



Ontogenic succession of thermokarst thaw ponds is linked to dissolved organic matter quality and microbial degradation potential

Sari Peura ^{1,2*} Maxime Wauthy ^{3,4} Domenico Simone,^{5,6} Alexander Eiler,^{2,7} Karólína Einarsdóttir,² Milla Rautio,^{3,4,8} Stefan Bertilsson ^{2,9}

¹Department of Forest Mycology and Plant Pathology, Science for Life Laboratory, Swedish University of Agricultural Science, Uppsala, Sweden

²Department of Ecology and Genetics, Limnology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden

³Département des Sciences Fondamentales, Université du Québec à Chicoutimi, Chicoutimi, Quebec, Canada

⁴Centre d'études nordiques (CEN), Université Laval, Quebec City, Quebec, Canada

⁵Department of Plant Biology, Swedish University of Agricultural Science, Uppsala, Sweden

⁶SLU Bioinformatics Infrastructure, Swedish University of Agricultural Science, Uppsala, Sweden

⁷Centre for Biogeochemistry in the Anthropocene, Department of Biosciences, Section for Aquatic Biology and Toxicology, University of Oslo, Oslo, Norway

⁸Group for Interuniversity Research in Limnology and Aquatic Environment (GRIL), Université de Montréal, Montreal, Quebec, Canada

⁹Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Science, Uppsala, Sweden

Abstract

Warming climate is thawing the permafrost in arctic and subarctic regions, leading to formation of thermokarst ponds. During the formation and geomorphological succession of these ponds, carbon that has been trapped in frozen soils for thousands of years is hydrologically mobilized and returned to the active carbon cycle. We sampled 12 thermokarst ponds representing three different stages of pond succession to study the potential of microbial communities to metabolize the organic carbon in the water. We investigated the quality of the dissolved organic carbon (DOC) in the water column based on the spectrophotometric and fluorometric properties of the chromophoric dissolved organic matter combined with parallel factor analysis and the potential of the microbial community for degrading these carbon compounds based on genetic markers related to carbon degradation. Our analysis showed a clear difference in the DOC quality across the different developmental stages. In the younger ponds, organic matter quality suggested that it was originating from the degrading permafrost and in the metagenomes collected from these ponds, the normalized abundance of genes related to degradation of carbon compounds was higher. There was also a shift in the degradation potential in the water column of the ponds, with higher potential for organic matter degradation in deeper, anoxic layers. In conclusion, our results show that the DOC quality and the genetic potential of the microbial community for carbon cycling change across the pond ontogeny, suggesting a capacity of the microbial communities to adapt to changing environmental conditions.

Permafrost soils contain one of the major carbon storages on Earth, with an estimated 950–1672 Gt of carbon (Zimov et al.

2006b; Schuur et al. 2008). Upon thawing, the bioavailability of this ancient carbon is very high (Waldrop et al. 2010; Vonk et al. 2013) and increasing temperature is expected to further increase the carbon release from the thawing permafrost (Dorrepal et al. 2009). Alaska alone is predicted to lose 50–100 Gt of carbon from the permafrost to the atmosphere by the end of the century (Stieglitz et al. 2003; Zhuang et al. 2006). In many regions, the thawing of the permafrost ice contributes to the creation of thermokarst ponds and lakes (Osterkamp et al. 2000; Luoto and Seppälä 2003) where the ancient carbon can be reintroduced in the active carbon cycle (Roehm et al. 2009). These waterbodies represent one of the most common and widespread aquatic ecosystem types in the arctic and subarctic

*Correspondence: sari.peura@slu.se

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Additional Supporting Information may be found in the online version of this article.

Special Issue: Linking Metagenomics to Aquatic Microbial Ecology and Biogeochemical Cycles. Edited by: Hans-Peter Grossart, Ramon Massana, Katherine McMahon and David A. Walsh.

landscape with high contribution to greenhouse gas emissions (Pienitz et al. 2008; Koch et al. 2014; Wik et al. 2016).

By successive erosion and thawing, these thaw ponds can expand vertically and horizontally and eventually merge, becoming larger and more numerous in some parts of the Arctic landscape, while in other northern regions they are disappearing by drainage, evaporation, or infilling (Vincent et al. 2017). During maturation, the ponds also deepen sufficiently to enable thermal stratification that may create distinct aquatic habitats for metabolic processes (Roiha et al. 2015). The persistence of these waterbodies is usually some hundreds of years (e.g., Fillion et al. 2014) but can differ, for example, in response to wind-exposure and resultant variability in the insulating snow cover (An and Allard 1995). The systems are also sensitive to stochastic effects such as fires that may be regarded as a key trigger for thaw/regeneration cycles in the permafrost (Zoltai 1993). The lifecycle of these ecosystems is expected to be shorter under warmer climate conditions (Vincent et al. 2013) and this will in turn expose a higher number of waterbodies to strong lateral influence from the adjacent carbon-rich permafrost erosion that will enrich the lakes and ponds and accordingly stimulate bacterial activity (Vonk et al. 2013). In the areas where thawing is ongoing, a shift toward increasing terrestrial dominance of the carbon in freshwaters has already been reported (Wauthy et al. 2018).

The massive amount of previously stabilized permafrost carbon that is now being routed back into the active carbon cycle can have global consequences for the biogeochemical cycling of carbon (Tranvik et al. 2009). A large share of the carbon in the permafrost is labile and readily available for microbial use (Waldrop et al. 2010), making the emerging thaw ponds hotspots for microbial production that may act as an important positive feedback to climate warming. Compared to freshwaters that are less influenced by disappearing permafrost, the molecular composition of the carbon in thaw ponds is characterized by a larger proportion of terrestrially derived dissolved organic matter (DOM) of higher molecular weights combined with a minor influence from compounds derived from autochthonous primary production (Wauthy et al. 2018). There is however no information on how DOM quality and the associated microbial composition change throughout the long-term succession of the permafrost thaw ponds. That said, carbon quality has been shown to change during the ontogenetic aging of boreal beaver ponds (Catalan et al. 2016) and such changes are likely also happening in thaw ponds.

Despite the documented importance of the microbial loop in the permafrost thaw pond carbon cycling, studies looking into microbiology of thermokarst waters are still scarce (Vonk et al. 2015) and majority of the microbiological studies have focused on the resident methanotrophs in these ponds (e.g., Rossi et al. 2013; Crevecoeur et al. 2015; Crevecoeur et al. 2017). This focus is likely because these waters represent a source of around 24 Tg of CH₄ emission per year (Walter et al. 2007), with a predicted increase of 54% by the end of the

century from high northern latitude lakes alone (Wik et al. 2016). However, prior to becoming CH₄, the carbon compounds coming from permafrost have to be extensively processed by various heterotrophic organisms (Tveit et al. 2013). Our current understanding of these processes is very limited, but using next generation sequencing methods, we can now study what kind of processes and substrate transformations the microbial communities have the potential to perform. Indeed, for permafrost soils, some information on microbial metabolic potential and carbon degradation already exists (Yergeau et al. 2010; Mackelprang et al. 2011; Tveit et al. 2013; Woodcroft et al. 2018). These preceding studies have shown the presence of genes related to plant polymer hydrolysis (Tveit et al. 2013) and organic matter degradation (Yergeau et al. 2010; Woodcroft et al. 2018) and even CH₄ cycling (Mackelprang et al. 2011) in the soils affected by permafrost and permafrost thaw. However, there are no analogous studies for the aquatic environments affected by permafrost thaw. We addressed this knowledge gap by studying the microbial pathways for degrading carbon compounds and the taxonomic composition of organisms involved in the degradation.

For this purpose, we used shotgun sequencing of the metagenomic DNA from 11 permafrost thaw ponds ranging in age from approximately 3-yr-old emerging ponds to 60-yr-old mature ponds (Wurzbacher et al. 2017). Due to the strong hydrological linkages with the surrounding soils and permafrost and the limited time for diagenetic “aging” of the soil organic matter before reaching the surface water, the carbon in young ponds was expected to be characterized by readily degradable terrestrial carbon fractions (Zimov et al. 2006a; Gao et al. 2018). Analogously, we expected the carbon compounds in the mature ponds to be more strongly influenced by autochthonous production and diagenetically aged terrestrial carbon compounds (Wauthy et al. 2018). Associated with this, we also hypothesized that the microbial genetic potential for carbon degradation would be tightly coupled with the carbon quality and that it would shift across the pond succession gradient and over depth in the stratified mature ponds. We further hypothesized that the taxonomic identities of the organisms contributing to carbon degradation across the lateral (among ponds) and vertical (within the depths of deep ponds) gradients would vary.

Materials and methods

Sample site and sampling

Twelve permafrost thaw ponds located on a palsa bog in a river valley in the sporadic permafrost zone (55°13'N, 77°42'W) close to the village of Whapmagoostui-Kuujuarapik (Nunavik, QC, Canada) were sampled once in the beginning of August 2014. Palsas are frost heaves containing permanently frozen ice lenses, with an ice core inside of the overlying soil (Seppälä 2006). The height of palsas at the study site is up to 3 m and the age of the organic material in the top 1.5 m

varies from 700 yr in the surface to 5000 yr at a depth of 1.6 m (Fillion et al. 2014). The vegetation is dominated by *Carex* and *Sphagnum* (Bhiry et al. 2011). The thermokarst ponds at the study site are formed after thawing of palsas. We chose thaw ponds representing three different phases of pond succession; (1) young ponds which were emerging at the edge of palsas and that had not been present a few years earlier (depth 0.10–0.30 m), (2) developing and highly dynamic ponds located around an existing but thawing palsas (depth 0.30–0.70 m), and (3) mature ponds where there was no palsa left (depth 2–4 m) (see Supporting Information Figs. S1, S2 for the size of the ponds). The first two pond types are still growing as the permafrost is thawing and the palsas are sinking to the pond. The mature ponds were approximately 50–60 yr old (dating based on satellite images) and thermally stratified. The mature phase of the ponds at this site has been shown to last for 300–400 yr (Fillion et al. 2014). The sampling was conducted as previously described (Wurzbacher et al. 2017). In short, surface waters from four ponds of each age category were sampled. When the pond depth exceeded 50 cm, separate samples were retrieved from the epilimnion, metalimnion, and hypolimnion. All ponds were located within a 100 m radius. In order to analyze dissolved organic carbon (DOC) concentrations and perform optical analyses of the chromophoric dissolved organic matter (CDOM), water samples were collected from the surface and filtered through prerinsed cellulose acetate filters (0.2 μm). This filter pore size removes small-size inorganic soil particles common in thaw ponds (Watanabe et al. 2011) and was chosen to be consistent with earlier studies carried out on circumpolar ponds (Roiha et al. 2015; Wauthy et al. 2018). The samples were stored in acid-washed and combusted glass vials at 4°C in the dark pending carbon characterization. Samples for chemical analyses (Fe^{2+} , Fe^{3+} , NH_4 , NO_2 , NO_3 , total N, PO_4 , SO_4 , pH) and CH_4 and CO_2 concentrations were collected from all targeted ponds and depth layers.

