Improved end-member characterization of modern organic matter pools in the Ohrid Basin (Albania, Macedonia) and evaluation of new palaeoenvironmental proxies

J. Holtvoeth[1,2], D. Rushworth[1], H. Copsey[2], A. Imeri[3], M. Cara[3], H. Vogel[4], T. Wagner[5], and G. A. Wolff[1]

[1]{Department of Earth, Ocean and Ecological Sciences, School of Environmental Sciences, University of Liverpool, Liverpool, UK}

[2]{School of Chemistry, Organic Geochemistry Unit, University of Bristol, Bristol, UK}

[3]{Faculty of Agriculture and Environment, Agricultural University of Tirana, Tirana, Albania}

[4]{Institute of Geological Sciences & Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland}

[5]{School of Civil Engineering & Geosciences, University of Newcastle, Newcastle-upon-Tyne, UK}

Correspondence to: J. Holtvoeth (j.holtvoeth@bristol.ac.uk)

Abstract

We present elemental, lipid biomarker and compound-specific isotope ($\delta^{13}C$, $\delta^{2}H$) data for soils and leaf litter collected in the catchment of Lake Ohrid (Albania, Macedonia), as well as macrophytes, particulate organic matter and sediments from the lake itself. Lake Ohrid provides an outstanding archive of continental environmental change of at least 1.2M years and the purpose of our study is to ground truth organic geochemical proxies that we developed in order to study past changes in the terrestrial biome. We show that soils dominate the lipid signal of the lake sediments rather than the vegetation or aquatic biomass. There is a strong imprint of suberin monomers on the composition of total lipid extracts and chain-length distributions of *n*-alkanoic acids, *n*-alcohols, ω-hydroxy acids and α,ω-dicarboxylic acids. Our end-member survey identifies that ratios of mid-chain length suberin-derived to long-chain length cuticular-derived alkyl compounds as well as their average chain length distributions can be used as new molecular proxies of organic matter sources to the lake. We tested these for the 8.2 ka event, a

pronounced and widespread Holocene climate fluctuation. In SE Europe climate became drier and cooler in response to the event, as is clearly recognizable in the carbonate and organic carbon records of Lake Ohrid sediments. Our new proxies indicate biome modification in response to hydrological changes, identifying two phases of increased soil OM supply, first from soils with moderately degraded OM and then from more degraded soils. Our study demonstrates that geochemical fingerprinting of terrestrial OM should focus on the main lipid sources, rather than the living biomass. Both can exhibit climate-controlled variability, but are generally not identical.

1 Introduction

Climate-controlled changes in hydrology are reflected in the development of the terrestrial biome, which determines the amount and the quality of organic matter (OM) supplied to environmental archives (Meyers and Lallier-Vergès, 1999). Currently, reconstructions of biome dynamics are largely based on pollen data, focussing on the development of the vegetation (Birks and Birks, 2006). The dynamics of the soil organic carbon pool, however, is poorly constrained, mainly because palynologists cannot easily discriminate between direct supply of pollen and intermediate storage of pollen in soils. Similarly, paleoclimatologists using established organic geochemical proxies such as the organic carbon to nitrogen ratio (C/N) or bulk organic carbon isotopes ($\delta^{13}C_{org}$) to identify changing proportions of OM from aquatic and terrestrial sources struggle to differentiate between the individual elements of the terrigenous OM pool. Particularly problematic is the variable composition of terrestrial OM resulting from changing proportions of soil and plant litter and the degree of biological degradation that is closely related to the moisture regime. Furthermore, potential end-member materials are often not properly defined and proxies are transferred between supposedly comparable settings untested, reducing their actual applicability and potentially leading to erroneous interpretations.

We focus on Lake Ohrid in the Western Balkans (Fig. 1), which provides an outstanding archive of continental environmental change dating back at least 1.2M years (Wagner et al., 2014). In 2013, almost 600 m of sediment core was recovered during a drilling campaign for the International Continental Drilling Program (ICDP). Initial analyses of core catcher material and logging data reveal a continuous limnic sequence, showing pronounced cyclicity that clearly relates to eccentricity and obliquity cycles (Wagner et al., 2014; Baumgarten et al., 2015). This exceptional continental sediment record thus appears very promising for high-resolution palaeoenvironmental reconstructions of orbitally controlled Northern Hemisphere climate

fluctuations. Earlier studies of pilot cores suggest a close relationship between the North Atlantic climate of the past 130,000 years and the sedimentation of inorganic as well as organic matter in the lake (Vogel et al., 2010; Holtvoeth et al., 2010). Whichever proxy is used to reconstruct changes in biogeochemical fluxes towards a sedimentary archive, the sources of each component have to be known in order to correctly interpret proxy variability. For distal marine sediment archives, mixing associated with long-distance transport results in averaged end-member characteristics for marine and terrestrially-sourced matter that are reasonably well defined. Terrestrial ecosystems, by contrast, show far greater complexity and diversity, with a great variety of local factors and processes that modify specific end-members. The local vegetation type, geology, geomorphology or hydrology may all reveal peculiarities that potentially interfere with the regional or global climate signal that paleoclimatologists are keen to extract from terrestrial archives.

As part of a longer term study of the Ohrid sedimentary record, we aim to improve the end-member definition of organic matter (OM) sources, and test new geochemical proxies across the 8.2 ka climate event. We provide lipid biomarker distributions and elemental compositions of modern materials from the Lake Ohrid Basin, including terrestrial plant litter, soils, macrophytes and lake water particles, all of which may influence the geochemical signature of the OM deposited in the lake.

A great advantage of the Lake Ohrid climate archive is its small, mountainous catchment (Fig. 1). Up to 55% of water inflow derives from karst springs fed by seepage from the higher altitude Lake Prespa. Although these waters do not deliver any sediment they do supply small amounts of nutrients to Lake Ohrid, e.g., 7% of the total phosphorous load (Matzinger et al., 2006), as well as calcium ions allowing carbonate precipitation. As there are no major rivers entering the Ohrid Basin, the supply of allochthonous sediment to the lake depends almost entirely on surface run-off from the surrounding mountain ranges in the form of small streams and gullies. The Triassic limestones forming the eastern, southeastern and northwestern slopes of the Ohrid Basin are highly permeable so that much of the precipitation is taken up by karstic systems before reaching the lake. Thus, the quantity and composition of terrestrial OM supplied to the lake will reflect hydrologically-controlled changes in vegetation and soil stocks of the immediate surroundings, while productivity in the lake reflects run-off-controlled nutrient supply. The steep morphology and the geological conditions in the catchment force a rapid response of the vegetation cover to precipitation decrease and are responsible for relatively low soil stability. Primary productivity in Lake Ohrid is low due to the low nutrient levels (e.g., total phosphorous: 0.15 ± 0.026 mmol m^{-3}, Matzinger et al., 2007). Short distances, rapid

response and little dilution through aquatic productivity are excellent preconditions for the reconstructions of biome dynamics using sedimentary OM composition.

While the use of biological markers as proxies for environmental change is well established, we present the first data set of this kind for the Ohrid Basin to improve the interpretation of sedimentary OM compositional changes. We incorporate observations made during the study of sediments from pilot cores taken prior to the 2013 ICDP drilling. Our catchment data then allows us to identify the main OM sources to the Lake Ohrid sediments.

We also determined the carbon and hydrogen stable isotope composition for *n*-alkanoic acids of modern terrestrial materials and macrophytes ($\delta^{13}C$, $\delta^{2}H$). Similar data from algal biomass could not be established and signals largely overlapped, therefore isotope-based proxies could not be developed. However, as the data produced may be of interest to scientists working on comparable subjects, it are reported and discussed in a supplementary section (see Supplement 2).

2 Sampling, materials and methods

2.1 Sample collection

Sampling was carried out along the southern and southeastern shores of Lake Ohrid (Fig. 1). We did not sample each component of the ecosystem, but instead focused on the major OM sources. Leaf litter and underlying topsoils, i.e. O2 and A horizons, respectively, were collected in June 2012 and 2013 on the Macedonian side of the lake at three sites: under light forest vegetation dominated by small oaks in the vicinity of the Saint Naum karst springs (SN, Fig. 1), from a similar site about 2 km north of Trpejca (TP) and in a high-altitude beech forest of the Galicica National Park (GN). In September 2014, two species of grass were sampled, one of which belongs to the evergreen genus *Festuca* and typically grows in shady locations on rocky substrates (GN-G). The generally most abundant but unidentified species was collected from the high-altitude grasslands at the top of Galicica Pass (GP, Fig. 1). Furthermore, we sampled a prominent type of subsoil (B horizon), a deep red Chromic Luvisol (Terra Rossa) that is frequently exposed in gullies or along road cuttings and typically sits above the Mesozoic carbonates. Two samples were collected from the same spot at a near vertical roadside soil profile (DG, Fig. 1) in June 2013 (DG-SS) and September 2014 (DG-SS 2). Both subsoil samples were taken from about 50 cm below the soil surface and about 30 cm below the A-horizon. A noteworthy difference is that sample DG-SS included exposed surface material

while DG-SS 2 was taken from after removing 10-15 cm of soil from the surface. In contrast to the topsoils, these samples do not visibly contain significant amounts of fine root material.

At two littoral sites, near the DG site and ~2 km north of Pogradec, we collected samples of reed (*Phragmites* spp.). Submerged macrophytes were collected from a small vessel along two transects from 4-16 m water depth off the town of Pogradec and the village of Tushemisht (Albania; Fig. 1) using a small Van Veen bottom sampler (Hydro-Bios, Kiel, Germany). From shallow to deep water, we collected benthic algae (*Cladophora* spp.), one species of pondweed (*Potamogeton perfoliatus*) and 3 species of charophytes (*Characea tomentosa*, *C. gymnophilia*, *C. contraria*).

Lake water samples were collected from two sites (DEEP and Co1202; Fig. 1) by deployment of Niskin Bottles (20 L). The water samples were filtered immediately after collection in the laboratory of the Hydrobiological Institute in Ohrid using an electrical pump and pre-combusted glass fiber filters with a nominal pore size of 0.7 µm. The filters were stored frozen (-20°C) until freeze-drying in the laboratory at Cologne University.

11 sediment samples were taken from core Lz1120 situated close to the southeastern shores of Lake Ohrid (Fig. 1). The detailed site description and age model are provided by Wagner et al. (2009). Accordingly, our samples cover the time span from 8.65 to 8.05 ka, sampled with a time resolution of 60 years. The sediment core sections were stored cold in the repository of the Institute of Geology and Mineralogy at Cologne University (Germany) prior to sub-sampling. Samples were freeze-dried before shipment and stored frozen (-20°C) prior to analysis. All terrestrial samples and macrophytes were stored cold in the field (4 °C) and then dried (70°C, 48 h).

2.2 Elemental analysis

Total carbon (TC) and total nitrogen (TN) contents were measured in duplicate (values were < 10% of the mean) using a CE Instruments NC 2500 elemental analyzer. Total organic carbon (TOC) was determined after acid vapor (HCl) digestion of the carbonate fraction (Yamamuro and Kayanne, 1995). Carbonate contents (assuming all carbonate was as $CaCO_3$) were then calculated from the difference between TC and TOC measurements using the equation $CaCO_3$ = (weight% TC – weight% TOC) * M_{CaCO3}/M_C (M = molar/atomic mass; $CaCO_3$ = 100, C = 12). Contents of total organic carbon (TOC) and carbonate ($CaCO_3$) are given as percentages of dry weight. The ratio of total organic carbon to total nitrogen (TOC/TN) is given as the molar ratio.

2.3 Lipid analysis

For extraction of the lipid fraction, about 0.7 g of plant matter and 2 g of soil and sediment were homogenized and sonicated (3 x 15 min) in a mixture of dichloromethane (DCM) and methanol (9:1, v:v; 10 mL). About 0.2 g of the grass samples and 5 g of the duplicate subsoil sample (DG-SS 2) were extracted by microwave-assisted solvent extraction (10 min at 70ºC) using the same solvent mixture. The total lipid extract (TLE) was concentrated and any water removed by passing it through a Pasteur pipette column packed with anhydrous sodium sulphate. Acid-catalysed transmethylation of *n*-alkanoic acids was achieved by adding a solution of acetyl chloride in methanol (1:30, v:v; 1 mL; 0°C) and then warming the samples (45°C, 12 h). This methylates the carboxyl group of free *n*-alkanoic acids and also breaks the ester bonds in bio-polyesters such as suberin and in cuticular wax esters, thus, releasing the (previously bound) alkyl monomers as methyl esters. The methylated TLEs were dried under nitrogen and re-dissolved in DCM and passed through a column loaded with potassium carbonate to remove excess acetic acid. Finally, N,O-bis-(trimethylsilyl)-trifluoroacetamide with 1% trimethylchlorosilane was added to the TLE and warmed (65°C, 30 min) to derivatise compounds with hydroxy groups (e.g., *n*-alcohols, sterols).

Water filtrates were extracted using a Dionex Accelerated Solvent Extraction system (ASE), with DCM and methanol (9:1, v:v) as solvent. Aliquots of the resulting total lipid extracts were separated into aliphatic, aromatic and hetero-compound fractions by column chromatography using deactivated SiO_2 (mesh size 60) and elution with hexane, DCM:hexane (2:1, v:v) and methanol, respectively. The polar fractions of the water filtrates were transmethylated and derivatised following the same protocol described above.

Aliquots of the transmethylated and derivatised extracts of samples collected in 2012/2013 were injected onto a Trace 2000 Series gas chromatograph (GC) fitted with an on-column injector and a fused high-temperature silica column (60 m × 0.25 mm i.d.; film: (5 % phenyl-) methylpolysiloxane; DB5-HT equivalent; J&W) with helium as the carrier gas (*ca.* 1.6 mL min^{-1}). A retention gap of deactivated fused silica (1 m × 0.32 mm i.d.) was used at the front of the column. The oven was programmed from 60°C to 170°C at 6°C min^{-1} after 1 minute, then to 315°C at 2.5°C min^{-1} and held for 10 minutes. The GC column was fed directly into the EI source of a Thermoquest Finnigan TSQ7000 mass spectrometer (ionisation potential 70 eV; source temperature 315°C; trap current 300 μA), operated in Full Data Acquisition mode, (50 - 600 Thompsons cycled every s). TLEs of the samples collected in 2014 (grasses and DG subsoil duplicate) were treated as above but analysed using a Trace 1300 Series GC fitted with

a different capillary column (fused silica; 50 m × 0.32 mm i.d.; film: dimethylpolysiloxane; Rtx-1, DB-1 equivalent; Thames Restek) and a programmed temperature vaporising (PTV) injector in splitless mode. The carrier gas was helium (2 mL min^{-1}). The GC was connected to a Thermo Fisher Scientific ISQ mass spectrometer in EI mode (ionisation potential 70 eV, source temperature 320ºC). Data were processed using Xcalibur software. Compounds were identified either by comparison of their mass spectra and relative retention indices with those available in the literature and/or by comparison with authentic standards. Quantitative data were calculated by comparison of peak area of the internal standard, 5α(H)-cholestane (spiked onto the samples before extraction), with those of the compounds of interest, using the total ion current chromatogram. The relative response factors of the analytes were determined individually for 36 representative compounds using authentic standards. For analytes for which authentic standards were not available, the response factors for similar compounds of the same class and/or similar structure were used. Response factors of standards were not determined for the grass samples (GN-G, GP-G) and the duplicate sample from site DG (DG-SS 2) and were assumed to be 1, hence data for those samples are semi-quantitative. Reproducibility of similar lipid analyses was determined to be < ±15 % (Kiriakoulakis et al., 2000). Data quality was checked regularly with procedural blanks for each extracted sample batch and organic contamination was subtracted from the sample values, although it typically was insignificant (<< 1 % of the sample values). Total lipids were calculated as the sum of all identified compounds from the total ion chromatograms (TIC).

2.4 Statistical methods

Statistical analyses (Multi-dimensional scaling, MDS; Analysis of Similarity, ANOSIM; Similarity Percentages, SIMPER) were conducted on lipid biomarkers, their concentrations being normalised to percentage of total identified lipids ($\%_{lipid}$). The analyses were carried out separately for soils, leaf litters, macrophytes, lake particulate organic matter (POM) and sediment samples from core Lz1120. 78 variables were chosen; these were saturated and branched fatty acids, hydroxy acids, alcohols, *n*-alkanes and amyrins. Where there were zero values, a minimum detection limit was chosen of one half of the lowest recorded concentration, namely 0.001 $\%_{lipid}$ (Yunker et al., 2005). Data were fourth root transformed and screened to confirm a normal distribution using a Draftsman's plot for all variables prior to analyses, which were carried out using PRIMER software (Primer-E Ltd., UK). All quantitative data can be found in the Supplementary Material.

