MS No.: **bg-2015-303** MS Type: Research Article Special Issue: Integrated perspectives on biological and geological dynamics in ancient Lake Ohrid.

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

Authors: J. Holtvoeth, D. Rushworth, A. Imeri, M. Cara, H. Vogel, T. Wagner, and G.A. Wolff

First of all, we would like to thank the reviewers for their critical reviews of our manuscript! We greatly appreciate the reviewer's comments and suggestions in order to improve the manuscript.

A general issue both referees have raised is the length of the manuscript. This is largely due to very detailed reporting of minor biomarkers and of isotope data, both of which do not ultimately contribute significantly to the development and evaluation of new palaeoenvironmental proxies for the Ohrid Basin. We have therefore decided to considerably streamline the manuscript and to shift complementary information on biomarkers and the discussion of the isotope data into a supplementary section (Supplement 2 - Additional information on compound-specific isotope data and specific biomarkers observed) since we believe these to be useful to scientists working on similar subjects. This and other points raised by the referees are addressed in greater detail below.

Reply to the comments by anonymous referee #1:

1) Referee: "Line 13, page 8: "Free and bound acids"? How were these obtained?"

   Authors: We did not actually determine the amounts of free and bound fatty acids. Fatty acids were transmethylated in order to produce GC-amendable fatty acid methyl esters (FAMEs). However, transmethylation with acetyl chloride/methanol also breaks ester bonds of wax esters. Thus, fatty acids that were bound in wax esters were released and could not be separated from free fatty acids that were present in the form of storage fats or oils, for example. We rephrased the relevant section in the Methods chapter for clarification.

2) Referee: "Line 9, page 15: The authors suggest that Taraxerol is 'quickly degraded' based on their data. This is very surprising since this compound has been widely applied as a marker for mangrove detritus, and has been shown to be quite stable in the environment. The authors should revise this comment as it is misleading."

   Authors: The referee unfortunately misread the name of the compound. This results from the very similar spelling of two distinct biomarkers: taraxerol and tarax**ast**erol. While taraxerol indeed is a prominent marker for mangrove input, tarax**ast**erol derives from asteraceae, a family of flowering land plants that apparently are present in the Ohrid catchment but were not sampled during our survey. The fact that this compound was found in the soil samples from one location, only, but not in the sediments led us to the conclusion that it is either not preserved or diluted by lipids from other sources and provided in amounts too low to be detected. Since taraxasterol does not appear to have potential as palaeoenvironmental indicator in the Ohrid Basin we have now shifted detailed information and references on this compound into Supplement 2.

3) Referee: "Page 19, lines 3-14: I suggest introducing the concept of P_{aq} here. In fact, I feel that the P_{aq} data should be shown in a Table or graph. The fact that the P_{aq} values in the sediments are so low is a clear indication of a very limited contribution of marcophytes and floating/submerged plant materials to the sediments. This is particularly important during time periods when the lake

water level decreased, which may suggest a vegetation shift along the lake edges to include larger macrophyte contributions.

However, this doesn't seem to be the case. So, may be making more intensive use of the P_{aq} information may be merited."

Authors: In the investigated macrophytes, *n*-alkanes are present in very small amounts, if at all, which is why P_{aq} can only be calculated for one *Potamogeton* sample and the *Phragmites* leaves (see Supplement 1 – Data). *n*-Alkanes were absent in lipid extracts of *Cladophora* and *Characeae* sp. and in the water filtrates as the non-polar fraction had been removed from the latter. We don't think it is necessary to insert a data table for P_{aq} with gaps only to show that the value is very low for those terrestrial and sediment samples for which it could be determined samples and higher for 2 out of 4 investigated macrophytes. However, we did insert the proxy into the data table of Supplement 2.

In the Ohrid Basin, changes in water level do not tend to increase the size of the macrophyte habitat, significantly. The reason for this is that the littoral zone in Lake Ohrid is generally very narrow due to the steep shores that characterise this tectonic basin. Furthermore, the lake is very deep, with a maximum depth of 290 m and an average depth of 150 m. Thus, even at low water levels littoral productivity will be low relative to terrestrial productivity.

