Quebec City, 28 October 2015

Dr. Jorien E. Vonk,
Guest Editor,
Biogeosciences,
Email: j.e.vonk@uu.nl

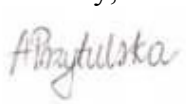
Dear Dr. Vonk,

Phototrophic pigment diversity and picophytoplankton abundance in permafrost thaw lakes by A. Przytulska, J. Comte, S. Crevecoeur, C. Lovejoy, I. Laurion & W.F. Vincent

Thank you for your continuing help with our manuscript and for providing the two sets of referee comments. We are pleased that the reviewers found the results to be valuable and interesting, and we greatly appreciated their detailed comments to strengthen the manuscript. Both reviewers requested a full revision of the Discussion to improve integration and clarity, and also more information concerning the methods; we are therefore paying special attention to these aspects. The revised manuscript will address each review comment as detailed below.

We look forward to further instructions from you concerning the upload of the revised manuscript.

Sincerely,



Anna Przytulska
cc all authors

PROPOSED REVISIONS IN RESPONSE TO REVIEW COMMENTS

Anonymous Referee #1
General comments:
The authors used a set of techniques (pigments, flow cytometry, epifluorescence microscopy, molecular analyses) to characterize and compare the phytoplankton of thaw lakes in northern Quebec, Canada. Although the effort is very valuable, it is necessary that the results obtained from the different techniques be better integrated to improve the discussion.

Thank you for this detailed evaluation and very helpful suggestions for improving the manuscript. These are being addressed in our revised manuscript as follows:

1. For example, the picophytoplankton fraction was analysed by flow cytometry. What was the relative contribution (%) of each fraction, pico-cyano and pico-eukaryotes, to the total pico fraction?

Information on % contribution is included in the revised manuscript.

*2. In Results, page 13, line 25, the authors said the picocyanobacteria abundance in KWK23 was 5.6 *10^5 cel/ml. Then, looking at figure 6 (biovolume), the biovolume for that sample/year was around 5 *10^5 um3/ml. Doing some simple calculations, and assuming 7 ug Chl-a/mm3 of picocyanobacteria biovolume (but please, see Reynolds 1984, The Ecology of Freshwater Phytoplankton), the concentration of chlorophyll a due to picocyanobacteria was in the order of 3.5 ug/L. Looking at the total Chl-a data, line 10, page 11, the relative contribution of picocyanobacteria to the total of the phytoplankton community is important (dominant?). Can the authors discuss this point more in detail?*

We have redone this calculation; if 7 µg Chl-*a* = 1 mm^3 = 10^9 μm^3 we get a value of only 0.0035 µg Chl-*a* L^{-1}, which implies that the picophytoplankton make only a very small contribution to total phytoplankton biomass. We will recheck these calculations and other analyses of Chl-*a* per unit biovolume, and will note this in the text.

Questions:
1. The molecular data needs to be better linked with the rest of the study. Why did the authors select the eukaryote fraction to do molecular taxonomy (excluding cyanobacteria, 16S RNA)? And, why is it relevant to describe and discuss the non-phototrophic taxa (predators: ciliates, fungi)? Most of the phytoplanktonic taxa identified by molecular analysis are in the fraction of nano to meso plankton: how does this information match with chlorophyll-a and the contribution of picoplankton to the community?

Our flow cytometry data indicated an abundance of eukaryotic picophytoplankton, but they could not be identified. The molecular data, although limited to two depths of one of the lakes (the site that has been focused upon by many collaborators in the overall program) provided unique insights into the taxonomic composition of this small size

fraction, and showed the potential importance of green algal picoeukaryotes. Additionally, the 18S data nicely complement and strengthen the pigment analysis by showing the relative abundance of reads in the different phyla. Although the information on the ciliate and heterotrophic nanoflagellates was simply a bi-product of the 18S analysis, there is great interest in these findings given their large grazing potential on picophytoplankton, and it seems appropriate to include them within the table. We are making these links and rationale clearer in the revised manuscript. Concerning 16S RNA, this is now published in Crevecoeur et al. (2015). We have now revised the text of the Discussion to connect to this study.

2. Why did the authors not analyse the nanophytoplantkon fraction with an optical microscope? And why was it necessary to use indirect techniques to infer the phytoplankton composition? Please, justify.

The aim here was to focus on HPLC pigment signatures, as in many limnological and oceanographic studies, as a first analysis of phytoplankton abundance and phylum composition in these lakes, which are representatives of an extremely abundant ecosystem type: thaw lakes in permafrost landscapes. This allowed an analysis of community structure (major phylogenetic groups) at many sites as a function of environmental gradients (degree of thawing of permafrost, DOC, TSS etc). It also allowed a comparison of light-capturing and photoprotective pigments. We have now strengthened this rationale in the Introduction. We did take samples for nanophytoplankton enumeration by microscopy, but this separate large dataset is still under analysis, and is beyond the scope of this manuscript, which already encompasses large data sets.