Chemical analyses

CH_4 and CO_2 concentrations were measured as in Kankaala et al. (2007) with the exception that instead of N_2 , ambient air was used to extract gas from the water. Shortly, 30 mL of water was collected to 60-mL syringes from each sampled layer and transported to the laboratory on ice. In the laboratory, the syringes were incubated in a 20°C water bath for 5 min and 30 mL of room air was added to each syringe. Following the addition of air, the syringes were shaken for 2 min to equilibrate the gas concentration and air was collected to glass vials. The CH_4 concentration of the ambient air was quantified in parallel and subtracted from the measured values to determine final concentrations. Nutrients were measured at Erken laboratory (Uppsala University, Sweden) using standard methods (https://www.ieg.uu.se/digitalAssets/437/c_437648-l_3-k_metodfo-rteckning-28.pdf) and pH was measured using Micro-pH 2001 (Crison Instruments, Barcelona, Spain).

Carbon characterization

DOC was measured from filtered water (GF/F glass fiber filter, Whatman, Little Chalfont, UK) using Shimadzu TOC-V analyzer (Shimadzu Corporation, Kyoto, Japan) and total organic carbon (TOC) from unfiltered water using TOC-L analyzer (Shimadzu Corporation, Kyoto, Japan). A set of carbon quality indicators was measured from the filtered water using spectrophotometric and spectrofluorometric methods. The absorbance spectra were measured in 1 cm quartz cuvettes using a UV-visible Cary 100 spectrophotometer (Agilent, Santa Clara, CA, U.S.A.), recording absorbances between 250 and 800 nm at 1 nm intervals. Milli-Q water was used as blank and spectra for samples were null-point adjusted by subtracting the mean blank value from 750 to 800 nm to correct for offsets due to instrument baseline effects (Helms et al. 2008). The absorbance measured at 254 nm (A_{254}) was normalized for DOC to calculate the specific ultraviolet absorbance at 254 nm (SUVA_{254}), which is used as a proxy for aromaticity of CDOM (Weishaar et al. 2003). Iron can bind to humic substances and form complexes that may increase both absorbance (Weishaar et al. 2003) and fluorescence (McKnight et al. 2001). To account for this, a correction factor developed for SUVA_{254} using concentrations of dissolved iron Fe^{3+} was applied as in Poulin et al. (2014). The absorption coefficient at 320 nm (a_{320}) was used as an index of CDOM concentration and was determined following the equation:

$$a_{320} = 2.303 \times A_{320}/L, \quad (1)$$

where a_{320} is the absorption coefficient (m^{-1}) at 320 nm, A_{320} the absorbance corrected at 320 nm, and L the path length of the cuvette (m) (Williamson et al. 1999). CDOM absorption spectra were further analyzed to calculate the spectral slopes S_i as described in Loiselle et al. (2009) for the intervals 279–299 nm (S_{289} , named by center wavelength), 275–295 nm (S_{285}), and 350–400 nm (S_{375}). All linear regressions were performed using SciLab v 5.5.2. (Scilab Enterprises, Orsay, France). S_{289} was used as a proxy to estimate the importance of fulvic and humic acids related to algal production (Loiselle et al. 2009), while the slope ratio (S_R) S_{285}/S_{375} is an index of CDOM molecular weight (Helms et al. 2008).

Fluorescence spectra were recorded on a Cary Eclipse spectrofluorometer (Agilent, Santa Clara, CA, U.S.A.) across the excitation wavelength range 250–450 nm (10 nm increments) with emission between 300 and 560 nm (2 nm increments) in order to construct excitation-emission matrices. The fluorescence spectra were corrected for Raman and Rayleigh scattering and inner filter effects, and standardized to Raman units using the FDOMcorr 1.4 toolbox (Murphy et al. 2010). In order to investigate the origin of fulvic acids, the fluorescence index (FI) was calculated as the ratio of fluorescence emission intensities at 450 nm and 500 nm with 370 nm excitation (McKnight et al. 2001). To identify and quantify the main DOM components, a parallel factor (PARAFAC) model was run

on the 12 excitation-emission matrices developed in this research combined with 117 extra-matrices collected from previous studies in 85 arctic and subarctic ponds. The PARAFAC modeling was performed on MATLAB v R2013a (MathWorks, Natick, MA, U.S.A.) as previously described (Murphy et al. 2013). The model was validated by split-half analysis (Supporting Information Fig. S3) using the drEEM toolbox (Murphy et al. 2013). The maximum fluorescence values [C_x] of each component x for each particular sample were summed to determine the total fluorescence (F_T) and the relative abundance of any component x was calculated according to the following equation:

$$\%C_x = ([C_x]/F_T) \times 100 \quad (2)$$

To identify the components of the model, their respective excitation and emission spectra were compared to published components from 70 models available in the OpenFluor database, as previously described (Murphy et al. 2014).

DNA extraction and sequencing

Microbial cells for DNA extraction and analyses were collected on Sterivex filters (Millipore, Massachusetts, U.S.A.) from each of the ponds and sampling depths by filtering water through the filter until it clogged (150–450 mL per sample). The filters were stored at –80°C until extraction, which was done using the Mobio PowerSoil DNA extraction kit (Mobio, Carlsbad, CA, U.S.A.). DNA concentrations were measured using PicoGreen (Invitrogen, Paisley, UK) and the DNA was fragmented using a Covaris E220 ultrasonicator (Covaris, Brighton, UK). Libraries were subsequently prepared using ThruPLEX FD Prep kit (Rubicon Genomics, Ann Arbor, MI, U.S.A.) according to the manufacturer's instructions and then sequenced at Science for Life Laboratory (Uppsala, Sweden) using an Illumina HiSeq 2500 system operating in high output mode. The raw data have been deposited to NCBI Short Read Archive under accession number PRJNA517163.

Processing of the metagenome data

The sequencing produced 163 Gb of total raw sequence data (range 4.3–11.4 GB per sample, Supporting Information Table S1). The raw reads were quality-trimmed using Sickle (v1.210) (Joshi and Fass 2011) with default parameters, then coassembled into contigs using Megahit (v1.1.2) (Li et al. 2015). This resulted in a combined assembly of 2.2 Gb of contigs longer than 1 kb. The raw reads were mapped against the assembly using BBMap (Bushnell 2016) to define coverage of individual samples in the assembly. Potential for carbon degradation was studied based on hidden Markov models (HMMs). HMMs are heuristic models of protein domains from groups of functionally related proteins (i.e., protein families). Each HMM is based on an alignment of the sequences from multiple proteins. These models have a robust way of handling sequence variation, such as insertions and deletions, providing higher

precision for protein annotations. Thus, the amount of false positives is lower than with most other methods, such as blast searches (Sonnhammer et al. 1998). For this study, we chose a set of HMMs from the PFAM-database (database for Protein Families; Finn et al. 2008) that have previously been used to study carbon degradation in permafrost soil (Tveit et al. 2013). These included HMMs associated with carbohydrate binding molecules, carbohydrate esterase, glycoside hydrolases, phenolic compound degrading domains, polysaccharide lyase, and a group of other carbohydrate active enzymes (Tveit et al. 2013). Presence of these marker genes in the dataset indicates that the microbial community has the potential for conducting the reaction catalyzed by the protein that the gene produces. To identify the marker genes in the data, protein-coding genes (CoDing Sequence, CDS) were predicted on the assembled contigs with Prodigal (v2.6.3) (Hyatt et al. 2010) in metagenomic mode. CDS were then annotated against the PFAM database (release 32) (El-Gebali et al. 2019) using hmmsearch (from the hmmer suite, version v3.1b2) (Durbin et al. 2002). For each sample, the abundance of CDS related to carbon degradation was calculated with htseq-count (v0.9.1) (Anders et al. 2015) in “–nonunique none” mode, as suggested by the authors to perform differential abundance (or gene expression) analysis. A separate calculation of the abundance of CDS related to carbon degradation in “–nonunique all” mode was performed for descriptive purposes and full results are reported as transcripts per million reads (TPM) in Supporting Information Table S2.

The taxonomic affiliation of the organisms engaged in carbon degradation was identified from the contigs using Kaiju (Menzel et al. 2016), with NCBI nr as reference database.

Statistical testing

All statistical analyses were done using R (R Core Team 2015). Differences in the environmental factors, carbon quality, and total degradation potential (total normalized abundance of degradation PFAMs) across different age classes were tested using one-way ANOVA with Tukey's post hoc test with *p* value adjustment. For these analyses, three epilimnion samples were included for emerging and mature ponds and four epilimnion samples for developing ponds. SUVA₂₅₄ and S_R were transformed using command “transformTukey” included in the R package rcompanion (Mangiafico 2015) in order for them to meet ANOVA assumptions. For the variables that did not fulfill the ANOVA assumptions even after transformation (O₂, NH₄, NO₂, NO₃, PO₄, SO₄, pH), differences were tested using Kruskal–Wallis tests and Wilcoxon post hoc tests with *p* value adjustment (Benjamini and Hochberg 1995; Benjamini and Yekutieli 2001).

To study the variation in abundances of PFAMs related to carbon degradation across different age classes and depth layers, PFAM abundances were rarefied to the sample with the lowest abundance, Bray–Curtis distances of the samples were calculated using the R package vegan (Oksanen et al. 2017) and these distances were visualized with nonmetric multidimensional scaling (NMDS). Differential abundance analysis

of PFAMs in the samples originating from ponds representing different ages (emerging, developing, mature), or different depth layers within the mature ponds (oxic vs. anoxic) was performed using the R package DESeq2 (Love et al. 2014).

The difference in the diversity of PFAMs related to carbon degradation across different age classes and depth layers was tested using permutational multivariate analysis of variance (PERMANOVA) of Bray-Curtis distances of the samples. This was used as a proxy for qualitative total potential for degrading carbon compounds. The difference in total normalized abundance of PFAMs related to carbon degradation between oxic and anoxic compartment of the mature ponds was tested from the sum of all normalized abundances of carbon degradation related PFAMs using Welch two-sample *t*-test, while the difference in total normalized abundance of PFAMs related to carbon degradation in the oxic layer across ponds representing different ages was tested using one-way

ANOVA. This was used as a proxy for quantitative total potential for degrading carbon compounds. Linkages between the carbon quality parameters and specific PFAM abundances were tested using Spearman correlations with Holm's correction for multiple inference (Holm 1979) (using the *rcorr.adjust* function from the R package RcmdrMisc) (Fox et al. 2018) and relationships with *p* value < 0.05 were visualized in a heat map.

