24 November 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, S. Crevecoeur, C. Lovejoy, J. Comte and W.F. Vincent

Thank you for your continuing help and for requesting that we now upload our manuscript. This has been fully revised in response to the two sets of referee comments, as per our response letter that we sent you previously and as further detailed and updated below.

Thank you for your ongoing consideration and editorial handling of this manuscript.

Sincerely,

Anna Przytulska
cc all authors

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 have been addressed in our revised manuscript as described below.

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 now included in the revised manuscript (p16).

*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 now consulted the literature as suggested and use an estimate of Chl *a* per picocyanobacterium cell (*Synechococcus*) as given in Moore et al. (1995). This allowed us to make a first order estimate (in the Discussion) of the contribution of picocyanobacteria to total phytoplankton, as requested. We note the large variation in cellular chlorophyll content as a function of strain and growth conditions (as shown in Moore et al. 1995) and conclude from this analysis that picocyanobacteria make a highly variable contribution among lakes to total phytoplankton biomass, from negligible to major (p. 21).

Moore, L.R., Goericke, R. and Chisholm, S. W., 1995. Comparative physiology of *Synechococcus* and *Prochlorococcus* - influence of light and temperature on growth, pigments, fluorescence and absorptive properties. *Marine Ecology Progress Series* 116; 259-275.

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 complements and strengthens 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) and Comte et al. (2015). We have now revised the text of the Discussion to connect to these studies (p. 24).

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 (p. 5). We did takes samples for nanophytoplankton enumeration by microscopy, but this separate large dataset is still under initial 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 (p. 16).

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, including new Table S1.

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 is known to be a poor descriptor of beta-diversity due to the presence of multiple zeros (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 (Replacement Figure 3 and p 15).

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 the suggestion by exploring using a db-RDA based 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. Here, cluster analyses, Permanova showed that no significant difference could be detected among valleys. Mantel tests further showed that no relationship between environmental conditions and pigment composition. We acknowledge however that 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 (replacement Figure 3).

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 micromoles as requested (Table 2) 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 (p. 13).

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?

The molar ratio of accessory pigments is now provided in revised Table 2 and further Discussed (p. 18).

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 picoplankton in 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 Answer: This is a new habitat type where bacteriochlorophyll *d* is described for the first time. We are revised the discussion to better integrate this information (p. 20).

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.

This figure has now been deleted.

Anonymous Referee #2

Anonymous Referee #2General 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 has been 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 (p. 20).

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

These terms are now defined in the revised Methods. Both are classes of accessory pigments, and the legend of Table 2 has been revised accordingly.

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 provided in new supplemental Table S1, with exact details.

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.

What about the distribution of picophytoplankton?

We do not have enough data and replicates for a full analysis of this, but note the large lake-to-lake variation (p. 20).

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 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, Crevecoeur et al. 2015. now cited, p. 18).

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 now mentioned on p. 23.

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 have 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. 3)

P. 125, L.10: Please add the information of sampling time and depths of each lake in Table 2. This sampling time and depth information is now included in the supplementary Table S1.

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 (Table 2).

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.

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.

This has been removed.

*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 (se new Table S1).

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 have modified 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.

As noted above, we have now consulted the literature as suggested and use an estimate of Chl *a* per picocyanobacterium cell (*Synechococcus*) as given in Moore et al. (1995). This allowed us to make a first order estimate (in the Discussion) of the contribution of picocyanobacteria to total phytoplankton, as requested. We note the large variation in cellular chlorophyll content as a function of strain and growth conditions (as shown in Moore et al. 1995) and conclude from this analysis that picocyanobacteria make a highly variable contribution among lakes to total phytoplankton biomass, from negligible to major (p. 21).

Moore, L.R., Goericke, R. and Chisholm, S. W., 1995. Comparative physiology of *Synechococcus* and *Prochlorococcus* - influence of light and temperature on growth, pigments, fluorescence and absorptive properties. *Marine Ecology Progress Series* 116; 259-275.

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 mention of low rRNA has now been deleted.4

Phototrophic pigment diversity and picophytoplankton in permafrost thaw lakes

A. Przytulska[1,2], J. Comte[1,2,3], S. Crevecoeur[1,2,3], C. Lovejoy[1,2,3], I. Laurion[4] and W. F. Vincent[1,2]

[1]Centre d'études nordiques (CEN) and Département de biologie, Université Laval, Québec City, QC G1V 0A6, CANADA

[2]Takuvik Joint International Laboratory, Centre National de la Recherche Scientifique (France, CNRS UMI 3376) and Département de Biologie, Université Laval, Québec City, QC G1V 0A6, CANADA.

[3]Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, Québec City, QC G1V 0A6, CANADA.

[4]Centre d'études nordiques (CEN) and Centre Eau Terre Environnement, Institut national de la recherche scientifique (INRS-ETE), Québec City, QC G1K 9A9, CANADA

Running title: Phototrophs in subarctic thaw lakes

Keywords: lakes; permafrost; phytoplankton; picocyanobacteria; picoeukaryotes; autotrophic picoplankton; pigments; protists; thermokarst

Biogeosciences special issue: 'Freshwater ecosystems in changing permafrost landscapes'.

Abstract

Permafrost thaw lakes (thermokarst lakes) are widely distributed across the northern landscape, and are known to be biogeochemically active sites that emit large amounts of carbon to the atmosphere as CH_4 and CO_2. However, the abundance and composition of the photosynthetic communities that consume CO_2 have been little explored in this ecosystem type. In order to identify the major groups of phototrophic organisms and their controlling variables, we sampled 12 permafrost thaw lakes along a permafrost degradation gradient in northern Québec, Canada. Additional samples were taken from 5 rock-basin reference lakes in the region to determine if the thaw waters differed in limnological properties and phototrophs. Phytoplankton community structure was determined by high performance liquid chromatography analysis of their photoprotective and photosynthetic pigments, and autotrophic picoplankton concentrations were assessed by flow cytometry. One of the black colored lakes located in a landscape of rapidly degrading palsas (permafrost mounds) was selected for high-throughput 18S rRNA sequencing complement conclusions based on the pigment and cytometry analyses. The results showed that the limnological properties of the thaw lakes differed significantly from the reference lakes, and were more highly stratified. However, both waterbody types contained similarly diverse phytoplankton groups, with dominance of the pigment assemblages by fucoxanthin-containing taxa, as well as chlorophytes, cryptophytes and cyanobacteria. Chlorophyll *a* concentrations (Chl *a*) were correlated with total phosphorus (TP), and both were significantly higher in the thaw lakes (overall means of 3.3 µg Chl *a* L^{-1} and 34 µg TP L^{-1}) relative to the reference lakes (2.0 µg Chl *a* L^{-1} and 8.2 µg TP L^{-1}). Stepwise multiple regression of Chl *a* against the other algal pigments showed that it was largely a function of alloxanthin, fucoxanthin and Chl *b* (R^2 = 0.85). The bottom waters of two of the thaw lakes also contained high concentrations of bacteriochlorophyll *d*, showing the presence of green photosynthetic

sulphur bacteria. The molecular analyses indicated a relatively minor contribution of diatoms, while chrysophytes, dinoflagellates and chlorophytes were well represented; the heterotrophic eukaryote fraction was dominated by numerous ciliate taxa, and also included Heliozoa, Rhizaria, chytrids and flagellates. Autotrophic picoplankton occurred in biovolume concentrations up to 3.1 × 10^5 μm^3 mL^{-1} (picocyanobacteria) and 1.9 × 10^6 μm^3 mL^{-1} (picoeukaryotes), and varied greatly among lakes. Both groups of picophytoplankton were positively correlated with total phytoplankton abundance, as measured by Chl *a*; picocyanobacteria were inversely correlated with dissolved organic carbon, while picoeukaryotes were inversely correlated with conductivity. Despite their net heterotrophic character, subarctic thaw lakes are rich habitats for diverse phototrophic communities.

1 Introduction

Degradation of ice-rich permafrost leads to the formation of thaw lakes, which are among the most abundant aquatic habitats in high latitude regions (Pienitz et al., 2008; Jones et al., 2012). These environments have attracted increasing scientific interest because of their biogeochemical reactivity. However, although there is rapidly increasing knowledge about their role in greenhouse gas (GHG) emissions (Laurion et al., 2010; Walter et al., 2006), little is known about their photosynthetic communities. Phototrophic organisms consume CO_2 and thereby reduce the net emission to the atmosphere; however, few studies have examined phytoplankton or other phototrophs in these abundant waters. Early studies in the U.S. Tundra Biome Program at Barrow, Alaska, recorded 105 species of algae in tundra lakes and ponds, with dominance of cryptophytes and chrysophytes (Alexander et al., 1980). More recent studies have focused on thaw lake diatoms as paleolimnological indicators, but the dominants in these records are often benthic taxa such as *Pinnularia* and *Fragilaria* (Bouchard et al., 2013). A lake survey in the western Hudson Bay lowlands, including in permafrost catchments, showed that the phytoplankton had diverse communities, primarily composed of cyanobacteria, chrysophytes, chlorophytes, cryptophytes, dinoflagellates and diatoms (Paterson et al., 2014).

Picophytoplankton (PP), consisting of picocyanobacteria and picoeukaryotes (nominally defined as cells 1 to 3 μm in diameter), contribute a major fraction of the total phototrophic biomass across a wide range of aquatic ecosystems (Richardson and Jackson, 2007), including northern lakes and rivers (Waleron et al., 2007; Vallières et al., 2008). In subarctic (Bergeron and Vincent, 1997) and high arctic (van Hove et al., 2008) lakes, picocyanobacteria may dominate the phytoplankton community in terms of biomass as well as cell abundance. For example, in

large oligotrophic Clear Water Lake (Lac à l'Eau Claire, Nunavik, Canada), small cell phytoplankton (cell fraction that passed through a 2 µm filter) accounted for 75% of the total phytoplankton Chl *a* (Bergeron and Vincent, 1997). However, the suitability of permafrost thaw lakes as a habitat for picophytoplankton has not been explored.

Our overall aim in the present study was to evaluate the major groups of phytoplankton in subarctic thaw lakes, and to relate this abundance and community structure to environmental variables. For this we employed phototrophic pigment analysis by high performance liquid chromatography (HPLC), an approach that has been applied with success to describe phytoplankton community structure at the phylum level in a wide range of freshwater (e.g., Fietz and Nicklisch 2004) and marine (e.g., Ansotegui et al., 2001) studies.