3. The concepts of abundance, concentration and biomass are confused in some paragraphs. In the objectives it is stated: "A secondary objective was to determine the abundance and distribution of picocyanobacteria and picoeukaryotes". Then, in the Results the authors present abundance and biovolume without a clear differentiation of both indicators. For example, in Results, page 13, subsection: "3.3 Picophytoplankton abundance", it is not clear when the authors describe information about abundance or biovolume. While the text refers to abundance, figure 6 presents biovolume (with no corresponding description in the text). Both variables are complementary but conceptually very different. In page 14, from line 11, is the analysis made with picoplankton abundance or biovolume? This needs to be specified. I would suggest that biovolume be used to explore correlations with environmental and biotic variables.

These complementary variables are being more clearly separated, with analysis of total picophytoplancton based also on biovolume.

4. The Material and Methods section has to be improved. The methodological design is complex and should be justified with more detail. Not all the analyses were performed for the same number of samples, lakes and dates. This makes it difficult to follow the results. For example: not all the analyses and sites were sampled on 2011 and 2012 at the two different depths (surface and bottom). It is necessary to explain how many samples, lakes, depths and dates where used for each analysis and why.

We now provide more detailed descriptions of study sides, sampling and statistics.

5. The statistical analysis section has to be described with more detail. Please, explain why PCA was selected (what was the gradient length of the data?). Using the pigment composition as a proxy of main phylogenetic phytoplankton groups, the authors could explore the % of variance of biological data explained by the environmental data (i.e.: multivariate analysis like CCA or RDA).

We acknowledge and understand the reviewer's concerns relative to the use of PCA. In Figure 3, we presented the environmental data only. PCA is a powerful and appropriate approach to visualize the clustering of sites based on a set of quantitative environmental data, and because the variables were expressed in different measurement scales, we computed a PCA on a correlation matrix that represented the covariances of standardized variables (as in Legendre & Legendre 1998). The questions we addressed were: how are these variables correlated? What can we learn from the ordination of the sites? In other words, are there specific environmental variables that are characteristic of a particular location. For example, our analysis indicated that DOC content was an important variable for the SAS valley.
Although PCA is a good way of exploring the distribution patterns in environmental data, we are aware that it has limitations for analysis of species matrices especially because it preserves Euclidean distances, which are known to be a poor descriptor of beta-diversity (Legendre & Legendre 1998). We thank the reviewer for suggesting using a canonical ordination technique to explore the link between environmental and species matrices. We agree that our HPLC pigment matrix could be seen as a composition or trait matrix, and we have followed this suggestion to use RDA analysis for the revised manuscript. This allowed us to investigate the extent to which the variance in the distribution of pigment traits can be explained by the measured environmental variables. For this analysis, we used a Bray Curtis similarity matrix as it bounds between 0 and 1 and therefore allowed comparison of the similarity among samples. This metric is not Euclidean, and therefore we performed a distance-based redundancy analysis (db-RDA). These results gave a significant pattern, but also reaffirmed the large lake-to-lake variation in each of the valleys.

Cluster analysis: I would suggest another kind of analysis to compare the sites defined by environmental and biological data (see above). I found the comparison of the two clusters too indirect and poorly supported in terms of statistical significance.

We thank the reviewer and we have followed this suggestion by exploring using a db-RDA how the pigment composition matrix was constrained by the ensemble of environmental variables. We agree that cluster analyses have limitations, however, the idea behind using this approach was to test whether the pigment composition patterns could be related to a particular configuration of the environmental conditions among sites. Cluster analyses are not statistically supported; therefore we performed permutations ANOVA and Mantel tests to validate whether the patterns detected in the clusters were significant. The cluster analysis permanova showed that no significant differences among valleys. Mantel tests further showed that there was no relationship between environmental

conditions and pigment composition. We acknowledge, however, that the Mantel test has limitations in terms of statistical power. We have therefore deleted this analysis from the revised draft in favor of the new db-RDA.

In any case, more information about the cluster analysis needs to be presented (which kind of cluster, distance or similarity, which index, which averaging method, which matrixdata, etc).

The cluster analysis is now deleted.

The authors compare two clusters built by two different indices "by eye"(distance: is it Euclidean?). Is it possible to identify different groups of lakes based on the environmental data, since the distances are very similar? Regarding the clusters based on biological data, and assuming 40% of similarity as a parsimonious cut point, it is possible to find only two groups and one outlier (2012SRB1).

The cluster analysis is now deleted.

