ORIGINAL ARTICLE

Ubiquity and quantitative significance of bacterioplankton lineages inhabiting the oxygenated hypolimnion of deep freshwater lakes

Yusuke Okazaki[1], Shohei Fujinaga[1], Atsushi Tanaka[2], Ayato Kohzu[2], Hideo Oyagi[3] and Shin-ichi Nakano[1]

[1]Center for Ecological Research, Kyoto University, Otsu, Japan; [2]National Institute for Environmental Studies, Tsukuba, Japan and [3]College of Humanities and Sciences, Nihon University, Tokyo, Japan

The oxygenated hypolimnion accounts for a volumetrically significant part of the global freshwater systems. Previous studies have proposed the presence of hypolimnion-specific bacterioplankton lineages that are distinct from those inhabiting the epilimnion. To date, however, no consensus exists regarding their ubiquity and abundance, which is necessary to evaluate their ecological importance. The present study investigated the bacterioplankton community in the oxygenated hypolimnia of 10 deep freshwater lakes. Despite the broad geochemical characteristics of the lakes, 16S rRNA gene sequencing demonstrated that the communities in the oxygenated hypolimnia were distinct from those in the epilimnia and identified several predominant lineages inhabiting multiple lakes. Catalyzed reporter deposition fluorescence *in situ* hybridization revealed that abundant hypolimnion-specific lineages, CL500-11 (*Chloroflexi*), CL500-3, CL500-37, CL500-15 (*Planctomycetes*) and Marine Group I (*Thaumarchaeota*), together accounted for 1.5–32.9% of all bacterioplankton in the hypolimnion of the lakes. Furthermore, an analysis of single-nucleotide variation in the partial 16S rRNA gene sequence (oligotyping) suggested the presence of different sub-populations between lakes and water layers among the lineages occurring in the entire water layer (for example, acI-B1 and acI-A7). Collectively, these results provide the first comprehensive overview of the bacterioplankton community in the oxygenated hypolimnion of deep freshwater lakes.

The ISME Journal (2017) **11,** 2279–2293; doi:10.1038/ismej.2017.89; published online 6 June 2017

Introduction

The hypolimnion is the dark cold water layer that lies below the thermocline in a thermally stratified lake. In a holomictic (that is, the water layers are not permanently stratified but mixed seasonally) oligo-mesotrophic lake, the hypolimnion remains oxygenated throughout the year, as heterotrophic oxygen demand does not exceed the stock of hypolimnetic oxygen. Although the oxygenated hypolimnion occurs in a limited number of lakes, it is a volumetrically significant realm because of its vertical expanse. For instance, Lake Baikal and the Laurentian Great Lakes together hold 36.6% of the freshwater in the global lakes (Tilzer and Serruya, 1990), and a majority of the water is kept in the oxygenated hypolimnion during stratification, as the average depths of Lake Baikal (740 m) and the Laurentian Great Lakes (53–148 m; except for Lake Erie = 18 m) are much deeper than the depth of the upper mixing layer (epilimnion) (generally < 20 m; Mazumder and Taylor, 1994). In such a lake, accumulation and remineralization of organic matter in the hypolimnion are important biogeochemical processes (Wetzel, 2001). Recent studies have suggested that dissolved organic matter in the oxygenated hypolimnion is enriched by the slowly consumed semilabile fraction (Maki *et al.*, 2010) and can be transformed into a more recalcitrant form by microbes (Thottathil *et al.*, 2013; Hayakawa *et al.*, 2016), as shown by the microbial carbon pump theory proposed for the ocean (Ogawa *et al.*, 2001; Yamashita and Tanoue, 2008; Jiao and Zheng, 2011; Hansell, 2013). Other studies have demonstrated that nitrification (Small *et al.*, 2013), dark carbon fixation (Callieri *et al.*, 2014) and methane oxidation (Murase and Sugimoto, 2005; Bornemann *et al.*, 2016) are also present in the oxygenated hypolimnion. The bacterioplankton inhabiting this realm are responsible for these important biogeochemical processes; thus, their ecophysiological characteristics should be studied.

Bacterioplankton affiliated with the phyla *Actinobacteria*, *Proteobacteria* and *Bacteroidetes* are globally predominant in freshwater systems (Zwart *et al.*, 2002; Newton *et al.*, 2011). However, these data

are based on surface water studies, and members of other phyla could dominate the oxygenated hypolimnion. Among them, the *Chloroflexi* CL500-11 clade (Urbach *et al.*, 2001, 2007; Okazaki *et al.*, 2013; Denef *et al.*, 2016) and the *Thaumarchaeota* Marine Group I (MGI) group (Urbach *et al.*, 2001, 2007; Auguet *et al.*, 2012; Vissers *et al.*, 2013; Coci *et al.*, 2015; Mukherjee *et al.*, 2016) are the most investigated. The relatively large cell size (1–2 µm) and high abundance (>15% of all bacteria and archaea) of CL500-11 suggest their quantitative importance in the oxygenated hypolimnion. MGI accounted for 8.7–19% of all bacterioplankton in the oxygenated hypolimnion of subalpine lakes (Callieri *et al.*, 2016). They are ammonia-oxidizing archaea, and members in the oxygenated hypolimnion are affiliated with either *Nitrosopumilus* or *Candidatus* Nitrosoarchaeum (Coci *et al.*, 2015). Other nitrifiers such as *Nitrosospira* and *Nitrospira* have also been found in the water column of the oxygenated hypolimnion (Fujimoto *et al.*, 2016; Mukherjee *et al.*, 2016; Okazaki and Nakano, 2016). Moreover, high-throughput sequencing of the 16S rRNA gene partial amplicon has highlighted members of *Planctomycetes* (for example, CL500-3, CL500-15 and CL500-37) as abundant lineages in the oxygenated hypolimnion (Rozmarynowycz, 2014; Okazaki and Nakano, 2016). These inhabitants are not negligible components of the microbial food web and biogeochemical cycling in the deep pelagic freshwater ecosystem. However, it remains unknown how ubiquitously and abundantly they are distributed in the oxygenated hypolimnion. Owing to a lack of quantitative data, their ecological importance remains poorly understood.

The present study investigated the bacterioplankton community in the oxygenated hypolimnia of 10 deep freshwater lakes with a variety of geochemical characteristics. Community composition was investigated by 16S rRNA gene amplicon sequencing, and several representative members were microscopically characterized and quantified by catalyzed reporter deposition fluorescence *in situ* hybridization (CARD-FISH). Taken together with data previously collected from other deep lakes, the results allowed us to create the first comprehensive overview of the bacterioplankton community inhabiting the oxygenated hypolimnion of deep freshwater lakes and identify abundant and ubiquitous lineages. Moreover, analyses of their habitat preference and microdiversification (oligotyping) facilitated hypotheses about their ecophysiological characteristics and potential diversified sub-populations.

Materials and methods

Field sampling
Water samples were collected at pelagic sites in 10 deep freshwater lakes in Japan from August to
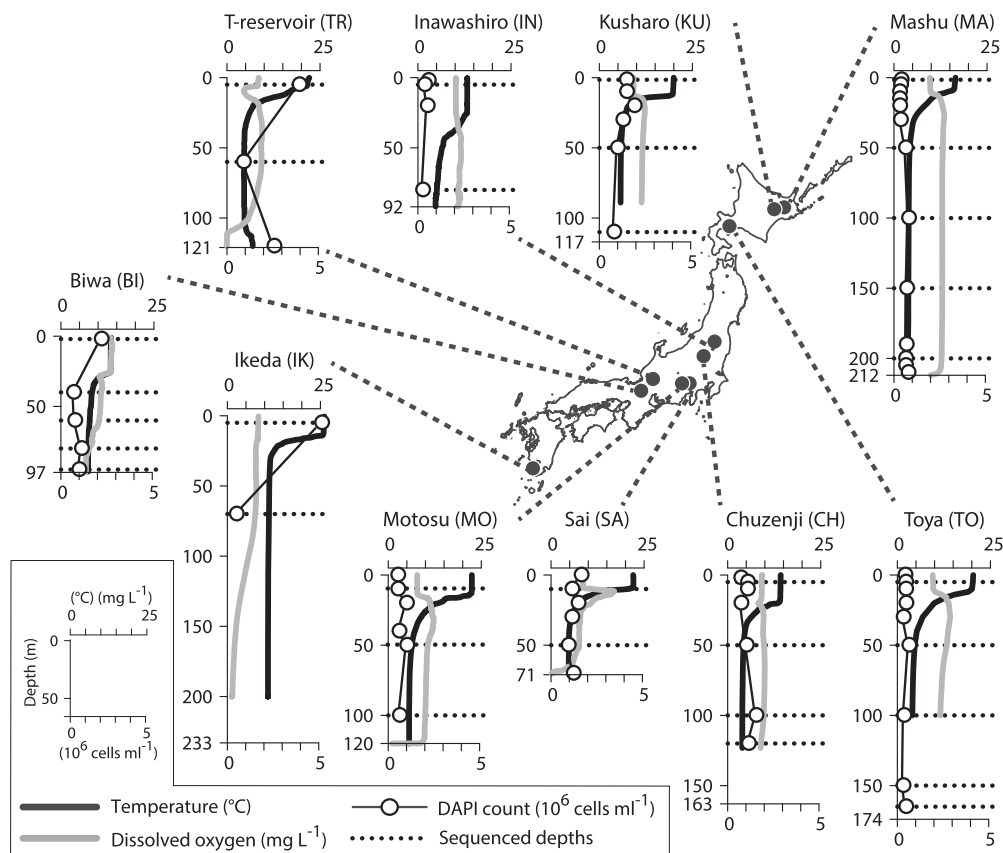

Figure 1 Locations and vertical profiles (water temperature, dissolved oxygen and DAPI-positive cells) of the sampling sites.

