

ORIGINAL ARTICLE

Distinct and diverse anaerobic bacterial communities in boreal lakes dominated by candidate division OD1

Sari Peura¹, Alexander Eiler², Stefan Bertilsson², Hannu Nykänen¹, Marja Tirola¹ and Roger I Jones¹

¹Department of Biological and Environmental Science, University of Jyväskylä, Jyväskylä, Finland and

²Department of Ecology and Genetics, Uppsala University, Uppsala, Sweden

Lakes have a central role in the carbon cycle of the boreal landscape. These systems typically stratify in summer and their hypolimnetic microbial communities influence burial of biogenic organic matter in sediments. The composition of bacterial communities in these suboxic habitats was studied by pyrosequencing of 16S rRNA amplicons from five lakes with variable dissolved organic carbon (DOC) concentrations. Bacterioplankton communities in the hypolimnetic waters were clearly different from the surface layer with candidate division OD1, *Chlorobi* and *Bacteroidetes* as dominant community members. Several operational taxonomic units (OTUs) affiliated with candidate division OD1 were abundant and consistently present in the suboxic hypolimnion in these boreal lakes. The overall representation of this group was positively correlated with DOC and methane concentrations. Network analysis of time-series data revealed contrasting temporal patterns but suggested similar ecological roles among the abundant OTUs affiliated with candidate division OD1. Together, stable isotope data and taxonomic classification point to methane oxidation and autotrophic denitrification as important processes in the suboxic zone of boreal lakes. Our data revealed that while hypolimnetic bacterial communities are less dynamic, they appear to be more diverse than communities from the oxic surface layer. An appreciable proportion of the hypolimnetic bacteria belong to poorly described phyla.

The ISME Journal (2012) 6, 1640–1652; doi:10.1038/ismej.2012.21; published online 15 March 2012

Subject Category: microbial population and community ecology

Keywords: 16S rRNA; 454 pyrosequencing; anoxic; diversity; hypolimnion; OD1

Introduction

Thousands of small forest ponds and lakes shape the boreal landscape, contributing significantly to regional and global nutrient cycling, redistribution of carbon (Battin *et al.*, 2009; Tranvik *et al.*, 2009) and greenhouse gas emissions (Bastviken *et al.*, 2011). While recent work has advanced our understanding of the composition and ecology of epilimnetic microbial communities in lakes (reviewed in Newton *et al.*, 2011), the microbial communities of suboxic hypolimnia have been much less studied, particularly in forest lakes.

Forest lakes differ from the well-studied clear-water counterparts in several key characteristics. High concentrations of humic substances give these lakes their characteristic colour, which limits light penetration, and hence also restricts phytoplankton

photosynthesis to the uppermost water column. At the same time, this chromophoric organic matter represents an energy subsidy for heterotrophic bacteria (for example, Cole *et al.*, 1994; del Giorgio and Peters, 1994; Kankaala *et al.*, 1996; Jones *et al.*, 2001; Ask *et al.*, 2009). Most boreal lakes are stratified with an oxic epilimnion and suboxic or even anoxic hypolimnion, separated by a transition layer (metalimnion). The stratification pattern may be disturbed by spring and autumn overturns, but many small boreal lakes are meromictic, with overturn only partial or totally lacking (Salonen *et al.*, 1984; Gudas *et al.*, 2010). Hypolimnetic microbial communities in these lakes may therefore evolve in the absence of oxygen over periods as long as several years. Other characteristic features of boreal lakes are low pH and low inorganic nutrient concentration in the epilimnion while nutrient availability is much higher in the hypolimnion. Furthermore, methane produced in the oxygen limited hypolimnion and sediment may supply energy to micro-organisms, not only in the oxygenated part of the lake, but in the whole water column (Taipale *et al.*, 2011).

Correspondence: S Peura, Department of Biological and Environmental Sciences, University of Jyväskylä, PO Box 35, 40014 Jyväskylä, Finland.

E-mail: sari.peura@jyu.fi

Received 7 November 2011; revised 10 February 2012; accepted 14 February 2012; published online 15 March 2012

From the major difference in the environmental conditions between epilimnion and hypolimnion, including oxygen and nutrient concentrations, it follows that their microbial communities are also different (Shade *et al.*, 2008; Barberan and Casamayor, 2011). Even though there are several studies addressing bacterial communities in humic lakes (for example, Lindström, 1998; Burkert *et al.*, 2003; Kent *et al.*, 2004), the microbial diversity and community composition of the suboxic hypolimnetic layers remain poorly explored and the few studies available are typically restricted to very specific questions, sediment communities, or are based on methods that provide little if any information on taxonomy (for example, Kleinstaub *et al.*, 2008; Lehours *et al.*, 2009; Berdjeb *et al.*, 2011; Taipale *et al.*, 2011) or diversity (Bent *et al.*, 2007). One exception is the recent study by Barberan and Casamayor (2011), which used observations from multiple lakes to demonstrate that microbial communities in the hypolimnia are distinct from those in the epilimnia. However, the lakes included in that study had either high salinity or low content of dissolved organic carbon (DOC) and it is not certain that those results can be extrapolated to the numerous lakes of the boreal zone. So far, only one study has used clone libraries and sequencing of taxonomic marker genes to describe microbial communities in the hypolimnion of boreal lakes (Taipale *et al.*, 2009). This study revealed the quantitative significance of members of the autotrophic green sulphur bacterial phylum *Chlorobi*, but was limited to a few samples and only probed the most abundant taxa as a result of limited sequencing depth.

En masse sequencing of 16S rRNA gene amplicons by 454 pyrosequencing now offers a tool for comprehensive studies of the members of bacterioplankton communities with the possibility of tracking their distribution patterns in time and space. When sequencing is coupled with other methods like analysis of stable isotopic composition of dissolved CH₄ ($\delta^{13}\text{C}_{\text{CH}_4}$) and statistical approaches such as local similarity analysis (LSA; Ruan *et al.*, 2006), valuable information concerning the ecology of individual populations present in the community may be gained. This also enables the formulation of hypotheses on metabolic capabilities and the use of specific process-related methods that require prior information on the bacterial community before they can be applied.

Here, we present an extensive analysis of bacterioplankton dynamics in humic Lake Alinen Mustajärvi. Temporal dynamics of the microbial communities were visualized in association networks and subnetworks illustrating the linkages among bacterial taxa and between specific groups and environmental driver variables, while stable isotope data were revealing microbial processes, such as methane oxidation in the suboxic hypolimnion. Epilimnetic and hypolimnetic microbial communities of the lake were compared over 5

consecutive years. For comparison, data from four additional lakes that reflect the full range of DOC concentrations typical for the boreal zone were also included in the study. A special focus was put on temporal and spatial changes and on poorly known but abundant microbial groups such as candidate division OD1.

Materials and methods

Study site and sampling

Time-series samples were collected from Alinen Mustajärvi in southern Finland (61°12'N, 25°06'E), a 0.007 km² head-water lake with maximum depth 6.5 m and an estimated volume of 31 × 10³ m³. The catchment area is <0.5 km² and consists of >90% coniferous forest and <10% peatland. The lake has icecover from late November until late April and it stratifies very steeply during summer. During the sampling period (from May 2006 until August 2010), the lake was only experiencing partial autumn overturns and no spring overturns. During stratifications, only the upper 1–2 m of the water column (the epilimnion) remained oxic. The natural DOC concentration in the epilimnion of the lake is around 10 mg Cl⁻¹, but in 2008 and 2009 it was experimentally elevated to 12 mg Cl⁻¹ by monthly additions of cane sugar. Effect of these additions is not in focus here and it is assumed to be equal for all layers. Water samples were taken with a 30-cm-long acrylic tube sampler (Limnos 2 L) and except for two samplings from under ice, samples were from open water season; thus, our results should not be extrapolated to winter season. Samplings were conducted twice in 2006, every other week in 2007–2008, once a month in 2009 and four times in 2010 making a total 44 time points. Samples for epilimnion, metalimnion and hypolimnion, respectively, were taken according to the oxygen concentration (Supplementary Figure 1). Epilimnion samples in 2006–2008 were pooled from 0, 1 and 2 m depths and in 2009–2010 from 0 and 1 m depths. Metalimnion samples in 2006–2008 were pooled from 3 and 4 m depths, and in 2009–2010 from 2 and 3 m depths. Hypolimnion samples in 2006 were from 5 m depth, in 2007–2008 pooled from 5 and 6 m depths and in 2009–2010 pooled from 4, 5 and 6 m depths. Water was passed through a 50- μm sieve to remove larger zooplankton. Pooled water samples were transported to the laboratory and kept at +4 °C until aliquots were processed by tangential flow filtration (Durapore cassette, pore size 0.22 μm ; Millipore, Billerica, MA, USA) within 6 h of sampling. Concentrated particles retained in this ultrafiltration were frozen and freeze dried with an Alpha 1-4 LD plus (Christ, Osterode, Germany).

