Novel autotrophic organisms contribute significantly to the internal carbon cycling potential of a boreal lake

Sari Peura [1,2,3#], Moritz Buck [1,4], Sanni L Aalto [2], Sergio E. Morales[5], Hannu Nykänen [2,6], Alexander Eiler [1,7]

[1] Department of Ecology and Evolution, Limnology, Science for Life Laboratory, Uppsala University, Uppsala, Sweden

[2] Department of Biological and Environmental Science, University of Jyväskylä, Jyväskylä, Finland

[3] Department of Forest Mycology and Plant Pathology, Science for Life Laboratories, Swedish University of Agricultural Sciences, Uppsala, Sweden

[4] National Bioinformatics Infrastructure for Life Sciences, Uppsala University, Uppsala, Sweden

[5] Department of Microbiology and Immunology, University of Otago, Dunedin, New Zealand

[6] Current address: Department of Environmental and Biological Sciences, University of Eastern Finland, Kuopio, Finland

[7] Current address: Department of Bioscience, University of Oslo, Oslo, Norway

[#] Corresponding author: Sari Peura

email: sari.peura@slu.se

tel. +46 72 269 4235

Department of Forest Mycology and Plant Pathology, Box 7026, 750 07 UPPSALA, Sweden

Running title: Functional stratification in a boreal lake

The authors declare no conflict of interest

Funding sources: the Academy of Finland, Science for Life Laboratories, Tryggers Foundation, the Swedish Research Council VR and the Swedish Foundation for strategic research

Abstract

Oxygen stratified lakes are typical for the boreal zone, and also a major source of greenhouse gas emissions in the region. Due to shallow light penetration, restricting the growth of phototrophic organisms, and large allochthonous organic carbon inputs from the catchment area, the lake metabolism is expected to be dominated by heterotrophic organisms. In this study we test this assumption and show that the potential for autotrophic carbon fixation and internal carbon cycling is high throughout the water column. Further, we show that during the summer stratification carbon fixation can exceed respiration in a boreal lake even below the euphotic zone. Metagenome assembled genomes and 16S profiling of a vertical transect of the lake revealed multiple organisms in oxygen depleted compartment belonging to novel or poorly characterized phyla. Many of these organisms were chemolithotrophic, deriving their energy from reactions related to sulfur, iron and nitrogen transformations. The community as well as the functions were stratified following the redox potentials. The autotrophic potential in the lake metagenome below the oxygenic zone was high, pointing towards a need for revising our concepts of internal carbon cycling in boreal lakes. Further, the importance of chemolithoautotrophy for the internal carbon cycling suggests that many predicted climate change associated changes in the physical properties of the lake, such as altered mixing patterns, likely have consequences for the whole lake metabolism even beyond the impact to the phototrophic community.

Importance

Autotrophic organisms at the base of the food web are the only life form capable of turning inorganic carbon into organic form, facilitating the survival of all other organisms. In certain environments the autotrophic production is limited by environmental conditions and the food web is supported by carbon coming from outside the ecosystem. One such environment is stratified boreal lakes, which are one of the biggest sources of greenhouse gas emissions in the boreal region. Thus, carbon cycling in these habitats is of outmost importance for the future climate. Here we demonstrate a high potential for internal carbon cycling via phototrophic and novel chemolithotrophic organisms in the dark and anoxic layers of a boreal lake. Our results significantly increase our knowledge on the microbial communities and their metabolic potential in oxygen depleted freshwaters and help to understand and predict how climate change induced alterations could impact the lake carbon dynamics.

Introduction

All life on Earth depends on carbon fixation, where microorganisms convert inorganic carbon dioxide into organic compounds and living biomass. Currently, oxygenic phototrophs, deriving their energy from the sun light, are regarded as the most important carbon fixers. However, on planetary time scales anoxygenic phototrophs and chemotrophs have been more prevalent (1). Even today chemolithoautotrophy is a major strategy in many environments, such as deep-sea vents and sediments (2-4). In these conditions carbon fixation is driven by a redox gradient (i.e. biogeochemical gradient of reductants and oxidants sorted by their redox potentials) often located in the border zone between oxic and anoxic conditions. These redox transition zones constitute a major share of the Earth´s biosphere, and have a significant impact on surrounding entities, such as elemental cycles and food webs (5-7).

Chemolithotrophy may have a large impact on the carbon cycle in many environments. For example, in boreal lakes there is a steep redox gradient at the oxic-anoxic border, which could facilitate carbon fixation via chemolithotrophy. Still, our knowledge on the chemolithotrophic energy generation in these habitats is poor. Some first studies have highlighted the importance of chemolithotrophy for carbon assimilation in freshwater lakes such as Kivu, a freshwater lake in central Africa with high methane (CH_4) concentration (2). Furthermore, two recent studies have shown genetic potential for autotrophy in microbial communities in the anoxic water masses of boreal lakes (8, 9). This suggests that chemolithotrophy could be a common process fueling assimilation of inorganic carbon in boreal lakes. Thus, dark

carbon fixation may play a significant role in the internal carbon cycling, and thus, modulate lake food webs and ultimately whole lake carbon balances.