A secondary objective was to determine the abundance and distribution of picocyanobacteria and picoeukaryotes. As a further guide to the composition of the eukaryotic plankton, and in support of the pigment and picoeukaryote observations, we also applied high throughput 18S rRNA sequencing to surface and bottom waters from one selected lake that was strongly influenced by permafrost degradation. Our study included a wide range of small lakes across the gradient of permafrost degradation in Subarctic Quebec, Canada, from sporadic permafrost landscapes in the south (less than 10% of the area containing permafrost) to discontinuous permafrost in the north (10-90% permafrost). We also took comparative samples from a set of shallow rock-basin lakes that are unaffected by thermokarst processes. Given their limnological variability, as indicated by the variety of water colors among thaw lakes, we hypothesized that there would be large variations in total phytoplankton pigment concentration, pigment diversity and

100 picophytoplankton biovolume. Degrading permafrost soils release dissolved organic carbon
101 (DOC) and fine inorganic particles into the thaw lakes, and these constituents determine the
102 attenuation of light down the water column and the variability in color (Watanabe et al., 2011).
103 DOC also influences the near surface thermal and stratification regime (Caplanne and Laurion,
104 2008), and temperature is known to exert a direct effect on phytoplankton community structure,
105 particularly favouring cyanobacterial dominance (Paerl and Huisman, 2008). We therefore
106 hypothesised that DOC and temperature would be the primary drivers of variations in
107 phytoplankton pigmentation and picophytoplankton biovolume.
108

109 **2 Materials and Methods**

110 **2.1 Study Sites**

111 Twelve thaw lakes (small perennial waterbodies created by thermokarst erosion of the
112 permafrost) were sampled in subarctic Québec during the period of warm open-water conditions,
113 in late summer (August) 2011 and 2012 (Table S1). The lakes were distributed along a north-
114 south permafrost degradation gradient and across four geographically distinct locations: the
115 Sasapimakwananisikw River valley (SAS) and the Kwakwatanikapistikw River valley (KWK)
116 near Whapmagoostui-Kuujjuarapik; and the Sheldrake River valley (BGR) and the Nastapoka
117 River valley (NAS) near Umiujaq. The KWK and SAS valleys occur within the sporadic
118 permafrost landscape, while the BGR and NAS valleys are located in the discontinuous
119 permafrost landscape (Fig. 1). Each valley is characterised by distinct vegetation cover and soil
120 structure. Lakes located within the KWK valley are situated on impermeable clay-silt beds where
121 the drainage basin is covered with dense shrub vegetation (Breton et al., 2009), whereas lakes in
122 the SAS valley are located in peatlands in which permafrost mounds (palsas) are thawing and

degrading rapidly (Bhiry et al., 2011). The lakes located in the northern valleys (BGR, NAS) are situated on marine clay-silt beds and are surrounded by forest and shrub tundra. In addition to twelve permafrost thaw lakes, a set of five shallow rock-basin lakes (SRB) located on basalt bedrock was sampled in the vicinity of Whapmagoostui-Kuujjuarapik. These provided a set of reference lakes that are located at the same latitude and climatic setting, but without the direct influence of degrading permafrost that is experienced by the thaw lakes. The dates of sampling are given in Table S1.

2.2 Physicochemical analyses

Profiles of temperature, dissolved oxygen, conductivity, and pH of the 17 lakes were recorded with a 600R multiparametric probe (Yellow Springs Instrument Co.). Additionally temperature and conductivity were recorded with RBR XR620 conductivity-temperature-depth profiler (Richard Brancker Research Ltd). Near surface water samples (0.2 m depth) were collected into dark polyethylene bottles, previously washed with 10% hydrochloric acid and rinsed in MQ water. The samples were stored in coolers and transported to laboratory within 4 h of collection. The total nitrogen (TN) and total phosphorus (TP) measurements were performed on unfiltered water samples collected in 125ml bottles, acidified with sulfuric acid (0.2% final concentration), and stored at 4°C until persulfate digestion. TN concentrations were then measured with a Lachat flow injection analyzer and TP concentrations were measured using a Genesys 10UV spectrophotometer (Thermo Spectronic) and standard techniques (Stainton et al., 1977). Total suspended solids (TSS) were collected onto pre-combusted and pre-weighed glass fiber filters (Advantec MFS) that were dried for 2 h at 60°C and weighed on a Sartorius high precision balance. Dissolved organic carbon (DOC), colored dissolved organic matter (CDOM), soluble

146 reactive phosphorus (SRP) and nitrate (NO_3^-) measurements were performed on water filtered
147 through 0.2 µm cellulose acetate filters (Advantec MFS). Samples for DOC analyses were stored
148 in 45 mL dark glass bottles that had been previously burned at 450°C for 4 h and rinsed with MQ
149 water to remove any traces of organic substances. The DOC analysis was with a Shimadzu TOC-
150 5000A carbon analyzer calibrated with potassium biphthalate. CDOM was determined by
151 spectrophotometric absorbance of the filtrates at 320 nm, blanked against filtered MQ water and
152 converted to absorption values. SRP and NO_3^- were measured in the filtrates using standard
153 colorimetric methods (Stainton et al., 1977), and major ions were measured using Dionex ICS
154 2000 ion chromatograph.
155
156 **2.3 Pigment analysis**
157 Near surface (0.2 m depth) and near-bottom water samples (0.2 m above sediments; 50-500 mL)
158 from each lake were filtered onto 25-mm diameter GF/F glass-fibre filters, and immediately
159 frozen and stored at -80°C until pigment extraction in methanol. Pigments were analyzed by high
160 performance liquid chromatography (HPLC) following the protocols and standards described in
161 Bonilla et al. (2005). For some of the statistical analyses, two groups of algal accessory pigments
162 were separated as in Bonilla (2005): photoprotective pigments (canthaxanthin, diadinoxanthin,
163 echinenone, lutein, violaxanthin and zeaxanthin) and light harvesting, photosynthetic pigments
164 (alloxanthin, Chl *b*, fucoxanthin and peridinin). Standards for identification and quantification of
165 pigments (Chl *a*, Chl *b*, Chl *c2*, alloxanthin, β,β-carotene, canthaxanthin, crocoxanthin,
166 diadinoxanthin, echinenone, fucoxanthin, lutein, peridinin, violaxanthin, and zeaxanthin) were
167 obtained from Sigma Inc. (St. Louis, MO, USA) and DHI Water & Environment (Hørsholm,
168 Denmark) to calibrate the HPLC. The photodiode array spectrum of each peak was checked

against the reference spectra in Roy et al. (2011). No standards were available for bacteriochlorophyll *d* and the primary peaks for this pigment at 428 nm were expressed as Chl *a* equivalent concentrations.

2.4 Picophytoplankton enumeration

Near surface (0.2 m depth), unfiltered water samples from each lake were transferred to 5mL Cryovials, fixed with glutaraldehyde (10% final concentration) and stored at -80°C until analysis for picophytoplankton abundance. The cells were enumerated using a Becton Dickinson flow cytometer (BD FacsCalibur), equipped with an argon laser. Analyses were done at the lowest flow rate (12 µL min^{-1}), using a solution of 1-µm diameter, yellow-green microspheres (Polysciences, Inc) as an internal standard. Bead concentrations in the calibration solution were controlled using TrueCountAbsolute counting tubes (BD biosciences). Picocyanobacteria and picoeukaryotes were distinguished based on their chlorophyll and phycoerythrin fluorescence. Detection of the two groups was performed by the comparison of flow cytograms where cells were discriminated based on their side scatter signals (SSC) and both red (FL3) and orange fluorescence (FL2) as well as FL3 versus FL2. Given the low oxygen conditions observed in the bottom layers of the thaw lakes, samples were also analysed for green sulfur bacteria (FL3 vs SCC). The cytograms were analyzed using the Cell Quest Pro software, with manual gating to discriminate the different populations. For the picophytoplankton biovolume estimates, the diameters of 20 cells of each group in a sample from thaw lake KWK12 were measured under epifluorescence microscopy at 1000x magnification, and were then converted to spherical biovolumes. The measured cell diameters (±SD) were 1.0 ± 0.2 µm for picocyanobacteria and 2.0 ± 0.5 µm for picoeukaryotes, giving biovolumes per cell of 0.52 and 4.19 μm^3, respectively.

192 **2.5 RNA sampling and analysis**

193 Water samples from the near surface (0.2 m depth) and near-bottom (0.2 m above sediments) of
194 the black palsa lake SAS2A were first prefiltered through a 20 µm mesh to remove larger
195 organisms and then filtered sequentially through a 3 µm pore size, 47 mm diameter
196 polycarbonate filter (DHI) and a 0.2 µm Sterivex unit (Millipore) with a peristaltic pump. From
197 100 to 300 mL of water were filtered and the filtration was stopped after 2 hours to minimize
198 RNA degradation. The 3 µm filter for larger cells (L fraction) and the 0.2 µm filter for the
199 smaller fraction (S fraction) were both preserved in RNAlater (Life Technologies) and then
200 stored at -80°C until extraction.
201
202 Samples were extracted with the AllPrep DNA/RNA Mini Kit (Qiagen). This protocol was
203 modified by the addition of cross-linked polyvinylpyrrolidone (PVP, Alfa Aesar) (UV light
204 sterilized) to a final concentration of 10% before loading the samples onto the lysate
205 homogenization column. For all samples, the extracted RNA was converted to cDNA
206 immediately with the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems-
207 Ambion) and stored at -80°C until analysis. The V4 region of the eukaryotic 18S rRNA that had
208 been converted to cDNA was amplified using the 454 primers as described in Comeau et al.
209 (2011). PCR was carried out in a total volume of 50 µL, the mixture contained HF buffer 1X
210 (NEB), 0.25 µM of each primer, 200 µM of each dNTPs (Life Technology), 0.4 mg mL^{-1} BSA
211 (NEB), 1 U of Phusion High-Fidelity DNA polymerase (NEB) and 1 µL of template cDNA. Two
212 more reactions with 5X and 10X diluted template were also carried out for each sample, to
213 minimize potential primer bias. Thermal cycling began with an initial denaturation at 98°C for
214 30 s, followed by 30 cycles of denaturation at 98°C for 10 s, annealing at 55°C for 30 s,

extension at 72°C for 30 s and a final extension at 72°C for 270 s. The three dilution reactions were pooled and purified with a magnetic bead kit Agencourt AMPure XP (Beckman Coulter) and then quantified spectrophotometerically with the Nanodrop 1000 (Thermo Fisher Scientific). The amplicons were sequenced on 1/8 plates of the Roche 454 GS-FLX using the "PLUS" chemistry at the IBIS/Laval University, plate-forme d'analyses Génomiques (Québec City, QC). The raw 454 sequences have been deposited in the NCBI database under the bioproject name PRJNA286764.

Sequences were analysed using the UPARSE pipeline (Edgar, 2013). For quality filtering, the sequences were truncated at 245 bp to keep 50% of the reads at the 0.5 expected error rate. Singletons as well as chimeras were then removed and operational taxonomic units (OTUs) were determined at the ≥ 98% similarity level. These OTUs were classified using the mothur classifier (Schloss et al., 2009) with a 0.8 confidence threshold based on the SILVA reference database (Pruesse et al., 2007) modified to include sequences from our in-house, curated northern 18S rRNA gene sequence database. In order to compare samples, the OTU tables were each subsampled 100 times at 2200 reads, which corresponded to the lowest number of reads per sample minus 10%; this subsampling used the command multiple_rarefaction_even_depth.py in Qiime (Caporaso et al., 2010). The most abundant and unclassified OTUs were subsequently submitted to a BLASTn search to the nr database in NCBI GenBank (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to identify the nearest match.

2.6 Statistical analysis

The normal distribution of environmental variables was tested using the Kolmogorov-Smirnov test, and right-skewed variables were normalized by natural logarithm transformation. Given the order of magnitude differences in picophytoplankton abundances and pigment concentrations among samples, the HPLC and flow cytometry data were also normalized by logarithmic transformation. Correlations within and among the phytoplankton, pigment and environmental variables were tested by Pearson correlation analysis, with correction for multi-testing using the false discovery rate procedure as in Benjamini and Hochberg (1995). To investigate the extent to which environmental variables drove the distribution of pigment diversity among the different water bodies, a redundancy analysis (RDA, Legendre and Legendre, 2012) was run. This was based on Bray-Curtis distances for the pigment matrix (db-RDA) and the data were log-transformed prior to analysis. The significance of the model was assessed via 1000 permutations, and the analysis was performed in RStudio (version 0.98.501) using the Vegan package (Oksanen et al., 2015). Stepwise multiple linear regression models were performed using Past 3.04, with secondary cross-correlated variables removed prior to these analysis.