3 Results and Discussion

3.1 Elemental analysis

3.1.1 Leaf litter and soils

Leaf litter samples from the near-shore low-altitude forests (sites SN and TP) have high TOC contents of 43 ± 2 % and TOC/TN ratios of ~27 ± 2 (Table 1). TOC contents of the Leptosol A-horizons at SN and TP range from 7.5 to 11 %, while their TOC/TN ratios vary little around 17 (Table 1). The TOC/TN ratio of the high-altitude Leptosol (GN) is slightly lower, with a value of 15.3 ± 0.8. For all Leptosols, carbonate contents are low and vary between 5 and 8 %. The surface-exposed material from the B-horizon of the Chromic Luvisol (Terra Rossa; DG-SS) has low TOC and carbonate contents of 0.9 % and < 0.5 %, respectively, within reported ranges for this soil type in the Mediterranean (e.g., Costantini et al., 2013: TOC 0.5 %, $CaCO_3$: 0.5 %, n = 48). The second sample (DG-SS 2) shows higher TOC and carbonate contents of 1.5 ± 0.1 % and 2.3 ± 1.5 %, respectively. The TOC/TN ratios of the Terra Rossa samples are almost identical, and have an average value of 11.6 ± 0.7, which is also significantly lower than the Leptosols (Table 1).

3.1.2 Macrophytes

The macrophyte samples can be separated into two types of plant that (a) do precipitate carbonate, i.e., *Characeae* spp., and (b) do not, i.e., *Cladophora* spp., *Potamogeton* spp.., *Cladophora* spp., *Potamogeton* spp. and the leaves of *Phragmites* spp.. Plants from group (b) have similar TOC contents (42 ± 2 %); TOC/TN ratios are also similar (13.5 ± 0.3; Table 2). The two specimens of *Characeae* (*Chara tomentosa*, *Chara gymnophyla*) have high amounts of carbonate (66 ± 7 %), diluting the TOC contents to 10 ± 0.3 %. The TOC/TN ratios are higher than those of the other macrophytes: 17.8 ± 1.2 %. High carbonate contents of >70 % are common for charophytes in freshwater systems (see Apolinarska et al., 2011 and references therein). Królikowska (1997) reports calcium contents of 234 mg g^{-1} and 246 mg g^{-1} dry weight for *Chara tomentosa* and *Chara contraria* from a shallow Polish lake, or carbonate contents of 59 ± 4 % and 61 ± 2 %, respectively, assuming that all calcium is present as carbonate, close to our value for *Chara tomentosa* (61.5 %).

3.1.3 Sediments

Carbonate contents of the Lz1120 sediments range from 20 to 56 % and TOC from 1.5 to 2.3 % (Fig. 2). The carbonate record of Lake Ohrid appears to be controlled by temperature and

terrestrial run-off, with precipitation and production of largely endogenic carbonate occurring mostly during the summer (Wagner et al., 2008; Vogel et al., 2010) and relying on the supply of calcium ions and nutrients from the catchment. Accordingly, minima in sedimentary carbonate represent periods of drier and cooler climate. Between 8.8 and 7.9 ka, the Lz1120 carbonate record features two pronounced minima at around 8.3 and 8.15 ka that appear to correspond to phases of the 8.2 ka event (Fig. 2) as documented in sediment records from the North Atlantic (MD99-2251) and Greenland ice cores (GISP 2; Ellison et al., 2006). TOC decreases from 2.3 to a minimum of 1.5 % at *ca.* 8.3 ka and remains below 2 % thereafter. Carbonate and organic carbon contents appear to co-vary apart from in the three youngest samples between 8.05 and 7.9 ka and the sample separating the two carbonate minima at 8.23 ka. In the remaining samples, TOC and carbonate correlate closely ($R^2 = 0.86$). TOC/TN ratios range from 7 to 10.7, with values > 10 between 8.9 and 8.3 ka (average: 10.5 ± 0.2) and < 10 thereafter (average: 8.1 ± 0.7). Higher TOC/TN values generally coincide with higher organic carbon contents ($R^2 = 0.91$). When TOC is plotted against TN, however, the samples are clearly separated into two groups (Fig. 2, insert): samples from before and after 8.3 ka show similar concentrations of TN but significantly different concentrations of TOC, with lower TOC relative to TN after 8.3 ka. Before 8.3 ka, there is a strong linear relationship between TOC and TN ($R^2 = 0.94$), with a y-axis intercept of 0.057 % TN for 0 % TOC, suggesting that around a quarter of the nitrogen is inorganic. The amount of organic nitrogen (N_{org}) that results from this assumption leads to an average TOC/N_{org} ratio of 13.6 for the samples before 8.3 ka. After 8.3 ka, the correlation between TOC and TN disappears, implying that higher nitrogen concentrations do not generally result from the supply of nitrogen-rich OM, but from variable amounts of inorganic nitrogen (N_{in}) such as ammonium (NH_4^+). The implications for the interpretation of the TOC/TN record of Lz1120 are discussed below (section 4.1 Assessing aquatic and terrestrial OM supply).

3.2 Lipid Biomarkers

3.2.1 Leaf litter, topsoils, grasses

3.2.1.1 Biomarker content and composition

Leaf litter under both types of vegetation, low-altitude oak-dominated forest (sites SN, TP) and high-altitude beech forest (GN) contain very similar total amounts of lipids: *ca.* 458 ± 63 μg g^{-1} dry weight (n = 4). Lipid contents of the topsoils are approximately 6x lower, with an average of 82 ± 37 μg g^{-1} dry weight (n = 6) for the Leptosols. Grasses revealed the by far highest

amounts of lipids, containing about 1.4 ± 0.2 mg g^{-1}. Both high- and low-altitude leaf litter do not contain visible contributions from grass but are dominated by leaf fragments of the main tree species, i.e. small oaks (SN, TP) and beeches (GN). Leaf litter can be regarded as a main contributor to soil organic matter and the compositional changes occurring during leaf litter incorporation into the topsoil are described in the following.

There are compositional differences in the major lipid compound classes, between leaf litter and corresponding topsoils and between high-altitude and low-altitude sites (see Figure 3 and Table 2). Lipids of the near-shore oak-dominated leaf litter (sites SN, TP) are dominated by saturated FAs that account for 37 ± 2 $\%_{lipids}$, while the other major compound classes, i.e. mono- and poly-unsaturated fatty acids (MUFAs, PUFAs) and sterols, vary between 13 and 18 $\%_{lipids}$. PUFAs and MUFAs are present in similar proportions, 9.3 ± 1.4 and 7.1 ± 0.3 $\%_{lipids}$, respectively. In high-altitude beech-dominated litter (GN), FAs make up only 23 $\%_{lipids}$ and the proportions of MUFAs and PUFAs are considerably higher, 21 and 17 $\%_{lipids,}$ respectively. Sterols make up 24 $\%_{lipids,}$ with β-sitosterol (24β-ethylcholest-5-en-3β-ol; 17 $\%_{lipids}$) and β-sitostenone (24β-ethylcholest-4-en-3−one; 5 $\%_{lipids}$) together accounting for 96 % of the total sterols. β-Sitosterol is also the dominant sterol in low-lying oak-dominated litter (SN, TP; 10.7 ± 0.3 $\%_{lipids}$). In SN leaf litter, the pentacyclic triterpenoid taraxasterol (urs-20(30)-en-3β-ol) is also abundant (11.3 ± 1.1 $\%_{lipids}$, on average). It also occurs in leaf litter at site TP (7.1 $\%_{lipids}$) but is absent in the litter and soil samples at GN. Details on the potential source of this compound are provided in Supplement 2.

The proportions of the major compound classes in the topsoils differ from those observed in the overlying leaf litter. While the relative amounts of FAs remain similar (35 ± 4.7 $\%_{lipids}$), the MUFAs increase in relative abundance from leaf litter to soil from ca. 7 to 20 $\%_{lipids}$ at TP and SN and from 12 to 25 $\%_{lipids}$ at GN with *n*-$C_{18:1}$(*cis*-9) being the dominant MUFA. Small amounts of *n*-$C_{20:1}$(*cis*-9) MUFA might derive from collembolans (springtails), small detritivorous and microbivorous hexapods feeding on nematodes as well as fungi in leaf litter (Ruess et al., 2005). PUFAs, on the other hand, decrease from leaf litter to soil in relative abundance at all sites, most notably at GN, to between 5 and 6 $\%_{lipids}$. The dominant PUFA in all leaf litter samples is $C_{18:2}$. Leaf litter at sites TP and GN contained $C_{18:2}$ only, albeit in higher amounts than at site SN with 8 $\%_{lipids}$ (TP) and 17 $\%_{lipids}$ (GN) *vs.* 8 $\%_{lipids}$ (SN). $C_{18:3}$ occurs only at site SN where it accounts for 2 $\%_{lipids}$. For $C_{18:2}$, this pattern is also preserved in the corresponding soils. The proportion of PUFA is also higher in the soil sample affected by white rot (TP-F), which is the only soil sample where we detected a small amount of $C_{18:3}$ and, notably

eicosapentaenoic acid (EPA; $C_{20:5}$). The latter probably has a fungal source as some soil fungi such as *Mortierella alpina* or plant pathogens such as *Pythioum ultimum* are known to biosynthesize EPA (Shimizu et al., 1988; Ghandi and Weete, 1991).

The higher amounts of *n*-$C_{18:1}$(*cis*-9) MUFA in the soils (58 % of all unsaturated FAs at TP and SN, up to 71 % at GN) are likely to derive from root suberin in addition to cutin. Besides *n*-C_{16} FA and *n*-C_{18} FA, the *n*-$C_{18:1}$ MUFA is one of the main monomers synthesized *de novo* to form both cutin and suberin (Mertz and Brutnell, 2014).

The relative amount of steroids also decreases from leaf litter to soil. At GN, their relative abundance nearly halves compared to the leaf litter (24 to 13 $\%_{lipids}$). At TP and SN the decrease is marginal, from 14 to 11 $\%_{lipids}$. The main sterol in both leaf litter and topsoils is β-sitosterol accounting for 60 to 80 % of the total sterols.

A notable difference in the total lipid composition between leaf litter and underlying soils is the significantly higher proportion of *n*-alcohols in the latter. *n*-Alcohols account for 3.1 ± 0.3 $\%_{lipids}$ in leaf litter at SN and TP and 6.3 $\%_{lipids}$ at GN. However, their proportions in the soils are 4 and 2.5-times higher, respectively.

Minor compound classes present in leaf litter and soils include hydroxy acids, *n*-alkanes and branched FA. Hydroxy acids (OH-FA) are more abundant in leaf litter at SN (3.9 ± 1.0 $\%_{lipids}$) than TP (2.1 $\%_{lipids}$) and GN (0.9 $\%_{lipids}$). At SN, relative concentrations of α-hydroxy acids (α-OH-FA) are *ca.* 10x higher than ω-hydroxy acids (ω-OH-FA absent in TP sample), while leaf litter at GN is dominated by ω-OH-FA. However, soils from all sites show similar relative amounts of ω- and α-OH-FAs (5 – 8 $\%_{lipids}$). The substantial increase in the amounts of ω-OH-FAs from litter to soil by an order of magnitude and the chain-length range from C_{16} to C_{26} (site GN: C_{16} - C_{28}) suggest their source is suberin, a protective biopolymer mainly found in the periderm of plant root and stem tissues, but also in bundle sheaths of grasses (Pollard et al., 2008; Molina et al., 2006). *n*-Alkanes account for 5 - 7 $\%_{lipids}$ in the leaf litter samples, but only 3 $\%_{lipids}$ in the topsoil samples at all three sites. The proportions of branched FAs increase by an order of magnitude from 0.2 to 2.7 $\%_{lipids}$ and 0.1 to 1.7 $\%_{lipids}$ at SN and GN, respectively; they are absent in leaf litter at TP. There is a difference in the composition of branched FAs between litter and topsoil samples. While the litter samples mainly contain *iso*- and *anteiso*-branched C_{15} FA as well as *iso*-C_{16} FA, the topsoil samples additionally contain *iso*- and *anteiso*-C_{17} FA as well as 10-methyl hexadecanoic acid. Finally, α,ω-dicarboxylic acids (α,ω-DiFAs) occur almost exclusively in the soil samples, the exception being one leaf litter sample from site SN where α,ω-DiFAs were detected in small amounts (0.3 $\%_{lipids}$). In all other leaf

litter samples, α,ω-DiFAs were not detectable. By contrast, α,ω-DiFAs accounted for 1 $\%_{lipids}$ (SN) and 0.7 $\%_{lipids}$ (TP) in the low-altitude soils and for 0.4 $\%_{lipids}$ in the high-altitude sample (GN). As for the ω-OH-FAs, their occurrence with a chain length range of C_{16} - C_{26} strongly suggest that the main source of the α,ω-DiFAs is suberin.

Compared to leaf litter and topsoils, TLEs of grasses show a far less complex composition, with the alkyl compounds being dominated by only a few individual biomarkers (Fig. 3b). Apart from representing single species samples, a further reason for this may be that the grasses were not sampled as litter from the ground but as upright, although dry, late-season plant matter. This means that degradation was at the earliest stage and contribution of lipids from degrading organisms likely minimal, which is also indicated by the absence of branched FAs. The grasses show significantly higher amounts of PUFAs (29 ± 2 $\%_{lipids}$) and a far lower proportion of MUFAs (3 ± 1 $\%_{lipids}$) than leaf litter and topsoils. The dominant PUFA and, in fact, the most abundant individual compound is $C_{18:3}$(*cis*-9), which accounts for 20% of the total lipids. The proportions of short-chain FAs, *n*-alkanes and *n*-alcohols are also elevated, which is due to three compounds: C_{16} FA, the C_{31} *n*-alkane and C_{26} *n*-alcohol. Together with $C_{18:3}$(*cis*-9) PUFA and $C_{18:2}$(*cis*-9) PUFA (8 $\%_{lipids}$) these compounds account for about 70 % of the total lipids in the grass samples. Long-chain FAs account for 6 $\%_{lipids}$, while the main sterol, sitosterol, contributes 4 $\%_{lipids}$ and β-amyrin 5 $\%_{lipids}$. The amounts of α- and ω-OH-FAs are very low: 0.3 $\%_{lipids}$ and 0.2 $\%_{lipids}$, respectively, comparable to leaf litter samples. Saturated C_{24} and mono-unsaturated $C_{24:1}$ α-OH-FA are the only α-OH-FA detected. In contrast to the leaf litter samples, the ω-OH-FAs in the grasses are dominated by C_{24} and C_{22} ω-OH-FA (~80 % of total ω-OH-FAs), with C_{16} ω-OH-FA accounting for the rest.

3.2.1.2 Carbon number distributions of *n*-alkyl compounds

Chain length distributions of *n*-alkyl compounds (saturated *n*-FA, *n*-alcohols, α-/ω-OH-FA and *n*-alkanes) show subtle differences between high-altitude beech-dominated (GN) and near-shore low-lying oak-dominated leaf litters (TP, SN), and between litter and underlying soils, which could nevertheless be useful in tracing the provenance of OM in the lake sediments. The main *n*-alkanoic acid in both leaf litter and topsoils is C_{16} FA. At TP and SN, the leaf-litter samples have a trimodal chain-length distribution, with long-chain and mid-chain length C_{30} FA and C_{22} FA, respectively, being the other modes (Fig. 3b). This trimodal distribution is more pronounced at GN, with C_{22} FA being the mid-chain mode, while the C_{28} FA is the long-chain mode. The proportion of mid-chain compounds, in particular, the amounts of C_{22}- and C_{24}-FA

increase significantly in the low-altitude topsoils (TP, SN), relative to the overlying leaf litters (Figs. 3, 4a). Accordingly, the average chain-length (ACL; Eq. 1) between C_{22} and C_{26}-FA shifts from 24.2 to 23.6 between litter and topsoil.

$$ACL = (22*C_{22}+24*C_{24}+26*C_{26})/(C_{22}+C_{24}+C_{26}). \quad (1)$$

At GN, by contrast, ACL increases from leaf litter to soil from 23.1 to 23.9, suggesting a relative decrease of the mid-chain compounds.

The *n*-alcohols show pronounced bimodal distributions of mid- and long-chain compounds in leaf litter and topsoils. In GN leaf litter C_{20} and C_{28} *n*-alcohols (OH) dominate, while the GN topsoil peaks at C_{22} and C_{28} OH. In leaf litter at TP and SN C_{24} OH and C_{30} OH are the dominant mid- and long-chain modes, respectively. The bimodal character in both litter and topsoils results from the *n*-alcohols deriving from two major sources: (a) cuticular wax esters, with chain lengths of C_{26} and C_{28} or, in some cases, C_{30} and C_{32} (e.g., Samuels et al., 2008 and references therein) and (b) suberin, with C_{22} and C_{24} OHs being the main monomers (Molina et al., 2006). The source of C_{20} OH in high-altitude leaf litter remains uncertain. Note that, although suberin is a major bio-polyester in root material, it is not exclusively present in roots but also forms part of other tissue types and is therefore present in leaf litter. In contrast to leaf litter and soils, the *n*-alcohols distribution in the grasses is unimodal and dominated by C_{26} OH, accounting for 85 % of the *n*-alcohols.