Small boreal lakes are important drivers of global greenhouse gas (GHG) emissions (10, 11) and their importance was highlighted in the recent report of the IPCC (12). Globally, water bodies smaller than 0.001 km^2 contribute 40% of all methane emissions from inland waters (13). Typically, these small lakes and ponds are characterized by high concentrations of dissolved organic carbon (DOC) and shallow light penetration depth, which leads to steep stratification of oxygen and other electron acceptors and donors through most of the year. The stratification coincides with a distinct set of bacterial and archaeal phyla organized according to the vertical redox gradient (9, 14, 15). The microbial communities in these lakes may harbor organisms that have the potential for photoautotrophy under low light intensity (14, 16, 17) and for chemoautotrophy throughout the water column (9). These predictions are based on taxonomic information of 16S rRNA genes combined with functional gene inventories and genomic data of related cultivated representatives, rather than a true reconstruction of the genetic makeup of individual organisms or whole communities along the redox tower. This lack of a detailed metabolic picture limits our understanding of the functional potential of microbes in boreal lakes.

We studied the potential of the microbial community for chemoautotrophy in lake Alinen Mustajärvi, a well-characterized boreal lake located in southern Finland (15) (18). This lake exhibits the typical characteristic features of boreal lakes including: 1) a high load of terrestrial organic carbon resulting in net heterotrophy

of the system (18), 2) a gradient of oxygen, temperature and light, and 3) stratification of the microbial community (9, 14, 15). We combined dark carbon fixation measurements with a survey of the functional potential of the microbial community (shotgun sequencing of the total DNA) from a vertical transect of the lake water column. Our aim was to link lake chemistry to the prevalence of genes related to energy generation via redox reactions and inorganic carbon assimilation. Moreover, we used annotated metagenome-assembled genomes (MAGs) to obtain metabolic reconstructions of uncharacterized and novel lake microbes and to identify the key chemoautotrophs in the lake. Our hypotheses were that i) the chemical stratification of the lake was linked to microbial activity and would be reflected in the functional potential and structure of the communities, ii) the lake harbors an abundant and diverse chemoautotrophic community, and iii) chemoautotrophic pathways would be enriched below the euphotic zone of the lake, leading to a high potential for internal carbon cycling.

Results and Discussion

Microbial community structure and CO_2 incorporation in the water column

The metabolic pathways and organisms that could be involved in autotrophic processes in the water column of lake Alinen Mustajärvi were studied based on metagenomic (shotgun sequencing and 16S rRNA gene amplicons) and geochemical analyses from samples taken in 2013. This vertical transect of the lake covered 13 depths. Additionally, the community composition was surveyed in 2008 using a

clone library covering 4 depths and 303 clones of 16S rRNA genes. We also measured inorganic carbon dynamics biweekly in 2008, monthly in 2009 and on 3 occasions in 2010. The carbon measurements indicated that at 3 meters depth the incorporation of inorganic carbon dominated over respiration during parts of the ice-free season (Fig. 1, Fig. S1). Time of net autotrophy coincided with stratification of the lake and vertical structuring of the microbial community along the vertical gradient (Fig. 2A). The maximum value for dark carbon assimilation at 3 m depth was 10 % of the average net primary production (carbon assimilation in the light) within the euphotic zone (Fig. 1, (18)). While these are point measurements and should be taken with caution, they are well in line with dark carbon fixation values measured from lake sediments (19).

The analysis of 16S rRNA genes indicated that the communities were dominated by similar phyla in 2008 and 2013 (Fig. 2A-B) and the community composition was consistent with previously published profiles from the same lake (15). *Actinobacteria* and *Alphaproteobacteria* were the major community members at the surface layer (epilimnion) and *Chlorobia* in the anoxic bottom layer (hypolimnion). In 2013 another major community member in the hypolimnion was the Parcubacteria, which was missing from the community in 2008 due to a mismatch in the primers used at the time (15). Community patterns were conserved across methods within the same year (Correlation in a symmetric Procrustes rotation of the 16S rRNA and metagenome communities in 2013 0.962, $p>0.001$).

The water column had a physico-chemical stratification with water temperature, oxygen (O_2) and sulfate (SO_4) concentrations decreasing with depth

168 whereas the concentrations of CH_4, DOC, ammonia (NH_4), nitrite (NO_2) and
169 phosphate (PO_4) were highest at the lake bottom (Fig. 3A-D). The microbial
170 community was stratified into distinct layers with positive correlation between
171 community composition in samples taken within 1 m distance of each other and
172 negative correlations among samples taken with more than 2.6 m apart (Fig. 3E).
173 While the bacterial community was homogenous throughout the epilimnion, below
174 there was a sharp change coinciding with the decrease in oxygen concentration in
175 the metalimnion and a subsequent succession of microbial taxa in the hypolimnion.
176
177 **Stratification of the microbial functional potential follows the redox gradient**
178 The functional potential of the microbial community was estimated based on
179 shotgun sequencing of the total DNA. Sequence coverage among samples (i.e. the
180 proportion of reads from each of the samples that could be mapped to contigs)
181 varied from 40.2 at the bottom to 75.5% in the surface, reflecting the diversity
182 gradient in the lake (Table 1). Thus, our functional analysis may be missing some of
183 the pathways present in the lake, but is expected to reflect the dominant metabolic
184 potential in the water column. The abundances of different genes were normalized
185 sample-wise using the abundance of 139 single copy genes (20). With this approach
186 the abundance of the markers is expressed as Occurrences per Genome equivalent
187 (GE), e.g. value of 1 suggests the presence of the gene in every genome of the
188 sample. It should be noted that the values suggesting abundances of 1 and above
189 could be explained by the presence of multiple gene copies in the genomes
190 harboring these markers.