3 Results

3.1 Environmental heterogeneity

The thaw lakes spanned a wide range of environmental conditions, including water color and CDOM, with the latter strongly correlated with DOC ($R = 0.67$, $p < 0.0001$). The highest DOC concentrations (up to 17 mg L^{-1}) and CDOM (up to 117 m^{-1}) were recorded in the SAS lakes (Table 1). These waters were black in color and also had the lowest pH values (6.0 – 6.6). The highest total nutrient concentrations (up to 125 µg TP L^{-1} and 4 mg TN L^{-1}) were recorded

in lakes located within the KWK and NAS valleys, and the values were lowest in the shallow rock-basin waters (minima of 1.6 µg TP L^{-1} and 0.1 mg TN L^{-1}). Nitrogen to phosphorus ratios varied greatly among the 17 lakes, from 4 to 131 (g g^{-1}), and total suspended solids were similarly variable, from 1 to 320 mg L^{-1} (Table 1). The NAS valley waters contained especially high concentrations of suspended clay particles, producing an opaque milky appearance. Despite their shallowness and small size, the thaw lakes were highly stratified in terms of temperature and oxygen (Fig. 2), with anoxic bottom waters in the SAS and KWK lakes. Some had pronounced thermal gradients, with temperature differences up to 10°C between the surface and bottom waters. In contrast, the reference lakes showed more homogenous conditions, indicative of mixing (Fig. 2).

3.2 Planktonic pigments

Phytoplankton abundance, as measured by Chl *a* concentrations, varied greatly among the waterbodies (Table 1), from 0.4 (SRB1) to 6.8 (KWK6) µg L^{-1} in 2011 and from 0.2 (SRB1) to 9.1 (KWK1) µg L^{-1} in 2012. There was also a small but significant difference in Chl *a* concentrations between years, with means of 3.7 and 2.6 µg L^{-1}, respectively (paired t-test, $t = 2.5$, $p = 0.02$). On average, Chl *a* was significantly higher in the thaw lakes than the reference rock-basin waters: the overall means were 3.3 and 2.0 µg Chl *a* L^{-1}, respectively.

The pigment analyses of the phytoplankton (Table 2) showed that there were diverse communities including fucoxanthin-containing groups (potentially diatoms, chrysophytes and certain dinoflagellates), chlorophytes (Chl *b*, lutein and violaxanthin), cryptophytes (alloxanthin), dinoflagellates (peridinin) and cyanobacteria (zeaxanthin, canthaxanthin,

284 echinenone). The pigments Chl c_1, c_2, c_3 and crocoxanthin were also present, but generally at
285 trace concentrations, and only in certain lakes. The abundance of cyanobacterial populations in
286 KWK, BGR and NAS lakes was indicated by high concentrations of zeaxanthin (e.g., NASH)
287 and echinenone (SRB3). The KWK lakes had high concentrations of zeaxanthin (up to
288 0.7 nmol L^{-1} in KWK23 lake), accompanied by high concentrations of fucoxanthin and green
289 algal pigments (lutein and violaxanthin), as well as high concentrations of diadinoxanthin (e.g.,
290 KWK1). In the SAS lakes, a dominance of dinoflagellates was indicated by high concentrations
291 of peridinin. Echinenone was present in KWK and SRB lakes and high concentrations of
292 violaxanthin were also recorded in BGR lakes. Fucoxanthin-groups were abundant in SRB and
293 SAS as well as in NASH and BGR2. The turbid thaw lakes within the NAS valley had high
294 concentrations of β,β-carotene. Relatively high levels of ancillary photosynthetic pigments were
295 present in NASA and SAS lakes as well as in some waters of shallow rock-basin lakes (Table 2).
296 Photoprotective pigments were relatively more abundant in KWK lakes (notably KWK1 and
297 KWK6) as well as in the SRB waters (violaxanthin), and less abundant in the DOC-rich SAS
298 lakes (Table 2). The bottom waters of the thaw lakes also contained diverse planktonic pigments,
299 including high levels of diadinoxanthin and alloxanthin in KWK lakes, fucoxanthin in BGR2 and
300 Chl *b* in SRB lakes. High levels of bacteriochlorophyll *d* indicating abundant populations of
301 green photosynthetic sulfur bacteria were recorded in the deeper, anoxic waters of KWK lakes
302 (Table 3, Fig. 4).
303
304 For the overall data set, Chl *a* concentrations were significantly correlated with TP ($R = 0.47$;
305 $p = 0.05$), and with TSS ($R = 0.55$; $P = 0.03$), which were themselves strongly correlated
306 ($R = 0.76$; $p < 0.0001$). A forward stepwise linear regression showed that Chl *a* was best

307 described by a combination of the accessory pigments alloxanthin ($p < 0.014$), fucoxanthin
308 ($p < 0.001$) and Chl *b* ($p < 0.001$): ln Chl *a* = 1.774 + 0.161 ln Allo + 0.380 ln Chl *b* + 0.341 ln
309 Fuco ($R^2 = 0.85$; $p < 0.001$).
310
311 Several pigments were highly cross-correlated. These included alloxanthin and lutein ($R = 0.81$,
312 $p < 0.001$) and both pigments with Chl *b* ($R = 0.71$; 0.92, $p < 0.001$). The chlorophyte pigment
313 violoxanthin was also correlated with Chl *b* ($R = 0.60$, $p < 0.001$) and fucoxanthin ($R = 0.58$, $p <$
314 0.001). The fucoxanthin itself was most strongly correlated with diadinoxanthin ($R = 0.77$, $p <$
315 0.001). The cyanobacterial pigments echinenone and canthaxanthin were significantly correlated
316 ($R = 0.57$, $p < 0.001$), but not with zeaxanthin ($p > 0.05$). The summations within the two
317 categories of pigments, photoprotective and photosynthetic, were also positively correlated ($R =$
318 0.63; $p < 0.001$). Consistent with the multivariate analyses, the accessory pigments were
319 uncorrelated with individual environmental variables (all corrected p values were > 0.05), with
320 the exception of lutein. This chlorophyte pigment was significantly correlated with TP ($R = 0.53$;
321 $p = 0.01$), but this may simply reflect the strong correlation between lutein and Chl *a* ($R = 0.79$;
322 $p < 0.0001$), which itself correlated with TP.
323
324 The db-RDA model of environmental variables and pigment composition for surface water
325 samples resulted in a significant pattern, but also reaffirmed the large lake-to-lake variation in
326 each of the valleys, even among nearby lakes. The first two canonical components; related to TP,
327 DOC and pH; explained 11% of the total variances in pigment composition (Fig. 3). The dbRDA
328 model as a whole explained 16.4% of the variance ($R^2 = 0.16$, $F = 1.71$, $df = 8$, $p = 0.012$).
329

3.3 Picophytoplankton abundance and biovolume

Picophytoplankton concentrations varied greatly among the lakes (Fig. 5). The picocyanobacterial abundances ranged from 1.8 x 10^3 cells mL^{-1} (SAS1B) to 5.9 x 10^5 cells mL^{-1} (KWK23), equivalent to biovolume concentrations of 9.5 × 10^2 (SAS1B) to 3.1 × 10^5 μm^3 mL^{-1} (KWK23), while the picoeukaryote abundances ranged from 1.35 × 10^2 cells mL^{-1} (SAS2B) to 4.6 × 10^5 cells mL^{-1} (KWK1), equivalent to biovolume concentrations of 5.6 × 10^2 (SAS2B) to 1.9 × 10^6 μm^3 mL^{-1} (KWK1). In general, the lakes located on marine clays (KWK and BGR) contained the highest cell concentrations and biovolume of total picophytoplankton. The shallow rock-basin (SRB) and peatland lakes (SAS) were apparently less favourable, with picocyanobacterial and picoeukaryote biovolume concentrations below 10^4 μm^3 mL^{-1}. Picoeukaryotes were generally less numerically abundant than picocyanobacteria, but because of their larger cell size, they dominated total picophytoplankton biomass. Picoeukaryotes accounted on average (±SD) for 65 (±28) % of total picophytoplankton biovolume, however there was a wide range among lakes in this contribution, from 8% (SAS2B) to 99% (SAS1A).

Total picophytoplankton biovolume increased with Chl *a* concentration (R = 0.52; p = 0.03), but this relationship was only significant for the eukaryotic component (R = 0.53; p = 0.02). Picocyanobacteria correlated negatively with DOC (R = -0.47; p = 0.05), while picoeukaryotes correlated negatively with conductivity (R = -0.48; p = 0.05). Picocyanobacteria were highly correlated with zeaxanthin (R = 0.72; p = 0.0002), and there was also a significant, albeit less strong, correlation between picoeukaryotes and zeaxanthin (R = 0.54; p = 0.02). Stepwise multiple linear regression analysis showed that picophytoplankton (picoeukaryotes, PEuk; picocyanobacteria, PCyan) biovolumes were statistically related to certain limnological variables

353 according to the relationships: PEuk = 14.9 + 2.9 × Chl *a* - 1.7 × TN (R^2 = 0.56, p = 0.001), and
354 PCyan = -2.9 + 4.3 Temp + 1.1 × Chl *a* - 1.1 × TSS + 1.5 × TP - 1.2 × DOC (R^2 = 0.67,
355 p = 0.001).
356
357 **3.4 Molecular analyses**
358 The 18S rRNA data set from the palsa thaw lake (SAS2A) contained large numbers of rotifer
359 sequences (400 to 1350 reads per sample, all with closest matches to the genus *Ascomorpha*) and
360 these were removed prior to further analysis. This left a total of 3857 and 3128 reads for the
361 surface L (> 3.0 µm) and S (< 3.0 µm) fractions, and 3522 and 2457 reads for the bottom L and S
362 fractions; 84 to 93% of these eukaryotic sequences could be assigned (≥ 98% identity) to phylum
363 in the modified SILVA database. The largest fraction of total reads was attributable to ciliates
364 (up to 33% in the surface waters and 74% in the bottom waters; Table 4), including the genus
365 *Stokesia,* especially in the surface waters, and the genera *Cryptocaryon, Halteria*, *Peniculida* and
366 *Cyclidium*, especially in the bottom waters (Table 5). Among the groups nominally considered as
367 phytoplankton were dinoflagellates, chrysophytes and chlorophytes, with lesser proportions of
368 reads associated with katablepharids, bacillariophytes (diatoms) and cryptophytes (Table 4).
369 Analysis of the dissimilarity distances (Bray-Curtis distance on the sub-sampled dataset) showed
370 that showed that community structure greatly differed with depth (Bray-Curtis dissimilarity
371 index of 0.795 for the large fraction and 0.820 for the small fraction), and to a much lesser extent
372 between large and small fractions (Bray-Curtis dissimilarity index of 0.423 for the surface
373 samples and 0.312 for the bottom samples). Chlorophytes, dinoflagellates, katablepharids and
374 diatoms were more represented in the large, surface water fraction.
375

4 Discussion

Each of the subarctic thaw lakes contained pigments from several phytoplankton phyla, revealing that these abundant waters provide habitats for diverse phototrophic groups., The most abundant accessory pigment (apart from β,β-carotene present in all algal groups) was fucoxanthin, indicating the possible presence of diatoms, chrysophytes or certain dinoflagellates. Peridinin and alloxanthin were also present in many of the samples, indicating the presence of dinoflagellates and cryptophytes, respectively (Jeffrey et al., 2011). Diatoms would be less favoured in these stratified waters given their relatively high sinking rates, while flagellated taxa including chrysophytes, dinoflagellates and cryptophytes would be able to maintain their position in the euphotic zone. Mixotrophic chrysophytes and dinoflagellates have been observed in many high latitude lakes (Charvet et al., 2012; and references therein), and may be additionally favored by the high biomass concentrations of heterotrophic bacteria that occur in some of these waters (Breton et al., 2009; Roiha et al., 2015). Green algae were also well represented at most sites, indicating that despite the strong light attenuation by CDOM and TSS (Watanabe et al., 2011), there is adequate light availability for obligate phototrophs. Another conspicuous feature of the pigment data was the large variation in pigment characteristics among lakes within the same valley, even between adjacent waterbodies. This large within-valley variation has also been observed in bacterial studies in the region (Crevecoeur et al., 2015; Comte et al., 2015).