4) Referee: "I am surprised how much effort went towards stable C and H isotope analyses, and how little was gained from this effort. I would imagine that even if there was a range of values in the end-members, that a significant climate shift would be reflected in a change in primary productivity (i.e. $\delta^{13}C$ shift) and a change in hydrology or water stress (i.e. shift in δD). While I found the discussion on the stable isotope data a bit confusing, I realize that it would take significant additional experimental efforts to constrain the findings and place them into a paleo-environmental context. However, I wonder if it would be worth it shortening this section considering that the paper is already quite extensive? Also, with this section in mind – what about effects of primary productivity changes to explain the data discussed in lines 15-20 on page 33? Or, what were the n-alkane specific δD values during 'warmer/humid' vs. 'cooler/drier' climates (Page 35, lines 25-28)? Do they agree with your statements?"

Authors: We agree with the referee that the contribution of the isotope data towards an isotope-based paleoproxy is limited. Unpulished data from a series of initial compound-specific hydrogen isotope measurements did not show statistically significant differences between samples from glacial and interglacial climate stages. This was a puzzling find considering that according to current knowledge such a difference certainly persisted in the hydrogen isotope composition of precipitation in the Ohrid Basin between these opposing climate regimes. However, climatically driven local factors such as changes in vegetation cover, evapotranspiration, shifts in the growth season or seasonality of precipitation may all have contributed to obscure the primary change in precipitation isotope composition. The apparent lack of understanding of how the hydrological system of the Ohrid Basin operates with regard to the hydrogen isotope composition of soil and plant lipids triggered the isotope survey presented in our study. The outcome provides ranges of isotope values for both carbon and hydrogen in the terrestrial environment of Lake Ohrid. As discussed in the original text, the results illustrate the complexity of interpreting isotope data resulting from multiple sources and shed light on the limits rather than providing a straightforward tool for palaeoenvironmental investigations. From a science philosophical point of view, however, defining methodological limits is as important as developing new

methodologies. For this reason, we think the current data set is worth reporting as it is likely useful to other scientists.

We also tend to agree with the referee that the presentation and discussion of the isotope data lengthens the manuscript considerably and may confuse the readers. In order to streamline the manuscript and to keep the main text focussed on the biomarker proxy approach, we have now shifted the description and discussion of the isotope data into Supplement 2.

With regard to the referee's question on *n*-alkane isotope data:
Unfortunately, we were not able to determine the isotope composition of the *n*-alkanes, as we had limited amounts of sample and the amounts of *n*-alkanes gained from the solvent extractions were too low.

5) Referee: "Line 14, page 32: remove 'always' – is confusing."

Authors: Done.

6) Referee: "Line 26, page 42: The authors suggest that adding bulk carbon and hydrogen stable isotope data to lipid biomarker data might help resolve issue regarding terrestrial vs. aquatic OM pools. I tend to agree with that, but in their case the stable isotope data did not seem to make a difference. As such it seems a bit awkward to state this here."

Authors: Lipid carbon represents only a minute fraction of the total organic carbon in aquatic and terrestrial environments. More importantly, the lipid fraction is almost certainly not quantitatively representative for the aquatic and terrestrial organic carbon pools. For example, under semi-arid climatic conditions terrestrial production of lipids may be higher relative to total organic carbon production than under more humid conditions when a greater proportion of carbon is incorporated into cellulose- and lignin-rich woody plant tissue. Proportions of above- and below-ground biomass may also vary considerably, which could affect the carbon isotope composition due to the intake of soil CO_2 and contribution from microbial respiration. We believe that bulk carbon isotopes may support assessments of terrestrial and aquatic organic carbon amounts, not despite but because of their more integrating quality.

Reply to the comments by anonymous referee #2:

General Comments

1) Referee: "I have one major concern about this manuscript directly related to the notion that an "improved endmember characterization" is provided."