6. Pigment results: Please, analyse pigment ratios to chlorophyll-a based on micromoles and not micro-grams. Micro-moles/L is not influenced by the molecular weight of each pigment and gives the information about the quantity of molecules of each signal pigment in the total. Since the authors wants to describe the composition of the community, I suggest using micro-moles instead of micro-grams.

The pigment data have been re-calculated in nanomoles and the statistical analyses redone.

It is noteworthy that chlorophyll c (any variety) was not detected when carotenoids such as fucoxanthin, diadinoxanthin and peridinin were found. What is the explanation?

Low concentrations of chlorophylls c1, c2, and c3 were indeed detected, but generally at trace levels. We have now noted this in the revised manuscript.

The classification of photoprotective and photosynthetic pigments, as presented in Table 2, is not clearly discussed. And what was the total photoprotective/total photosynthetic pigment ratio? What are the consequences in these differences?

A more detailed explanation for both groups of pigments is being provided in the revised text, with discussion of the observed ratios, which are now presented in new Table 3.

Specific comments:

1. Doing some quick calculations for 2011 data presented in figure 6, the individual size of picocyanobacteria cells in SAS1 was very big (~ 2.3 um3) in comparison with KWK23 (0.89 um3). It would be interesting to explore and discuss these differences.

All data for picophytoplankton in the studied lakes are now reported as biovolumes in the text and figures.

2. All the information presented in table 3 (bacterio-chlorophyll) is not well discussed and it does not flow with the rest of the article. I suggest removing this section.

We were surprised by this finding and given the magnitude of this pigment concentration we would prefer to report it. It is true that it is a non-eukaryotic pigment, but it complements the information about another group of prokaryotic phototrophs, the picocyanobacteria This is a new habitat type where bacterio-chlorophyll is described for the first time. We are revising the Discussion to better integrate this information.

3. Figure 4: I suggest reformatting this figure. It is not easy to follow the differences between carotenoids and sites. The legend of this figure needs to be improved so as to give more information.

In response to this comment and also to comments by the second reviewer we have removed this figure; all this information is now given in Tables 2 and 3.

Anonymous Referee #2
General comments:
The manuscript by Przytulska et al. studied the phototrophic communities in permafrost thaw lakes of subarctic Quebec, mainly through specific pigments analysis, flow cytometry and molecular methods. It is suggested that the diverse phototrophic groups and abundant picophytoplankton in those special ecosystems could potentially contribute to higher trophic levels and lessen the release of GHGs. While the sampling design is sound and the results are interesting, I have some comments and suggestions on improving the quality of the manuscript.

Thank you for this critical evaluation and for your very helpful suggestions for improving the manuscript. These are being addressed in our revised manuscript as follows:

Questions:
1. There's a general lack of information on methodological description. For example, what analysis system, scanning atlas and quantification calculation is used for the HPLC analysis? What is the relationship between phytoplankton groups and specific pigments? To what extend the CHEMTAX is applied or not at all?

The description of methods is being improved in the revised version of the manuscript, including more detailed description of the HPLC analysis, reference spectra and standards. CHEMTAX requires a very good cross calibration with phytoplankton enumerations, which is not available at this time for this ecosystem type; we therefore did not apply CHEMTAX and note the opportunities for such an approach in the future.

There's no clarification on the terms of "photoprotective, photosynthetic, and accessory pigments".

These terms are now defined in the revised Methods.

Unclear what sampling dates and layers (surface and/or bottom) were at each location, and this makes it hard to follow the results.

This information is now in the text.

No information on specific samples used for each analysis, e.g. What samples are used to run the correlation analysis between picocyanobacteria and temperature? Are the bottom waters included as well? Please at least include the information of P value and observation numbers for each statistical analysis.

More detailed descriptions of the statistics is now included in the revised manuscript, including N and p values.

2. Another issue is the inconsistency and complexity of samples and methods chosen for different statistical analysis. Could this be a potential cause for the "insignificant" results/relationship of variables?

We have now clarified these aspects, and have used the same dataset for the statistical analyses throughout the paper.

For instance, it is not fully convincing that no grouping of pigment characteristics were found among sites, especially knowing the significant environment heterogeneity between thaw lakes and SRB reference.

We agree that some of the results were unexpected relative to the hypothesis that certain lake types (e.g. the palsa thaw lakes) would select for a unique subset of phytoplankton phyla with a few dominant taxa. In fact, most of the pigments were detected in all the lakes, suggesting that these environments are favourable for phylogenetically diverse taxa rather than the expected dominance. Additionally, our results point to the large lake-to-lake variations within each valley, even among nearby lakes. These novel findings are now given more attention in the revised manuscript.

What about the distribution of picophytoplankton?

