Towards understanding ecological disaster in the Harz Mountains (Central Germany) by carbon tracing: pyrolysis-GC-MS of biological tissues and their water-extractable organic matter (WEOM)

JOERI KAAL[1,2,3], CÉSAR PLAZA[2], MARTA PÉREZ RODRÍGUEZ[1], AND HARALD BIESTER[1]

[1] *Institut fur Geo-Ökologie, Abt. Umweltgeochemie, Technische Universität Braunschweig, Braunschweig, Germany*
[2] *Instituto de Ciencias Agrarias (ICA), Consejo Superior de Investigaciones Científicas (CSIC), Madrid, Spain*
[3] *Pyrolyscience, Madrid, Spain*

Compiled March 3, 2020

Streams and reservoirs in the Harz National Park experience high dissolved organic matter (DOM) concentrations, the cause of which is unknown. We studied potential sources of DOM by means of pyrolysis-GC-MS (Py-GC-MS). The biological materials include vegetation samples (spruce, birch, blueberry, heather, sedge, grass, peat moss, epiphytic moss), microbial sources (epilithic biofilm, lichen, fungi) and excremental fabric. In addition to ground biological samples (bulk organic matter; BOM), their leachates (water extraction of BOM samples followed by filtration; WEOM) were analyzed, to obtain knowledge on solid-liquid transfer effects. Results of BOM showed potential for molecular provenancing on the basis of relative proportions of major biopolymers (lignin, polysaccharides, tannin, resin, etc.) and specific products (biomarkers). Even though WEOM had lower molecular diversity than BOM, identified molecular features of WEOM can be useful to identify potential sources of environmental DOM. Unsubstituted guaiacol prevails among the pyrolysis products of the WEOM of spruce samples (especially wood), which probably originates from pyrolytically decarboxylated vanillic acid moieties. The other plant materials produced a more diverse fingerprint of phenols, guaiacols and syringols, which can be used to distinguish gymnosperm-derived DOM, peatland vegetation and angiosperm trees and shrubs. Several other O-containing moieties such as benzoic acids, benzaldehydes and indanones (not present in BOM) were also frequent products of WEOM. These results lay the foundation for future interpretations of environmental DOM samples' molecular fingerprints.

Keywords *Harz Mountains; dissolved organic matter; Py-GC-MS; biological sources.*

1. INTRODUCTION

The Harz National Park in Central Germany is suffering major ecological alterations which are best reflected by two phenomena. Firstly, the massive deterioration of the Norway spruce (*Picea abies*) stands that dominate the forest biome. It is widely accepted that this is due mainly to the damage caused by the bark beetle (*Ips typographus*), which attacks and eventually kills trees that are in a physiologically weak state (Overbeck and Schmidt, 2012). Heat waves reduce the fitness of Norway spruce and increase its vulnerability to bark beetles (Arnberger et al., 2018), linking the forest degradation to global warming. Secondly, there has been a strong increase in the amount of dissolved organic matter (DOM) that is being transported by streams and rivers into the large number of freshwater reservoirs in the Harz (Broder and Biester 2015; Broder et al., 2017), giving rise to profound brown/red colours of the surface waters. These reservoirs are used as a source of drinking water. In the water treatment pipeline, DOM and organically bound pollutants require drastic use of chemicals, which also create environmentally hazardous substances such as chlorine disinfection by-products including trihalomethanes (Matilainen et al., 2011).

The cause of the increase of DOM and associated pollutants fluxes has not been determined yet. It is commonly thought that enhanced decay of the mountain peatlands is the main source. The underlying mechanism could be climate change (warmer conditions) accelerating peat decay by lowering the water table, or a stimulated peat decay rate due to the general cessation of acid rain impacts since the late 1990s (hence, re-equilibration after environmental deterioration) (Vogt, 2003). However, enhanced DOM release from forest soils due to either decomposing spruce debris or enhanced microbial alteration of soil organic matter (SOM), possibly related to climate change as well, can also play a significant role (Kaal et al., 2017).

The project **DOMtrace** (DFG 419258863), of which this is the first contribution, uses the DOM in the streams and reservoirs of the Harz Mountains to identify what ecological factor is stimulating DOM release and which habitat is subjected to the strongest environmental change in terms of re-equilibrating

OM dynamics (Amon et al., 2012). There are numerous biological (vascular plants, phytoplankton, zooplankton, bacteria, fungi, arthropods, etc.) and thus biomolecular (polysaccharides, protein, lipids, lignin, chitin, peptidoglycan, terpenoids, etc.) precursors of freshwater DOM and many processes involved in its formation and alteration. As a consequence, DOM is one of the most complex and challenging types of natural organic substances (Hedges et al., 2000). Due to this complexity, a detailed analyses of possible sources of DOM is needed to create a reference molecular database to which the stream and reservoir DOM can be compared eventually.

Pyrolysis techniques such as Py-GC-MS are valuable tools for rapid screening of organic matter in environmental samples (Derenne and Quénéa, 2015; Kögel-Knabner and Rumpel, 2018). For the present study, a series of biological materials were selected on the basis of their hypothetical role in the catchment. Hence, we selected various materials of dominant tree species in the forests (Norway spruce and birch), blueberry (an important member of understorey vegetation in both spruce and birch stands), heather (important member of the peatland vegetation), peat moss (*Sphagnum*), herbaceous material from the floodplain of the Ecker stream and a random sedge sample. Furthermore, samples from an epiphytic moss, lichen, fungi (fruit body), epilithic biofilm (from the Ecker stream), green algae and deer excrement were considered as representatives of several other sources.

Besides the examination of the bulk biological materials (BOM), we study the dissolved fraction of water-extractable organic matter (i.e., WEOM, filtered leachates). The comparison of the molecular composition of BOM and WEOM fractions, by means of Py-GC-MS, aims to create the necessary framework to identifying the sources of DOM in the environment (catchments, streams and reservoir), and estimating their relative contributions. This is the first time that a detailed ecosystem-targeted fingerprinting of potential sources has been performed using BOM and WEOM as reference materials to eventually study environmental DOM.

2. MATERIALS AND METHODS

A. Sample selection

Samples were collected during several DOM sampling campaigns in the Ecker catchment in 2018 (Table 1). Due to the possible importance of Norway spruce, not only in terms of plant cover, but also in terms of litter production (vast amounts of decaying dead spruce logs), the biological materials of this species was studied in more detail than the other sources, separating different plant organs (wood, needles, bark and cones). Furthermore, two samples of dead spruce trunks were taken, both of which had a reddish appearance and the wood could be easily crumbled, indicative of a profound effect of decomposition. The managers of the National Park leave felled logs of spruce trees, presumably killed by the bark beetle, on site for decay. Of the spruce logs, a qualitative degree of alteration was based on the fragility of the material (dead log 2 is more decomposed in that sense; Table 1). The other higher plant sources were sampled by mixing wood/stem, leaf/needle and bark materials. Some microbial sources were included as well, most of which have not been identified precisely. As for the Poaceae, Cyperaceae and moss epiphyte samples, the uncertainty in source is acceptable because the fingerprints are assumed to be only roughly representative of the order or class organism they belong to. The examination of useful molecular features from the pyrolyzates does not solely rely on these pyrolyzates, but also on Py-GC-MS literature where multiple plant types were compared (e.g., Schellekens et al., 2012).

Samples were dried at 50 °C for ten days and shredded to a homogenous and manageable sample; the goal is to obtain a finely divided sample and increase its surface area in contact with the aqueous phase during the leaching process.

B. Production and isolation of WEOM

Water-extractable OM was obtained by creating a suspension of ground biological materials, i.e. 0.4 to 0.8 g dry BOM [with the exception of the epilithic biofilm and green algae samples (0.05 g) due to limited sample availability] and distilled water (25 ml), in 50 ml polyethylene tubes. The suspension was maintained horizontally on a linear shaking device for 2 hrs, followed by 10 min of sonication, aiming to disrupt the fabrics and enhance the release of organic matter. This sequence of homogenization/sonication was performed three times, after which the suspensions were left overnight on the shaking instrument. The next day, the suspensions were centrifuged (3000 rpm, 10 min) filtered through 0.45 µm cellulose nitrate membrane filters (by means of vacuum), isolating the WEOM. Two ml of the resultant WEOM solutions were used to measure total organic carbon content (data not available yet). The remaining solution was evaporated at 40 °C. For clarity, the sonication-stimulated WEOM is thought to mimic the abiotically produced leachate DOM in the natural environment. Given that it is performed in plastic material, the sonication step is expected to have very little effect on WEOM release but it might have stimulated some cell lysis as it caused a mild change in the colour of some of the suspensions.

The experimental procedure for generating leachates of biological materials is very diverse, including differences in temperature (hot-water extraction is more efficient), salt used to enhance the extraction efficiency (here avoided because salts will affect decomposition pathways during the pyrolytic reactions) and filter size (e.g., Zsolnay, 1996; Chantigny, 2003). Hence, any WEOM is operationally defined and our procedure aimed to have as little effect on molecular composition as possible. The WEOM samples have also been incubated to examine microbial decay effects (*cf.* Moore and Dalva, 2001), but those attempts so far produced insufficient DOM (suffice it to mention that only the residue of sedge materials generated enough material for Py-GC-MS analysis, and that the resultant chromatogram resembled that of the WEOM leachate).

C. Pyrolysis-GC-MS

Pyrolysis-GC-MS was performed using the same method as that used by Kaal et al. (2017) for the study of DOM in the Odersprung system. For BOM, approximately 1 mg of dried material was inserted into quartz wool-containing quartz tubes. The WEOM was redissolved by addition of 1 ml of H_2O and a plug of pre-combusted quartz wool. Then, the solvent was evaporated again to concentrate the WEOM on the quartz wool, which was then inserted into quartz tubes. The BOM and WEOM were analyzed using a pyrolyzer from CDS (Pyroprobe 5000) that was coupled to an Agilent 6890 gas chromatograph, which in turn was interfaced with a 5975 mass selective detector from Agilent. The pyrolysis set-point temperature was 650 °C. The gas chromatograph was equipped with a HP-5MS (non-polar) column. For the WEOM, sample concentration efficiency on the quartz wool could not be controlled, giving rise to large differences in signal intensity of the pyrolysis chromatograms. For some

Table 1. List of biological sources and reference to figures of corresponding chromatograms.

