Review on bg-2018-116 "Long-chain diols in rivers: distribution and potential biological sources" by Lattaud et al.

This manuscript investigates the distribution of long-chain diols (LCDs) in suspended particulate matter (SPM) of three river systems in relation with season, precipitation, temperature, and source catchments. Confirming previous results, riverine LCDs show a striking difference from marine LCDs, with the dominance of the C_{32} 1,15-diol in all investigated river systems. Higher concentrations of the C_{32} 1,15-diol are also observed in stagnant water and during seasonal river low stands. 18S rRNA gene sequencing of SPM from the Rhine and isotope incubation of the river water suggest that the LCD-producers in rivers predominantly reside in lakes or side ponds that are part of the river system.

General comments:

This paper is a valuable contribution to the understanding of LCD genesis and relationships with environmental variables in fluvial environments. Indeed, there is currently only a limited body of literature on LCD source organisms in general and even more so in fluvial environments. The authors address relevant questions and provide novel concepts and data. The writing style is clear and precise, especially in the methods section, which provides sufficient information to replicate the results. The experiments were conducted with rigour and on substantial numbers of samples per river. The interpretations are most of the time supported by the data. This manuscript is thus suitable for Biogeosciences.

However, the current manuscript can be improved before publication. The authors discuss their finding in context of seasonality but sampling was only done during two separate months per year, which makes difficult to make inferences about the seasonality. It would be also nice to show the non-correlation between temperature and LCD fractional abundances of different isomers in the data section and not as a minor mentioning within the discussion. This aspect could have been discussed a bit more in the context of stronger temperature gradients between lakes and air temperature versus rivers and air temperature. Those two points become even more relevant since the abstract suggests a focus on LCD in relation to season, precipitation and temperature.

It is also unclear why GDGTs (BIT index) and Chlorophyll a are relevant for the outlined questions or are helpful in understanding spatial distributions of LCD and their source organisms. Although Chlorophyll a is sometimes used as indicator of primary productivity, it has been shown that there are more suitable parameters (Lyngsgaard et al., 2017). The result description could also be a bit more concise and part of the results shown in tables. Why did the authors not take SPM samples integrated over a greater part of the water column? Villanueva et al. (2014) showed maximum LCD concentration a few meters below the surface, and even though this is based on a lake, stagnant parts of the river systems could have a similar vertical distribution.

Specific comments:

