Reviewer #1:

The manuscript is well written, and the data analysis is very thorough and state of the art. The authors attempted to connect geochemical characteristics, DOM molecule formulas, and metagenomics based on 7 rivers across a large spatial scale. They basically concluded that DOM properties can be distinguished by river type and geochemistry of the rivers differed. They also showed that the microbes of the rivers shared core functional potential. While I appreciate the data and particularly the excellent data analysis, I feel that the conclusions are not that novel and I also have some concerns that the authors need to address.

Thank you very much for taking the time to review this manuscript.

We agree that our results were largely not presented in a way that emphasizes the novelty of our observations but were presented as providing context to a developing comprehensive conceptual model within the field. Despite this structure, we do have noteworthy observations that may have been lost in the text. Specifically, the observed glycoside hydrolase niche conservation, interactions between environmental conditions and DOM assembly, and relationships between organic matter richness and glycoside hydrolase diversity were all previously unreported. Importantly, each of these observations have potential implications when attempting to predict organic matter dynamics within and across watersheds

We will plan to improve our language and rearrange text to better emphasize our novel observations while balancing the limitations inherent to these large-scale metagenomic/DOM characterization studies.

The several rivers seemed to be randomly selected without any justification. It needs to be clearly stated why these rivers are selected and how representative they are in terms of the world's rivers. It is also stated the river types, wastewater vs headwater, but more concrete data or logic connections are needed. What are the nutrient data? Other than the RDA in Figure 2, there is no nutrient data reported in the manuscript. Also, how exactly can you connect DOM to the wastewater, directly input of DOM or through nutrient-inspired algal blooms? In general, statistical analysis is fancy but there is little or no mechanistical connection. I feel that this shortcoming is throughout the manuscript, such as the connection between DOM and metagenomics.

These rivers were selected from a previous study based on the following criteria: 1) each of these rivers experience some sort of diel dynamic, and 2) these collaborators volunteered to collect samples. The analyses in this manuscript were performed on data collected for a different purpose as described in the methods (see Lines 168-172 in the original manuscript – "*Samples from the seven rivers were collected between July and October 2018 as part of a WHONDRS initiative to study diel river dynamics (Stegen and Goldman, 2018). As such, the sampling scheme was similar to the methodology described in Danczak et al., 2021 and was originally devised to investigate temporal patterns despite this study focusing on spatial*

variation.”). We will attempt to add in extra context throughout the manuscript to ensure that readers understand the limitations of how these data can be interpreted.

We are not completely certain “nutrient data” means in this context (e.g., we read this as nitrogen or phosphorous loadings). However, we think there might be confusion in the text – we were not trying to assert from the DOM that the rivers were a given categorization or to connect DOM to different input types. Instead, rivers based upon their historical and environmental contexts as described in Section 2.1 under the Methods section. For example, the Erpe River is a wastewater impacted river because samples were collected downstream from the effluent from a wastewater plant. Furthermore, we do not believe that these rivers are intrinsically representative of these classifications broadly – for us, these were simply categorical variables that we could use in our conceptualization to understand divergent/convergent patterns.

Finally, we were not attempting to develop mechanistic connections specifically. Instead, we were focused on ecological connections as these tend to be more predictable across scales. As such, we were not focused on examining explicit metabolic connections (in part due to the absence of metatranscriptomic or metaproteomic). We will elevate this ecological perspective through general language changes to prevent confusion.

A set of geochemical parameters were selected for the work, including Cl, Mg, TN F, Fe etc., but why? For example, why not Chla and why not dissolved oxygen? Chla would be very straightforward to connect to DOM and microbes. I am not saying you have to include Chla, but need to justify why you chose this specific set of parameters among the numerous choices.

These parameters were selected due to available instrumentation and given that they provide a broad context to local geochemistry. These parameters allowed us to deeply contextualize each location in a continuous sense rather than rely purely on *a priori* assumptions. Given that the main goal of the dataset originally was to understand the association between DOM diel cycling and environmental context, these provided a broad suite of parameters to evaluate DOM. We will add a sentence stating that these variables offer insight into general geochemical processes.

The authors used metabolome for the DOM characterization. I am not very sure this is an accurate definition as it is assumed all the formulas obtained from FTICR-MS are metabolites. I don’t think this is true because there could be contribution from abiotic reactions or selective preservation.

We appreciate this concern – to address this, we will only use “metabolome” when referring to the meta-metabolome ecology to eliminate confusion with the method. Outside of this context, we will refer use either “DOM” or “DOM assemblage” depending on the situation (e.g., chemical vs. ecological contexts, respectively).

FTICR-MS is a non-quantitative technique; thus, it is great that the authors chose to use 'presence or absence" to process the data. But this is still tricky if you don't inject the same amount of carbon (or DOC) in samples when you are trying to compare them. In other words, the absence or presence of a specific molecule may depend on the matrix or DOC concentration. Some QA/QC on this aspect needs to be added.

We agree with these concerns, though we do inject samples at a standardized carbon concentration (Line 213: "*samples were standardized to a given carbon concentrations (NPOC 0.69 – 1.5 mg C/L)"*) and our replicate approach should act as a control for the presence/absence of molecular formulas. That being said, we will try to add context of some inherent limitations in FTICR-MS-based analyses.