December in 2015, including Lake Mashu, Kusharo, Toya, Inawashiro, Chuzenji, Sai, Motosu, Biwa, Ikeda and T-Reservoir (hereafter referred as MA, KU, TO, IN, CH, SA, MO, BI, IK and TR, respectively) (Figure 1). Given that the total water volume in Japanese lakes is estimated to ca. 90 km^3 (Mori and Sato, 2015), the studied 10 lakes (total water volume = 50.0 km^3) comprise more than a half of the lake volume in the country. The profiles of the sampling locations are summarized in Table 1. In all, 3 to 13 depths were sampled in each lake, and the temperature and dissolved oxygen vertical profiles were measured using a conductivity, temperature, depth (CTD) probe *in situ*. These lakes have permanently oxygenated hypolimnia and were thermally stratified when the sampling was carried out (Figure 1). Total prokaryotic abundance was determined by microscopic enumeration of 4',6-diamidino-2-phenylindole (DAPI)-stained cells (Porter and Feig, 1980) with at least 900 cells counted in each filter.

Partial 16S rRNA gene amplicon sequencing
In total, 33 DNA samples were collected (Figure 1). Immediately after sampling, 70–300 ml the water collected was filtered through a 0.2 μm polycarbonate filter (47 mm diameter; Whatman, Maidstone, UK). The filter was stored at −20 °C until DNA was extracted using the PowerSoil DNA Isolation Kit (MoBio Laboratories, Carlsbad, CA, USA). The V4 and V5 regions of the 16S rRNA gene were amplified using the modified 530F and 907R primers (Nunoura *et al.*, 2009) and then an 8-bp DNA tag (for postsequencing sample identification) and 454 adaptors were conjugated by second PCR, as described previously (Okazaki and Nakano, 2016). The PCR products from the samples were pooled in equimolar quantities and sequenced in the 1/8 regions of a sequencing reaction on the Roche 454 GS-FLX Titanium Sequencer (Roche Science, Mannheim, Germany). The nucleotide sequence data are available in the Sequence Read Archive database under accession numbers DRX062810–DRX062842 (BioProject: PRJDB5151).

Analysis of sequencing reads
The sequence data were processed using the UPARSE pipeline (Edgar, 2013) following the author's instructions (http://www.drive5.com/usearch/manual/upp_454.html). We used a fastq_maxee value of 1.0, a truncated length of 350 bp and an operational taxonomic unit (OTU) creation identity threshold of 97%. Thereafter, respective OTUs were taxonomically assigned with the SINA 1.2.11 (Pruesse *et al.*, 2012) online tool (https://www.arb-silva.de/aligner/) referring to the 123 Ref database (Quast *et al.*, 2013) and using the default parameter settings. Subsequently, non-prokaryotic OTUs (that is, chloroplast, eukaryote and unclassified domains) were removed. The resulting 96 149 reads were used for subsequent analyses.

Before calculating community diversity, coverage-based rarefaction (Chao and Jost, 2012) was applied. Reads were randomly discarded from each sample until coverage was <96.59% (that is, slope of the rarefaction curve was >0.0341), which was the minimum value recorded among samples. Subsequently, α- (the inversed Simpson index) and β-diversities (non-metric multidimensional scaling based on Bray–Curtis dissimilarity) were calculated. These analyses were performed using the vegan 2.4-0 package (Oksanen *et al.*, 2016) in the R 3.3.1 software (http://www.R-project.org/).

The OTUs were classified following the nomenclature proposed by Newton *et al.* (2011). Using the '--search' option in the SINA 1.2.11 (Pruesse *et al.*, 2012) stand-alone tool, the representative sequences of the respective OTUs (generated by the UPARSE pipeline) were searched and classified against the original ARB (Ludwig *et al.*, 2004) database provided by Newton *et al.* (2011). Dominant OTUs that failed to be classified by this procedure were named by the following schemes: if closely related (>99% identity) sequences with the prefixes 'CL' (Urbach *et al.*, 2001) or 'LiUU-' (Eiler and Bertilsson, 2004) were in the public sequence database, their names were preferentially used. The *Planctomycetes* phylogenetic clades were newly defined for the collective descriptions (Supplementary Figure S1). In other cases, the OTU was shown by its taxonomic affiliation (for example, genus) based on the SILVA nomenclature.

The sequences were aligned using the SINA 1.2.11 (Pruesse *et al.*, 2012) stand-alone tool against the SILVA 123 Ref. NR 99 database to construct a phylogenetic tree (Quast *et al.*, 2013). Maximum-likelihood trees were calculated by the RAxML 8.2.4 software (Stamatakis, 2014) with the GTR substitution model and the GAMMA rate heterogeneity model. The trees were drawn in the MEGA 7 software (Kumar *et al.*, 2016).

Oligotyping
Intra-OTU microdiversity of representative lineages was analyzed by oligotyping, which facilitated detection of single-nucleotide variation by excluding the effects of sequencing errors based on the Shannon entropy values (Eren *et al.*, 2013). Following the author's instructions (http://merenlab.org/software/oligotyping/), the quality filtered FASTA file generated by the UPARSE pipeline was split at individual OTUs and aligned against the Greengenes (DeSantis *et al.*, 2006) alignment database using the SINA 1.2.11 (Pruesse *et al.*, 2012) stand-alone tool. Uninformative columns were removed by subsequently applying the 'o-trim-uninformative-columns-from-alignment' and 'o-smart-trim' scripts. Several rounds of oligotyping were repeated by manually choosing the most informative (that is,

Table 1 Profiles of the sites studied

	Max. depth (m)	Depth of sampling point (m)	Surface area (km²)	Water volume (10⁶ m³)	Surface altitude (m)	Lake origin	Water retention time (year)	Trophic status	pH[a]	Total nitrogen (mg l⁻¹)[a]	Total phosphorus (mg l⁻¹)[a]	Sampling location	Sampling date
Mashu (MA)	211	211	19.2	2640	351	Caldera	200	Oligotrophic	7.0–8.0	0.06–0.08	<0.003–0.004	43°35′16.0″N, 144°31′50.0″E	2015.8.26–29
Kussharo (KU)	118	117	79.6	2261	121	Caldera	12	Oligotrophic	7.0–7.5	0.07–0.24	<0.003–0.004	43°35′26.0″N, 144°19′48.0″E	2015.8.25
Toya (TO)	180	174	70.7	8272	84	Caldera	9	Mesotrophic	7.1–7.8	0.24–0.38	0.003–0.01	42°36′15.1″N, 140°49′07.7″E	2015.9.24
Inawashiro (IN)	93.5	92	103.3	5320	514	Tectonic	5.4	Oligotrophic	6.6–7.0	0.17–0.34	<0.003–0.004	37°28′27.6″N, 140°05′40.5″E	2015.10.30
Chuzenji (CH)	163	163	11.8	1116	1269	Dammed	5.9	Oligotrophic	7.7–8.8	0.07–0.27	<0.003–0.009	36°43′55.7″N, 139°28′25.9″E	2015.10.21
Sai (SA)	71.7	71	2.1	81	900	Dammed	1.6	Oligotrophic	7.4–8.8	0.06–0.32	0.003–0.008	35°29′49.5″N, 138°41′09.5″E	2015.9.8
Motosu (MO)	122	120	4.7	319	900	Dammed	6.5	Oligotrophic	6.8–7.9	0.15–0.25	<0.003–0.006	35°28′06.5″N, 138°34′57.8″E	2015.9.7
T-Reservoir (TR)	121	121	13	660	390	Artificial reservoir	0.27	Mesotrophic	7.2–9.2	0.18–0.46	<0.003–0.008	35°40′26.0″N, 136°29′57.0″E	2015.9.16
Biwa (BI)	104	97	670.3	27616	85	Tectonic	5.5	Mesotrophic	7.7–8.4	0.15–0.32	0.005–0.009	35°23′21.0″N, 136°07′51.0″E	2015.12.9
Ikeda (IK)	233	233	10.9	1368	66	Caldera	33.5	Mesotrophic	7.1–8.6	0.14–0.32	<0.003–0.006	31°14′14.0″N, 130°33′52.0″E	2015.9.17

Data were collected from Mori and Sato (2015) and the Japanese Ministry of the Environment public water database.
[a]Annual range recorded in the epilimnion.

Table 2 CARD-FISH probes, helpers and formamide concentrations (at 46 °C) used in this study

Probe (helper)	Sequence (5'–3')	Length (base)	Formamide (%)	Target	Reference
CLGNS_584	GCCGACTTGCCCAACCTC	18	20	CL500-11	Okazaki et al. (2013)
CLGNS_567h	CTACACGCCCTTTACGCC	18			
CL500-3_829	CGTTCAACCCCCTCGACC	18	20	CL500-3	This study
CL500-3_847h	ATCAGTTTTCCTACGGCCGAGAAGG	25			
CL500-3_804h	AAGTGCGCATCGTTTACGGCGTGGA	25			
CL500-37_658	CGTTCCACCCGCCTCTACCTG	21	30	CL500-37	This study
CL500-15_826	ACAGGGAGACCCCAAACTAG	20	30	CL500-15	This study
CL500-15_846h	TCACTTTTGTTTCGCCTGGGATCCC	25			
CL500-15_801h	TGCCCATCGTTTACAGCTAGGACTA	25			
MGI_535	TCCTGACCACTTGAGGTGCTGG	22	25	MGI	Coci et al. (2015)

Helpers are indicated by the suffix 'h' in their names.
Abbreviations: CARD-FISH, catalyzed reporter deposition fluorescence *in situ* hybridization; MGI, Marine Group I.

the highest entropy) column until all oligotypes with >100 reads exceeded the purity score of 0.90. The minimum substantive abundance parameter (option-M) was always set to 10.