Despite the shallow light penetration depth (Fig. S2), markers indicative for phototrophy could be found throughout the water column (Fig. 4). The genes encoding for oxygen-evolving photosystem II, and aerobic anoxygenic photosynthesis (AAP) decreased rapidly with depth with combined abundance of these pathways at the surface being 1 GE. A second peak in photosynthetic potential was observed in the hypolimnion, where the abundance of the marker for anaerobic anoxygenic photosynthesis (bacterial photosynthesis) was up to 0.32 GE (Fig. S3).

An iron gradient in the water column suggested that redox reactions related to iron could be important pathways for energy acquisition (Fig. 3D) and putative iron oxidizers (order *Ferrovales*) were abundant in the metalimnion. However, there were no specific markers for this pathway, thus while we have support for the potential for using iron as an energy source, we cannot visualize the trend for this pathway in the water column. For hydrogen oxidation the potential increased towards the lake bottom with the nickel dependent hydrogenase being more abundant right below the metalimnion and iron dependent at the lake bottom (Fig. 4), where the abundance of hydrogenases was close to 1 GE.

Based on their concentrations, the most important inorganic electron donors in the water column appeared to be CH_4 in the upper hypolimnion, and sulfide/sulfur, that spanned almost throughout the hypolimnion. The peak of the marker indicative for sulfide/sulfur oxidation coincided with the maxima in phototrophic sulfur-oxidizing *Chlorobia*, a major community member in the hypolimnion. At this depth, the abundance of the markers for sulfide oxidation was 0.39 GE. The marker specific for sulfate reduction suggested highest potential in the

lower hypolimnion. Markers for bacteria using anaerobic ammonia oxidation (ANAMMOX) were not detectable in the dataset. We could find markers for ammonia monooxygenase, but a manual inspection of the hits to these HMMs indicated that these were in fact to its paralog methane monooxygenase. Thus, these HMMs were used as markers for methanotrophy instead of nitrification.

The potential for aerobic respiration was stable from epilimnion to upper hypolimnion, whereas the highest potential for microaerophilic respiration was right above the depth where oxygen concentration dropped below detection limit and the potential for anaerobic respiration was highest in the upper hypolimnion. Potentials for using alternative electron acceptors for respiratory reactions followed the classic redox tower being nitrous oxide (N_2O), NO_2, NO_3, Fe^{3+}, SO_4 and CO_2 from metalimnion to the lake bottom. The profile suggested that the reactions in the denitrification pathway (reduction of nitrate to N_2O or further to nitrogen gas N_2) were divided between multiple organisms inhabiting different redox zones in the oxic-anoxic boundary layer, as the highest potential for different parts of the pathway were found at different depths (Fig. 4). Previous results on stratified lakes suggest that N_2O commonly accumulates at the oxycline (21), as it is produced through both anaerobic nitrate reduction in the anoxic hypolimnion and aerobic nitrification in the oxic water layers. Furthermore, the N_2O accumulation has been explained with the potential for nitrite reduction being higher than the potential for N_2O reduction (9, 21). This is due to the expression of N_2O reductase gene (nosZ) being more sensitive to oxygen than expression of nitrite reductase gene (nirK) (22). In our pfam/tigrfam profiles, the potential for nitrate reduction occurred in the

upper hypolimnion and N_2O reduction in the lower metalimnion, while markers for nitrification were sparse, suggesting that N_2O dynamics is driven by anaerobic processes, potentially using NO_3 as electron acceptor. The general patterns for Fe^{3+} and SO_4 reduction could not be assessed due to lack of pathway specific HMMs, but these were investigated using metagenome assembled genomes (MAGs; Table S2).

Unrecognized metabolic and chemoautotrophic potential in novel bacterial taxa

In general the microbial community had potential for three different pathways of autotrophic carbon fixation (Fig. 5). In the epilimnion, Calvin-Benson-Bassham cycle (CBB; also called reductive pentose phosphate pathway) was highly abundant, while reductive citric acid cycle (rTCA) and Wood-Ljungdahl pathway (WL; also called reductive acetyl CoA cycle) were mainly found in the hypolimnion. For the fourth carbon assimilation pathway present in the communities, 3-hydroxypropionate cycle (3HP), no specific HMMs could be found. However, protein annotations from Prokka suggested that multiple organisms possessed an almost full 3HP pathway, but certain genes in the pathway, such as mcr (malonyl CoA reductase), were not present in the dataset. The data suggested increasing autotrophic potential towards the bottom of the lake with a peak right below the oxycline and the highest potential at lake bottom with the total abundance of pathways related to carbon fixation close to 1 GE (Fig. 5).

We were able to construct a total of 270 metagenome assembled genomes (MAGs) of which 93 fulfilled our criteria for a high quality MAG (> 40 %

