microbial community structure and enzyme activity is therefore needed to clarify these observations.

## Conclusions

Untargeted analysis using high resolution NMR and MS techniques was performed to characterise the transformations of hot water extracts of green tea buried in grassland, woodland and peatlands as a proxy to labile SOM. Our results demonstrate that the molecular composition of green tea changes upon burial in a soil type specific manner. The abiotic and biotic factors that caused these changes were discussed. NMR and MS reported on major and minor products of the biochemical transformations of the buried tea, respectively. Furthermore, the use of a simpler mixture allowed the identification of specific molecules that represent the differences in decomposition in individual soil types. Differences were seen in the HWE tea buried between peatland sites that were damaged and 'under restoration', which shows that the molecular characterisation of buried tea could be used to examine the effect of land-use or management changes. The data from this study, together with their chemometric interpretation, provide a basis on which to extend our knowledge of soil decomposition and C-cycling through a widely used litter material – green tea – as a proxy to the transformation of labile SOM.

## Methods

**Tea bag burial sites.** Four sites were investigated in this study (Supplementary Fig. S1). The first two were on a raised peatland (called Red Moss of Balerno) located in Balerno outside of Edinburgh, Scotland (55°51'38.4"N 3°20'15.4"W). This is a designated Site of Specific Scientific Interest as one of the remaining raised bogs in Scotland. The bog has undergone historic damage from grazing, burning and drainage. In the last 30 years its western part has undergone restoration in the form of drain blocking using dams. Two sites were examined, one nearest to the dams (RMII, 55°51'37.7"N 3°20'17.4"W) and a site on the NE side, which is currently unrestored (RMI, 55°51'43.6"N 3°20'04.1"W). The third was a woodland located 100 m to the west of the Red Moss peatland

near Threipmuir reservoir (55°51'37.0"N 3°19'59.9"W). The soil type according to the Soil map of Scotland is a mineral gley<sup>72</sup>. The forth site was a species rich grassland (55°52'23.3"N 3°12'11.8"W) located on the south side of the Pentland hills; samples were placed on the edge of arable plots. The soil type according to the Soil map of Scotland is brown soil<sup>72</sup>. The average pH of each soil is given in Table SI.

**Tea bag burial and extraction.** Eleven sets of non-woven rooibos (Lipton, EAN 87 22700 18843 8) and green tea (Lipton, EAN 87 22700 05552 5) bags were weighed (average weight 1.98 g) and buried at the two peatland sites and the woodland site, while 12 sets were buried at the grassland site in November 2017. Winter burial was chosen to prevent influence of high temperatures and root penetration on the tea decomposition. In addition, the winter months generally provide easier access to the sites, considering conservation activities and staff availability. One rooibos tea bag and one green tea bag were buried 15 cm apart, both at a depth of 8 cm. Three sets of bags were buried 2 m apart and the rest diagonally outwards. Four sets of tea bags were buried 15 m from the centre and four sets 30 m. The samples were collected after 96 days. Out of the 90 tea bags buried, 82 were recovered. To enable submission of data to the Tea Bag Index project, 4–6 bags were taken from each set and processed according to the Tea Bag Index protocol<sup>8</sup>.

The remaining tea bags were dried and weighed with any soil particles on the outside removed. The tea taken out of each bag and weighed, including string plus label, separately before being balled milled. The average tea bag weight loss per site are provided in Table SI.

Samples for NMR and MS analysis were prepared by HWE using the following protocol: 150 mg of ground tea was heated to 80 °C in Milli-Q water for 30 minutes using a microwave. The supernatants were filtered using glass fibre filters (GF/C, Whatman) and freeze dried. On average 15 mg of the freeze dried green tea extract was obtained per tea bag.

**Elemental analysis.** The elemental composition (C, H, N) of ground tea samples was determined using an ThermoScientific FlashSmart analyser. C, H, N was determined for selected tea bags with duplicates from each burial site. Elemental analysis of tea bags that were unburied but HWE and unburied and not HWE was also conducted (Table SII).