Type	Source	Plant organ	Lab code	Figure
Plant	*Picea*	spruce wood (living)	H0480a	1
Plant	*Picea*	spruce wood (log 1)	H0479	1
Plant	*Picea*	spruce wood (log 2)	H0484	1
Plant	*Picea*	spruce bark	H0480c	2
Plant	*Picea*	spruce needles	H0480b	2
Plant	*Picea*	spruce cone	H0480d	2
Plant	*Betula*	birch mixed litter	H0485	3
Plant	*Vaccinium*	blueberry mixed litter	H0476	3
Plant	*Calluna*	heather mixed litter	H1506	3
Plant	Poaceae	grass mixed litter	H0477	4
Plant	Cyperaceae	sedge mixed litter	H0487	4
Plant	*Sphagnum*	peat moss mixed litter	H1507	5
Plant	Epiphytic moss	mixed litter	H0478	5
Microbial	Lichen	-	H0481	5
Microbial	Excrement (deer)	-	H0483	4
Microbial	Mushroom	-	H0482	6
Microbial	Epilithic biofilm	-	H0486	6
Microbial	Green algae	-	H1511	6

samples of which poor data was obtained, the quartz wool was rewetted and rubbed against the bottom of the centrifuge tube, and re-analyzed.

D. Data evaluation

Pyrolysis-GC-MS chromatograms of the BOM and WEOM were evaluated qualitatively. A list of 150 products was obtained by denoting the main peaks in each of the chromatograms of the BOM samples (Table 2, see end of document). The pyrolysis product numbers in Table 2 are used to label the peaks in the chromatograms of both BOM and WEOM. For several of the WEOM samples signal intensities were too low and those chromatograms were discarded. Finally, a qualitative comparison was made between the results of environmental DOM from the Odersprung stream (Broder and Biester, 2015; Kaal et al., 2017), located nearby the Ecker stream, and the results from the BOM and WEOM from the potential source materials in the Ecker catchment.

3. RESULTS AND DISCUSSION

A. Vascular plant materials

A.1. Norway spruce materials

For spruce, the BOM samples of xylem from living spruce wood produced a Py-GC-MS fingerprint typical of gymnosperm wood in general, i.e. dominated by the products of polysaccharides (acetic acid, cyclopentenones, furans, furaldehydes, pyrans and anhydrosugars) and of lignin of the guaiacyl (G) type (guaiacol and various guaiacols with diverse substitutions on the C1 carbon atom) (Figure 1). Furthermore, multiple products of abietane diterpenes were identified, such as dehydroabietic acid (DHA) derivatives and retene. Guaiacyl dimer products could also be identified, such as the often observed m/z 272 and m/z 298 compounds (Evans et al., 1986; van der Heyden and Boon, 1994; Dijkstra et al., 1998; Dufour et al., 2013). The link between the guaiacyl groups of these dimers can be of B-*O*-4, phenylcoumarane or diarylpropane nature (Evans et al., 1986; Kuroda et al., 2010). Guggenberger and Zech (1994) also found these compounds using Py-FIMS of spruce forest litter. These compounds have not been identified in the pyrolyzates of the other lignin-containing biological materials (see below), or are present in much lower abundances (e.g., birch), suggesting that the molecular conformation of spruce lignin favours the production of dimers during pyrolysis. These products may be useful to evaluate the proportion of spruce-derived guaiacyl groups to total guaiacyl groups, and to determine the alteration state of this lignin fraction.

The WEOM from the branch wood of Norway spruce (Figure 1) reveals a remarkable dominance of guaiacyl lignin derivatives with no side-chain substitution, i.e. guaiacol. This was also observed for environmental DOM from the Odersprung (Kaal et al., 2017), and in that case the THM-GC-MS data ("pyrolysis-derivatization" using tetramethylammonium hydroxide (TMAH)) showed that the 3,4-dimethoxybenzoic acid methyl ester (G6) prevailed. This compound G6 is the methylated (by TMAH) product of vanillic acid (guaiacol with a carboxylic acid group at C1 position). Hence, it is concluded that the Py-GC-MS product guaiacol formed mainly upon decarboxylation of vanillic acid. The minor peaks with a guaiacyl group showed another interesting difference with the G-products of the

BOM of the same material. These peaks correspond to vanillin (4-formylguaiacol), vanillic acid, vanillic acid methyl ester, 4-acetylguaiacol, 4-propan-2-one-guaiacol and an unidentified product with m/z 151 (base peak), 123, 108 and probably 194 (M^+), which could be guaiacol with a $C_3H_5O_2$ group (hence, coniferyl alcohol with an additional ketone group). The predominance of guaiacol and detection of compounds like vanillic acid methyl ester, suggest that the guaiacol in the WEOM is largely deacylated vanillic acid that was esterified to the lignin backbone. Indeed, absence of G units with propanoid groups (which are formed upon pyrolysis of B-*O*-4 bonds) is in agreement with a negligible contribution of intact macromolecular lignin to the WEOM samples. It seems likely that the production of the leachates favoured the hydrolysis of acylated vanillic acid rather than mobilization (depolymerization) of the lignin itself.

Besides G-based products, the WEOM showed several peaks of carbohydrates, in particular 3/2-furaldehyde and 5-methyl-2-furaldehyde and perhaps benzofuran is a product of carbohydrates as well. Phenol was also a significant peak. Another product gave m/z 115, 145 and 160 and could correspond to a dimethylindanone or isopropenylacetophenone (i.e. acetylstyrene), the former option of which is considered more likely due to detection of other indanones. This set of compounds shows the pyrolysis fingerprint of spruce lignin-derived WEOM, which is clearly very different from that of the pyrolyzate of the POM, in great detail. It is concluded that signal of spruce lignin is probably relatively easily identified in DOM on the basis of the guaiacol dominance and vanillic acid products. The release of G6 upon (abiotic) leaching due to deacetylation may also explain why DOM samples tend to have such high G6/G4 ratios (e.g. van Heemst et al., 2000; Jiang et al., 2017). This proxy may therefore not reflect the degree of biological alteration of lignin (the WEOM is assumed to be unaffected by biological decay). Lignin dimers and resin products were below the detection limit. Both the BOM and the WEOM are virtually devoid of long-chain aliphatic products and nitrogen moieties.

The chromatograms of the BOM samples from the decomposing spruce logs clearly show the loss of polysaccharide signals and the relative enrichment of the peaks from guaiacyl lignin (Figure 1), when compared to the results of the extant wood (*cf.* Filley et al., 2002). In addition, products of oxidized lignin seemed to increase relative to the other G products in the decomposed wood samples, especially vanillin but also 4-acetylguaiacol (Guggenberger and Zech, 1994). However, contrary to the latter study, a decrease in coniferyl alcohol was not apparent, and in general terms the chromatograms of the BOM samples of the logs were quite similar to that obtained from living wood. The chromatograms of the WEOM of the decomposing wood samples were of poor quality, but as was observed for fresh branch wood, the fingerprints were dominated by unsubstituted guaiacol. The poor quality of the chromatograms could be an indication of (1) the imperfection of the WEOM sample introduction method and thus need for use of larger amounts of material for the isolation of WEOM, and/or (2) that early decomposition eliminates preferentially those lignin groups that tend to be mobilized into the WEOM fraction so that the decomposing logs release less WEOM than fresh wood.

The pyrolyzate of the BOM sample of bark material (Figure 2) of living Norway spruce shows the same dominance of G-lignin products as observed for the wood samples, and large peaks for multiple products of polysaccharides including levoglucosan. The peaks of fatty acids are higher for bark than for wood, and not limited to the omnipresent C_{16} and C_{18} fatty acids but also, and even in higher abundances, of C_{20}, C_{22} and C_{24} fatty acid. This signal of long-chain fatty acids can be ascribed to the high suberin content of bark materials, which act as a protective barrier (Kolattukudy, 2001). The peak for catechol is considerably higher, probably due to the condensed tannin. Diterpene products (DHAs) and an unidentified product with m/z 396 and 381 are abundant, whereas lignin dimers are scarce. The WEOM of the bark material is dominated by unsubstituted guaiacol, as was observed for the wood, indicating that lignin derivatives prevail in the WEOM. Other peaks were phenol, multiple carbohydrate products (including pyrans), 4-methylguaiacol, 4-ethylguaiacol, 4-vinylguaiacol, 4-methylbenzoic acid (a rare product in pyrolyzates, with m/z 91, 119, 136), the aforementioned tentatively identified dimethylindanone, all of which probably represent the lignin/polysaccharide complex. The WEOM of bark material was prolific of a homologous series of alkenes, of C_{16} fatty acid, of alkylnaphthalenes (C_3 and C_5), retene and several abietane products, which originate from the suberin and resin constituents. Catechol, possibly from tannin, was also identified and clearly more abundant than in the pyrolyzate of the WEOM from the wood materials. It is concluded that the spruce bark would contribute a more diverse set of G-type products from lignin-derived WEOM and a signature of tannin, resin and suberin.

The importance of spruce needles as a potential source of DOM is evidenced by the loss of needles during recent summer heat waves (e.g., 2019), when the spruce forest lost massive amounts of leaf materials (green needles that would not be expelled in that time of the year under normal conditions). The Py-GC-MS chromatogram (Figure 2) shows G lignin products, relatively intense signal of fatty acids but only C_{16} and C_{18} (including unsaturated fatty acids), a stigmasterol, and catechol. Clearly, needle materials are composed mainly of polysaccharides, G lignin, tannin and fatty acids. The fatty acid pattern probably indicates cutin as a partial source. The abundance of nitrogen-containing compounds is higher in the chromatograms of needles than in those from other anatomical parts of spruce. The WEOM of spruce needles has a significant peak for guaiacol, but it is overshadowed by catechol. This probably indicates that the WEOM from the needles is composed primarily of derivatives of condensed tannin. Other peaks were 2-methylfuran, toluene, indanone (m/z 104 and 132), 4-vinylguaiacol, 4-propan-2-one-guaiacol, 4-hydroxyacetophenone (m/z 121, 136) and homovanillic acid (m/z 137 and 182). Products of resins and fatty acids/alkenes were not detected. Hence, the needles fingerprint is recognized from mainly catechol due to the high tannin content, but other phenolic products are indicative of a patterns of G-based products other than vanillic acid derivatives.