260 completeness and < 4 % contamination). These were identified using PhyloPhlan
261 (Fig. S4) (23). The high quality MAGs had the potential to use a wide range of
262 different electron donors and acceptors and for a full range of microbial metabolic
263 pathways from chemolithoautotrophy to photoheterotrophy (Table S2). Here we
264 concentrate on the most abundant organisms with chemotrophic potential.
265 As stated above, the chemotrophic organisms appeared to be deriving their
266 energy mainly from oxidation of sulfur, iron and hydrogen compounds. The most
267 abundant MAGs with potential for sulfur oxidation (sox operon typically including
268 genes soxACXYZ) included organisms related to *Chlorobia* (bin 4), *Polynucleobacter*
269 (bins 15 and 120), *Comamonadaceae* (bin 108), *Ferrovales* (bin 139a), *Rhizobiales*
270 (bin 139b) and *Acetobacteraceae* (bins 72 and 168). However, only the *Chlorobia*
271 and *Ferrovales* MAGs had potential for autotrophy. *Chlorobia* was located in the lake
272 hypolimnion and had a near-complete complex for oxidation of reduced sulfur
273 compounds and also a near-complete CBB pathway for autotrophic carbon
274 assimilation. Further, two other *Chlorobium* MAGs (bins 52 and 92) had dsr-operon,
275 which is used in reverse in *Chlorobium* to oxidize sulfur (24). *Chlorobia* are known
276 to be able to grow both autotrophically and mixotrophically (25). They derive the
277 energy for autotrophy from the sunlight using sulfur oxidization to produce
278 reducing equivalents for growth. Thus, they are photoautotrophic organisms rather
279 than chemoautotrophs.
280 The other MAG with capacity for sulfur oxidation and inorganic carbon
281 assimilation was a MAG closely related to betaproteobacterial *Ferrovales*-order (bin
282 139a). This MAG was 4.6 Mb in size (91.3% complete with 0.67% contamination)

283 and included all of the genes in the CBB and rTCA pathways, suggesting the potential
284 for autotrophy. While the MAG had potential for sulfur oxidation, it also encoded
285 genes for iron oxidation: cytochrome cyc1, cytochrome-c oxidase (ctaCDE),
286 ubiquinol-cytochrome-c reductase (petABC) and NADH:quinone oxidoraductase
287 (nuoABCDEFGHIJKLMN). Further, the genome included an iron oxidase, which was
288 located on the same contig as cytochrome cyc1, completing the pathway. This is
289 consistent with the closest cultivated organism, iron oxidizing *Ferrovum mycofaciens*
290 (26), which has previously been shown to thrive in acidic mine drainage (26). The
291 MAG also included markers indicative of aerobic, anaerobic and microaerophilic
292 respiration. Thus, the electron acceptor could be O_2 when it is available. However,
293 we could not identify any other electron acceptors for this organism. Thus, the
294 genomic features together with the abundance distribution in the lake suggest that
295 this organism is a chemoautotroph inhabiting suboxic to anoxic environments, using
296 iron or sulfur oxidation as an energy source. In our data this organism was rather
297 abundant in both 16S rRNA amplicon data, and in the metagenomes, but has not
298 been previously found in boreal lakes. This particulate taxon was abundant in a
299 narrow zone (between 2.5-3.6m), thus, a possible reason why it has been previously
300 missed is the sparse sampling schemes of many experiments. Moreover, it
301 represents a recently established bacterial order (26) and this taxon may have been
302 classified as "uncultured *Betaproteobacteria*" in previous studies.

303 MAGs with autotrophic potential also included organisms that would appear
304 to acquire energy by combining oxidation of hydrogen to sulfate or nitrate
305 reduction. Hydrogenases found in the data represented FeFe- and NiFe-type

306 hydrogenases as described in (27), with the latter type being more prevalent among
307 the MAGs. For example, a MAG closely related to *Gallionellaceae* (bin 129) carried a
308 1e type hydrogenase, which is specifically used for electron input to sulfur
309 respiration, and the MAG did have potential for sulfur reduction. It also had a full
310 CBB pathway. Bin 129 was closely related to betaproteobacterial *Sideroxydans*
311 *lithotropicus*, which has previously been found thrive in the same environment with
312 *Ferrovales* (28). Similar to bin 129, this organism has the potential for CBB cycle,
313 however, it has been suggested to derive its energy from iron rather than hydrogen
314 oxidation (28). Another MAG, closely related to *Desulfobulbaceae* (bin 93), had a 1c
315 type hydrogenase, also typically related to sulfate respiration, and an almost
316 complete pathway for sulfate reduction. It also appeared to have the potential for
317 CO_2 fixation through reductive TCA cycle. The closest relative to this bin,
318 *Desulfotalea psychrophila*, is also a sulfate reducer, but has been reported to be
319 heterotrophic instead of autotrophic (29). However, *D. psychrophila* inhabits cold
320 enviroenments, which is consistent with bin 93, which was most abundant in the
321 deep layer of the lake where the water temperature is around 4 ˚C throughout the
322 year.

323 In accordance with the redox potentials in the water column, and literature
324 (30, 31), the reduction of nitrate to N_2 was dispersed among multiple organisms.
325 Candidatus Methyloumidiphilus alinensis (bin 10) (32) and a MAG closely affiliated
326 with *Chrenotrix* (bin 149) both had a complete narGHIJ operon for nitrate reduction,
327 and also genes for methane oxidation, as has been previously described for a
328 member of *Methylobacter* family (33). Gene NosZ, coding for N_2O reductase, was

329 present in two high quality MAGs, which were taxonomically assigned to
330 *Myxococcales* (bin 233) and *Bacteroidetes* (bin 64). However, these did not appear to
331 be autotrophic organisms. NorCB operon, encoding NO reductase, was complete in
332 two MAGs, in Candidatus M. alinensis and in a MAG affiliated to *Comamonadaceae*
333 (bin 239). The latter was also carrying the potential for CBB cycle. The possible
334 electron acceptors for the organism were sulfur and hydrogen. At the very bottom of
335 the lake we could identify three archeal MAGs. Two of these were hydrogen
336 oxidizing (hydrogenotrophic methanogenesis; bins 133 and 155, closely related to
337 *Metanolinea* and *Methanoregula*, respectively), while the third one was using acetate
338 (acetogenic methanogenesis; bin 74 closely related to *Methanosaeta*).
339