The concentrations of photoprotective pigments were conspicuously high in NASH relative to photosynthetic pigments (Table 2). This was unexpected given that it contained elevated concentrations of suspended solids, which indicate a low light availability for photosynthesis, and a lack of need for protection against bright light. It is possible, however, that in this lake,

cells suspended in the mixed layer are adapted to intermittent exposure to bright light rather than the average water column irradiance. Such conditions have been observed in a turbid estuarine environment, where the phytoplankton were photosynthetically adapted to high near-surface irradiances rather than the overall shade or dark conditions experienced on average by the cells as they were circulated by turbulent mixing through the water column (Vincent et al., 1994). In contrast, the ratio of photosynthetic to photoprotective pigments was high in NASA and the SAS lakes, indicating acclimation to low irradiances in their strongly light-attenuating water columns.

The pigment analyses also indicated the abundant presence of cyanobacteria. Echinenone and canthaxanthin are well known photoprotective pigments in cyanobacteria, with the latter especially prevalent in Nostocales, which may suggest the presence of nitrogen-fixing taxa. These results are consistent with bacterial 16S rRNA analyses, which showed the presence of cyanobacterial taxa in some of these lakes that had strong affinities (> 99% sequence similarity) to the Nostocalean taxon *Dolichospermum curvum* (Crevecoeur et al., 2015). Zeaxanthin can potentially occur in high cellular concentrations in cyanobacteria, but it also is found in eukaryotic algal groups. This pigment is a component of photoprotective xanthophyll cycles, and may co-occur with other components of these cycles. For example, studies on the diatom *Phaedactylum tricornutum* have shown the co-occurrence of the diadinoxanthin cycle and the violaxanthin cycle (Lohr and Wilhelm 1999). Consistent with this co-occurrence, we found a strong correlation between diadinoxathin and violaxanthin in the studied lakes ($R = 0.72$, $p < 0.001$). We also observed high concentrations of zeaxanthin, which is often associated with cyanobacteria but also chlorophytes (Jeffrey et al., 2011). Given the molecular analyses results of thaw lake bacterial communities (Crevecoeur et al., 2015) and our flow cytometry data,

zeaxanthin was likely to at least in part be associated with the abundant picocyanobacteria in the order Synechococcales. The strong correlation between picocyanobacteria and zeaxanthin further supports this relationship.

HPLC analysis has been used with success in a variety of aquatic ecosystems to not only identify major algal groups, but also to quantify their proportional representation using the software program CHEMTAX (Mackay et al., 1996). However, given the large known variation in pigment ratios in algal cells, this method requires extensive calibration on each class of waters. For example, in shallow a eutrophic lake, CHEMTAX gave a reliable estimation of cyanobacterial and chlorophyte biomass, but not chrysophytes and dinoflagellates (Tamm et al., 2015). The latter were two of the dominant groups in the permafrost thaw lakes, and although CHEMTAX offers a potentially useful approach for future analyses of these waters, further work will be required before it can be calibrated and reliably applied.

The presence of bacteriochlorophyll *d* in high concentrations in KWK lakes containing anoxic bottom waters indicate that these environments are favourable habitats for photosynthetic sulfur bacteria. These results were unexpected given the strong attenuation of light by the CDOM and suspended particles in these waters, and the low photosynthetically available radiation at depth. However, the results are consistent with molecular analyses of the bacterial assemblages. 16S rRNA gene clone library analysis of KWK lakes detected the presence of green sulfur bacteria (Rossi et al., 2013), and high throughput 16S rRNA sequencing revealed that the green sulfur bacterium *Pelodictyon* (*Chlorobi*) was one of the most abundant sequences in KWK waters (Crevecoeur et al., 2015). The high concentrations of bacteriochlorophyll *d* suggest that these

445 populations could play an important role in the overall primary production of certain thaw lakes,
446 although restricted to deeper water, anoxic conditions, and our observations extend the range of
447 environments in which this pigment has been detected.
448
449 Picophytoplankton occurred in all of the sampled lakes, but with large differences among waters
450 in terms of the relative abundance of prokaryotes versus eukaryotes, and probably also in terms
451 of their contribution to the total phytoplankton community biomass. As a first estimate of the
452 relative contribution of picocyanobacteria, their cell concentrations may be converted to
453 equivalent Chl *a* by an appropriate cell conversion factor. Analysis of *Synechococcus* in culture
454 under different irradiance regimes gave a median value around 7.5 fg Chl a per cell (Moore et al.
455 1995), and applying this value as a first order estimate, picocyanobacteria would contribute 0.3%
456 (SAS1B) to 80% (KWK23) of the measured total community Chl *a*. Such estimates are highly
457 approximate given the known variation in the cellular content of this pigment among strains and
458 with growth conditions; for example, by a factor of 4 as a function of irradiance (Moore et al.
459 1995). However these calculations imply large lake-to-lake variations in the percentage
460 contribution of picophytoplankton to the total community, ranging from negligible to major.
461
462 In general, the concentrations of both picocyanobacteria and picoeukaryotes increased with
463 increasing total phytoplankton biomass, as measured by Chl *a* concentrations. However, the two
464 groups differed in their correlative relationships with other limnological variables. In partial
465 support of our initial hypothesis that DOC would be a controlling variable, picocyanobacteria,
466 but not eukaryotes, were negatively correlated with DOC. An inverse relationship with DOC was
467 also found for picophytoplankton in Swedish lakes (Drakare et al., 2003). Similarly in Lake

468 Valkea-Kotinen, in the boreal zone of Finland, variations in autotrophic picoplankton were most
469 closely correlated with water column stability, which in turn was strongly regulated by DOC
470 concentration (Pelromaa and Ojala, 2012). High DOC waters are often characterized by low pH,
471 which may be a constraint on certain cyanobacteria, however acid-tolerant picocyanobacteria are
472 known (Jasser et al., 2013). Even in the low pH SAS waters picocyanobacteria were always
473 present, although in low concentrations (e.g., the minimum of 1.8×10^3 mL^{-1} in SAS2B in 2011).
474 Other factors such as zooplankton grazing may also have played a role in controlling
475 picocyanobacteria (Rautio and Vincent, 2006), although this seems less likely for
476 picocyanobacteria given that they are a nutritionally deficient food source for zooplankton in
477 thaw lakes (Przytulska et al., 2015).
478
479 Picocyanobacteria did not show the expected relationship with temperature in the correlation
480 analyses, although temperature was one of the variables retained in the multiple linear regression
481 analysis. Temperature has often been identified as a key variable for cyanobacterial growth and
482 dominance in lakes elsewhere. For example, in reservoirs in the southeastern USA, there was a
483 strong, positive correlation between picocyanobacterial cell concentrations and temperature,
484 while picoeukaryotes showed an inverse correlation, and dominance of the picophytoplankton
485 community shifted from picoeukaryotes in winter to picocyanobacteria in summer (Ochs and
486 Rhew, 1997). Similarly, increasing temperature favoured picocyanobacteria over picoeukaryotes
487 in German lakes (Hepperle and Krienitz, 2001). In experiments with subarctic lake and river
488 water at 10 and 20°C, the concentration of Chl *a* in the picoplankton fraction increased
489 substantially at the warmer temperature (Rae and Vincent, 1998). The temperature range in the
490 present study may have been too restricted to observe such effects.

The molecular data provided insight into the large diversity of microbial eukaryotes that occur within the plankton of thaw lake ecosystems, including heterotrophic components such as ciliates and flagellates that may exert grazing pressure on some of the phototrophs. When rotifers were excluded from the analyses, ciliates were dominant in the RNA sequences of SAS2A. This likely reflects not only their cellular abundance but also their large cell sizes with a concomitantly large number of ribosomes; for example, *Stokesia vernalis,* the most abundant sequence identified in the surface waters (Table 5), can be >100 µm in length. Ciliates are also known to be fragile cells that are easily broken up during manipulation like pre-filtration through a 20 µm mesh, which could account for their highly abundant sequences in the S as well as L fractions.

Chrysophytes and chlorophytes were well represented in the RNA sequences, particularly in the surface water L fraction, consistent with their abundance as indicated by the pigment data. Dinoflagellates constituted the dominant fraction of the phytoplankton sequences, yet were not detected as peridinin in SAS2A, although this pigment was present in two other SAS lakes. This may indicate the presence of large dinoflagellate cells, for example rigid Ceratium cells that can extend up to 100 µm in length, but may also be due to the presence of dinoflagellates that lack the accessory pigment peridinin. It should also be noted that the relative abundance of taxa in these analyses may additionally reflect PCR biases in amplification. Diatoms can also include large cell types, but their representation in the sequences was small, suggesting that most of the fucoxanthin that we measured was associated with chrysophytes such as *Uroglena* (Table 5) rather than diatoms. It is of interest that diatoms from the genus *Urosolenia* were the closest match following a BLAST search. This diatom is known to be lightly silicified and may be less

514 susceptible to sedimentation in these well stratified waters. The cryptophyte pigment alloxanthin
515 was in high concentration in SAS2A, as in the other thaw lakes, yet cryptophyte sequences
516 accounted for < 1.5% of the total RNA reads. This might reflect the small cell-size of certain
517 cryptophyte taxa, for example *Chroomonas*.
518
519 The molecular data also provided insight into the nature of the picoeukaryotic communities in
520 the SAS lakes. The taxonomic identities (Table 5) indicated the presence of several chlorophyte
521 genera that are known to produce small cells, notably *Choricystis*, *Lemmermannia*
522 *Monoraphidium* and *Chlorella*. For example, in subalpine Lake Tahoe (USA), *Choricystis*
523 *coccoides* produces cells that are only 0.5 μm^3 in volume, too small to be grazed by calanoid
524 copepods in that lake (Vincent, 1982). Among the chrysophytes, *Spumella* and related genera are
525 known to produce small cells. For all of these analyses, the many unidentified eukaryotic reads
526 add an extra element of uncertainty to the interpretation, but collectively these data underscore
527 the eukaryotic diversity of the thaw lake ecosystem. This parallels the large prokaryotic diversity
528 that has been observed in these lakes, including bacterial phototrophs (Crevecoeur et al. 2015;
529 Comte et al. 2015).
530
531 Permafrost thaw lakes receive large quantities of allochthonous organic carbon from their
532 surrounding catchments and this is reflected in their high DOC and CDOM concentrations, as
533 observed in the present study. These waters have high respiratory oxygen demands and are net
534 heterotrophic, resulting in prolonged hypoxia or anoxia in the bottom waters during summer, and
535 anoxia throughout the water column once ice covers the lake in winter (Deshpande et al., 2015).
536 The abundant cilate and nanoflagellate sequences in our molecular analyses also point to high

productivity by bacterial heterotrophs, their likely prey in these waters. However, despite these multiple signs of intense heterotrophy, the pigment, cytometry and molecular results in the present study show that these ecosystems are also the habitats for abundant phototrophs from diverse taxonomic groups.