Comparison of the relative abundances of individual FAs and OHs in the two samples from site TP provides clues as towards the effect of white rot on chain-length distribution (Fig. 4b, c). Fungal biomass contribution and breakdown of plant material appears to increase the proportion of short-chain C_{16} and C_{18} FA and of C_{30} FA (Fig. 4b) as well as the proportions of C_{30} and C_{32} OH, relative to mid- and short-chain OHs (Fig. 4c).

Chain lengths of ω-hydroxy acids in all leaf litter samples range from 16 to 22 but are dominated by short-chain C_{16} ω-OH-FA (Fig. 3). This distribution is typical for cutin monomers (Matzke and Riederer, 1991), the lipid polymer being part of the protective hydrophobic layer of plant cell walls (Pollard et al., 2008). By contrast, the dominance of C_{24} and C_{22} ω-OH-FAs in the grasses point towards suberin incorporated in bundle sheaths as their main source (Mertz and Brutnell, 2014). In the topsoil samples, ω-OH-FAs range from C_{16} to C_{26}, the dominant compounds being C_{22} ω-OH-FA and C_{24} ω-OH-FA, which together account for ~ 50 % of the total ω-OH-FA in the TP and SN soils and 65 % of the total ω-OH-FA in the GN topsoil, where the proportion of C_{24} ω-OH-FA is higher. C_{22} ω-OH-FA and C_{24} ω-OH-FA typically are the

main ω-OH-FAs in suberin from *Quercus robur L.* and *Fagus sylvatica* L. (Matzke and Riederer, 1991) or *Quercus suber* (Graça and Santos, 2007) and a similar ω-OH-FA distribution has been reported for an oak forest soil by Nierop et al. (2005). α-Hydroxy acids range from C_{22} to C_{24} in GN litter and C_{15} to C_{26} in TP and SN litters, with C_{24} α-OH-FA being dominant, followed by C_{22} α-OH-FA; there are also high abundances of odd-numbered α-OH-FAs, C_{23} and C_{25} in particular. The distribution of α-OH-FAs is similar in the underlying soils.

n-Alkanes in the leaf litter samples (sites TP, SN, GN) range from *n*-C_{23} to *n*-C_{31} and reveal a clear compositional difference between high-altitude beech-dominated and low-lying oak-dominated leaf litters (Fig. 3): the main *n*-alkane at SN and TP is *n*-C_{29} while litter at GN almost exclusively contains *n*-C_{27}, accounting for 90 % of the total *n*-alkanes. Thus, the average chain length (*n*-C_{27} to *n*-C_{31}; $ACL_{27\text{-}31}$) is 28.9 at TP and SN and 27.1 for the GN site. In the grasses, the *n*-alkanes range from *n*-C_{23} to *n*-C_{33}. The distribution is dominated the C_{31} *n*-alkane, which accounts for 75 % of the total *n*-alkanes, while *n*-C_{29} accounts for 15 %, giving an $ACL_{27\text{-}33}$ value of 30.8. In the soils, *n*-alkane chain lengths also range from C_{23} to C_{33}. The dominant *n*-alkane in all low-altitude topsoil samples is *n*-C_{29}. Compared to the leaf litter samples, the ratio of *n*-C_{29} relative to *n*-C_{31} (C_{29}/C_{31} = *n*-C_{29}/(*n*-C_{29} + *n*-C_{31})) decreases from 0.78 in leaf litter to 0.63 in the corresponding topsoils at SN, TP. In the sample affected by white rot (TP-F) this ratio is even lower (0.56). In GN leaf litter *n*-C_{31} was absent. Notably, *n*-alkanes in the underlying topsoil at site GN show a bimodal distribution, with *n*-C_{31} and *n*-C_{27} being the main *n*-alkanes and the C_{29}/C_{31} ratio being 0.36. Since grass is currently neither present at the sampling site nor a visible part of the leaf litter this suggests that it has contributed to the soil lipids in the past, i.e. before the establishment of the present mature beech canopy.

The α,ω-dicarboxylic acids (α,ω-DiFAs) in the topsoils reveal chain-length distributions that are similar to those of the ω-hydroxy acids, i.e. bimodal, with peaks at C_{16} and C_{22}. Notably, the increased proportions of C_{24} relative to C_{22} ω-OH-FA that is observed in the high-altitude soil when compared to the low-lying soils is reflected in the DiFA chain-length distribution confirming that both ω-OH-FAs and α,ω-DiFAs most likely derive from suberin. The presence of α,ω-DiFAs in roots but not in leaves and stems of land plants, led Mendez-Millan et al. (2011) to conclude that α,ω-DiFAs represent subterranean biomass.

3.2.2 Subsoil (Terra Rossa)

As for TOC, the subsoil samples, DG-SS and DG-SS 2, contain the lowest amounts of lipids: 1.3 μg g^{-1} and 2.3 μg g^{-1} dry weight, respectively, which is an order of magnitude lower than

the lipid content of the topsoils. Notably, there is a significant difference in lipid composition between the sample collected from the exposed surface of the soil profile (DG-SS) and the sample collected from slightly greater depth, for simplicity distinguished as 'weathered' and 'un-weathered' subsoil samples. The lipid composition of the un-weathered sample closely resembles those of the low-altitude topsoils, but appears to contain significantly lower amounts of sterols (3 $\%_{lipids}$ *vs*. 11 $\%_{lipids}$ in the topsoils). The proportion of ω-OH-FAs (2 $\%_{lipids}$) is half compared to the topsoils while α,ω-DiFAs were not detected, which is consistent with the low visual abundance of root material. The amount of *n*-alcohols, on the other hand, is slightly enhanced, with 20 $\%_{lipids}$ compared to 15 $\%_{lipids}$ in the low-altitude topsoils. While these differences set the un-weathered subsoil sample slightly apart from the topsoils (also see Fig. 5), some of the variation may arise from the different extraction method applied (microwave), the different analytical equipment (see methods) and, the semi-quantitative data determined for this sample.

The lipid composition of the weathered sample contrasts with the un-weathered subsoil sample and the topsoils. Almost half of the lipids (48 %) are *n*-alcohols, followed by unsaturated FAs (20 $\%_{lipids}$), saturated FAs (15 $\%_{lipids}$) and steroids (11 %), with β-sitosterol being the dominant sterol (9 $\%_{lipids}$). Minor components include *n*-alkanes (3.5 $\%_{lipids}$), OH-FAs (1.4 $\%_{lipids}$) and branched FAs (0.6 $\%_{lipids}$). As in the un-weathered subsoil sample, α,ω-DiFAs were absent. Notably, almost all compounds with carboxylic groups (i.e. fatty acids) appear depleted relative to *n*-alcohols, *n*-alkanes and sterols when compared to the un-weathered subsoil sample and the topsoils (Fig. 3), with the exception of PUFAs. Short-chain FAs and MUFAs appear less severely depleted than mid-chain and long-chain saturated FAs. The fact that the amount of unsaturated FAs in the weathered subsoil sample is higher than the amount of saturated FAs suggests that a large proportion of the unsaturated FAs in this sample is produced *in situ*, with the dominant *n*-$C_{18:1}$(*cis*-9) MUFA (45 % of unsaturated FA) most likely deriving from micro-organisms.

The dominant *n*-alcohol in the un-weathered Terra Rossa sample is C_{24} OH. In the weathered/degraded Terra Rossa sample, the *n*-alcohols show a bimodal chain-length distribution peaking at 28 and 24, with C_{28} OH accounting for 25 % and C_{24} OH for 19 % of all *n*-alcohols. The saturated *n*-alkanoic acids are dominated by short-chain (C_{16}, C_{18}) and mid-chain (C_{24}, C_{22}) FAs, with C_{16} FA in the degraded Terra Rossa sample showing the highest relative amounts compared to the longer chain *n*-alkanoic acids in all other soil samples (36 $\%_{FA}$ *vs*. 20 $\%_{FA}$). As for the unsaturated compounds, this may reflect *in-situ* contributions from

micro-organisms. The chain-length distributions of ω-OH-FAs and *n*-alkanes in both Terra Rossa samples appear very similar. As in the topsoils, the dominant ω-OH-FA is the C_{22} compound, suggesting suberin as their main source. The only α-OH-FA is the C_{24} homologue. Chain lengths of the *n*-alkanes range from 23 to 33, with *n*-C_{31} and *n*-C_{29} being dominant compounds and being present in almost equal amounts (C_{29}/C_{31} = 0.49) accounting for 68 % of the total *n*-alkanes. This suggests higher *n*-alkane input from grasses (*n*-C_{31}) compared to the low-altitude topsoils under oak-dominated vegetation (*n*-C_{29}).

The differences between the un-weathered and the weathered/degraded Terra Rossa sample may be indicative of the compositional changes soil lipids may undergo when the soil dries out frequently and becomes aerated. *n*-Alcohols appear selectively preserved relative to *n*-alkanoic acids to an even greater extent than during the degradation process in the transition from plant litter to soil organic matter. As leaf litter and weathered Terra Rossa have highest and lowest concentrations of total organic carbon and total lipids, these may represent the end-members of continuous plant lipid transformation determined for this initial survey of lipid biomarker sources in the SE Ohrid Basin. However, the projection towards the weathered/degraded stage is solely based on one pair of samples and potential controls such as soil moisture, oxygen availability, pH or temperature are unknown.

3.2.3 Macrophytes

Submerged macrophytes contain between 0.2 and 1.1 mg g^{-1} dry wt. of lipids, which is in the same range as the leaf litter. Highest concentrations are found in *Cladophora* spp. (1.07 ± 0.33 mg g^{-1}). *Potamogeton* spp. and *Characeae* spp. contain lower amounts (0.22 ± 0.13 mg g^{-1} and 0.53 ± 0.34 mg g^{-1}, respectively), which is likely due to a higher proportion of supportive tissue in the latter and, in case of *Characeae* spp., carbonate incrustations. Sub-aquatic and sub-aerial parts of *Phragmites* spp. were analysed separately; lipid concentrations in the root and stem (sub-aquatic) are in a similar range to the submerged macrophytes (0.44 and 0.34 mg g^{-1}, respectively). The highest lipid concentrations were in the (sub-aerial) *Phragmites* spp. leaves (1.7 ± 0.3 mg g^{-1}).

The macrophyte lipids are dominated by unsaturated and saturated fatty acids that, together, account from 64 $\%_{lipids}$ (*Potamogeton*) to 95 $\%_{lipids}$ (*Cladophora*). Unsaturated FAs dominate, with slightly higher proportions in *Phragmites* sp. leaves (59 $\%_{lipids}$) and *Characeae* spp. (57 $\%_{lipids}$) and lower proportions in *Potamogeton* spp. (41 $\%_{lipids}$) and the *Phragmites* sp. roots (48 $\%_{lipids}$). The $C_{18:2}$ PUFA is the dominant unsaturated FA in the *Phragmites* sp. stem (21 $\%_{lipids}$)

and roots (27 $\%_{lipids}$), while in the leaves it is the $C_{18:3}$ PUFA (48 $\%_{lipids}$). The latter also dominates in *Potamogeton* sp. (23 $\%_{lipids}$). $C_{16:1}$ (*cis*-9) MUFA is most abundant in *Cladophora* sp. (29 $\%_{lipids}$), which also contains the highest amounts of $C_{18:1}$ (*cis*-9) MUFA (8 $\%_{lipids}$). $C_{16:1}$ (*cis*-9) MUFA was also present in traces in *Phragmites* stems (13 $\%_{lipids}$) but absent in roots and leaves.

Saturated *n*-alkanoic acids have highest relative abundance in *Cladophora* spp. (45 $\%_{lipids}$) and lowest concentrations in *Phragmites* sp. leaves (15 $\%_{lipids}$). They are dominated by C_{16} FA, which accounts for about three quarters of all saturated FA, with lower values only in the *Phragmites* stem and root (67 and 55 % of saturated FA, respectively), where there are slightly elevated amounts of C_{20} FA and C_{24} FA. The remaining compounds are mainly sterols, accounting for 3.6 - 12 % of the total lipids and with highest amounts in *Potamogeton* spp. (33 $\%_{lipids}$) and elevated proportions in the *Phragmites* sp. root and stem (12 and 11 $\%_{lipids}$, respectively). The main sterol present in all samples is β-sitosterol followed by stigmasterol (absent in *Cladophora* sp.). Small amounts of cholesterol (0.1 – 2.3 $\%_{lipids}$) were detected in all samples except for the *Phragmites* sp. stem.

n-Alcohol concentrations are generally very low (< 1.5 $\%_{lipids}$), with the exception of the *Phragmites* leaves where long-chain C_{28} OH and C_{30} OH account for 19 % of the total lipids and likely to derive from cuticular waxes (*Samuels* et al., 2008). In most samples, hydroxy acids are either absent (*Characeae* spp.) or detected in small amounts of less than 0.4 $\%_{lipids}$ (*Cladophora* spp., *Potamogeton* spp., *Phragmites* spp. leaves). Higher amounts were detected in the *Phragmites* stem and root: 1.4 and 3.5 $\%_{lipids}$, respectively. Thus, *Phragmites* spp. roots, in particular, are a potential macrophytic source for sedimentary α- and ω-OH-FA. As in leaf litter, the amount of α-OH-FA is higher than the amount of ω-OH-FA in all samples where OH-FAs are present. The α-OH-FA chain-length distribution is bimodal in all *Phragmites* spp. tissues, with α-C_{24} OH-FA dominant followed by the α-C_{20} OH-FA. In *Potamogeton* spp. and *Cladophora* spp. α-C_{24} OH-FA dominates, followed by α-C_{22} OH-FA, and only *Potamogeton* spp. contained a small amount of α-C_{16} OH-FA. The main ω-OH-FA is the ω-C_{24} OH-FA compound, which is present in *Potamogeton* spp. (traces) and the *Phragmites* spp. stem and root samples (0.2 and 0.9 $\%_{lipids}$, respectively). The latter also contain ω-C_{22} OH-FA, ω-C_{26} OH-FA and ω-C_{20} OH-FA. Although ω-C_{16} OH-FA was not detected, this distribution suggests that the ω-OH-FA in *Phragmites* spp. also derive from suberin, which is a known biopolymer in *Phragmites australis* (Soukup et al., 2007).

n-Alkanes were either absent (*Characeae*, *Phragmites* root and stem) or found in very low

concentrations, e.g., 0.25 $\%_{lipids}$ in *Phragmites* sp. leaves (dominated by *n*-C_{29}, ACL_{23-31} = 28.3) and 1 $\%_{lipids}$ in *Potamogeton* spp. (dominated by *n*-C_{25}, ACL_{23-31} = 25.6). The differences in chain-length distributions between *Potamogeton* spp. and the leaves of *Phragmites* sp. are consistent with observations made by Ficken et al. (2000) on the *n*-alkane composition of macrophytes from lakes on Mt. Kenya. There, *n*-alkanes produced by submerged macrophytes were dominated by mid-chain compounds (C_{23}, C_{25}), whereas *n*-alkanes produced by emergent macrohpytes were dominated by long-chain compounds (> C_{29}) as in land plants. Ficken et al. (2000) suggested a proxy (eq. 2) to determine varying contributions from submerged and emergent macrophytes defined as:

$$P_{aq} = (n\text{-}C_{23} + n\text{-}C_{25})/(n\text{-}C_{23} + n\text{-}C_{25} + n\text{-}C_{29} + n\text{-}C_{31}), \quad 2)$$

with values between 0.1 and 0.4 suggested to correspond to emergent macrophytes and values between 0.4 and 1 to submerged macrophytes. Accordingly, the P_{aq} value of 0.78 for *Potamogeton perfoliatus* in Lake Ohrid matches the values reported by Ficken et al. (2000) for *Potamogeton thumbergii* (0.73 and 0.92) while the value for *Phragmites* leaves (0.24) does fall in the suggested range for emergent macrophytes.

3.2.4 Lake water particulate organic matter (POM)

The total lipid composition of the lake particulate organic matter (POM) was not determined, as extracts had previously been separated into non-polar and polar fractions. Here we report the results from analyses of the polar components that include fatty acids, alcohols and steroids of the water filtrates taken at Co1202 and the DEEP site. At both sites, these were dominated by short-chain FAs (C_{16} FA and C_{18} FA), accounting for 26 % of polar lipids (% $_{polar\ lipids}$), C_{18} MUFAs (21 % $_{polar\ lipids}$) and some sterols (20 % $_{polar\ lipids}$), mainly cholesterol (8 % $_{polar\ lipids}$), sitosterol (6 % $_{polar\ lipids}$) and stigmastanol (5 % $_{polar\ lipids}$). Minor polar compounds were $C_{18:2}$ PUFA (7 % $_{polar\ lipids}$), *n*-alcohols (4 % $_{polar\ lipids}$), mid- and long-chain *n*-FAs (3 % $_{polar\ lipids}$), branched FAs (2 % $_{polar\ lipids}$) and OH-FAs (1.3 % $_{polar\ lipids}$). The proportion of $C_{18:2}$ PUFA increased with water depth at site Co1202, from 4 to 12 $\%_{polar\ lipids}$ suggesting that most recently produced labile OM had sunk to depth at the time of sampling. Long-chain *n*-alkanoic acids and *n*-alcohols were detected in trace amounts, only. Notably, the *n*-alcohol distribution in the water filtrates suggests a bacterial source (Fig. 3b). Yang et al. (2014) observed a very similar *n*-alcohol distribution in *Erythrobacter* sp. isolated from the South China Sea. Albaigés et al. (1984) noted that C_{18} OH and C_{22} OH dominated the free and bound *n*-alcohol fraction in lacustrine sediments from Spain and assigned them to bacterial inputs.