Page	Line	Comment
1	20, 26	Not only SPM but also sediment samples were investigated; "…in relation with season, precipitation, temperature, and source catchments" may be misunderstood as you making statements in all three rivers about seasonality even though SPMs were only sampled once during spring and once during autumn per location. Stating in the abstract that the relationship between LCD and temperature/precipitation was investigated and then only mentioning no correlation in the discussion may be perceived as misleading.
2	6	It would be clearer if you write that those culture experiments have been made on marine, lacustrine, soil and in snow living species.
2	7	Please be more specific. I think what you mean is that LCD signature of marine core top samples differ significantly from those of marine and lacustrine eustigmatophyte algae cultures. It would be good to state here that in marine versus lacustrine environments the C_{32} 1,15 is less abundant that the C_{30} 1,15-diol.
2	19	See comment page 2, line 6
3	16	SPM of which water depth interval?
5	1-2	Why were the filters of the Rhine and the incubation experiment base hydrolyzed (Page 5, Lines 1-2) and not the other samples? It would be worth explaining the reason why in the methods.
5	19-20	It is said that the C_{22} 7,16-diol was used as internal standard, while in Page 6, Lines 23-24, the C_{22} 5,17-diol is indicated as internal standard. Is it a typo?
7	28	Please use the names of the primers provided by Stoeck et al. (2010). Why has this primer been used when it only yielded low quality V4 reads in the original paper? Organism-specific abundances may be biased by the quality of primer annealing to the template. It should be mentioned that denovo sequencing has been done. These constraints should be discussed. More specific primers could probably be used in future work instead of the universal eukaryotic primer. Since so far LCD producers all belong to the heterokonts, a primer specific to this algae group adapted to NGS would potentially yield better results (Coolen et al., 2004, Bittner et al., 2012).
9	8	Why writing the unit with a dot ("ng.L-1")? More commonly written without the dot.
9	9	Since figure 1c is referred to before 1b, I would change the numbering of figures to match the order of their mentioning.
9	26-28	The LCDs from Black Sea sediments have been quantified but only the fractional abundances are discussed in the text and shown in Figure 2c. All the other data (from the Basin, Reservoir and Delta) are absolute quantifications. Why?
10	15	Why not providing a cumulative column diagram for the different groups found in the DNA analysis?
12	11, 28	Discrepancies between text and ternary plot labels. Plot suggests all C28, C30 and C32 diols but in the text it is written as C30 1,15 diols, C32, 1,15 diols…
13	11	Why does the DNA work in lake Challa by Villanueva et al. (2014) suggest a role of novel uncultivated eustigmatophytes in LCD production in riverine ecosystems?
14	2	It is known that different algae have different chlorophyll signatures and chlorophyll a is very common. It was therefore unlikely that a relationship could have been found. Additionally, chlorophyll a is not necessary the best indicator of primary productivity.
14	7	Did the authors also do incubation experiments on waters in dead arms? ^{13}C may have been unsuccessful because LCD may have been produced in situ during blooms and incubation experiments may have been done on post-blooming waters. Please include those aspects in discussion.

Notes on figures		
	1a	It would be good to extent the white frame of the overview map further so that the labels are all within white background. Since reading the actual elevations is irrelevant and within the work area also not changing, I would reduce the labels on the scale to 0 to 4500 m.
	1b,c	Since fig b and c use the same x scale, it would be better to put them closer together. Please write in the caption what those error bars represent. Standard error? 95%CI? Variability? (Same for Figures 1 and 3).
	2a	Scale as discussed in 1a.
	2b,c	Why BIT index is shown far away from marine influence? I guess there are no error bars for the Reservoir samples because there was only one sample. This should be specified in the figure caption.
	3a,b	Please write if the samples are from sediment or SPM. Why no error bars here? Why BIT index?
	4a,b	This Figure could be improved by using different symbols, for example in 4b different symbols could be used to distinguish more easily the culture samples from the SPM samples. The description of Figure 4a in Page 12, Lines 11-12 is different from the figure caption. The same applies for Figure 4b (Page 12,

		Lines 28-29). Please clarify. As the 1,14-diols were excluded, it would be appropriate to specify on the Figure "C_{30} 1,13+1,15-diols" instead of only "C_{30} diols".
	5a,b,c	Please harmonise axis labels concerning diol concentrations. Why do gene copies have sometimes error bars and diols concentrations never have them? Why is chlorophyll a concentration shown without error bars? Why does the LCD concentration sometimes have error bars and sometimes not?

References:

Bittner, L., et al. (2012). "Diversity patterns of uncultured Haptophytes unravelled by pyrosequencing in Naples Bay." Molecular Ecology **22**(1): 87-101.

Coolen, M. J. L., et al. (2004). "Combined DNA and lipid analyses of sediments reveal changes in Holocene haptophyte and diatom populations in an Antarctic lake." Earth and Planetary Science Letters **223**(1–2): 225-239.

Lyngsgaard, M. M., et al. (2017). "How Well Does Chlorophyll Explain the Seasonal Variation in Phytoplankton Activity?" Estuaries and Coasts **40**(5): 1263-1275.

Stoeck, T., et al. (2010). "Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water." Molecular Ecology **19**(s1): 21-31.

Villanueva, L., et al. (2014). "Potential biological sources of long chain alkyl diols in a lacustrine system." Organic Geochemistry **68**: 27-30.

Julien Plancq and Mike M. Zwick