Drapeau et al., Response to Reviews
BG-2023-121 Revision 2

We would like to thank the editor and anonymous reviewer for their comments and guidance, which have substantially improved our manuscript. Responses to comments are provided below, in blue, italicized, text.

line 97: space after "the"
Done.

line 105: "a priori" repeated in both sentences – suggest rewording. It should also be in italics here and throughout, I think (along with 'in situ')
Thanks for pointing out this redundancy in wording. Edits made at this point in the text, and both a priori and in situ have been italicized throughout.

Figure 1: text on both panel of figure 1, but especially panel A, should be enlarged, as it is almost impossible to read them. Also, perhaps it could be possible to overlay the path of a given stream with a line (maybe in different colors by stream?) so that we can see their placement on the map?
We have re-made Figure 1 as a stacked figure, to allow for larger text sizes. Hopefully this enables the lines denoting the streams to be a bit clearer, too.

line 236: maybe explicitly state here what samples are being pooled? It's a bit confusing as written
This is now clarified in the text: Any samples showing an expected band size of ~400bp were purified using a bead cleanup protocol. The cleaned bands were then pooled to make the final library for sequencing using 5 uL of each sample; if a sample was more faint than other bands, 10uL of that sample was added.

line 244: BSA in parentheses?
Done.

line 243-245: seems like this info should come before the sequencing information?
Agreed. We have moved this information to the beginning of the section to help with flow.

line 254-257: Im not familiar with PhiX…..would it be possible to briefly explain what this is and what it does? Also, would this need to be also outlined above somewhere?
The following information on PhiX has been added at this point in the text:
[PhiX] is a quality control reagent used commonly in sequencing runs to optimize cluster generation, sequencing, alignment and calibration control throughout the run. Because PhiX is a well-defined bacteriophage genome, it has a diverse base composition that provides the balanced fluorescent signal that low diversity sample libraries, like ours, lack, during each sequencing cycle (Illumina).
We also add a reference to this added text where PhiX is briefly mentioned, above.

line 257: Although 5,000 is not terribly many at the end of the day, so would probably make sense that they do not reach a plateau at this point? What kinds of range did you have with your sequence numbers per sample? Lastly, be careful in calling these ASVs 'species'
Yes, agreed. We have added a reference to the supplemental figure that shows rarefaction curves to give the reader a sense of range of sequence numbers per sample. Thanks for pointing out that we use species erroneously here: this has been changed to ASVs.

line 289: were alpha diversity measures calculated on rarefied data? And what was the rarefaction level? I saw that you threw out everything below 5000, so was it 5000? this should be explicitly stated somewhere (although I realize that rarefaction may not be necessary if all samples plateau, in which case im not sure why rarefaction is mentioned here? just some clarification necessary)
Alpha diversity was calculated on untransformed data; this has been clarified in the text. We double checked our text on this point; rarefaction only comes up at the end of this paragraph, where we state that we compared our Hellinger-transformed NMDS and perMANOVA analyses using a rarefied dataset. We've added a few words for clarity at this point in the text.

line 294: I thought that the mid sites were excluded from the analysis?
Thank you, this is a typo. Mid should read near and has been changed.

line 323: headwater and near sites? Try to keep all terminology similar throughout the paper so that the reader can keep it all straight
This is correct as written: our distance bins are headwater, near, mid (for chemical analyses only), and far. No edit made at this point in the text.

line 365: A-ha….here are the read details….might they be better above? I leave this to the discretion of the editor and authors. Also, is 'Rarefaction curves' a proper noun?
Agreed that these details are better in the methods section. We have moved and integrated this section into lines 321-323 (tracked changes version). We have also removed the capitalization for "rarefaction".

line 378: might it be worthwhile to give the genera names for some of these common ASVs? Family names (for example) do not provide a lot of information, although its better than nothing of course
Thank you for your comment. For amplicon sequencing runs such as ours, we are rarely able to resolve down to the genus level, even when using the fairly high confidence threshold of 0.8, when aligning our sequences to the taxonomy database. This kind of reporting is usually rare in 16S environmental publications, especially for oligotrophic environments such as ours. We also lose some confidence the further down the taxonomic tree we go. As it is, resolving down to the family level is usually the best we can do. Indeed, only 2 of the top 10 families actually resolve into genera. For this reason, we will not specify genera further for these top 10 families.

