



EDITOR'S
CHOICE

Microbial biogeography of permafrost thaw ponds across the changing northern landscape

Jérôme Comte, Adam Monier, Sophie Crevecoeur, Connie Lovejoy and Warwick F. Vincent

J. Comte (jerome.comte@takuvik.ulaval.ca), S. Crevecoeur, C. Lovejoy and W. F. Vincent, Centre d'études nordiques (CEN), Univ. Laval, Québec, QC G1V 0A6, Canada. – A. Monier, JC, SC, CL and WFV, Takuvik Joint International Laboratory, Centre National de la Recherche Scientifique (France, CNRS UMI 3376) and Dépt de Biologie, Univ. Laval, Québec, QC G1V 0A6, Canada. JC, AM, SC and CL also at: Inst. de Biologie Intégrative et des Systèmes (IBIS), Univ. Laval, Québec, QC G1V 0A6, Canada. AM and CL also at: Québec Océan, Univ. Laval, Québec, QC G1V 0A6, Canada. AM also at: Biosciences, Univ. of Exeter, Exeter, EX4 4QD, UK.

Microbial diversity varies at multiple spatial scales, but little is known about how climate change may influence this variation. Here we assessed the free-living bacterioplankton composition of thaw ponds over a north-south gradient of permafrost degradation in the eastern Canadian subarctic. Three nested spatial scales were compared: 1) among ponds within individual valleys 2) between two valleys within each landscape type, and 3) between landscape types (southern sporadic versus northern discontinuous permafrost). As a reference point, we sampled rock-basin lakes whose formation was not related to permafrost thawing. β -diversity was low at the smallest scale despite marked differences in limnological properties among neighboring ponds. β -diversity was high among valleys, associated with greater environmental heterogeneity. The largest differences were between landscape types and appeared to reflect the concomitant effects of environmental filtering and dispersal limitation. Raup–Crick β -diversity indicated that community assembly was driven by both stochastic (random extinction, dispersal, ecological drift) and deterministic (environmental filtering) processes. Communities sampled in the most degraded valley appeared primarily assembled through stochastic processes, while environmental filtering played a greater role at the other valleys. These results imply that climate warming and ongoing permafrost degradation will influence microbial community assembly, which in turn is likely to affect the functioning of thaw pond ecosystems.

Microorganisms display biogeographic patterns across a broad range of spatial scales, from < 20 m (Lear et al. 2014) and hundreds of m (Langenheder and Ragnarsson 2007) to regional (Souffreau et al. 2014), global (Fuhrman et al. 2008) and or multiple scales (Martiny et al. 2011), but the processes underlying such patterns are still poorly understood. Community β -diversity is shaped by the interplay between deterministic non-random, niche-based mechanisms, and stochastic processes (random extinction and colonization, ecological drift, priority effects) (Vellend 2010, Chase and Myers 2011). Spatial dynamics within a metacommunity, defined as a set of local communities linked by dispersal, are influenced by the local environment, dispersal and stochasticity (Leibold et al. 2004), and over the last decade the relative influence and interactions of these effects have been investigated (Hanson et al. 2012, Lindström and Langenheder 2012). However, the majority of these studies have focused on community assembly processes at relatively small scales, and the role of large-scale forces in shaping microbial metacommunity structure has received little attention.

Climate change is causing large-scale transformations of the Arctic landscape, with the loss of summer sea ice and rapid permafrost thawing (Vincent et al. 2011). The latter has resulted in accelerated erosion and the increased formation of 'thaw ponds' (thermokarst lakes and ponds) in

subarctic regions (Pienitz et al. 2008). Since there are vast quantities of organic carbon stored in tundra soils (Tarnocai et al. 2009), which can potentially be mobilized through aquatic microbial processes (Walter et al. 2007, Deshpande et al. 2015), these ponds may play a major role in global carbon cycles (Laurion et al. 2010). Given their widespread distribution across the northern landscape thaw ponds are excellent models to investigate the spatial structure of natural microbial communities as well as the influence of climate related changes on community assembly processes. In subarctic Quebec, Canada, permafrost thaw ponds are expanding both in size and number, and their total extent is increasing, particularly in the transition zone from northern discontinuous (underlying 50–90% of the landscape) to southern sporadic (10–50%), highly degraded permafrost landscapes (Bouchard et al. 2014). The landscape and latitudinal drivers may be tempered by local limnological conditions; for example dissolved organic carbon, suspended sediments and optical properties (Watanabe et al. 2011) are variable, even among neighboring ponds along this latitudinal gradient of permafrost degradation.

The aim of the present study was to determine patterns of bacterial community structure in thaw ponds occurring over different valleys under varied permafrost degradation states. A major goal was to identify factors controlling these

patterns over a range of spatial and environmental scales that could be applicable to other systems. We hypothesized that thaw pond communities would show greater compositional differences among communities over increasing geographical distances as a result of greater environmental variation over larger distances and dispersal limitation. Conversely, at smaller scales, local environmental filtering would drive bacterial community composition due to lower dispersal limitation among nearby ponds. We also hypothesized that permafrost thawing would influence metacommunity structure with the dominance of distinct assembly processes along the degradation gradient; i.e. environmental filtering should be stronger in the most degraded area of permafrost distribution. To test these hypotheses, we determined bacterial community structure via high-throughput sequencing of the 16S rRNA gene from ponds across the subarctic forest-tundra region of Nunavik (northern Quebec, Canada). Ponds were selected across a latitudinal gradient and a range of permafrost conditions. Five rock-basin lakes were also sampled and analyzed, as reference lakes unaffected by permafrost degradation.

Material and methods

Sampling sites and procedures

A total of 29 thermokarst ponds from four different valleys, of the eastern Canadian subarctic, were sampled

from 1 to 13 August 2012. The valleys are near the villages of Whapmagoostui-Kuujuarapik and Umiujaq, in Nunavik, Quebec. This region has experienced rapid environmental change over the last five decades, with a significant increase in annual mean air temperatures and an associated degradation of permafrost landscapes (Payette et al. 2004, Bhiry et al. 2011). Two of the valleys lie in sporadic, highly degraded permafrost landscapes: the Sasapimakwananisikw River valley (SAS; 55°13.13'N, 77°42.48'W) and the Kwakwatanikapistikw River valley (KWK; 55°19.95'N, 77°30.13'W). The two other valleys lie in the discontinuous, less degraded, permafrost zone: the Sheldrake River valley (BGR; 56°36.66'N, 76°12.93'W) and the Nastapoka River valley (NAS; 56°55.42'N, 76°22.72'W) (Fig. 1a). The SAS ponds derived from the thawing and erosion of organic permafrost mounds (palsas) whereas ponds in the 3 other valleys were derived from mineral permafrost mounds (lithalsas) (Calmels et al. 2008, Bhiry et al. 2011). These thaw ponds are typically small (10 to 100 m wide) and shallow (1 to 4 m depth) (Watanabe et al. 2011, Crevecoeur et al. 2015). In addition, a set of 5 shallow, rock-basin lakes near Whapmagoostui-Kuujuarapik were sampled (RBL, Fig. 1a) as 'reference lakes'. These waterbodies are of glacial origin and experienced the same subarctic climate conditions, but not the effects of permafrost erosion. Surface water samples for chemical and microbiological analysis were collected 1 m away from the shoreline.