Analyses of pH, total and inorganic P and N, alkalinity, as well as chlorophyll-a (chl_a) were done using standard methods (<http://www.sfs.fi/>). Particulate organic carbon, DOC and methane concentration were measured as described in Kankaala

et al. (2010), bacterial production as in Tulonen (1993) and $\delta^{13}\text{C}\text{-CH}_4$ as in Kankaala *et al.* (2007). Samples for methane concentration and $\delta^{13}\text{CH}_4$ measurement were taken every 1 m in 2008–2010. Methane concentration in the water column may diminish due to methane oxidation by methanotrophic bacteria, or because of diffusion to the atmosphere. Methane oxidation, like other microbial processes, prefers light isotope ($\delta^{12}\text{C}$), which leads to enrichment of the residual methane with heavier isotope ($\delta^{13}\text{C}$) (Whiticar, 1999). By determining the fraction of $\delta^{13}\text{C}$ in methane, we could evaluate which of these processes contributed to the decrease in methane concentration in Alinen Mustajärvi and in which depths methane oxidation was occurring. Zooplankton samples were preserved with formaldehyde and the species composition and abundance of zooplankton were determined using an inverted microscope at $\times 100$ magnification. The other four humic lakes were sampled twice in 2006 (May and October) from epilimnion, metalimnion and hypolimnion according to the oxygen concentration. The sampling and environmental characteristics of lakes Nimetön, Mekkojärvi, Valkea Kotinen, and Valkea Mustajärvi have been described in Kankaala *et al.* (2010). Summary of environmental data is given in Supplementary Table 1.

DNA extraction, PCR amplification, pyrosequencing and sequence quality control

The DNA extraction procedure was modified from protocol described by Griffiths *et al.* (2000). Briefly, 0.4 mg of freeze-dried material was homogenized with glass beads in a mixture of phenol-chloroform-isoamylalcohol (25:24:1) and hexadecyltrimethylammonium bromide. After 5 min incubation on ice to allow humic acids dissolve into phenol-chloroform-isoamylalcohol, tubes were centrifuged. The upper aqueous phase was then re-extracted with chloroform-isoamylalcohol (24:1), precipitated with polyethylene glycol and dissolved in 50 μl of TE buffer (10 mM Tris (pH 8.0), 1 mM EDTA). DNA concentrations were measured with PicoGreen in a Qubit fluorometer (Invitrogen Corporation, Carlsbad, CA, USA) showing no difference between layers ($\chi^2 = 0.0667$, $P = 0.9672$). Amplification of bacterial 16S rRNA genes (*E. coli* positions 341–805) was conducted using general bacteria primers 341F (5'-CCTACGGGNGGCWGCAG-3') and 805R (5'-GACTACHVGGGTATCTAATCC-3') (Herlemann *et al.*, 2011). Primer 341F carried a 454FLX adaptor B at the 5' end and primer 805R carried a 5-bp molecular barcode specific for each sample followed by a 454FLX adaptor A at the 5' end. PCR and amplicon processing before sequencing was performed as described in Eiler *et al.* (2012), except for purification of PCR products with Agencourt AMPure XP purification system (Beckman Coulter, Danvers, MA, USA) and amplicon quantification with PicoGreen in a Qubit fluorometer. After PCR,

amplicon concentrations were similar for all layers ($\chi^2 = 4.8343$, $P = 0.3047$). Equal concentrations of amplicons from each sample were sequenced with Titanium chemistry from adaptor A, using a 454 GS-FLX system (454 Life Sciences, Branford, CT, USA) at the Institute of Biotechnology hosted by the University of Helsinki, Finland. The resulting reads carried the sample-specific molecular barcode and covered the entire V4 region of the 16S rRNA gene as well as flanking regions. The sequence runs also included samples that were not part of this study. A total of 380 512 reads were assigned to samples from Alinen Mustajärvi (called hereafter as time-series data) and 77 146 were assigned to samples from the other four lakes (Nimetön, Mekkojärvi, Valkea Kotinen and Valkea Mustajärvi, called hereafter as four lakes data). Ambiguous sequences were removed from the data set including reads with low quality as inferred from their flowcharts and those that did not carry the exact primer sequence. After implementation of these quality control criteria, remaining reads were denoised using AmpliconNoise Version 1.22 (Quince *et al.*, 2011). AmpliconNoise implements algorithms that remove PCR and 454 pyrosequencing noise as well as the novel chimera removal tool Perseus. A total of 6201 reads was removed by Perseus using default settings ($\alpha = -7.5$, $\beta = 0.5$) for false positive detection. The 454 sequences have been deposited in the NCBI Short Read Archive under accession number SRP007933.

Operational taxonomic unit assignments

To assign reads into operational taxonomic units (OTUs), average linkage clustering was applied with a 97% sequence similarity cutoff. A representative sequence from each of the resulting 11 496 OTUs was classified using the naive Bayesian classifier (Wang *et al.*, 2007) implemented in MOTHUR (Schloss *et al.*, 2009) in combination with the Silva 102 database and the rdp6 taxonomic framework. In addition, the reads were annotated against a local freshwater bacterial sequence database that included almost 12 000 sequence entries (Newton *et al.*, 2011). This classification system was designed to maintain the phylogenetic context by which epilimnetic freshwater bacterial gene sequences historically have been identified, clustered and named. Sequences that could not be annotated to at least the phylum level with the naive Bayesian classifier were aligned in MOTHUR using kmer for finding the template sequence and Needleman for aligning sequences against the SILVA102 database (Schloss *et al.*, 2009). Aligned sequences were imported into ARB (Ludwig *et al.*, 2004) and the quick parsimony option was used to add the aligned sequences to the small subunit reference tree included in the SILVA106 small subunit rRNA gene database. The resulting position in tree was used to annotate sequences to the phylum level.

Statistical analyses

To compare and perform statistics across samples, we used the perl script daisychopper.pl (available at <http://www.genomics.ceh.ac.uk/GeneSwytch/Tools.html>; Gilbert *et al.*, 2009) to randomly resample all samples to the same size based on the sample with smallest sampling size. To include all samples and increase the number of reads after resampling, the number of reads in two samples with <1000 reads (E080826 and M070814 in the time-series data set) were multiplied by two after which the minimum read number was 1139. Simulation of the impact of this doubling was conducted (data not shown) in which all samples were resampled to 570 reads and then doubled; this had no significant affect on the Bray–Curtis indexes compared with the original data resampled to 1139 reads.

All statistical analyses were conducted using R (R Development Core Team, 2011; <http://www.R-project.org/>) and the vegan package. Non-metric multidimensional scaling plots of a Bray–Curtis distance matrix were used to visualize dynamics in community structure (β diversity) using an OTU abundance matrix based on 757 OTUs, each represented by at least 20 reads in the total data set. Kruskal–Wallis test, Spearman's Rank correlation, paired Wilcoxon's test, and permutational MANOVA were used to determine relationships between and within community components and physical and spatial variables. For Spearman's Rank correlation tests between DOC and taxonomic groups, only samples from 2006 from Alinen Mustajärvi and the four other lakes were included. The inverse Simpson index was used for diversity evaluation.

Using the resampled data matrix from the time-series data set and cutoff value $P < 0.05$, time-dependent paired occurrences of each OTU with >20 reads in epilimnion and hypolimnion were analysed with Wilcoxon's signed-rank tests (Wilcoxon, 1945). False discovery rate was used to measure the proportion of false positives in multiple testing (Storey *et al.*, 2004). This non-parametric test allowed identification of OTUs characteristic for the epilimnion or hypolimnion. To compare the dynamic properties of the communities among layers, permutational analysis of multivariate dispersions (also called MJ Anderson's permutated analysis of beta-dispersion) was applied on Bray–Curtis based dissimilarity matrices (Anderson *et al.*, 2004). This analysis tests for homogeneity of multivariate dispersion among groups of samples, hence reveals if communities from different groups differ in their taxa variability. To summarize seasonal patterns among OD1-related OTUs, all OTUs were divided into 10 clusters using k-means clustering with the Hartigan–Wong algorithm (Hartigan and Wong, 1979).