Results

Chemical parameters

Environmental conditions at the surface layer of the ponds were similar across the full set of ponds that represented different stages of thaw pond formation. There was no difference in concentrations of O₂, CH₄, CO₂, iron, NH₄, NO₂, NO₃, PO₄, and SO₄. However, pH was lowest in the emerging and highest

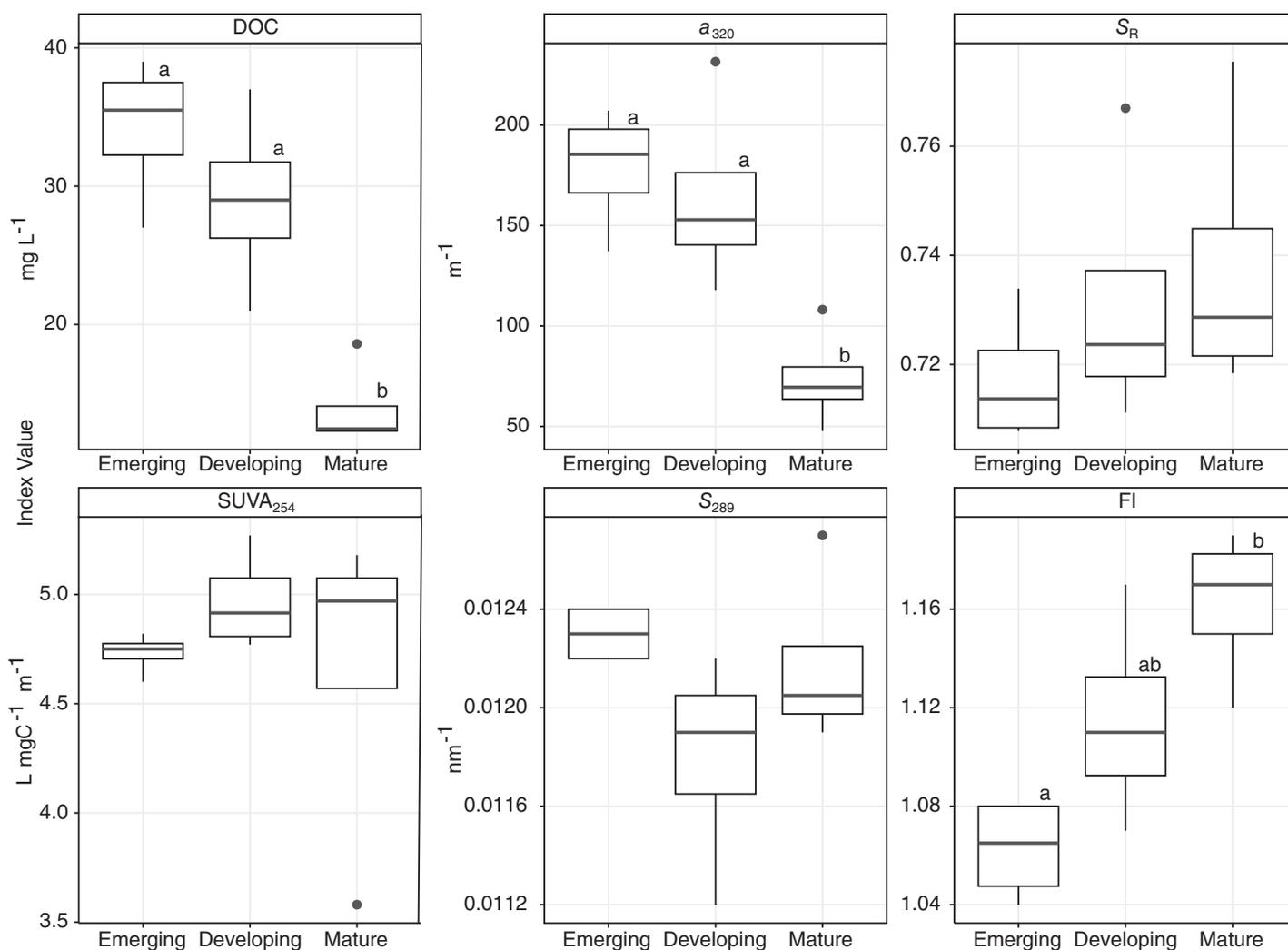


Fig. 1. Boxplots of the concentration of DOC (mg L^{-1}) and optical proxies of carbon in ponds according developmental stages. Optical proxies include a_{320} , S_R , $SUVA_{254}$, S_{289} , and FI. The letters indicate significant differences between the different developmental stages. $N = 4$ in each category.

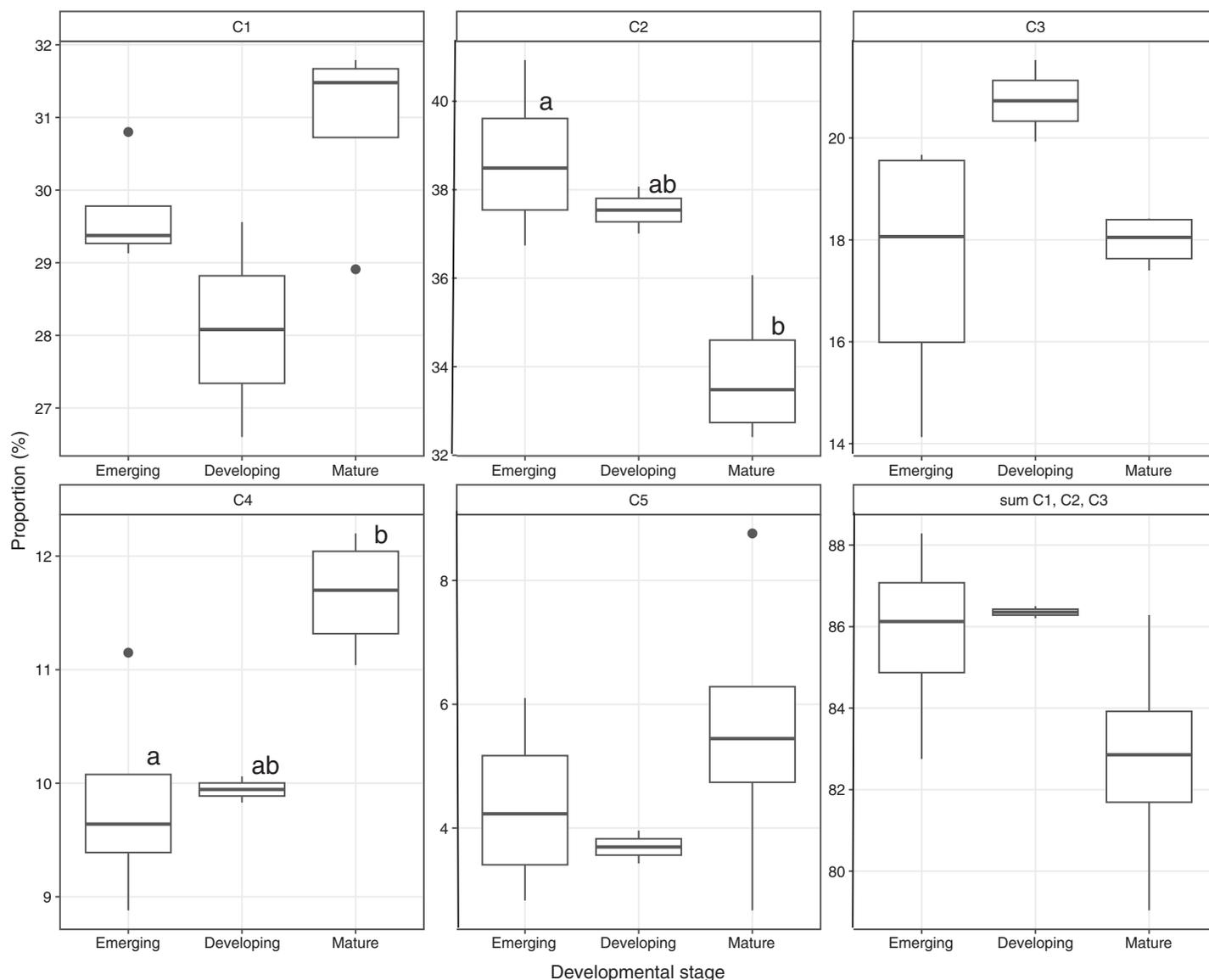


Fig. 2. Boxplots of the five components identified by the PARAFAC model in ponds according developmental stages. The letters indicate significant differences between the different developmental stages. $N = 4$ in emerging and mature category and $N = 2$ in developing ponds category.

in the mature ponds, while TOC and total N were higher in emerging and developing ponds than in the mature and temperature was higher in emerging than in developing ponds (Supporting Information Fig. S4, Table S3; Wurzbacher et al. 2017). In the mature thaw ponds, the environmental factors were also tested for different water depth layers. As expected, most chemical variables were significantly different between the water depth layers (Supporting Information Fig. S5). Oxygen and temperature decreased from surface to bottom, while for CH_4 , CO_2 , TOC, total N, and Fe^{2+} the trend was opposite and pH was higher in the epilimnion than in the metalimnion. For Fe^{3+} , NH_4 , NO_2 , NO_3 , PO_4 , and SO_4 , there was no difference between the depth layers.

Carbon characterization

The characterization of carbon quality revealed a strong effect of the thaw pond developmental stage on DOC quality and quantity (Fig. 1), with significant differences for DOC, a_{320} , and FI (Supporting Information Table S3). The DOC concentrations varied from 12.3 to 39.0 mg L^{-1} (mean = 25.7 mg L^{-1} , SD = 10.1), with the highest values found in the emerging (mean = 34.3 mg L^{-1} , SD = 5.3) and developing ponds (mean = 29.0 mg L^{-1} , SD = 6.6). CDOM concentration (a_{320}) followed the same trend. In contrast, S_R and FI were lowest in emerging thaw ponds, but the difference was only significant for FI. Moreover, the highest values of FI were found in mature thaw ponds, suggesting that these

systems had a higher content of microbial-derived fulvic acids (McKnight et al. 2001). No significant differences or patterns were observed for $SUVA_{254}$ and S_{289} , suggesting that the developmental stage had no strong effect on the aromaticity and the humic and fulvic content originating from algal production.

The PARAFAC model identified five fluorescence components (Supporting Information Fig. S6), of which four shared fluorescence properties with humic materials from terrestrial (C1–C3) and microbial origin (C4). The fifth component (C5) presented spectra similar to amino acids or proteins and is known to be associated with algal, microbial or terrestrial production (more details in Supporting Information Table S4). The combined percentage of terrestrial humic components (C1–C3) showed the highest values in emerging and developing thaw ponds ($85.8\% \pm 2.3\%$ and $86.4\% \pm 0.2\%$, respectively) (Fig. 2). However, these differences were significant only for C2. Furthermore, along the developmental stages gradient, C1 showed an inverse trend compared to C3. Moreover, mature thaw ponds had the highest mean value for C4 ($11.7\% \pm 0.5\%$), indicating increasing representation of humic-dominated microbial material in these ponds. Finally, the proportion of C5 varied between 2.7% and 8.8% but was not significantly different among pond types.