Finally, the BOM of spruce cones shows elevated proportions of condensed tannin products (catechol and methylcatechols upon Py-GC-MS; Figure 2). Nevertheless, lignocellulose is also clearly present. The peak of DHA (abietane resin) is larger than in any other sample. The abundance of diterpene resin is also the main feature of the Py-GC-MS fingerprint of the WEOM from the cone materials, dominated by retene and di- and tetrahydroretenes, but with numerous other aromatic diterpene products (m/z combinations 235/310, 237/312, 251/266, 253/268, 232/202). Besides the resin signal, phenolic products are mainly guaiacol and derivatives (4-vinylguaiacol), isopropenylacetophenone/dimethylindanone, and importantly, 4-vinylphenol. This latter compound is a decarboxylation product of *p*-coumaric acid (and perhaps also of phloretic acid), which is abundant in herbaceous lignin structures but also in sporopollenin (Wehling

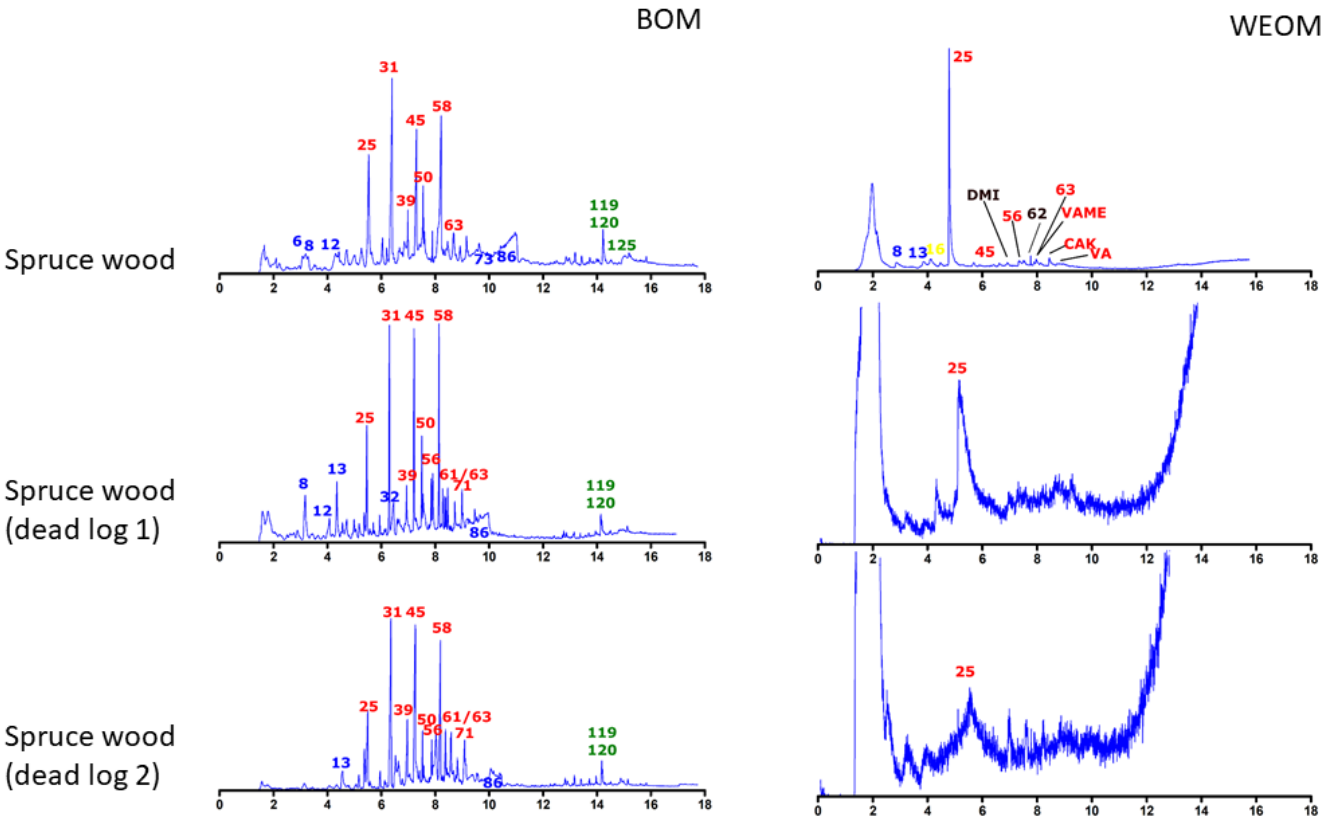

Fig. 1. Pyrolysis chromatograms (pyrograms) of BOM and WEOM samples. Numbers refer to peak labels in Table 2, which was based on the BOM samples. For WEOM products that were not identified in the BOM as well, the following abbreviations were used: 4HAP=4-hydroxyacetophenone (*m*/*z* 121/136); 4MBA=4-methylbenzoic acid (91/136); a= alkane fragment; ab312= unidentified abietane compound (237/312)(2 isomers); aP= C_2-C_3 alkylphenols; BD= benzendiol (hydroquinone, resorcinol); CAK= coniferyl alcohol (keton substitution); dhRet= dihydroretene (221/236); DMI=dimethylindanone (or 4-isopropenylacetophenone) (115/145/160); e= alkene fragment; HVA= homovanillic acid (137/182); Indanone=indanone; m180= unidentified product with M^+ 180 (tentatively identified as phenazine); MB= methylbenzaldehyde (91/119/120); pQ= *p*-quinone (54/82/108); pypy= pyridine/pyrrole; thRet= tetrahydroretene (223/238); un220= unidentified compound (91/150/177/220), VA= vanillic acid, VAME= vanillic acid methyl ester. Colours of labour help in the identification of pyrolysis product groups, i.e. guaiacyl (red), syringyl (orange), phenols (yellow), long-chain aliphatic products (alkanes, alkenes, fatty acids; magenta), carbohydrate (blue), diagnostic terpenoids (green), other compounds (black).

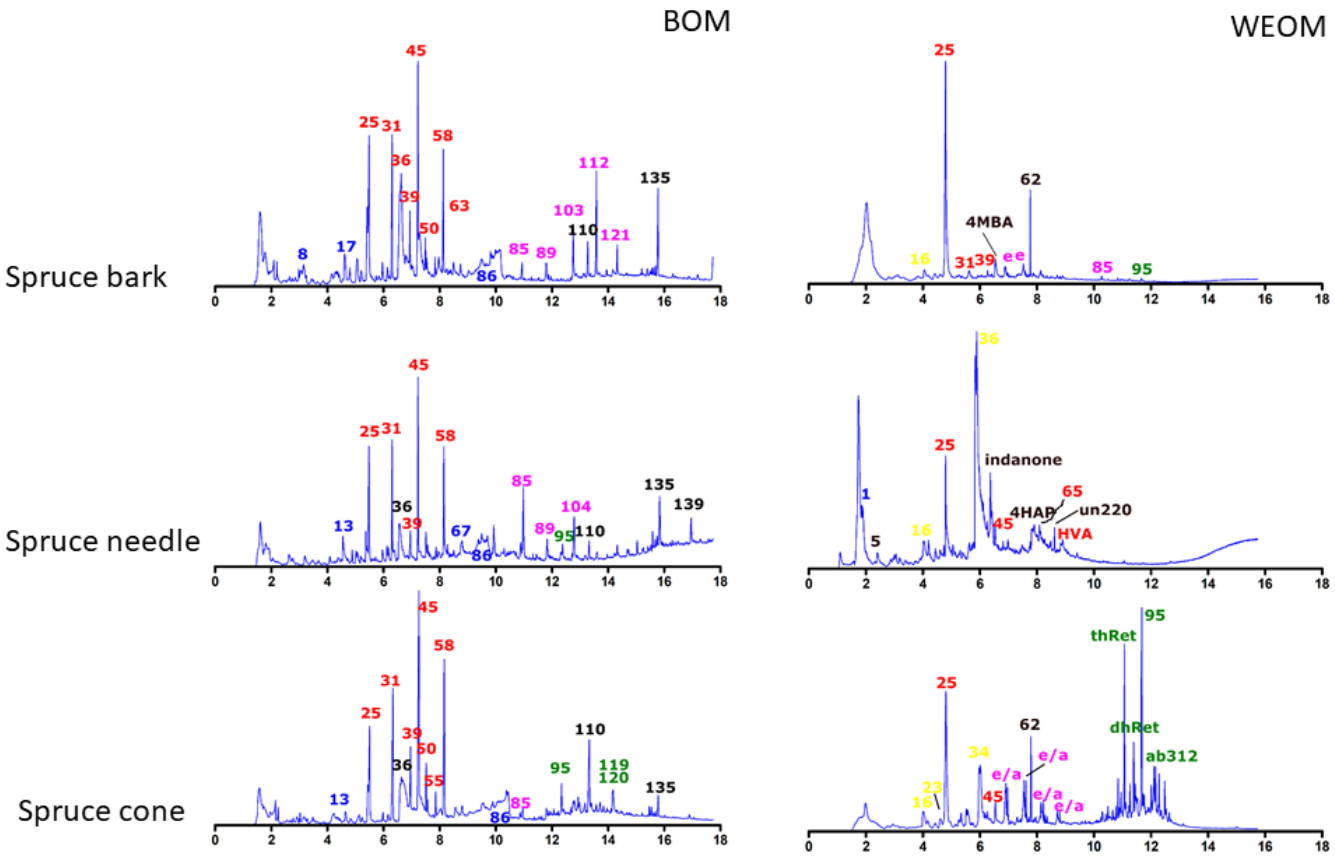

Fig. 2. Pyrolysis chromatograms (pyrograms) of BOM and WEOM samples. For explanations of labels see Figure 1 caption.

et al., 1989; Montgomery et al., 2016). Alkanes and alkenes (short-chain, i.e. <C_{18}) are also conspicuous (possible also from sporopollenin; de Leeuw et al., 2006) and phenols (including C_1- and C_2-alkylphenols). Carbohydrate products are relatively scarce but clearly distinguishable. The cone material is enriched in resin and sporopollenin (from pollen), the former of which seems to be efficiently transferred to the WEOM fraction, but lignin and polysaccharides can be clearly recognized. Sporopollenin is extremely hydrophobic and recalcitrant and its transfer to WEOM is not expected (spruce pollen cannot pass the membrane filter) but from the available data it cannot be discarded.