340 **Conclusions**

341 Consistent with our expectations, the microbial community included a variety
342 of different chemotrophic pathways. Further, the microbial community driving
343 these processes contained abundant novel organisms with the potential for
344 autotrophy. The assembly of the functional potentials was in accordance with redox
345 potentials of the electron acceptors in the lake (Fig. 6). Further, consistent with our
346 hypothesis, there was a diverse set of abundant chemoautotrophic organisms below
347 the euphotic zone of the lake. The autotrophic community had an unexpected major
348 community member closely related to *Ferrovales*, presumably thriving via iron
349 oxidation. We also identified other novel autotrophs, such as an organism related to
350 recently identified autotrophic *Sideroxydans lithotropicus*.

351 The fact that many of the autotrophic microbes were among the most

352 abundant microbes in the lake and the abundance of photoautotrophic *Chlorobia,*
353 emphasize the potential role of internal carbon cycling as a process that mitigates
354 the flow of CO_2 from boreal lakes to the atmosphere. Our results suggest that
355 autotrophic iron, sulfur and hydrogen oxidizing microbes have a high potential to
356 significantly contribute to inorganic carbon fixation in the lake. In fact, our
357 measurements of inorganic carbon incorporation suggested that a significant
358 amount of CO_2 originating from degradation of autochthonous and terrestrial
359 carbon can be re-incorporated into biomass in the poorly illuminated, anoxic layer
360 of the lake. This is also well in line with results showing that chemolithoautotrophy
361 significantly contributes to carbon and energy flow in meromictic lake Kivu (2), and
362 with recent results regarding the autotrophic potential in boreal lakes (8). These
363 processes are strongly dependent on the prevalent environmental conditions, which
364 have been predicted to change following the warming of the climate. Thus, our
365 results suggest that if the predicted alterations in lake environment, such as changes
366 in mixing patterns, should happen, we may expect reorganization of the metabolic
367 processes in the lake, which would have unknown implications to the carbon flow in
368 the water column.
369

370 **Materials and Methods**

371 **Site description and sampling**

372 The study lake, Alinen Mustajärvi, is situated in southern Finland (61°12´N,
373 25°06´E). It is a 0.007 km^2 head-water lake with maximum depth 6.5 m and an
374 estimated volume of 31 × 10^3 m^3. The catchment area is <0.5 km^2 and it consists of

375 >90% coniferous forest and <10% peatland. The lake is characterized by steep
376 oxygen stratification during summer and also during ice cover period, which lasts
377 from late November until late April. The stratification is disrupted by regular
378 autumn and irregular spring mixings. As such Alinen Mustajärvi is a representative
379 for the millions of lakes and ponds in the arctic and boreal zones.

380 The metagenome sampling was conducted in the beginning of September
381 2013, at the end of stratification period. The lake was sampled at 13 depths; the oxic
382 epilimnion was sampled at 0.1, 1.1, 1.6, 2.1 and 2.3 m; the metalimnion at 2.5 and
383 2.9 and the hypolimnion at 3.6 m, 4.1, 4.6, 5.1, 5.6 and 6.1 m. Water samples were
384 taken with a 20-cm-long acrylic tube sampler (Limnos vol. 1.1 L) and subsequently
385 analyzed for nutrients (NO_2/NO_3, PO_4, NH_4, total N, total P, SO_4), gasses (CH_4, CO_2)
386 and dissolved organic carbon (DOC) concentration. Nutrient analyses were
387 conducted using standard methods (http://www.sfs.fi/). Gas and DOC analyses
388 were done as in (34) and iron as in (35). The dark carbon fixation was measured in
389 2008 as an increase or decrease of dissolved inorganic carbon (DIC) concentration
390 during 24 h incubation at the depth of 3 m in two foil-covered 50-ml glass stoppered
391 BOD bottles. The measurements were conducted every second week from beginning
392 of May until end of October and the changes in DIC concentration were analyzed
393 according to (36). In August 2008 a clone library was created from depths 0.5, 2.5,
394 4.5 and 5.5 m, consisting of a total of 303 sequences (37).

395

396 *Amplicon and Metagenome analysis*

The samples for metagenomic analysis of lake microbiota were taken by filtering water through a 0.2 µm polycarbonate filters which were then frozen at -78 °C until further analysis. The DNA was extracted from the filters using Mobio PowerSoil DNA extraction kit (MO BIO Laboratories). Sample preparation for 16S rRNA gene analysis and the following sequence processing were conducted as previously described (38).

Shotgun metagenomic libraries were prepared from 10 ng of genomic DNA. First, the genomic DNA was sheared using a focused-ultrasonicator (Covaris E220) and subsequently, sequencing libraries were prepared with the Thruplex FD Prep kit from Rubicon Genomics according to the manufactures protocol (R40048-08, QAM-094-002). Library size selection was made with AMPure XP beads (Beckman Coulter) in 1:1 ratio. The prepared sample libraries were quantified using KAPA Biosystem´s next-generation sequencing library qPCR kit and run on a StepOnePlus (Life Technologies) real-time PCR instrument. The quantified libraries were then prepared for sequencing on the Illumina HiSeq sequencing platform with a TruSeq paired-end cluster kit, v3, and Illumina´s cBot instrument to generate a clustered flowcell for sequencing. Sequencing of the flow cell was performed on the Illumina HiSeq2500 sequencer using Illumina TruSeq SBS sequencing kits, v3, following a 2x100 indexed high-output run protocol.