Authors: Sediments record contributions of organic matter derived from a wide range of sources and pools, either from within the overlying water body or from the surrounding terrestrial environment. Palaeoclimatologists have frequently referred to organic matter from aquatic and terrestrial sources as the marine/lacustrine and terrestrial end-members, and still do so. However, in our opinion there is no such thing as *the* terrestrial end-member since there is great heterogeneity within the soil organic matter pool alone, as the referee quite rightly points out. Rather, there is a range of terrestrial sources for organic matter that can roughly be subdivided into vegetation and soils, both of which may provide organic matter of greatly variable chemical composition. In large-size catchments, e.g., of marine sampling sites, local or regional heterogeneity is often averaged out to various extents through fluvial or eolian long-distance transport and mixing. However, in small catchments and in lacustrine catchments, in particular,

only a subset of terrestrial organic matter sources is typically present and will contribute to the sedimentary organic matter. Thus, chemical characterisation of this subset of organic matter sources is required if molecular data from the sedimentary archive were to be interpreted with some confidence. For the Ohrid Basin, no data on the molecular composition of potential organic matter sources has so far been available. In our view, any such data contributing to the characterisation of potential organic matter sources, compared to none, is an improvement, and we would therefore like to stick with this term. We do point out in the manuscript that this source characterisation is incomplete (see Summary and conclusions: "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.").

We would like to highlight that our results, even from the limited number of investigated sites, have provided an explanation for the variable amounts of mid-chain C_{22} and C_{24} compounds that are observed in sediments from other sites in Lake Ohrid (further manuscript in preparation) as well as in comparable lacustrine settings (see examples in Chapter 4.2, e.g., Matsuda and Koyama, 1977). Previously, and due to the lack of such evidence, an interpretation as towards root material as the source for these compounds could not be provided.

Referee: "The soil OM data set is based on and all conclusions […] are drawn from a single deeper soil sample and 3 topsoils. Soils are extremely heterogeneous and even samples taken within closest proximity differ considerably depending on a multitude of factors. [For example …] Accordingly, information from a single Terra Rossa B-horizon sample is likely biased and not representative for the Terra Rossa soils in the catchment; I am very sceptical that meaningful inferences about the relative contributions of Terra Rossa vs. topsoils to Ohrid sediments (4.3, e.g., p. 13010 l.28 – p. 13011 l. 2) can be made based on this limitation"

Authors: We agree with the referee that soil organic matter composition is generally patchy. However, the increase in the proportion of long-chain *n*-alcohols in soil relative to leaf litter is substantial and appears to be a very robust feature even though the number of investigated samples is admittedly small.

After we received the reviews, we have analysed another Terra Rossa sample that was taken from the very same spot as the first one, the difference being that it was not taken from the exposed surface of the soil profile but from 10-15 cm depth (see revised methods chapter). Interestingly, we found that the lipid composition of this "un-weathered" sample closely resembles the composition of the topsoil samples, including the relative enrichment of mid-chain C_{22}, C_{24} fatty acids. It now appears as if the increase in *n*-alcohols from leaf litter to soil to weathered soil results from selective lipid degradation/preservation. We therefore refrained from making the distinction between topsoil and subsoil lipid profiles but discuss moderately and strongly degraded soil organic matter, instead.

(FYI: Our current findings and conclusions are confirmed by follow-up research that is currently carried out using soil and plant litter samples from further sites, the results of which will be published in the future.)

Referee: "Another question arising in this context is whether leaf litter and soil OM can truly be considered as two entities? Is there evidence that leaf litter is directly transported into the lake in sufficient amounts to count as separate endmember? I would assume that the majority of the

litter in the catchment will be remineralized and become incorporated into the soil, which than carries its modified biomarker signal."