Potential for carbon degradation in the surface layer

The microbial carbon degradation potential, as determined from diversity of PFAMs indicative of carbon degradation (Tveit et al. 2013), differed between ponds representing different developmental stages. Significant differences in diversity were also seen for the different sampling depths (Fig. 3, PERMANOVA, $p < 0.01$ and $p < 0.05$ for age and depth, respectively). There was no significant difference in the total degradation potential among pond types (the sum of normalized abundances of all carbon degradation PFAMs per sample; one-way ANOVA $p = 0.0654$, $F = 3.29$). This suggests that while the total normalized abundance of genes related to carbon degradation were similar across the ponds, the distribution of different PFAMs varied between the age classes. In addition, several individual PFAMs indicative of degradation potential had a significantly different abundance across age classes (Fig. 4, Supporting information Table S5). In general, carbon degradation genes with significantly different abundance were typically more abundant in the two younger pond classes. These differences were particularly pronounced for genes related to degradation of polymeric carbohydrates (glycosyl hydrolases, polysaccharide lyases) with overall higher abundances in the younger ponds.

In general, the organisms carrying degradation genes represented a diverse set of taxa, including mostly bacterial organisms, but also eukaryotes such as fungi, and viruses (Fig. 5, Supporting Information Fig. S7). Many of the individual PFAMs with differential abundance across ponds, such as PF00150, a cellulase belonging to glycosyl hydrolase family 5, were present in a wide variety of different taxa, whereas some, such as PF03663, a member of glycosyl hydrolase

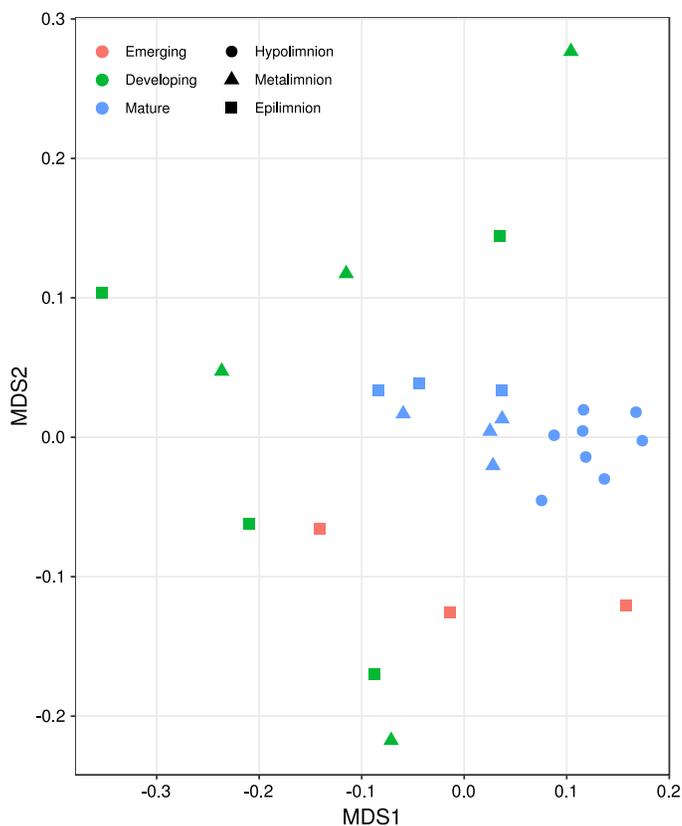


Fig. 3. NMDS plot visualizing the variation between the ponds in the diversity of potential for carbon degradation based on their PFAM abundances.

family 76, had a very limited taxonomic range. A number of fungal taxa contributed to the degradation potential, including organisms related to Ascomycota, Basidiomycota, and Mucoromycota. For some PFAMs, such as PF00150 (glycoside hydrolase family 5 [GH5]; Supporting Information Fig. S7), taxonomic classification suggested significant contribution from viruses to the degradation potential. It also appeared that there was a shift in the type of carbohydrate binding domains across the age classes, with emerging and developing ponds having the same dominant domains, while in the mature ponds, the dominant domains were different (Fig. 4A).

Shifting potential for carbon degradation across the depth gradient

For the mature ponds, the total potential for carbon degradation (the sum of the normalized abundances of all carbon degradation PFAMs per sample) was higher in the anoxic layer compared to the oxic ($p < 0.001$, $t = -12.82$). The normalized abundance of several PFAMs was significantly different between the oxic and anoxic layers (Fig. 6). For example, carbohydrate-binding modules were much more abundant in the surface layer (Fig. 6A) and this was also seen for pectate lyase (PF12708) suggested to participate in the degradation of

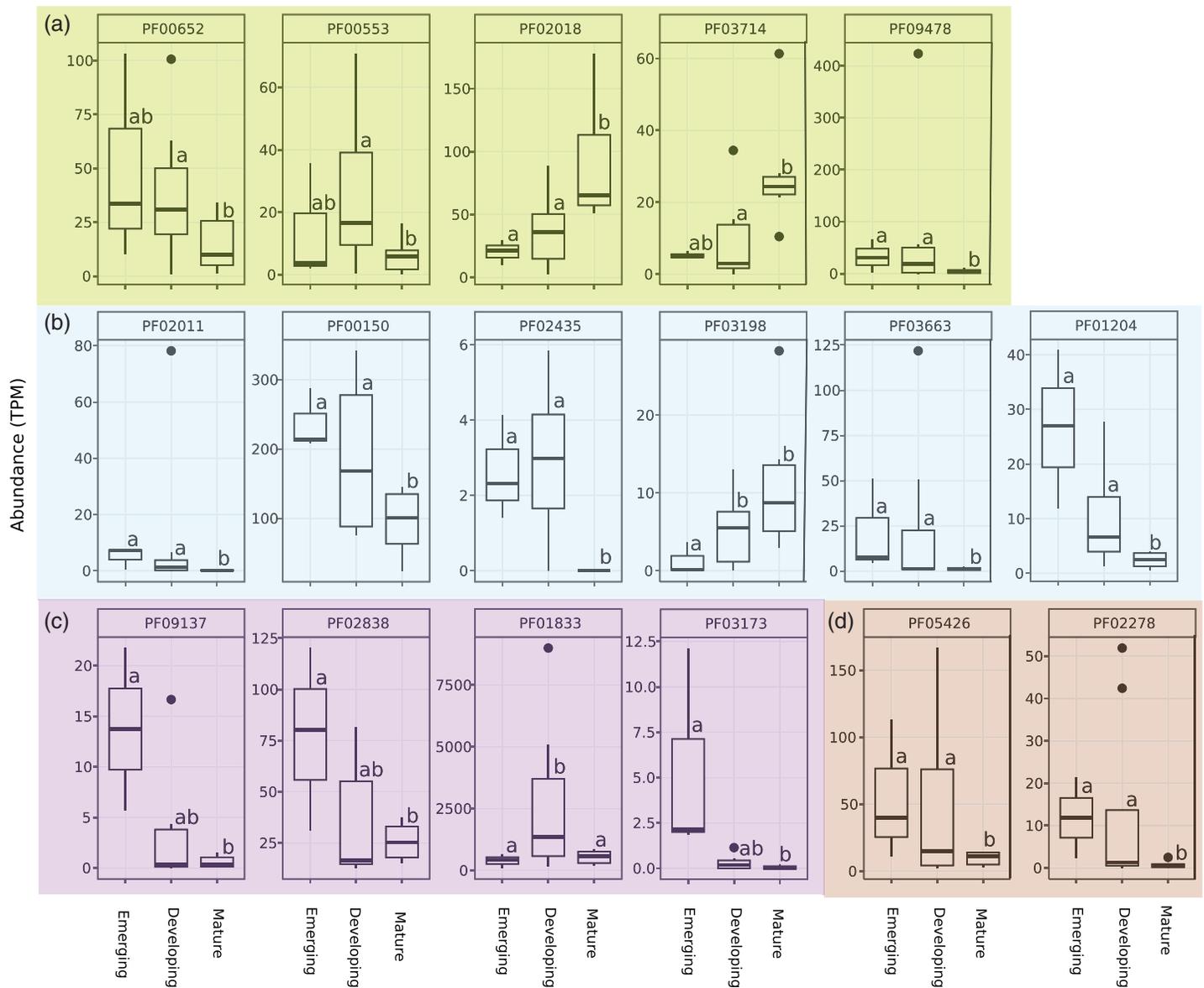


Fig. 4. The normalized abundances of significantly different epilimnetic PFAMs between pond age classes. PFAMs were related to (a) carbohydrate binding modules, (b) glycoside hydrolases, (c) other carbohydrate active enzymes, and (d) polysaccharide lyases, respectively. The letters denote statistical differences between the ponds representing different age classes.

plant tissue (Hugouvieux-Cotte-Pattat et al. 2014) and genes that are involved in chitin degradation (PF01607 and PF00182). Many of the glycosyl hydrolases were much more abundant in the oxic layer, suggesting higher potential for degrading complex carbohydrates under such conditions.

Taxonomically the potential for degradation was also in this case related to a wide variety of bacterial taxa as well as fungi and viruses (Fig. 5, Supporting Information Fig. S8). Also some taxonomic differences in the organisms carrying degradation potential were apparent between the different layers. In the oxic layer, many PFAMs involved in degradation were affiliated with Proteobacteria, and these genes were also often significantly different in abundance between the layers.

Example of this is PF00182 (chitin degradation) that was abundant in Alpha- and Gammaproteobacteria. In the anoxic layer, especially Chlorobi and candidate phylum Omnitrophica contributed to the degradation potential via multiple PFAMs, such as PF01510 (N-acetylmuramoyl-L-alanine amidase) participating in cell wall degradation and PF2578 (laccase) which acts on compounds such as phenols and lignin. Also three PFAMs related to glycosyl hydrolase family 65 (PF03632, PF03633, and PF03636) were affiliated with Chlorobi in the anoxic compartment. Enzymes in this family act on substrates containing α -glucosidic linkages. Phylum Bacteroidetes contributed to most PFAMs and was found in both oxic and anoxic layers.