Network analysis was conducted with 138 OTUs (minimum 50 reads), 33 zooplankton taxa and 15 environmental state variables which were normalized using normal score transformation before

analysis. Significant correlations were identified based on a P -value of <0.01 and a false discovery rate of <0.05, coefficient greater than 0.3 or smaller than –0.3 among OTUs and between OTUs and environmental variables (biological and physicochemical data) using LSA as implemented in R. Other parameters for the LSA were set to 1000 permutations and a maximum delay of 0 as sampling points were not evenly spaced. The resulting LSA matrix was translated into an association network using Cytoscape 2.6.3 (Shannon *et al.*, 2003). Cytoscape depicts data sets as nodes (environmental variables and OTUs) connected by edges (significant correlations) that denote the character (positive or negative correlation) of the relationship. Network analyses have been used in previous studies on both marine and freshwater bacterioplankton communities (see, for example, Fuhrman and Steele, 2008; Fuhrman, 2009; Shade *et al.*, 2010; Steele *et al.*, 2011; Eiler *et al.*, 2012).

Results

Community composition in different layers and among lakes

Sequencing of 122 time-series samples from Alinen Mustajärvi yielded a total of 287 123 high quality reads (range 1139–4945 sequences per sample). For the 24 samples from the four lakes data set, 49 225 high quality reads were obtained (range 1680–2356 sequences per sample).

After resampling both data sets by daisychopper, 1139 sequences remained for each sample. Even though this diminished the read number per sample to only approximately one order of magnitude higher than in typical clone libraries, in total there were still 166 000 reads kept in the analysis, which is two to three orders of magnitude higher than collected by most clone library surveys. In the resampled time-series data set, these sequences clustered into 5055 OTUs (range 74–392 OTUs per sample). The number of OTUs detected in the hypolimnion was higher than in the epilimnion (paired Wilcoxon's test; $P < 0.001$, inverse Simpson's index was on average 18.1, 26.6 and 38.0 for epilimnion, metalimnion and hypolimnion, respectively). In the four lakes data set, resampling left a total of 3564 OTUs (range 66–484 per sample). The highest number of OTUs was found in Lake Nimetön and the lowest in Lake Valkea Mustajärvi. These lakes also had the highest and lowest DOC concentration, respectively. Correlation analysis revealed a positive relationship between DOC concentration and the number of OTUs found for each sample (Spearman's $\rho = 0.39$; $P < 0.05$). More specific OTU distributions can be found in Supplementary Table 2.

As expected, the microbial community composition changed gradually from epilimnion to metalimnion and hypolimnion as visualized in the

non-metric multidimensional scaling plot (Figure 1). Disparity between layers was especially clear in the time-series data set (Lake Alinen Mustajärvi), but was also obvious in the four other lakes. Permutational MANOVA confirmed the presence of distinct communities in each layer (pseudo- $F = 13.4$, $P < 0.001$) and that communities differed between lakes (pseudo- $F = 5.1$, $P < 0.001$) as well as along season (pseudo- $F = 10.7$, $P < 0.001$).

Taxonomic assignments and distinction between epilimnion- and hypolimnion-associated OTUs

The main focus in comparing the communities in different layers was put on differences between epilimnion and hypolimnion as they present highly different environments. After classification with MOTHUR, a considerable proportion of reads from both data sets remained unclassified (Figure 2a). The proportion of unclassified reads was much higher in the hypolimnia than in the epilimnia (paired Wilcoxon's test, $P < 0.001$). With ARB, half of the unclassified reads were annotated to candidate divisions, most of them belonging to candidate division OD1, which constituted about 70% and 40% of unclassified reads in time-series and four lakes data sets, respectively (Figure 2b). In the time-series data set, the most abundant unclassified OTU (OTU 1237) after classification in MOTHUR (including 38.4% of all unclassified reads and 4% of all reads) was affiliated with OD1 using the

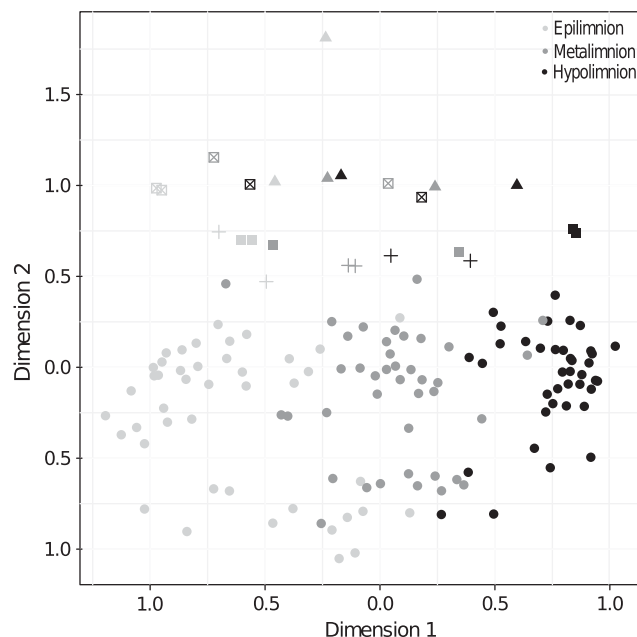


Figure 1 Non-metric multidimensional scaling ordination visualizing the Bray–Curtis distances of all 122 samples from the time-series data set from Lake Alinen Mustajärvi and 24 samples from the four lakes data set. The five lakes are represented by different symbols (■ = Nimetön, △ = Mekkojärvi, + = Valkea Kotinen, ● = Alinen Mustajärvi, ⊠ = Valkea Mustajärvi) and the three separate depth layers are indicated by grey scale. The stress value is 17.3.

positions of the sequences in the ARB reference tree. All other abundant OD1 OTUs, such as OTU1357, OTU1422, OTU1526 and OTU1498, could also be annotated as OD1 using their position in the phylogenetic tree after parsimony insertion.

OTUs were also annotated using a freshwater-specific sequence database to look for the typical freshwater community members in the five lakes (Figure 2c). In the time-series data set, the most abundant freshwater tribes in the epilimnion were acI-B2 (16.3% of all epilimnetic reads), acV-A2 (8.5%) and PnecC (6.9%) and for hypolimnion PnecC (3.0%), acI-B2 (2.3%) and acV-A1 (1.7%). Among lakes, DOC concentration (Figure 3a) was found to have a positive correlation with the proportion of reads affiliated with candidate divisions (Spearman's $\rho = 0.40$, $P < 0.05$) and OD1 (Spearman's $\rho = 0.37$, $P < 0.05$) and a negative correlation with freshwater tribes (taxa) (Spearman's $\rho = -0.64$, $P < 0.001$). This indicates that generally high DOC lakes are neglected in studies of freshwater microbial communities. OD1 was also positively correlated with methane concentration (Figure 3b; Spearman's $\rho = 0.61$, $P < 0.001$) among lakes as well as in time-series data set (Spearman's $\rho = 0.43$, $P < 0.01$).

With combined MOTHUR and ARB classification, the main phyla in the epilimnion of time-series data set were *Actinobacteria* and *Betaproteobacteria*, while the main phyla in the hypolimnion were candidate division OD1, *Chlorobi* and *Bacteroidetes* (Figure 4a). Even though many of the taxa found from metalimnion were shared with the other layers, the community still had distinct features. For example, *Cyanobacteria* were mainly found in this layer. OTUs typical for either epilimnion or hypolimnion were identified with Wilcoxon's signed-rank test, using all OTUs with > 20 reads. OTUs that had significantly more reads in epilimnion or hypolimnion were defined as epilimnion-associated or hypolimnion-associated. OTUs with different occurrences between epilimnion and hypolimnion contained 43.1% of reads included into the analysis. In the epilimnion by far the most abundant phylum was *Actinobacteria*, which comprised more than half of all epilimnion-associated reads (Figure 4b). Reads affiliated to *Betaproteobacteria* were also much more abundant in the epilimnion, constituting almost one quarter of all epilimnion-associated reads. In all, there were only representatives from seven phyla among the epilimnion-associated OTUs. Among the OTUs that were more abundant in the hypolimnion, there were representatives from 20 different phyla (Figure 4c). A third of all hypolimnion-associated reads were affiliated with candidate division OD1, while *Chlorobi* and *Bacteroidetes* were also well represented. The only phylum with equal proportion of epilimnion- and hypolimnion-associated reads was *Verrucomicrobia*.