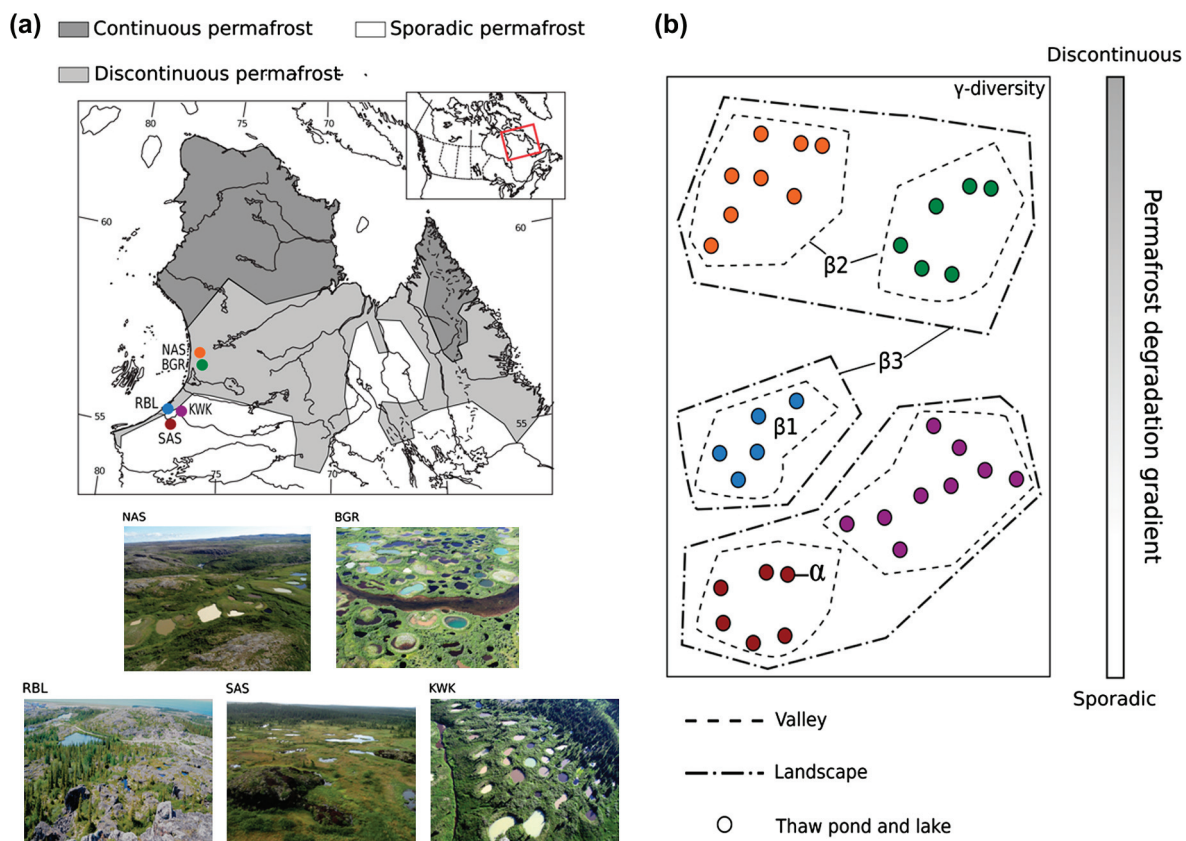


Figure 1. (a) Location and aerial photographs of the five valleys sampled in Nunavik (Québec, Canada) along a north-south permafrost degradation gradient; the base map is adapted from Fortier et al. (2011). (b) Schematic representation of the different spatial scales selected to investigate microbial β -diversity. The valley origin is indicated by color: orange, NAS; green, BGR; purple, KWK; brown, SAS; and blue, RBL.

Environmental variables

Surface water temperature, conductivity, dissolved oxygen concentration and pH were measured in situ using a multiparametric probe (Yellow Springs Instrument profiler, model 600R) and the position (latitude and longitude) of each sampling station was determined by GPS.

All chemical analyses were conducted at Inst. National de Recherche Scientifique – Eau Terre Environnement (INRS-ETE). Dissolved organic carbon (DOC) concentrations were measured using a TOC-5000A carbon analyzer (Shimadzu) calibrated with potassium biphthalate. Colored dissolved organic matter (CDOM) absorbance scans were performed from 200 to 800 nm on a Cary 100 dual beam spectrophotometer (Varian) and absorbance at 320 nm (a_{320}) was used as a proxy for water colour. The dissolved aromatic carbon content was determined using the SUVA₂₅₄ index (Weishaar et al. 2003), which is the ratio of absorbance at 254 nm to dissolved organic carbon concentration. Total phosphorus and nitrogen samples (TP and TN respectively) were fixed with H₂SO₄ (0.15% final concentration) and digested with potassium persulfate. TP was measured by spectrophotometry and TN by flow injection analysis (Lachat Instruments).

Chlorophyll-*a* (Chl-*a*) concentrations were determined from water samples filtered onto GF/F filters (Whatman). Filters were kept frozen at -80°C until pigments were analysed by high performance liquid chromatography using a ProStar HPLC system (Varian, Palo Alto, CA, USA) following procedure described in (Bonilla et al. 2005).

Bacterial community composition

Samples processing

For microbial DNA collection, water was sequentially filtered through a 20 μm mesh net to remove larger organisms, a 47-mm diameter 3 μm pore size polycarbonate filter (Whatman Nuclepore, USA) and a 0.2 μm pore size Sterivex unit (EMD Millipore, Billerica, MA, USA) using a peristaltic pump. All filters were immersed in 1.8 ml of RNAlater (Life Technologies, USA) to stabilize nucleic acids and then frozen at -80°C until further processing. For this study we analysed the 0.2 to 3 μm fraction retained in the Sterivex units. Extractions were performed using the PowerWater Sterivex DNA Isolation Kit (MO BIO Laboratories, USA) following the manufacturer instructions.