3.2.5 Sediments

Total lipid contents of sediments from core Lz1120 range from 7.6 to 79 μg g^{-1}, with an average of 32 μg g^{-1}. Notably, the amounts of lipids normalized to TOC correlate closely to the lipid concentrations of the dry samples (R^2 = 0.92), suggesting that, for example, a drop in TOC corresponds to a drop in the lipid concentration of the OM supplied.

With regard to the major lipid compound classes, there are clear compositional differences between the samples with low and high carbonate contents, i.e. between the samples dated to 8.29, 8.17 and 8.11 ka, representing cool, dry episodes associated with the 8.2 ka event, and those from the rest of the investigated time slice. In particular, the proportion of *n*-alkanoic acids is significantly lower in the samples from 8.17 and 8.11 ka compared to the high-carbonate samples (16 $\%_{lipids}$ *vs*. 40 $\%_{lipids}$, on average), while the proportions of sterols and *n*-alcohols are higher (24 $\%_{lipids}$ *vs*. 15 $\%_{lipids}$ and 37 $\%_{lipids}$ *vs*. 21 $\%_{lipids}$, respectively). In the sample from 8.29 ka, both *n*-alkanoic acids (27 $\%_{lipids}$) and *n*-alcohols (18 $\%_{lipids}$) are present in lower amounts but similar proportion relative to each other as in the high-carbonate samples (see 'Supplementary Data').

The concentrations of OH-FAs range from 1 to 6.7 $\%_{lipids}$, with averages of 3.3 and 4.6 $\%_{lipids}$ for low- and high-carbonate samples, respectively. These values are similar to those of the topsoil (average: 5.7 $\%_{lipids}$) and leaf litter samples (average: 2.5 $\%_{lipids}$), but significantly higher than in the macrophytes, with the exception of *Phragmites* stems (1.4 $\%_{lipids}$) and roots (3.5 $\%_{lipids}$). As in the soils, but unlike the leaf litters and macrophytes, the majority of the OH-FA in the sediments are ω-OH-FA, except at 8.17 ka where α-OH-FAs dominate (2.1 *vs.* 1.1 $\%_{lipids}$). Branched FAs account for 0.9 $\%_{lipids}$ on average in the high-carbonate samples. Of the low-carbonate samples, the sample at 8.11 ka shows an exceptionally high value of 2.5 $\%_{lipids}$ while the sample at 8.29 ka contains the lowest amount of branched FAs of all sediment samples (0.2 $\%_{lipids}$). *n*-Alkane concentrations are more than twice as high in the three low-carbonate samples than in the high-carbonate samples (3.3 $\%_{lipids}$ *vs*. 1.5 $\%_{lipids}$).

In contrast to macrophytes, leaf litter and soils, the early Holocene sediment samples contain almost no unsaturated fatty acids. The only detectable MUFA is $C_{18:1}$(*cis*-9) accounting for about 0.1 to 0.5 $\%_{lipids}$, if present. PUFAs are not preserved in the sediments. On the other hand, the C_{21} - C_{33} methyl ketones were identified only in the sediments (0.1 to 1.1 $\%_{lipids}$; with the exception of C_{27} in SN leaf litter).

Other compounds account for 17 % of the total lipids in both high- and low-carbonate samples.

While most of these compounds are minor components (<2 $\%_{lipids}$), long-chain diols, lanosterol and 17β(H),21β(H)-bis-homohopanoic acid, on average, account for 3, 4 and 8 $\%_{lipids}$, respectively (Table 2). Like many of the minor compounds they were absent in the terrestrial samples and imply production in the water column or sediment. Further details and references as towards potential sources are provided Supplement 2.The distributions of *n*-alkanoic acids reveal significant differences in chain length between samples with high and low carbonate contents. The input of long-chain *n*-alkanoic acids from terrestrial plant waxes appears strongly reduced relative to short-chain C_{16}-FA and mid-chain C_{22}- and C_{24}-FAs during periods of reduced or diluted carbonate sedimentation. By contrast, the chain-length distribution of *n*-alcohols shows rather subtle differences that, in fact, result from the depletion of long-chain *n*-alcohols in one sample, only (8.29 ka, Fig. 5). *n*-Alkanes show a unimodal distribution peaking at *n*-C_{29} and do not show a significant difference between high- and low-carbonate samples. *n*-C_{31} is slightly more abundant relative to *n*-C_{29}. The ratio of these two compounds (C_{29}/C_{31}) changes marginally, with values of 0.53 and 0.56 for low- and high-carbonate samples, respectively. These values match the value of the soil sample influenced by white rot (0.56) but would also result from a hypothetical 3:4 mix of low-altitude and high-altitude woodland topsoils, for example. The methyl ketone distribution is unimodal peaking at C_{27}. Methyl ketones appear to correlate with *n*-alkanes with regard to concentrations ($R^2 = 0.73$) and chain-length distribution ($R^2 = 0.94$). Microbial oxidation of *n*-alkanes has widely been suggested to explain the presence of methyl ketones in soils and sediments (e.g., van Bergen et al., 1998; Cranwell et al., 1987; Albaigés et al., 1984) and their increase in relative concentration with sediment depth (Cranwell et al., 1987). We did not find methyl ketones to be a lipid fraction in the soil samples and so we assume methyl ketones to be *in situ* products of microbial *n*-alkane breakdown in the lake sediments.

4 Implications for proxy records

The principle behind paleoenvironmental reconstructions using organic-geochemical proxies is that OM fluxes from aquatic and terrestrial sources within the catchment of a given sedimentary archive change in response to changes in regional or global climate. The main purpose of this study is to characterize the organic geochemical composition of potential sources of OM buried in sediments of Lake Ohrid in order to improve the interpretation of organic geochemical proxy data in reconstructions of hydrology-controlled environmental change in the Ohrid Basin. Our initial data set provides significant insight towards an improved interpretation of changes in

elemental (TOC/TN) and biomarker composition.

4.1 Assessing aquatic and terrestrial OM supply

4.1.1 Elemental composition of aquatic and terrestrial OM

TOC/TN ratios of bulk sediment are often used for assessment of aquatic and terrestrial contributions, with two values frequently applied in simple 2-endmember models: the Redfield Ratio of 6.6 for the aquatic/algal end member and the value of 20 for the terrestrial end-member (see, e.g., Meyers, 1994). As such, this approach appears of limited value in case of the Lz1120 sediments since there is a range of terrestrial OM of different levels of degradation present in the Ohrid Basin, reaching from plant litter to degraded (or "weathered") soil. This degradation gradient coincides with a substantial range of TOC/TN signatures, while plant matter/litter and soils are mobilized from a great variety of topographical settings (mountain slopes to flood plains) at different rates under specific circumstances. Furthermore, macrophytes have intermediate TOC/TN ratios of 15 ± 2.4 and produce a significant amount of biomass in the littoral, intercepting nutrient supply from terrestrial run-off and directly tapping into the sedimentary nutrient pool. For example, *Potamogeton perfoliatus* L. produces a biomass of 2,747 x 10^6 g y^{-1} (Talevska, 2006), while the stock of charophytes in Lake Ohrid is estimated at 10,500 x 10^6 g (Trajanovska et al., 2012). Lake Ohrid is oligotrophic and the biomass from the epilimnion may well be the smallest source of sedimentary OM. A maximum phytoplankton biomass of 131 mg m^{-3} in spring 2001 (Petrova et al., 2008) leads us to estimate a peak phytoplankton biomass in the upper 50 m of the water column of 2,300 x 10^6 g for the spring season (calculated for lake surface area of 358 km^2), which is comparable to the annual production of *Potamogeton perfoliatus* L. and a quarter of the standing stock of charophytes. This OM nevertheless represents high-value food for the lake's heterotrophic consumers and is likely largely remineralised in the water column.

Straightforward interpretation of the Lz1120 TOC/TN record based on the conventional two end members therefore is not possible, but the variability of the TOC/TN record does provide valuable information on changes in OM sources in the given paleoenvironmental context. Before 8.3 ka, the estimated TOC/N_{org} ratio is 13.6 (see section 3.1.3). This is higher than the value of Terra Rossa (11.6 ± 0.8), within the range of macrophytes (15 ± 2.4), and may result from any mix of these with leaf litter, topsoils and a small amount of algal material. Still, the majority of the nitrogen appears to be organic and to vary with organic carbon supply. After 8.3 ka, by contrast, TOC/TN ratios are very low, with a value of 7 at 8.05 ka close to the

Redfield ratio, suggesting that the majority of the sedimentary OM is of algal origin. This, however, is not supported by the lipid biomarkers, which are dominated by terrestrially-derived compounds throughout this period. An alternative explanation for the very low TOC/TN ratios in the Lz1120 sediments after 8.3 ka other than algal matter would be elevated amounts of inorganic nitrogen (N_{in}), for which there is some evidence (see also section 3.1.3). Within the terrestrial biome, the highest proportions of N_{in} are likely to be found in the clay fraction of mineral soils. This is due to the fact that ammonium (NH_4^+), in particular, can firmly bind to clay minerals and strongly influence the soil TN pool (e.g., Elmacı et al., 2002; Kothawala and Moore, 2009; Nieder et al., 2011). The low TOC/TN value for Terra Rossa is likely due to enhanced NH_4^+ levels. Soils developing from former lake sediment during lake level low stands are also likely to be charged with NH_4^+, in particular. It has frequently been reported that the fine fraction of mineral soils contains substantially higher amounts of ammonium than the coarser soil fractions (see Nieder et al., 2011 and references therein). Size fractionation during transport of eroded mineral soil may therefore result in the deposition of clay-rich material with TOC/TN ratios that are lower than those determined for the bulk soils *in situ*. Hence, the low TOC/TN ratios in Lz1120 sediments after 8.3 ka probably result from enhanced supply of clay-bound ammonium following mineral soil erosion rather than algal OM.

4.1.2 Aquatic and terrestrial biomarkers

Nearly all of the biomarkers preserved in the sediments can be ascribed to terrestrial OM, or have a non-specific source. The exception is dinosterol (from dinoflagellates), but this is a very minor component of the total lipid extracts (see Table 2, Supplement 2) and we consider it not to be a robust quantitative proxy for lake-derived OM.

The amount of lipids relative to the total OM of individual sources is likely to vary with environmental conditions. Thus, without knowledge on how representative individual biomarkers are for the size of the OM pool they derive from under specific hydrological conditions, assessments of relative contributions remain speculative unless they are backed up by other proxy data (e.g., elemental data, petrology, bulk isotopes).

4.2 Identifying the changing composition of terrestrial OM using lipid biomarkers

As evident from the pie diagrams in Figure 3, topsoils contain significantly higher amounts of *n*-alcohols (OHs) relative to *n*-alkanoic acids (FAs) than leaf litter. The weathered/degraded subsoil sample contains even higher proportions of OHs, accounting for almost half of the total

lipids, and very low amounts of saturated FAs. Thus, the ratio of sedimentary *n*-alkanoic acids over *n*-alcohols (FA/OH) is indicative of shifts in the relative amounts of moderately to strongly degraded OM from the major terrestrial sources, with higher ratios indicating higher proportions of plant litter-type OM relative to soil-type OM and lower ratios indicating enhanced contributions of stronger degraded OM from young topsoils and/or degraded subsoils. Since a significant amount of short-chain FAs, *n*-C_{16} FA in particular, may also derive from aquatic sources we consider the ratio of long-chain FA over long-chain OH of certain terrestrial origin (terr FA/OH) to be a more reliable measure. Notably, both ratios, FA/OH and terr FA/OH, correlate well in the sediment samples ($R^2 = 0.99$ for all samples, $R^2 = 0.95$ for high-carbonate samples) suggesting that aquatic sourced short-chain FAs are of minor importance.

The chain-length distributions in Figure 3 also reveal another important difference between leaf litter and soils. In the near-shore oak-dominated setting, the proportion of mid-chain FAs, C_{22} and C_{24} FA in particular, is higher relative to long-chain FAs ($\geq C_{26}$) in the topsoil and in Terra Rossa than in the corresponding leaf litter. The ratio of sedimentary mid-chain C_{22} and C_{24} FAs over long-chain C_{26}, C_{28} and C_{30} FAs should therefore be sensitive to soil OM supply from the near-shore setting. Similar increases in the proportion of mid-chain FAs in soils relative to litter have also been observed in topsoils of a high-altitude beech forest in Central France (Marseille et al., 1999) and of an evergreen oak forest in Central Spain (Almendros et al., 1996). In a forest topsoil from the surroundings of Lake Suwa in Japan (Matsuda and Koyama, 1977), which represents a setting very similar to the Ohrid Basin (mid-latitude, high-altitude intra-mountainous lake basin with moist forest biome), FA distributions in the topsoils are very similar to that at site SN ($R^2 = 0.90$), providing evidence that, generally, similar natural processes operate under similar environmental conditions. The chain-length distributions of the *n*-alcohols also reveal differences between the terrestrial sources (Fig. 3), which are most pronounced between high-altitude and low-altitude sources when expressed, for example, as the average chain length (ACL) of C_{24} to C_{28} OHs ($ACL_{24\text{-}28}$). The amounts of *n*-alcohols are significantly higher in soils than in leaf litter. Changes in $ACL_{24\text{-}28}$ in the sediments therefore potentially indicate variable contributions of soil OM from high-altitude beech-dominated or low-altitude oak-dominated habitats. *n*-Alkane distributions, finally, allow for a clear distinction between C_{27}-dominated and C_{29}-dominated leaf litter from the high and low-altitude forests, respectively. In the soils, the input of C_{31} *n*-alkane significantly increases (Fig. 3) and either reflects over-proportional incorporation of lipids from the grassy undergrowth or from an earlier, more open type of vegetation with higher proportions of grass. In the sediments, the amounts of *n*-C_{27} alkane are always significantly lower than the *n*-C_{29} alkane, suggesting that

contribution from the high-altitude habitat is minor. In fact, the *n*-alkane distribution of a hypothetical 1:1 mix of *n*-alkanes from low-altitude topsoil and weathered Terra Rossa correlates very well with the average sedimentary *n*-alkane distribution ($R^2 = 0.95$). Similarly, the high-proportion of the *n*-C_{31} alkane indicates that direct inputs of leaf litter from beech or oak are minor relative to soil OM. Two observations suggest that the popular application of specific *n*-alkane ratios (*n*-C_{27} or *n*-C_{29} relative to *n*-C_{31}) as proxies for the development of the vegetation (mainly for trees *vs.* grass/shrubs; e.g., Schwark et al., 2002, Zhang et al., 2006; Leider et al., 2013) is compromised in the Ohrid Basin: a) the bimodal *n*-alkane distribution in high-altitude topsoil and b) the similarity of the distributions in soils and sediments that distinctly differ from the pronounced unimodal leaf litter distributions. The ambiguity of proxies based on these ratios in the Ohrid Basin is highlighted by the fact that, for example, the C_{29}/C_{31} *n*-alkane ratio in both high-carbonate samples (warmer, more humid climate) and low-carbonate samples (cooler, more arid climate) are very similar, 0.56 and 0.53, respectively, even though the terrestrial biome certainly changed significantly through the pronounced climatic fluctuation of the 8.2 ka event.

MDS shows separation of sample types, namely soils, leaf litter, grasses, macrophytes, lake POM and sediments, which 1-way ANOSIM shows to be significantly different. The "weakest" differences are between subsoil and grass (R = 0.25, P = 33.3 %) and leaf litter and grass (R = 0.5, P = 20 %), they not being significantly different. This is also evident in the MDS plot (Figure 5), which shows that the soils are close in Euclidian space to the sediments and implies that the composition of the sediments is dominated by soil-derived lipids. If the analysis is restricted to the lake sediments only, then again, there is a significant difference between the low carbonate and other sediments (R=0.734, P=0.6%). This is particularly clear for the sediments that are interpreted to reflect the 8.2 ka event (Fig. 2 and Fig 6), which are on the other hand, much closer in MDS space to the single sample of degraded Terra Rossa soil. This finding would support the notion that soils eroded around 8.11 and 8.17 ka contained more strongly degraded OM. However, this assumption is based on the findings from a single sample, only.