Since the presence of OD1 in lakes is rarely reported in the literature, the fit of the most common

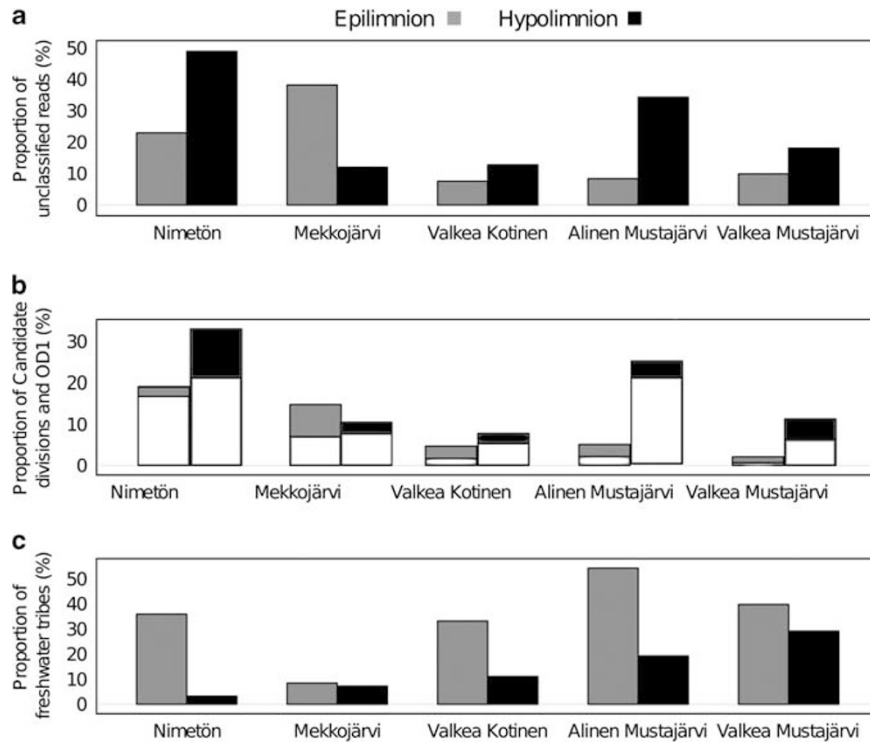


Figure 2 The mean proportion of unclassified reads using the naive Bayesian classifier in MOTHUR (a). The mean proportion of reads annotated with candidate divisions after classification with ARB; the proportion of OD1 reads is indicated in white (b). The proportion of reads annotated with typical freshwater tribes as described in Newton *et al.* (2011) (c). All panels present proportions from all sequences.

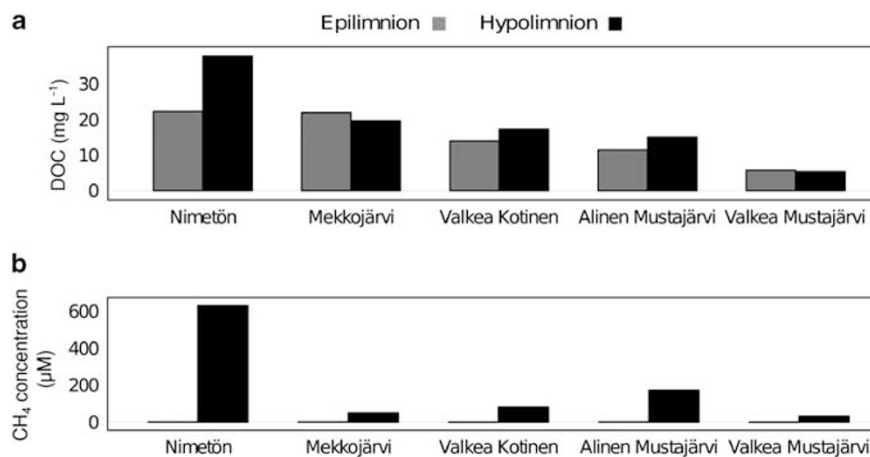


Figure 3 Mean DOC concentration (mg l⁻¹) (a) and CH₄ concentration (μM) (b) in epilimnion and hypolimnion of Lakes Nimetön, Mekkojärvi, Valkea Kotinen, Alinen Mustajärvi and Valkea Mustajärvi.

16S rRNA-targeting primers used for bacterial community analyses to all OD1-related sequences in RDP database (release 10, update 27) was tested with RDP Probe Match (<http://rdp.cme.msu.edu/probematch/search.jsp>) (Table 1). The primers used in our study (341F and 805r) exactly matched a much higher percentage of OD1 sequences (64% and 60%, respectively) compared with other commonly used primers (from 0% to 22%).

Temporal dynamics of OTUs in Lake Alinen Mustajärvi
In the time-series data set, the bacterial communities in epilimnion and metalimnion were more dynamic in composition than in the hypolimnion (average Bray–Curtis distances 0.75 (±0.15), 0.74 (±0.15) and 0.67 (±0.15), respectively), which is also apparent in the non-metric multidimensional scaling plot (Figure 1). Reads affiliated with OD1 constituted a considerable and persistent proportion

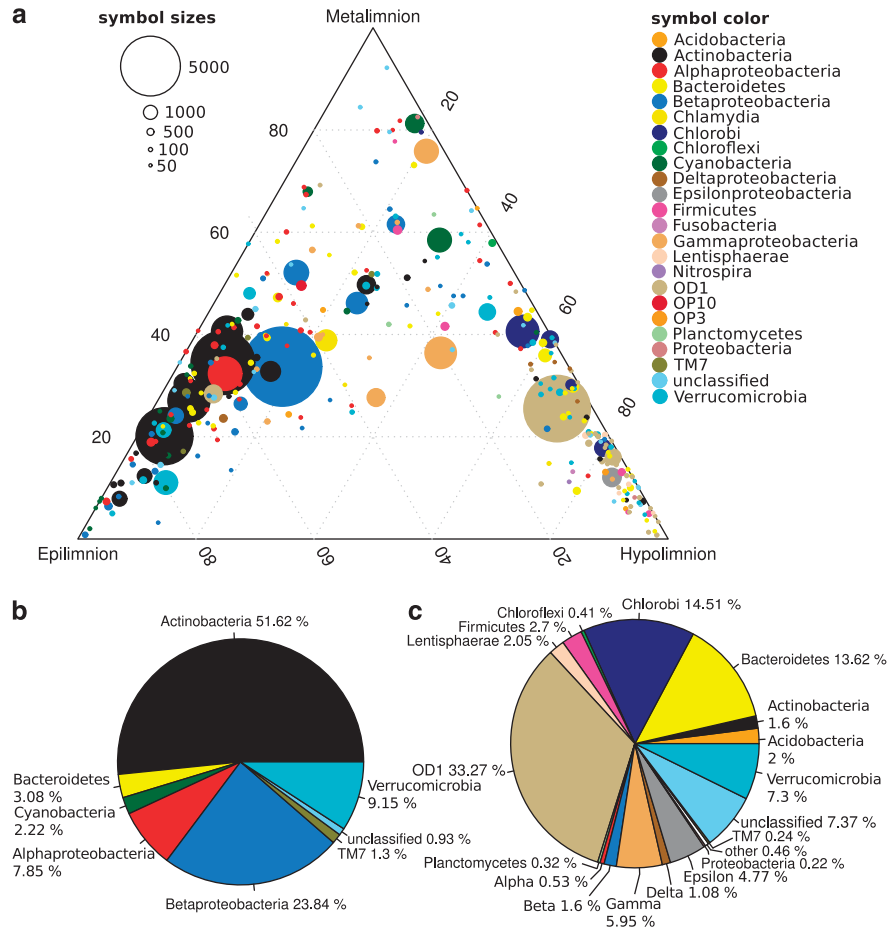


Figure 4 Ternary plot showing the distribution of OTUs in the time-series data from Lake Alinen Mustajärvi (a). Axes represent the three depth layers of the lake (epilimnion, metalimnion and hypolimnion) and the percentage of reads associated with each layer for each OTU. The size of the symbol indicates number of reads associated with each OTU and taxonomic affiliation are indicated by colours. Pie charts indicating the proportion of reads affiliated with OTUs associated with epilimnion (b) and hypolimnion (c) as determined by Wilcoxon's signed-rank test using the time-series data set (false discovery rate $q=0.021$). Colours represent different phyla as in (a). In (a), OTUs represented by at least 50 reads are included, panels (b) and (c) include all OTUs represented by at least 20 reads in the time-series data set.