Community composition was investigated by amplifying the V6–V8 regions of 16S rRNA gene using the forward primer 969F and reverse primer 1406R and with sample-specific tags; primer details are given in Comeau et al. (2011). PCR reactions were performed in three independent 20 μl reaction volumes comprising 0.4 U Phusion high-fidelity DNA polymerase (New England Biolabs, USA), 1X Phusion HF reaction buffer (New England Biolabs, USA), 200 mM of each dNTP (Invitrogen, USA), 500 nM of each primer (Invitrogen), 0.4 mg ml⁻¹ BSA (New England Biolabs, USA) and finally 1 μl of the extracted DNA. Three concentrations (1, 0.5, and 0.2 \times) of template were used for each sample to avoid primer bias from a single concentration. Amplification was carried out on a C1000 Thermal Cycler (Bio-Rad Laboratories, USA) following an initial

denaturation step at 98°C for 30 s, with 25 cycles of denaturation at 95°C for 10 s, annealing at 50°C for 30 s and extension at 72°C for 30 s, and a final 7 min extension step at 72°C . Amplicons were purified using a PCR purification kit from Feldan (Canada) and were quantified spectrophotometrically (Nanodrop, ND-1000). The total PCR products were divided into 5 pooled samples (one per valley) in which equal concentration of amplicons (30 ng μl^{-1}) with different sample-specific tag were mixed. Pooled samples were sequenced using Roche/454 GS FLX Titanium technology at Plate-forme d'Analyses Génomiques, Inst. de biologie intégrative et des systèmes, Univ. Laval (QC, Canada).

Sequence processing

All sequence data were processed using QIIME ver. 1.8.0 (Caporaso et al. 2010b). Low quality and ambiguous reads were discarded, specifically those with a length outside the bounds of 300 to 500 nucleotides, or those that did not match a sample tag and primer sequence. The remaining reads were then processed through the QIIME denoiser and reverse primer sequences were removed. Denoised sequences were processed through the UPARSE pipeline (Edgar 2013). Chimeras were detected using the 16S Silva gold database (Pruesse et al. 2007), operational taxonomic units (OTUs) were clustered at a 97% sequence similarity cutoff and singletons were removed. OTU sequence representatives (the most abundant sequence within the same OTU cluster) were aligned using PyNAST (Caporaso et al. 2010a) with the Greengenes 16S core set as an alignment template (DeSantis et al. 2006). OTU representatives were taxonomically classified via the Mothur Bayesian classifier (Schloss et al. 2009) against a curated in house reference database based on the SILVA taxonomical hierarchy (ver. 108), which includes previously generated clone library sequences from northern sites (available upon request). Sequences from OTUs classified as plastid or mitochondrial 16S rRNA gene sequences were removed and were not further analyzed. The 454 sequences have been deposited in the NCBI Sequence Read Archive under accession number SRP044372.

Statistics

Statistical analyses and graphs were performed in R 3.0.3 (R Development Core Team). The matrix of OTU abundance data was Hellinger transformed (Legendre and Gallagher 2001) prior to distance-based analyses. All analyses were performed on a subsampled dataset with 4000 sequences per sample. Adjustments for multiple testing were applied using the Bonferroni correction.

Two main factors, as well as their factorial effect on composition patterns were tested by permutational MANOVA (PERMANOVA; Anderson 2001): 1) landscape type with three levels (discontinuous permafrost, sporadic permafrost and rock-basin); 2) valley identity with five levels that corresponded to the locations of the five different valleys. PERMANOVA analyses were computed using 1000 permutations.

A matrix of spatial variables was produced to evaluate the spatial relationships among ponds. This matrix was constructed based on Moran eigenvector maps (MEM)

eigenvectors derived from the geographic coordinates (longitude and latitude) of the ponds and lakes. We used the block approach of MEM variables as described in Declerck et al. (2011), and these variables were used in RDA analyses to test for their effect on bacterial composition and in the variation partitioning analyses. Spatial autocorrelation within each valley was tested using Mantel correlograms.

Environmental heterogeneity was calculated as the mean of all pairwise dissimilarities, based on the Euclidean distance matrix at each spatial scale. β -diversity was estimated at each spatial scale as the total variance in the OTU dataset following the procedure described in Legendre and De Cáceres (2013). Briefly, a Euclidean distance matrix was computed from the Hellinger transformed OTU data. The total sum of squares was then estimated and from which the total variance (β -diversity) was computed.

In order to investigate changes in community structure across spatial scales, we further computed β -diversity both in terms of species richness and Shannon diversity. A full hierarchical diversity partitioning analysis was made (Crist et al. 2003) to decompose the γ -diversity into α -diversity and components of β -diversity for each level of spatial scale (Fig. 1b). Both diversity attributes were then partitioned in the additive way proposed by Lande (1996): $\gamma = \alpha + \beta_1 + \beta_2 + \beta_3$, where γ and α refer to the total regional diversity and local communities (ponds) respectively. The different levels of β refer to the level of β -diversity at different spatial scales: β_1 = among ponds within valleys, β_2 = among valleys within landscape types, β_3 = among landscape types. The observed patterns in γ -diversity partitioning were compared with hierarchical null models built with 10 000 randomizations per level in order to test whether the observed components of diversity could have been obtained by the random distribution of bacterial taxa among samples at all hierarchical levels. Diversity partitions were computed using `adipart` implemented in the `vegan` R package (Oksanen et al. 2013). Relationships between β -diversity environmental heterogeneity and trophic status (indicated by TP concentrations) were analyzed by least square regression models.

Variable clustering analysis was applied to assess the redundancy of the environmental variables. Only environmental variables with Spearman correlation Rho values lower than 0.60 were selected. CDOM was therefore removed from the matrix of environmental variables because of a high correlation with DOC (Spearman Rho = 0.87). The best subset of environmental variables that correlated with the compositional patterns was identified as those generating maximum rank correlations between the environmental and community distance matrices using the `BIOENV` procedure developed by Clarke and Ainsworth (1993) and implemented in `vegan` R package (Oksanen et al. 2013). Environmental variables were log-transformed with the exception of pH. Distance-based RDA analysis was performed to test the influence of environmental variables on the composition matrix.

Variance was partitioned among and within landscape types to examine the relative effect of spatial (defined by MEM variables) and environmental variables (best environmental subset) on microbial community composition. Individual fractions were tested with redundancy analyses; due to missing values, this used the data from 31 of the 34 ponds/lakes.

We addressed the question of whether permafrost degradation had effects on community assembly mechanisms using a null model approach based on the Raup–Crick metric of β -diversity (β_{RC}) following the Chase et al. (2011) algorithm. Comparison of β_{RC} among ponds provided a guide to whether deterministic processes had occurred during the assembly of the community; the results indicated whether the communities were either less similar ($0.95 < \beta_{RC} < 1$) or more similar ($-0.95 > \beta_{RC} > -1$) than expected by chance, or stochastically assembled ($-0.95 < \beta_{RC} < 0.95$). β_{RC} values were calculated for each pair of local communities after a total of 1000 iterations. To identify the dominant process affecting community assembly, the β_{RC} metric was converted to a binary number for each of the 3 possibilities: 1 when the β_{RC} was within a specific β_{RC} interval and 0 when it was outside that interval. For example, in the scenario that the communities were more similar than by chance, pairwise comparisons of ponds communities were given the value 1 when $-0.95 > \beta_{RC} > -1$, and otherwise were assigned the value 0. The proportions of each of the main ecological community assembly processes were estimated as the ratio between the sum of all positive pair-wise tests (comparisons with values equal to 1) and the total number of possible pair-wise comparisons. Differences in the proportions were tested with a Chi-squared equal proportions test of the null hypothesis that the proportions were the same. β_{RC} patterns were visualized with multidimensional scaling (MDS) and compositional heterogeneity across spatial scales was estimated using multivariate homogeneity of group dispersion (Anderson et al. 2006).