We have now included ANOVA results that compare different picophytoplankton variables among valleys. Again the is highlighted the large lake-to-lake variation in each valley.

Also, is it common that the variation of environmental parameters and pigments composition between lakes of the same type is so big (see the thaw lakes on marine clays for example)?

Like the reviewer, we were surprised by this large variability within each valley, and we now discuss this unexpected result in the Discussion. In fact, this finding is consistent with new data on the bacterial communities that have shown that variability among ponds within the same valley can exceed differences among ponds from specific valleys (Comte et al. 2015, now cited).

I suggest to also re-analyse the molecular data exclusive of heterotrophic eukaryotes such as ciliates and fungi. Amplification biases should be addressed in more details.

We have rearranged the tables to place emphasis on the phytoplankton phyla, and to give much less attention to non-phototrophic groups. However, as noted above, the data for the heterotrophs are unique observations and extremely interesting, with relevance to grazer control of the picophytoplankton; for these reasons we are reluctant to completely remove this information that nicely completes the 18S rRNA records. The question of PCR biases is important and longstanding, and we have now included discussion of this in the revised Discussion.

3. I suggest the author to strength the discussions, in a more direct manner detailing the similarities and differences of phototropic community found between thaw lakes and reference lakes, and their contributions to the microbial community compared to heterotrophs. As written, it is currently difficult to recognize the key information of the results and evaluate the ecological significance phototrophic plankton have in the heterotrophic thaw lakes (e.g. in terms of lessen the emission of GHGs). It would be interesting to count and calculate the abundance and biomass ratios between heterotrophs and autotrophs in the thaw lakes, or even compare the ratio of picocyanobacteria to heterotrophic bacteria.

We are now revising the Discussion, placing greater emphasis on the similarities and differences in phototrophic communities between the ensemble of thaw lakes and the reference rock basin lakes. We have also modified the Introduction and Study site sections to indicate this comparison. We do not have a full set of data for heterotrophic bacteria to allow comparisons with picocyanobacteria.

Specific comments:

P. *123, L.6: Should be ". . ., while picoeukaryotes were inversely correlated with conductivity."*
Thank you for this correction, now made on the revised draft.

P. 125, L.10: Please add the information of sampling time and depths of each lake in Table 2.
This sampling time information is now included in the supplementary Table. The depth information is now placed in the legend for this table.

P. 130, L. 5-20: Please also mention the temperature differences among the lakes.
This point is now mentioned in the revised draft.

P. 131, L. 16: Please clarify the sampling year described in the manuscript and the Table title.
The sampling year is now inserted.

P. 131, L. 16-25: I found it very hard to follow the pigment results present in Table 2 and Figure 4, especially when there're 10 different pigments from 17 sampling sites at 4 different environments. I would suggest the authors to, 1). Unify the legends/terms for pigments in Table 2 and Fig. 4, and be consistent using them in the results and discussion section.
To address this concern we have removed Figure 4 from the manuscript and have explicitly defined the classes of pigments in Table 2.

2) If the special purpose of Table 2 is to compare the different contribution of photosynthetic and photoprotective pigments, please add a few columns in Table 2 to calculate the total percentage of each at different stations.
Thank you for this suggestion. We have addressed this by providing a new Table that gives total pigments in each category and the molar ratio between the two categories (new Table 3).

P. 133, L.11-13: This result seems too speculative. Also, it should be Figure S1.
This statement about the importance of photosynthetic sulphur bacteria has now been rephrased in the revised manuscript to avoid speculation. The figure reference has been now corrected to Figure S1, thank you.

P. 135, L. 20: Inconsistent information on the prevalence of diatom (see L. 21-22 of P. 140). Please clarify.
This has been edited for consistency.

P.136, L. 5-6: Please add a reference here.
The citation will be added here.

P. 136, L. 7-8: "The concentrations of β, β-carotene, were conspicuously high in the NAS lakes." This was only found during summer season of year 2012?
Yes, the NAS site was sampled only in 2012. We are modifying the graph to make this more obvious, and also the dates are specified in the supplementary table of sites.

P. 136, L. 26-29: How is this related to the occurrence of zeaxanthin? In any case, this information is useful but maybe fits somewhere else better?
This information relates to the pigment composition of cyanobacteria, and we are modifying this section for clarity.

P. 137, L. 27-30: The fraction/contribution of picoplankton to total phytoplankton community (especially in lake KWK and NAS), in terms of either pigments or biomass, should be also discussed.
The contribution of picophytoplankton to the total Chl-*a* will be included in the revised version of the manuscript (please see the calculation above).

P. 139, L. 10-14: Did the authors have a closer look at the dominating dinoflagellate species?
We are unable to address this question at this time.

P. 139, L. 24: Please add a reference here.
The appropriate citation will be added.