Table 1 Comparison of fit of commonly used primers with sequences affiliated with candidate division OD1

Primer	Sequence	Fit to OD1	Reference
8F	AGAGTTTGATCCTGGCTCAG	14.6/24.2/27.9	Edwards <i>et al.</i> (1989)
27F	AGAGTTTGATCMTGGCTCAG	21.5/26.6/28.6	Lane (1991)
*341F	CCTACGGGNGGCWGCAG	63.6/91.8/93.8	Herlemann <i>et al.</i> (2011)
967f	CAACGCGAAGAACCTTACC	0/0/12.2	Sogin <i>et al.</i> (2006)
338r	TGCTGCCTCCCGTAGGAGT	0.2/57.9/59.9	Francés <i>et al.</i> (2004)
518r	ATTACCGCGGCTGCTGG	2.0/49.9/81.8	Neefs <i>et al.</i> (1990)
534r	ATTACCGCGGCTGCTGGCA	2.0/49.4/81.6	Muyzer <i>et al.</i> (1996)
*805r	GACTACHVGGGTATCTAATCC	59.6/72.1/74.3	Herlemann <i>et al.</i> (2011)
907r	CCGTCGAATTCMTTGTGAGTTT	8.0/50.1/51.9	Johnson (1994)
1046r	CGACAGCCATGCANACCT	1.1/25.1/37.9	Sogin <i>et al.</i> (2006)
1387R	GGGCGGWGTGTACAAGGC	1.6/4.7/14.2	Marchesi <i>et al.</i> (1998)
1492R	GTTACCTTGTTCGACTT	2.2/2.7/3.8	Lane (1991)

Third column indicates the proportion of OD1 sequences matching the primer sequence with 0, 1 and 2 mismatches allowed.

*Indicate the primers used in this study.

of hypolimnetic reads, the average representation being 24.8% (Figure 5a). Still, each year the OD1 population apparently crashed upon autumn overturn. The most common OD1, OTU1237, contributed almost half of all reads annotated as OD1.

The abundant OTUs in the hypolimnion were less dynamic when compared with rarer OTUs, as suggested by a significant negative correlation between their relative abundances and corresponding standardized variances (Spearman's rank

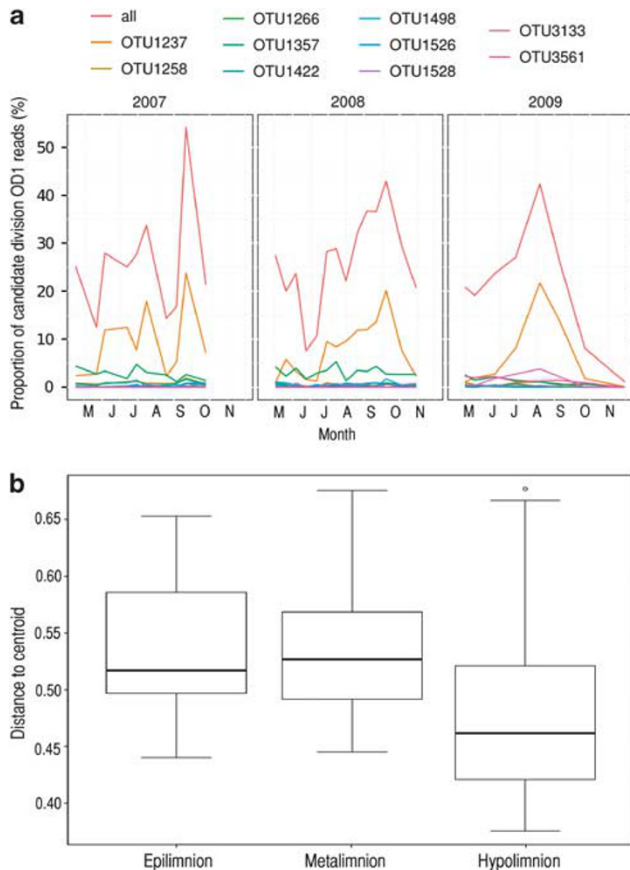


Figure 5 Proportion of all OD1-related sequences and of the 10 most abundant OD1-affiliated OTUs (a). The distribution of the distance from the centroid (b) determined from the temporal dynamics of the bacterial communities in the corresponding layers. The distance from the centroid was estimated by MJ Anderson's permutated analysis of beta-dispersion and represents a statistical measure for the variability of the communities over time.

correlations $R = -0.41$, $P < 0.001$). Results of the permutated analysis of beta-dispersion revealed that there was a significant difference in the temporal variability of taxa over time ($F = 10.468$, $P < 0.001$), with the hypolimnion showing a significant higher homogeneity (lower distance to centroid) than the other two layers (Figure 5b).

Association networks of the hypolimnetic bacterioplankton community

Complete results from time-series data analysed by LSA were visualized in association networks (Supplementary Figure 2; Supplementary Table 3). From all possible interactions, 9.2% were significant at significance level $P < 0.01$ and with a false discovery rate $q < 0.05$. OTUs annotated as OD1 were often significantly positively correlated with each other, most likely indicating functional interdependencies or similar environmental preferences. Further, many OD1-related OTUs were among the

most connected OTUs, for example, OTU1357, for which 28% of all possible interactions were significant. *Chlorobi* OTU1239 (second most abundant OTU in the hypolimnion) was negatively related to the OD1 OTU1357 and OTU1422, suggesting competition or differences in growth constraints among these OTUs. From the complete network, two subnetworks were derived to visualize only the OTUs connected to CH_4 (Figure 6a) and to NH_4^+ and combined NO_2^- and NO_3^- (Figure 6b). OTU1357, OTU1237 and OTU1462 annotated as OD1 were positively related to CH_4 concentrations, as was OTU727, annotated as *Verrucomicrobia*. OD1 OTUs were also overrepresented in the number of positive correlations with ammonium including six of the most abundant OD1 OTUs. Only two OTUs, one OD1 and one *Betaproteobacteria*, were significantly correlated with the combined concentration of nitrite and nitrate.

Nutrient and DOC concentrations in Alinen Mustajärvi increased from epilimnion to hypolimnion while the oxygen concentration decreased (Table 2). The $\delta^{13}\text{C}\text{-CH}_4$ values and methane concentrations indicated that the methane was oxidized in the hypolimnion (Supplementary Figure 3). Methane concentration in the epilimnion and metalimnion was lower than in hypolimnion ($\chi^2 = 89.7556$, $P < 0.001$) while the $\delta^{13}\text{C}\text{-CH}_4$ value was higher ($\chi^2 = 84.7$, $P < 0.001$). There was a drastic decrease in methane concentration between bottom and 4 m depth coinciding with a minor enrichment in $\delta^{13}\text{C}\text{-CH}_4$. The change in $\delta^{13}\text{C}$ did not seem drastic since the more CH_4 there is, the smaller the change in isotope value. However, calculation of CH_4 oxidation using fractionation factor determined for methane oxidation in polyhumic boreal lakes (1.037; Kankaala *et al.*, 2007), showed that this change between 4 and 6 m corresponded to half up to almost all of the oxidized CH_4 in the water column. Thus, even though we could not determine whether the oxidation was aerobic or anaerobic, a major portion of CH_4 was oxidized in suboxic hypolimnion. Several OTUs annotated as OD1, such as OTU28, OTU1237 and OTU1357, were highly abundant in these depth horizons and as mentioned above correlated positively with methane concentrations.

Discussion

A key finding of our study was that the representation of tribes previously described as typical for freshwater (Newton *et al.*, 2011) was low in the hypolimnia while the proportion of candidate divisions and other poorly described phyla was very high. At the same time, these hypolimnetic microbial communities were more stable compared with the more dynamic communities in the epilimnion, corroborating the previously findings from other types of lakes (Shade *et al.*, 2008; Barberan and Casamayor, 2011; Berdjeb *et al.*, 2011; Pollet