Authors: We agree with the referee that leaf litter and topsoil organic matter are not independent entities. Rather, leaf litter represents the starting material from which a major fraction of soil organic matter at various levels of degradation derives (see above), bar subterraneous biomass input, that is. We rephrased some sections of the manuscript in order to avoid the impression that we regard leaf litter and soil organic matter as distinct terrestrial end-members. Nevertheless, it is important to realise that leaf litter and soil organic matter represent different sources: while leaf litter consists of material from the present vegetation cover, soil organic matter integrates over time and incorporates material from potentially different past vegetation as the bimodal *n*-alkane in the high-altitude topsoil nicely demonstrates. We also agree with the referee that direct supply of leaf litter towards the lake is probably small compared to its contribution towards soil organic matter. In fact, we conclude that, despite leaf litter having significantly higher concentrations of lipids, soil organic matter is the major source of sedimentary lipids rather than plant litter, which is also illustrated by the statistical analysis.

2) Referee: "The manuscript is really long and particularly paragraph 3.2 is quite excessive in listing all investigated biomarkers and their relative abundances. From Sections 4.2 and 4.3 as well as the figures it is evident that the vast majority of conclusions are based on a subset of the biomarkers analysed (notably the ones shown in Figure 3). The authors may want to consider moving a substantial amount of paragraph 3.2 into a supplementary discussion chapter for ease of conveying their most important conclusions."

Authors: we followed the referee's advice and shifted into a supplementary file.

3) Referee: "When quantifying absolute lipid concentrations they should be normalized to g TOC rather than g DW to avoid bias due to, e.g., preservation effects or changes in sedimentation rates (dilution effects) and for better comparison of the different materials analyzed - dry weight leaf litter and macrophytes biomass is not directly comparable to soils and sediments, which consist of large amounts of non-organic material. This might also change the estimation of aquatic contributions to, e.g., the sedimentary FA inventory (paragraph 4.3)."

Authors: The referee highlights one of the fundamental problems in the investigation of degradable organic matter in relation to inorganic/recalcitrant matter, i.e. the fact that TOC differs between the sources of sedimentary material and that it can be modified due to microbial OM degradation. However, this is equally true for the proportion of lipids relative to the total organic carbon (TOC) in that lipids represent only a minute and highly variable proportion of the total organic matter, with greatly variable biodegradability relative to TOC. For example, soils with high amounts of black carbon (soot, charred material) have high TOC values but low lipid concentrations. Thus, unless further information on the majority of the organic matter is available from, e.g., Rock-Eval pyrolysis or pyrolysis GC-MS there is little gain from lipid concentration normalised to TOC as the reasons behind any observed changes in this relation remain unknown. Therefore, the main aim of our study was to identify compositional changes within the lipid fraction in order to provide a biomarker fingerprint for the main terrestrial organic matter sources. We think that changing proportions of individual compounds classes (*n*-alkyl compounds) and individual biomarkers within the lipid fraction is the best way to illustrate qualitative changes of labile OM.

Nevertheless, we did compare the concentrations of lipids per g sediment and normalised to TOC in the sedimentary sequence covering the 8.2 ka event. Notably, there is a very high correlation between these (R^2 = 0.92), suggesting that a drop in organic matter supply corresponds to drop in the lipid concentration of the organic matter. We have now included this in the revised text.

4) Referee: "The potential of the compound-specific isotope data could be exploited further. It seems that cross-plotting $\delta^{13}C$ and δD values for the different samples and different FAs might reveal that the different endmembers have characteristic values/ranges."

Authors: Following the advice of referee 1, we have decided to shift the discussion of the isotope data into a separate text (Supplement 2). Further research is currently carried out to improve the data base for a future publication that will include the current data.

Specific Comments:

Referee: "p. 12978 l. 9-12: This is a rather limited representation of the tools paleoclimatologists use; many use established molecular tools as well."