CARD-FISH

CARD-FISH was performed based on Pernthaler et al. (2002) with some modifications as described previously (Okazaki et al., 2013). Specifically, the hybridization, amplification and washing steps were carried out at 46 °C. AlexaFluor 488 (Life Technologies, Carlsbad, CA, USA) was used as the tyramide-conjugated fluorescent dye. The oligonucleotide probes were designed previously for CL500-11 (Okazaki et al., 2013) and MGI (Coci et al., 2015), and newly designed for the other targets (Table 2). The probes were constructed using the 'Design Probes' function in the ARB 6.0.3 software (Ludwig et al., 2004) against the SILVA 123 Ref. NR 99 database (Quast et al., 2013). Specificity of the probes was confirmed by an NCBI BLAST online search and the Test Probe 3.0 tool against the SILVA 123 Parc database (https://www.arb-silva.de/search/testprobe/). The probes designed in this study targeted the V4 or V5 regions of the 16S rRNA gene, which were targeted by amplicon sequencing. Since there could be probe-matching (false-positive) sequences that are not present in the database, we confirmed that the probes perfectly and exclusively match their target lineages in all our sequenced reads. To enhance the fluorescent signal, oligonucleotide helpers (Fuchs et al., 2000) were used for several probes (Table 2). The helpers were designed to target the adjacent or opposite loci of the probe target site to loosen the secondary rRNA structure (Fuchs et al., 2000) by confirming that all probe targets in the database were not mismatched with their corresponding helpers. The hybridization buffer contained 0.5 μg ml^{-1} probe and 0.1 μg ml^{-1} of each helper. The formamide concentration in the buffer was determined for each probe (Table 2) by testing a series of concentrations to obtain the best stringency (that is, highest concentration without signal loss). The stringency test was performed in samples with a high read proportion of the target determined by amplicon sequencing. The hybridized filters were counterstained with DAPI and enumerated under an epifluorescence microscope. At least 300 DAPI-positive cells and the corresponding CARD-FISH-positive cells were enumerated three times per sample (the same filter piece). A negative NON338 probe control (Wallner et al., 1993) confirmed that no false-positive cells were present.

Results

16S rRNA gene partial amplicon sequencing generated 666 OTUs from the 96 149 reads originating from the 33 samples collected from the water columns of the 10 lakes. The complete data set generated by the pipeline is available in Supplementary Data 1 and 2.

At the phylum-level phylogenetic resolution, *Actinobacteria* dominated throughout the water column, *Bacteriodetes* and *Betaproteobacteria* showed preference to the epilimnion, and *Chloroflexi* and *Planctomycetes* preferentially occurred in the hypolimnion (Figure 2 and Supplementary Figure S2). The epilimnetic and hypolimnetic communities were compared at the OTU level by averaging the samples from each layer (Figure 2). The dominant members in the epilimnion were generally shared between the lakes (for example, acI-B1, acI-A7, Lhab-A1 and bacI-A1). In addition to the lineages common to the epilimnion, *Chloroflexi* and *Planctomycetes* were also ranked as dominant OTUs in the hypolimnion (Figure 2). CL500-11 alone accounted for most of the *Chloroflexi* reads, whereas *Planctomycetes* consisted of diverse OTUs (for example, CL500-3, CL500-15, CL500-37 and plaI-A) (Figure 2), which were affiliated with three classes in the phylum (Supplementary Figure S1). The α-diversity was higher in deeper samples in the lakes, except for KU, CH and BI (Figure 2). The β-diversity analysis (NMDS) clearly separated the hypolimnetic communities from the epilimnetic ones (Supplementary Figure S3).

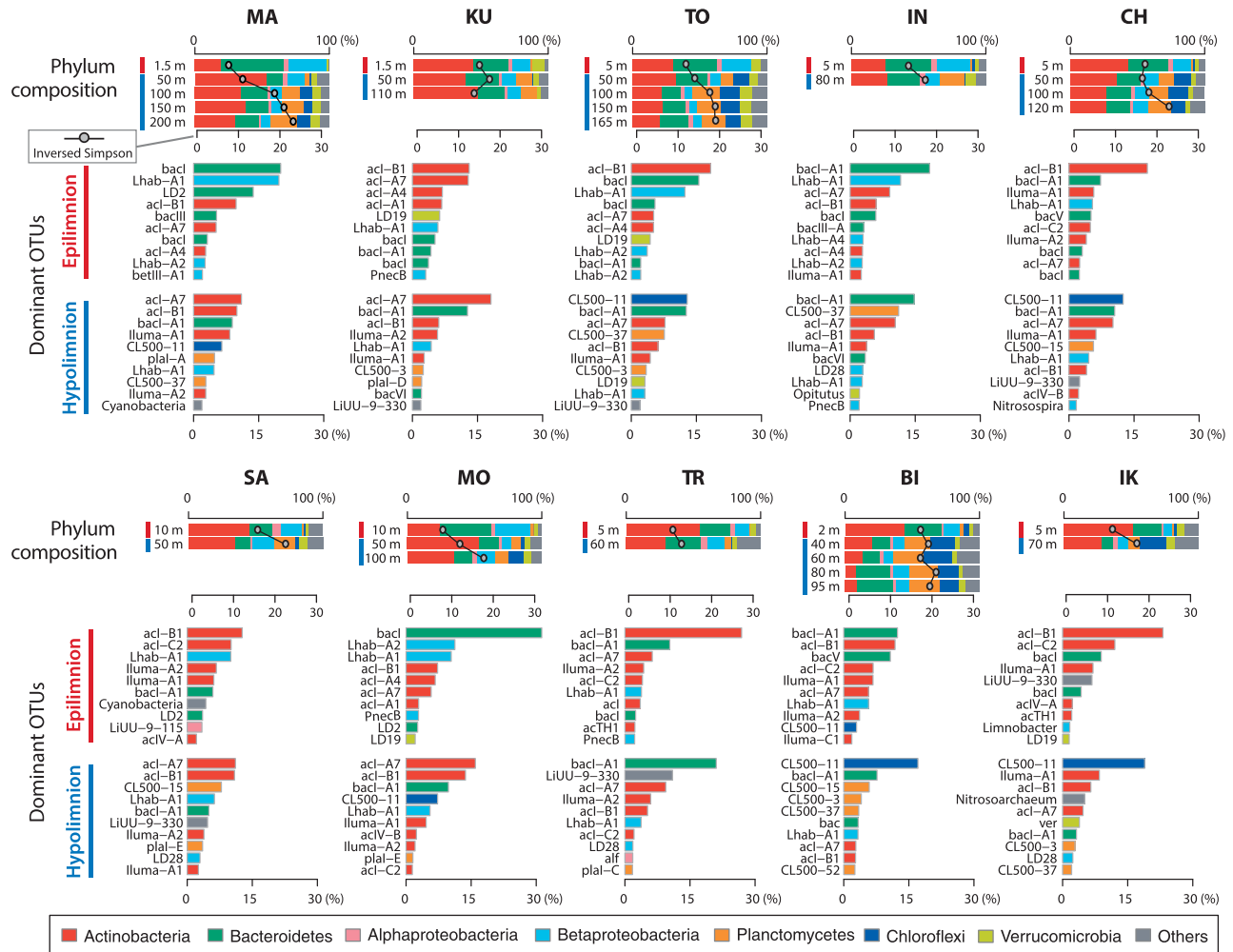Vertical preference of the bacterioplankton in each OTU was examined based on the number

Figure 2 Composition of the 16S rRNA gene amplicon reads. The top panel (band graphs) displays phylum-resolved community composition at each depth for each lake, with an overwritten line graph indicating the α-diversity (the inversed Simpson index). The two lower panels show the 10 dominant OTUs in the epilimnion and hypolimnion, composed of averaged data for each layer (the depths averaged are illustrated by red and blue lines in the top panel). Bar colors indicate phyla to which individual OTUs were assigned.

of lakes where it accounted for >1% of all amplicon reads in each water layer (Figure 3). This facilitated to separate epilimnion inhabitants, hypolimnion inhabitants and whole-layer inhabitants on the basis of the occurrence patterns (Figure 3). Results produced by other abundance thresholds (>0.5% and >2%) are shown in Supplementary Figure S4. Twenty-nine OTUs that were ubiquitous or specific to the hypolimnion were selected for further analysis (indicated by a gray dotted box in Figure 3). While they showed preferences between the lakes, only few environmental parameters significantly explained the patterns (Figure 4). The oligo-typing analysis successfully detected conserved single-nucleotide intra-OTU variants (that is, base positions with high entropy, which are not likely sequencing errors) (Figure 4), and revealed that there were oligotypes that disproportionally distributed among the water layers or lakes (Figure 5, Supplementary Figure S5 and Supplementary Data 4).