Conclusions

The wide range of thaw lakes sampled in the present study significantly differed from the reference rock basin lakes in their limnological properties. On average, they contained higher phytoplankton (Chl *a*) and TP concentrations than the reference lakes, but had a comparable diversity of pigments, dominated by chlorophyte, chrysophyte and dinoflagellate pigments. Cyanobacteria and cryptophytes were also well represented, but the thaw waters appeared to be less favorable for diatoms, at least during the highly stratified late-summer period. Picophytoplankton occurred in all of the thaw lakes, in some of the waters at high biovolume up to 10^6 μm^3 mL^{-1}, but the proportion of eukaryotes and prokaryotes and their contribution to total phytoplankton biomass varied greatly among the lakes. Molecular analysis of samples from one of lake types indicated that small cell chlorophytes may be among the dominants in the picoeukaryotic fraction. Despite the heterotrophic nature of these organic-rich ecosystems, with respiration likely exceeding photosynthesis throughout the year, permafrost thaw lakes contain abundant, diverse phototrophs that potentially support higher trophic levels, and that will lessen the net CO_2 release from these waters to the atmosphere.

Author contributions

A. P., W. F. V. and I. L. designed the study; A. P. led the field sampling; S. C. with input from J. C. undertook the molecular analyses under the supervision of C. L.; laboratory analyses were overseen by A. P., W. F. V. and I. L.; flow cytometry analyses were by J. C. and A. P., under the supervision of I. L.; and data analysis was by A. P. with input from J. C., S. C. and W. F. V. A. P. prepared the manuscript with contributions from all co-authors.

Acknowledgments

We thank Marie-Josée Martineau for laboratory assistance with the HPLC analyses, and ADAPT colleagues Paschale Bégin, Bethany Deshpande and Alex Matveev, for assistance in the field. We are also grateful to Maciej Bartosiewicz for his help throughout this project and to Sylvia Bonilla and two anonymous referees for their insightful comments on the manuscript. This study was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) including the Discovery Frontiers ADAPT grant to W. F. V. and an EnviroNord fellowship to J. C., Fonds de recherche du Québec, the Network of Centres of Excellence ArcticNet and the Canada Research Chair program.

References

Alexander, V., Stanley, D. W., Daley, R. J. and McRoy, C. P.: Primary producers. In: Hobbie, J. E. (ed.) Limnology of tundra ponds, Barrow, Alaska, Dowden, Hutchinson and Ross Inc, Stroudsburg, PA. U.S./IBP synthesis series, 13, 179–250, 1980.

Ansotegui, A., Trigueros, J. M., and Orive, E.: The use of pigment signatures to assess phytoplankton assemblage structure in estuarine waters, Estuarine, Coastal and Shelf Science, 52, 689–703, 2001.

Benjamini, Y., and Hochberg, Y.: Controlling the false discovery rate: a practical and powerful approach to multiple testing, J. R. Stat. Soc., Series B, 57, 289–300, 1995.

Bergeron, M. and Vincent, W. F.: Microbial food web responses to phosphorus and solar UV radiation in a subarctic lake, Aquat. Microb. Ecol., 12, 239–249, 1997.

Bhiry, N., Delwaide, A., Allard, M., Bégin, Y., Filion, L., Lavoie, M., … and Vincent, W. F.: Environmental change in the Great Whale River region, Hudson Bay: Five decades of multidisciplinary research by Centre d'études nordiques (CEN), Ecoscience, 18, 182–203, doi:10.2980/18-3-3469, 2011.

Bonilla, S., Villeneuve, V., and Vincent, W. F.: Benthic and planktonic algal communities in a high arctic lake: pigment structure and contrasting responses to nutrient enrichment, J. Phycol., 41, 1120–1130, doi:10.1111/j.1529-8817.2005.00154.x, 2005.

Bouchard, F., Pienitz, R., Ortiz, JD., Francus, P., and Laurion, I.: Palaeolimnological conditions inferred from fossil diatom assemblages and derivative spectral properties of sediments in thermokarst ponds of subarctic Quebec, Canada, BOREAS 42, 575–595, doi:10.1111/bor.12000, 2013.

Breton, J., Vallières, C., and Laurion, I.: Limnological properties of permafrost thaw ponds in northeastern Canada, Can. J. Fish. Aquat. Sci., 66, 1635–1648, doi:10.1139/F09-108, 2009.

Caplanne, S., and Laurion, I.: Effects of chromophoric dissolved organic matter on epilimnetic stratification in lakes, Aquat. Sci., 70, 123–133, doi:10.1007/s00027-007-7006-0, 2008.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., and Fierer, N.: Qiime allows analysis of high-throughput community sequencing data, Nat. Methods, 7, 335–336, doi:10.1038/nmeth.f.303.QIIME, 2010.

Charvet, S., Vincent, W. F., and Lovejoy, C.: Chrysophytes and other protists in high Arctic lakes: molecular gene surveys, pigment signatures and microscopy, Polar Biol., 35, 733–748, doi:10.1007/s00300-011-1118-7, 2012.

Comeau, A. M., Li, W. K. W., Tremblay, J. É., Carmack, E. C., and Lovejoy, C.: Arctic Ocean microbial community structure before and after the 2007 record sea ice minimum, PloS ONE, 6, e27492, doi:10.1371/journal.pone.0027492, 2011.

609 Comte, J., Monier, A., Crevecoeur, S. M., Lovejoy, C., and Vincent, W. F.: Bacterial
610 biogeography of permafrost thaw lakes in the changing northern landscape, Ecography, 38,
611 doi: 10.1111/ecog.01667, 2015.

612 Crevecoeur, S., Vincent, W. F., Comte, J., and Lovejoy, C.: Bacterial community structure
613 across environmental gradients in permafrost thaw ponds: Methanotroph-rich ecosystems,
614 Front. Microbiol., 6, 192, doi:10.3389/fmicb.2015.00192, 2015.

615 Deshpande, B. N., MacIntyre S., Matveev A., and Vincent, W. F.: Oxygen dynamics in
616 permafrost thaw lakes: Anaerobic bioreactors in the Canadian subarctic. Limnol. Oceanogr.
617 60: doi:10.1002/lno.10126, 2015.

618 Drakare, S., Blomqvist, P., Bergström, A. K., and Jansson, M.: Relationships between
619 picophytoplankton and environmental variables in lakes along a gradient of water colour and
620 nutrient content, Freshwater Biol., 48, 729–740, doi:10.1046/j.1365-2427.2003.01049.x,
621 2003.

622 Edgar, R. C.: UPARSE: highly accurate OTU sequences from microbial amplicon reads, Nat.
623 Methods,10, 996–998, doi:10.1038/nmeth.2604, 2013.

624 Fietz, S., and Nicklisch, A.: An HPLC analysis of the summer phytoplankton assemblage in
625 Lake Baikal, Freshwater Biol., 49, 332–345, doi: 10.1111/j.1365-2427.2004.01183.x, 2004.

626 Hepperle, D., and Krienitz, L.: Systematics and ecology of chlorophyte picoplankton in German
627 inland waters along a nutrient gradient, Int. Rev. Hydrobiol., 86, 269–284, 2001.

628 Jasser, I., Karnkowska-Ishikawa, A., and Chróst, R. J.: Do acid-tolerant picocyanobacteria exist?
629 A study of two strains isolated from humic lakes in Poland, Hydrobiologia, 707, 209–218,
630 doi:10.1007/s10750-012-1428-y, 2013.

631 Jeffrey, S. W., Wright, S. W., and Zapata, M.: Microalgal classes and their signature pigments.
632 In: Roy, S., Llewellyn, C. A., Skarstad Egeland, E., and Johnsen, G. (eds) Phytoplankton
633 pigments: characterization, chemotaxonomy and aplications in oceanography, Cambridge
634 University Press. (SCOR), 52–53, 2011.

635 Jones, M., Grosse, G., Jones, B. M., and Walter Anthony, K. M.: Peat accumulation in a
636 thermokarst-affected landscape in continuous ice-rich permafrost, Seward Peninsula, Alaska,
637 J. Geophys. Res., 117, G00M07, doi:10.1029/2011JG001766, 2012.

638 Mackey, M. D., Mackey, D. J., Higgins, H. W., and Wright, S. W.: CHEMTAX—a program for
639 estimating class abundances from chemical markers: application to HPLC measurements of
640 phytoplankton, Mar. Ecol.-Prog. Ser., 144, 265–83, doi:10.3354/meps144265, 1996.

641 Moore, L. R., Goericke, R. and Chisholm, S. W.: Comparative physiology of *Synechococcus* and
642 *Prochlorococcus* - influence of light and temperature on growth, pigments, fluorescence and
643 absorptive properties, Marine Ecology Progress Series, 116; 259–275, 1995.

644 Laurion, I., Vincent, W. F., MacIntyre, S., Retamal, L., Dupont, C., Francus, P., and Pienitz, R.:
645 Variability in greenhouse gas emissions from permafrost thaw ponds, Limnol. Oceanogr., 55,
646 115–133, doi:10.4319/lo.2010.55.1.0115, 2010.

647 Legendre, P. and Legendre, L.: Numerical Ecology, 3rd English Edition, Elsevier, 2012.

648 Lohr, M., and Wilhelm, C.: Algae displaying the diadinoxanthin cycle also possess the
649 violaxanthin cycle, Proc. Natl. Acad. Sci. USA, 96, 8784–8789, 1999.

650 Ochs, C. A., and Rhew, K.: Population dynamics of autotrophic picoplankton in a southeastern
651 US reservoir, Int. Rev. Hydrobiol. Hydrogr., 82, 293–313, 1997.

652 Oksanen, J., Blanchet, G.F., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson,
653 G.L., Solymos, P., Henry, M., Stevens, H. and Wagner H.: Vegan: Community Ecology
654 Package, R package version 2.2-1, http://CRAN.R-project.org/package=vegan, 2015.

655 Paterson, A. M., Keller, W., Rühland K. M., Jones C. F., and Winter J. G.: An exploratory
656 survey of summer water chemistry and plankton communities in lakes near the Sutton river,
657 Hudson Bay lowlands, Ontario, Canada, Arct. Antarct. Alp. Res., 46, 121–138,
658 doi:10.1657/1938-4246-46.1.121, 2014.

659 Peltomaa, E., and Ojala, A.: Meteorological drivers of the dynamics of autotrophic picoplankton,
660 Freshwater Biol., 57, 1005–1016, doi:10.1111/j.1365-2427.2012.02761.x, 2012.

661 Paerl, H. W., and Huisman, J.: Blooms like it hot, Science, 320, 57–58,
662 doi:10.1126/science.1155398, 2008.

663 Pienitz, R., Doran, P. T., and Lamoureux S. F.: Origin and geomorphology of lakes in the polar
664 regions. Polar lakes and rivers, limnology of Arctic and Antarctic aquatic ecosystems, Oxford
665 University Press, Oxford, UK 25–41, 2008.

666 Pruesse, E., Quast, C., Knittel, K., Fuchs, B. M., Ludwig, W., Peplies, J., … and Glockner, F. O.:
667 SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA
668 sequence data compatible with ARB, Nucleic Acids Res., 35, 7188–7196,
669 doi:10.1093/nar/gkm864, 2007.

670 Przytulska, A., Bartosiewicz, M., Rautio, M., Dufresne, F., and Vincent, W. F.: Climate effects
671 on high latitude *Daphnia* via food quality and thresholds, PLoS ONE, 10, e0126231,
672 doi:10.1371/journal.pone.0126231, 2015.