It is concluded that there are plenty of differences to distinguish the different parts of spruce debris on the basis of relative abundances of lignin, resin, suberin, cutin, tannin and sporopollenin. The spruce debris in general can be identified from the guaiacol that dominates the Py-GC-MS chromatograms. The wood has exceptional dominance of G lignin derivatives including vanillic acid-based products, the bark WEOM shows the fingerprint of some resin and tannin but with dominance of G-type lignin, the needles are marked by the release of catechol and several phenols that are not abundant in the wood-derived WEOM, and the cones show the fingerprints of sporopollenin (from pollen) and of the diterpene resin. It is remarkable that the vast majority of the typical G-lignin products of gymnosperm wood, including the often dominant $C_{3:1}$ (propylene) guaiacols are barely detectable in the WEOM. This implies that their absence in environmental DOM (Kaal et al., 2017) is not a result of biological decomposition, and neither of mineral-catalyzed decomposition during pyrolysis, but basically of the lack of unaltered macromolecular lignin in the leachates. Note that the tannin fingerprint of the living materials is strongly concentrated on catechol, which is also a product of demethylated G lignin products in spruce litter that is affected by fungal decay (Filley et al., 2002). This implies that for environmental DOM samples, demethylated lignin cannot be distinguished from condensed tannin when Py-GC-MS is used.

A.2. Other vascular plants

The relatively high abundance of both fatty acids –including unsaturated $C_{18:1}$, $C_{18:2}$, $C_{22:1}$ and probably $C_{22:2}$ fatty acid– and triterpenoids in the BOM sample of birch mixed litter, is indicative of a high aliphatic content, predominantly from the bark and/or leaf organs rather than xylem (Figure 3). The lignin is recognized from the simultaneous presence of guaiacyl (G) and syringyl (S) compounds, the latter of which are represented primarily by 4-methylsyringol, 4-vinylsyringol and *trans* $C_{3:1}$-syringol (Figure 3). Polysaccharides are well-represented by levoglucosan. Several high molecular weight products produced intense peaks, i.e. alpha-tocopherol (M^+ 430) and unidentified compounds with characteristic m/z 163, 190, 394 and with m/z 409 and 424, the latter of which probably corresponding to pentacyclic triterpenoids such as lupenones and betulones. Tannins are possibly reflected by catechol and methylcatechols. The WEOM pyrolyzate is prolific of carbohydrate products (acetic acid, 3/2-furaldehyde, 1,4:3,6-dianhydro-a-D-glucose, 2,3-dimethylcyclopent-2-en-1-one), guaiacol, syringol (both unsubstituted), alkylphenols, dimethylindanone and alkenes. Besides

the typical woody angiosperm fingerprint of the lignin, birch litter may be recognized in DOM from the unidentified high molecular weight compounds and perhaps the relatively high proportion of 1,4:3,6-dianhydro-a-D-glucose. However, it may be antipicated that this fingerprint will not be easily identified in environmental DOM pyrolyzates unless birch-derived DOM is a large proportion of total DOM.

The sample of blueberry (*Vaccinium*) was a mixture of branch, shoot and buds, and the molecular fingerprints are indicative of a diverse mixture of biopolymeric constituents (Figure 3). The lignin products dominate the Py-GC-MS chromatogram, and include many G- and S-type moieties, especially of guaiacol, syringol and the 4-methyl-, 4-vinyl- and $C_{3:1}$-(propenyl-)substituted analogues, reflecting the angiosperm wood component. Homosyringic acid was identified as well. This sample produced significant peaks of catechol, methylcatechols and dimethyl/ethylcatechol, which probably originate from tannins. Carbohydrate products are also abundant, including levoglucosan. Minor constituents were fatty acids (C_{16}, C_{18}), pentacyclic triterpenoids (probably amyrins; Neto, 2011) and a homologous series of *n*-alkanes, maximizing at C_{29}. Proteins or alkaloids are represented by diketodipyrrole and other N-containing compounds. The WEOM pyrolyzate has large contributions of furan, 2-methylfuran, 2,5-dimethylfuran and 5-methyl-2-furaldehyde, representing carbohydrates. Lignin derivatives are represented by guaiacol and syringol, the latter of which is more abundant (largest peak of the chromatogram) and suggests that the S lignin (or acylated S groups) in blueberry debris is more easily released to the WEOM fraction than the S lignin in birch materials. Hence, similar to observation for G products in spruce lignin, the diversity of S products is also much lower for WEOM than for BOM. 4-Vinylphenol also produced a significant peak. Catechol and methylcatechols may be indicative of a significant tannin content, but the peak for 1,3- or 1,4-benzenediol (hydroquinone or resorcinol, i.e. an isomer of catechol) may point towards a non-tannin source of these compounds. In short, the blueberry fingerprint could be recognized on the basis of angiosperm wood lignin, with a relatively high proportion of S groups, and several phenolic products that are scarce or absent in the other higher plant pyrolyzates.

The signal of G and S moieties from lignin, and polysaccharides, dominate the chromatogram of BOM from peatland heather (Figure 3). Of the BOM samples of the vascular plant sources, the similarities are strongest for the samples of the Ericaceae, i.e. *Vaccinium myrtillus* and *Calluna vulgaris*. In terms of polyphenols, the sample of heather is less prolific of catechols and a methoxycatechol, but such minor differences can easily be explained by differences in the balance between the different anatomical parts of the sampled plant and can therefore not be used for distinguishing the materials. The aliphatic products of heather included large peaks for a series of *n*-2-methylketones (C_{25}-C_{29}, with a clear maximum at C_{27}) and the *n*-alkanes are accompanied by significant peaks of *n*-alkenes and *n*-alkanols. Heather BOM was also prolific of a triterpenoid compound with m/z 95, 109 and 410. For the WEOM, 3/2- and 5-methyl-2-furaldehyde are the principal carbohydrate products. The peak of *p*-quinone (m/z 54, 82, 80, 108) may be a particular feature of heather-derived DOM. Phenol and benzoic acid are among the prominent peaks, indicative of H lignin structures or acylated H groups, probably. Guaiacol and 4-vinylguaiacol were also identified and a large broad peak of catechol, methylcatechols, indanone and 4-vinylphenol was detected. Syringyl structures were not identified, and, again, the aliphatic moieties that are abundant in the pyrolyzates of the BOM are scarce in those of the WEOM. As such, the WEOM of heather will probably cause enrichment of H lignin products, *p*-quinone and perhaps carbohydrate products when a significant part of DOM originates from it.

Comparing the results of the three WEOM samples obtained from angiosperm shrubs and trees, the birch fingerprint is enriched in alkenes, that of blueberry in syringol and hydroquinone, and that of heather of phenol, *p*-quinone and benzoic acid. Interestingly, the three considered sources have different carbohydrate patterns (furaldehydes for birch and heather, 1,4:3,6-dianhydro-a-D-glucose for birch, furans for blueberry).

The Poaceae are represented by a riverbank grass from the Ecker stream (near its spring). The pyrolysis fingerprint (Figure 4) is dominated by 4-vinylguaiacol, which originates from ferulic acid. 4-Vinylphenol, from *p*-coumaric acid, was also abundant. These compounds reflect the herbaceous lignin and lignin-like phenolics (glycosylated cinnamyl compounds) typical of grasses. Polysaccharides are also well-represented in the chromatogram. The results of the BOM serve as an example of the easy recognition of herbaceous lignin-derived materials. The abundance of fatty acids, triterpenoids and tannin are very small in this kind of samples. The WEOM has significant peaks for pyrrole and pyridine, suggesting that the grass-derived WEOM is relatively protein-rich. Alkylbenzenes (toluene, C_2-alkylbenzenes) and phenols (phenol, alkylphenols) generated significant peaks that cannot be assigned to a specific biopolymer source but a source in protein cannot be discarded, and alkanes, alkenes and fatty acids were also identified. Guaiacol, 4-ethylguaiacol, 4-vinylguaiacol, syringol and 4-vinylsyringol reflect angiosperm lignin. 4-Vinylphenol was also present but not as dominant as in the analogous BOM.

The BOM sample of a Cyperaceae stalk material (*Carex* ssp.) was also prolific of 4-vinylphenol and 4-vinylguaiacol (Figure 4). The polysaccharide products are relatively scarce. Significant peaks are also found for catechol, C_{16}-fatty acid and a triterpenoid. WEOM produced a chromatogram of minor quality, in which toluene, phenol, alkylphenols, 2,3-dimethylcyclopent-2-en-1-one, guaiacol, 4-vinylguaiacol, syringol and indole were found. Furthermore, a product with M^+ 180 shows a spectral fingerprint that is very similar to that of phenazine in the NIST library, probably indicating the presence of alkaloids in the WEOM (as pyrrole, pyridine and diketopiperazines, usually from vegetal protein, were not detected). Alkaloids could explain the large peak for indole as well. The most typical feature of this chromatogram is, however, the abundance of C_{14}- and especially C_{12}-fatty acid. This fatty acid pattern, different from that of the other source candidates (with prevalence of C_{16} and C_{18}), in addition to the alkaloid products, could be useful for identifying DOM derived from Cyperaceae.