Results

Spatial and environmental heterogeneity

The mean geographic distance among ponds and lakes within each valley ranged from 0.1 to 6.5 km (average 0.6 km) depending on the valley (Supplementary material Appendix 1, Table A1). The mean Euclidean distances for environmental variables were highly variable, both among landscape types as well as among the different valleys (Fig. 2a). We detected no tendency for higher heterogeneity at the landscape compared to the valley scale.

Taxonomic composition and β -diversity

The vast majority of the communities inhabiting subarctic ponds were dominated by Actinobacteria and β -proteobacteria, which collectively represented up to 63% of the total number of reads (Fig. 3). However, cluster analysis based on Hellinger distances indicated that ponds and lakes from the same valley were more similar to each other than to those from other valleys. There was taxonomic variability among ponds and lakes, with Bacteroidetes and Flavobacteria being locally abundant in some valleys; these classes represented up to 9% in BGR valley. In contrast, several bacterial groups were particularly common in a single landscape type; for example, *Synechococcophycidae* (Cyanobacteria) were mainly detected in the rock-basin lakes (37% of the total reads were assigned to Cyanobacteria in these waters).

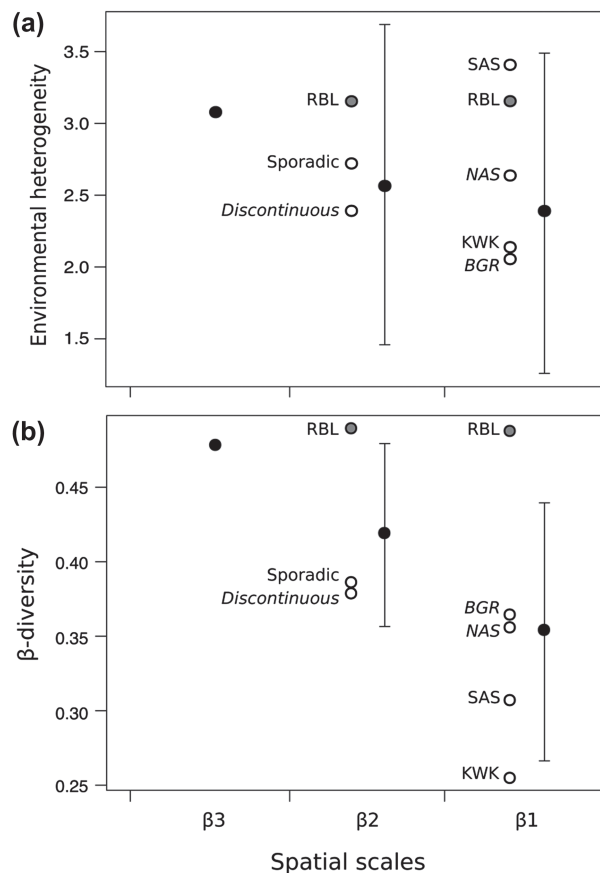


Figure 2. Average environmental heterogeneity (a) and β -diversity (b) among ponds/lakes across spatial scales (β_3 : among landscape types, β_2 : among valleys within landscape types and β_1 : among ponds within valleys). Environmental heterogeneity was calculated as the mean of all pairwise dissimilarities based on the environmental Euclidean matrix and β -diversity was computed as the total variance in OTUs Hellinger distance matrix. Valleys located in the permafrost landscape are represented by white circles. BGR and NAS valleys that lie within the discontinuous permafrost landscape are indicated in italic. SAS and KWK are located in the sporadic permafrost landscape. Reference lakes (grey circles) are located in a bedrock valley that is unaffected by permafrost degradation. The black circles represent averages and error bars denote the standard deviation.

Community patterns

The Hellinger dendrogram showed evidence of clustering of community composition according to landscape types and valleys (Fig. 3). This was further tested with PERMANOVA analysis, which showed a significant landscape effect (discontinuous, sporadic and rock-basin, $F = 6.1$, $R^2 = 0.22$, $p < 0.005$) and valley effect (RBL, BGR, NAS, KWK, SAS, $F = 4.2$, $R^2 = 0.16$, $p < 0.005$) on bacterial community composition. The results further indicated that there were significant differences in community composition between valleys within the sporadic ($F = 7.8$, $R^2 = 0.34$, $p < 0.001$) but not in the discontinuous permafrost landscapes ($F = 1.8$, $R^2 = 0.10$, $p = 0.07$).

β -diversity showed some degree of scale dependency with a tendency towards higher values at larger spatial scales (Fig. 2b). The results further showed greater variability in

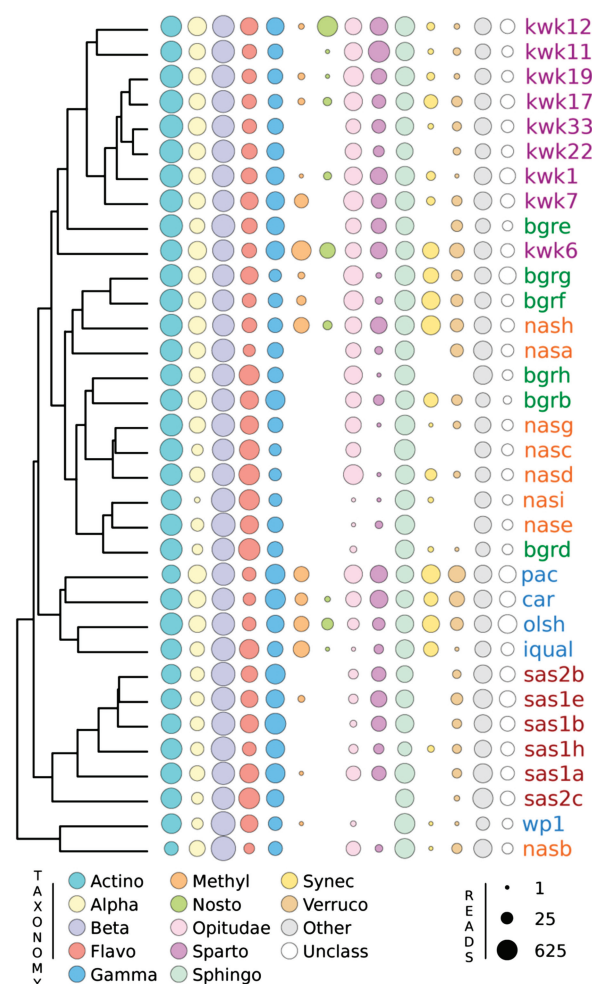


Figure 3. Clustering of the sites based on their compositional dissimilarity (Hellinger distances). The circles associated with each site show the 16S reads (from the subsampled 4000 sequences per sample dataset) classified by best matching bacterial class, with rare groups in the 'other' category ($< 3\%$ of total read abundance). The valley origin is indicated by color: orange, NAS; green, BGR; purple, KWK; brown, SAS; and blue, RBL.

community composition among sites within each valley than among waterbodies within each landscape.