673 Rae, R. and Vincent, W. F.: Effects of temperature and ultraviolet radiation on microbial
674 foodweb structure: potential responses to global change, Freshwater Biol., 40, 747–758,
675 doi:10.1046/j.1365-2427.1998.00361.x, 1998.

676 Rautio, M., and Vincent, W. F.: Benthic and pelagic food resources for zooplankton in shallow
677 high latitude lakes and ponds, Freshwater Biol., 51, 1038–1052, doi:10.1111/j.1365-
678 2427.2006.01550.x, 2006.

679 R Core Team R: A language and environment for statistical computing, R Foundation for
680 Statistical Computing, Vienna, Austria, http://www.R-project.org/, 2014.

681 Richardson, T. L., and Jackson, G. A.: Small phytoplankton and carbon export from the surface
682 ocean, Science, 315(5813), 838–840, doi:10.1126/science.1133471, 2007.

683 Roiha, T., Rautio, M. and Laurion I.: Carbon dynamics in highly net heterotrophic subarctic thaw
684 ponds, Biogeosciences, in press, 2015

685 Rossi P., Laurion I., and Lovejoy C.: Distribution and identity of bacteria in subarctic permafrost
686 thaw ponds, Aquat. Microb. Ecol., 69, 231–245, doi:10.3354/ame01634, 2013.

687 Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., and Hartmann, M.: Introducing mothur:
688 Open-source, platform-independent, community-supported software for describing and
689 comparing microbial communities. Appl. Environ. Microbiol., 75, 7537–7541,
690 doi:10.1128/AEM.01541-09, 2009.

691 Roy, S., Llewellyn, C. A., Egeland, E. S., and Johnsen, G. (Eds.): Phytoplankton pigments:
692 characterization, chemotaxonomy and applications in oceanography, Cambridge University
693 Press., United Kingdom, 845 pp, 2011.

694 Stainton, M., Capel, M., and Armstrong, F.: The chemical analysis of fresh water, 2nd ed.
695 Special Publication 25 Canada Fisheries and Marine Service, 1977.

696 Tamm, M., Freiberg, R., Tõnno, I., Nõges, P., and Nõges ,T.: Pigment-based chemotaxonomy - a
697 quick alternative to determine algal assemblages in large shallow eutrophic lake?, PLoS ONE,
698 10, e0122526, doi:10.1371/journal.pone.0122526, 2015.

699 Vallières, C., Retamal, L., Ramlal, P., Osburn, C. L., and Vincent, W. F.: Bacterial production
700 and microbial food web structure in a large arctic river and the coastal Arctic Ocean,
701 J. Marine Syst., 74, 756–773, doi:10.1016/j.jmarsys.2007.12.002, 2008.

702 Van Hove, P., Vincent, W. F., Galand, P. E. and Wilmotte, A.: Abundance and diversity of
703 picocyanobacteria in high arctic lakes and fjords, Algological Studies, 126, 209–227,
704 doi:10.1127/1864-1318/2008/0126-0209, 2008.

705 Vincent, W. F.: The autecology of an ultraplanktonic shade alga in Lake Tahoe, J. Phycol., 18,
706 226–232, 1982.

707 Vincent, W.F., Bertrand, N., and Frenette, J-J.: Photoadaptation to intermittent light across the
708 St. Lawrence Estuary freshwater-saltwater transition zone, Mar. Ecol.-Prog. Ser., 110, 283–
709 292, 1994.

710 Waleron, M., Waleron, K., Vincent, W. F., and Wilmotte, A.: Allochthonous inputs of riverine
711 picocyanobacteria to coastal waters in the Arctic Ocean, FEMS Microbial. Ecol., 59, 356–
712 365, doi:10.1111/j.1574-6941.2006.00236.x, 2007.

713 Walter, K. M., Zimov, S.A., Chanton, J. P., Verbyla, D., and Chapin, F. S.: III Methane bubbling
714 from Siberian thaw lakes as a positive feedback to climate warming, Nature, 443, 71–75,
715 doi:10.1038/nature05040, 2006.

716 Watanabe, S., Laurion, I., Chokmani, K., Pienitz, R., and Vincent, W. F.: Optical diversity of
717 thaw ponds in discontinuous permafrost: A model system for water color analysis,
718 J. Geophys. Res.-Biogeo., 116, doi:10.1029/2010JG001380, 2011.

719

Table 1. Limnological characteristics including surface values for temperature (T), pH, dissolved organic carbon concentration (DOC), colored dissolved organic matter (CDOM), total suspended solids (TSS), soluble reactive phosphorus (SRP), total phosphorus (TP), total nitrogen (TN), nitrate (NO_3) and Chlorophyll *a* (Chl *a*) in studied subarctic lakes. Mean values from 2011 and 2012 (+/- range in brackets; nd = no data from 2011).

Sites	T (°C)	pH	DOC (mg L^{-1})	CDOM (m^{-1})	TSS (mg L^{-1})	SRP (µg L^{-1})	TP (µg L^{-1})	TN (mg L^{-1})	NO_3 (mg N L^{-1})	Chl *a* (µg L^{-1})
Thaw lakes on marine clays										
BGR1	15.1 (0.7)	7.5 (0.2)	3.9 (0.4)	6.1 (1.7)	2.9 (0.5)	1.8 (0.6)	17.3 (3.5)	0.1 (0.0)	0.05 (0.0)	1.8 (1.0)
BGR2	14.5 (0.4)	7.0 (0.3)	9.0 (0.3)	39.1 (5.7)	19.2 (6.1)	2.4 (1.0)	45.7 (0.2)	0.3 (0.0)	0.05 (0.0)	3.4 (1.4)
NASA	15.6 (nd)	7.0 (nd)	3.0 (nd)	12.3 (nd)	319.3 (nd)	2.9 (nd)	124.5 (nd)	3.7 (nd)	0.25 (nd)	4.1 (nd)
NASH	18.3 (nd)	7.6 (nd)	4.1 (nd)	22.6 (nd)	18.2 (nd)	6.2 (nd)	28.5 (nd)	0.4 (nd)	0.04 (nd)	1.7 (nd)
Thaw lakes on mineral clays										
KWK1	17.1 (4.4)	6.1 (0.8)	9.2 (2.8)	39.0 (16.4)	14.1 (12.0)	2.2 (1.5)	36.8 (26.7)	0.3 (0.1)	0.05 (0.0)	8.0 (1.2)
KWK6	14.9 (0.8)	6.8 (0.4)	5.2 (0.0)	10.7 (0.7)	9.3 (1.1)	1.0 (0.4)	30.9 (3.0)	0.2 (0.0)	0.06 (0.0)	4.4 (1.6)
KWK12	16.8 (0.8)	8.0 (1.1)	8.6 (0.7)	39.6 (3.7)	13.8 (2.6)	1.6 (0.0)	27.7 (2.1)	0.2 (0.0)	0.08 (0.0)	2.5 (0.1)
KWK23	14.9 (0.2)	6.7 (0.2)	7.2 (0.6)	35.8 (1.9)	10.2 (1.9)	4.9 (0.6)	47.7 (5.6)	0.2 (0.0)	0.04 (0.0)	3.5 (1.9)
Thaw lakes on peatlands										
SAS1A	14.3 (0.2)	6.6 (0.3)	10.7 (0.8)	68.7 (2.7)	7.6 (2.6)	1.6 (0.3)	14.3 (0.9)	0.5 (0.1)	0.17 (0.0)	3.6 (1.2)
SAS1B	13.6 (0.1)	6.3 (0.3)	15.9 (0.4)	109.5 (3.9)	21.8 (5.4)	1.9 (0.4)	12.7 (2.2)	0.6 (0.1)	0.07 (0.0)	4.3 (0.1)
SAS2A	19.9 (nd)	6.2 (nd)	14.9 (nd)	98.4 (nd)	2.6 (nd)	3.1 (nd)	9.6 (nd)	0.7 (nd)	0.04 (nd)	1.2 (nd)
SAS2B	16.0 (nd)	6.0 (nd)	17.1 (nd)	116.9 (nd)	5.2 (nd)	1.3 (nd)	10.3 (nd)	0.5 (nd)	0.11 (nd)	0.9 (nd)
Shallow rocky basins										
SRB1	15.8 (2.2)	7.6 (0.6)	9.9 (0.0)	46.0 (6.5)	2.6 (1.2)	2.1 (0.6)	7.9 (2.8)	0.2 (0.1)	0.06 (0.0)	0.3 (0.1)
SRB2	13.8 (0.5)	7.6 (1.1)	13.2 (2.7)	68.8 (21.5)	1.3 (0.3)	1.4 (0.5)	11.2 (2.6)	0.2 (0.1)	0.14 (0.1)	1.4 (0.4)
SRB3	15.6 (0.2)	6.6 (0.2)	7.8 (1.4)	34.5 (9.5)	5.4 (2.0)	1.1 (0.1)	13.4 (2.8)	0.3 (0.1)	0.05 (0.0)	5.2 (0.6)
SRB4	15.0 (0.3)	7.9 (0.5)	10.4 (1.6)	20.0 (0.5)	5.0 (1.9)	0.8 (0.1)	5.7 (2.6)	0.8 (0.4)	0.44 (0.4)	2.4 (1.0)
SRB5	18.7 (1.8)	7.1 (0.9)	3.7 (0.1)	9.4 (2.7)	0.7 (0.1)	0.5 (0.2)	2.9 (1.3)	0.1 (0.0)	0.32 (0.2)	0.8 (0.0)

Table 2. The dominant photosynthetic and photoprotective accessory pigments (nmol L^{-1}), their sum (Σ) and ratio in the subarctic water bodies sampled in 2012.

Sites	Photosynthetic					Photoprotective							Ratio
	Allo	Chl b	Fuco	Perid	Σ	Cantha	Diadino	Echin	Lut	Viola	Zea	Σ	
Thaw lakes on marine clays													
BGR1	0.282	0.108	0.174	0.034	0.598	0.000	0.113	0.045	0.126	0.132	0.080	0.497	1.2
BGR2	0.053	0.159	0.611	0.343	1.167	0.073	0.483	0.088	0.372	0.442	0.677	2.134	0.5
NASA	2.078	0.401	0.101	0.000	2.580	0.000	0.000	0.000	0.745	0.308	0.000	1.053	2.4
NASH	0.161	0.204	0.740	0.000	1.106	0.184	0.464	0.208	0.208	0.568	1.692	3.324	0.3
Thaw lakes on mineral clays													
KWK1	0.641	1.681	0.702	0.458	3.481	0.147	3.836	0.202	1.435	0.847	0.375	6.843	0.5
KWK6	0.248	0.760	0.429	0.075	1.513	0.093	0.398	0.000	1.335	0.771	0.561	3.158	0.5
KWK12	0.290	0.202	0.569	0.151	1.213	0.046	0.358	0.089	0.337	0.354	0.111	1.295	0.9
KWK23	0.281	0.236	0.499	0.057	1.073	0.046	0.443	0.026	0.644	0.450	0.688	2.296	0.5
Thaw lakes on peatlands													
SAS1A	0.498	0.130	1.080	0.110	1.817	0.083	0.291	0.000	0.161	0.366	0.092	0.993	1.8
SAS1B	1.066	0.271	1.484	0.363	3.184	0.044	0.435	0.000	0.294	0.420	0.083	1.275	2.5
SAS2A	0.342	0.054	0.199	0.000	0.594	0.000	0.047	0.000	0.047	0.000	0.000	0.093	6.4
SAS2B	0.606	0.163	0.502	0.000	1.271	0.036	0.049	0.045	0.220	0.171	0.000	0.521	2.4
Shallow rock-basin lakes													
SRB1	0.018	0.038	0.051	0.017	0.124	0.000	0.026	0.008	0.045	0.069	0.026	0.175	0.7
SRB2	0.196	0.220	0.360	0.040	0.816	0.037	0.078	0.057	0.229	0.183	0.062	0.646	1.3
SRB3	1.026	0.352	2.011	0.261	3.650	0.103	0.460	0.219	0.557	0.924	0.530	2.792	1.3
SRB4	0.094	0.203	0.557	0.050	0.905	0.016	0.159	0.080	0.333	0.340	0.101	1.028	0.9
SRB5	0.066	0.048	0.420	0.004	0.537	0.010	0.085	0.031	0.064	0.199	0.087	0.477	1.1

Key: Allo, alloxanthin; Chl *b*, chlorophyll *b*; Fuco, fucoxathin; Perid, peridinin; Cantha, canthaxanthin; Diadino, diadinoxanthin; Echin, echinenone; Lut, lutein; Viola, violaxanthin; Zea, zeaxanthin.