Authors: Palaeoclimatologists certainly apply a wide range of molecular proxies to address environmental change. When it comes to assessing contributions from aquatic and terrestrial sources, however, this narrows down dramatically. We are not aware of a reliable molecular proxy suited to this particular purpose. Molecular approaches such as ratios of branched over isoprenoid glycerol-dialkyl-glycerol tetraethers (BIT index) from soil bacteria and aquatic archaea, respectively, cover only a specific part of the terrestrial organic matter pool, i.e. soil organic matter, assuming the amount of bacterial biomass in the soils remains more or less constant. Assessments based on ratios of short-chain (C_{16}, C_{18}) over long-chain ($>C_{20}$) *n*-fatty acids are similarly biased as this ratio varies greatly within terrestrial organic matter alone. Neither tetraethers nor *n*-fatty acids or, in fact, any specific compound class has been shown to quantitatively represent their specific sources, which is one reason why proxies such as bulk $\delta^{13}C_{org}$ and C/N remain popular for first order assessments of aquatic and terrestrial contributions, and with some justification.

Referee: "p. 12978 l. 25: Is there no reference in this special issue, which can be referred to?"

Authors: Baumgarten et al. (2015) added.

Referee: "p. 12980 l. 26: As per USDA taxonomy (which was used in the manuscript), all aerated horizons with <20% TOC are mineral soils as well. Accordingly, all "topsoils" investigated here are mineral soils and not O-horizons."

Authors: The O-horizon in the methods chapter refers to the leaf litter layer, which is in accordance with the USDA soil taxonomy. The relevant sentence has been rephrased to clarify this. An erroneous use of O-horizon in the results chapter has been removed. Furthermore, we have removed the term "mineral soil" from the text and now refer to "topsoils" for the investigated A-horizons and "subsoil" for the investigated B-horizon.

Referee: "p. 12980 l. 24-26: It would be much easier for readers if the sampling locations in Fig. 1 would be marked with the same abbreviations introduced in this paragraph and used throughout the text."

Authors: Abbreviations for the sampling locations have been added to Fig. 1.

Referee: "p. 12981 l. 15, 22: It should probably read "stored frozen" in both sentences."

Authors: Corrected accordingly.

Referee: "p. 12981 l. 19: "8647–8049 ± 130 cal years BP" is an odd way of presenting calibrated ages. […] In order to not confuse readers, I would just omit this detail and refer to the respective reference."

Authors: Corrected accordingly.

Referee: "p. 12982 l. 12: How were the bound fatty acids released? Saponification?"

Authors: We use anhydrous hydrogen chloride (produced by adding acetyl chloride to methanol) for the acid catalysed transmethylation of fatty acids. Apart from methylation of the carboxyl groups of free *n*-fatty acids this method also breaks the ester bonds in polyesters such as suberin and in cuticular wax esters, i.e. compounds that were bound in these esters are released. Thus, we ultimately analyse both free and (previously) bound compounds. The relevant section in the methods chapter has been rephrased in order to avoid confusion.

Referee: "p. 12983 l. 16: Delete 'by'."

Authors: Done.

Referee: "Paragraph 3.2.1: I cannot follow the majority of %lipid values throughout this paragraph since they do not match the values given in Table 2. E.g., p. 12988 l. 18: according to the main text saturated FAs in SN and TP leaf litter "account for 35±1.4%" while Table 2 shows that n-FAs represent 35.0% (SN) and 38.3% (TP), respectively. How is the value "35±1.4%" calculated? Or p. 12988 l. 22: according to the main text saturated FAs in GN leaf litter "make up only 24%", MUFAs and PUFAs 22% and 18% while Table 2 shows n-FAs represent 22.5% and MUFAs and PUFAs 20.6% and 17.3%, respectively. P. 12989, l. 11: text states topsoil relative FA amounts "24-29%", table shows 29.0 to 39.6%? All numbers should be carefully checked for consistency and rounding errors throughout the manuscript."

Authors: Numbers were checked and corrected in the text and Table 2 where necessary.

Referee: "p. 12992 l. 3-9: This paragraph is confusing and should be rephrased. It is not entirely clear whether the n-OH distribution is discussed for the leaf litter or the topsoils. It seems only the litter is discussed which implies that suberin monomers are not a source for the OHs (or the definition "leaf litter" is wrong). I guess this source assignment also includes the soils, which should be mentioned in the text."

Authors: We have modified the paragraph for clarification.