Sphagnum moss from the Torfhausmoor peatland produced a large amount of levoglucosan and other polysaccharide products that are relatively abundant in comparison with the other samples, such as levogalactosan, 5-hydroxymethyl-2-tetrahydrofuraldehyde-3-one and 1,4-dideoxy-D-glycero-hex-1-enopyranose-3-ulose (see also Kuder and Kruge, 2001) (Figure 5). Some of these compounds may have been formed partially from the relatively decay-resistant carbohydrate in *Sphagnum* known as sphagnan. The phenolic products were dominated by 4-vinylphenol and 4-isopropenylphenol, the latter of which is widely considered as a marker of sphagnum acid (*p*-hydroxy-b-[carboxymethyl]-cinnamic acid) in pyrolyzates (e.g. Kracht and Gleixner, 2000). Phenol is also abundant, in this case proba-

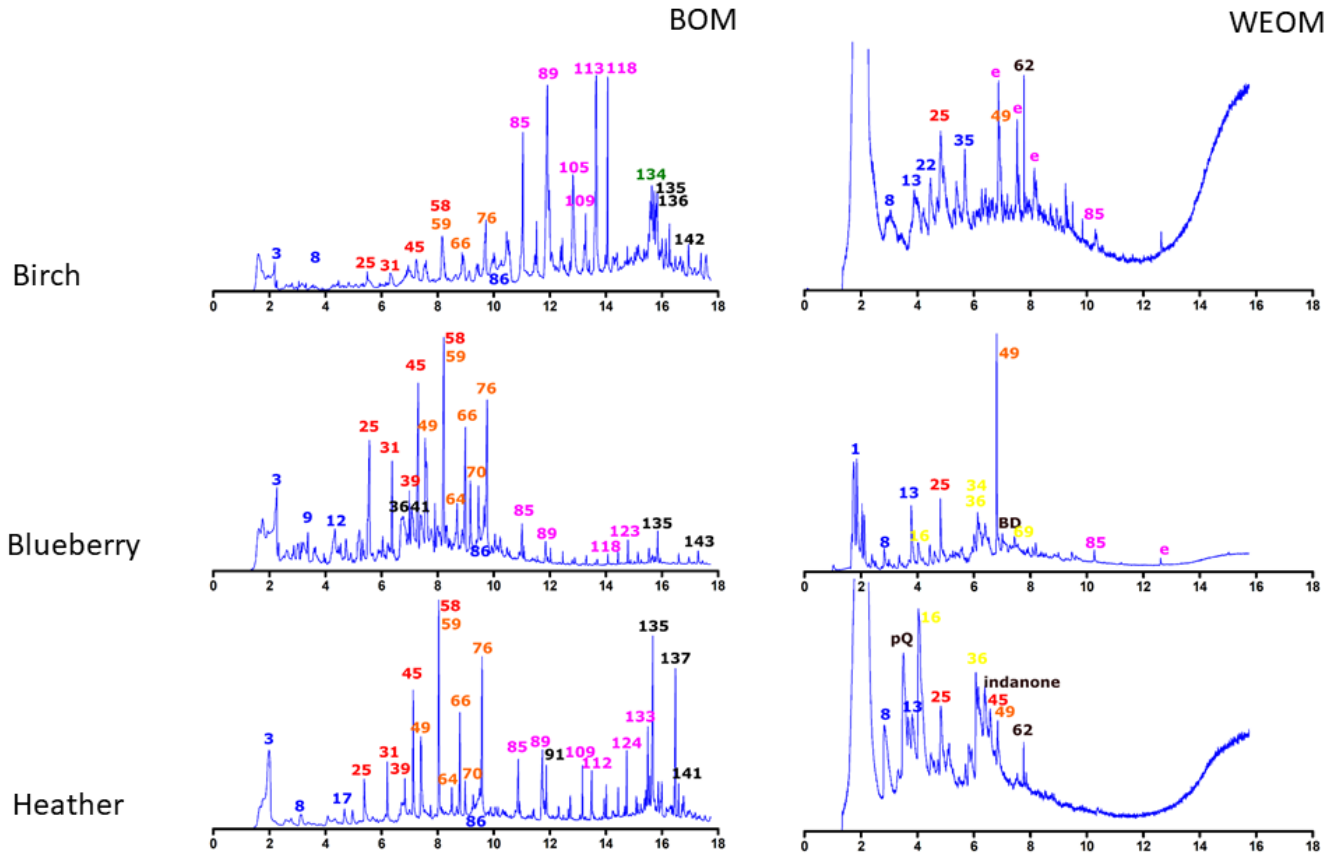

Fig. 3. Pyrolysis chromatograms (pyrograms) of BOM and WEOM samples. For explanations of labels see Figure 1 caption.

bly formed after side-chain elimination of sphagnum acid during pyrolysis. The *Sphagnum* sample did not produce enough WEOM to generate a meaningful chromatogram, but it may be anticipated that *Sphagnum*-derived DOM carries a signal of isopropenylphenol accompanied with 4-vinylphenol, as this was identified in the headwater DOM from the Odersprung bog (Kaal et al., 2017).

Contrary to *Sphagnum*, the epiphytic moss (growing on a trunk of dead spruce wood) did not produce sphagnum acid products. However, the polysaccharide fingerprint, which is dominant in the Py-GC-MS chromatogram, is quite similar (Figure 5). This sample is prolific of fatty acids, a homologous series of *n*-alkanes and a big hump of alkane and alkene products which could not be resolved. The WEOM pyrolyzate was of low intensity and dominated by alkenes and alkylbenzenes, representing a fingerprint of little diagnostic value.

B. Other materials

The sample of an unidentified lichen was profilic of polysaccharides products, including anhydrosugars (Figure 5). The BOM produced 3-methoxy-5-methylphenol, which is one of the most omnipresent markers of lichen (Schellekens et al., 2015). Another prominent peak that was not observed in other sources is that of m/z 254 and 226, probably corresponding to an anthraquinone (possibly chrysophanol). Anhydrosugars and G-type lignin products as well (in particular guaiacol, 4-methylguaiacol and 4-ethylguaiacol, but not 4-vinylguaiacol or 4-propenyl-substituted guaiacols). The fungus (or fungi) that compose(s) the lichen is probably lacking chitinase activity (non chitinous cell walls, known for many lichenised ascomycota) as markers of chitin were below the detection limit. Protein-derived N-containing were slightly more abundant than in vascular plant sources in general. The 3-methoxy-5-methylphenol was also one of the largest peaks in the chromatogram obtained from the WEOM, supporting the value of this moiety as a marker of lichens (or lichinized fungi) in DOM. A compound with m/z 152, 137 and 121, that eluted at the same retention time as 4-ethylguaiacol from the other samples (which also produces m/z 137 and 152 but in different relative proportions), could correspond to a dimethoxytoluene. Alkenes, guaiacol, benzofuran, phenols, furaldehydes and toluene were also detected but in minor proportions.

The fruit body of an unidentified fungus barely produced products of chitin such as acetamide and acetamidofurans (not listed), suggesting that the proportion of chitin was low. The majority of the BOM's Py-GC-MS signal can be ascribed to polysaccharides (levoglucosan, levogalactosan, 5-methyl-2-furaldehyde, 2-hydroxy-3-methyl-2-cyclopenten-1-one, dianhydrorhamnose, 1,4-dideoxy-D-glycero-hex-1-enopyranose-3-ulose and 5-hydroxymethylfurfural) (Figure 6). An unidentified compound (allegedly an alkylcyclohexane) was also abundant and another unidentified product with m/z 244, 382 and 422 may be a marker of this fungus. A hump of alkene fragments at the end of the chromatogram could not be resolved. Methylnaphthalenes were also relatively abundant. Phenol and fatty acids were among the minor compounds (no methoxyphenols detected). With exception of the possible high molecular weight markers and several unusually abundant carbohydrate prod-

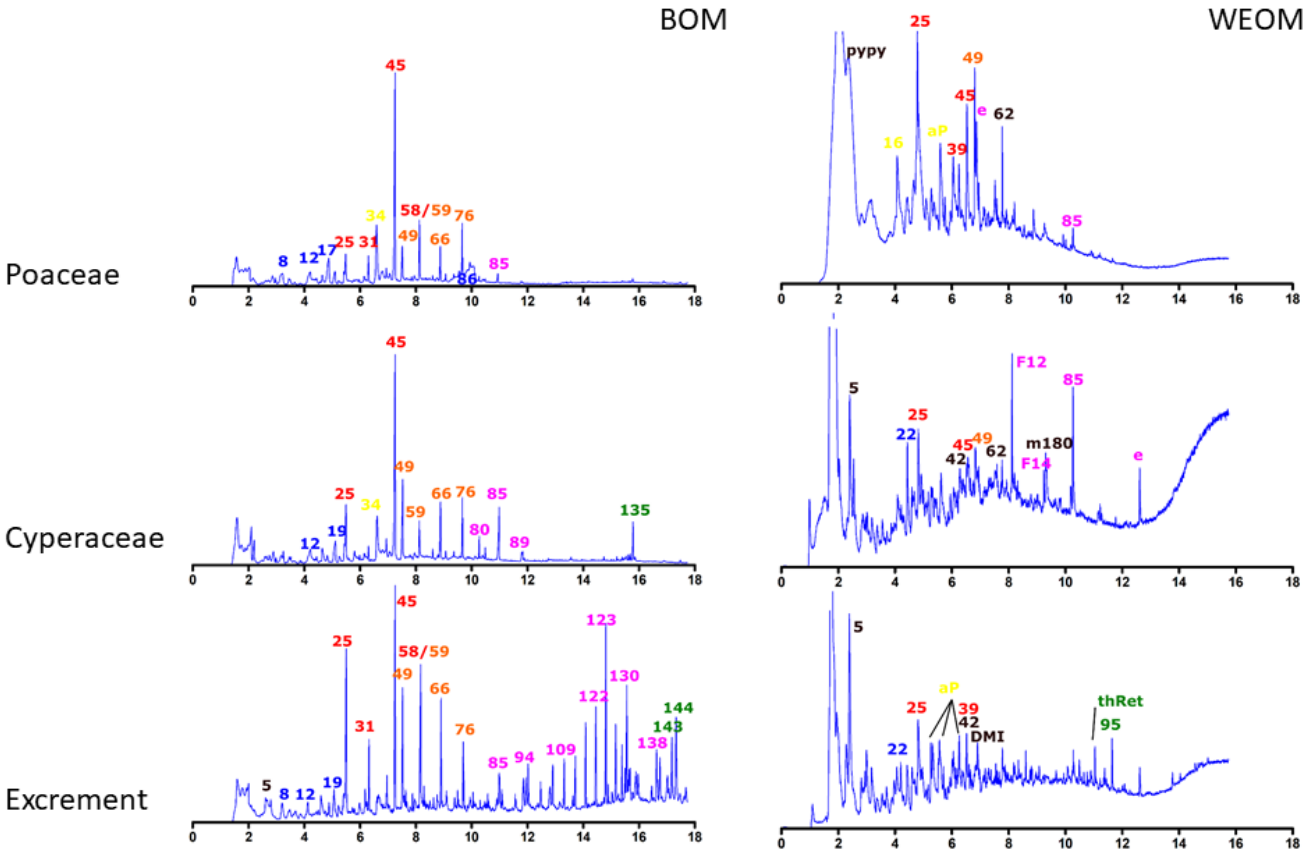

Fig. 4. Pyrolysis chromatograms (pyrograms) of BOM and WEOM samples. For explanations of labels see Figure 1 caption.

ucts, these results are not very useful for source identification. Furthermore, it may be anticipated that the fruit bodies of this fungus are not representative of the soil or stream fungi in the system, as the microbial DOM from the Odersprung did produce acetamide and the main N-containing products of DOM, i.e. pyrroles and pyridines (Kaal et al., 2016, 2017). The WEOM samples produced some alkenes upon pyrolysis but the chromatogram was of poor quality.