Spatial versus environmental effects

At the highest spatial scale level, MEM analyses showed evidence of differences in bacterial communities among landscape types ($F = 1.79$, $R^2 = 0.14$, $p = 0.005$), whereas specific within-landscape MEM models showed no evidence for spatial relationships. Within-valleys, multivariate Mantel correlograms showed no evidence of spatial autocorrelation.

Seven environmental variables had maximum correlation with the community dissimilarity matrix: conductivity, dissolved organic carbon (DOC), total phosphorus and total nitrogen concentrations, dissolved oxygen content, $SUVA_{254}$ and Chlorophyll-*a* concentration (Chl-*a*). The RDA analysis based on this best environmental subset indicated that community composition was partially related to key

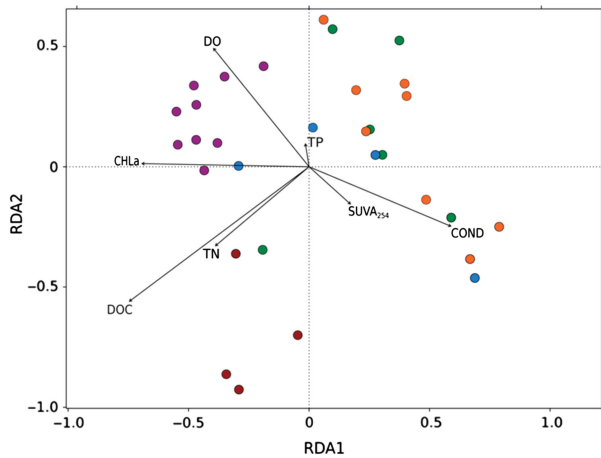


Figure 4. Distant-based redundancy analysis of the bacterial community composition of 31 thaw ponds and rock-basin lakes. Each circle represents individual waterbodies, and the color coding is for geographic location as in Fig. 1: (green, BGR; orange, NAS; purple, KWK; brown, SAS; and blue, RBL). The arrows represent the significant environmental vectors resulting from the rank correlation analysis, and the vectors correspond to total phosphorus (TP), total nitrogen (TN), dissolved oxygen (DO), dissolved organic carbon (DOC), chlorophyll-*a* (CHLa), organic matter aromaticity (SUVA₂₅₄) and water conductivity (COND). Note that some samples were excluded from analyses due to missing environmental data.

environmental variables that vary among the different valleys ($R^2=0.21$, $p=0.005$). In particular, the southern sites clusters followed Chl-*a* and DOC, whereas the northern sites clusters followed conductivity (Fig. 4). Significant relationships were detected between β -diversity and both environmental heterogeneity and trophic status (Supplementary material Appendix 1, Fig. A2).

We identified compositional variation among landscapes: variation partitioning analyses showed that environmental heterogeneity ($R^2=0.25$, $p=0.005$) and MEM variables ($R^2=0.06$, $p=0.03$) explained a significant part of the variation in bacterial community structure at the largest spatial scale. Only environmental heterogeneity significantly explained the compositional variation between valleys both in sporadic (DOC, conductivity, TP: $R^2=0.43$, $p=0.001$) and discontinuous (DOC, conductivity, dissolved oxygen: $R^2=0.21$, $p=0.03$) landscapes. There was no significant interaction effect between the MEM and environmental variables on bacterial composition.

The full hierarchical diversity partitioning analysis showed that within- and among-ponds components contributed 33% of the total species richness (α and β_1 , Fig. 5). The results revealed a substantial contribution of both the smallest ($\beta_1=23\%$) and largest spatial scales ($\beta_3=48\%$) to species richness. β -diversity between valleys within landscapes ($\beta_2=19\%$) was relatively smaller. Additive partitioning of Shannon diversity showed that α -diversity comprised 77% of the bacterial total diversity, whereas the different components of β -diversity had a lower combined contribution to γ -diversity (23%) (Fig. 5). In particular, most of the contribution from the β components of γ -diversity came from β_1 (12%) whereas the contributions from β_2 (4%) and β_3 (7%) were smaller. The β -diversity

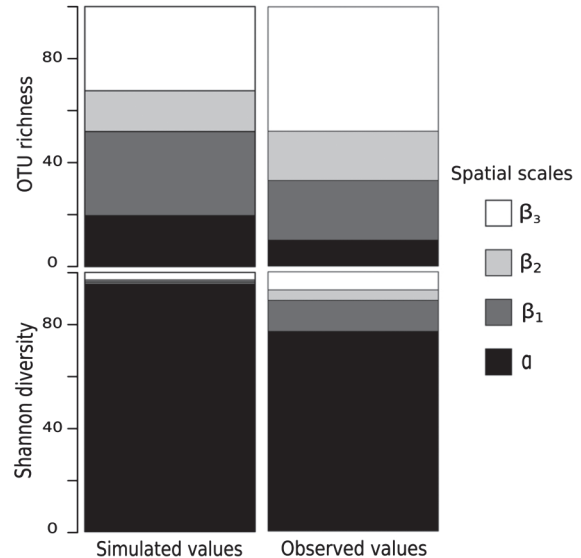


Figure 5. Species richness and Shannon diversity across four spatial scales. The values are expressed as the percent of the γ -diversity of bacterial communities explained by each hierarchical level. α : local pond diversity; β_1 : compositional variation among ponds within valleys; β_2 : compositional variation among valleys within landscapes; β_3 : compositional variation among landscapes.

components at all three spatial scales were larger than would be expected by chance.

Thawing permafrost and community assembly processes

Raup–Crick measure of β -diversity (β_{RC}) showed that ponds were more similar to each other than by chance, suggesting regional invariance across this area of the subarctic (Table 1). Thawing permafrost led to distinct communities relative to the reference rock-basin lakes, which clustered together and were isolated from the ensemble of thaw ponds (Supplementary material Appendix 1, Fig. A3a). Further, β_{RC} increased with increasing permafrost degradation state (Supplementary material Appendix 1, Fig. A3b). Ponds located in the discontinuous landscape appeared to be governed (85%) by deterministic processes ($\beta_{RC} < -0.95$, Table 1). β_{RC} among sites located in the sporadic permafrost landscape suggested that both deterministic (more similar than by chance: 44%) and stochastic (no difference from null model: 30%) processes operated. Ponds from the same valley (with the exception of SAS ponds) were more similar to each other than expected by chance indicating the prevalent role of deterministic community assembly processes. In contrast, β_{RC} among ponds from SAS showed a higher influence of stochastic events (Table 1).

