Manuscript Review *Biogeosciences BGD-12/12975/2015*

AUTHORS: Holtvoeth et al.

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

REVIEW COMMENTS

BGD manuscript 12/12975/2015 submitted by Holtvoeth et al. reports biomarker assemblages and isotope data for a suite of materials from the Ohrid catchment (including soils and leaf litter) and the lake itself (such as macrophytes and POC) attempting to decipher the endmembers contributing to the Ohrid sedimentary OC. The data are new and original, and within the scope of *Biogeosciences*. While it should be acknowledged that compiling this large data set certainly took a lot of analytical as well as processing effort, I think the manuscript would benefit from focusing on the relevant biomarker data used for paleoenvironmental reconstructions, thus, making the manuscript much more concise (without hampering the major conclusions). In addition, I have some general and specific comments (see below), which should be addressed to warrant publication in *Biogeosciences*.

GENERAL COMMENTS

1) I have one major concern about this manuscript directly related to the notion that an "improved endmember characterization" is provided. The soil OM data set is based on and all conclusions (particularly paragraph 4.2 and 4.3) 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, Zocatelli et al. 2014 (doi:10.1016/j.orggeochem.2014.07.016) demonstrate the spatial biomarker variability along an approximately 700m long peat transect showing that the immediate local vegetation cover strongly influences the soil biomarker inventories within one biome. Spielvogel et al. 2014 (doi:10.1007/s11104-014-2103-z) demonstrate the spatial heterogeneity of cutin- and suberin-specific monomers of soil horizons within 2.5m distance to trees showing highly variable concentrations per g TOC (implying changes of relative abundance vs. other biomarkers). 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 skeptical 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. The authors need to clearly address this problem in the manuscript and provide convincing arguments regarding the validity of the approach.

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.

2) 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 analyzed (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.

3) 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).

4) 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.

SPECIFIC COMMENTS

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

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

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.

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.

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

p. 12981 l. 19: “8647–8049 ± 130 cal years BP” is an odd way of presenting calibrated ages. It gives the impression as if either a 2σ probability distribution of a single sample is presented (with a wrong error) or two distinct samples (of which one is missing its error unless they were identical). In order to not confuse readers, I would just omit this detail and refer to the respective reference.

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

p. 12983 l. 16: Delete “by”.

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.

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.

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.

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}.

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").

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

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 artifact 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.

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

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.

p. 13003 l. 1-2: Change "Fig. 6" to "Fig. 5".

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.

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 8.3ka an effect of dilution by $\%CaCO_3$ (which is highest after 8.3ka)?

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?

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).

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

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.

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?

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.

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