The epilithic biofilm was taken from a slimy layer on a rock substrate in the Ecker stream. Such materials are known to be hotspots of in-stream DOM degradation (Kamjunke et al., 2016). Fatty acids (C_{16}, C_{18}) dominate the Py-GC-MS chromatogram (Figure 6) but markers of polysaccharides (e.g. levoglucosan and pyrans) and proteins (e.g. indoles, pyrrole, pyridine, and in this case probably toluene) are abundant as well. Phytadienes from chlorophyll probably indicate green algae in this material. However, traces of abietane diterpenes suggests that gymnosperm needles may have deposited in the film, which could also be a source of chlorophyll. Short chain alkanes/alkenes were also relatively abundant, possibly from phytoplankton or bacteria. The WEOM sample gave a chromatogram of moderate quality, containing the series of alkenes (C_{14}-C_{16} chain length range), phenol, guaiacol (probably from spruce debris inclusions) and a large peak (the largest peak) for a compound with m/z 91, 119 and 120, tentatively identified as methylbenzaldehyde. This compound was identified as smaller peaks in many other pyrolyzates. A series of unidentified high molecular weight compounds was also detected.

The BOM sample of green algae (Figure 6) gave largest peaks for C_{16} and C_{18} fatty acid, phytadienes and several unidentified products with characteristic m/z 79 and 91. The alkane/alkene pattern had a noteworthy maximum at C_{17}. An aromatic compound with m/z 157, 142 and 172 probably corresponds to a trimethyldihydronaphthalene. Toluene, carbohydrate products and some N-containing products were identified as well. The results suggest that fresh green algae may be recognized by high proportions of phytadiene but this signal is unlikely to be transferred to, and preserved in, DOM samples. No data was obtained for WEOM (insufficient BOM).

Finally, the excremental fabric (Figure 4) showed a clear sign of stanols from the animal tract. However, the majority of the signal can be ascribed to lignin: guaiacol, 4-methylguaiacol, 4-vinylguaiacol, $C_{3:1}$-guaiacol (*trans*) and the syringyl analogues represent the lignin fibre, from the animals' predominantly herbaceous diet. Nitrogen-containing compounds are also clearly present in the chromatograms. Amyrins represent vegetal-derived pentacyclic terpenoids (amyrins; some of which also identified in the pyrolyzate of blueberry BOM), long-chain alkanes may correspond to plant waxes (including *Vaccinium*), long-chain methylketones might originate mainly from blueberry as well, traces of retene represent gymnosperm diterpene (probably spruce) and tocopherol was also identified. The WEOM signal revealed (tetrahydro)retene peaks indicative of diterpene resins, probably of spruce (spruce is known to be eaten by deer, but it is not their favourite food; e.g., Duncan et al., 1994). Multiple peaks of compounds with high molecular

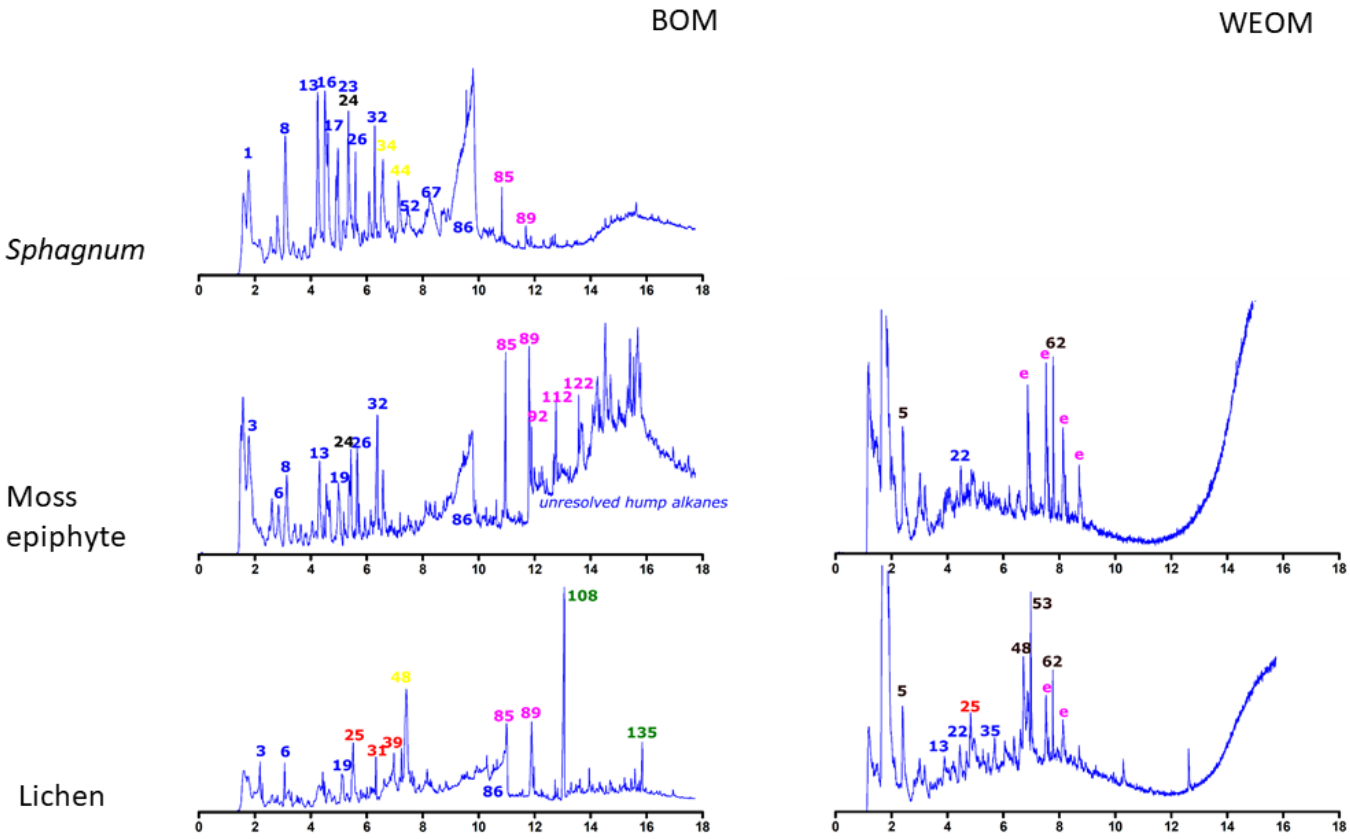

Fig. 5. Pyrolysis chromatograms (pyrograms) of BOM and WEOM samples. For explanations of labels see Figure 1 caption.

weight could not be identified and might correspond largely to stanols. The dimethylindanone was also detected, as were phenols (including abundant C_3-alkylphenols), guaiacols, alkylhydronaphthalenes and indole. This chromatogram reflects primarily grasses, spruce needles and *Vaccinium* tissues (also browsed by roe deer; Bergquist et al., 2001).

C. Use of BOM and WEOM data to evaluate environmental DOM chemistry

The Py-GC-MS fingerprints of the environmental DOM from the Odersprung system are dominated by phenols, especially phenol, and alkylbenzenes (Kaal et al., 2017). Polycyclic aromatic hydrocarbons (PAHs, non-resinous) are much more abundant than in any of the WEOM samples studies, and this might be indicative of a significant effect of Fe-DOM complexes (DOM-bound free or mineral Fe), deteriorating the Py-GC-MS fingerprints (Miltner and Zech, 1997). It could clearly be established that the samples from the forest environment are enriched in guaiacol and microbial products (N-compounds), whereas DOM samples from the peatland environment contained larger proportions of phenol, syringols and sphagnum acid products.

Using the results of BOM and WEOM from the source candidiates, we can hypothesize on the likely sources of the pyrolysis products from the Odersprung DOM. Firstly, the abundance of 2-methylfuran and 2,5-dimethylfuran might indicate a source in polysaccharides from *Vaccinium*. The proportion of these products decreases downstream. The products 4-vinylphenol and 4-isopropenylphenol also diminish at the boundary from peat to forest environment, and both can be explained by a reduction in the proportion of DOM from *Sphagnum*, and possibly also of peatland graminoids (grasses/sedges). The predominance of guaiacol among the guaiacols, and virtual absence of syringols in all samples except for those from the peatland, indicate that it is very likely that woody spruce debris is the main source of lignin-derived DOM in all the samples taken from the forest environment. This is confirmed by the presence of the same DHA (diterpene) products as identified in particular from spruce needles. The microbial DOM was prolific of furaldehydes (can be of any source though, including many vascular plants analyzed) but also of diketodipyrrole. This latter compound was not a significant product in the source candidates' chromatograms (WEOM), suggesting that they may originate from microbial soil organic matter which has no representatives in the present reference set of fresh source candidates. Hence, so far there is no indication that could be used to link this microbial DOM to WEOM fingerprints of the source candidates.

