Interactive comment on "Co-occurrence patterns in aquatic bacterial communities across changing permafrost landscapes" *by* J. Comte et al.

Anonymous Referee #1

Received and published: 18 August 2015

I was very pleased to read the manuscript "Co-occurrence patterns among aquatic bacterial communities across changing permafrost landscape" by J. Comte et al. The paper is well written and it takes an advantage of the statistical tools that are available for this type of data. The paper describes similarities and differences in microbial communities residing in thaw ponds within and between 5 regions in the subarctic Canada, along permafrost gradient. It concludes that these communities are heterogeneous, just as these environments are, but still share many key taxa. Further, they suggest that bacterial communities within a valley are more similar at the areas of extensive melting of the permafrost and that the properties of the bacterial network make these communities robust against environmental change. Authors have done an extensive statistical analysis on the community data. What I was missing was the environmental data. It is stated multiple times that especially DOC and conductivity had a big impact,

but would be good for the reader also to see how big differences there actually were among and within valleys. Also please include to the table column indicating how the valleys are situated in relation to permafrost degradation gradient. Another issue that should be taken into account through out the ms is that usually classifications of 16S genes, especially for such a short fragments as here, only suggest that the OTU is closely related to certain organism. Thus, authors should be very careful in making conclusions about the metabolic capacity of an OTU based on 16S.

Specific comments

L65 "Specifically, patterns of co-occurrence may reveal to what extent groups of microbes share habitat preferences, to what extent there may be ecological linkages among bacterial taxa and with other planktonic organisms, and the extent of phylogenetic closeness of co-occurring bacterial taxa." I'm not sure if the patterns of co-occurrence per se reveal the extent of phylogenetic closeness of co-occurring bacterial taxa. Please explain.

L271 I have problems understanding this sentence. Do the authors mean that communities within SAS were similar and different from communities from other valleys, or that within SAS communities were more similar than within other valleys? I'm assuming it is meant to be the latter option, how ever; I'm not fully convinced that SAS and KWK had more similar communities than others. I get the impression that this conclusion is based on visual inspection of the figures. Was, for example, average distance to centroid shorter for those valleys than for others?

L279 This sounds like the result was depending on the location in the analysis; please rewrite.

L293 Were these ponds physically connected to each other? Or is there some other explanation for close relatedness?

L295 Could you please state the indices that were used to estimate species richness

and diversity?

L319 How reliable is this classification? Are you sure it is Opitutus, or is it a bacterium closely related to it? How were the oxygen conditions in the ponds? Methods only state "surface water", which specific depth is this and were the oxygen conditions comparable among the ponds? Oxygen is a major determinant for microbia, and if some of the ponds have oxic and others anoxic conditions in the sampled depth, how do you justify mixing these different redox conditions in a single co-occurrence analysis?

L326 I was not able to interpret which OTUs in table 2 are these discontinuous and sporadic –associated OTUs?

L332 How big proportion of the community did these OTUs comprise?

L341 Do you have any explanation why the community in BGR appeared to be randomly distributed? Any environmental factors that could explain this?

L347 Sentence starting with "In general. . .". I don't understand what the authors' message in this sentence is.

L364 What about relationship with the actual abundance of these OTUs? I think the number of organisms within each of these OTUs is much more relevant than the number of OTUs from the same phylum. Were the most connected OTUs also the most abundant ones?

L381 Could the authors please clarify how the reduced network was constructed. The text mentions removal of the most connected nodes while the table legend says "network centered on the 24 most connected OTUs".

L432 DOC concentrations?

L435 The order in which these valleys are on the permafrost gradient is very important for interpreting the results. Please include it to a table.

L474 Overdispersed?

L515 Again, how were the O2 conditions among the ponds? Is it realistic that this OTU really is an anaerobe? Or just closely related to one?

L563 "The bacterial networks similarly showed differences among valleys, corresponding at least in part to differences in DOC and conductivity." Maybe I missed something, but I got the impression that there was only 3 networks; based on real data, reduced real data and a random network. Also table 4 shows data only for these 2. Now I'm a bit confused which are these networks that show differences among valleys?

Figure 1. A legend illustrating the colors in panels a and b would improve the readability.

Table 2. Legend and data in conflict regarding numbers presented for each site. Legend states top 10 per site. How ever, NAS only has 9 taxa while SAS and KWK have 14.

Table S1. Does number of taxa correspond to number of OTUs?

Interactive comment on Biogeosciences Discuss., 12, 10233, 2015.