Discussion

Separating the effects of stochastic and deterministic processes has received increasing attention in microbial ecology (Martiny et al. 2011), but the influence of spatial variability on these processes remains a challenging question (Barton

Table 1. Proportion of β -diversity influenced by deterministic or stochastic processes across different spatial scales. Values represent the proportion (%) of differences in bacterial community composition among sites that is related to deterministic or stochastic processes as estimated by Raup–Crick measure of β -diversity (β_{RC}). Sites are more and less similar than expected by chance with values of β_{RC} below -0.95 and above 0.95 respectively. Differences in composition are stochastic when β_{RC} estimates are close to 0. X^2 refers to the Chi-square value of test for equality of proportions. Values in bold represent the significantly dominant community assembly process (two-tailed tests) with significant thresholds indicated as: $p < 0.001^{***}$, $p < 0.01^{**}$, $p < 0.05^*$, $p > 0.05^{NS}$ after Bonferroni correction for multi-testing. † No significant difference for pairwise comparison between the proportions of the two main assembly processes.

Across spatial scales	Raup–Crick β -diversity			X^2
	> 0.95	0	< -0.95	
All sites	17	31	52	159 ^{***}
Within landscapes	20	30	50	46.8 ^{***}
Within discontinuous permafrost	0	15	85	166.4 ^{***}
Within sporadic permafrost	26	30	44†	8.3 [*]
Within BGR valley	0	0	100	45 ^{***}
Within NAS valley	0	14	86	53.1 ^{***}
Within KWK valley	0	0	100	108 ^{***}
Within SAS valley	0	87	13	29.4 ^{***}
Within RBL valley	0	30	70	11.1 ^{**}

et al. 2013). The hierarchical design of our study allowed us to distinguish the effects of spatial isolation versus local environmental filtering, and provided insight into the major drivers of community assembly. The variation in β -diversity across different spatial scales was mainly related to environmental filtering at each of the studied spatial scales, whereas the influence of spatial variables was limited and restricted to the largest spatial scale. Raup–Crick analyses of β -diversity showed that both stochastic processes and deterministic processes drive community assembly, and that these processes can operate over multiple spatial scales. The implication is that a large-scale environmental gradient, specifically climate via its effect on the degradation state of the permafrost landscape, affected the mixture of processes determining community assembly.

Environmental heterogeneity

The distance among ponds/lakes across multiple spatial scales varied over five orders of magnitude, from a few m to hundreds of km (Supplementary material Appendix 1, Table A1). Distinct bacterial assemblages inhabited different valleys and landscape types, consistent with the cluster analysis (Fig. 1) and validated by permutation tests. Although β -diversity increased at increasing spatial scales, this was only partially associated with an increase in environmental heterogeneity (Fig. 2, Supplementary material Appendix 1, Fig. A2). There was a high level of environmental heterogeneity across spatial scales. In particular, environmental heterogeneity at the smallest scale (among ponds within a valley) could be greater than among ponds and lakes within each landscape type (Fig. 2a). Ponds within valleys varied in their limnological properties such as nutrients, oxygen and dissolved organic carbon; valleys and landscapes varied further by soil

type, vegetation cover, latitude, and permafrost degradation state.

Variation partitioning revealed that environmental conditions explained a substantial portion of compositional variance across multiple spatial scales: 25% among landscapes, 43% within sporadic permafrost and 21% within discontinuous permafrost. The RDA analyses based on the best set of environmental variables measured, further showed that thaw pond communities were driven by distinct environmental variables depending on their origin. At the landscape scale, certain environmental variables best explained β -diversity depending on whether ponds were located within the sporadic (Chl-*a*) or discontinuous (conductivity) permafrost, whereas other variables (DOC and TP) were significant in both. DOC concentration appeared to be of greatest importance for the SAS ponds. In this valley, thawing and eroding palsas provided the aquatic microbial communities with a rich source of organic matter. This would be consistent with the nature and source of organic matter affecting microbial community composition (Kritzberg et al. 2006, Perez and Sommaruga 2006).

Our results point to environmental filtering as the main structuring factor for permafrost thaw pond microbial communities, consistent with reports of aquatic bacterial communities elsewhere (Hanson et al. 2012, Lindström and Langenheder 2012). However, RDA and variation partitioning analyses showed that environmental heterogeneity did not account for all of the observed β -diversity and other processes may also operate including spatial variables.

Spatial structure of thaw pond communities

The MEM models and variation partitioning showed a significant contribution of spatial variables to differences in community composition, but only between different landscape types. These results indicate that some of the environmental variables were spatially structured; for example, conductivity was higher in northern sites in comparison to southern sites. Previous studies on microbial β -diversity over large distances have documented the effect of spatially structured abiotic variables on community composition patterns (Souffreau et al. 2014). However, the results presented here did not support such a pattern. The variability in the identity of main environmental drivers among the different valleys as well as the high heterogeneity among ponds within a given valley, may explain the absence of a significant shared fraction between spatial and environmental variables. This in turn suggests that a part of the variation observed in community composition across landscape types may relate to dispersal limitation of thaw pond bacteria. Similar results have been reported for bacterial, phytoplankton and zooplankton communities elsewhere (Martiny et al. 2006, Soininen et al. 2011).

The discontinuous and sporadic permafrost landscapes differ in geomorphological characteristics, and are separated by a distance of 163 km. Exchange of living cells among landscape pools would be limited to wind dispersal and migrating animals due to the lack of direct hydrological connections. Some microbes have the potential to disperse over large distances (Hervàs et al. 2009), however microbes may fail to colonize a suitable ecological niche, due to

niche overlap with competing taxa (Soininen et al. 2007) and priority effects in recipient habitats (Chase 2007). At the smallest spatial scale (within a valley) we found no evidence of spatial autocorrelation; i.e. there was no difference in bacterial community composition with increasing distance between ponds. This in turn implies that community assembly processes related to dispersal are unlikely to have occurred. Although dispersal limitation has been reported to operate at small spatial scales (Lear et al. 2014), the short distances between ponds and the high dispersal capabilities of bacteria should prevail and overcome any dispersal limitation at this local scale. There is evidence from elsewhere, however, that dispersal rates must be extremely high to cause the homogenisation of bacterial community composition by immigration (Logue and Lindström 2010).

Community structure

Clear differences in the spatial effects of β -diversity were observed on community structure, with contrasting patterns in species richness and Shannon diversity (Fig. 5). Diversity partitioning analyses showed a greater influence of β -components in determining OTU richness than Shannon diversity. This contribution of β -components to species richness may relate to the high level of environmental heterogeneity across spatial scales, which in turn may influence the size of the species pool. The Shannon diversity partitioning suggested that the pond scale was dominated by common species; i.e. the same common OTUs comprised most of the α -diversity. Most of the ponds were dominated by the β -proteobacteria *Polynucleobacter* and *Variovorax*, along with the actinobacterium ACK-M1. Collectively these represented 46% of the total reads. This is consistent with the idea of regional invariance among dominant microbes (Östman et al. 2010), where locally abundant taxa are also widespread, in this case the freshwater bacterium *Polynucleobacter* (Hahn et al. 2015). This might indicate that these generalist taxa are preferentially selected in these aquatic ecosystems given their intrinsic ecological characters such as a wide niche breadth or competitive abilities, or that neutral processes may operate in the assembly of these bacterial communities, as reported in lakes elsewhere (Langenheder and Székely 2011).