The newly constructed CARD-FISH probes targeted a monophyletic clade of the target lineages (Supplementary Figure S6). Enumeration of the positive cells revealed that CL500-11 accounted for >10% of all prokaryotic cells in four lakes (maximum was 25.9% at 60 m in BI). CL500-3, CL500-37, and MGI respectively accounted for >3% in several lakes. CL500-15 were less abundant but still detectable, with a maximum percentage of 1.6% (Figure 6). The cells detected in each target shared identical morphology between the lakes: CL500-11 were curved rods 1–2 μm long; CL500-3, CL500-37, and CL500-15 were cocci approximately 1 μm diameter; and MGI were rods around 1 μm long (Figure 6). The cells were mostly planktonic, except for CL500-15, which were often found to be particle-associated; At 50 m in SA, 51% of the CL500-15 cells were particle-associated (determined by enumerating at least 300 positive cells for three times) (Figure 6). The correlation between relative abundance determined by amplicon sequencing and CARD-FISH was
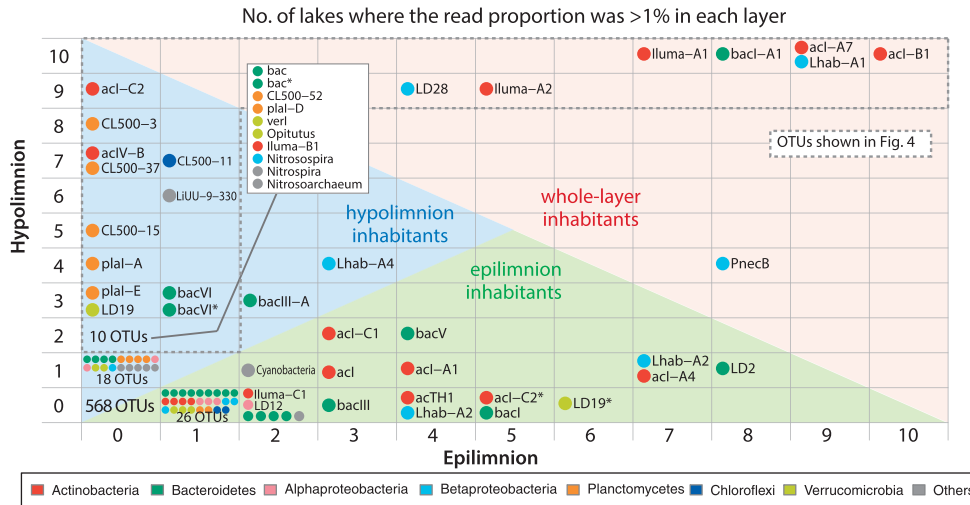
Figure 3 Vertical preferences of individual OTUs, mapped by the number of lakes where individual OTUs accounted for >1% of all amplicon reads in each water layer. Data for the hypolimnion were generated by averaging the data at multiple depths in the hypolimnion. The OTUs were categorized into three groups (epilimnion-, hypolimnion- and whole-layer inhabitants) based on their distribution patterns. A gray dotted box highlight the 29 OTUs selected for further analyses shown in Figure 4. Point color indicates the phylum to which an individual OTU was assigned. Asterisks in the group name distinguish the different OTUs assigned to the same group. Results produced by other abundance thresholds (>0.5 and >2%) are shown in Supplementary Figure S4.

significant for all targets, but the CARD-FISH estimates tended to be lower than those of amplicon sequencing (Supplementary Figure S7). The ratio of CARD-FISH to amplicon sequence estimates was 0.57 on average (range, 0.14–1.02) for the five targeted lineages (Supplementary Figure S7). The proportions of the respective targets in each sample (that is, raw data for Figure 6) are available in Supplementary Data 5.

Discussion

Structure of the bacterioplankton community in the oxygenated hypolimnion
The present study investigated lakes with a variety of environmental characteristics, ranging from a oligotrophic lake with the cold (4 °C) hypolimnion (MA) to a mesotrophic lake with the relatively warm (11 °C) hypolimnion (IK) (Figure 1 and Table 1). Nevertheless, the β-diversity analysis indicated that the bacterioplankton communities in the oxygenated hypolimnia were separated from the communities in the epilimnia (Supplementary Figure S3), indicating that the oxygenated hypolimnion is an independent habitat for microbial communities. The hypolimnetic community consisted of members present across the water layers (whole-layer inhabitants), and members that preferentially occurred in the oxygenated hypolimnion (hypolimnion inhabitants) (Figure 3). The whole-layer inhabitants were composed of several commonly known freshwater bacterioplankton lineages (for example, acI, Iluma-A1, Iluma-A2, Lhab-A1, LD28 and bacI-A1), whereas the hypolimnion inhabitants were represented by phyla that were not common to the epilimnion, including *Chloroflexi* CL500-11, members of

Planctomycetes (for example, CL500-3, CL500-15, CL500-37 and plaI-A), and *Ca.* Nitrosoarchaeum in *Thaumarchaeota* (Figures 2 and 3). Their abundance and ubiquity in the hypolimnion of the studied lakes were demonstrated by the CARD-FISH analysis, showing that they collectively accounted for 1.5% (TR) to 32.9% (BI) of all bacterioplankton in the hypolimnion (Figure 6 and Supplementary Data 5). Our analysis also revealed that these groups were not always present in the studied lakes and were absent in some lakes (for example, IN and KU for CL500-11, CH for CL500-3 and MO for CL500-37) (Figures 3, 4 and 6), suggesting that they have respective habitat preferences.

Ubiquity, quantitative importance and potential ecophysiology of the hypolimnion inhabitants
The predominance (~25% of all bacterioplankton) of planktonic CL500-11 cells in several lakes (Figure 6) suggests that their resources are diffuse, abundant and ubiquitous. The metagenome-assembled genome and *in situ* transcriptional evidence of CL500-11 in Lake Michigan suggests their importance in peptide turnover (Denef *et al.*, 2016). Peptides in aquatic systems are mainly derived from peptidoglycans in the bacterial cell wall (Nagata *et al.*, 2003; McCarthy *et al.*, 2013) and from proteins released by other bacteria (Tanoue *et al.*, 1995) or phytoplankton (Nguyen and Harvey, 1997; Yamada *et al.*, 2012). Previous studies in BI have collectively demonstrated that N-rich (by stoichiometry) or protein-like (by fluorescence properties) semilabile dissolved organic matter derived from autochthonous phytoplankton production that accumulates in the hypolimnion is slowly remineralized during stratification (Kim *et al.*, 2006; Maki *et al.*, 2010; Thottathil *et al.*, 2013).
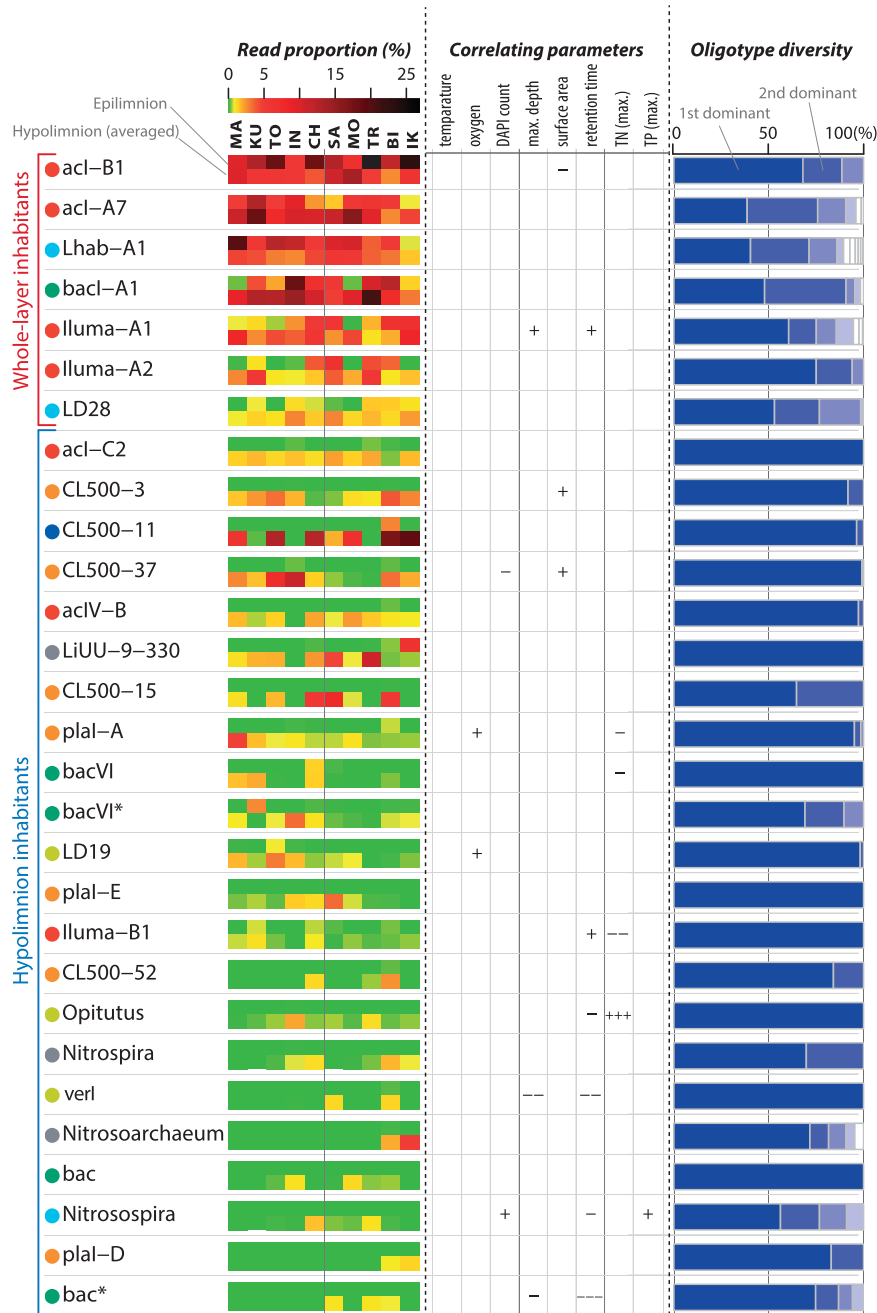
Figure 4 Distribution patterns (left column), correlating environmental parameters (center) and oligotype diversity (right) of the whole-layer inhabitants and the hypolimnion inhabitants. The distribution pattern is illustrated by the read proportion to total amplicon reads. The hypolimnion data were generated by averaging the data at multiple depths in the hypolimnion. The correlation between read proportion and environmental parameters in the hypolimnion was evaluated by Spearman's test. For a positive correlation, $^{+++}P<0.005$; $^{++}P<0.01$; $^{+}P<0.05$. For a negative correlation, '−' was shown instead of '+'. For TN and TP, maximum values recorded in the lake (Table 1) were used for the correlation analysis. The right column (band graphs) indicates composition of the oligotypes among all amplicon reads assigned to each OTU in the present study. Point colors indicate phyla to which individual OTUs were assigned (See Figures 2 and 3 for legend). Asterisks in the group name distinguish the different OTUs assigned to the same group.