The possible effects of mineral-induced alterations complicate the further analysis of the sources of the DOM samples from the Odersprung, especially in absence of a detailed numerical evaluation of the WEOM pyrolysis fingerprints. Kaal et al. (2017) found that THM-GC-MS provided a much better signal-to-noise ratio of the chromatograms than the Py-GC-MS chromatograms. However, Py-GC-MS provided a more reliable fingerprint of the microbial-derived DOM (probably from the forest soils). The next step in the research project is to elaborate the THM-GC-MS data of the BOM and WEOM of the source candidates in detail (using semi-quantitative data instead of visual inspection of chromatograms) and use that information for

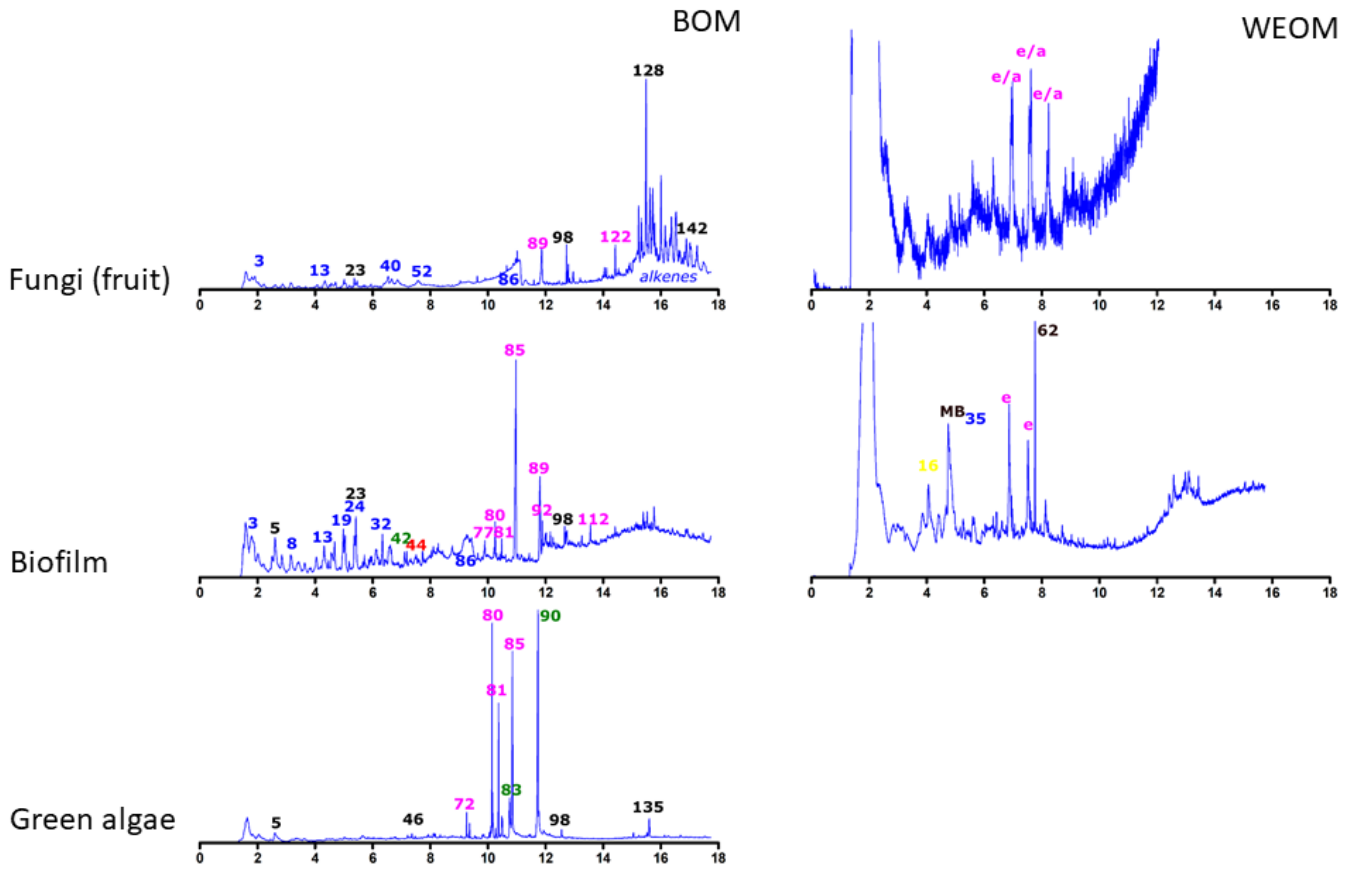

Fig. 6. Pyrolysis chromatograms (pyrograms) of BOM and WEOM samples. For explanations of labels see Figure 1 caption.

source identification, similar to the Py-GC-MS data discussed here. In addition, the Odersprung THM-GC-MS data will be re-quantified using the much longer compound list and an attempt is made to isolate the signal of microbial DOM in the samples from the Odersprung (not done for Py-GC-MS as Py-GC-MS produces fewer compounds that can be used to identify sources, and is more prone to secondary rearrangements), partly by means of quantification of a series of microbial products not used in the published study of the Odersprung (Kaal et al., 2017).

4. CONCLUSIONS

The Py-GC-MS results of the BOM samples analyzed were in line with what can be expected of the different types of potential sources analyzed, which is useful for tracing the origin of environmental DOM samples in the future. Analysis of BOM also enabled to identify the link between BOM and WEOM chemistries. In that regard, most BOM samples with a high aliphatic content produced a WEOM that was depleted in such compounds. WEOM fractions were enriched in phenols of diverse nature and other O-containing aromatic compounds that were not produced by their BOM counterparts, so that they likely reflect effects of mobilization and or different pyrolytic pathways between BOM and WEOM. The remarkable contrast in the distribution of likely lignin products could also be identified on the basis of comparing BOM and WEOM. This was found for syringols, but especially for guaiacols. For guaiacyl products, dominance of guaiacol and presence of vanillic acid methyl ester, possibly indicate an important source in acylated vanillic acid (esterified to the lignin backbone) rather than the lignin macromolecule as well. The abundance of dimers (high in POM, low in WEOM) may also relate to the balance of aryl ether B-O-4 polymeric lignin and vanillic ester (G-G acylation) side-chains. This will be further explored using THM-GC-MS.

Acknowledgements. This research was funded by the Deutsche Forschungsgemeinschaft (DFG) Projektnummer 419258863, entitled "The Role of Molecular Dissolved Organic Matter (DOM) Composition to Identify Sources and Release of DOM and Trace Elements in Catchments of Drinking Water Reservoirs in Mid-Latitude Mountain Areas – (DOMtrace) A Solid Phase Analytic Pyrolysis (Py-GC-MS, THM-GC-MS) Approach".

5. REFERENCES

Amon, R.M.W., Rinehart, A.J., Duan, S., Louchouarn, P., Prokushkin, A., Guggenberger, G., Bauch, D., Stedmon, C., Raymond, P.A., Holmes, R.M., McClelland, J.W., Peterson, B.J., Walker, S.A., Zhulidov, A.V., 2012. Dissolved organic matter sources in large Arctic rivers. Geochimica et Cosmochimica Acta 94, 217-237.

Arnberger, A., Ebenberger, M., Schneider, I.E., Cottrell, S., Schlueter, A.C., von Ruschkowski, E., Venette, R.C., Snyder, S.A., Gobster, P.H., 2018. Visitor preferences for visual changes in bark

beetle-impacted forest recreation settings in the United States and Germany. Environmental Management 61, 209–223.

Bergquist, J., Kullberg, Y., Örlander, G., 2001. Effects of shelterwood and soil scarification on deer browsing on planted Norway spruce *Picea abies* L. (Karst) seedlings. Forestry: An International Journal of Forest Research 74, 359–367.

Broder, T., Biester, H., 2015. Hydrologic controls on DOC, As and Pb export from a polluted peatland-the importance of heavy rain events, antecedent moisture conditions and hydrological connectivity. Biogeosciences 12, 4651–4664.

Broder, T., Knorr, K.-H., Biester, H., 2017. Changes in dissolved organic matter quality in a peatland and forest headwater stream as a function of seasonality and hydrologic conditions. Hydrology and Earth System Sciences 21, 2035–2051.

Chantigny, M.H., 2003. Dissolved and water-extractable organic matter in soils: a review on the influence of land use and management practices. Geoderma 113, 357-380.

De Leeuw, J.W., Versteegh, G.J.M., van Bergen, P.F., 2006. Biomacromolecules of algae and plants and their fossil analogues. Plant Ecology 182, 209-233.

Derenne, S., Quénéa, K., 2015. Analytical pyrolysis as a tool to probe soil organic matter. Journal of Analytical and Applied Pyrolysis 111, 108–120.

Dijkstra, E.F., Boon, J.J., van Mourik, J.M., 1998. Analytical pyrolysis of a soil profile under Scots pine. European Journal of Soil Science 49, 295–304.

Dufour, A., Weng, J., Jia, L., Tang, X., Sirjean, B., Fournet, R., Le Gall, H., Brosse, N., Billaud, F., Mauviel, G., Qi, F., 2013. Revealing the chemistry of biomass pyrolysis by means of tunable synchrotron photoionisation-mass-spectrometry. RSC Advances 3, 4786.

Duncan, A.J., Hartley, S.E., Iason, G.R., 1994. The effect of monoterpene concentrations in Sitka spruce (*Picea sitchensis*) on the browsing behaviour of red deer (*Cervus elaphus*). Canadian Journal of Zoology 72, 1715-1720.

Evans, R.L., Milne, T.A., Soltys, M.N., 1986. Direct mass-spectrometric studies of the pyrolysis of carbonaceous fuels. III. Primary pyrolysis of lignin. Journal of Analytical and Applied Pyrolysis 9, 207-236.

Filley, T.R., Cody, G.D., Goodell, B., Jellison, J., Noser, C., Ostrofsky, A., 2002. Lignin demethylation and polysaccharide decomposition in spruce sapwood degraded by brown rot fungi. Organic Geochemistry 33, 111-124.

Guggenberger, G., Zech, W., 1994. Dissolved organic carbon in forest floor leachates: simple degradation products or humic substances? The Science of the Total Environment 152, 37-47.

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

Jiang, T., Kaal, J., Liang, J., Zhang, Y., Wei, S., Wang, D., Green, N.W., 2017 Quality of dissolved organic matter (DOM) from periodically submerged soils in the Three Gorges Reservoir areas as determined by elemental and optical analysis, infrared spectroscopy, pyrolysis-GC-MS and thermally assisted hydrolysis and methylation. The Science of the Total Environment 603–604, 461–471.

Kaal, J., Wagner, S., Jaffé, R., 2016. Molecular properties of ultrafiltered dissolved organic matter and dissolved black carbon in headwater streams as determined by pyrolysis-GC–MS. Journal of Analytical and Applied Pyrolysis 118, 181–191.

Kaal, J. Martínez-Cortizas, A., Biester, H. 2017. Downstream changes in molecular composition of DOM along a headwater stream in the Harz mountains (Central Germany) as determined by FTIR, Pyrolysis-GC–MS and THM-GC–MS. Journal of Analytical and Applied Pyrolysis 126, 50-61.

Kamjunke, N., Oosterwoud, M.R., Herzsprung, P., Tittel, J., 2016. Bacterial production and their role in the removal of dissolved organic matter from tributaries of drinking water reservoirs. The Science of the Total Environment 548–549, 51–59.