The sequencing produced a total of 120.5 Gb of sequence data. The raw data has been deposited to NCBI Sequence Read Archive under accession number SRP076290. Reads were filtered based on their quality scores using sickle (version v1.33) (39) and subsequently assembled with Ray (version v2.3.1) (40). Assembled

420 contigs from kmer sizes of 51, 61, 71 and 81 were cut into 1000 bp pieces and
421 scaffolded with Newbler (454 Life Sciences, Roche Diagnostics). Mapping of the
422 original reads to the Newbler assembly was done using bowtie2 (version v2.15.0)
423 (41), while duplicates were removed using picard-tools (version 1.101;
424 https://github.com/broadinstitute/picard), and for computing coverage, bedtools
425 (42) was used. Details on the assembly results are presented in Table 1. The data
426 was then normalized using the counts of 139 single copy genes as previously
427 described (20). Assembled contigs were binned with MetaBAT (version v0.26.3)
428 (43) to reconstruct genomes of the most abundant lake microbes (metagenome
429 assembled genomes (MAGs)). The quality of the MAGs was evaluated using CheckM
430 (version v1.0.6) (44). Cutoffs for high quality MAGs were set to ≥ 40% for
431 completeness and ≤ 4% for contamination. The placement of the MAGs in the
432 microbial tree of life was estimated using Phylophlan (version v1.1.0) (23).

433 The functional potential of the metagenomes was assessed from assembled
434 data using the hidden Markov models (HMM) of the Pfam and TIGRFAM databases
435 (45, 46) and the HMMER3 software (version v3.1b2) (47). Special attention was
436 paid to pathways linked to energy metabolism and carbon cycle. To assure pathway
437 specificity, marker HMMs were chosen to be unique to specific pathways (Table S1).
438 Normalized coverage information of the contigs combined with HMMs of specific
439 marker genes (Table S1) was used to predict protein domains related to energy
440 metabolism and carbon incorporation to biomass. Only marker genes that were
441 found to exhibit a significantly different distribution between layers are reported
442 (p-values in Table S1). All of the MAGs were also annotated using Prokka (version

v1.11) (48). The metabolic potentials of all high quality MAGs were evaluated based on Prokka annotations (Table S2). In case of novel functional combinations, such as combination of AAP to CBB, the contigs were blasted against the NCBI nr database to verify that the closest relatives of the overall genes in these contigs were matching to the Phylophlan annotation. MAGs with special functional properties were also visualized with non-metric multidimentional scaling to check the placement of the contigs with markers within all the contigs comprising the MAG in question. All statistical analyses were done using R software (http://www.R-project.org (49)) and packages vegan (50) and mpmcorrelogram (51). Differences in the abundance of marker HMMs between layers was tested using permutation test (1000 permutations) on the t-statistics with package MetagenomeSeq (52).

Acknowledgements

We thank Sainur Samad for his help with processing the DNA samples. Lammi biological station is acknowledged for facilities and equipment during the sampling as well as for nutrient analyses. Sequencing was performed at the SNP & SEQ technology platform at Science for Life Laboratory, Uppsala. IT Center for Science (CSC; Espoo, Finland) and Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX; Uppsala, Sweden) provided the computational and data storage resources. Funding was provided by the Academy of Finland (Grant Number 265902 to SP, and for 2008 sampling grant number 114604 to Roger I Jones), the Tryggers Foundation (Grant to SP and AE), the Swedish Research

Council VR (Grant 2012-4592 to AE) and the Swedish Foundation for strategic research (Grant ICA10-0015 to AE).

The authors declare no competing financial interests.

References

1. Oschmann W, Grasshof, M. & Gudo, M. 2002. The early evolution of the planet earth and the origin of life. Senckenbergiana lethaea 82:284-294.
2. Morana C, Roland FAE, Crowe SA, Lliros M, Borges AV, Darchambeau F, Bouillon S. 2016. Chemoautotrophy and anoxygenic photosynthesis within the water column of a large meromictic tropical lake (Lake Kivu, East Africa). Limnol Oceanogr 61:1424-1437.
3. Bell JB, Woulds C, Oevelen DV. 2017. Hydrothermal activity, functional diversity and chemoautotrophy are major drivers of seafloor carbon cycling. Sci Rep 7:12025.
4. Middelburg JJ, Mueller CE, Veuger B, Larsson AI, Form A, van Oevelen D. 2015. Discovery of symbiotic nitrogen fixation and chemoautotrophy in cold-water corals. Sci Rep 5:17962.
5. Canfield DE, Stewart FJ, Thamdrup B, De Brabandere L, Dalsgaard T, Delong EF, Revsbech NP, Ulloa O. 2010. A Cryptic Sulfur Cycle in Oxygen-Minimum-Zone Waters off the Chilean Coast. Science 330:1375-1378.
6. Fuchsman CA, Kirkpatrick JB, Brazelton WJ, Murray JW, Staley JT. 2011. Metabolic strategies of free-living and aggregate-associated bacterial communities inferred from biologic and chemical profiles in the Black Sea suboxic zone. Fems Microbiol Ecol 78:586-603.
7. La Cono V, Smedile F, Bortoluzzi G, Arcadi E, Maimone G, Messina E, Borghini M, Oliveri E, Mazzola S, L'Haridon S, Toffin L, Genovese L, Ferrer M, Giuliano L, Golyshin PN, Yakimov MM. 2011. Unveiling microbial life in new deep-sea hypersaline Lake Thetis. Part I: Prokaryotes and environmental settings. Environ Microbiol 13:2250-2268.
8. Alfreider A, Baumer A, Bogensperger T, Posch T, Salcher MM, Summerer M. 2017. CO2 assimilation strategies in stratified lakes: Diversity and distribution patterns of chemolithoautotrophs. Environ Microbiol 19:2754-2768.
9. Peura S, Sinclair L, Bertilsson S, Eiler A. 2015. Metagenomic insights into strategies of aerobic and anaerobic carbon and nitrogen transformation in boreal lakes. Sci Rep 5:12102.
10. Downing JA, Prairie YT, Cole JJ, Duarte CM, Tranvik LJ, Striegl RG, McDowell WH, Kortelainen P, Caraco NF, Melack JM, Middelburg JJ. 2006. The global