**<sup>1</sup>H NMR spectroscopy.** For all NMR analysis of unburied and buried green tea samples (Bruker 600/800 MHz AVANCE III with 5 mm TCI cryoprobes), 5 mg of freeze dried HWE tea was dissolved in 600 µL of D<sub>2</sub>O (100% deuterated, Sigma Aldrich). A 1D NOESY based water suppression pulse sequence was used to acquire the 1D <sup>1</sup>H spectra. The following parameters were used: 64 scans, 9 s relaxation delay, 1.3 s acquisition time, 40 ppm spectral width and 80% pulsed field gradients during a 100 ms the mixing time. Overall experimental time was 12 minutes per sample. Line broadening of 1 Hz and zero filling were applied prior to Fourier transformation. Altogether 34 spectra were acquired including samples of the unburied tea (3), and samples buried in woodland (8), grassland (9), the damaged peatland (RMI, 8) and the peatland under restoration (RMII, 6).

**Parameters of additional NMR experiments.**  $T_1$  relaxation experiments were performed on representative green tea samples using a modified inversion recovery pulse sequence on the 800 MHz spectrometer equipped with a TCI cryoprobe at 300 K. The following parameters were used: inversion recovery times of  $\tau_{min} = 1$  ms and  $\tau_{max} = 1400$  ms and 16 relaxation points, 4 dummy scans, 128 scans, 40 ppm spectral width, 2 s relaxation delay and 1 s acquisition time. Water presaturation was applied during the prescan relaxation delay and relaxation intervals.

$T_2$  experiments were acquired on the same sample set using a CPMG sequence on the 800 MHz spectrometer. The parameters used were identical to those used for the  $T_1$  experiment apart from the following: relaxation intervals:  $\tau_{min} = 20$  ms,  $\tau_{max} = 564$  ms, number of relaxation points = 12.

2D DOSY spectra were acquired using the convection compensated pulse sequence (Bruker pulse program ledbpgp2s) at 300 K using the 800 MHz NMR spectrometer. The gradient strength of the pulse field gradients was increased from 5% to 95% of a set maximum of 53 G/cm over 16 points. The diffusion time was set to 100 ms. The following additional parameters were used: 2 dummy scans, 256 scans, 2 s relaxation delay, 40 ppm sweep width and 2 s acquisition time.

1D CSSF TOCSY<sup>73</sup> spectra were acquired using the Bruker pulse program selcssfdizs at 300 K using the 800 MHz NMR spectrometer. The carrier was positioned on a chemical shift of the protons of interest (H-1b/6b of mannitol, H-2 trehalose and H-1 amylopectin). The CSSF was set to eliminate signals at a distance of 12.5 Hz. A 40 ms Gaussian pulse was used. 12 increments were acquired with 2 scans each. A mixing time of 100 ms and 120 ms was used for mannitol/trehalose and amylopectin, respectively.

2D <sup>1</sup>H, <sup>13</sup>C HSQC spectra were acquired using the Bruker pulse program hsqcedetgpsisp2.3 and the following parameters: 2048 and 2048 complex points in  $t_2$  and  $t_1$ , respectively, spectral widths of 10 and 150 ppm in  $F_2$  and  $F_1$ , yielding  $t_2$  and  $t_1$  acquisition times of 128 and 34 ms, respectively. Sixteen scans were acquired for each  $t_1$  increment using a relaxation time of 1.5 s. The polarisation transfer was optimised for  $^1J_{CH} = 150$  Hz. Forward linear prediction to 4096 points was applied in  $F_1$ . A zero filling to 4096 was applied in  $F_2$ . A cosine squared window function was used for apodization prior to Fourier transformation in both dimensions. The overall acquisition time was 15 hours per spectrum.