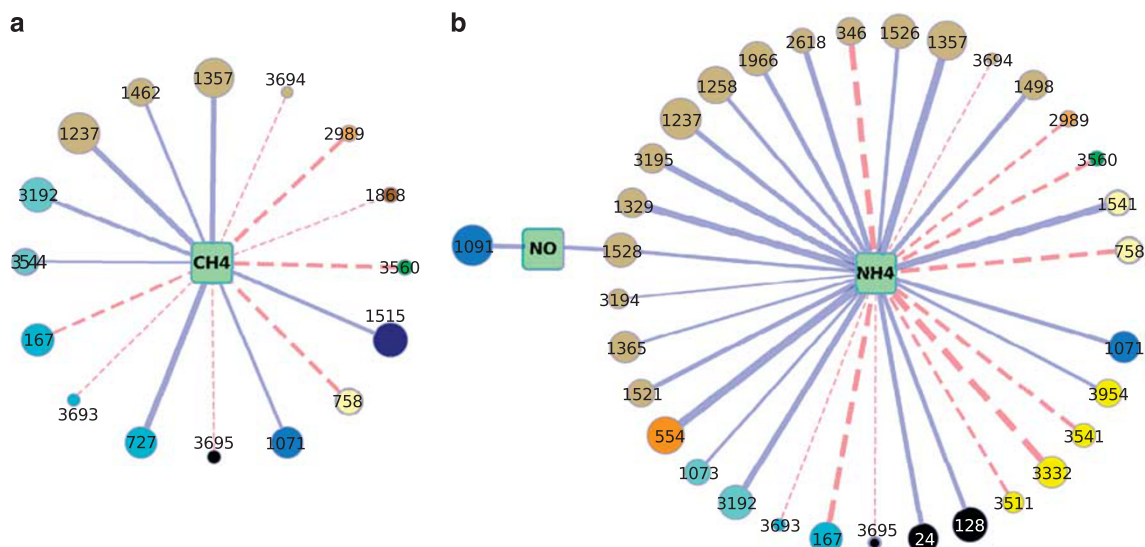


Figure 6 Subnetworks organized around methane (a) and inorganic nitrogen (b) compounds. The subnetworks were extracted from the entire association network (see Supplementary Figure 2) and show all OTUs (nodes) correlating with the respective compound. Circles represent bacterioplankton OTUs where colours represent their taxonomic annotations at the phyla level as in Figure 3a. The size of the circles reflects the number of samples where each OTU could be detected. Blue solid lines (edges) indicate positive correlations and red dashed lines (edges) indicate negative correlations, whereas the width of the line reflects the LSA coefficient. Abbreviations for the environmental variables are translated as following: NH_4 , ammonium concentrations, NO, combined nitrate and nitrite concentrations and CH_4 , methane concentrations.

Table 2 Mean $\delta^{13}\text{C}_4$ values and mean concentrations of O_2 , DOC, CH_4 , NH_4^+ , combined NO_3^- and NO_2^- and PO_4^- for each layer in Alinen Mustajärvi

Layer	O_2 (mg l^{-1})	DOC (mg l^{-1})	CH_4 (nM)	$\delta^{13}\text{C}_4$	NH_4^+ ($\mu\text{g l}^{-1}$)	$\text{NO}_3^- + \text{NO}_2^-$ ($\mu\text{g l}^{-1}$)	PO_4^- ($\mu\text{g l}^{-1}$)
Epilimnion	6.6	11.2	3.6×10^3	-50.7	39.4	21.3	1.4
Metalimnion	1.4	11.9	2.0×10^4	-52.5	136.0	16.3	1.4
Hypolimnion	0.5	18.2	2.0×10^5	-69.1	1222.4	18.9	54.6

Abbreviation: DOC, dissolved organic carbon.

et al., 2011). This may simply reflect more pronounced seasonal shifts in temperature (range 0.9–18 °C in the epilimnion compared with 3.6–7.5 °C in the hypolimnion), light and nutrient concentrations in the epilimnion compared with the more stable deeper strata. Other possible reasons for higher variability in the epilimnetic community over time may be selective feeding by heteronano-flagellates (Fu *et al.*, 2003; Massana *et al.*, 2009) or pulsed inputs of organic compounds from allochthonous production that are more labile in the surface layers, while quenched before reaching the deeper waters via sedimentation. Additional explanations for the low variability in hypolimnetic community may arise from the fact that anaerobic metabolic pathways have a lower energy yield compared with aerobic metabolism and that the mineralization of organic compounds in anaerobic systems is a multistep process that must involve complex communities (Magonigal *et al.*, 2005). Lower energy yield and the need for a complex community may also facilitate the higher diversity in the

bacterioplankton communities of the hypolimnion compared with the epilimnion.

There was a clear shift in bacterial community composition from epilimnion to hypolimnion with metalimnion sharing features with both neighbouring layers. Low proportion of typical freshwater tribes in the hypolimnion emphasizes that the microbial communities in this type of waters are generally underexplored. The proportion of freshwater taxa was also decreasing while the proportion of candidate divisions increased with increasing DOC concentration in both the epilimnion and hypolimnion. The epilimnion-associated OTUs were mainly annotated as acI-B2, acV-A2, xip-A1, PncC and *Novosphingobium* (Novo-A2), which have previously been described as abundant tribes in the epilimnion of humic lakes (Newton *et al.*, 2011). In the hypolimnion, the most abundant OTU (OTU1237) was annotated as OD1. In previous studies, *Chlorobi* (Taipale *et al.*, 2009; Barberan and Casamayor, 2011) and *Bacteroidetes* (Barberan and Casamayor, 2011) have been found to be

major groups in the anoxic hypolimnion, but the dominance of OD1-related sequences has been reported only once (Briée *et al.*, 2007) in a suboxic pond, where it comprised 25% of the sequences from a clone library. We suggest that OD1 may well have been overlooked and neglected as a result of insufficient coverage in previous bacterial PCR surveys. Most commonly used primers for such bacterial community surveys do not target this group, whereas the bacterial primers used in our study (341f and 805r) have a much higher coverage to known sequences of candidate division OD1. We found the proportion of OD1 sequences in Lake Mekkojärvi to be 7.0%, whereas Taipale *et al.* (2009) did not detect them in the same lake using the bacterial 16S rRNA primer set 27f-518r. Similar underestimation by most commonly used bacterial primers was recently observed also for another poorly characterized phylum, *Verrucomicrobia* (Bergmann *et al.*, 2011).

Other abundant OTUs in the hypolimnion were annotated as *Proteobacteria* such as the epsilonbacterial genus *Sulfurimonas* (OTU1271) and the gammaproteobacterial genus *Methylobacter* (OTU1103). Members of the genus *Methylobacter* are group I methanotrophs (Bowman *et al.*, 1993) and they were likely contributing to the methane oxidation as indicated by the stable isotope data. In addition, anaerobic denitrification is suggested to be a common process in the studied lakes as indicated by the presence of *Sulfurimonas* and the oversaturation of N₂ in the suboxic hypolimnia (Tirola *et al.*, 2011). *Sulfurimonas* isolates represent sulphur-oxidizing chemolithoautotrophs (for example, Inagaki *et al.*, 2003; Takai *et al.*, 2006) that can perform autotrophic denitrification. This chemolithotrophic process couples denitrification with the oxidation of reduced inorganic sulphur compounds and members of the genus *Sulfurimonas* have been identified as the main denitrifiers in chemoclines of the Baltic Sea (Glaubitz *et al.*, 2010). OTUs identified as *Chlorobi* (OTU1239 and OTU1268) were also abundant; isolates belonging to the phylum *Chlorobi* have been described as anaerobic phototrophs that require sulphur and can photoassimilate acetate or pyruvate directly into the amino acids leading to cell growth (Hoare and Gibson, 1964; Hell *et al.*, 2008). OTU727 correlating positively with methane concentration was affiliated with phylum *Verrucomicrobia*. Recently, *Verrucomicrobia* have been implicated with methanotrophic metabolism in extremely acidic environment (Dunfield *et al.*, 2007; Pol *et al.*, 2007). These organisms harbour a novel methylotrophic pathway, with *PmoA* genes distinct from *Proteobacteria*, which were considered to be the only phylum capable of methanotrophy before the discovery of these new methanotrophs.

Metabolic pathways in candidate division OD1 have not been specified as the members of this candidate phylum have never been cultivated and only a partial genome has been recovered (Elshahed

et al., 2005). Nevertheless, even the scarce data available for this group may still give a hint about the ecology and metabolic capabilities of organisms within this division. OD1 has mainly been detected in sulphur-rich, anoxic environments (for example, Elshahed *et al.*, 2005; Briée *et al.*, 2007; Barberan and Casamayor, 2011), and the partial genome revealed genes known only from anaerobic or facultatively anaerobic microorganisms presenting a similar genomic organization as methanotrophic *Archaea* (Elshahed *et al.*, 2005) found in anoxic marine sediments (Hallam *et al.*, 2004). Network analysis revealed mostly significant positive correlations with methane and ammonium concentrations, and stable isotope data were indicating methane oxidation in the same depth where OD1 OTUs were highly abundant. Furthermore, the distribution of OD1-related reads in the studied lakes suggests preference towards environments with high DOC and CH₄ concentrations. Together, these findings suggest that members of the division are very likely anaerobic, and might have a role in anaerobic methane oxidation. In addition, OTUs within the OD1 candidate division were mainly positively correlated with each other, suggesting that these OTUs share many ecological properties with considerable niche overlap. Furthermore, OD1 OTUs represent abundant and persistent populations with low proportional variability over the ice-free season. Together with network theory this suggests that several OD1 OTUs occupy key roles in the environment. OTUs with a high number of associations (such as OTU1357 which had the highest number of significant correlations) can be regarded as a 'hub' in the complex network (see Montoya *et al.*, 2006) that anaerobic lake bacterioplankton communities form in the hypolimnion of boreal lakes.