Table 3. The relative concentration of bacteriochlorophyll *d* (BChl *d*, μg L^{-1}) based on the maximum peak area at 430 nm. The lakes have been arranged from lowest to highest concentrations.

Site	Date	Depth (m)	BChl *d* (μg L^{-1})
KWK6	4 Aug 2012	3.1	1.2
KWK6	21 Aug 2011	3.1	1.3
KWK23	4 Aug 2012	2.0	1.9
KWK12	3 Aug 2012	2.0	8.2
KWK12	19 Aug 2011	2.5	24.6
KWK1	19 Aug 2011	2.0	28.9
KWK23	4 Aug 2012	3.3	36.2
KWK23	21 Aug 2011	3.3	44.3
KWK1	3 Aug 2012	2.0	44.7
KWK12	3 Aug 2012	2.5	47.3

Table 4. RNA sequence analysis of eukaryotes in samples from permafrost thaw lake SAS2A, sampled in 2012. Each value is the % number of reads of the total for each sample (total number of reads minus rotifer sequences). The large fraction was retained on a 3 μm filter and the small fraction was on a 0.2 μm filter after filtration through the 3 μm pre-filter.

Taxonomic group	Percentage of reads			
	Surface		Bottom	
	Large	Small	Large	Small
Phytoplankton groups				
Dinophyta	17.11	4.02	8.71	1.90
Chrysophyta[a]	14.47	14.81	9.07	3.60
Chlorophyta[b]	9.97	2.03	2.11	0.40
Cryptophyta	3.10	1.71	2.39	0.65
Katablepharidophyta	2.71	1.00	0.11	0.20
Bacillariophyta	2.47	0.44	0.96	0.45
Raphidophyceae	0.51	0.33	0.00	0.04
Pavlovales	0.44	0.19	0.18	0.04
Prymnesiales	0.15	0.03	0.06	0.00
Other groups				
Ciliophora	22.94	43.55	62.95	83.90
Cercozoa	6.17	12.37	2.50	1.22
Fungi[c]	1.47	0.61	0.32	0.04
Centroheliozoa	1.44	0.80	0.11	0.12
Choanoflagellida	1.32	3.12	0.23	0.20
Perkinsea	0.00	0.00	0.03	0.12
Unknown affinities	15.75	15.00	10.28	7.11

[a]includes Chrysophyceae, Synurophyceae and Bicosoecida
[b]includes Chlorophyceae and Trebouxiophyceae
[c]includes Chytridiomycota, Oomycota and Ascomycota

Table 5. Closest identity (ID) of eukaryotic RNA sequences from permafrost thaw lake SAS2A to GenBank sequences (following a BLASTn search), at the lowest taxonomic level identified.

GenBank Taxonomy	Accession number	Isolation source	% ID	Percentage of reads	
				Surface	Bottom
Phytoplankton groups					
Chrysophyta					
Uroglena sp.	EU024983	FU44-26	99	6.03	0.03
Paraphysomonas sp.	JQ967316	Freshwater	99	0.67	3.78
Dinobryon divergens	KJ579346	WO33_4	99	0.40	0.00
Spumella-like flagellate	AY651098	Lake Mondsee	99	0.03	0.76
Cryptophyta					
Cryptomonas tetrapyrenoidosa	KF907407	Deokam032610	99	1.71	0.87
Cryptomonas pyrenoidifera	KF907397	CNUCRY 166	99	0.33	0.00
Cryptomonas curvata	KF907377	CNUCRY 90	99	0.06	0.28
Dinophyta					
Dinophyceae sp.	GQ423577	Lake Baikal	99	1.31	0.58
Peridinium wierzejskii	KF446619	Baikal region	99	1.03	2.07
Gyrodiniellum shiwhaense	FR720082	Shiwha Bay	98	0.49	0.01
Bacillariophyta					
Urosolenia eriensis	HQ912577	Y98-8	98	1.22	0.58
Chlorophyta					
Lemmermannia punctata	JQ356704	SAG 25.81	99	1.07	0.08
Chlorella sp.	Y12816	OvS/Ger1	99	1.00	0.00
Choricystis sp.	AY195972	AS-29	99	0.89	0.11
Koliella longiseta	HE610126	SAG 470-1	99	0.72	0.04
Monoraphidium sp.	KP017571	LB59	99	0.39	0.00
Raphidophyta					
Gonyostomum semen	KP200894	Freshwater	100	0.23	0.02
Prymnesiales					
Chrysochromulina parva	EU024987	FU44-40	100	0.09	0.03
Other groups					
Ciliophora					
Stokesia vernalis	HM030738	Freshwater	99	8.65	0.04
Cryptocaryon sp.	JF317699	Drinking water	99	0.89	5.61
Peniculida sp.	GQ330632	Peat bog water	98	0.83	1.85
Halteria sp.	GU067995	Lake water	99	0.49	6.38
Cyclidium marinum	JQ956553	Marine coast	99	0.00	34.42
Rhizaria					
Cercozoa	AB771834	Lake Kusaki	99	3.83	0.28
Fungi					
Saprolegnia sp.	FJ794911	Lake (parasite)	99	0.11	0.00
Penicillium brevicompactum	KP981369	ATCC 16024	99	0.00	0.10

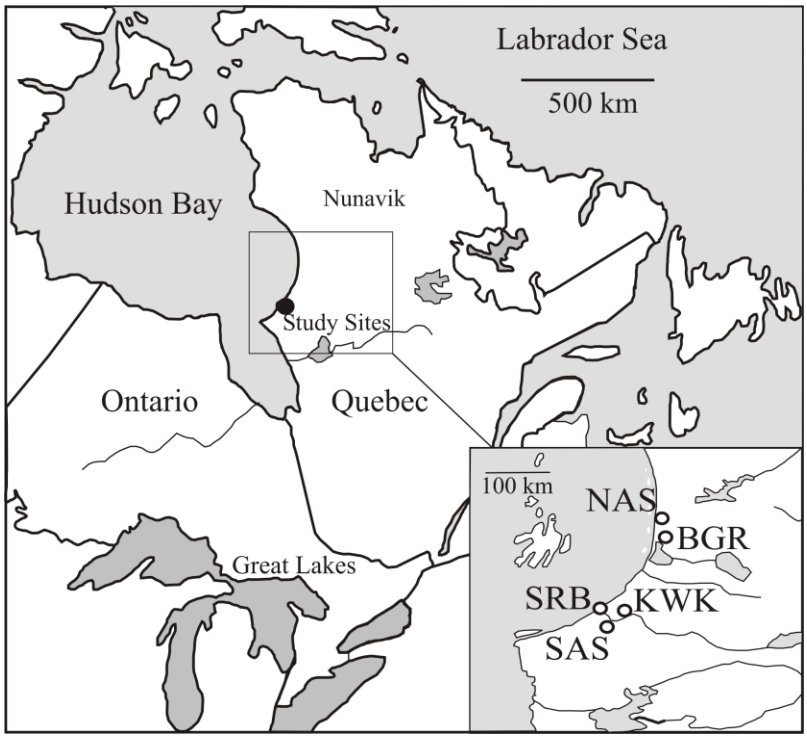

Figure 1. The location of the study area in Subarctic Quebec.

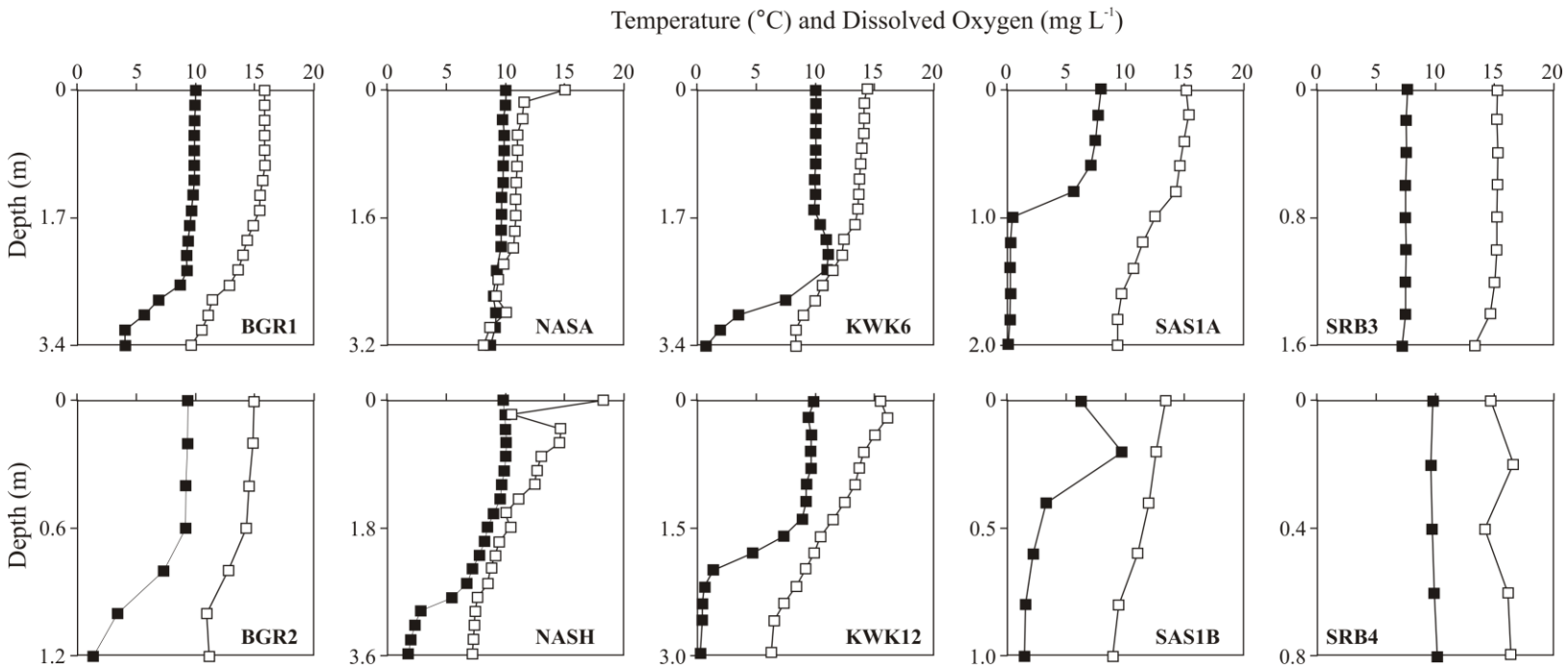

Figure 2. Temperature (white squares) and oxygen (black squares) stratification in permafrost thaw lakes (BGR, KWK, NAS, SAS) and shallow rock-basin lakes (SRB) during the summer 2012.