**FT-ICR-MS.** 1 mg samples of HWE freeze dried tea extracts were prepared in 50:50 MeOH/H<sub>2</sub>O and diluted to a final concentration of 0.1 mg/ml. Analysis was performed on a 12 T SolariX (Bruker Daltonics) with an Infinity cell. Negative mode Electrospray Ionization was used and the sample was sprayed at a rate of 200 µl/hour. Ion accumulation time and time of flight, were set to 0.2 s and 0.6 ms, respectively. 200 transients at 4 MW were summed. Spectra of 3 unburied (in triplicate) and 28 buried green tea samples were collected and calibrated using the formulae of compounds present in green tea<sup>22,40</sup>. A signal to noise ratio > 5 and an error threshold of ±0.5

ppm were used to assign molecular formulae in Formularity<sup>74</sup> with the following elemental constraints <sup>12</sup>C0–80 <sup>1</sup>H0–160 <sup>16</sup>O0–40 <sup>14</sup>N0–8.

**Chemometric analysis of 1D 1H NMR and FT-ICR-MS data.** The chemometric analysis (PCA, PLS-DA, HCA) was performed using SIMCA 14.1. A <sup>1</sup>H NMR spectral region of 0.5–10 ppm, excluding the interval between 4.65–4.75 ppm that contains the residual HOD signal, was used. The bin width was 0.01 ppm and regional spectral alignment prior to integration was done using home written scripts based on icoshift<sup>75</sup>. Individual masses were used for the analysis of the MS data with sum total normalisation applied to each sample. Pareto scaling was used for both NMR and MS data.

### Data availability

The datasets generated during and analysed during the current study are available from the corresponding author.

Received: 12 April 2019; Accepted: 9 January 2020;