Despite a substantial level of regional invariance among thaw pond bacterial communities, there were significant differences among valleys in taxonomic composition, and in particular, a high degree of uniqueness of the communities within the SAS valley ponds. These were rich in dissolved organic material and had low dissolved oxygen through the water column (less than 1 mg l⁻¹ of dissolved oxygen at 20 cm depth). These features may impose a strong selective environmental filter on bacterial community structure, resulting in the observed local dominance.

Community assembly and climate change

The β_{RC} analysis indicated that deterministic processes ($\beta_{RC} < -0.95$ and $\beta_{RC} > 0.95$) dominated community assembly in the permafrost thaw ponds across this study region (Table 1). In particular, the bacterial communities appeared to be mainly assembled through environmental filtering (β_{RC}

< -0.95). Previous studies have shown that environmental perturbation can promote deterministic processes that impose a niche-selection filtering from the regional species pool (Chase 2007). In these aquatic systems, environmental filters among valleys as well as local sites within a valley have resulted in communities that are more similar than expected by chance. The variation partitioning and canonical analyses also pointed to environmental control of β -diversity. This influence may be the result of bottom-up control via resource availability, as implied by the relationship between bacterial β -diversity and total phosphorus concentration, an index of trophic status (Supplementary material Appendix 1, Fig. A3). There may also be top-down trophic effects (e.g. bacterivory, viral lysis) in the patterns described above; previous studies have documented their importance in microbial community structure and assembly (Winter et al. 2013, Berga et al. 2014, Souffreau et al. 2014), but these were not assessed in the present study.

On a global scale, Hudson Bay region of the eastern Canadian subarctic is experiencing more rapid warming than most locations elsewhere in the circumpolar North (Bhiry et al. 2011, Rühland et al. 2013), and its southern permafrost margin is shifting northwards. The southern ponds lie in valleys that have lost more than 90% of its permafrost over the last few decades (Bouchard et al. 2014). The community assembly processes in these environments varied between the two valleys. The higher similarity among KWK ponds in comparison to null models, suggest a prominent role of niche-selection in the assembly of these communities. In contrast, in the SAS valley more divergent community composition were observed which may indicate the influence of stochastic processes such as ecological drift; i.e. random chance in species relative abundance. This may also indicate high heterogeneity in environmental variables among the ponds. In particular, SAS ponds are fed by rapidly thawing and eroding permafrost (Bhiry et al. 2011). It is therefore possible that variability in carbon subsidies among the ponds in combination with other environmental variables such as dissolved oxygen and bacterivores could have resulted in the dominance of stochastic processes in the assembly of bacterial communities in this valley. In the discontinuous, cooler, much less degraded permafrost landscape further northward, environmental filtering processes dominated. In combination, these results imply that climate, through its effects on permafrost landscapes, can alter community assembly processes and thereby influence microbial community structure, which in turn is likely to affect ecosystem function.

Acknowledgements – We are grateful to Alex Matveev, Bethany Deshpande, Paschale N. Bégin and Anna Przytułska-Bartosiewicz for field and laboratory work assistance and Claude Tremblay for managing the Centre d'études nordiques (CEN) research station at Whapmagoostui-Kuujuarapik. We thank INRS-EET for chemical analyses. Aerial pictures were taken by Isabelle Laurion (KWK) Alex Matveev (RBL, SAS) and Jérôme Comte (NAS, BGR). Computing support was provided by the bioinformatic service from IBIS-Univ. Laval. We acknowledge Mercè Berga for her valuable comments and inputs in previous versions of the manuscript. This study was funded by the Natural Sciences and Engineering Council (NSERC) of Canada from Discovery and Discovery Frontier (ADAPT) grants, the Network of Centres of Excellence program ArcticNet, and the Canadian Research Chair Program.

Additional support from Fonds de Recherche du Québec Nature et Technologies (FRQNT) to CEN is acknowledged. JC was partially supported by a FRQNT postdoctoral fellowships and the EnviroNorth CREATE program from NSERC.