4.3 Application of new proxies across the 8.2 ka climate event

Figure 6 compares the elemental records across the 8.2 ka event to biomarker proxies that are based on our geochemical fingerprinting of modern OM sources. The carbonate record shows two pronounced minima in response to changes in temperature and hydrology (Wagner et al., 2009) that likely reflect the Northern Hemisphere climate development controlled by two major

phases of freshwater release into the North Atlantic (Lajeunesse and St-Onge, 2008; Roy et al., 2011). In the following, we refer to the timing of these minima as Phase 1 and Phase 2 according to their chronological succession. While the carbonate record is closely linked to climate forcing, the proxies that incorporate an organic component, i.e. TOC, TOC/TN, the concentration of total lipids and the biomarker ratios, are controlled by the response of the biome and associated changes in OM fluxes to the lake sediments.

Between 8.7 and 8.35 ka, TOC and TOC/TN values are relatively high as is the terr FA/OH ratio. These data are consistent with moderately degraded OM from high- and low-altitude topsoils (terr FA/OH 0.7 and 0.9, respectively) and little degraded leaf litter (terr FA/OH 0.5 and 0.7, respectively) being the main sources of sedimentary OM between 8.7 and 8.35 ka, while the contribution from weathered (degraded) Terra Rossa-type soil (terr FA/OH 0.1) is minimal. Notably, the concentration of the total lipids steadily decreases from 8.7 ka onwards, indicating a decline in the supply of lipid-rich material and a gradual change in the make-up of the terrestrial biome. A similar trend can be seen in the $ACL_{22\text{-}26}$ FA between 8.7 and 8.5 ka, but not in $ACL_{24\text{-}28}$ OH indicating an early change in the lipid sources. This subtle shift can also be discerned in the TOC/TN ratios. While the ultimate driver behind these early trends is certainly a slowly changing regional climate, more immediate forcing factors may have included:

i. A slowly decreasing lake level, exposing the flood plains to the south of site Lz1120 and reducing the flux of lipid-rich terrestrial OM to the lake.
ii. A change in the soil moisture budget leading to increased degradation of lipids.
iii. A slow decline in terrestrial productivity.

Due to its tectonic origin, the Ohrid Basin is characterized by a pronounced terraced morphology, particularly at its northern and southern ends where today's agricultural plains were once marsh land and/or lake floor. As a result, lake level change can introduce a threshold-type control of biome characteristics and, thus, cause abrupt changes in material fluxes within the catchment of site Lz1120.

A major change occurred shortly after 8.35 ka, with the onset of Phase 1, when values for carbonate, TOC and TOC/TN dropped sharply, indicating a substantial change in the deposition of OM quantity and composition. While the carbonate record suggests a short "recovery" followed by another, slightly prolonged but less pronounced minimum during Phase 2, the

records of TOC and TOC/TN reveal that the non-carbonate sediment delivered during Phases 1 and 2 contained bulk OM of significantly different quantity and composition. During Phase 1, an increased supply of non-carbonate material apparently contributed little OM and diluted both carbonate and TOC while the sedimentary OM shows low TOC/TN ratios. During Phase 2, by contrast, TOC/TN ratios were higher and TOC was either less diluted and/or the non-carbonate material supplied to the lake contained higher amounts of OM. This difference in OM composition between Phases 1 and 2 is confirmed by the composition of the total lipid extracts that generally contain high proportions of both long-chain terrestrial *n*-alkanoic acids and *n*-alcohols, but show a substantial drop in the terr FA/OH ratio in Phase 2, only. Conversely, the average chain length of *n*-alcohols ($ACL_{24\text{-}28}$ OH) does not show any significant change except for a minimum in Phase 1. The ratio of largely suberin-derived mid-chain over plant wax-derived long-chain *n*-alkanoic acids (m-FA/l-FA), on the other hand, shows maxima in both phases. In order to explain these patterns, variable contributions from terrestrial OM pools containing little, moderately and strongly degraded OM from the low- and high-altitude habitats need to be considered. For Phase 1, higher values of m-FA/l-FA and a lower $ACL_{22\text{-}26}$ FA and $ACL_{24\text{-}28}$ OH are consistent with the lipid fingerprint of a low-altitude topsoil. However, the shifts in the elemental data require substantial supply of a material with very low amounts of carbonate and TOC, lower TOC/TN ratios and lowest lipid concentrations. An enhanced supply of un-weathered Terra Rossa-type soil would explain this pattern. An alternative OM source with low TOC/TN ratios and TOC values could be former lake sediment. This would have been deposited in the shallow areas to the south and southwest of Lz1120 at lake level high-stands and would have been spiked with suberin-derived mid-chain FA and OH during a slowly falling lake level and the development of marshy vegetation while the high groundwater table would have dissolved the carbonate. In fact, the lipid composition of the sample from 8.29 ka (Phase 1), despite showing different values for the m-FA/l-FA ratio and $ACL_{22\text{-}26}$ FA, overall does not significantly differ from the total lipid composition of samples with high carbonate contents (Fig. 5). Silty agricultural soils of brown-grey color that appear most likely to have developed from former lake sediment can be found, for example, in the fields to the south of Pogradec, with variable amounts of fine sand incorporated into the clayey silt matrix deriving from former beach lines. The shift in the geochemical records in Phase 1 may thus be explained by relatively enhanced contribution from (un-weathered) Terra Rossa-type subsoils and erosion of material from the exposed plains, combined with some contribution from surrounding low-altitude and moderately degraded topsoils. This also implies that the lake level must have dropped to roughly the modern level or below at the time. Unfortunately, we do not have the geochemical

fingerprint for the marshy element of the biome to verify this hypothesis. By contrast, while an even more pronounced peak in the m-FA/l-FA ratio and a minimum in $ACL_{22\text{-}26}$ FA again suggest increased supply of suberin-derived lipids, the significant drop in the terr FA/OH ratio observed in Phase 2 clearly suggests substantial contributions from soils with a lipid fingerprint comparable to weathered (or degraded) Terra Rossa. MDS analysis confirms the assumption of significant contribution from degraded Terra Rossa-type soil as the samples from 8.17 and 8.11 ka plot closest to the weathered Terra Rossa sample (Fig. 5).

Degraded soils as the main source of sedimentary lipids and little contribution from topsoils or of terrestrial vegetation suggests that the latter were both diminished, leaving the deeper soil layers vulnerable to erosion. A depleted topsoil pool and, possibly, a vegetation cover that had not yet fully recovered from a period of at least seasonally increased aridity would also explain why the samples from 8.23 ka (separating Phases 1 and 2) and from 8.05 ka show TOC values that appear low relative to their carbonate contents, with these two proxies otherwise correlating closely ($R^2 = 0.88$ between 8.8 and 8.0 ka, without samples from 8.05 and 8.23 ka). By contrast, carbonate supply from primary production and carbonate precipitation would have increased rapidly under warmer and more humid conditions and with increased nutrient supply.

Finally, the composition of *n*-alkanes in Figure 6 apparently follows the concentration of the total lipids, with higher proportions of grass-derived *n*-C_{31} alkane relative to the tree-derived *n*-C_{27} and *n*-C_{29} alkanes coinciding with lower concentrations of sedimentary lipids. The supply of Terra Rossa-type soil may contribute to this pattern, as it contains very low amounts of lipids and significantly higher proportions of the C_{31} alkane. Considering the fact that the incorporation of lipids into the subsoil is likely lagging the development of the overlying topsoil and vegetation makes the application of *n*-alkane distributions for vegetation reconstructions in the Ohrid Basin challenging. By contrast, biomarker proxies based on chain-length distributions and proportions of the major lipid compound classes, FA and OH, appear to follow a pattern that is consistent with the climatic development and likely biome and lake level changes.

Since Lake Ohrid appears to have been an oligotrophic ecosystem for most of its history, it is not surprising that soil OM is the major source of the sedimentary lipids. Our lipid-based proxies suggest that soil OM supply was relatively enhanced during the 8.2 ka event in response to a dryer and cooler climate and the associated recession of terrestrial vegetation. This contradicts the conclusions of Lacey et al. (2014) who assume a reduction in primary productivity and an even more profound reduction in the contribution of soil-derived OM for the period between 8.5 and 8.0 ka, based on their interpretation of proxy data from bulk organic

carbon isotope measurements and Rock-Eval pyrolysis of sediments from the western part of Lake Ohrid. The authors postulate lower soil-OM contributions on the basis of heavier bulk carbon isotopes and higher values of the oxygen index (OI) from Rock-Eval pyrolysis. However, soil OM is frequently characterized by the same pattern, i.e. bulk carbon isotope ratios that are heavier by 1 – 3 ‰ compared to the original vegetation (Lichtfouse et al., 1995; Ehleringer et al., 2000 and references therein) and have higher OI values (Disnar et al., 2003). This results from initial OM degradation in the soils rather than in the water column as assumed by Lacey et al. (2014). We therefore suggest that the data reported by Lacey et al. (2014) actually supports our conclusion of relatively increased amounts of soil-derived OM in the Lake Ohrid sediments between about 8.5 and 8.0 ka.

5 Summary and conclusions

Lake Ohrid is an outstanding sedimentary archive of SE European continental climate change, recording environmental changes in the Ohrid Basin continuously for at least 1.2 million years. OM buried in Lake Ohrid sediments can deliver information on the response of the ecosystem to climatically controlled changes in hydrology. In order to optimize the reconstruction of environmental changes we carried out the first organic geochemical survey of potential sources of sedimentary OM in the Ohrid Basin.

We determined the organic geochemical fingerprints of leaf litter, topsoil, un-weathered and weatherd subsoil, macrophytes and filtrates of suspended and slow-sinking particles from the Ohrid water column. Comparison with the TLE composition of Early Holocene sediments from site Lz1120 in the southeast of Lake Ohrid reveals that little of the sedimentary lipids appear to derive from aquatic OM sources such as macrophytes of phytoplankton. Labile compounds including MUFAs and PUFAs, accounting for a third of the polar lipids in water filtrates and up to half of the total lipids in macrophytes, are not preserved in the sediments, while the dominant saturated fatty acid in these samples, the *n*-hexadecanoic acid, is not source specific. By contrast, the majority of the sedimentary lipids are from land sources. Apart from long-chain (> C_{24}) *n*-alkanoic acids (FAs) and *n*-alcohols (OHs) that derive from cuticular waxes, mid-chain C_{22} and C_{24} FAs, ω-hydroxy FAs (ω-OH-FAs) and OHs most likely are from suberin, a bio-polyester forming protective tissue mainly in roots but also barks and bundle sheaths of grasses. The C_{22} and C_{24} ω-OH-FAs appear to be particularly reliable indicators for soil-derived OM as they were present in all topsoil samples but largely absent in leaf litter, grasses, macrophytes and water filtrates. With proportions of ω-OH-FA showing an almost identical

range in soils and sediments of about 4 ± 2 $\%_{lipids}$ on average, each, we conclude that the majority of the extracted lipids originates from soil OM with considerable inputs from root material. This is supported by the generally high amounts of suberin-derived mid-chain FA and OH. Terra Rossa, the most abundant type of subsoil in the East and South of the Ohrid Basin, contains little organic carbon compared to the topsoils (1 – 2 % *vs*. 8 - 11 %) and, accordingly, only small amounts of lipids. Notably, the terr FA/OH ratio is significantly lower in the weathered Terra Rossa, making the ratio of these compound classes a useful indicator for soil degradation and erosion. The lipid composition of the sediments is closest to the composition of topsoils and, in some cases, likely represents a mixture of lipids from moderately and strongly degraded soils.

We tested a set of new proxies based on our improved end-member characterizations for a section of sediment core Lz1120 that includes the prominent 8.2 ka event, a 400- to 500-year episode of cold and dry climate in the North Atlantic realm affecting much of the Northern Hemisphere, including SE Europe. Both the carbonate record and the biomarker proxies highlight two distinct phases of the 8.2 ka event affecting the ecosystems of the Ohrid Basin. Changing proportions of suberin-derived mid-chain compounds relative to cuticular long-chain compounds as expressed in ratios (m-FA/l-FA) or average chain lengths (ACL_{C22-26} FA, ACL_{C24-28} OH) combined with the proportion of terrestrial *n*-alkanoic acids over *n*-alcohols (terr FA/OH) indicating contributions from degraded soils also reveal significant differences between the two phases with regard to lipid fluxes from the major OM sources. Accordingly, Phase 1 (~ 8.3 ka) is characterized by relatively enhanced contribution of lipids from moderately degraded (top)soils while Phase 2 (8.2 - 8.05 ka) shows clear evidence for a substantial proportion of lipids derived from degraded soil, suggesting that the moderately degraded soil OM pool had already been depleted in Phase 1, or that soils were less protected by the modified vegetation cover. Notably, the biomarker proxies show a consistent trend of biome modification starting early from 8.65 ka onwards that is consistent with a similar trend in carbonate, TOC and TOC/TN and occurs well before the substantial change at the onset of Phase 1. These trends result from a continuous change in the hydrology of the Ohrid Basin and underline the sensitivity of the biomarker approach.

Overall, it appears that soil-derived lipids have the greatest potential to be preserved in the sedimentary archive of the Ohrid Basin. However, the extent as to which terrigenous and aquatic biomarkers are representative of the terrestrial and aquatic OM pools still is a matter of uncertainty. Multi-proxy approaches combining lipid analyses with, for example, bulk carbon and hydrogen stable isotope data, pyrolysis GC-MS analyses or Rock-Eval analyses of both

OM sources and sediments may help resolve this issue. Further improvements of biomarker-based reconstructions of the terrestrial environment may be achieved by completing the geochemical characterization of elements of the Ohrid Basin's habitats that are currently missing including, e.g., marshland soils and soils developing on the ultrabasic rocks along the western shores. However, our first survey of major OM sources in the Ohrid Basin demonstrates that organic-geochemical fingerprinting and the development of proxies adjusted to the local OM inventory can lead to an improved understanding of terrestrial biome dynamics and OM fluxes towards the lake sediments, in particular, when short-distance OM transport in a heterogeneous and variable catchment significantly increase complexity. When applied to the sedimentary record of a well-known climate fluctuation, the 8.2 ka event, the adjusted proxies demonstrably improve the interpretation of bulk-proxy data such as TOC and TOC/TN.

Appendix A: Abbreviations used for lipid biomarkers in the main text and supplements and their IUPAC names

n-Alkanes

These are abbreviated by their carbon number, for example:

n-C_{27}	*n*-heptadocosane

Fatty acids

n-Alkanoic acids are abbreviated by their carbon number, for example:

C_{16} FA	hexadecanoic acid
C_{18} FA	octadecanoic acid

Monounsaturated fatty acids (MUFAs) are abbreviated using IUPAC recommended numerical symbols, for example:

$C_{16:1}$ (*cis*-9)	*cis*-9-hexadecenoic acid
$C_{18:1}$ (*cis*-9)	*cis*-9-octadecenoic acid
$C_{20:1}$ (*cis*-9)	*cis*-9-docasenoic acid

Polyunsaturated acids (PUFAs) are abbreviated in the same way, but where double bond positions are unsure, no designated position is given, for example:

$C_{18:2}$	octadecanadienoic acid

Branched fatty acids are abbreviated as follows:

iso-C_{15}	13-methyltetradecanoic acid
anteiso-C_{15}	12-methyltetradecanoic acid
iso-C_{16}	14-methylpentadecanoic acid
iso-C_{17}	15-methylhexadecanoic acid

anteiso-C_{17} 14-methylhexadecanoic acid

Hydroxy fatty acids, either α– or ω– are abbreviated as follows:

C_{22} α–OH-FA 2-hydroxydocosanoic acid

C_{22} ω–OH-FA 22-hydroxydocosanoic acid

Diacids are denoted α,ω-DiFAs, for example:

C_{16} α,ω-DiFA hexadecadioic acid

Fatty alcohols

n-Alcohols are abbreviated by their carbon number, for example:

C_{16} OH 1-hexadecanol

C_{18} OH 1-octadecanol

Branched alcohols are abbreviated as follows:

iso-C_{15} OH 13-methyltetradecan-1-ol

anteiso-C_{15} OH 12-methyltetradecan-1-ol

Long chain diols are abbreviated as follows:

C_{30} 1,15-diol 1,15-triacontadiol

C_{32} 1,15-diol 1,15-dotriacontadiol

Long-chain ketones

Long-chain hydroxy-ketones are abbreviated as follows:

1,15(ω16)C_{30} keto-ol 1-hydroxy-tetradecan-15-one

Terpenes	
Isoprenoids	
α-tocopherol	2,5,7,8-Tetramethyl-2-(4,8,12-trimethyltridecyl)-6-chromanol
Sterols	
β-sitosterol	24β-ethylcholest-5-en-3β-ol
β-sitostenone	24β-ethylcholest-4-en-3β-one
stigmasterol	24β-ethylcholesta-5,22*E*-dien-3β-ol
stigmastanol	24-ethyl-5α(H)-cholestan-3β-ol
cholesterol	cholest-5-en-3β-ol
cholestanol	5α(H)-cholestan-3β-ol
coprostanol	5β(H)-cholestan-3α-ol
epicoprostanol	5β(H)-cholestan-3β-ol
epicholestanol	5α(H)-cholestan-3α-ol
brassicasterol	24β-methylcholesta-5,22*E*-dien-3β-ol
campesterol	24β-methylcholest-5-en-3β-ol
ergosterol	ergosta-5,7,22-trien-3β-ol
dinosterol	4α,23,24-trimethylcholest-22*E*-dien-3β-ol
dinostanol	4α,23,24-trimethylcholestan-3β-ol
lanosterol	5α-lanosta-8,24-en-3β-ol
Triterpenoids	
taraxasterol	5α–taraxast-20(30)-en-3β-ol
tetrahymanol	gammaceran-3β-ol
β-amyrin	olean-12-en-3β-ol

Acknowledgements

We are grateful to D. Kurti (Pogradec) for logistical support during macrophyte sampling, O. Cara (Durrës) and B. Muceku (Tirana) for support during soil sampling and treatment and B. Wagner and A. Francke (Cologne) for providing the sediment samples. Finally, we thank A. Thompson, S. Blackbird (Liverpool) and P. Donohoe (Newcastle) for support in the laboratories. The study was funded by the Leverhulme Trust (grant F/00 025/AU).