Published online: 30 January 2020

### References

- Schimel, D. S. Terrestrial ecosystems and the carbon-cycle. *Global Change Biology* **1**, 77–91, <https://doi.org/10.1111/j.1365-2486.1995.tb00008.x> (1995).
- Trofymow, J. A. *et al.* Rates of litter decomposition over 6 years in Canadian forests: influence of litter quality and climate. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere* **32**, 789–804, <https://doi.org/10.1139/x01-117> (2002).
- Briones, M. J. I., McNamara, N. P., Poskitt, J., Crow, S. E. & Ostle, N. J. Interactive biotic and abiotic regulators of soil carbon cycling: evidence from controlled climate experiments on peatland and boreal soils. *Global Change Biology* **20**, 2971–2982, <https://doi.org/10.1111/gcb.12585> (2014).
- Moore, T. R. Litter decomposition in a subarctic spruce-lichen woodland, eastern Canada. *Ecology* **65**, 299–308, <https://doi.org/10.2307/1939482> (1984).
- Wieder, R. K. & Lang, G. E. A Critique of the analytical methods used in examining decomposition data obtained from litter bags. *Ecology* **63**, 1636–1642 (1982).
- Djukic, I. *et al.* Early stage litter decomposition across biomes. *Science of the Total Environment* **628–629**, 1369–1394, <https://doi.org/10.1016/j.scitotenv.2018.01.012> (2018).
- Didion, M. *et al.* Towards Harmonizing Leaf Litter Decomposition Studies Using Standard Tea Bags—A Field Study and Model Application. *Forests* **7**, <https://doi.org/10.3390/f7080167> (2016).
- Keuskamp, J. A., Dingemans, B. J. J., Lehtinen, T., Sarneel, J. M. & Hefting, M. M. Tea Bag Index: a novel approach to collect uniform decomposition data across ecosystems. *Methods in Ecology and Evolution*. **4**, 1070–1075, <https://doi.org/10.1111/2041-210x.12097> (2013).
- Becker, J. N. & Kuzyakov, Y. Teatime on Mount Kilimanjaro: Assessing climate and land-use effects on litter decomposition and stabilization using the Tea Bag Index. *Land Degradation & Development* **29**, 2321–2329, <https://doi.org/10.1002/ldr.2982> (2018).
- Mikola, J. *et al.* Spatial variation and linkages of soil and vegetation in the Siberian Arctic tundra - coupling field observations with remote sensing data. *Biogeosciences* **15**, 2781–2801, <https://doi.org/10.5194/bg-15-2781-2018> (2018).
- MacDonald, E. *et al.* Using the Tea Bag Index to characterize decomposition rates in restored peatlands. *Boreal Environment Research* **23**, 221–235 (2018).
- Berg, B. & McClaugherty, C. Plant Litter. Decomposition, Humus Formation, Carbon Sequestration. (2014).
- Dungait, J. A. J., Hopkins, D. W., Gregory, A. S. & Whitmore, A. P. Soil organic matter turnover is governed by accessibility not recalcitrance. *Global Change Biology* **18**, 1781–1796, <https://doi.org/10.1111/j.1365-2486.2012.02665.x> (2012).
- Gregorich, E. G., Carter, M. R., Angers, D. A., Monreal, C. M. & Ellert, B. H. Towards a minimum data set to assess soil organic-matter quality in agricultural soils. *Canadian Journal of Soil Science* **74**, 367–385, <https://doi.org/10.4141/cjss94-051> (1994).
- Landgraf, D., Leinweber, P. & Makeschin, F. Cold and hot water-extractable organic matter as indicators of litter decomposition in forest soils. *Journal of Plant Nutrition and Soil Science* **169**, 76–82, <https://doi.org/10.1002/jpin.200521711> (2006).