Thus, it is possible that CL500-11 is scavenging protein-like debris that accumulates in the lake because of its relatively recalcitrant nature. Given that the water retention time of a system affects dissolved organic matter composition (Kellerman et al., 2014; Catalán et al., 2016) and that autochthonous dissolved proteins can accumulate in a lake even at centennial timescales (Goldberg et al., 2015), we expected that lakes with a longer water retention time would contain more bacterioplankton lineages specialized to consume relatively recalcitrant dissolved organic matter. In the present study, the water retention time of the lakes ranged from 0.27 years (TR) to 200 years (MA) (Table 1). However, most of the
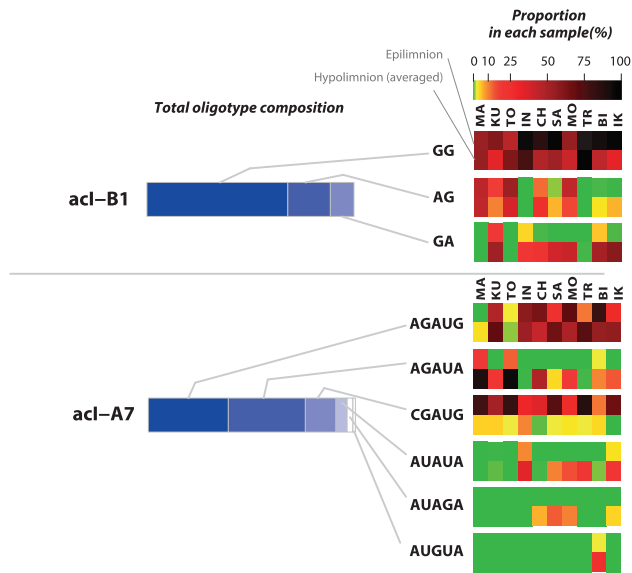
Figure 5 Composition and distribution of oligotypes for two abundant whole-layer inhabitants, acI-B1 and acI-A7. The left bars indicate the total oligotype composition in the present study (data shown in Figure 4), whereas right panels indicate the proportion of the respective oligotypes in each sample. The complete data set including all OTUs is available in Supplementary Figure S5 and Supplementary Data 4.

hypolimnion inhabitants, including CL500-11, were not distributed in a manner that was associated with water retention time (Figure 4). As CL500-11 is a large cell (Figure 6), protistan size-selective grazing (Pernthaler, 2005) may be a factor controlling CL500-11 dynamics, and little information is available on the grazer communities inhabiting the deep oxygenated hypolimnion (Masquelier et al., 2010; Mukherjee et al., 2015). These assumptions should be verified by future studies, given their ubiquity and quantitative importance in deep freshwater systems. Indeed, the dominance of CL500-11 has been reported in the two largest deep freshwater systems on Earth, the Laurentian Great Lakes (Rozmarynowycz, 2014; Denef et al., 2016) and Lake Baikal (Kurilkina et al., 2016).

CL500-3 and CL500-37 were the two most abundant *Planctomycetes* in the present study (Figure 6). Each was affiliated with their respective phylogenetic clade in the class *Phycisphaerae* (Supplementary Figure S1). Aquatic *Planctomycetes* are often associated with algal blooms (Morris et al., 2006; Pizzetti et al., 2011), and genomic evidence indicates their potential to aerobically consume sulfated polysaccharides derived from algae (Glöckner et al., 2003; Woebken et al., 2007; Erbilgin et al., 2014). In a marine study, sequences closely related to CL500-3 were enriched in DNA extracted from bacterioplankton that assimilate protein secreted by phytoplankton (Orsi et al., 2016). Consequently, it can be hypothesized that CL500-3 and CL500-37 consume polysaccharides or proteins derived from phytoplankton. The two

lineages did not always co-occur, and disproportional dominance of CL500-3 was found in MO and TR, and of CL500-37 in CH and IN (Figure 6). Amplicon sequencing data taken from the oxygenated hypolimnion in Lake Michigan (Fujimoto et al., 2016) showed that only CL500-37 was abundant (Supplementary Figure S8). These observations indicate that the ecological niches of CL500-3 and CL500-37 are not the same. Given that algal exudates from different phytoplankton species select different bacterial communities (Šimek et al., 2011; Paver et al., 2013), the difference might be attributable to differences in the phytoplankton species in a lake. It should also be noted that their closely related sequences were not necessarily retrieved from the oxygenated hypolimnion and have been found in an Antarctic lake (Karlov et al., 2017), arctic lake (Ntougias et al., 2016) and Baltic Sea ice (Eronen-Rasimus et al., 2015) (Supplementary Figure S6A). More information is needed to elucidate the ecophysiological characteristics of these widespread and abundant yet largely overlooked bacterial lineages.

The CL500-15 clade belonged to the uncultured OM190 class (Supplementary Figure S1) with only three sequences in the database: two from deep freshwater lakes (Urbach et al., 2001; Pollet et al., 2011) and one from deep sea sediment (Zhang et al., 2013) (Supplementary Figure S6B). Another sequence was reported from the littoral water of Lake Baikal (Parfenova et al., 2013). In the present study, CL500-15 was detected in half of the lakes (Figure 3), indicating that they are one of the most common lineages in the oxygenated hypolimnion. The CARD-FISH analysis revealed a high proportion of particle-associated cells and microscopic observations revealed that the particles were not cells but rather looked like transparent exopolymer particles, which are gel-like sticky particles mainly composed of polysaccharides (Passow, 2002). Members of the OM190 class in marine systems have been preferentially detected in the particle-associated fraction (Bižić-Ionescu et al., 2015; Salazar et al., 2015), suggesting that their particle-associated form is preserved across the class. As particle-associated bacteria can contribute disproportionally to total bacterial activity (Lemarchand et al., 2006; Grossart et al., 2007), CL500-15 may have a substantial role in substrate remineralization in the oxygenated hypolimnion.

Other representative *Planctomycetes* were affiliated with class *Planctomycetacia* (plaI-A–F) (Supplementary Figure S1). They usually had a smaller proportion of reads than those of CL500-3, CL500-37, and CL500-15 (Figure 2). However, plaI-A was the most represented *Planctomycetes* based on the read proportion in the oxygenated hypolimnion of MA (Figure 2) and Lake Michigan (Supplementary Figure S8). Each of the other members of plaI (for example, plaI-B [CL500–52], plaI-D, and plaI-E) showed their respective distribution patterns between the lakes (Figure 4). Overall, *Planctomycetacia* (plaI group) was generally less ubiquitous and
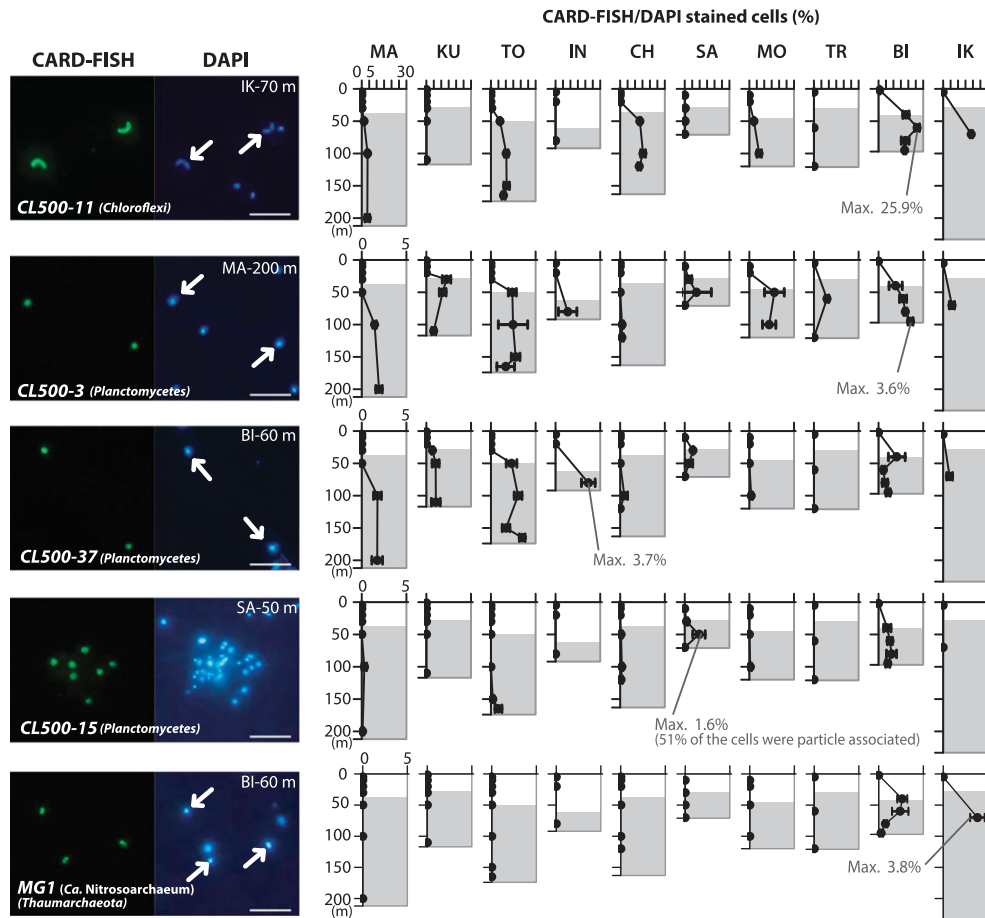
Figure 6 CARD-FISH images and the enumeration results. Positive cells and the corresponding DAPI-stained image are shown in each micrograph. Arrows in the DAPI-image indicate cells with CARD-FISH-positive signals. Scale bar = 5 μm. The line graphs show percentages of CARD-FISH-positive cells to DAPI-positive cells, and the error bars indicating the standard deviation determined by triplicate enumeration of an identical filter. The maximum value recorded for each lineage is designated. Gray background illustrates the hypolimnion (that is, below the thermocline). The raw data are available in Supplementary Data 5.