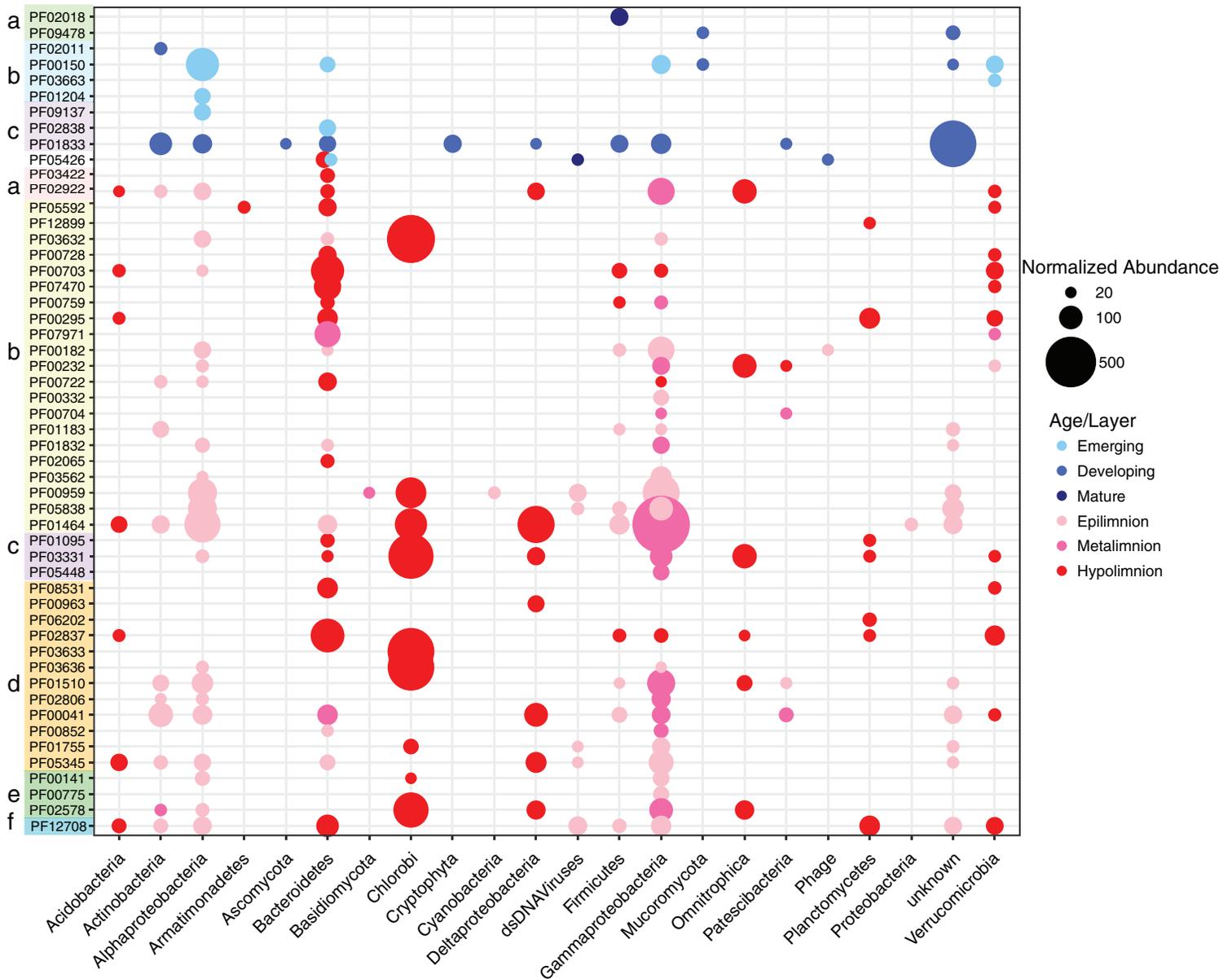


Fig. 5. Taxonomic distribution of organisms with highest contribution to the PFAMs with differential normalized abundance across age groups (blue shades) and vertical layers (red shades). The plot includes all taxa with minimum of 20 TPM average coverage in at least one of the age groups or vertical layers. In the figure, for each PFAM and taxon, only the value of the age group/vertical layer with highest TPM is shown, representing the environmental preference of the taxon. Full taxonomic ranges of all PFAMs are presented in Supporting Information Figs. S7, S8. Letters on the y-axis highlight the groups specified in Figs. 4, 6 with the exception of PF05426, which is present in both figures (panels d and f, respectively).

We also explored the possible connections between the observed carbon degradation potential and carbon quality using a correlation analysis (Supporting Information Fig. S9). In the analysis, several PFAMs had significant correlations with the various carbon quality measures. For example, three PFAMs (PF03512, PF01229, and PF02156) related to xylan degradation were positively correlated with fraction C2, an indicator for terrestrial carbon. Xylan is a hemicellulose originating from plants, suggesting a link between carbon quality and genetic potential of the microbial

community. However, several other PFAMs also related to xylan degradation (e.g., PF00722 and PF00457) were positively correlated with carbon fractions not as directly associated with plant production (C4), suggesting that the quality measures used here are too broad for linking organic matter substrates to individual PFAMs. In future studies, application of high resolution mass spectrometry and other high-resolution methods for organic matter characterization may shed more light on whether or not there are such direct linkages.

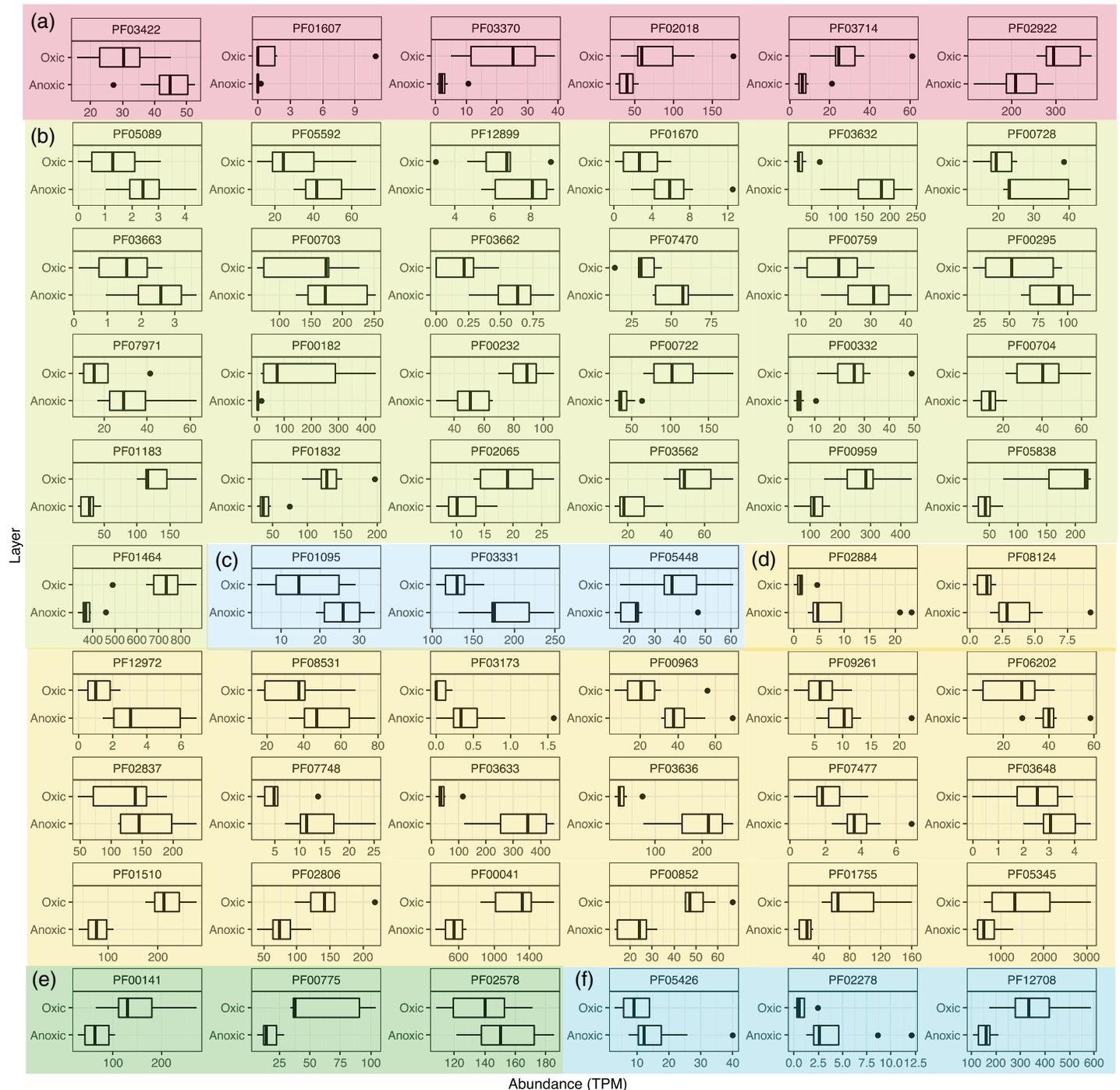


Fig. 6. The normalized abundances of PFAMs with significantly different distribution across the depth gradient in the mature ponds. PFAMs related to (a) carbohydrate binding modules, (b) glycoside hydrolases, (c) carbohydrate esterase, (d) other carbohydrate active enzymes, (e) phenolic compound degradation domains and (f) polysaccharide lyases.

Discussion

High terrestrial inputs to emerging ponds

In this study, we assessed how DOC quality and potential for microbial organic carbon degradation change along the ontogenetic development of permafrost thaw ponds in a

subarctic site exposed to on-going permafrost thaw. This natural gradient in thawing conditions provides a more realistic strategy for studying the impact of climate warming on permafrost compared to widely used short-term laboratory incubations (Monteux et al. 2018). The studied ponds were in close proximity to each other, eliminating the possible impact

of many other environmental factors such as vegetation, weather, and geology, thus, ensuring that the observed differences are most likely due to the different stages of permafrost thaw and pond development. The ontogenetic stage of the pond had a strong influence on the carbon quantity and quality, as well as on the carbon degradation potential of the microbial communities. Our proxies suggested the highest average molecular weight as well as absolute DOC and CDOM concentrations in the emerging and developing thaw ponds, likely reflecting a higher supply of fresh allochthonous carbon at the onset of permafrost thaw, but possibly also the closer physical proximity to the sediment. However, the TOC concentrations in the bottom samples from the mature ponds (Supporting Information Fig. S5), taken equally close to the sediment, were significantly lower than the TOC in the emerging ponds (Supporting Information Fig. S4), suggesting that the carbon flow from the thawing palsa had a stronger influence on the carbon concentration than the proximity to the sediment.

Despite the fact that the TOC concentration suggested higher carbon flow from the thawing palsas to the younger ponds, the quality of this supposedly allochthonous organic carbon in the younger ponds did not seem to adhere to the regular characteristics of highly chromophoric humic substances, for example, no significant differences in $SUVA_{254}$ could be observed compared to the old ponds. Also, S_{289} did not vary across the age gradient, which could suggest that the differences according to developmental stage were moderate. Previously both proxies have been shown to differentiate, for example, the carbon source in thaw ponds compared to bedrock and tundra ponds (Wauthy et al. 2018), but it is possible, that these proxies are not sensitive enough to capture the variation in the carbon quality within this group of thaw ponds. Nevertheless, the higher sum of terrestrial humic components C1–C3 in the younger thaw ponds, although not significant, supported the assumption that there is stronger terrestrial input of organic compounds in the early stages of permafrost thaw.