abundance and size distribution of lakes, ponds, and impoundments. Limnol Oceanogr 51:2388-2397.

11. Negandhi K, Laurion I, Whiticar MJ, Galand PE, Xu XM, Lovejoy C. 2013. Small Thaw Ponds: An Unaccounted Source of Methane in the Canadian High Arctic. Plos One 8:e78204.

12. Settele J, Scholes R, Betts R, Bunn SE, Leadley P, Nepstad D, Overpeck JT, Taboada MA. 2014. Terrestrial and inland water systems, p 271-359. *In* Field CB, Barros VR, Dokken DJ, Mach KJ, Mastrandrea MD, Bilir TE, Chatterjee M, Ebi KL, Estrada YO, Genova RC, Girma B, Kissel ES, Levy AN, MacCracken S, Mastrandrea PR, White LL (ed), Climate Change 2014: Impacts, Adaptation, and Vulnerability Part A: Global and Sectoral Aspects Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel of Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

13. Holgerson MA, Raymond PA. 2016. Large contribution to inland water CO2 and CH4 emissions from very small ponds. Nature Geosci 9:222-226.

14. Taipale S, Jones RI, Tiirola M. 2009. Vertical diversity of bacteria in an oxygen-stratified humic lake, evaluated using DNA and phospholipid analyses. Aquat Microbial Ecol 55:1-16.

15. Peura S, Eiler A, Bertilsson S, Nykanen H, Tiirola M, Jones RI. 2012. Distinct and diverse anaerobic bacterial communities in boreal lakes dominated by candidate division OD1. ISME J 6:1640-1652.

16. Eiler A, Beier S, Sawstrom C, Karlsson J, Bertilsson S. 2009. High Ratio of Bacteriochlorophyll Biosynthesis Genes to Chlorophyll Biosynthesis Genes in Bacteria of Humic Lakes. Appl Environ Microbiol 75:7221-7228.

17. Karhunen J, Arvola L, Peura S, Tiirola M. 2013. Green sulphur bacteria as a component of the photosynthetic plankton community in small dimictic humic lakes with an anoxic hypolimnion. Aquat Microbial Ecol 68:267-272.

18. Peura S, Nykanen H, Kankaala P, Eiler A, Tiirola M, Jones RI. 2014. Enhanced greenhouse gas emissions and changes in plankton communities following an experimental increase in organic carbon loading to a humic lake. Biogeochemistry 118:177-194.

19. Santoro AL, Bastviken D, Gudasz C, Tranvik L, Enrich-Prast A. 2013. Dark carbon fixation: an important process in lake sediments. PLoS One 8:e65813.

20. Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng JF, Darling A, Malfatti S, Swan BK, Gies EA, Dodsworth JA, Hedlund BP, Tsiamis G, Sievert SM, Liu WT, Eisen JA, Hallam SJ, Kyrpides NC, Stepanauskas R, Rubin EM, Hugenholtz P, Woyke T. 2013. Insights into the phylogeny and coding potential of microbial dark matter. Nature 499:431-437.

21. Saarenheimo J, Rissanen AJ, Arvola L, Nykanen H, Lehmann MF, Tiirola M. 2015. Genetic and Environmental Controls on Nitrous Oxide Accumulation in Lakes. Plos One 10:e0121201.

22. Dalsgaard T, Stewart FJ, Thamdrup B, De Brabandere L, Revsbech NP, Ulloa O, Canfield DE, DeLong EF. 2014. Oxygen at Nanomolar Levels Reversibly Suppresses Process Rates and Gene Expression in Anammox and