- Joubert, E. & de Beer, D. Rooibos (*Aspalathus linearis*) beyond the farm gate: From herbal tea to potential phytopharmaceutical. *South African Journal of Botany* **77**, 869–886, <https://doi.org/10.1016/j.sajb.2011.07.004> (2011).
- Houben, D., Faucon, M. P. & Mercadal, A. M. Response of Organic Matter Decomposition to No-Tillage Adoption Evaluated by the Tea Bag Technique. *Soil Systems* **2**, <https://doi.org/10.3390/soilsystems2030042> (2018).
- Wahyuni, D. S. C., Kristanti, M. W., Putri, R. K. & Rinanto, Y. In International Conference on Science and Applied Science (Engineering and Educational Science) ((ICSAS). (2017).
- Lee, J. E. *et al.* H-1 NMR-based metabolomic characterization during green tea (*Camellia sinensis*) fermentation. *Food Research International* **44**, 597–604, <https://doi.org/10.1016/j.foodres.2010.12.004> (2011).
- Le Gall, G., Colquhoun, I. J. & Defernez, M. Metabolite profiling using H-1 NMR spectroscopy for quality assessment of green tea, *Camellia sinensis* (L.). *Journal of Agricultural and Food Chemistry* **52**, 692–700, <https://doi.org/10.1021/jf034828r> (2004).
- Graham, H. N. Green tea composition, consumption, and polyphenol chemistry. *Preventive Medicine* **21**, 334–350, [https://doi.org/10.1016/0091-7435\(92\)90041-f](https://doi.org/10.1016/0091-7435(92)90041-f) (1992).
- van der Hooft, J. J. *et al.* Structural Annotation and Elucidation of Conjugated Phenolic Compounds in Black, Green, and White Tea Extracts. *Journal of Agricultural and Food Chemistry* **60**, 8841–8850, <https://doi.org/10.1021/jf300297y> (2012).
- Nagle, D. G., Ferreira, D. & Zhou, Y. D. Epigallocatechin-3-gallate (EGCG): Chemical and biomedical perspectives. *Phytochemistry* **67**, 1849–1855, <https://doi.org/10.1016/j.phytochem.2006.06.020> (2006).
- Sparling, G., Vojvodic-Vukovic, M. & Schipper, L. A. Hot-water-soluble C as a simple measure of labile soil organic matter: the relationship with microbial biomass C. *Soil Biology & Biochemistry* **30**, 1469–1472, [https://doi.org/10.1016/s0038-0717\(98\)00040-6](https://doi.org/10.1016/s0038-0717(98)00040-6) (1998).
- Schulz, E. & Korschens, M. Characterization of the decomposable part of soil organic matter (SOM) and transformation processes by hot water extraction. *Eurasian Soil Science* **31**, 809–813 (1998).
- Chodak, M., Khanna, P. & Beese, F. Hot water extractable C and N in relation to microbiological properties of soils under beech forests. *Biology and Fertility of Soils* **39**, 123–130, <https://doi.org/10.1007/s00374-003-0688-0> (2003).
- Hertkorn, N. *et al.* spectroscopy and FTICR mass spectrometry: powerful discovery tools for the molecular level characterization of marine dissolved organic matter. *Biogeosciences* **10**, 1583–1624, <https://doi.org/10.5194/bg-10-1583-2013> (2013).
- Tfaily, M. M. *et al.* Investigating dissolved organic matter decomposition in northern peatlands using complementary analytical techniques. *Geochimica Et Cosmochimica Acta* **112**, 116–129, <https://doi.org/10.1016/j.gca.2013.03.002> (2013).
- Sharma, R., Gogna, N., Singh, H. & Dorai, K. Fast profiling of metabolite mixtures using chemometric analysis of a speeded-up 2D heteronuclear correlation NMR experiment. *Rsc Advances* **7**, 29860–29870, <https://doi.org/10.1039/c7ra04032f> (2017).