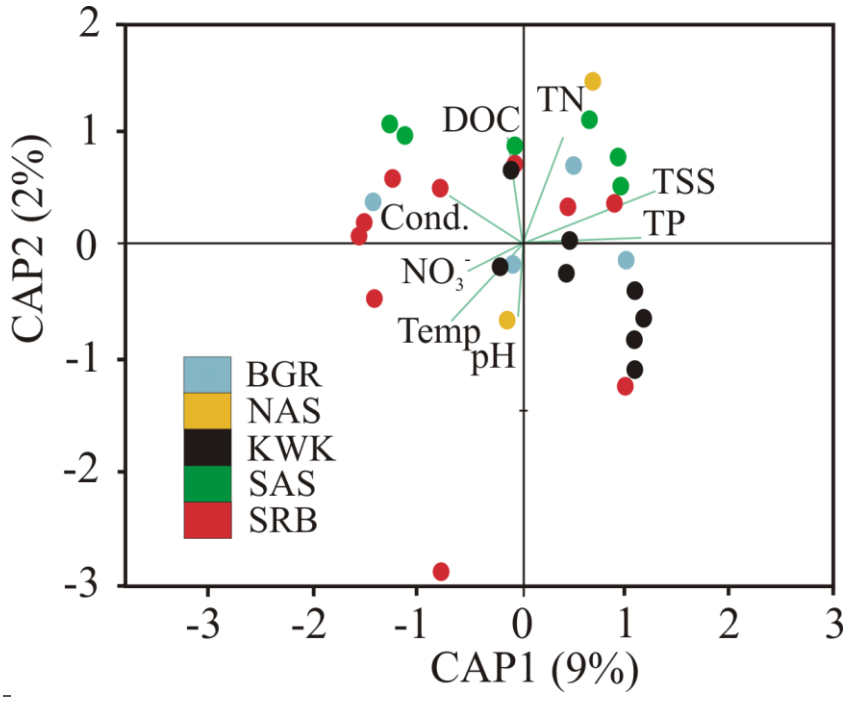

Figure 3. dbRDA of the pigment composition in the thaw and SRB lakes. Each circle represents individual lakes. The labelled lines represent the significant environmental vectors resulting from correlation analysis.

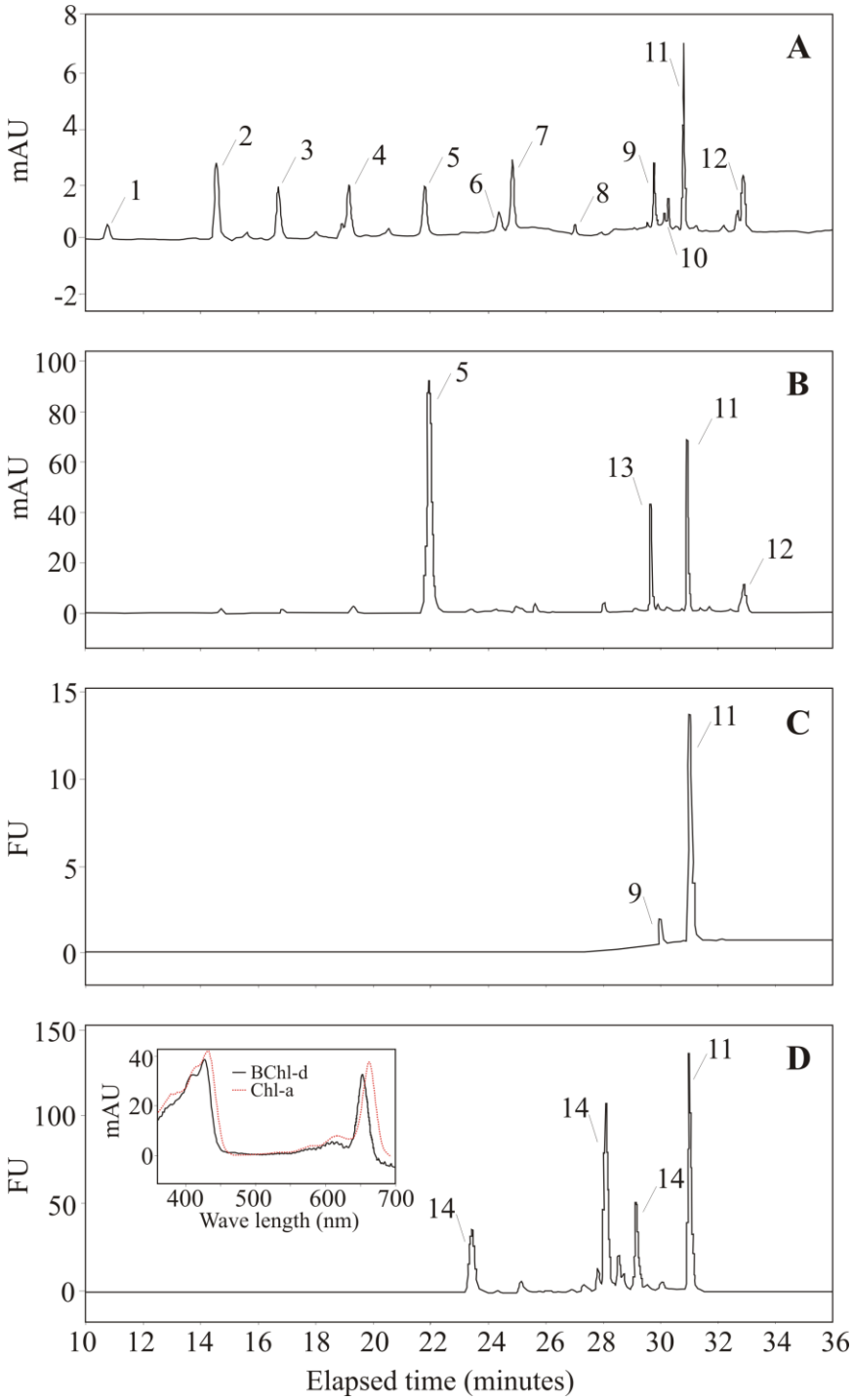

Figure 4. High-performance liquid chromatograms of lake KWK12 sampled 3 August 2012: absorbance for the surface (A) or bottom (B) water layers and fluorescence for the surface (C) or bottom (D) water layers. Pigments from left to right: 1. Perid, 2. Fuco, 3. Viola, 4. Diadino, 5. Allo, 6. Zea, 7. Lut, 8. Cantha, 9. Chl *b*, 10. Echin, 11. Chl *a*, 12. β,β-Carotene, 13. Croco and 14. BChl *d*. Insert in panel D: Bacteriochlorophyll Chl *d* (BCHl *d*, black line) and Chl *a* (red line) absorption spectra (mAU = measured absorption units).

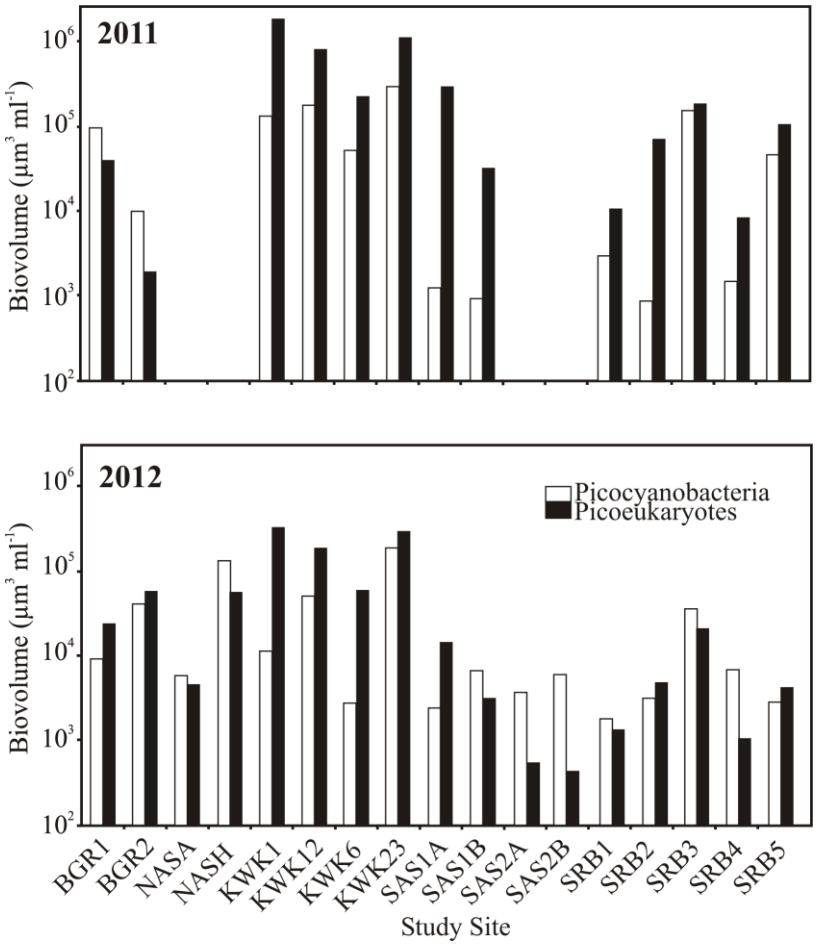

Figure 5. Picophytoplankton biovolume in the surface water of shallow rock-basin (SRB) and permafrost thaw lakes located on marine clays (KWK, BGR, NAS) and peatlands (SAS).

Supplementary material

Table S1. Location (latitude and longitude), maximum depth (Z) of the subarctic lakes and sampling dates. Water was sampled ca. 20 cm below the surface and ca. 20 cm above the maximum depth. The shallow rock basin lakes have been referred elsewhere as follows: WP1 (SRB1), WP2 (SRB2), Olsha (SRB3), 4 KM (SRB4), Iqalusiuvik (SRB5). – no sampling.

Sites	Latitude	Longitude	Z	Sampling dates	
			(m)	2011	2012
Thaw lakes on marine clays					
BGR1	56°36.650'N	76°12.900'W	3.5	20 Aug	9 Aug
BGR2	56°36.632'N	76°12.937'W	1.0	20 Aug	9 Aug
NASA	56°55.434'N	76°22.708'W	3.2	7 Aug	–
NASH	56°55.452'N	76°22.636'W	3.6	7 Aug	–
Thaw lakes on mineral clays					
KWK1	55°19.890'N	77°30.241'W	2.1	19 Aug	3 Aug
KWK6	55°19.937'N	77°30.117'W	3.2	21 Aug	4 Aug
KWK12	55°19.808'N	77°30.239'W	2.6	19 Aug	3 Aug
KWK23	55°19.947'N	77°30.131'W	3.4	21 Aug	4 Aug
Thaw lakes on peatlands					
SAS1A	55°13.128'N	77°42.477'W	1.9	23 Aug	5 Aug
SAS1B	55°13.143'N	77°42.475'W	1.7	23 Aug	5 Aug
SAS2A	55°13.591'N	77°41.815'W	2.6	–	13 Aug
SAS2B	55°13.600'N	77°41.806'W	2.0	–	13 Aug
Shallow rock-basin lakes					
SRB1	55°16.982'N	77°44.187'W	0.4	24 Aug	11 Aug
SRB2	55°16.970'N	77°44.122'W	0.8	24 Aug	11 Aug
SRB3	55°16.958'N	77°44.387'W	1.6	24 Aug	14 Aug
SRB4	55°19.907'N	77°41.959'W	0.7	16 Aug	8 Aug
SRB5	55°22.262'N	77°37.072'W	1.8	12 Aug	8 Aug