Referee: "p. 12992 l. 10-14: The conclusions in this paragraph and Fig. 4 should be carefully re-assessed. The caption of figure 4 summarizes that "The patterns illustrate enrichment of mid-chain FA ($C_{22,24}$) from leaf litter to soil and of long-chain n-alcohols ($C_{26–32}$) 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." However, I think this is an over-interpretation of the data. Looking at the original data in the supplement, one can see that the soil samples as well as the SN litter sample were extracted in duplicates. In my opinion, the reproducibility of the biomarker data does not allow to draw meaningful conclusions. For example, the C_{22} FA (one of the highlighted FAs) relative abundances in the two TP-S replicates amount to 14.5% and 10.8. In the TP-S/F sample, they are 9.4% and 10.4% and in the one TP-LL sample 8.6%. I would argue that the C_{22} FA abundances do not differ between these samples, but are all within the measurement

uncertainty (the amount of data points in one sample is not statistically significant, to deduce a real error all replicate biomarker data should be used). This uncertainty gets even worse for the less abundant biomarkers. For the C_{32} FA in sample TP-S or the C_{30} OH in sample TP-S/F the relative abundances are 2.8% and 7.0% or 22.9% and 10.8%, respectively, in each of the two replicate samples. While the replicate measurements are better for the SN-S and SN-LL samples, the overall measurement uncertainty - most certainly due to the fact that GC-MS runs were used for quantification instead of GC-FID runs - seems too large to allow tracing enrichment or depletion of biomarkers between samples.

Authors: As pointed out be the referee, the individual values for C_{22} FA (calculated by the referee as $\%_{FA}$ from the data in Supplement 2) do not differ significantly between soil and leaf litter. One has to consider, though, that compounds such as C_{16} FA are ubiquitous and may also be produced by microbial organisms living in the soil. In order to test for the assumed input of mid-chain C_{22} and C_{24} FA from suberin one has to compare the amount of mid-chain compounds relative to cuticular long-chain compounds such as C_{26}, C_{28} and C_{30} FA that were reliably determined (> $1\%_{lipids}$). Then, the ratios of $C_{22,24}$ over $C_{26,28,30}$ are greater for the topsoil material at sites TP and SN than for the corresponding leaf litter. They are slightly greater or near identical for the white rot-affected sample at site TP. This sample might be regarded as a special case, though. At site GN, the ratio is actually lower in the leaf litter than in the topsoil. High-altitude beech litter thus appears to contain high amounts of mid-chain compounds from the start, and we now highlight this discrepancy in the revised text. This observation may be explained by high initial amounts of suberin in beech leaves. (One has to remember that suberin is also present in plant tissue other than root material.) Altogether, we think it reasonable to argue that mid-chain suberin-derived compounds appear to increase from leaf litter to soil in the low-altitude habitat, at least. Whether or not the result from the high-altitude habitat (GN) is a robust feature is currently tested in a follow-up study using new samples.

We regard the uncertainty for biomarker quantification arising from the use of the GC-MS system as compared to a GC-FID system as very low. In order to maintain a high level of accuracy, we were running a standard mix containing a range of saturated and mono- and poly-unsaturated n-alkanoic acids, n-alcohols, sterols and 5α-cholestane as the internal standard of choice (42 individual compounds, altogether) after every 4-5 samples. This allowed to determine the response factors for the individual biomarkers in the major compound classes for each small batch of analysed samples, which is indeed absolutely necessary for the production of quantitative data using GC-MS. In fact, using compound-specific ions such as m/z=74 for n-alkanoic acids, peak areas from the total ion count can be corrected for co-elution of other compounds – which do occasionally occur during total lipid extract (TLE) analyses. Such a correction cannot be carried out for GC-FID data. We therefore believe that GC-MS data from TLE analyses in some cases can achieve even greater accuracy than GC-FID data. Separating compound classes through flash chromatography and individual analysis of collected fractions will produce semi-quantitative data for total lipid composition (e.g., due to variable losses on the silica columns) and is therefore unsuitable for the determination of changes in TLE composition.