30. Yuan, Y. F. *et al.* Simultaneous determination of caffeine, gallic acid, theanine, (-)-epigallocatechin and (-)-epigallocatechin-3-gallate in green tea using quantitative H-1-NMR spectroscopy. *Analytical Methods* **6**, 907–914, <https://doi.org/10.1039/c3ay41369a> (2014).
31. Bell, N. G. A. & Sakas, J. Reduced dimensionality hyphenated NMR experiments for the structure determination of compounds in mixtures. *Faraday Discussions* <https://doi.org/10.1039/x0xx00000x> (2019).
32. Hansen, P. I. *et al.* Structure and Hydration of the Amylopectin Trisaccharide Building Blocks-Synthesis, NMR, and Molecular Dynamics. *Biopolymers* **89**, 1179–1193, <https://doi.org/10.1002/bip.21075> (2008).
33. Wang, H. J., Shi, S. S., Bao, B., Li, X. J. & Wang, S. C. Structure characterization of an arabinogalactan from green tea and its anti-diabetic effect. *Carbohydrate Polymers* **124**, 98–108, <https://doi.org/10.1016/j.carbpol.2015.01.070> (2015).
34. Krasensky, J. & Jonak, C. Drought, salt, and temperature stress-induced metabolic rearrangements and regulatory networks. *Journal of Experimental Botany* **63**, 1593–1608, <https://doi.org/10.1093/jxb/err460> (2012).
35. Tibbett, M., Sanders, F. E. & Cairney, J. W. G. Low-temperature-induced changes in trehalose, mannitol and arabitol associated with enhanced tolerance to freezing in ectomycorrhizal basidiomycetes (*Hebeloma* spp.). *Mycorrhiza* **12**, 249–255, <https://doi.org/10.1007/s00572-002-0183-8> (2002).
36. Lex, A., Gehlenborg, N., Strobel, H., Vuilleumot, R. & Pfister, H. UpSet: Visualization of Intersecting Sets. *Ieee Transactions on Visualization and Computer Graphics* **20**, 1983–1992, <https://doi.org/10.1109/tvcg.2014.2346248> (2014).
37. Koch, B. P. & Dittmar, T. From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter. *Rapid Communications in Mass Spectrometry* **20**, 926–932, <https://doi.org/10.1002/rcm.2386> (2006).
38. Bhuyani, L. P., Tamuly, P. & Mahanta, P. K. Lipid-content and fatty-acid composition of tea shoot and manufactured tea. *Journal of Agricultural and Food Chemistry* **39**, 1159–1162, <https://doi.org/10.1021/jf00006a034> (1991).
39. Chung, T. Y. *et al.* Analysis of lipophilic compounds of tea coated on the surface of clay teapots. *Journal of Food and Drug Analysis* **23**, 71–81, <https://doi.org/10.1016/j.jfda.2014.05.005> (2015).
40. Cadiz-Gurrea, M. D., Fernandez-Arroyo, S. & Segura-Carretero, A. Pine Bark and Green Tea Concentrated Extracts: Antioxidant Activity and Comprehensive Characterization of Bioactive Compounds by HPLC-ESI-QTOF-MS. *International Journal of Molecular Sciences* **15**, 20382–20402, <https://doi.org/10.3390/ijms151120382> (2014).
41. Jiang, H. Y., Engelhardt, U. H., Thrane, C., Maiwald, B. & Stark, J. Determination of flavonol glycosides in green tea, oolong tea and black tea by UHPLC compared to HPLC. *Food Chemistry* **183**, 30–35, <https://doi.org/10.1016/j.foodchem.2015.03.024> (2015).
42. Kuhnert, N., Drynan, J. W., Obuchowicz, J., Clifford, M. N. & Witt, M. Mass spectrometric characterization of black tea thearubigins leading to an oxidative cascade hypothesis for thearubigin formation. *Rapid Communications in Mass Spectrometry* **24**, 3387–3404, <https://doi.org/10.1002/rcm.4778> (2010).
43. Yassin, G. H., Koek, J. H., Jayaraman, S. & Kuhnert, N. Identification of Novel Homologous Series of Polyhydroxylated Theasinensins and Theanaphthoquinones in the SII Fraction of Black Tea Thearubigins Using ESI/HPLC Tandem Mass Spectrometry. *Journal of Agricultural and Food Chemistry* **62**, 9848–9859, <https://doi.org/10.1021/jf502220c> (2014).
44. Morton, J. T. *et al.* Uncovering the Horseshoe Effect in Microbial Analyses. *Msystems* **2**, <https://doi.org/10.1128/mSystems.00166-16> (2017).
45. Lee, L. S., Kim, S. H., Kim, Y. B. & Kim, Y. C. Quantitative Analysis of Major Constituents in Green Tea with Different Plucking Periods and Their Antioxidant Activity. *Molecules* **19**, 9173–9186, <https://doi.org/10.3390/molecules19079173> (2014).
46. Xin, Z. Q. *et al.* UPLC-Orbitrap-MS/MS combined with chemometrics establishes variations in chemical components in green tea from Yunnan and Hunan origins. *Food Chemistry* **266**, 534–544, <https://doi.org/10.1016/j.foodchem.2018.06.056> (2018).