Among the PARAFAC components, C1 and C3 have both been previously associated with humic materials of terrestrial origin, but in our data, they responded in contrasting ways to the aging of the ponds (Fig. 2). This has previously been explained by their interrelationship where one of these two components could be a product from the biological transformation of the other (Jørgensen et al. 2011). Furthermore, the higher content of fulvic (F1) and humic (C4) acids of microbial origin in mature thaw ponds suggests that this material has experienced longer exposure to microbial degradation and accordingly feature diagenetic changes in DOM quality. Because of the microbial origin of C4, we cannot determine if this material is from algal or terrestrial sources (Wauthy et al. 2018). It should however be noted that coupled to microbial degradation, photochemical transformation of terrigenous and autochthonous DOM is expected to be very important in thermokarst systems (Cory et al. 2013). Thus, this may also affect the DOM pool quality particularly in the more exposed mature thaw ponds. Consistent with a previous study, the amino acid- or protein-like component C5 featured very low

values in all thaw ponds, highlighting the overwhelming inputs of terrestrial humic and fulvic acids from the degrading permafrost (Wauthy et al. 2018).

Bacterial, viral, and fungal contributions to carbon degradation potential

The degradation potential was linked to various bacterial taxa. In general, the taxonomic composition of the microbial communities in the thaw ponds was in agreement with previous studies in peatland environments, with high contribution of Proteobacteria, Bacteroidetes, Verrucomicrobia, and Actinobacteria to organic matter degradation potential (Tveit et al. 2013; Ivanova et al. 2016; Wurzbacher et al. 2017; Woodcroft et al. 2018). Bacteroidetes was a major contributor to degradation potential in all age categories and different layers of the ponds. This agrees with a previous study on carbon degradation in Arctic peat soils (Tveit et al. 2013). Bacteroidetes are also well documented to participate in carbon degradation in other environments, such as anoxic marine sediments (Julies et al. 2010) and rumen (Solden et al. 2017). In the anoxic layer of the mature ponds, Chlorobi appeared to contribute to degradation potential. Typically, members of Chlorobi are associated with photoautotrophic energy metabolism and this is most likely also the case here as the Chlorobi represented the typical phototrophic taxa. Recently, Chlorobi was also found to have a major contribution to nitrogen fixation potential in the same thaw ponds studied here (Fernandez et al. 2019), and members of this phylum are also known to carry potential for degradation of aromatic compounds (Fuchs et al. 2011). While Chlorobi was specific for the anoxic layer, also other bacterial taxa contributed to the higher degradation potential in the anoxic compartment.

We could not classify any of the found degradation marker to a specific viral taxon, but for several PFAMs, there was a clear viral signal. A previous study, specifically addressing the viral community, showed that in one of the mature ponds (SAS2A), chloroviruses, typically infecting algae, and myoviruses, which are bacteriophages, were dominating the viral community (Lévesque et al. 2018). The authors also suggested that the DOC concentration was the primary driver of viral community composition and diversity, further linking viruses to carbon cycle. In fact, viruses have been suggested to impact the carbon cycle in aquatic ecosystems (Worden et al. 2015). A recent study from a similar permafrost thaw site in northern Sweden showed that in peat, viruses contributed directly to carbon cycling by carrying genes for enzymes that degrade plant-derived polymers, as well as indirectly by infecting major carbon degrading bacteria (Emerson et al. 2018; Trubl et al. 2018). Furthermore, an enzyme belonging to GH5 (PF00150), and found from a virus, was tested and found to be fully functional (Emerson et al. 2018). These results suggest that viruses influence carbon degradation in more complex ways than just via host lysis (Emerson et al. 2018). The same PFAM PF00150 was found also in our study to have viral contribution suggesting

that viruses could have a direct as well as indirect role in carbon degradation also in the ponds investigated by us.

Our data also included several PFAMs that appeared to be of fungal origin. Studies from northern peatlands have shown a significant influence of fungi in carbon cycling (Thormann 2006) and in the tundra, rising temperature has been shown to increase the abundance of fungi (Xue et al. 2016). Here, one of the fungal taxa found was Mucoromycota, which are known to participate in cellulose degradation in, for example, agricultural soil (Koechli et al. 2019). However, for both, viruses and fungi the databases lack behind, limiting the taxonomic information available and further studies are needed to better map the taxonomic identities of viral and fungal organisms in permafrost environments.

The impact of pond age and water layer to carbon degradation potential

The potential of the microbial community for degradation of complex organic compounds was higher in the younger ponds (Fig. 4), which was in concordance with what was observed for the carbon quality. Individual PFAMs with significantly differential abundances across the developmental stages were also typically more abundant in emerging and/or developing ponds. Additionally, PFAMs with significantly higher abundances in the younger developmental stages included enzymes related to degradation of carbohydrate compounds likely originating from terrestrial production, such as plant polymers (e.g., polysaccharides). For example, PF00150, referring to GH5, was abundant in the surface layer of all the ponds, while still significantly more abundant in the younger stages than in the mature ponds (Fig. 4B). This family of enzymes includes a variety of substrate specificities, such as endoglucanases (cellulase), exoglucanases, and high specificity xyloglucanases. These enzymes target polymers typically found in plant tissue, which is likely to represent a major component of the bulk permafrost organic carbon (Schirmer et al. 2011). Accordingly, this cellulase family has been suggested to be the most abundant type of cellulases in arctic soil (Tveit et al. 2013). Enzymes belonging to GH5 are found widely distributed in archaea, bacteria, fungi, and plants (Davies and Attia 2018). This wide taxonomic range was also seen in our results, as organisms hosting PF00150 included alphaproteobacterial orders Rhodospirillales and Rhizobiales, but also fungal taxon Mucoromycota had a high contribution to GH5 abundance, as well as viruses.

For the mature ponds, we also compared the carbon degradation potential between the surface waters and the deeper anoxic layer. Overall, the potential for carbon degradation was higher in the anoxic water masses. One option for the higher potential could be input of dead cell material and their DNA from the upper layers of the ponds. However, the taxonomic analysis of the microbial community in the bottom water of the ponds showed taxa typical for anoxic environments (Wurzbacher et al. 2017) suggesting that the impact of possible sinking organisms from upper layers is minor and likely

not affect the functional analysis. Other options for the higher degradation potential in the bottom water include the impact of higher total carbon concentration in the bottom layer (Supporting Information Fig. S5), or the fact that the prevailing conditions at the bottom with low oxygen concentration and low irradiance are limiting the autotrophic processes and together with the abundance of carbon substrates, the environment is more favorable for heterotrophic metabolism. We do not have data on differences in organic carbon quality between the upper and lower water masses, but in the literature anoxia and increasing depth have been linked with increased levels of more recalcitrant carbon compounds (Sachse et al. 2001; Jiao et al. 2010). $SUVA_{254}$ and S_{289} values from thaw ponds located 40 km east of the sites we studied have also revealed that carbon in the hypolimnion is highly chromophoric and depleted of fulvic and humic acids from autochthonous production (Roiha et al. 2015).

For PFAMs, there was no clear trend toward higher potential for degradation of more recalcitrant compounds in the anoxic layer, but rather a more general difference between the compartments, where the same functional capacity was present in both layers but mediated by different enzyme groups. For example, while both compartments had several glycoside hydrolase families, those were clearly different in their taxonomic affiliations between oxic and anoxic environments. Nevertheless, certain genes related to degradation of more recalcitrant compounds were more prevalent in the bottom layer. An example of this is laccase which was highly present in the anoxic layer and related to, for example, Candidate Phylum Omnitrophica. This is a fairly recently established bacterial phylum for which functional information is still scarce. Interestingly, members of this phylum have been suggested to carry potential for methane oxidation (Momper et al. 2017), nitrogen fixation, and magnetotactic ability (Kolinko et al. 2015). Here, we add to this list a capacity for organic carbon degradation.

Another enzyme worth mentioning is trehalase that was also more abundant in the younger ponds. Trehalose is a compound that is used by organisms to survive in subzero conditions, but it may also serve as a source of energy and carbon (Strom and Kaasen 1993; Elbein et al. 2003). In arctic soil, trehalase has been shown to be more abundant in the active layer than in the permafrost (Yergeau et al. 2010), suggesting use of trehalose as a substrate as the environments warms up. In our study, the highest potential for trehalose degradation was found from the early stages of the permafrost thaw suggesting that organisms in permafrost may have depended on this cryopreservant to survive the extended frozen conditions.

Conclusions

Our study shows that both the DOC quality and the microbial potential for carbon cycling shift over time as the permafrost thaws and the ponds develop from small emerging puddles to deep and stratified mature ponds. We show that the

carbon in the early stages of pond formation has a strong terrestrial signal and that while the majority of the degradation potential is in the bacterial community, also eukaryotes and viruses have a role in gauging the permafrost carbon to the active food web. Our results highlight the dynamic nature of the thaw ponds, which is reflected in both abiotic and biotic characteristics of these ponds. More studies are needed in order to show how many of the observed processes are in fact active in these ponds, as well as more detailed characterization of the carbon compounds found in these waterbodies. However, our results are an important first step in disentangling the relationship between carbon quality and microbial processes.