Still, there is clearly a need to verify these hypotheses on metabolic processes and the identity of the involved key players by functional metagenomic and single cell approaches. Such approaches may support the hypotheses postulated from our field survey on the metabolic capabilities of the diverse and abundant candidate division OD1 and other poorly known groups in boreal lakes. In conclusion, this study shows that the previously described freshwater-related tribes are not abundant in suboxic hypolimnia, which reflects the current focus on aerobic freshwater microbiology. Further, the persistent dominance and high network connectivity of candidate division OD1 suggests that these previously overlooked microbes have an important ecological role in these mostly anoxic environments. Overall, our results for bacterioplankton in boreal lakes revealed a higher diversity and more persistent community composition in the hypolimnion compared with epilimnion and metalimnion. Together with high abundance of poorly characterized bacteria this points to distinct ecological properties of these suboxic communities.

Acknowledgements

This work was supported by Academy of Finland (grant 114604 to RIJ) and the Swedish Research Council (grant to SB). SP was funded by a grant from the rector of the University of Jyväskylä. We are grateful to Drs Paula Kankaala and Sami Taipale for providing additional data for some of the lakes. We thank Lammi Biological Station, University of Helsinki, for the facilities available and all personnel assisting with the experiments. We also want to thank the Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) for help with data storage and analysis, and Minna Hiltunen for providing zooplankton data. Biogeochemistry group at University of Eastern Finland is acknowledged for providing tools for CH₄ isotope analyses.

References

- Anderson MJ, Ellingsen KE, McArdle BH. (2004). Multivariate dispersion as a measure of beta diversity. *Ecol Lett* **6**: 683–693.
- Ask J, Karlsson J, Persson L, Ask P, Byström P, Jansson M. (2009). Terrestrial organic matter and light penetration: effects on bacterial and primary production in lakes. *Limnol Oceanogr* **54**: 2034–2040.
- Barberan A, Casamayor EO. (2011). Euxinic freshwater hypolimnia promote bacterial endemism in continental areas. *Microb Ecol* **61**: 465–472.
- Bastviken D, Tranvik LJ, Downing JA, Crill PM, Enrich-Bas A. (2011). Freshwater methane emissions offset the continental carbon sink. *Science* **331**: 50.
- Battin TJ, Luysaert S, Kaplan LA, Aufdenkampe AK, Richter A, Tranvik LJ. (2009). The boundless carbon cycle. *Nat Geosci* **2**: 598–600.
- Bent SJ, Pierson JD, Forney LJ. (2007). Measuring species richness based on microbial community fingerprints: the emperor has no clothes. *Appl Environ Microbiol* **73**: 2399–2401.
- Berdjeb L, Ghiglione JF, Jacquet S. (2011). Bottom-up versus top-down control of hypo- and epilimnion free living bacterial community structures in two neighboring freshwater lakes. *Appl Environ Microbiol* **77**: 3591–3599.
- Bergmann GT, Bates ST, Eilers KG, Lauber CL, Caporaso JG, Walters JA *et al.* (2011). The under-recognized dominance of Verrucomicrobia in soil bacterial communities. *Soil Biol Biochem* **43**: 1450–1455.
- Bowman JP, Sly LI, Nichols LPD, Hayward AC. (1993). Revised taxonomy of the methanotrophs: description of *Methylobacter* gen nov., emendation of *Methylococcus*, validation of *Methylosinus* and *Methylocystis* species, and a proposal that the family *Methylococcaceae* includes only the group I methanotrophs. *Int J Syst Bacteriol* **43**: 735–753.
- Briée C, Moreira D, López-Gracia P. (2007). Archaeal and bacterial community composition of sediment and plankton from a suboxic freshwater pond. *Res Microbiol* **158**: 213–227.
- Burkert U, Warnecke F, Babenzien D, Zwirnmann E, Pernthaler J. (2003). Members of a readily enriched β -proteobacterial clade are common in surface waters of a humic lake. *Appl Environ Microbiol* **69**: 6550–6559.
- Cole JJ, Caraco NF, Kling GW, Kratz TK. (1994). Carbon dioxide supersaturation in the surface waters of lakes. *Science* **265**: 1568–1570.
- del Giorgio PA, Peters RH. (1994). Patterns in planktonic P:R ratios in lakes: influence of lake trophy and dissolved organic C. *Limnol Oceanogr* **39**: 772–787.
- Dunfield PF, Yuryev A, Senin P, Smirnova AV, Stott MB, Hou S *et al.* (2007). Methane oxidation by an extremely acidophilic bacterium of the phylum Verrucomicrobia. *Nature* **450**: 879–882.
- Edwards U, Rogall T, Blöcker H, Emde M, Böttger EC. (1989). Isolation and direct sequencing of entire genes. Characterization of a gene coding for 16S ribosomal RNA. *Nucleic Acids Res* **17**: 7843–7853.
- Eiler A, Heinrich F, Bertilsson S. (2012). Coherent dynamics and association networks among lake bacterioplankton taxa. *ISME J* **6**: 330–342.
- Elshahed MS, Najjar FZ, Aycocock M, Qu C, Roe BA, Krumholz LR. (2005). Metagenomic analysis of the microbial community at Zodletone Spring (Oklahoma): insights into the genome of a member of the novel Candidate Division OD1. *Appl Environ Microbiol* **71**: 7598–7602.
- Francés R, Benlloch S, Zapater P, González JM, Lozano B, Muñoz C *et al.* (2004). A sequential study of serum bacterial DNA in patients with advanced cirrhosis and ascites. *Hepatology* **39**: 484–491.
- Fu Y, O’Kelly C, Sieracki M, Distel DL. (2003). Protistan grazing analysis by flow cytometry using prey labeled by *in vivo* expression of fluorescent proteins. *Appl Environ Microbiol* **69**: 6848–6855.
- Fuhrman JA, Steele JA. (2008). Community structure of marine bacterioplankton: patterns, networks, and relationships to function. *Aquat Microb Ecol* **53**: 69–81.
- Fuhrman JA. (2009). Microbial community structure and its functional implications. *Nature* **459**: 193–199.
- Gilbert JA, Field D, Swift P, Newbold L, Oliver A, Smyth T *et al.* (2009). The seasonal structure of microbial communities in the Western English Channel. *Environ Microbiol* **11**: 3132–3139.
- Glaubitx S, Labrenz M, Jost G, Jürgens K. (2010). Diversity of active chemolithoautotrophic prokaryotes in the sulfidic zone of a Black Sea pelagic redoxcline as determined by rRNA-based stable isotope probing. *FEMS Microbiol Ecol* **74**: 32–41.
- Griffiths RI, Whiteley AS, O’Donnell AG, Bailey MJ. (2000). Rapid method for coextraction of DNA and RNA from natural environments for analysis of ribosomal DNA- and rRNA-based microbial community composition. *Appl Environ Microbiol* **66**: 5488–5491.
- Gudasz C, Bastviken D, Steger K, Premke K, Sobek S, Tranvik LJ. (2010). Temperature-controlled organic carbon mineralization in lake sediments. *Nature* **466**: 478–481.
- Hallam SJ, Putnam N, Preston CM, Detter JC, Rokhsar D, Richardson PM *et al.* (2004). Reverse methanogenesis: testing the hypothesis with environmental genomics. *Science* **305**: 1457–1462.
- Hartigan JA, Wong MA. (1979). A K-means clustering algorithm. *Appl Stat* **28**: 100–108.
- Hell R, Dahl C, Knaff DB, Leustek T (eds). (2008). Sulfur metabolism in phototrophic organisms. In: *Advances in Photosynthesis and Respiration* 27. Springer: Dordrecht, p 516.