abundant, but more diverse than the other two classes (that is, *Phycisphaerae* and OM190) inhabiting the oxygenated hypolimnion.

In the present study, the MGI group was detected only from two lakes (BI and IK) (Figure 4), and their maximum percentage determined by CARD-FISH was 3.8% (Figure 6). These numbers are lower than those of a previous study that detected 8.7–19% MGI in the oxygenated hypolimnion of all six subalpine lakes investigated (Callieri *et al.*, 2016). In the present study, MGI was exclusively affiliated with *Ca.* Nitrosoarchaeum (Blainey *et al.*, 2011), but *Nitrosopumilus*, another predominant MGI member in the oxygenated hypolimnion (Berdjeb *et al.*, 2013; Vissers *et al.*, 2013; Coci *et al.*, 2015), was not detected. *Ca.* Nitrosoarchaeum has been reported in the oxygenated hypolimnion of Crater Lake (Urbach *et al.*, 2001, 2007), Lake Redon (Auguet *et al.*, 2012) and Lake Superior (Mukherjee *et al.*, 2016), which are oligotrophic lakes with a hypolimnetic temperature of 4 °C. The occurrence of *Ca.* Nitrosoarchaeum in BI and IK, which are mesotrophic lakes with the hypolimnetic temperatures of 8 and 11 °C,

respectively (Figure 1), revealed their broad habitat spectrum. In other lakes (for example, CH and TR), ammonia-oxidizing bacteria, *Nitrosospira*, were detected in the oxygenated hypolimnion (Figure 4), in agreement with the previously suggested niche separation between ammonia-oxidizing archaea and bacteria (Coci *et al.*, 2015; Mukherjee *et al.*, 2016). Another nitrifier, *Nitrospira*, was also detected in the oxygenated hypolimnion (Figure 4), in line with previous studies (Small *et al.*, 2013; Okazaki and Nakano, 2016). However, no nitrifiers were detected in the three northern lakes (MA, KU and TY) (Figure 4). Notably, their absence should not be concluded by the present data, as the nitrifier community can shift over seasons (Okazaki and Nakano, 2016). Nevertheless, our results indicate the potential diversity of the nitrification systems in the oxygenated hypolimnion, yet information remains scarce to conclude the cause and effects of the diversity.

The discussion above is based on the assumption that each of the bacterioplankton lineages prefer their particular suitable habitat. However, it is also

possible that the occurrence of the members is controlled by occurrence of other lineages, as bacterioplankton often have streamlined genomes and are dependent on each other for lost metabolic functions (Morris et al., 2012; Garcia et al., 2015; Mas et al., 2016). In the present study, many pairs of hypolimnion specialists showed correlating distribution patterns between the lakes (Supplementary Figure S9). For example, CL500-3 was positively correlated with CL500-37 and Ca. Nitrosoarchaeum, and negatively correlated with Nitrosospira (Supplementary Figure S9). Although these results do not directly support an interaction, they suggest that some hypolimnion specialists are presumably co-occurring or sharing similar ecological niches.

Notable but less represented lineages

Although several lineages originally described for Crater Lake (with the prefix 'CL') (Urbach et al., 2001) were identified as representative hypolimnion inhabitants (Figures 3 and 4), several other members dominant in Crater Lake were not highly represented in our study: CL120-10 of Verrucomicrobia (OTU65 in the present study; see Supplementary Data 1), CL0-1 of Armatimonadetes (OTU77) and CL500-9 of Chloroflexi (not detected in the present study). It is plausible that further investigations in lakes on other continents or those with depths > 250 m will detect bacterioplankton not found in the present study. Methanotrophs, such as Methylococcaceae and Methylocystaceae, also accounted for a very minor fraction of all amplicon reads (Supplementary Data 1). A more intriguing result is the limited distribution and low relative abundance of LD12 (Figures 2 and 3), which is one of the most dominant and ubiquitous freshwater bacterioplankton (Zwart et al., 1998; Newton et al., 2011; Salcher et al., 2011). Notably, data produced using another sequencing platform (Miseq) from a part of the DNA samples used in the present study showed a higher read proportion of LD12 (Fujinaga et al., unpublished data), despite the fact that both analysis used PCR primers that perfectly matched the LD12 16S rRNA gene sequence. Although the reason behind this discrepancy is unknown, it is possible that the relative abundance of LD12 was underestimated in the present study, which, in turn, might have overestimated other lineages among the reads, as in the discrepancy with the CARD-FISH results (Supplementary Figure S7). Direct cell enumeration using FISH should be considered an accurate abundance estimate.

Intra-OTU diversification revealed by the oligotyping analysis

The oligotyping analysis revealed the intra-OTU diversification of the whole-layer inhabitants (Figure 4 and Supplementary Figure S5). Some of the oligotypes were disproportionally distributed among depths or lakes (Figure 5 and Supplementary Data 4), suggesting that their ubiquity was achieved collectively by heterogeneous oligotypes that specialized in a respective niche. Such cryptic microdiversification within a ubiquitous freshwater lineage with an almost identical 16S rRNA gene sequence has previously been reported in Limnohabitans (Jezbera et al., 2013; Kasalický et al., 2013) and Polynucleobacter (Jezbera et al., 2011; Hahn et al., 2015, 2016). These studies reported diversification between habitats with different temperatures, pHs, organic and inorganic substrate availability, and geography. Recently, a horizontal oligotype profile in Lake Michigan indicated oligotype diversification within predominant bacterioplankton lineages between estuarine and pelagic sites (Newton and McLellan, 2015). The present study discovered microdiversification between the epilimnion and hypolimnion, suggesting the presence of hypolimnion-specific sub-populations within the common freshwater bacterioplankton (for example, acI-B1, acI-A7; Figure 5). Since oligotypes may not always reflect ecologically or evolutionarily cohesive populations (Berry et al., 2017), further comparative studies of individual oligotypes are required to clarify their diversification to the respective water layers, which differ considerably in temperature, substrate availability and grazing pressure.

The low oligotype diversity of several hypolimnion inhabitants (for example, CL500-11, CL500-3, CL500-37; Figure 4, and Supplementary Figure S5) was intriguing, as we expected that the oligotype would be diverse between the hypolimnia of different lakes, which are physically isolated and differ in physicochemical properties (Table 1). The occurrence of CL500-3 in TR (Figure 6), which is a reservoir constructed just 10 years before the sampling, support the idea that hypolimnion inhabitants migrate between lakes; thus, diversification is limited. On the other hand, diverse sequence types are likely present at least on a continental scale, as CL500-11 from North American lakes have a conserved single-nucleotide difference from the Japanese ones, whereas European ones have both sequence types (Supplementary Figure S10).

Conclusion

This study provides the first comprehensive overview of the bacterioplankton community inhabiting the oxygenated hypolimnion by investigating 10 deep freshwater lakes with a variety of environmental characteristics. Partial 16S rRNA gene amplicon sequencing and CARD-FISH discovered the presence and quantitative significance of the hypolimnion-specific lineages, represented by Chloroflexi, Planctomycetes and Thaumarchaeota. An oligotyping

analysis suggested the presence of hypolimnion-specific sub-populations within the whole-layer inhabitants (for example, acI-B1, acI-A7). These results revealed the uniqueness, ubiquity and quantitative significance of bacterioplankton lineages inhabiting the oxygenated hypolimnion of deep freshwater lakes. Collectively, the present study provides valuable information for further studies on deep freshwater ecosystems and motivate future works to focus on ecophysiological characteristics of the individual bacterial lineages.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

References

Auguet J-C, Triadó-Margarit X, Nomokonova N, Camarero L, Casamayor EO. (2012). Vertical segregation and phylogenetic characterization of ammonia-oxidizing Archaea in a deep oligotrophic lake. *ISME J* **6**: 1786–1797.