References

- An, W., and M. Allard. 1995. A mathematical approach to modelling palsa formation: Insights on processes and growth conditions. *Cold Reg. Sci. Technol.* **23**: 231–244. doi:10.1016/0165-232X(94)00015-P
- Anders, S., P. T. Pyl, and W. Huber. 2015. HTSeq—a python framework to work with high-throughput sequencing data. *Bioinformatics* **31**: 166–169. doi:10.1093/bioinformatics/btu638
- Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B* **57**: 289–300. <http://www.jstor.org/stable/2346101>
- Benjamini, Y., and D. Yekutieli. 2001. The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* **29**: 1165–1188. doi:10.1007/s11749-008-0129-3
- Bhiry, N., and others. 2011. 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
- Bushnell, B. 2016. BBMap short read aligner. [accessed 2019 January 5]. Available from <https://www.osti.gov/servlets/purl/1241166>
- Catalan, N., and others. 2016. Effects of beaver impoundments on dissolved organic matter quality and biodegradability in boreal riverine systems. *Hydrobiologia* **793**: 135–148. doi:10.1007/s10750-016-2766-y
- Cory, R. M., B. C. Crump, J. A. Dobkowski, and G. W. Kling. 2013. Surface exposure to sunlight stimulates CO₂ release from permafrost soil carbon in the Arctic. *Proc. Natl. Acad. Sci. USA* **110**: 3429–3434. doi:10.1073/pnas.1214104110
- Crevecoeur, S., W. F. Vincent, J. Comte, and C. Lovejoy. 2015. Bacterial community structure across environmental gradients in permafrost thaw ponds: Methanotroph-rich ecosystems. *Front. Microbiol.* **6**: 15. doi:10.3389/fmicb.2015.00192
- Crevecoeur, S., W. F. Vincent, J. Comte, A. Matveev, and C. Lovejoy. 2017. Diversity and potential activity of methanotrophs in high methane-emitting permafrost thaw ponds. *PLoS One* **12**: e0188223. doi:10.1371/journal.pone.0188223
- Davies, G., and M. Attia. 2018. Glycoside hydrolase family 5. [accessed 2019 January 5]. Available from <http://www.cazy.org/GH5.html>
- Dorrepaal, E., S. Toet, R. S. P. van Logtestijn, E. Swart, M. J. van de Weg, T. V. Callaghan, and R. Aerts. 2009. Carbon respiration from subsurface peat accelerated by climate warming in the subarctic. *Nature* **460**: 616–619. doi:10.1038/nature08216
- Durbin, R., S. Eddy, A. Krogh, and G. Mitchison. 2002. *Biological sequence analysis: Probabilistic models of proteins and nucleic acids*. Cambridge Univ. Press. doi: 10.1110/ps.8.3.695
- Elbein, A. D., Y. T. Pan, I. Pastuszak, and D. Carroll. 2003. New insights on trehalose: A multifunctional molecule. *Glycobiology* **13**: 17R–27R. doi:10.1093/glycob/cwg047
- El-Gebali, S., and others. 2019. The Pfam protein families database in 2019. *Nucleic Acids Res.* **47**: D427–D432. doi:10.1093/nar/gky995
- Emerson, J. B., and others. 2018. Host-linked soil viral ecology along a permafrost thaw gradient. *Nat. Microbiol.* **3**: 870–880. doi:10.1038/s41564-018-0190-y
- Fernandez, L., S. Bertilsson, and S. Peura. 2019. Non-cyanobacterial diazotrophs dominate nitrogen-fixing communities in permafrost thaw ponds. *Limnol. Oceanogr.* doi: 10.1002/lno.11243
- Fillion, M.-E. v., N. Bhiry, and M. Touazi. 2014. Differential development of two Palsa fields in a peatland located near Whapmagoostui-Kuujuarapik, Northern Québec, Canada. *Arct. Antarct. Alp. Res.* **46**: 40–54. doi:10.1657/1938-4246-46.1.40
- Finn, R. D., and others. 2008. The Pfam protein families database. *Nucleic Acids Res.* **36**: D281–D288. doi:10.1093/nar/gkm960
- Fox, J., R. Muenchen, and D. Putler. 2018. R commander miscellaneous functions. R package version 1.0-7. [accessed 2019 January 5]. Available from <https://CRAN.R-project.org/package=RcmdrMisc>
- Fuchs, G., M. Boll, and J. Heider. 2011. Microbial degradation of aromatic compounds—from one strategy to four. *Nat. Rev. Microbiol.* **9**: 803–816. doi:10.1038/nrmicro2652
- Gao, L., Z. Zhou, A. V. Reyes, and L. Guo. 2018. Yields and characterization of dissolved organic matter from different aged soils in northern Alaska. *J. Geophys. Res. Biogeosci.* **123**: 2035–2052. doi:10.1029/2018JG004408
- Helms, J. R., A. Stubbins, J. D. Ritchie, E. C. Minor, D. J. Kieber, and K. Mopper. 2008. Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photo-bleaching of chromophoric dissolved organic matter. *Limnol. Oceanogr.* **53**: 955–969. doi:10.4319/lo.2008.53.3.0955
- Holm, S. 1979. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* **6**: 65–70. <https://www.jstor.org/stable/4615733>
- Hugouvieux-Cotte-Pattat, N., G. Condemine, and V. E. Shevchik. 2014. Bacterial pectate lyases, structural and

- functional diversity. *Environ. Microbiol. Rep.* **6**: 427–440. doi:10.1111/1758-2229.12166
- Hyatt, D., G.-L. Chen, P. F. LoCascio, M. L. Land, F. W. Larimer, and L. J. Hauser. 2010. Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**: 119. doi:10.1186/1471-2105-11-119
- Ivanova, A. A., C.-E. Wegner, Y. Kim, W. Liesack, and S. N. Dedysh. 2016. Identification of microbial populations driving biopolymer degradation in acidic peatlands by meta-transcriptomic analysis. *Mol. Ecol.* **25**: 4818–4835. doi:10.1111/mec.13806
- Jiao, N., and others. 2010. Microbial production of recalcitrant dissolved organic matter: Long-term carbon storage in the global ocean. *Nat. Rev. Microbiol.* **8**: 593–599. doi:10.1038/nrmicro2386
- Jørgensen, L., C. A. Stedmon, T. Kragh, S. Markager, M. Middelboe, and M. Søndergaard. 2011. Global trends in the fluorescence characteristics and distribution of marine dissolved organic matter. *Mar. Chem.* **126**: 139–148. doi:10.1016/j.marchem.2011.05.002
- Joshi, N. A., and J. N. Fass. 2011. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files. [accessed 2018 November 15]. Available from <https://github.com/najoshi/sickle>
- Julies, E. M., B. M. Fuchs, C. Arnosti, and V. Bruchert. 2010. Organic carbon degradation in anoxic organic-rich shelf sediments: Biogeochemical rates and microbial abundance. *Geomicrobiol. J.* **27**: 303–314. doi:10.1080/01490450.903410464
- Kankaala, P., S. Taipale, H. Nykanen, and R. I. Jones. 2007. Oxidation, efflux, and isotopic fractionation of methane during autumnal turnover in a polyhumic, boreal lake. *J. Geophys. Res. Biogeosci.* **112**: 7. doi:10.1029/2006jg000336
- Koch, J. C., K. Gurney, and M. S. Wipfli. 2014. Morphology-dependent water budgets and nutrient fluxes in Arctic thaw ponds. *Permafr. Periglac. Process.* **25**: 79–93. doi:10.1002/ppp.1804
- Koehler, C., A. N. Campbell, C. Pepe-Rannek, and D. H. Buckley. 2019. Assessing fungal contributions to cellulose degradation in soil by using high-throughput stable isotope probing. *Soil Biol. Biochem.* **130**: 150–158. doi:10.1016/j.soilbio.2018.12.013
- Kolinko, S., M. Richter, F.-O. Glöckner, A. Brachmann, and D. Schüler. 2015. Single-cell genomics of uncultivated deep-branching magnetotactic bacteria reveals a conserved set of magnetosome genes. *Environ. Microbiol.* **18**: 21–37. doi:10.1111/1462-2920.12907
- Lévesque, A. V., W. F. Vincent, J. Comte, C. Lovejoy, and A. I. Culley. 2018. Chlorovirus and myovirus diversity in permafrost thaw ponds. *Aquat. Microb. Ecol.* **82**: 209–224. doi:10.3354/ame01893
- Li, D., C. M. Liu, R. Luo, K. Sadakane, and T. W. Lam. 2015. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**: 1674–1676. doi:10.1093/bioinformatics/btv033
- Loiselle, S. A., L. Bracchini, A. Cózar, A. M. Dattilo, A. Tognazzi, and C. Rossi. 2009. Variability in photobleaching yields and their related impacts on optical conditions in subtropical lakes. *J. Photochem. Photobiol. B* **95**: 129–137. doi:10.1016/j.jphotobiol.2009.02.002
- Love, M. I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**: 550. doi:10.1186/s13059-014-0550-8
- Luoto, M., and M. Seppälä. 2003. Thermokarst ponds as indicators of the former distribution of palsas in Finnish Lapland. *Permafr. Periglac. Process.* **14**: 19–27. doi:10.1002/ppp.441
- Mackelprang, R., M. P. Waldrop, K. M. DeAngelis, M. M. David, K. L. Chavarria, S. J. Blazewicz, E. M. Rubin, and J. K. Jansson. 2011. Metagenomic analysis of a permafrost microbial community reveals a rapid response to thaw. *Nature* **480**: 368–371. doi:10.1038/nature10576
- Mangiafico, S. S. 2015. An R companion for the handbook of biological statistics. Rutgers Cooperative Extension, version 1.3.2. [accessed 2019 January 5]. Available from <http://rcompanion.org/rcompanion/>
- McKnight, D. M., E. W. Boyer, P. K. Westerhoff, P. T. Doran, T. Kulbe, and D. T. Andersen. 2001. Spectrofluorometric characterization of dissolved organic matter for indication of precursor organic material and aromaticity. *Limnol. Oceanogr.* **46**: 38–48. doi:10.4319/lo.2001.46.1.0038
- Menzel, P., K. L. Ng, and A. Krogh. 2016. Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat. Commun.* **7**: 11257. doi:10.1038/ncomms11257
- Momper, L., S. P. Jungbluth, M. D. Lee, and J. P. Amend. 2017. Energy and carbon metabolisms in a deep terrestrial subsurface fluid microbial community. *ISME J.* **11**: 2319–2333. doi:10.1038/ismej.2017.94
- Monteux, S., and others. 2018. Long-term in situ permafrost thaw effects on bacterial communities and potential aerobic respiration. *ISME J.* **12**: 2129–2141. doi:10.1038/s41396-018-0176-z
- Murphy, K. R., K. D. Butler, R. G. Spencer, C. A. Stedmon, J. R. Boehme, and G. R. Aiken. 2010. Measurement of dissolved organic matter fluorescence in aquatic environments: An interlaboratory comparison. *Environ. Sci. Technol.* **44**: 9405–9412. doi:10.1021/es102362t
- Murphy, K. R., C. A. Stedmon, D. Graeber, and R. Bro. 2013. Fluorescence spectroscopy and multi-way techniques. *PARAFAC. Anal. Methods* **5**: 6557–6566. doi:10.1039/C3AY41160E
- Murphy, K. R., C. A. Stedmon, P. Wenig, and R. Bro. 2014. OpenFluor—an online spectral library of auto-fluorescence by organic compounds in the environment. *Anal. Methods* **6**: 658–661. doi:10.1039/C3AY41935E
- Oksanen, J., and others. 2017. Vegan: Community ecology package. R package version 2.4-4. [accessed 2019 January 5].