References

Albaigés, J., Algaba, J., and Grimalt, J.: Extractable and bound neutral lipids in some lacustrine sediments, *Org. Geochem.*, 6, 223-236, 1984.

Almendros, G., Sanz, J., and Velasco, F.: Signatures of lipid assemblages in soils under continental Mediterranean forests, *Eur. J. Soil Sci.*, 47, 183-196, 1996.

Apolinarska, K., Pełechaty, M., and Pukacz, A.: $CaCO_3$ sedimentation by modern charophytes (Characeae): Can calcified remains and carbonate $\delta^{13}C$ and $\delta^{18}O$ record the ecological state of lakes? – A review, *Stud. Lim. Tel.*, 5(2), 55-66, 2011.

Barber, D. C., Dyke, A., Hillaire-Marcel, C., Jennings, A. E., Andrews, J. T., Kerwin, M. W., Bilodeau, G., McNeely, R., Southon, J., Morehead, M. D., and Gagnon, J.-M.: Forcing of the cold event of 8,200 years ago by catastrophic drainage of Laurentide lakes, *Nature*, 400, 344-348, 1999.

Baumgarten, H., Wonik, T., Tanner, D. C., Francke, A., Wagner, B., Zanchetta, G., Sulpizio, R., Giaccio, B., and Nomade, S.: Age depth-model for the past 630 ka in Lake Ohrid (Macedonia/Albania) based on cyclostratigraphic analysis of downhole gamma ray data, *Biogeosciences Discussions*, 12, 7671-7703, 2015.

Birks, H. H. and Birks, H. J. B.: Multi-proxy studies in palaeolimnology. *Veget. Hist. Archaeobot.,* 15, 235-251, 2006.

Costantini, E. A. C., Barbetti, R., Fantappiè, M., L'Abate, G. L., Lorenzetti, R., and Magini, S.: Pedodiversity. In: Costantini, E. A. C. and Dazzi, C. (eds.): *The Soils of Italy*, World Soils Book Series, 105-178, doi: 10.1007/978-94-007-5642-7_6, Springer, Dordrecht, 2013.

Cranwell, P. A., Eglinton, G., and Robinson, N.: Lipids of aquatic organisms as potential contributors to lacustrine sediments – II, *Org. Geochem*, 11, 513-527, 1987.

Disnar, J. R., Guillet, B., Keravis, D., Di-Giovanni, C., and Sebag, D.: Soil organic matter (SOM) characterization by Rock-Eval pyrolysis: scope and limitations, *Org. Geochem.*, 34, 327-343, 2003.

Ehleringer, J. R., Buchmann, N., and Flanagan, L. B.: Carbon isotope ratios in belowground carbon cycle processes, *Ecol. Appl.*, 10, 412-422, 2000.

Ellison, C. R. W., Chapman, M. R., and Hall, I. R.: Surface and deep ocean interactions during the cold climate event 8200 years ago, *Science*, 312, 1929-1932, 2006.

Elmacı, Ö. L., Seçer, M., Erdemir, O., and Iqbal, N.: Ammonium fixation properties of some arable soils from the Aegean region of Turkey, *Europ. J. Agron.*, 17, 199-208, 2002.

Ficken, K. J., Li, B., Swain, D. L., and Eglinton, G.: An *n*-alkane proxy for the sedimentary input of submerged/floating freshwater aquatic macrophytes, *Org. Geochem.*, 31, 745-749, 2000.

Ghandi, S. R. and Weete, J. D.: Production of the polyunsaturated fatty acids arachidonic acid and eicosapentaenoic acid by the fungus *Pythium ultimum*, *J. Gen. Microbiol.*, 137, 1825-1830, 1991.

Graça, J. and Santos, S.: Suberin: a biopolyester of plants' skin, *Macromol. Biosci.*, 7, 128-135, 2007.

Grootes, P. M., Stuiver, M., White, J. W. C., Johnsen, S., and Jouzel, J.: Comparison of oxygen isotope records from the GISP2 and GRIP Greenland ice cores, *Nature*, 366, 552-554, 1993.

Holtvoeth, J., Vogel, H., Wagner, B., and Wolff, G. A.: Lipid biomarkers in Holocene and glacial sediments from ancient Lake Ohrid (Macedonia, Albania), *Biogeosciences*, 7, 3473–3489, doi:10.5194/bg-7-3473-2010, 2010.

Kiriakoulakis K., Marshall, J. D., and Wolff, G. A.: Biomarkers in a Lower Jurassic concretion from Dorset (U.K.), *J. Geol. Soc. London*, 157, 207-220, 2000.

Kothawala, D. N. and Moore, T. R.: Adsorption of dissolved nitrogen by forest mineral soils, *Can. J. For. Res.*, 39, 2381-2390, 2009.

Królikowska, J.: Eutrophication processes in a shallow, macrophyte-dominated lake – species differentiation, biomass and the distribution of submerged macrophytes in Lake Łuknajno (Poland), *Hydrobiologia*, 342/343, 411-416, 1997.

Lacey, J. H., Francke, A., Leng, M. J., Vane, C. H., and Wagner, B.: A high‐resolution Late Glacial to Holocene record of environmental change in the Mediterranean from Lake Ohrid (Macedonia/Albania), *Int. J. Earth. Sci*., 103, DOI 10.1007/s00531-014-1033-6, 2014.

Lajeunesse, P. and St-Onge, G.: The subglacial origin of the Lake Agassiz-Ojibway final outburst flood, *Nat. Geosc.*, 1, 184-188, 2008.

Leider, A., Hinrichs, K.-U., Schefuβ, E., and Versteegh, G. J. M.: Distribution and stable isotopes of plant wax derived *n*-alkanes in lacustrine, fluvial and marine surface sediments along an Eastern Italian transect and their potential to reconstruct the hydrological cycle, *Geochim. Cosmochim. Ac.*, 117, 16-32, 2013.

Lichtfouse, É., Dou, S., Girardin, C., Grably, M., Balesdent, J., Behar, F., and Vandenbroucke, M.: Unexpected ^{13}C-enrichment of organic components from wheat crop soils: evidence for the *in situ* origin of soil organic matter, *Org. Geochem.*, 23, 865-868, 1995.

Marseille, F., Disnar, J. R., Guillet, B., and Noack, Y.: *n*-Alkanes and free fatty acids in humus and A1 horizons of soils under beech, spruce and grass in the Massif-Central (Mont-Lozère), France, *Eur. J. Soil Sci.*, 50, 433-441, 1999.

Matsuda, H. and Koyama, T.: Early diagenesis of fatty acids in lacustrine sediments-II. A statistical approach to changes in fatty acid composition from recent sediments and some source materials, *Geochim. Cosmochim. Ac.*, 41, 1825-1834, 1977.

Matzinger, A., Schmid, M., Veljanoska-Sarafiloska, E., Patceva, S., Guseska, D., Wagner, B., Müller, B., Sturm, M., and Wüest, A.: Eutrophication of ancient Lake Ohrid: Global warming amplifies detrimental effects of increased nutrient inputs, *Limnol. Oceanogr.*, 52 338-353, 2007.

Matzinger, A., Jordanoski, M., Veljanoska-Sarafiloska, E., Sturm, M., Müller, B., and Wüest, A.: Is Lake Prespa jeopardizing the ecosystem of ancient Lake Ohrid? *Hydrobiologia*, 553, 89-109, DOI 10.1007/s10750-005-6427-9, 2006.

Matzke, K. and Riederer, M.: A comparative study into the chemical constitution of cutins and suberins from *Picea abies* (L.) Karst., *Quercus robur* L., and *Fagus sylvatica* L., *Planta*, 185, 233-245, 1991.

Mendez-Millan, M., Dignac, M.-F., Rumpel, C., and Derenne, S.: Can cutin and suberin biomarkers be used to trace shoot and root-derived organic matter? A molecular and isotopic approach, *Biogeochemistry*, 106, 23-28, 2011.

Mertz, R. A. and Brutnell, T. P.: Bundle sheath suberization in grass leaves: multiple barriers to characterization, *J. Exp. Bot.*, 65, 3371-3380, 2014.

Meyers, P. A.: Preservation of elemental and isotopic source identification of sedimentary organic matter, *Chem. Geol.*, 114, 289-302, 1994.

Meyers, P. A. and Lallier-Vergès, E.: Lacustrine sedimentary organic matter records of Late Quaternary paleoclimates. *J. Paleolimnol.*, 21, 345-372, 1999.

Molina, I., Bonaventure, G., Ohlrogge, J., and Pollard, M.: The lipid polyester composition of *Arabidopsis thaliana* and *Brassica napus* seeds, *Phytochem*., 67, 2597-2610, 2006.

Nieder, R., Benbi, D. K., and Scherer, H. W.: Fixation and defixation of ammonium in soils: a review, *Biol. Fertil. Soils*, 4, 1-14, 2011.

Nierop, K. G. J., Naafs, D. F. W., and van Bergen, P. F.: Origin, occurrence and fate of extractable lipids in Dutch coastal dune soils along a pH gradient, *Org. Geochem.*, 36, 555-566, 2005.

Petrova, D., Patceva, S., Mitic, V., Shtereva, G., and Gerdzhikov, D.: State of phytoplankton community in the Bulgarian and Macedonian lakes, *J. Env. Protect. Ecol.*, 9, 501-512, 2008.

Pollard, M., Beisson, F., Li, Y., and Ohlrogge, J. B.: Building lipid barriers: biosynthesis of cutin and suberin, *Trends Plant Sci.*, 13, 236-246, 2008.

Roy, M., Dell'Oste, F., Veillette, J. J., de Vernal, A., Hélie, J.-F., and Parent, M.: Insights on the events surrounding the final drainage of Lake Ojibway based on James Bay stratigraphic sequences, *Quaternary Sci. Rev.*, 30, 682-692, 2011.

Ruess, L., Schütz, K., Haubert, D., Häggblom, M. M., Kandeler, E., and Scheu, S.: Application of lipid analysis to understand trophic interactions in soil, *Ecology*, 86, 2075-2082, 2005.

Samuels, L., Kunst, L., and Jetter, R.: Sealing plant surfaces: cuticular wax formation by epidermal cells, *Annu. Rev. Plant Biol.*, 59, 683-707, 2008.

Schwark, L. Zink, K., and Lechterbeck, J.: Reconstruction of postglacial to early Holocene vegetation history in terrestrial Central Europe via cuticular lipid biomarkers an pollen records from lake sediments, *Geology*, 30, 463-466, 2002.

Shimizu, S., Kawashima, H., Shinmen, Y., Akimoto, K., and Yamada, H.: Production of eicosapentaenoic acid by *Mortierella* fungi, *J. Am. Oil Chem. Soc.*, 65, 1455-1459, 1988.

Soukup, A., Armstrong, W., Schreiber, L., Franke, R., and Votrubová, O.: Apoplastic barriers to radial oxygen loss and solute penetration: a chemical and functional comparison of the exodermis of two wetland species, *Phragmites australis* and *Glyceria maxima*, *New Phytol*., 173, 264-278, 2007.

Talevska, M.: Annual biomass and production of Red Head Grass (*Potamogeton Perfoliatus* L.) from Lake Ohrid, paper presented at the 2nd conference of the Balkan Water Observation and Information System (BALWOIS), Balwois Network, Ohrid, Rep. of Macedonia, 2006.

Trajanovska, S., Blaženčić, J., Trajanovski, S., and Budzakoska-Gjoreska, B.: Distribution, morphological variability, ecology and the present state of *Nitella* from Lake Ohrid and its surroundings, *Arch. Biol. Sci.*, 64(2), 549-556, doi:10.2298/ABS1202549T, 2012.

van Bergen, P. F., Nott, C. J., Bull, I. D., Poulton, P. R., Evershed, R. P.: Organic geochemical studies of soils from the Rothamsted Classical Experiments - IV. Preliminary results from a study of the effect of soil pH on organic matter decay, *Org. Geochem.*, 29, 1779-1795, 1998.

Vogel, H., Wagner, B., Zanchetta, G., Sulpizio, R., and Rosén, P.: A paleoclimate record with tephrochronological age control for the last glacial-interglacial cycle from Lake Ohrid, Albania and Macedonia, *J. Paleolimnol*., 44, 295-310, 2010.

Wagner, B., Wilke, T., Krastel, S., Zanchetta, G., Sulpizio, R., Reicherter, K., Leng, M., Grazhdani, A., Trajanovski, S., Francke, A., Lindhorst, K., Levkov, Z., Cvetkoska, A., Reed, J., Zhang, X., Lacey, J. H., Wonik, T., Baumgarten, H., and Vogel, H.: The SCOPSCO drilling project recovers more than 1.2 million years of history from Lake Ohrid, *Sci. Dril.*, 17, 19-29, 2014.

Wagner, B., Lotter, A. F., Nowaczyk, N., Reed, J. M., Schwalb, A., Sulpizio, R., Valsecchi, V., Wessels, M., and Zanchetta, G.: A 40 000-year record of environmental change from ancient Lake Ohrid (Albania and Macedonia), *J. Paleolimnol.*, 41, 407-430, 2009.

Yamamuro, M. and Kayanne, H.: Rapid direct determination of organic-carbon and nitrogen in carbonate-bearing sediments with a yanaco mt-5 chn analyzer. *Limnol. Oceanogr.,* 40**,** 1001-1005, 1995.

Yang H, Ding, W. H., and Xie, S. C.: Distribution of microbial fatty acids and fatty alcohols in soils from an altitude transect of Mt. Jianfengling in Hainan, China: Implication for paleoaltimetry and paleotemperature reconstruction, *Sci. China Ser. D*, 57, 999-1012, 2014.

Yunker, M. B., Belicka, L. L., Harvey, H. R., and Macdonald, R. W.: Tracing the inputs and fate of marine and terrigenous organic matter in Arctic Ocean sediments: A multivariate analysis of lipid biomarkers. *Deep-Sea Res. Pt. II,* 53, 3478-3508, 2005.

Zhang, Z., Zhao, M., Eglinton, G., Lu, H., and Huang, C.-Y.: Leaf wax lipids as paleovegetational and paleoenvironmental proxies for the Chinese Loess Plateau over the last 170 kyr, *Quaternary Sci. Rev*., 25, 575-594, 2006.