- Herlemann DPR, Labrenz M, Jürgens K, Bertilsson S, Waniek JJ, Andersson AF. (2011). Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME J* **5**: 1571–1579.
- Hoare DS, Gibson J. (1964). Photoassimilation of acetate and the biosynthesis of amino acids by *Chlorobium thiosulphatophilum*. *Biochem J* **91**: 546–559.
- Inagaki F, Takai K, Kobayashi H, Neelson KH, Horikoshi K. (2003). *Sulfurimonas autotrophica* gen. nov., sp. nov., a novel sulfur-oxidizing ϵ -proteobacterium isolated from hydrothermal sediments in the Mid-Okinawa Trough. *Int J Syst Evol Microbiol* **53**: 1801–1805.
- Johnson JL. (1994). Similarity analysis of DNAs. In: Gerhardt P, Murray RGE, Wood WA, Krieg NR (eds), *Methods for General and Molecular Bacteriology*. American Society for Microbiology: Washington, DC, pp 655–682.
- Jones RI, Grey J, Quarmby C, Sleep D. (2001). Sources and fluxes of inorganic carbon in a deep, oligotrophic lake (Loch Ness, Scotland). *Global Biogeochem Cycles* **15**: 863–870.
- Kankaala P, Arvola L, Tulonen T, Ojala A. (1996). Carbon budget for the pelagic food web of the euphotic zone in a boreal lake (Lake Pääjärvi). *Can J Fish Aquat Sci* **53**: 1663–1674.
- Kankaala P, Taipale S, Li L, Jones RI. (2010). Diets of crustacean zooplankton, inferred from stable carbon and nitrogen isotope analyses, in lakes with varying allochthonous dissolved organic carbon content. *Aquat Ecol* **44**: 781–795.
- Kankaala P, Taipale S, Nykänen H, Jones RI. (2007). Oxidation, efflux, and isotopic fractionation of methane during autumnal turnover in a polyhumic, boreal lake. *J Geophys Res* **112**: G02003.
- Kent AD, Jones SE, Yannarell AC, Graham JM, Lauster GH, Kratz TK *et al.* (2004). Annual patterns in bacterioplankton community variability in a humic lake. *Microb Ecol* **48**: 550–560.
- Kleinsteuber S, Mueller F-D, Chatzinotas A, Wendt-Potthoff K, Harms H. (2008). Diversity and *in situ* quantification of Acidobacteria subdivision 1 in an acidic mining lake. *FEMS Microbiol Ecol* **63**: 107–117.
- Lane DJ. (1991). 16S/23S rRNA sequencing. In: Stackebrandt E, Goodfellow M (eds), *Nucleic Acid Techniques in Bacterial Systematics*. Wiley: New York, pp 115–175.
- Lehours A-C, Bardot C, Pelisson P-F, Guedon A, Pesce S, Demeure G *et al.* (2009). Successional changes in bacterial community assemblages following anoxia in the hypolimnion of a eutrophic lake. *Aquat Microb Ecol* **54**: 71–82.
- Lindström E. (1998). Bacterioplankton community composition in a boreal forest lake. *FEMS Microbiol Ecol* **27**: 163–174.
- Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar *et al.* (2004). ARB: a software environment for sequence data. *Nucleic Acids Res* **32**: 1363–1371.
- Marchesi JR, Sato T, Weightman AJ, Martin TA, Fry JC, Hiom SJ *et al.* (1998). Design and evaluation of useful bacterium-specific PCR primers that amplify genes coding. *Appl Environ Microbiol* **64**: 795–799.
- Massana R, Unrein F, Rodríguez-Martínez R, Forn I, Lefort T, Pinhassi J *et al.* (2009). Grazing rates and functional diversity of uncultured heterotrophic flagellates. *ISME J* **3**: 588–596.
- Megonigal JP, Hines ME, Visscher PT. (2005). Anaerobic metabolism: linkages to trace gases and aerobic processes. In: Schlesinger WH (ed), *Biogeochemistry*. Elsevier: Oxford, pp 317–424.
- Montoya JM, Pimm SL, Solé RV. (2006). Ecological networks and their fragility. *Nature* **442**: 259–264.
- Muyzer G, Hottentrager S, Teske A, Wawer C. (1996). Denaturing gradient gel electrophoresis of PCR-amplified 16S rDNA. A new molecular approach to analyze the genetic diversity of mixed microbial communities. In: Akkermans ADL, van Elsas JD, de Bruijn FJ (eds), *Molecular Microbial Ecology Manual*. Kluwer Academic Publishing: Dordrecht, pp 3.4.4.1–3.4.4.22.
- Neefs J-M, Van de Peer Y, Hendriks L, De Wachter R. (1990). Compilation of small ribosomal subunit RNA sequences. *Nucleic Acids Res* **18**: 2237–2317.
- Newton RJ, Jones SE, Eiler A, McMahon KD, Bertilsson S. (2011). A guide to the natural history of freshwater lake bacteria. *Microbiol Mol Biol Rev* **75**: 14–49.
- Pol A, Heijmans K, Harhangi HR, Tedesco D, Jetten MSM, Op den Camp HJM. (2007). Methanotrophy below pH 1 by a new Verrucomicrobia species. *Nature* **450**: 874–878.
- Pollet T, Tadonlélé RD, Humbert JF. (2011). Spatiotemporal changes in the structure and composition of a less-abundant bacterial phylum (Planctomycetes) in two perialpine lakes. *Appl Environ Microbiol* **77**: 4811–4821.
- Quince C, Lanzen A, Davenport RJ, Turnbaugh PJ. (2011). Removing noise from pyrosequenced amplicons. *BMC Bioinformatics* **12**: 38.
- R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, p 2630.
- Ruan Q, Dutta D, Schwalbach MS, Steele JA, Fuhrman JA, Sun F. (2006). Local similarity analysis reveals unique associations among marine bacterioplankton species and environmental factors. *Bioinformatics* **22**: 2532–2538.
- Salonen K, Arvola L, Rask M. (1984). Autumnal and vernal circulation of small forest lakes in Southern Finland. *Verh Int Verein Limnol* **22**: 103–107.
- Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB *et al.* (2009). Introducing Mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**: 7537–7541.
- Shade A, Chiu CY, McMahon KD. (2010). Differential bacterial dynamics promote emergent community robustness to lake mixing: an epilimnion to hypolimnion transplant experiment. *Environ Microbiol* **12**: 455–466.
- Shade A, Jones SE, McMahon KD. (2008). The influence of habitat heterogeneity on freshwater bacterial community composition and dynamics. *Environ Microbiol* **10**: 1057–1067.
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D *et al.* (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**: 2498–2504.
- Sogin ML, Morrison HG, Huber JA, Welch DM, Huse SM, Neal RP *et al.* (2006). Microbial diversity in the deep sea and the underexplored “rare biosphere”. *PNAS* **103**: 12115–12120.
- Steele JA, Countway PD, Xia L, Vigil PD, Beman JM, Kim DY *et al.* (2011). Marine bacterial, archaeal and

- protistan association networks reveal ecological linkages. *ISME J* **5**: 1414–1425.
- Storey JD, Taylor JE, Siegmund D. (2004). Strong control, conservative point estimation, and simultaneous conservative consistency of false discovery rates: a unified approach. *J R Stat Soc B* **66**: 187–205.
- Taipale S, Jones RI, Tiirola M. (2009). Vertical diversity of bacteria in an oxygen-stratified humic lake, evaluated using DNA and phospholipid analyses. *Aquat Microb Ecol* **55**: 1–16.
- Taipale S, Kankaala P, Hahn MW, Jones RI, Tiirola M. (2011). Methane-oxidizing and photoautotrophic bacteria are major producers in a humic lake with a large anoxic hypolimnion. *Aquat Microb Ecol* **64**: 81–95.
- Takai K, Suzuki M, Nakagawa S, Miyazaki M, Suzuki Y, Inagaki F *et al.* (2006). *Sulfurimonas parvalvinellae* sp. nov., a novel mesophilic, hydrogen- and sulfur-oxidizing chemolithoautotroph within the Epsilon-proteobacteria isolated from a deep-sea hydrothermal vent polychaete nest, reclassification of *Thiomicrospira denitrificans* as *Sulfurimonas denitrificans* comb. nov. and emended description of the genus *Sulfurimonas*. *Int J Syst Evol Microbiol* **56**: 1725–1733.
- Tiirola MA, Rissanen AJ, Sarpakunnas M, Nykänen H. (2011). Stable isotope profiles of nitrogen gas indicate denitrification in oxygen-stratified humic lakes. *Rapid Commun Mass Spectrom* **25**: 1497–1502.
- Tranvik LJ, Downing JA, Cotner JB, Loiselle SA, Striegl RG, Ballatore TJ *et al.* (2009). Lakes and reservoirs as regulators of carbon cycling and climate. *Limnol Oceanogr* **54**: 2298–2314.
- Tulonen T. (1993). Bacterial production in a mesohumic lake estimated from [¹⁴C] leucine incorporation rate. *Microb Ecol* **26**: 201–217.
- Wang Q, Garrity GM, Tiedje JM, Cole JR. (2007). Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* **73**: 5261–5267.
- Whiticar M. (1999). Carbon and hydrogen isotope systematics of bacterial formation and oxidation of methane. *Chem Geol* **161**: 291–314.
- Wilcoxon F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin* **1**: 80–83.

Supplementary Information accompanies the paper on The ISME Journal website (<http://www.nature.com/ismej>)