Referees: “p. 12993 l. 5-8: According to Fig. 3b the leaf litter (and by inference ACL) of GN is dominated by n-C_{27} and SN/TP by n-C_{29}.”

Authors: Accidentally swapped in the text, corrected accordingly.

Referee: “p. 12994 l. 5-6: This sentence should be modified. According to Fig. 3a, low- and medium molecular weight n-FA proportions in the B-horizon are higher than in the topsoils (contrary to the notion of “very low amounts of saturated FAs”).”

Authors: The description is actually correct. As can be seen in the pie diagrams in Figure 3a, the amounts of s-, m- and l-FA (i.e. short-, mid- and long-chain FA, represented by green colours) together are clearly lower in the original B-horizon (“weathered subsoil” in revised figure) than in the topsoils.

Referee: “p. 12994 l. 11: In Fig. 3a n-C_{28} OH accounts for approximately 25%, not 45%.”

Authors: Typo corrected.

Referee: “p. 12999 l. 19-22: How big were the original water filtrate samples (this is missing in the supplement file)? I wonder whether the absence of the above mentioned compounds is an artefact of sample size and detection limit rather than it is indicative of in situ production in the sediments. Normally, huge amounts of water need to be filtered in order to obtain sufficient material for biomarker analyses. Even if several 20L Niskin bottles were deployed and filtered, this might just not be enough material (few mg?) to detect minor compounds. I would certainly expect that dinosterol (dinoflagellates) and the long-chain diols (eustigmatophytes) are produced in the euphotic zone, which is the authors point out in the subsequent paragraphs.”

Authors: We agree with the referee that the absence of these compounds is likely a result of the small sample size and included this consideration (now in Supplement 2).

Referee: “p. 13000 l. 19: The reference to Fig. 5 seems wrong here.”

Authors: Corrected.

Referee: “p. 13002 l. 17-29: This entire paragraph seems misleading. According to the first sentence, the $\delta^{13}C$ values of HMW FAs is discussed for the investigated soil samples (and only for these were HMW FAs analyzed). Soils, however, integrate a very heterogeneous FA leaf wax signal, i.e., a FA pool from a multitude of vascular plant species and over unknown timescales. Accordingly, the $\delta^{13}C$ difference between various FAs might simply reflect different leaf wax input sources and/or an input source (locally integrated) over the course of decades (or even longer, likely variable in their FA $\delta^{13}C$). I don’t think that the observations of Lockheart et al. (1997) can easily be applied here.”

Authors: The referee is correct in pointing out that soils integrate plant isotope signals over long periods of time. The key point here is that ubiquitous C_{16} FA and partially suberin-derived C_{22}, C_{24} FA integrate the plant isotope signature over the full growing season while the long-chain compounds from leaf waxes ($C_{26\text{-}30}$ FA) represent the autumn isotope signature, only. As Lockheart et al. (1997) describe a trend towards lighter values for leaf wax compounds towards autumn we hypothesise that this may contribute to the lighter values for the long-chain wax-derived compounds.

Referee: “p. 13003 l. 1-2: Change “Fig. 6” to “Fig. 5”.”

Authors: Not applicable (anymore); figures shifted to Supplement 2.

Referee: “p. 13004 l. 13-14: According to Fig. 6 the statement “while C_{16} FA is the lightest compound (−192 ± 10 ‰) in all soil and litter samples” is not correct. E.g., for GN-S C_{16} FA is the heaviest homologue.”

Authors: We have modified the text in Supplement 2 accordingly.

Referee: "p. 13006 l. 23 - p.13007 l. 2: What exactly is the climatic/environmental scenario prior to and following the dry and cool 8.2ka event? Prior to 8.3ka, TOC/N ratios indicate "macrophytes (15 ± 2.4), and may result from any mix of these with leaf litter, topsoils and a small amount of algal material", but following the event, TOC/N ratios rather indicate ammonium-rich soil input (to be consistent with the biomarker data) - but would such an input not fuel primary productivity in Ohrid (including algae and macrophytes) as indicated by the $CaCO_3$ record? Is there other evidence for enhanced surface runoff after 8.2ka implying climate was perturbated for longer timescales? Or are the low %TOC (driving TOC/N) after 8.3 ka an effect of dilution by %$CaCO_3$ (which is highest after 8.3ka)?"