Table 1a. Total organic carbon contents (TOC), organic carbon to total nitrogen ratio (TOC/TN) and carbonate ($CaCO_3$) contents of soils and leaf litter in the southeastern Ohrid Basin (TS = topsoil, TS/F = topsoil with white rot, SS = subsoil, LL = leaf litter). For leaf litter, total nitrogen equals organic nitrogen.

sample ID	**location; coordinates**	**soil type; vegetation**	**TOC** (%)	**TOC/TN** (atomic)	**$CaCO_3$** (%)
SN-TS	above St. Naum springs; 40°54'43.78" N	Leptosol/Rendzina; small oak & beech,	**11.0** ± 0.8	**17.5** ± 0.1	**7.7** ± 3.0
SN-LL	20°44'35.18" E, 708 m.a.s.l.	shrubs	**41.1** ± 0.6	**28.8** ± 0.4	-
TP-TS	near Trpejca; 40°58'45.46" N	Leptosol/Rendzina; small oak & beech,	**7.5** ± 0.2	**17.3** ± 0.0	**5.4** ± 0.8
TP-LL	20°47'46.62" E, 808 m.a.s.l.	shrubs	**44.2** ± 0.3	**25.5** ± 0.2	-
TP-TS/F		with white rot	**10.1** ± 0.7	**17.4** ± 0.3	**5.4** ± 4.9
GN-TS	Galicica National Park; 40°57'59.64" N	Leptosol/Rendzina; mature beech forest	**8.9** ± 0.3	**15.3** ± 0.8	**7.4** ± 1.6
GN-LL	20°48'52.10" E, 1380 m.a.s.l.		n.d	n.d.	-
DG-SS	near border (dogana); 40°54'29.47" N	Chromic Luvisol/Terra Rossa (B-horizon);	**1.0** ± 0.1	**11.4** ± 0.8	**<0.5**
DG-SS 2	20°43'48.04" E, 708 m.a.s.l.	shrubs	**1.5** ± 0.1	**11.8** ± 0.6	**2.3** ± 1.5

Table 1b. Total organic carbon contents (TOC), organic carbon to organic nitrogen ratios (TOC/N_{org}) and carbonate contents ($CaCO_3$) of major Lake Ohrid macrophytes. Reed leaves (RL) were sampled near the Albanian-Macedonian border post (DG = Dogana). Submerged macrophytes were collected off Tushemisht (TU) and Pogradec (PG).

sample ID	**species**	**water depth** (m)	**TOC** (%)	**TOC/N_{org}** (atomic)	**$CaCO_3$** (%)
DG-RL	*Phragmites* leaves	0.5	44.2	13.3	-
TU-3	*Cladophora*	3	40.3	13.8	-
TU-4	*Potamogeton perfoliatus*	4	41.0	14.7	-
TU-6	*Potamogeton perfoliatus*	6	40.7	12.3	-
TU-12	*Chara tomentosa*	12	9.9	18.7	61.5
PG-14-II	*Chara gymnophyla*	14	10.4	17.0	71.2

Table 2. Composition of the total lipid extracts (TLEs) and amounts of compound classes, sub-categories and individual lipids of soils, sediments, leaf litter, macrophytes and water filtrates from the Lake Ohrid Basin. Values are given as percentages of the total lipids ($\%_{lipids}$); SN = St Naum, TP = Trpejca, GN = Galicica National Park, DG = Dogana; ME = Memëlisht; ll = leaf litter, ts = topsoil, ss = subsoil, hc = high-carbonate sediment, lc = low-carbonate sediment, clad. = *Cladophora*, potam. = *Potamogeton*, chara. = *Characeae* spp., phrag-r, -s, -l = *Phragmites* spp. roots, stem, leaves. Grey shading indicates the main compound in each category, boxes indicate the two most abundant compounds where the amount of the second-most abundant compound is less than 20 % below the amount of the most abundant compound. * Indicates only polar fractions were analysed. For the full list of compounds see 'supplementary data'.

site	SN	TP	GN	SN	TP	TP/F	GN	DG	DG	GN	GP	Lz1120		TU/PG	PG	TU/PG	DG	DG	DG/
sample type/species	**ll**	**ll**	**ll**	**ts**	**ts**	**ts**	**ts**	**ss**	**ss ll**	**g**	**g**	**hc**	**lc**	**clad.**	**potam.**	**chara.**	**phrag-r**	**phrag-s**	**phra**
number of samples	*2*	*1*	*1*	*2*	*2*	*2*	*2*	*1*	*1*	*1*	*1*	*8*	*3*	*2*	*2*	*3*	*1*	*1*	*2*
total lipids (μg/g dry weight)	643	471	653	63	97	89	79	1.3	2.3	1599	1298	40	13	1068	224	533	443	345	172
lipid fractions (%)																			
n-alkanoic acids (FA)	35.0	38.3	22.5	34.9	35.9	39.6	29.0	14.7	37.8	30.4	27.3	39.6	19.6	44.8	22.7	32.7	31.8	35.8	15.
hydroxy acids (OH-FA)	3.9	2.1	0.9	5.6	7.9	5.3	5.4	1.4	4.2	0.3	0.2	4.6	3.3	0.2	0.3	-	3.5	1.4	0.
branched alkanoic acids	0.2	-	0.1	2.6	1.5	1.4	1.5	0.6	2.3	-	-	0.9	1.2	0.6	0.04	1.7	0.4	0.1	-
mono-unsaturated alkenoic acids	6.9	7.5	20.6	19.4	22.1	17.1	25.4	14.7	21.9	2.0	3.3	0.2	0.2	40.6	6.0	20.7	4.7	15.5	1.
poly-unsaturated alkenoic acids	9.8	8.2	17.3	3.4	5.4	5.6	5.4	5.4	4.7	27.6	29.7	-	-	9.5	35.3	36.1	43.6	34.6	57.
n-alcohols (OH)	3.2	3.1	6.3	15.3	13.6	15.8	16.6	47.6	20.2	20.2	16.3	21.1	37.0	0.8	1.5	0.1	0.6	1.4	19.
n-alkanes	4.8	6.7	5.4	3.4	2.2	1.7	3.1	3.5	5.1	10.7	8.5	1.5	3.3	0.05	1.0	-	-	-	0.
methyl ketones	0.04	-	-	-	-	-	-	-	-	-	-	0.4	0.9	-	-	-	-	-	-
sterols	13.1	15.7	23.7	13.2	9.7	11.0	12.8	11.0	3.4	3.1	6.2	14.7	24.2	3.6	32.9	8.7	12.3	10.6	4.
others	23.1	18.5	3.2	2.1	1.7	2.5	0.9	1.2	-	5.5	8.2	17.0	16.6	-	0.3	0.1	3.2	0.7	0.
***n*-alkanoic acids**																			
Σ C_{14} - C_{19} FA (short-chain)	15.5	17.1	13.8	10.5	9.9	13.2	6.1	7.1	12.3	18.2	22.4	7.3	7.2	44.3	19.2	31.7	19.4	28.1	12.
Σ C_{20} - C_{25} FA (mid-chain)	6.4	10.0	5.7	11.3	13.3	11.4	9.5	5.7	14.4	2.4	2.2	13.7	7.3	0.5	2.8	1.0	10.0	5.6	1.
Σ C_{26} - C_{34} FA (long-chain)	13.2	11.2	2.9	13.1	12.7	15.0	13.4	1.8	11.1	9.8	2.7	18.6	5.1	0.1	0.8	0.01	2.3	2.0	1.
hydroxy acids																			
Σ ω-hydroxy acids	0.3	0.1	0.6	4.4	6.1	4.0	4.7	1.1	1.9	0.2	0.1	4.4	2.6	-	0.02	-	1.6	0.4	-
Σ α-hydroxy acids	3.5	2.0	0.2	1.2	1.7	1.3	0.7	0.3	2.3	0.1	0.1	0.3	0.7	0.2	0.3	-	2.0	1.0	0.
branched alkanoic acids																			
C_{14} - C_{16} branched FA	0.2	-	0.1	1.7	0.8	0.9	1.0	0.6	0.5	-	-	0.9	1.1	0.4	-	0.7	-	0.1	-
C_{17} - C_{19} branched FA	-	-	-	0.9	0.7	0.5	0.5	-	1.7	-	-	0.01	0.1	0.1	-	1.0	0.4	-	-
C_{20} - C_{25} branched FA	-	-	-	-	-	-	-	-	-	-	-	0.01	-	-	-	-	-	-	-
mono-unsaturated alkenoic acids																			
$C_{16:1}$ FA	0.7	0.9	1.0	4.4	5.2	8.4	2.7	4.0	2.1	-	-	-	-	29.3	2.7	11.8	-	13.1	1.
$C_{18:1}$ FA	6.1	6.3	19.4	14.5	16.6	8.5	22.5	10.2	19.7	1.9	3.3	0.2	0.5	11.3	2.7	7.1	0.5	0.7	0.
$C_{19:1}$ - $C_{24:1}$ FA	0.1	0.3	0.2	0.5	0.2	0.2	0.2	0.5	0.1	0.2	-	-	-	0.02	0.6	1.8	4.1	1.2	0.
poly-unsaturated alkenoic acids																			
$C_{18:2}$ FA	7.9	8.2	17.3	3.4	5.4	5.3	5.4	5.4	4.7	8.1	7.7	-	-	2.5	11.2	8.8	27.0	21.1	10.
$C_{18:3}$ FA	2.0	-	-	0.02	0.1	0.3	0.02	-	-	19.4	21.9	-	-	3.7	22.9	11.3	16.4	12.7	47.
	T E R R E S T R I A L											**S E D I M.**				**A Q U A T I C**			

site	SN	TP	GN	SN	TP	TP/F	GN	DG	DG	GN	GP	Lz1120		TU/PG	PG	TU/PG	DG	DG	DG/
sample type/species	II	II	II	ts	ts	ts	ts	ss	ss II	g	g	hc	lc	clad.	potam.	chara.	phrag-r	phrag-s	phra
poly-unsaturated alkenoic acids																			
$C_{20:5}$ FA (EPA)	-	-	-	-	-	0.1	-	-	-	-	-	-	-	1.5	0.6	5.5	-	0.7	-
$C_{22:6}$ FA (DHA)	-	-	-	-	-	-	-	-	-	-	-	-	-	0.1	0.2	2.2	-	-	-
***n*-alcohols**																			
Σ C_{12} - C_{19} OH (short-chain)	0.2	0.1	1.4	0.8	0.7	0.5	0.6	0.6	0.9	0.1	-	1.1	1.3	0.2	0.1	0.1	0.1	-	0.
Σ C_{20} - C_{25} OH (mid-chain)	1.8	1.5	3.8	7.5	6.8	6.3	5.3	16.6	15.9	17.1	15.2	7.6	11.1	0.1	0.8	0.01	0.3	0.4	0.
Σ C_{26} - C_{34} OH (long-chain)	1.2	1.4	1.1	7.0	6.1	9.0	10.7	30.4	3.3	2.9	1.0	12.4	18.1	0.4	0.6	-	0.2	1.1	18.
***n*-alkanes**																			
Σ C_{19} - C_{20} (short-chain)	-	-	-	-	-	-	-	-	-	-	-	-	0.02	-	-	-	-	-	-
Σ C_{21} - C_{26} (mid-chain)	0.5	0.3	0.3	0.2	0.1	0.1	0.1	0.3	1.1	0.6	0.1	0.2	0.4	-	0.4	-	-	-	0.0
Σ C_{27} - C_{33} (long-chain)	4.3	6.4	5.1	3.2	2.1	1.6	2.9	3.2	4.0	10.1	8.5	1.3	3.0	0.05	0.3	-	-	0.3	-
sterols																			
cholesterol	-	0.1	-	0.5	0.5	0.2	0.9	0.6	-	-	-	1.5	1.9	0.6	0.4	2.3	0.1	-	0.
cholestanol	-	-	-	-	-	-	-	-	-	-	-	0.8	2.2	-	-	-	-	-	-
ergosterol	0.2	-	0.4	0.1	0.2	0.1	0.1	-	-	-	-	-	-	-	-	-	-	-	-
stigmasterol	0.03	0.2	0.3	0.4	0.6	0.6	0.3	-	0.6	-	-	0.3	0.5	-	8.1	0.1	1.6	1.7	0.
sitosterol	10.5	11.0	16.6	9.5	5.9	7.8	9.4	9.1	2.9	2.8	5.1	1.7	3.0	2.8	15.3	3.4	10.3	8.6	3.
stigmastanol	0.3	0.6	0.4	0.7	0.5	0.5	0.8	0.9	-	-	0.5	3.3	6.2	-	-	-	0.3	0.3	-
dinosterol	-	-	-	-	-	-	-	-	-	-	-	1.2	1.1	-	-	-	-	-	-
dinostanol	-	-	-	-	-	-	-	-	-	-	-	1.4	1.8	-	-	-	-	-	-
β-sitostenone	1.6	2.9	5.1	1.0	0.9	0.6	0.8	0.3	-	-	-	-	-	-	-	-	-	-	-
lanosterol	-	-	-	-	-	-	-	-	-	-	-	3.3	5.9	-	-	-	-	-	-
others																			
dicarboxylic acids (DiFA)	0.2	-	-	1.0	0.5	1.0	0.4	-	-	-	-	-	-	-	-	-	0.2	-	-
α-tocopherol	2.6	3.6	3.2	0.7	0.4	0.3	0.3	-	-	0.4	0.2	0.7	0.4	-	0.2	0.1	0.1	0.1	0.
α+β-amyrin	1.2	1.8	-	0.7	0.8	0.7	0.1	1.2	-	5.1	8.0	0.7	1.2	-	-	-	-	-	-
tetrahymanol	-	-	-	-	-	-	-	-	-	-	-	0.7	1.5	-	-	-	-	-	-
1,15(ω16)C_{30} diol	-	-	-	-	-	-	-	-	-	-	-	2.0	2.6	-	-	-	-	-	-
1,15(ω16)C_{30} keto-ol	-	-	-	-	-	-	-	-	-	-	-	1.0	1.3	-	-	-	-	-	-
1,15(ω18)C_{32} diol	-	-	-	-	-	-	-	-	-	-	-	0.8	0.8	-	-	-	-	-	-
17β(H),21β(H)-bishomohopanoic acid	-	-	-	-	-	-	-	-	-	-	-	9.3	6.1	-	-	-	-	-	-
	TERRESTRIAL											SEDIM.				AQUATIC			

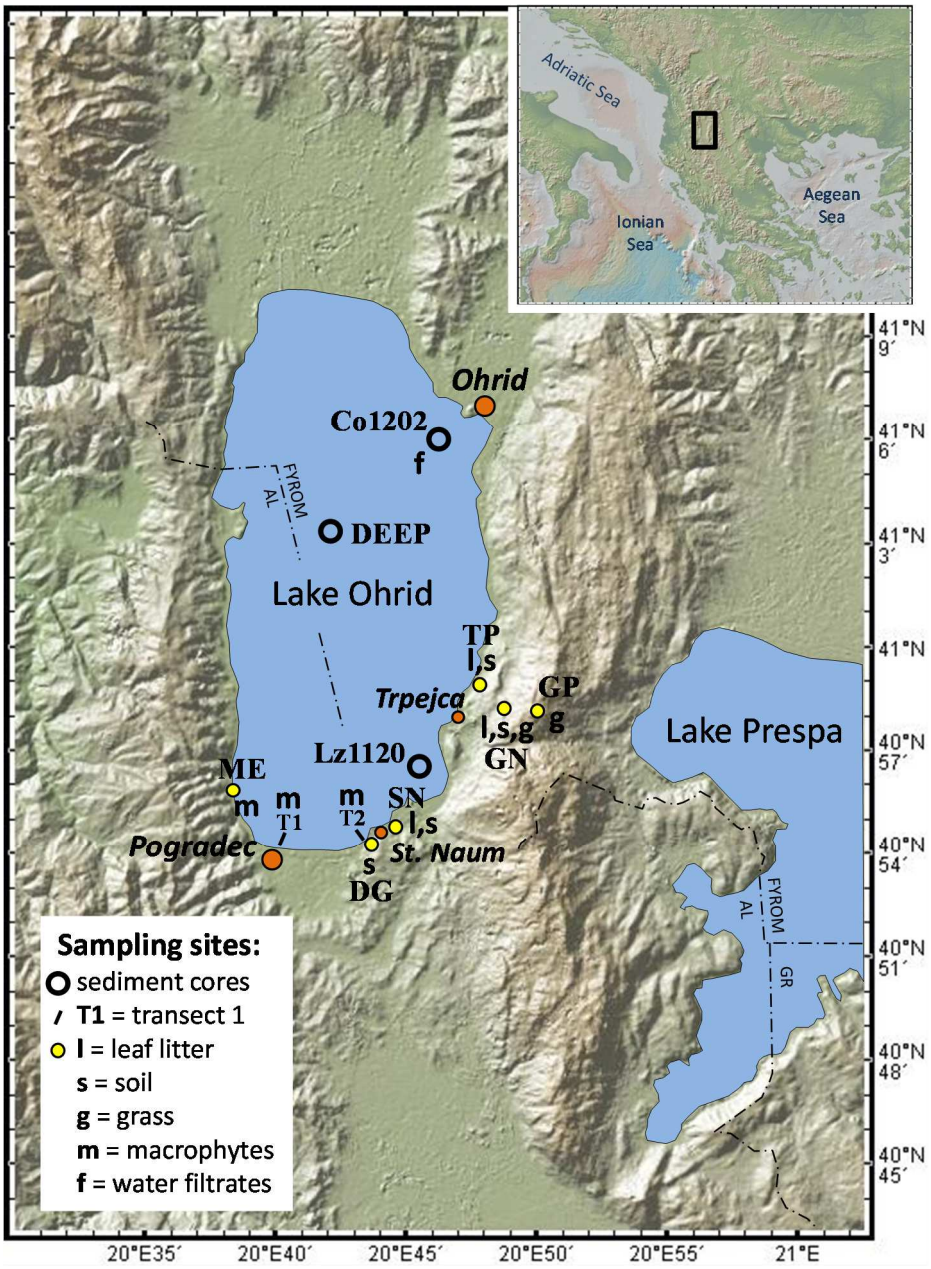


Figure 1. Locations of sediment core Lz1120 and sampling sites for modern materials in the Ohrid Basin. Lake Prespa contributes 28% of the inflow of meteoric water through karst systems. Inset: Location of Lake Ohrid in the Balkans region of South Eastern Europe.