47. Ohno, T., He, Z. Q., Sleighter, R. L., Honeycutt, C. W. & Hatcher, P. G. Ultrahigh Resolution Mass Spectrometry and Indicator Species Analysis to Identify Marker Components of Soil- and Plant Biomass-Derived Organic Matter Fractions. *Environmental Science & Technology* **44**, 8594–8600, <https://doi.org/10.1021/es101089t> (2010).
48. Guigue, J. *et al.* Ultrahigh-resolution FT-ICR mass spectrometry for molecular characterisation of pressurised hot water-extractable organic matter in soils. *Biogeochemistry* **128**, 307–326, <https://doi.org/10.1007/s10533-016-0209-5> (2016).
49. Duboc, O. *et al.* Decomposition of European beech and Black pine foliar litter along an Alpine elevation gradient: Mass loss and molecular characteristics. *Geoderma* **189**, 522–531, <https://doi.org/10.1016/j.geoderma.2012.06.018> (2012).
50. Geraei, D. S., Hojati, S., Landi, A. & Cano, A. F. Total and labile forms of soil organic carbon as affected by land use change in southwestern Iran. *Geoderma Regional* **7**, 29–37, <https://doi.org/10.1016/j.geodrs.2016.01.001> (2016).
51. Mirsky, S. B., Lanyon, L. E. & Needelman, B. A. Evaluating soil management using particulate and chemically labile soil organic matter fractions. *Soil Science Society of America Journal* **72**, 180–185, <https://doi.org/10.2136/sssaj2005.0279> (2008).
52. Lee, S. B. *et al.* Changes of soil organic carbon and its fractions in relation to soil physical properties in a long-term fertilized paddy. *Soil & Tillage Research* **104**, 227–232, <https://doi.org/10.1016/j.still.2009.02.007> (2009).
53. Meentemeyer, V. Macroclimate and lignin control of litter decomposition rates. *Ecology* **59**, 465–472, <https://doi.org/10.2307/1936576> (1978).
54. Bauer, J. *et al.* Temperature response of wheat decomposition is more complex than the common approaches of most multi-pool models. *Soil Biology & Biochemistry* **40**, 2780–2786, <https://doi.org/10.1016/j.soilbio.2008.07.024> (2008).
55. Sierra, C. A., Trumbore, S. E., Davidson, E. A., Vicca, S. & Janssens, I. Sensitivity of decomposition rates of soil organic matter with respect to simultaneous changes in temperature and moisture. *Journal of Advances in Modeling Earth Systems* **7**, 335–356, <https://doi.org/10.1002/2014ms000358> (2015).
56. Freeman, C., Ostle, N. & Kang, H. An enzymic 'latch' on a global carbon store - A shortage of oxygen locks up carbon in peatlands by restraining a single enzyme. *Nature* **409**, 149–149, <https://doi.org/10.1038/35051650> (2001).
57. Motavalli, P. P., Palm, C. A., Parton, W. J., Elliott, E. T. & Frey, S. D. Soil pH and organic C dynamics in tropical forest soils: Evidence from Laboratory and simulation studies. *Soil Biology & Biochemistry* **27**, 1589–1599, [https://doi.org/10.1016/0038-0717\(95\)00082-p](https://doi.org/10.1016/0038-0717(95)00082-p) (1995).
58. Sinsabaugh, R. L. *et al.* Stoichiometry of soil enzyme activity at global scale. *Ecology Letters* **11**, 1252–1264, <https://doi.org/10.1111/j.1461-0248.2008.01245.x> (2008).
59. Pind, A., Freeman, C. & Lock, M. A. Enzymatic degradation of phenolic materials in peatlands - measurement of phenol oxidase activity. *Plant and Soil* **159**, 227–231, <https://doi.org/10.1007/bf00009285> (1994).
60. Trasar-Cepeda, C., Leiros, M. C. & Gil-Sotres, F. Hydrolytic enzyme activities in agricultural and forest soils. Some implications for their use as indicators of soil quality. *Soil Biology & Biochemistry* **40**, 2146–2155, <https://doi.org/10.1016/j.soilbio.2008.03.015> (2008).
61. Sinsabaugh, R. L. Phenol oxidase, peroxidase and organic matter dynamics of soil. *Soil Biology & Biochemistry* **42**, 391–404, <https://doi.org/10.1016/j.soilbio.2009.10.014> (2010).
62. Zhang, Q. *et al.* Variations in carbon-decomposition enzyme activities respond differently to land use change in central China. *Land Degradation & Development* **30**, 459–469, <https://doi.org/10.1002/ldr.3240> (2019).
63. Cepakova, S. & Frouz, J. Changes in chemical composition of litter during decomposition: a review of published C-13 NMR spectra. *Journal of Soil Science and Plant Nutrition* **15**, 805–815 (2015).
64. Tanaka, T. & Kouno, I. Oxidation of tea catechins: Chemical structures and reaction mechanism. *Food Science and Technology Research* **9**, 128–133, <https://doi.org/10.3136/fstr.9.128> (2003).
65. Reina-Bueno, M. *et al.* Role of trehalose in heat and desiccation tolerance in the soil bacterium *Rhizobium etli*. *Bmc Microbiology* **12**, <https://doi.org/10.1186/1471-2180-12-207> (2012).