- Available from <https://CRAN.R-project.org/package=vegan>
- Osterkamp, T. E., L. Viereck, Y. Shur, M. T. Jorgenson, C. Racine, A. Doyle, and R. D. Boone. 2000. Observations of thermokarst and its impact on boreal forests in Alaska, U.S. *A. Arct. Antarct. Alp. Res.* **32**: 303–315. doi:10.1080/15230430.2000.12003368
- Pienitz, R., P. T. Doran, and S. F. Lamoureux. 2008. Origin and geomorphology of lakes in the polar region. p. 25–41. *In* W. F. Vincent and J. Laybourn-Parry [eds.], *Polar lakes and rivers: Limnology of arctic and antarctic aquatic ecosystems*. Oxford Univ. Press. doi:10.1093/acprof:oso/9780199213887.001.0001
- Poulin, B. A., J. N. Ryan, and G. R. Aiken. 2014. Effects of iron on optical properties of dissolved organic matter. *Environ. Sci. Technol.* **48**: 10098–10106. doi:10.1021/es502670r
- R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing.
- Roehm, C. L., R. Giesler, and J. Karlsson. 2009. Bioavailability of terrestrial organic carbon to lake bacteria: The case of a degrading subarctic permafrost mire complex. *J. Geophys. Res.* **114**: G03006. doi:10.1029/2008JG000863
- Roiha, T., I. Laurion, and M. Rautio. 2015. Carbon dynamics in highly heterotrophic subarctic thaw ponds. *Biogeosciences* **12**: 7223–7237. doi:10.5194/bg-12-7223-2015
- Rossi, P. G., I. Laurion, and C. Lovejoy. 2013. Distribution and identity of bacteria in subarctic permafrost thaw ponds. *Aquat. Microb. Ecol.* **69**: 231–245. doi:10.3354/ame01634
- Sachse, A., D. Babenzien, G. Ginzler, J. Gelbrecht, and C. E. W. Steinberg. 2001. Characterization of dissolved organic carbon (DOC) in a dystrophic lake and an adjacent fen. *Biogeochemistry* **54**: 279–296. doi:10.1023/A:1010649227510
- Schirmer, L., G. Grosse, S. Wetterich, P. P. Overduin, J. Strauss, E. A. G. Schuur, and H. W. Hubberten. 2011. Fossil organic matter characteristics in permafrost deposits of the northeast Siberian Arctic. *J. Geophys. Res. Biogeosci.* **116**: G00M02. doi:10.1029/2011JG001647
- Schuur, E. A. G., and others. 2008. Vulnerability of permafrost carbon to climate change: Implications for the global carbon cycle. *BioScience* **58**: 701–714. doi:10.1641/B580807
- Seppälä, M. 2006. Palsa mires in Finland. *Finnish Environ.* **23**: 155–162.
- Solden, L. M., and others. 2017. New roles in hemicellulosic sugar fermentation for the uncultivated Bacteroidetes family BS11. *ISME J.* **11**: 691–703. doi:10.1038/ismej.2016.150
- Sonnhammer, E. L. L., S. R. Eddy, E. Birney, A. Bateman, and R. Durbin. 1998. Pfam: Multiple sequence alignments and HMM-profiles of protein domains. *Nucleic Acids Res.* **26**: 320–322. doi:10.1093/nar/26.1.320
- Stieglitz, M., S. J. Déry, V. E. Romanovsky, and T. E. Osterkamp. 2003. The role of snow cover in the warming of arctic permafrost. *Geophys. Res. Lett.* **30**: 1721. doi:10.1029/2003GL017337
- Strom, A. R., and I. Kaasen. 1993. Trehalose metabolism in *Escherichia coli*: Stress protection and stress regulation of gene expression. *Mol. Microbiol.* **8**: 205–210. doi:10.1111/j.1365-2958.1993.tb01564.x
- Thormann, M. N. 2006. Diversity and function of fungi in peatlands: A carbon cycling perspective. *Can. J. Soil Sci.* **86**: 281–293. doi:10.4141/S05-082
- Tranvik, L. J., and others. 2009. Lakes and reservoirs as regulators of carbon cycling and climate. *Limnol. Oceanogr.* **54**: 2298–2314. doi:10.4319/lo.2009.54.6_part_2.2298
- Trubl, G., and others. 2018. Soil viruses are underexplored players in ecosystem carbon processing. *mSystems* **3**: e00018–e00076. doi:10.1128/mSystems.00076-18
- Tveit, A., R. Schwacke, M. M. Svenning, and T. Urich. 2013. Organic carbon transformations in high-Arctic peat soils: Key functions and microorganisms. *ISME J.* **7**: 299–311. doi:10.1038/ismej.2012.99
- Vincent, W. F., M. Lemay, M. Allard, and B. B. Wolfe. 2013. Adapting to permafrost change: A science framework. *Eos Trans. AGU* **94**: 373–375. doi:10.1002/2013EO420002
- Vincent, W. F., M. Lemay, and M. Allard. 2017. Arctic permafrost landscapes in transition: Towards an integrated earth system approach. *Arct. Sci.* **3**: 39–64. doi:10.1139/as-2016-0027
- Vonk, J. E., and others. 2013. High biolability of ancient permafrost carbon upon thaw. *Geophys. Res. Lett.* **40**: 2689–2693. doi:10.1002/grl.50348
- Vonk, J. E., and others. 2015. Reviews and syntheses: Effects of permafrost thaw on Arctic aquatic ecosystems. *Biogeosciences* **12**: 7129–7167. doi:10.5194/bg-12-7129-2015
- Waldrop, M. P., K. P. Wickland, R. White III, A. A. Berhe, J. W. Harden, and V. E. Romanovsky. 2010. Molecular investigations into a globally important carbon pool: Permafrost-protected carbon in Alaskan soils. *Glob. Chang. Biol.* **16**: 2543–2554. doi:10.1111/j.1365-2486.2009.02141.x
- Walter, K. M., L. C. Smith, and F. S. Chapin. 2007. Methane bubbling from northern lakes: Present and future contributions to the global methane budget. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **365**: 1657–1676. doi:10.1098/rsta.2007.2036
- Watanabe, S., I. Laurion, K. Chokmani, R. Pienitz, and W. F. Vincent. 2011. Optical diversity of thaw ponds in discontinuous permafrost: A model system for water color analysis. *J. Geophys. Res.* **116**: G02003. doi:10.1029/2010JG001380
- Wauthy, M., M. Rautio, K. S. Christoffersen, L. Forsström, I. Laurion, H. L. Mariash, S. Peura, and W. F. Vincent. 2018. Increasing dominance of terrigenous organic matter in circumpolar freshwaters due to permafrost thaw. *Limnol. Oceanogr. Lett.* **3**: 186–198. doi:10.1002/lo12.10063
- Weishaar, J. L., G. R. Aiken, B. A. Bergamaschi, M. S. Fram, R. Fujii, and K. Mopper. 2003. Evaluation of specific ultraviolet absorbance as an indicator of the chemical composition and reactivity of dissolved organic carbon. *Environ. Sci. Technol.* **37**: 4702–4708. doi:10.1021/es030360x

- Wik, M., R. K. Varner, K. W. Anthony, S. MacIntyre, and D. Bastviken. 2016. Climate-sensitive northern lakes and ponds are critical components of methane release. *Nat. Geosci.* **9**: 99–105. doi:[10.1038/ngeo2578](https://doi.org/10.1038/ngeo2578)
- Williamson, C. E., D. P. Morris, M. L. Pace, and O. G. Olson. 1999. Dissolved organic carbon and nutrients as regulators of lake ecosystems: Resurrection of a more integrated paradigm. *Limnol. Oceanogr.* **44**: 795–803. doi:[10.4319/lo.1999.44.3_part_2.0795](https://doi.org/10.4319/lo.1999.44.3_part_2.0795)
- Woodcroft, B. J., and others. 2018. Genome-centric view of carbon processing in thawing permafrost. *Nature* **560**: 49–54. doi:[10.1038/s41586-018-0338-1](https://doi.org/10.1038/s41586-018-0338-1)
- Worden, A. Z., M. J. Follows, S. J. Giovannoni, S. Wilken, A. E. Zimmerman, and P. J. Keeling. 2015. Rethinking the marine carbon cycle: Factoring in the multifarious lifestyles of microbes. *Science* **347**: 1257594. doi:[10.1126/science.1257594](https://doi.org/10.1126/science.1257594)
- Wurzbacher, C., R. H. Nilsson, M. Rautio, and S. Peura. 2017. Poorly known microbial taxa dominate the microbiome of permafrost thaw ponds. *ISME J.* **11**: 1938–1941. doi:[10.1038/ismej.2017.54](https://doi.org/10.1038/ismej.2017.54)
- Xue, K., and others. 2016. Tundra soil carbon is vulnerable to rapid microbial decomposition under climate warming. *Nat. Clim. Chang.* **6**: 595–600. doi:[10.1038/NCLIMATE2940](https://doi.org/10.1038/NCLIMATE2940)
- Yergeau, E., H. Hogues, L. G. Whyte, and C. W. Greer. 2010. The functional potential of high Arctic permafrost revealed by metagenomic sequencing, qPCR and microarray analyses. *ISME J.* **4**: 1206–1214. doi:[10.1038/ismej.2010.41](https://doi.org/10.1038/ismej.2010.41)
- Zhuang, Q., and others. 2006. CO₂ and CH₄ exchanges between land ecosystems and the atmosphere in northern high latitudes over the 21st century. *Geophys. Res. Lett.* **33**: L17403. doi:[10.1029/2006GL026972](https://doi.org/10.1029/2006GL026972)
- Zimov, S., S. Davydov, G. Zimova, A. Davydova, E. Schuur, K. Dutta, and F. S. Chapin III. 2006a. Permafrost carbon: Stock and decomposability of a globally significant carbon pool. *Geophys. Res. Lett.* **33**: L20502. doi:[10.1029/2006GL027484](https://doi.org/10.1029/2006GL027484)
- Zimov, S. A., E. A. G. Schuur, and F. S. Chapin III. 2006b. Permafrost and the global carbon budget. *Science* **312**: 1612–1613. doi:[10.1126/science.1128908](https://doi.org/10.1126/science.1128908)
- Zoltai, S. 1993. Cyclic development of permafrost in the peatlands of northwestern Alberta, Canada. *Arct. Alp. Res.* **25**: 240–246. doi:[10.1080/00040851.1993.12003011](https://doi.org/10.1080/00040851.1993.12003011)

Acknowledgments

We thank the CEN Whapmagoostui-Kuujuarapik Station and Claude Tremblay for the use of laboratory space and transportation to the thermokarstic site. David-Alexandre Gauthier, Félix Faucher, and Julien Lebrun are acknowledged for their help with the sampling and Pilar Lopez-Hernandez for her help with the molecular laboratory work. The funding was provided by the Academy of Finland (grant 265902 to SP), the International Network for Terrestrial Research and Monitoring in the Arctic (INTERACT; grant to SP), the Swedish Research Council (grants to SB and AE) and the Knut and Alice Wallenberg foundation via the KAWater project awarded to Lars Tranvik. Financial support was further provided by the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canada Research Chair program, and the Canada Foundation for Innovation funds. The computations were performed on resources provided by SNIC through Uppsala Multi-disciplinary Center for Advanced Computational Science (UPPMAX) under project snic2018-3-22. For sequencing, we acknowledge Science for Life Laboratory.

Conflict of Interest

None declared.

Submitted 28 January 2019

Revised 19 June 2019

Accepted 10 September 2019

Associate editor: Ramon Massana Molera