Authors: With regard to the low TOC/N ratios of the post-8.2 ka sediments we can only speculate on the cause. Lake Ohrid is oligotrophic and has probably been ultra-oligotrophic, at times, and productivity is not only nitrate- but also phosphate-limited. There is no indication for enhanced productivity after the 8.2 ka event as the amounts of water-column derived lipids do not increase. Like before the 8.2 ka event, the biomarkers are dominated by terrestrial supply, i.e. the only post-8.2 ka sample suggests a return to previous conditions on the terrestrial side, with supply of less-degraded OM. Increased erosion of nitrogen-rich soil that formed on former lake sediment during lake level rise could be an explanation for the low TOC values and TOC/N ratios. TOC would be diluted while the nitrogen cannot be consumed without an increase in phosphate. Further research is required in order to establish the biogeochemical fingerprint of soils forming on lake sediments as substrate.

Referee: "p. 13008 l. 1-2: "Developing a mixing model for aquatic vs. terrestrial OM in the sediments using $\delta^{13}C$ and $\delta^{2}H$ therefore appears challenging". I am missing a plot showing $\delta^{13}C$ vs. $\delta^{2}H$. The combined isotope data for the C_{16} FA show that the different sources (LL, S, macrophytes) do plot in distinct ranges. Maybe this will allow for better constraints?"

Authors: Following the advice of referee 1, we have decided to shift the discussion of the isotope data into a separate text (Supplement 2). Further research is currently carried out to improve the data base for a future publication that will include the current data.

Referee: "p. 13011 l. 2-4: The statement "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." is based on total lipid concentrations per g DW. To judge the supply of lipids or trace a change in the biome, the concentrations must be normalized to TOC, otherwise OM dilution (by clastics?) and OM preservation cannot be ruled out to be controlling the sedimentary composition (not only during phases 1 and 2)."

Authors: see response to comment 3) above

Referee: "Figure 2: Why are some symbols filled and others are not? It is not explained in the caption."

Authors: Explanation added.

Referee: "Figure 3: For quick reference it would be helpful to add the station IDs to the litter, soil, and filtrate panels. Also, are the bar charts for the low altitude leaf litter sites representative for TP or SN? Or do they represent averaged distributions? This information should also be added for the macrophyte, water filtrate, and sediment data. In case of the sediment data, are the panels

representing the average of time slices 8.29, 8.17, and 8.11ka (which could be deduced from p. 12998, l. 4-5) omitting the 8.22ka (or so) data point? In both figures, the n-alkane y-axis labels and the C_{30}-hydroxy acid bars overlay and x-axis font sizes are inconsistent. In Fig. 3b the n-FA x-axis labels are missing."

Authors: Suggested changes in Figures 3a,b were carried out, information added and errors corrected.

Referee: "Overall, I wonder why the figure is divided into a) and b). Would it not be easier for the reader to compare the different samples if all panels were combined in one figure?"

Authors: We will check with the editors if merging the figures is suitable with regard to typesetting.

Referee: "Figure 4: Why do the graphs contain regression lines of a suite of FAs or OHs instead of 1:1 mixing lines? If these plots aim at assessing preservation/degradation of biomarkers between litter and soil or soil and soil with fungus, respectively, a 1:1 mixing line should be shown."

Authors: The regression lines in Fig. 4 represent the best fit for those compounds that are assumed to represent the leaf litter end-member, thus, singling out those compounds that appear to be enriched in soil OM. A 1:1 mixing line would show the very same divergence but would actually not represent either the soil or the litter end-member.

Referee: "Supplement: It would be nice to include the isotope data in the attached file."

Authors: We provide a table with the isotope data in Supplement 3.