Reply to the interactive comment of the reviewer 2 on "Long-chain diols in rivers: distribution and potential biological sources"

We thank Dr. Julien Plancq for his helpful comments on our manuscript. Below follows our reply to the main comments.

General comments

-''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."*

We agree with the reviewer that only two months of sampling does not reflect a whole year of primary production. However the climate of the Godavari River system is characterized by only two distinct seasons, i.e a wet and a dry season. The sampling months have been chosen to reflect these two seasons. Samples from the Rhine River have been collected in spring and autumn, and thus does not reflect all seasons of the area. This will be specified in a revised version of the manuscript. We will also extend the discussion on the non-existent link between LCDs distributions 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."*

We have used the BIT index as an indicator for marine input, especially in the lower part of the river. We have used chlorophyll a as indicator for primary production, as the data was readily available. Indeed, Lynsgaard et al. (2017) suggest that chlorophyll a failed to reflect the summer primary productivity in the Baltic Sea and that phytoplankton carbon biomass concentration is a better proxy. Unfortunately, we did not fix the phytoplankton and did not perform any biovolume counts. Therefore, we cannot obtain these phytoplankton biomass concentration values and chose to use chlorophyll a, imperfect as it is, instead.

The SPM from the water column of the Godavari River were taken at the surface (0 m), at the middle of the water column (2 or 4 m) and as close as the bottom as possible (8 m) to cover as much as possible the water column. The non relation with depth will be more emphasized in the revised manuscript.

Specific comments-Page 1, line 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."

For the Godavari River, the two months sampled are reflecting the two seasons existing in this area, i.e. the wet and the dry season. For the Rhine River we agree that sampling in spring and autumn does not reflect all seasons of the area, and this will be nuanced in a revised version of the manuscript. The Danube samples covers a large temperature gradient where we could have detected a temperature/diols relation. The relationship between temperature and LCDs distribution is only briefly mentioned as no correlation found but we will clarify and somewhat extend that discussion in the revised manuscript. The relationship between precipitation and LCDs distribution was especially investigated in the Godavari River system as the two months of sampling differ mainly in precipitation, and is discussed on page 9 lines 8-18 and pages 10-11 lines 31-32 and 1-3 of the initial manuscript.

-Page 2, line 6 *"It would be clearer if you write that those culture experiments have been made on marine, lacustrine, soil and in snow living species.*

We will add this to the text in a revised manuscript.

-Page 2, line 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 C32 1,15 is less abundant that the C30 1,15-diol."*

We will clarify the sentence and specify the difference between the LCD signature in marine and lacustrine sediments.

-Page 2, line 19 *"See comment page 2, line 6.*

We will specify that the cultivated isolates are derived from different environments.

-Page 3, line 16 *"SPM of which water depth interval?"*

We will add that most of the SPM were collected at the surface and that water column profiles were sampled at two locations in the wet season and one in the dry season.

-Page 5, lines 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."*

The samples for the Godavari and Danube systems have been primarily studied for GDGTs, and have been treated as such. The standard protocol for LCDs extraction includes a base hydrolysis step to allow enhance recovery. However, a recent study (Reiche et al., in preparation) showed that base hydrolysis does not yield much more LCDs than direct extraction. Hence we expect that there will be little difference in the results between the two work-up procedures.

-Page 5, lines 19-20 *"It is said that the C22 7,16-diol was used as internal standard, while in Page 6, Lines 23-24, the C22 5,17-diol is indicated as internal standard. Is it a typo?"*

This is a typo, it should be the C_{22} 7,16-diol, and will be corrected.

-Page 7, line 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)."

The names of the primers are TAReuk454FWD1 and TAReuk454REV3 (Stoeck et al., 2010). The original paper compares V4 sequencing and V9 sequencing. While V9 primers seem indeed to capture the highest level of diversity, the V4 primers do not underperform at all (see Fig. 4 on Stoeck paper). These primers (Tareuk) have been extensively used for many years after the original publication and they are considered among the best primers for targeting microbial eukaryotes in the environment. Furthermore, the V4 region is a longer fragment than the V9, this allows a higher resolution in discriminating species from the same genus or clades within the same species. More specific primers can be used, like in Villanueva et al. (2014) where they used specific eustigmatophyte primers. However, we wanted, in this study, to look at all potential diol producers that includes organisms outside of the heterokont group (like *Azolla filliculoides*, Jetter et al., 2009).

-Page 9, line 8 "9 Why writing the unit with a dot ("ng.L-1")? More commonly written without the dot."

We will remove the dot in the revised manuscript.

-Page 9, line 9 "Since figure 1c is referred to before 1b, I would change the numbering of figures to match the order of their mentioning."

We will change the numbering of the figures accordingly.

-Page 9, lines 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?"

The Black Sea sediments have only been used to identify the marine end member in the ternary plots which is based on fractional abundances. We do have the absolute concentrations of the LCDs in the Black Sea and will add them to the results section of the revised manuscript.

-Page 10, line 15 "Why not providing a cumulative column diagram for the different groups found in the DNA analysis?

We will add a cumulative diagram in the supplement with the DNA sequencing results.

-Page 12, lines 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…"

The legend of figure 4 is correct, we will correct the text to indicate that the three axes are the C_{32} 1,15-diol, the C_{28} 1,13-diol and the C_{30} 1,13 + 1,15-diols.

-Page 13, line 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?"

Sequences of novel uncultivated eustigmatophytes were found in lake Challa and their gene copies distribution fit with the LCD distribution, suggesting they may be a source for LCDs.

-Page 14, line 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."

We agree with the comments, but unfortunately we do not have the total chlorophyll available.

-Page 14, line 7 "Did the authors also do incubation experiments on waters in dead arms? 13C 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."

We did not perform any incubations in dead arms of rivers. We agree with the reviewer that the incubation could have failed because of the timing of the experiment (post-bloom) rather than because of the absence of the producers although it should be noted that we did detect incorporation in other algal biomarkers. We will add this aspect in the discussion.

Note 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."

We will make the white frame bigger and reduce the elevation scale.

-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)."

We will put the figures closer together to make the reading clearer. We will add to the figure legend that the error bars represent standard deviation.

-2a "Scale as discussed in 1a."

We will also extend the white frame of the overview map to make the label more visible.

-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."

The BIT values for the parts removed from the sea are only to show that the riverine end member is very different from the marine end member. We will specify in the figure caption that there is only one sample for the reservoir.

-3a, b "Please write if the samples are from sediment or SPM. Why no error bars here? Why BIT index?"

We will add that the samples are from SPM. The BIT index here allows us to compare with the Danube River to clearly show the difference between marine and riverine environments.

-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 "C30 1,13+1,15-diols" instead of only "C30 diols".

We will change the symbols for culture samples for figure 4b. The figure caption is right, the description should be C_{32} 1,15-diol; C_{30} 1,13 + 1,15-diols and C_{28} 1,13-diol. We will correct this in the text, and make a more detailed caption.

-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?"

Diol concentrations are only reflecting one point and one measurement, in contrast to the gene copies which are from triplicate analysis. Only one replica has been done to the chlorophyll concentration measurements, which is why there is no error bar. LCDs concentrations have error bars when they represent the average of a river section (for the Godavari and Danube systems). The LCD concentrations for the Rhine reflect only one site. We will specify this in the figure caption.