References

- Anderson, M. J. 2001. Permutation tests for univariate or multivariate analysis of variance and regression. – *Can. J. Fish. Aquat. Sci.* 58: 626–639.
- Anderson, M. J. et al. 2006. Multivariate dispersion as a measure of beta diversity. – *Ecol. Lett.* 9: 683–693.
- Barton, P. S. et al. 2013. The spatial scaling of beta diversity. – *Global Ecol. Biogeogr.* 22: 639–647.
- Berga, M. et al. 2014. Combined effects of zooplankton grazing and dispersal on the diversity and assembly mechanisms of bacterial metacommunities. – *Environ. Microbiol.* doi: 10.1111/1462-2920.12688
- Bhiry, N. et al. 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.
- Bonilla, S. et al. 2005. Benthic and planktonic algal communities in a high arctic lake: pigment structure and contrasting responses to nutrient enrichment. – *J. Phycol.* 41: 1120–1130.
- Bouchard, F. et al. 2014. Subarctic thermokarst ponds: investigating recent landscape evolution and sediment dynamics in thawed permafrost of northern Québec (Canada). – *Arct. Antarct. Alp. Res.* 46: 251–271.
- Calmels, F. et al. 2008. Development and decay of a lithalsa in northern Québec: a geomorphological history. – *Geomorphology* 97: 287–299.
- Caporaso, J. G. et al. 2010a. PyNAST: a flexible tool for aligning sequences to a template alignment. – *Bioinformatics* 26: 266–267.
- Caporaso, J. G. et al. 2010b. QIIME allows analysis of high-throughput community sequencing data. – *Nat. Methods* 7: 335–336.
- Chase, J. M. 2007. Drought mediates the importance of stochastic community assembly. – *Proc. Natl Acad. Sci. USA* 104: 17430–17434.
- Chase, J. M. and Myers, J. A. 2011. Disentangling the importance of ecological niches from stochastic processes across scales. – *Phil. Trans. R. Soc. B* 366: 2351–2363.
- Chase, J. M. et al. 2011. Using null models to disentangle variation in community dissimilarity from variation in α -diversity. – *Ecosphere* 2: art24.
- Clarke, K. R. and Ainsworth, M. 1993. A method of linking multivariate community structure to environmental variables. – *Mar. Ecol. Prog. Ser.* 92: 205–205.
- Comeau, A. M. et al. 2011. Arctic Ocean microbial community structure before and after the 2007 record sea ice minimum. – *PLoS One* 6: e27492.
- Crevecoeur, S. et al. 2015. Bacterial community structure across environmental gradients in permafrost thaw ponds: methanotroph-rich ecosystems. – *Front. Microbiol.* 6: 192.
- Crist, T. O. et al. 2003. Partitioning species diversity across landscapes and regions: a hierarchical analysis of alpha, beta, and gamma diversity. – *Am. Nat.* 162: 734–743.
- Declerck, S. A. et al. 2011. Scale dependency of processes structuring metacommunities of cladocerans in temporary pools of High Andes wetlands. – *Ecography* 34: 296–305.
- DeSantis, T. Z. et al. 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. – *Appl. Environ. Microbiol.* 72: 5069–5072.
- Deshpande, B. N. et al. 2015. Oxygen dynamics in permafrost thaw lakes: anaerobic bioreactors in the Canadian subarctic. – *Limnol. Oceanogr.* 60: 1656–1670.
- Edgar, R. C. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. – *Nat. Methods* 10: 996–998.
- Fortier, R. et al. 2011. Impacts of permafrost degradation on a road embankment at Umiujaq in Nunavik (Quebec). – *Can. Geotech. J.* 48: 720–740.
- Fuhrman, J. A. et al. 2008. A latitudinal diversity gradient in planktonic marine bacteria. – *Proc. Natl Acad. Sci. USA* 105: 7774–7778.
- Hahn, M. W. et al. 2015. Global phylogeography of pelagic Polynucleobacter bacteria: restricted geographic distribution of subgroups, isolation by distance and influence of climate. – *Environ. Microbiol.* 17: 829–840.
- Hanson, C. A. et al. 2012. Beyond biogeographic patterns: processes shaping the microbial landscape. – *Nat. Rev. Microbiol.* 10: 497–506.
- Hervás, A. et al. 2009. Viability and potential for immigration of airborne bacteria from Africa that reach high mountain lakes in Europe. – *Environ. Microbiol.* 11: 1612–1623.
- Kritzberg, E. S. et al. 2006. Influence of dissolved organic matter source on lake bacterioplankton structure and function – implications for seasonal dynamics of community composition. – *FEMS Microbiol. Ecol.* 56: 406–417.
- Lande, R. 1996. Statistics and partitioning of species diversity, and similarity among multiple communities. – *Oikos* 76: 5–13.
- Langenheder, S. and Ragnarsson, H. 2007. The role of environmental and spatial factors for the composition of aquatic bacterial communities. – *Ecology* 88: 2154–2161.
- Langenheder, S. and Székely, A. J. 2011. Species sorting and neutral processes are both important during the initial assembly of bacterial communities. – *ISME J.* 5: 1086–1094.
- Laurion, I. et al. 2010. Variability in greenhouse gas emissions from permafrost thaw ponds. – *Limnol. Oceanogr.* 55: 115.
- Lear, G. et al. 2014. Fine-scale spatial patterns in bacterial community composition and function within freshwater ponds. – *ISME J.* 8: 1715–1726.
- Legendre, P. and Gallagher, E. D. 2001. Ecologically meaningful transformations for ordination of species data. – *Oecologia* 129: 271–280.
- Legendre, P. and De Cáceres, M. 2013. Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. – *Ecol. Lett.* 16: 951–963.
- Leibold, M. A. et al. 2004. The metacommunity concept: a framework for multi-scale community ecology. – *Ecol. Lett.* 7: 601–613.
- Lindström, E. S. and Langenheder, S. 2012. Local and regional factors influencing bacterial community assembly. – *Environ. Microbiol. Reports* 4: 1–9.
- Logue, J. B. and Lindström, E. S. 2010. Species sorting affects bacterioplankton community composition as determined by 16S rDNA and 16S rRNA fingerprints. – *ISME J.* 4: 729–738.
- Martiny, J. B. H. et al. 2006. Microbial biogeography: putting microorganisms on the map. – *Nat. Rev. Microbiol.* 4: 102–112.
- Martiny, J. B. et al. 2011. Drivers of bacterial β -diversity depend on spatial scale. – *Proc. Natl Acad. Sci. USA* 108: 7850–7854.
- Oksanen, J. et al. 2013. Package “vegan.” – *Community ecology package ver. 2.2-1.*
- Östman, Ö. et al. 2010. Regional invariance among microbial communities. – *Ecol. Lett.* 13: 118–127.
- Payette, S. et al. 2004. Accelerated thawing of subarctic peatland permafrost over the last 50 years. – *Geophys. Res. Lett.* 31: L18208.
- Perez, M. T. and Sommaruga, R. 2006. Differential effect of algal- and soil-derived dissolved organic matter on alpine lake bacterial community composition and activity. – *Limnol. Oceanogr.* 51: 2527–2537.
- Pienitz, R. et al. 2008. Origin and geomorphology of lakes in the polar regions. – In: Vincent, W. F. and Laybourn-Parry, J. (eds), *Polar lakes and rivers: limnology of Arctic and Antarctic aquatic ecosystems.* Oxford Univ. Press, pp. 25–41.

- Pruesse, E. et al. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. – *Nucl. Acids Res.* 35: 7188–7196.
- Rühland, K. M. et al. 2013. Global warming triggers the loss of a key Arctic refugium. – *Proc. R. Soc. B* 280: 20131887.
- Schloss, P. D. et al. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. – *Appl. Environ. Microbiol.* 75: 7537–7541.
- Soininen, J. et al. 2007. A multivariate analysis of beta diversity across organisms and environments. – *Ecology* 88: 2830–2838.
- Soininen, J. et al. 2011. Disentangling the spatial patterns in community composition of prokaryotic and eukaryotic lake plankton. – *Limnol. Oceanogr.* 56: 508–520.
- Souffreau, C. et al. 2014. Environmental rather than spatial factors structure bacterioplankton communities in shallow lakes along a >6,000 km latitudinal gradient in South America. – *Environ. Microbiol.* doi: 10.1111/1462-2920.12692
- Tarnocai, C. et al. 2009. Soil organic carbon pools in the northern circumpolar permafrost region. – *Global Biogeochem. Cycles* 23: GB2023.
- Vellend, M. 2010. Conceptual synthesis in community ecology. – *Q. Rev. Biol.* 85: 183–206.
- Vincent, W. F. et al. 2011. Ecological implications of changes in the Arctic cryosphere. – *Ambio* 40: 87–99.
- Walter, K. M. et al. 2007. Methane bubbling from northern lakes: present and future contributions to the global methane budget. – *Phil. Trans. R. Soc. A* 365: 1657–1676.
- Watanabe, S. et al. 2011. Optical diversity of thaw ponds in discontinuous permafrost: a model system for water color analysis. – *J. Geophys. Res.* 116: GO2003.
- Weishaar, J. L. et al. 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.
- Winter, C. et al. 2013. Effects of environmental variation and spatial distance on bacteria, archaea and viruses in sub-polar and arctic waters. – *ISME J.* 7: 1507–1518.

Supplementary material (Appendix ECOG-01667 at <www.ecography.org/appendix/ecog-01667>). Appendix 1.