Berdjeb L, Pollet T, Chardon C, Jacquet S. (2013). Spatio-temporal changes in the structure of archaeal communities in two deep freshwater lakes. *FEMS Microbiol Ecol* **86**: 215–230.

Berry MA, White JD, Davis TW, Jain S, Johengen TH, Dick GJ et al. (2017). Are oligotypes meaningful ecological and phylogenetic units? A case study of microcystis in freshwater lakes. *Front Microbiol* **8**: 1–7.

Bižić-Ionescu M, Zeder M, Ionescu D, Orlić S, Fuchs BM, Grossart HP et al. (2015). Comparison of bacterial communities on limnic versus coastal marine particles reveals profound differences in colonization. *Environ Microbiol* **17**: 3500–3514.

Blainey PC, Mosier AC, Potanina A, Francis CA, Quake SR. (2011). Genome of a low-salinity ammonia-oxidizing archaeon determined by single-cell and metagenomic analysis. *PLoS One* **6**: e16626.

Bornemann M, Bussmann I, Tichy L, Deutzmann J, Schink B, Pester M. (2016). Methane release from sediment seeps to the atmosphere is counteracted by highly active Methylococcaceae in the water column of deep oligotrophic Lake Constance. *FEMS Microbiol Ecol* **92**: fiw123.

Callieri C, Coci M, Eckert EM, Salcher MM, Bertoni R. (2014). Archaea and bacteria in deep lake hypolimnion: In situ dark inorganic carbon uptake. *J Limnol* **73**: 31–38.

Callieri C, Hernández-Avilés S, Salcher MM, Fontaneto D, Bertoni R. (2016). Distribution patterns and environmental correlates of Thaumarchaeota abundance in six deep subalpine lakes. *Aquat Sci* **78**: 215–225.

Catalán N, Marcé R, Kothawala DN, Tranvik LJ. (2016). Organic carbon decomposition rates controlled by water retention time across inland waters. *Nat Geosci* **9**: 1–7.

Chao A, Jost L. (2012). Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. *Ecology* **93**: 2533–2547.

Coci M, Odermatt N, Salcher MM, Pernthaler J, Corno G. (2015). Ecology and distribution of Thaumarchaea in the deep hypolimnion of Lake Maggiore. *Archaea* **2015**: 590434.

Denef VJ, Mueller RS, Chiang E, Liebig JR, Vanderploeg HA. (2016). Chloroflexi CL500-11 populations that predominate deep-lake hypolimnion bacterioplankton rely on nitrogen-rich dissolved organic matter metabolism and C 1 compound oxidation. *Appl Environ Microbiol* **82**: 1423–1432.

DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K et al. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* **72**: 5069–5072.

Edgar RC. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* **10**: 996–998.

Eiler A, Bertilsson S. (2004). Composition of freshwater bacterial communities associated with cyanobacterial blooms in four Swedish lakes. *Environ Microbiol* **6**: 1228–1243.

Erbilgin O, McDonald KL, Kerfeld CA. (2014). Characterization of a planctomycetal organelle: a novel bacterial microcompartment for the aerobic degradation of plant saccharides. *Appl Environ Microbiol* **80**: 2193–2205.

Eren AM, Maignien L, Sul WJ, Murphy LG, Grim SL, Morrison HG et al. (2013). Oligotyping: differentiating between closely related microbial taxa using 16S rRNA gene data. *Methods Ecol Evol* **4**: 1111–1119.

Eronen-Rasimus E, Lyra C, Rintala JM, Jürgens K, Ikonen V, Kaartokallio H. (2015). Ice formation and growth shape bacterial community structure in Baltic Sea drift ice. *FEMS Microbiol Ecol* **91**: 1–13.

Fuchs BM, Glöckner FO, Wulf J, Amann R. (2000). Unlabeled helper oligonucleotide increase the in situ accessibility of 16S rRNA of fluorescently labeled oligonucleotide probes. *Appl Environ Microbiol* **66**: 3603–3607.

Fujimoto M, Cavaletto J, Liebig JR, McCarthy A, Vanderploeg HA, Denef VJ. (2016). Spatiotemporal distribution of bacterioplankton functional groups along a freshwater estuary to pelagic gradient in Lake Michigan. *J Great Lakes Res* **42**: 1036–1048.

Garcia SL, Buck M, McMahon KD, Grossart HP, Eiler A, Warnecke F. (2015). Auxotrophy and intrapopulation complementary in the interactome of a cultivated freshwater model community. *Mol Ecol* **24**: 4449–4459.

Glöckner FO, Kube M, Bauer M, Teeling H, Lombardot T, Ludwig W et al. (2003). Complete genome sequence of the marine planctomycete *Pirellula* sp. strain 1. *Proc Natl Acad Sci USA* **100**: 8298–8303.

Goldberg SJ, Ball GI, Allen BC, Schladow SG, Simpson AJ, Masoom H et al. (2015). Refractory dissolved organic

nitrogen accumulation in high-elevation lakes. *Nat Commun* **6**: 6347.

Grossart HP, Tang KW, Kiørboe T, Ploug H. (2007). Comparison of cell-specific activity between free-living and attached bacteria using isolates and natural assemblages. *FEMS Microbiol Lett* **266**: 194–200.

Hahn MW, Jezberova J, Koll U, Saueressig-Beck T, Schmidt J. (2016). Complete ecological isolation and cryptic diversity in Polynucleobacter bacteria not resolved by 16S rRNA gene sequences. *ISME J* **10**: 1–14.

Hahn MW, Koll U, Jezberová J, Camacho A. (2015). Global phylogeography of pelagic Polynucleobacter bacteria: restricted geographic distribution of subgroups, isolation by distance and influence of climate. *Environ Microbiol* **17**: 829–840.

Hansell DA. (2013). Recalcitrant dissolved organic carbon fractions. *Ann Rev Mar Sci* **5**: 421–445.

Hayakawa K, Kojima R, Wada C, Suzuki T, Sugiyama Y. (2016). Distribution and characteristics of ultraviolet absorption and fluorescence of dissolved organic matter in a large lake (Lake Biwa, Japan). *J Great Lakes Res* **42**: 2–10.

Jezbera J, Jezberová J, Brandt U, Hahn MW. (2011). Ubiquity of *Polynucleobacter necessarius* subspecies *asymbioticus* results from ecological diversification. *Environ Microbiol* **13**: 922–931.

Jezbera J, Jezberová J, Kasalický V, Šimek K, Hahn MW. (2013). Patterns of Limnohabitans microdiversity across a large set of freshwater habitats as revealed by reverse line blot hybridization. *PLoS One* **8**: e58527.

Jiao N, Zheng Q. (2011). The microbial carbon pump: from genes to ecosystems. *Appl Environ Microbiol* **77**: 7439–7444.

Karlov DS, Marie D, Sumbatyan DA, Chuvochina MS, Kulichevskaya IS, Alekhina IA *et al.* (2017). Microbial communities within the water column of freshwater Lake Radok, East Antarctica: predominant 16S rDNA phylotypes and bacterial cultures. *Polar Biol* **40**: 823–836.

Kasalický V, Jezbera J, Hahn MW, Šimek K. (2013). The diversity of the Limnohabitans genus, an important group of freshwater bacterioplankton, by characterization of 35 isolated strains. *PLoS One* **8**: e58209.

Kellerman AM, Dittmar T, Kothawala DN, Tranvik LJ. (2014). Chemodiversity of dissolved organic matter in lakes driven by climate and hydrology. *Nat Commun* **5**: 1–8.

Kim C, Nishimura Y, Nagata T. (2006). Role of dissolved organic matter in hypolimnetic mineralization of carbon and nitrogen in a large, monomictic lake. *Limnol Oceanogr* **51**: 70–78.

Kumar S, Stecher G, Tamura K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* **33**: 1870–1874.

Kurilkina MI, Zakharova YR, Galachyants YP, Petrova DP, Bukin YS, Domysheva VM *et al.* (2016). Bacterial community composition in the water column of the deepest freshwater Lake Baikal as determined by next-generation sequencing. *FEMS Microbiol Ecol* **92**: fiw094.

Lemarchand C, Jardillier L, Carrias JF, Richardot M, Debroas D, Sime-Ngando T *et al.* (2006). Community composition and activity of prokaryotes associated to detrital particles in two contrasting lake ecosystems. *FEMS Microbiol Ecol* **57**: 442–451.

Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar A *et al.* (2004). ARB: a software environment for sequence data. *Nucleic Acids Res* **32**: 1363–1371.

Maki K, Kim C, Yoshimizu C, Tayasu I, Miyajima T, Nagata T. (2010). Autochthonous origin of semi-labile dissolved organic carbon in a large monomictic lake (Lake Biwa): carbon stable isotopic evidence. *Limnology* **11**: 143–153.

Mas A, Jamshidi S, Lagadeuc Y, Eveillard D, Vandenkoornhuyse P. (2016). Beyond the Black Queen Hypothesis. *ISME J* **10**: 2085–2091.

Masquelier S, Lepe C, Domaizon I, Curie M, Lepère C, Masquelier S *et al.* (2010). Vertical structure of small eukaryotes in three lakes that differ by their trophic status: a quantitative approach. *ISME J* **4**: 1509–1519.

Mazumder A, Taylor WD. (1994). Thermal structure of lakes varying in size and water clarity. *Limnol Oceanogr* **39**: 968–976.

McCarthy MD, Hedges JI, Benner R. (2013). Major bacterial contribution to marine dissolved organic nitrogen. *Physics (College Park, MD)* **6**: 35.

Mori K, Sato Y. (2015). *Japanese Lakes.* Asakura Publishing: Tokyo, Japan.