66. Crowe, J. H. Trehalose as a “chemical chaperone”: Fact and fantasy. *Molecular Aspects of the Stress Response: Chaperones, Membranes and Networks* **594**, 143–158 (2007).
67. Willers, C., van Rensburg, P. J. J. & Claassens, S. Phospholipid fatty acid profiling of microbial communities—a review of interpretations and recent applications. *Journal of Applied Microbiology* **119**, 1207–1218, <https://doi.org/10.1111/jam.12902> (2015).
68. Zelles, L. & Bai, Q. Y. Fatty-acid patterns of phospholipids and lipopolysaccharides in environmental-samples. *Chemosphere* **28**, 391–411, [https://doi.org/10.1016/0045-6535\(94\)90136-8](https://doi.org/10.1016/0045-6535(94)90136-8) (1994).
69. Ivanova, A. A., Kulichevskaya, I. S., Merkel, A. Y., Toshchakov, S. V. & Dedys, S. N. High Diversity of Planctomycetes in Soils of Two Lichen-Dominated Sub-Arctic Ecosystems of Northwestern Siberia. *Frontiers in Microbiology* **7**, <https://doi.org/10.3389/fmicb.2016.02065> (2016).
70. Moore, E. K. *et al.* Abundant Trimethylornithine Lipids and Specific Gene Sequences Are Indicative of Planctomycete Importance at the Oxidic/Anoxic Interface in Sphagnum-Dominated Northern Wetlands. *Applied and Environmental Microbiology* **81**, 6333–6344, <https://doi.org/10.1128/aem.00324-15> (2015).
71. Pinsonneault, A. J., Moore, T. R. & Roulet, N. T. Temperature the dominant control on the enzyme-latch across a range of temperate peatland types. *Soil Biology & Biochemistry* **97**, 121–130, <https://doi.org/10.1016/j.soilbio.2016.03.006> (2016).
72. Staff, S. S. o. S. (Macaulay Institute for Soil Research, Aberdeen, 1970–1987).
73. Robinson, P. T., Pham, T. N. & Uhrin, D. In phase selective excitation of overlapping multiplets by gradient-enhanced chemical shift selective filters. *Journal of Magnetic Resonance* **170**, 97–103, <https://doi.org/10.1016/j.jmr.2004.06.004> (2004).
74. Tolic, N. *et al.* Formularity: Software for Automated Formula Assignment of Natural and Other Organic Matter from Ultrahigh-Resolution Mass Spectra. *Analytical Chemistry* **89**, 12659–12665, <https://doi.org/10.1021/acs.analchem.7b03318> (2017).
75. Savorani, F., Tomasi, G. & Engelsen, S. B. icoshift: A versatile tool for the rapid alignment of 1D NMR spectra. *Journal of Magnetic Resonance* **202**, 190–202, <https://doi.org/10.1016/j.jmr.2009.11.012> (2010).
76. Kellerman, A. M., Dittmar, T., Kothawala, D. N. & Tranvik, L. J. Chemodiversity of dissolved organic matter in lakes driven by climate and hydrology. *Nature Communications* **5**, <https://doi.org/10.1038/ncomms4804> (2014).

## Acknowledgements

Authors would like to acknowledge L. MacKay, F. Cruikshank for their maintenance of the FT-ICR-MS facility and L. Murray and J. Bella for maintaining the NMR facility. We also acknowledge J. Marshall from SRUC and J. Warman from SWT for granting access to the studied sites. N. G. A. Bell would like to acknowledge the support of a NERC Soil Security Programme Fellowship (NE/N020227/1) to conduct this work.

## Author contributions

N.G.A.B. led the research design, data interpretation and manuscript preparation. A.S. led the FT-ICR-MS data collection, analysis and produced all the MS figures, Y.Z. contributed to the field work, 2D NMR data collection and analysis. Y.Z. also produced Figures 2 and 3. W.B., D.F., K.C., E.K. and Y.J. conducted the field work, tea bag sample preparation and 1D NMR data collection and analysis. All authors reviewed the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-58325-8>.

**Correspondence** and requests for materials should be addressed to N.G.A.B.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020