Kögel-Knabner, I., Rumpel, C., 2018. Advances in molecular approaches for understanding soil organic matter composition, origin, and turnover: A historical overview. Advances in Agronomy 149, 1-48.

Kolattukudy, P.E., 2001. Polyesters in higher plants. Advances in Biochemical Engineering/Biotechnology 71, 1–49.

Kracht, O., Gleixner, G., 2000. Isotope analysis of pyrolysis products from *Sphagnum* peat and dissolved organic matter from bog water. Organic Geochemistry 31, 645–654.

Kuder, T., Kruge, M.A., 2001. Carbon dynamics in peat bogsL Insights from substrate macromolecular chemistry. Global Biogeochemical Cycles 15, 721-727.

Kuroda, K., Ashitani, T., Fujita, K., 2010.Tetramethylammonium hydroxide (TMAH) thermochemolysis of lignin: Formation of (*E*)-5-formyl-2,3,3′,4′-tetramethoxystilbene and its origins. Journal of Analytical and Applied Pyrolysis 89, 233–238

Matilainen, A., Gjessing, E.T., Lahtinen, T., Hed, L., Bhatnagar, A., Sillanpää, M., 2011. An overview of the methods used in the characterisation of natural organic matter (NOM) in relation to drinking water treatment. Chemosphere 83, 1431-1442

Miltner, A., Zech, W., 1997. Effects of minerals on the transformation of organic matter during simulated fire-induced pyrolysis. Organic Geochemistry 26, 175-182.

Montgomery, W., Potiszil, C., Watson, J.S., Sephton, M.A., 2016. Sporopollenin, a natural copolymer, is robust under high hydrostatic pressure. Macromolecular Chemistry and Physics 217, 2494-2500.

Moore, T.R., Dalva, M., 2001. Some controls on the release of dissolved organic carbon by plant tissues and soils. Soil Science 166, 38-47.

Neto, C.C., 2011. Ursolic acid and other pentacyclic triterpenoids: Anticancer activities and occurrence in berries. In: G.D. Stoner, N.P. Seeram (Eds.), Berries and Cancer Prevention, Springer, pp. 41-49.

Overbeck, M., Schmidt, M., 2012. Modelling infestation risk of Norway spruce by *Ips typographus* (L.) in the Lower Saxon Harz Mountains (Germany). Forest Ecology and Management 266, 115-125.

Schellekens, J., Buurman, P., Kuyper, T.W., 2012. Source and transformations of lignin in *Carex*-dominated peat. Soil Biology Biochemistry 53, 32-42.

Schellekens, J., Bradley, J.A., Kuyper, T.W., Fraga, I., Pontevedra-Pombal, X., Vidal-Torrado, P., Abbott, G.D., Buurman, P., 2015. The use of plant-specific pyrolysis products as biomarkers in peat deposits. Quaternary Science Reviews 123, 254-264.

Van der Heyden, E., Boon, J.J., 1994. A combined pyrolysis mass spectrometric and light microscopic study of peatified *Calluna* wood isolated from raised bog peat deposits. Organic Geochemistry 22, 903-919.

Van Heemst, J.D.H., del Río, J.C., Hatcher, P.G., de Leeuw, J.W., 2000. Characterization of estuarine and fluvial dissolved organic matter by thermochemolysis using tetramethylammonium hydroxide. Acta Hydrochimica et Hydrobiologica 28, 69–76.

Vogt, R.D., 2003. Increase in colour and amount of organic matter in surface waters. NORDTEST Position paper 009. Available at http://www.nordtest.org. 11 p.

Wehling, K., Niester, C., Boon, J.J., Willemse, M.T., Wiermann, R., 1989. *p*-Coumaric acid - a monomer in the sporopollenin skeleton. Planta 179, 376-380.

Zsolnay, A., 1996. Dissolved humus in soil waters. *In*: A. Piccolo (Ed.), Humic Substances in Terrestrial Ecosystems, Elsevier, Amsterdam, pp. 171-223.

Table 2. List of pyrolysis products. Labels refer to peak labels in the chromatograms depicted in Figures 1 to 6.

Label	RT (min)	Compound
1	1.800	2-methylfuran
2	2.008	benzene
3	2.033	acetic acid
4	2.200	hydroxypropanone
5	2.609	toluene
6	2.874	(2*H*)-furan-3-one
7	2.910	propionaldehyde
8	3.170	3/2-furaldehyde
9	3.237	2-cyclopenten-1-one
10	3.517	unidentified carbohydrate product
11	3.616	tetrahydrofuran-3-one/pentanal
12	4.098	2,3-dihydro-5-methylfuran-2-one
13	4.358	5-methyl-2-furaldehyde
14	4.540	benzofuran
15	4.555	benzonitrile
16	4.597	phenol
17	4.684	4-hydroxy-5,6-dihydro-(2*H*)-pyran-2-one
18	4.887	limonene
19	4.980	2-hydroxy-3-methyl-2-cyclopenten-1-one
20	4.998	indene
21	5.053	dianhydrorhamnose
22	5.141	2,3-dimethylcyclopent-2-en-1-one
23	5.364	4-methylphenol
24	5.421	2-propan-2-one-tetrahydrofuran
25	5.505	guaiacol
26	5.650	levoglucosenone
27	5.728	3-hydroxy-2-methyl-(2*H*)-pyran-4-one
28	5.831	C_1-indole
29	6.060	unidentified carbohydrate product
30	6.151	naphthalene
31	6.304	4-methylguaiacol
32	6.359	5-hydroxymethyl-2-tetrahydrofuraldehyde-3-one
33	6.532	3,5-dihydroxy-2-methyl-(4*H*)-pyran-4-one
34	6.610	4-vinylphenol
35	6.677	1,4:3,6-dianhydro--D-glucose
36	6.772	catechol
37	6.781	ethylbenzaldehyde

Continued on next page

Continued from previous page

Label	RT (min)	Compound
38	6.859	5-hydroxymethyl-2-furaldehyde
39	6.942	4-ethylguaiacol
40	6.989	C_1-naphthalene
41	7.072	3-methoxycatechol
42	7.108	indole
43	7.112	C1-naphthalene
44	7.134	4-isopropenylphenol
45	7.248	4-vinylguaiacol
46	7.362	alkylhydronaphthalene compound (C_3-(4*H*)-naphthalene)
47	7.414	2-coumaranone/*trans*-4-propenylphenol
48	7.429	3-methyl-5-methoxyphenol
49	7.518	syringol
50	7.532	4-(prop-1-enyl)guaiacol
51	7.575	biphenyl
52	7.695	1,4-dideoxy-D-glycero-hex-1-enopyranose-3-ulose
53	7.699	2,5-dimethoxytoluene
54	7.741	C_1-indole
55	7.897	4-(prop-2-enyl)guaiacol (*cis*)
56	7.907	4-formylguaicol (vanillin)
57	8.005	Vanillic acid/C2-catechol
58	8.102	4-(prop-2-enyl)guaiacol (*trans*)
59	8.130	4-methylsyringol
60	8.135	fatty acid C_{10}
61	8.286	homovanillin
62	8.348	2,6-di-*tert*-butylphenol
63	8.467	4-acetylguaiacol
64	8.696	4-ethylsyringol
65	8.737	4-propan-2-one-guaiacol
66	8.872	4-vinylsyringol
67	8.882	levogalactosan
68	8.939	fluorene
69	9.002	C_1-catechol
70	9.069	4-(prop-1-enyl)syringol
71	9.090	coniferyl alcohol
72	9.261	alkene C_{17}
73	9.292	levomannosan
74	9.360	4-(prop-2-enyl)syringol (*cis*)
75	9.360	alkane C_{17}
76	9.661	4-(prop-2-enyl)syringol (*trans*)

Continued on next page

Continued from previous page

Label	RT (min)	Compound
77	9.894	fatty acid C_{14}
78	10.182	anthracene/phenanthrene
79	10.237	4-propan-2-one-syringol
80	10.268	phytadiene 1
81	10.486	phytadiene 2
82	10.496	phytadiene 3
83	10.755	unidentified product
84	10.885	C_1-anthracene
85	11.046	fatty acid C_{16}
86	11.093	levoglucosan
87	11.477	unidentified compound
88	11.606	unidentified compound
89	11.798	fatty acid $C_{18:1}$
90	11.840	unidentified product
91	11.881	unidentified compound
92	11.887	fatty acid C_{18}
93	11.975	unidentified compound
94	12.032	fatty acid $C_{18:2}$
95	12.353	retene
96	12.411	unidentified compound
97	12.649	unidentified compound
98	12.675	unidentified compound
99	12.722	unidentified compound
100	12.769	unidentified compound
101	12.779	unidentified compound
102	12.789	unidentified gymnosperm diterpene
103	12.790	fatty acid C_{20}
104	12.795	unsaturated fatty acid
105	12.815	unsaturated fatty acid
106	12.862	unidentified compound
107	12.961	unidentified compound
108	13.059	unidentified compound
109	13.278	*n*-alkane
110	13.329	guaiacol dimer (or abietance)
111	13.412	unidentified compound
112	13.568	fatty acid C_{22}
113	13.660	unsaturated fatty acid
114	13.713	*n*-alkane
115	13.717	guaiacol dimer

Continued on next page

Continued from previous page

Label	RT (min)	Compound
116	13.972	guaiacol dimer
117	14.019	long-chain 2-methylketone
118	14.060	*n*-alkane
119	14.133	guaiacol dimer
120	14.180	guaiacol dimer
121	14.341	fatty acid C_{24}
122	14.455	*n*-alkane
123	14.813	*n*-alkane
124	14.839	long-chain 2-methylketone
125	15.031	unidentified compound
126	15.119	unidentified compound
127	15.166	*n*-alkane
128	15.493	unidentified compound
129	15.529	fatty acid C_{26}
130	15.560	*n*-alkane
131	15.586	unidentified compound
132	15.586	alkanol
133	15.643	long-chain 2-methylketone
134	15.731	unidentified compound
135	15.794	stigmastan-3,5-diene
136	15.799	alpha-tocopherol
137	16.478	unidentified compound
138	16.629	long-chain 2-methylketone
139	16.722	unidentified compound
140	16.748	unidentified compound
141	16.779	unidentified compound
142	16.894	unidentified compound
143	17.174	pentacyclic triterpenoid (amyrin)
144	17.330	pentacyclic triterpenoid (amyrin)