line 382: would it be possible to give some hard numbers on the diversity values here? Its great to know that they change from upstream to downstream, but neither the text nor the figures really gives the reader a sense of the magnitude of change, or what is there in the first place.
Thank you for the suggestion. We have added median diversity values pre and post-melt to illustrate the difference in alpha diversity between near and far sites throughout the melt season.

line 402: have you used this ISA acronym before?
Yes, ISA (indicator species analysis) is first defined at L304. No change made at this point in the text.

line 405: again, be careful in the use of 'species'
Thank you; changed in both cases to ASV.

line 451: consistent use of 'carbon' versus "C"
C is switched to "carbon"

Figure 7: In panel A, both the Far sites and the intersection of Far sites with Headwaters have the same values (73.94%, 306,996).....this is a mistake, no?
Yes, this is a typo – thanks for catching our error! It has now been corrected (now Figure 6).

line 540: Again, while true, these are pretty coarse/vague taxonomic entities
Yes, we agree, but we purposefully retained this text in the last iteration to allow for broad comparison with other studies, which often present results at this taxonomic resolution. Later in the paragraph, we provide results at a finer taxonomic resolution. No edit made.

line 597: should mass effects be explicitly defined?
We have added a definition as "homogenization by high rates of dispersal", similar to the definition provided in the abstract.

line 614: assemblages rather than communities?
We have replaced community with assemblage at this point in the text, and have done one last "find and replace" to ensure consistency throughout.

Conclusions: this sections seems a bit long for me for a conclusions section – also don't know if references are appropriate here, but will leave that to the discretion of the journal
We have left the conclusions section as-is, but are happy to take further suggestions on length or suitability of references.

Figure 2: match letter cases with text and figure (i.e. 'a' versus 'A')
This has been changed in the figure caption.

GENERAL COMMENT ON FIGURES: if there would be any way to create some sort of color / shape scheme that would be possible to unify across figures, that would greatly aid in interpretation for the reader. Right now, each figure really has its own key in terms of what color / shape means, and it is hurting my head a bit going from one figure to the next
Yes, one unfortunate side effect of the move to the mixed effects models was that we had to expand on our previously unified colour scheme. However, we have put some thought into this, to make sure that our scheme was consistent, with different base colours to represent the different factors that we sample across, as appropriate for the figure in question. Our scheme is as follows:

- *Our figures primarily differentiate near/mid/far sites, which was our original schema. For these plots (Figures 1a, 4-9), we use colours across a range of green, tan, and brown, often with symbol shape differentiating between hydrologic periods. Notably, we retained this plotting for all microbial plots, and for plots where binning was useful to show distance downstream, but analysis was not via a mixed effects model (compositional PCA, 13C-14C biplots)*
- *With the move to the mixed effects model and edits implemented following the initial review, we moved some plots away from distance bins, to enable plotting of data along a distance gradient. This approach was implemented for all plots associated with a mixed effects model. Here, the plotting scheme de-emphasizes river and year (random effects in our model).*
    - *For Figure 2, we plot to emphasize fixed effects either on the x-axis (distance) or via faceting (hydrologic period across columns), and then use colours (river) and symbols (year) for the random effects.*
    - *For Figure 3 (which now combines previous figures 3 and 5) we plot to emphasize distance (x-axis), and use colour (rather than faceting) to illustrate hydrologic period. Figure 1b uses this same colour scheme to show hydrologic period.*

Figure 3: would it be possible to break this figure up by river (i.e. make separate panels by river)? Right now there is so much going on that its really hard to see any pattern. If there is not a pattern you want to highlight, could alternatively merge with figure 5 to create a new panel?
We've chosen to not break up by river, because we control for this (as a random effect) in our models. Given that we don't focus too much on the original figure 3 in the text, we now combine both Figure 3 and Figure 5 into a new, 3-panel Figure 3.

Figure 4: these are for all sites and seasons combined? Should specify this…
Yes. This is now specified in the figure caption.

Figure 8: what exactly are the different groupings shown by the circles? Is there any way to show river name as well? 'holms' should probably be capitalized, and the 2 in R2 superscript
(Now Figure 7). We've tried to clarify further in the caption: circles show that the far sites differ significantly from the headwater and near sites, but that headwater and near do not differ from one another.
Plotting by river (and, also year) is shown in the Supplement, and we now point the reader to these plots in the figure caption.