Morris JJ, Lenski RE, Zinser ER. (2012). The Black Queen Hypothesis: evolution of dependencies through adaptive gene loss. *MBio* **3**: e00036-12–e00036-12.

Morris RM, Longnecker K, Giovannoni SJ. (2006). Pirellula and OM43 are among the dominant lineages identified in an Oregon coast diatom bloom. *Environ Microbiol* **8**: 1361–1370.

Mukherjee I, Hodoki Y, Nakano S. (2015). Kinetoplastid flagellates overlooked by universal primers dominate in the oxygenated hypolimnion of Lake Biwa, Japan. *FEMS Microbiol Ecol* **91**: fiv083.

Mukherjee M, Ray A, Post AF, McKay RM, Bullerjahn GS. (2016). Identification, enumeration and diversity of nitrifying planktonic archaea and bacteria in trophic end members of the Laurentian Great Lakes. *J Great Lakes Res* **42**: 39–49.

Murase J, Sugimoto A. (2005). Inhibitory effect of light on methane oxidation in the pelagic water column of a mesotrophic lake (Lake Biwa, Japan). *Limnol Oceanogr* **50**: 1339–1343.

Nagata T, Meon B, Kirchman DL. (2003). Microbial degradation of peptidoglycan in seawater. *Limnol Oceanogr* **48**: 745–754.

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.

Newton RJ, McLellan SL. (2015). A unique assemblage of cosmopolitan freshwater bacteria and higher community diversity differentiate an urbanized estuary from oligotrophic Lake Michigan. *Front Microbiol* **6**: 1–13.

Nguyen RT, Harvey HR. (1997). Protein and amino acid cycling during phytoplankton decomposition in oxic and anoxic waters. *Org Geochem* **27**: 115–128.

Ntougias S, Polkowska Ż, Nikolaki S, Dionyssopoulou E, Stathopoulou P, Doudoumis V *et al.* (2016). Bacterial community structures in freshwater polar environments of Svalbard. *Microbes Environ* **31**: 401–409.

Nunoura T, Takaki Y, Kazama H, Hirai M, Ashi J, Imachi H *et al.* (2009). Microbial diversity in deep-sea methane

seep sediments presented by SSU rRNA gene tag sequencing. *Microbes Environ* **27**: 382–390.

Ogawa H, Amagai Y, Koike I, Kaiser K, Benner R. (2001). Production of refractory dissolved organic matter by bacteria. *Science* **292**: 917–920.

Okazaki Y, Hodoki Y, Nakano SI. (2013). Seasonal dominance of CL500-11 bacterioplankton (phylum *Chloroflexi*) in the oxygenated hypolimnion of Lake Biwa, Japan. *FEMS Microbiol Ecol* **83**: 82–92.

Okazaki Y, Nakano SI. (2016). Vertical partitioning of freshwater bacterioplankton community in a deep mesotrophic lake with a fully oxygenated hypolimnion (Lake Biwa, Japan). *Environ Microbiol Rep* **8**: 780–788.

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D et al. (2016). vegan: Community Ecology Package. version 2.4-0. Available at: https://cran.r-project.org/package = vegan (accessed 25 May 2017).

Orsi WD, Smith JM, Liu S, Liu Z, Sakamoto CM, Wilken S et al. (2016). Diverse, uncultivated bacteria and archaea underlying the cycling of dissolved protein in the ocean. *ISME J* **10**: 2158–2173.

Parfenova VV, Gladkikh AS, Belykh OI. (2013). Comparative analysis of biodiversity in the planktonic and biofilm bacterial communities in Lake Baikal. *Microbiology* **82**: 91–101.

Passow U. (2002). Transparent exopolymer particles (TEP) in aquatic environments. *Prog Oceanogr* **55**: 287–333.

Paver SF, Hayek KR, Gano KA, Fagen JR, Brown CT, Davis-Richardson AG et al. (2013). Interactions between specific phytoplankton and bacteria affect lake bacterial community succession. *Environ Microbiol* **15**: 2489–2504.

Pernthaler J. (2005). Predation on prokaryotes in the water column and its ecological implications. *Nat Rev Microbiol* **3**: 537–546.

Pernthaler A, Pernthaler J, Amann R. (2002). Fluorescence *in situ* hybridization and catalyzed reporter deposition for the identification of marine bacteria fluorescence *in situ* hybridization and catalyzed reporter deposition for the identification of marine bacteria. *Appl Environ Microbiol* **68**: 3094–3101.

Pizzetti I, Fuchs BM, Gerdts G, Wichels A, Wiltshire KH, Amann R. (2011). Temporal variability of coastal Planctomycetes clades at Kabeltonne station, North Sea. *Appl Environ Microbiol* **77**: 5009–5017.

Pollet T, Tadonléké 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.

Porter KG, Feig YS. (1980). The use of DAPI for identifying and counting aquatic microflora1. *Limnol Oceanogr* **25**: 943–948.

Pruesse E, Peplies J, Glöckner FO. (2012). SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* **28**: 1823–1829.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**: D590–D596.

Rozmarynowycz MJ. (2014). Spatio-temporal distribution of microbial communities in the Laurentian Great Lakes. Doctoral dissertation, Bowling Green State University, Bowling Green, OH.

Salazar G, Cornejo-Castillo FM, Borrull E, Díez-Vives C, Lara E, Vaqué D et al. (2015). Particle-association lifestyle is a phylogenetically conserved trait in bathypelagic prokaryotes. *Mol Ecol* **24**: 5692–5706.

Salcher MM, Pernthaler J, Posch T. (2011). Seasonal bloom dynamics and ecophysiology of the freshwater sister clade of SAR11 bacteria 'that rule the waves' (LD12). *ISME J* **5**: 1242–1252.

Šimek K, Kasalický V, Zapomělová E, Horňák K. (2011). Alga-derived substrates select for distinct betaproteo-bacterial lineages and contribute to niche separation in Limnohabitans strains. *Appl Environ Microbiol* **77**: 7307–7315.

Small GE, Bullerjahn GS, Sterner RW, Beall BFN, Brovold S, Finlay JC et al. (2013). Rates and controls of nitrification in a large oligotrophic lake. *Limnol Oceanogr* **58**: 276–286.

Stamatakis A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**: 1312–1313.

Tanoue E, Nishiyama S, Kamo M, Tsugita A. (1995). Bacterial membranes: possible source of a major dissolved protein in seawater. *Geochim Cosmochim Acta* **59**: 2643–2648.

Thottathil SD, Hayakawa K, Hodoki Y, Yoshimizu C, Kobayashi Y, Nakano S. (2013). Biogeochemical control on fluorescent dissolved organic matter dynamics in a large freshwater lake (Lake Biwa, Japan). *Limnol Oceanogr* **58**: 2262–2278.

Tilzer MM, Serruya C. (1990). *Large Lakes*. Springer Berlin Heidelberg: Berlin, Heidelberg, Germany.

Urbach E, Vergin KL, Larson GL, Giovannoni SJ. (2007). Bacterioplankton communities of Crater Lake, OR: dynamic changes with euphotic zone food web structure and stable deep water populations. *Hydrobiologia* **574**: 161–177.

Urbach E, Vergin KL, Young L, Morse A, Larson GL, Giovannoni SJ. (2001). Unusual bacterioplankton community structure in ultra-oligotrophic Crater Lake. *Limnol Oceanogr* **46**: 557–572.

Vissers EW, Blaga CI, Bodelier PLE, Muyzer G, Schleper C, Sinninghe Damsté JS et al. (2013). Seasonal and vertical distribution of putative ammonia-oxidizing thaumarchaeotal communities in an oligotrophic lake. *FEMS Microbiol Ecol* **83**: 515–526.

Wallner G, Amann R, Beisker W. (1993). Optimizing fluorescent in sit hybridization with rRNA-targeted oligonucleotide probes for flow cytometric identification of microorganisms. *Cytometry* **14**: 136–143.

Wetzel RG. (2001). *Limnology: Lake and River Ecosystems*. Academic Press: San Diego, CA, USA.

Woebken D, Teeling H, Wecker P, Dumitriu A, Kostadinov I, DeLong EF et al. (2007). Fosmids of novel marine Planctomycetes from the Namibian and Oregon coast upwelling systems and their cross-comparison with planctomycete genomes. *ISME J* **1**: 419–435.

Yamada E, Hirota T, Hatori N, Kitao Y, Fuse Y, Aoki S et al. (2012). Characterization of protein-like fluorophores released from lake phytoplankton on the basis of fractionation and electrophoresis. *Anal Sci* **28**: 595–600.

Yamashita Y, Tanoue E. (2008). Production of bio-refractory fluorescent dissolved organic matter in the ocean interior. *Nat Geosci* **1**: 579–582.

Zhang H, Wu H, Wang G, Xiang W, Yan W. (2013). Prokaryote diversity in the surface sediment of northern South China Sea. *Wei Sheng Wu Xue Bao* **53**: 915–926.

Zwart G, Crump BC, Kamst-van Agterveld MP, Hagen F, Han SK. (2002). Typical freshwater bacteria: an analysis of available 16S rRNA gene sequences from plankton of lakes and rivers. *Aquat Microb Ecol* **28**: 141–155.

Zwart G, Hiorns WD, Methé BA, van Agterveld MP, Huismans R, Nold SC *et al.* (1998). Nearly identical 16S rRNA sequences recovered from lakes in North America and Europe indicate the existence of clades of globally distributed freshwater bacteria. *Syst Appl Microbiol* **21**: 546–556.

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