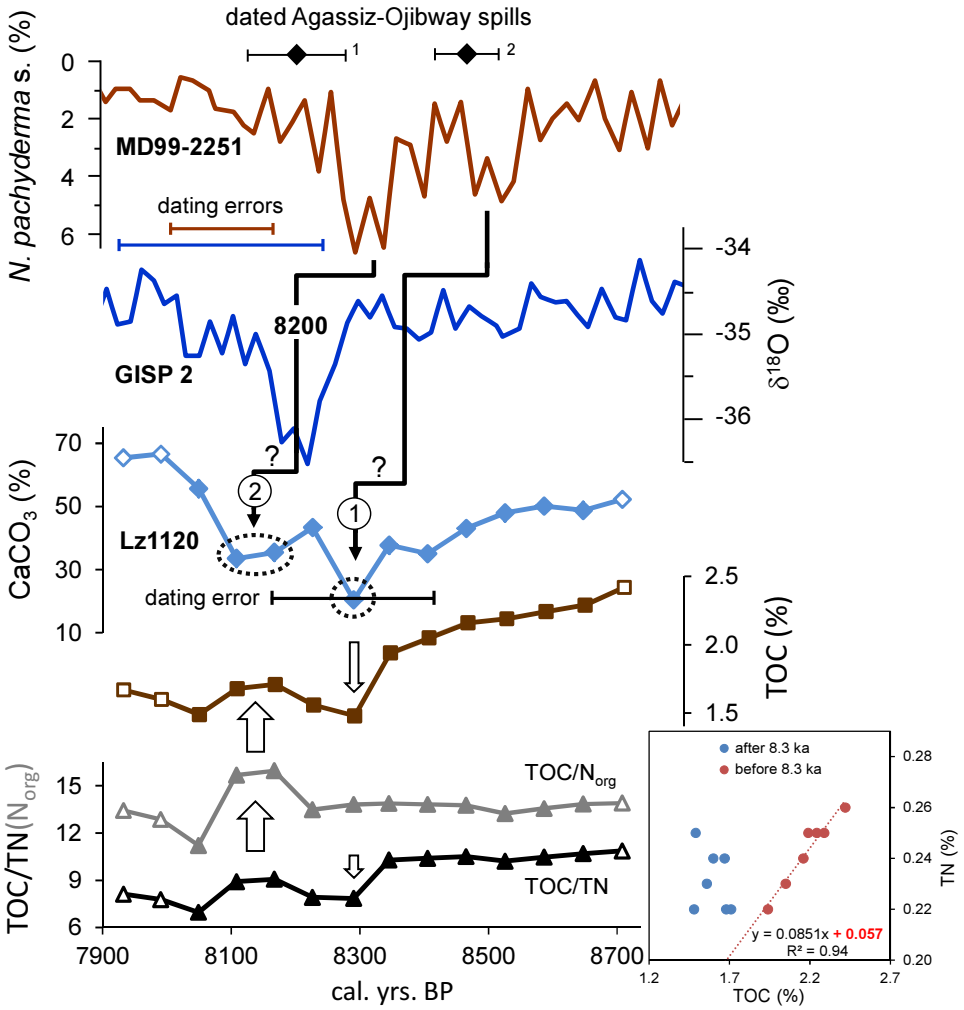
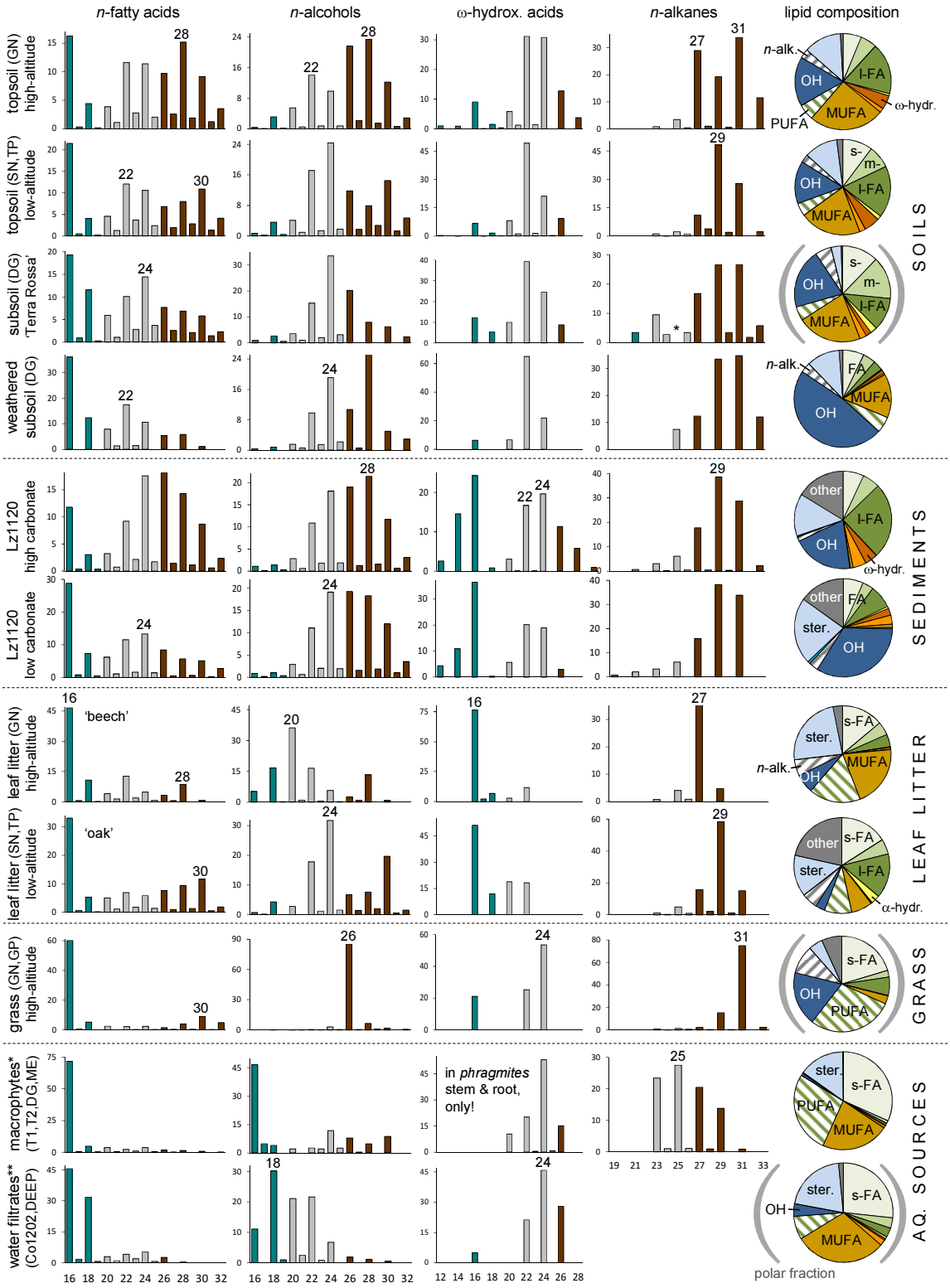


Figure 2. Records of carbonate ($CaCO_3$), total organic carbon (TOC) and the organic carbon to total nitrogen ratio (TOC/TN) of core Lz1120 compared to proxy data from the North Atlantic. Phases of increased percentages of *N. pachyderma* s. in MD99-2251 indicate at least two southward shifts of colder surface waters (Ellison et al., 2006) while the excursion in the $\delta^{18}O$ record of Greenland ice cores (GISP), which first defined the 8.2 ka event, indicates a drop in atmospheric temperature over the Greenland Ice Sheet (Grootes et al., 1993). The dashed circles mark those samples that are discussed as low-carbonate samples in carbonate minima 1 and 2 and are possibly corresponding to the climatic deteriorations occurring over the course of the 8.2 ka event caused by at least two catastrophic freshwater spills into the North Atlantic ([1]Roy et al., 2011; [2]Lajeunesse and St-Onge, 2008). Closed symbols mark samples analysed for their lipid biomarker composition. White arrows highlight different fluxes of organic carbon and total nitrogen during the carbonate minima. Insert: TOC *vs*. TN for samples from before (red) and after (blue) 8.3 ka.



1

2 Figure 3. Chain-length distributions of *n*-alkyl compounds (bar diagrams) in high- and low-
3 altitude topsoil, subsoil (un-weathered and weathered Terra Rossa), early Holocene sediments,
4 in high- and low-altitude leaf litter, high-altitude grasses, macrophytes and water filtrates and
5 the composition of GC-amendable lipids (pie diagrams). Chain-length distribution y-axis

values are percentages of the total amount of each compound class, i.e. $\%_{FA}$, $\%_{OH}$ *etc*.; s-, m-, l-FA = short-, mid- and long-chain FA; MUFA = mono-unsaturated FA; PUFA = poly-unsaturated FA; ster. = sterols; hydr. = hydroxy acids; br. FA = branched FA; asterix (*): co-elution of *n*-C_{25} with contaminant; lipid composition for un-weathered Terra Rossa is likely less accurate as response factors for main compounds were not determined. Note the distinct shift towards C_{22} and C_{24} in chain-length distributions of FAs and towards C_{24} in chain-length distributions of OHs from high to low carbonate sediment samples. C_{12} and C_{14} ω-hydroxy acids in the sediments appear to derive from an *in-situ* source. Macrophyte data derives from submerged species, except for the ω-hydroxy acids, which exclusively occur in (emergent) *Phragmites* sp.. Accuracy of the lipid composition of the grasses is limited since response factors could not be determined for the main compound classes. Water filtrate data does not include non-polar compounds (e.g., *n*-alkanes) as these were separated from the polar TLE fractions prior to further sample processing.

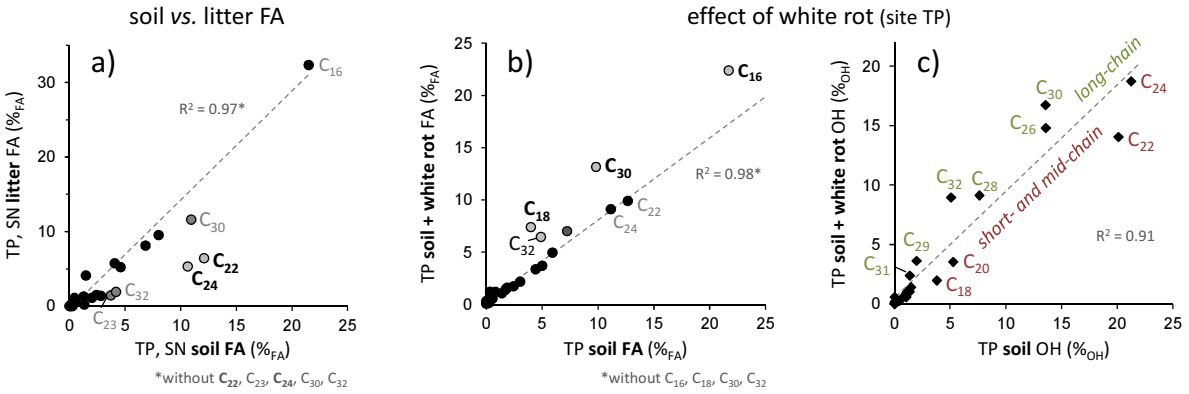


Figure 4. Correlations of compound class-specific relative abundances of a) *n*-alkanoic acids in soils and leaf litter of the low-altitude, oak-dominated sites (TP, SN), b) *n*-alkanoic acids and c) *n*-alcohols in the two soil samples from site TP, one of which is affected by white rot. The patterns illustrate enrichment of mid-chain FA ($C_{22, 24}$) from leaf litter to soil and of long-chain *n*-alcohols ($C_{26\text{-}32}$) as well as long-chain FA ($C_{30, 32}$) and short-chain FA ($C_{16,18}$) due to visibly enhanced fungal degradation (white rot) within the soil at site TP. Chain-length shifts in the latter suggest either selective degradation or biosynthesis of specific compounds by fungi.

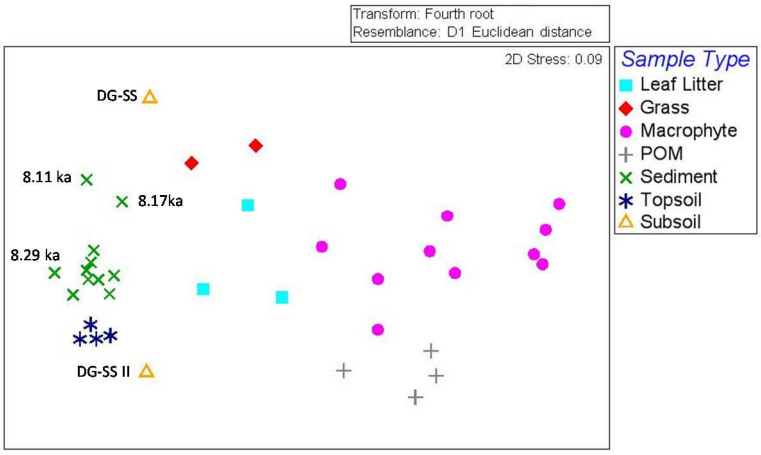


Figure 5. MDS plot showing separation of sample types in Euclidean space. For details see methodology section 2.4. Low-carbon sediment samples are labelled with their age (ka). DG-SS and DG-SS II refer to the weathered and un-weathered subsoil samples (Terra Rossa) from site DG.

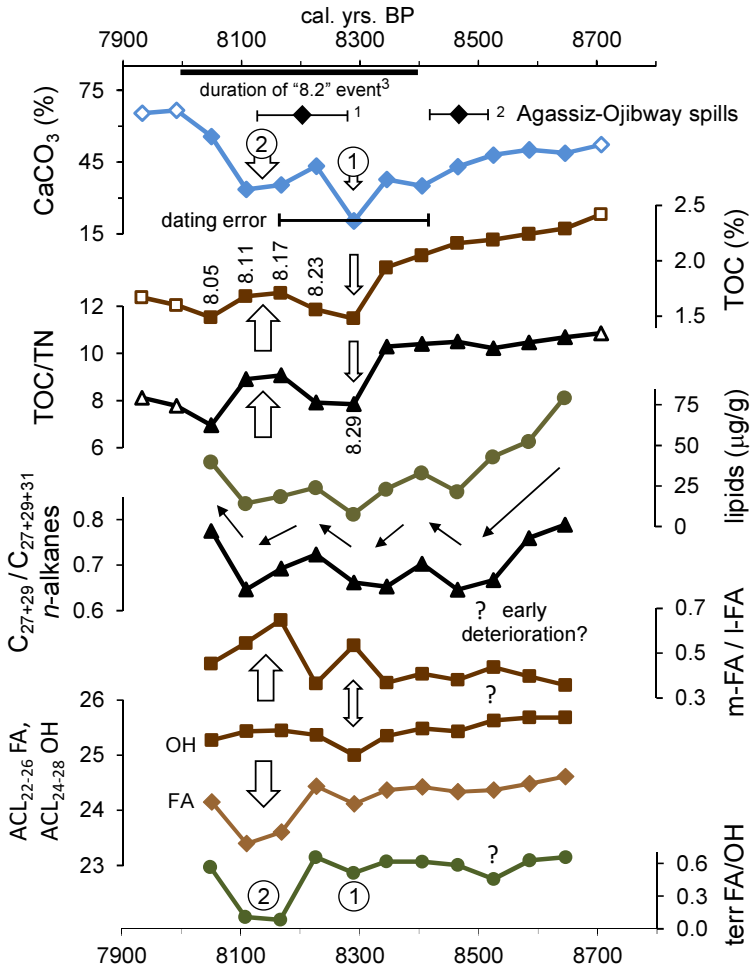


Figure 6. Records from elemental ($CaCO_3$, TOC, TOC/TN) and lipid biomarker analyses (closed symbols) of sediment core Lz1120 over the course of the 8.2 ka climate deterioration. White arrows indicate contrasting response of elemental and biomarker proxies in response to the reduced carbonate supply, i.e. in carbonate minima. Black diamonds indicate Agassiz-Ojibway freshwater spillages dated by [1]Roy et al. (2011) and [2]Lajeunesse and St-Onge (2008), [3]duration of 8.2 climate deterioration (Barber et al., 1999). While *n*-alkane composition (C_{27} and C_{29} relative to C_{31}) appears to co-vary with the total concentration of lipids in the sediment (indicated by the black arrows), proxies based on *n*-alkanoic acid and *n*-alcohol distribution appear to correspond to the environmental changes indicated by the variable supply of carbonate and non-carbonate material.