549 Denitrification in the Oxygen Minimum Zone off Northern Chile. mBio
550 5:e01966.
551 23. Segata N, Bornigen D, Morgan XC, Huttenhower C. 2013. PhyloPhlAn is a new
552 method for improved phylogenetic and taxonomic placement of microbes.
553 Nat Comm 4:2304.
554 24. Holkenbrink C, Barbas SO, Mellerup A, Otaki H, Frigaard NU. 2011. Sulfur
555 globule oxidation in green sulfur bacteria is dependent on the dissimilatory
556 sulfite reductase system. Microbiology 157:1229-1239.
557 25. Tang KH, Blankenship RE. 2010. Both forward and reverse TCA cycles
558 operate in green sulfur bacteria. J Biol Chem 285:35848-35854.
559 26. Johnson DB, Hallberg KB, Hedrich S. 2014. Uncovering a Microbial Enigma:
560 Isolation and Characterization of the Streamer-Generating, Iron-Oxidizing,
561 Acidophilic Bacterium "Ferrovum myxofaciens". Appl Environ Microbiol
562 80:672-680.
563 27. Søndergaard D, Pedersen CN, Greening C. 2016. HydDB: A web tool for
564 hydrogenase classification and analysis. Sci Rep 6:34212.
565 28. Mühling M, Poehlein A, Stuhr A, Voitel M, Daniel R, Schlömann M. 2016.
566 Reconstruction of the Metabolic Potential of Acidophilic. Front Microbiol
567 7:2082.
568 29. Rabus R, Ruepp A, Frickey T, Rattei T, Fartmann B, Stark M, Bauer M, Zibat A,
569 Lombardot T, Becker I, Amann J, Gellner K, Teeling H, Leuschner WD,
570 Glöckner FO, Lupas AN, Amann R, Klenk HP. 2004. The genome of
571 Desulfotalea psychrophila, a sulfate-reducing bacterium from permanently
572 cold Arctic sediments. Environ Microbiol 6:887-902.
573 30. Green SJ, Prakash O, Gihring TM, Akob DM, Jasrotia P, Jardine PM, Watson DB,
574 Brown SD, Palumbo AV, Kostka JE. 2010. Denitrifying Bacteria Isolated from
575 Terrestrial Subsurface Sediments Exposed to Mixed-Waste Contamination.
576 Appl Environ Microbiol 76:3244-3254.
577 31. Jones CM, Stres B, Rosenquist M, Hallin S. 2008. Phylogenetic analysis of
578 nitrite, nitric oxide, and nitrous oxide respiratory enzymes reveal a complex
579 evolutionary history for denitrification. Mol Biol Evol 25:1955-1966.
580 32. Rissanen AJ, Saarenheimo J, Tiirola M, Peura S, Aalto SL, Karvinen A,
581 Nykänen H. 2018. Gammaproteobacterial methanotrophs dominate
582 methanotrophy in aerobic and anaerobic layers of boreal lake waters. Aquat
583 Microbial Ecol doi:10.3354/ame01874.
584 33. Kits KD, Klotz MG, Stein LY. 2015. Methane oxidation coupled to nitrate
585 reduction under hypoxia by the Gammaproteobacterium Methylomonas
586 denitrificans, sp. nov. type strain FJG1. Environ Microbiol 17:3219-3232.
587 34. Kankaala P, Taipale S, Li L, Jones RI. 2010. Diets of crustacean zooplankton,
588 inferred from stable carbon and nitrogen isotope analyses, in lakes with
589 varying allochthonous dissolved organic carbon content. Aquat Ecol 44:781-
590 795.
591 35. Viollier E, Inglett PW, Hunter K, Roychoudhury AN, Van Cappellen P. 2000.
592 The ferrozine method revisited: Fe(II)/Fe(III) determination in natural
593 waters. Appl Geochem 15:785-790.

36. Salonen K. 1981. Rapid and precise determination of total inorganic carbon and some gases in aqueous-solutions. Water Res 15:403-406.
37. Peura S, Eiler A, Hiltunen M, Nykänen H, Tiirola M, Jones RI. 2012. Bacterial and phytoplankton responses to nutrient amendments in a boreal lake differ according to season and to taxonomic resolution. PLoS One 7:e38552.
38. Sinclair L, Osman OA, Bertilsson S, Eiler A. 2015. Microbial Community Composition and Diversity via 16S rRNA Gene Amplicons: Evaluating the Illumina Platform. Plos One 10:e116955.
39. Joshi NA, Fass, J.N. 2011. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files. https://github.com/najoshi/sickle
40. Boisvert S, Laviolette F, Corbeil J. 2010. Ray: Simultaneous Assembly of Reads from a Mix of High-Throughput Sequencing Technologies. J Comp Biol 17:1519-1533.
41. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357-9.
42. Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841-842.
43. Kang DWD, Froula J, Egan R, Wang Z. 2015. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. Peerj 3:e1165.
44. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Research 25:1043-1055.
45. Finn RD, Tate J, Mistry J, Coggill PC, Sammut SJ, Hotz HR, Ceric G, Forslund K, Eddy SR, Sonnhammer ELL, Bateman A. 2008. The Pfam protein families database. Nucleic Acids Res 36:D281-D288.
46. Selengut JD, Haft DH, Davidsen T, Ganapathy A, Gwinn-Giglio M, Nelson WC, Richter AR, White O. 2007. TIGRFAMs and Genome Properties: tools for the assignment of molecular function and biological process in prokaryotic genomes. Nucleic Acids Res 35:D260-D264.
47. Durbin R, Eddy, S., Krogh, A. & Mitchison, G. 2002. Biological sequence analysis: probabilistic models of proteins and nucleic acids. Cambridge University Press.
48. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068-2069.
49. Team RC. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
50. Jari Oksanen FGB, Michael Friendly, Roeland Kindt, Pierre Legendre, Dan, McGlinn PRM, R. B. O'Hara, Gavin L. Simpson, Peter Solymos, M. Henry H. Stevens,, Eduard Szoecs HW. 2017. vegan: Community Ecology Package.
51. S. Matesanz TEG, M. de la Cruz, A. Escudero and F. Valladares. 2011. Competition may explain the fine-scale spatial patterns and genetic structure of two co-occurring plant congeners. J Ecol 99:838-848.
52. Paulson JN, Stine OC, Bravo HC, Pop M. 2013. Differential abundance analysis for microbial marker-gene surveys. Nat Methods 10:1200-1202.

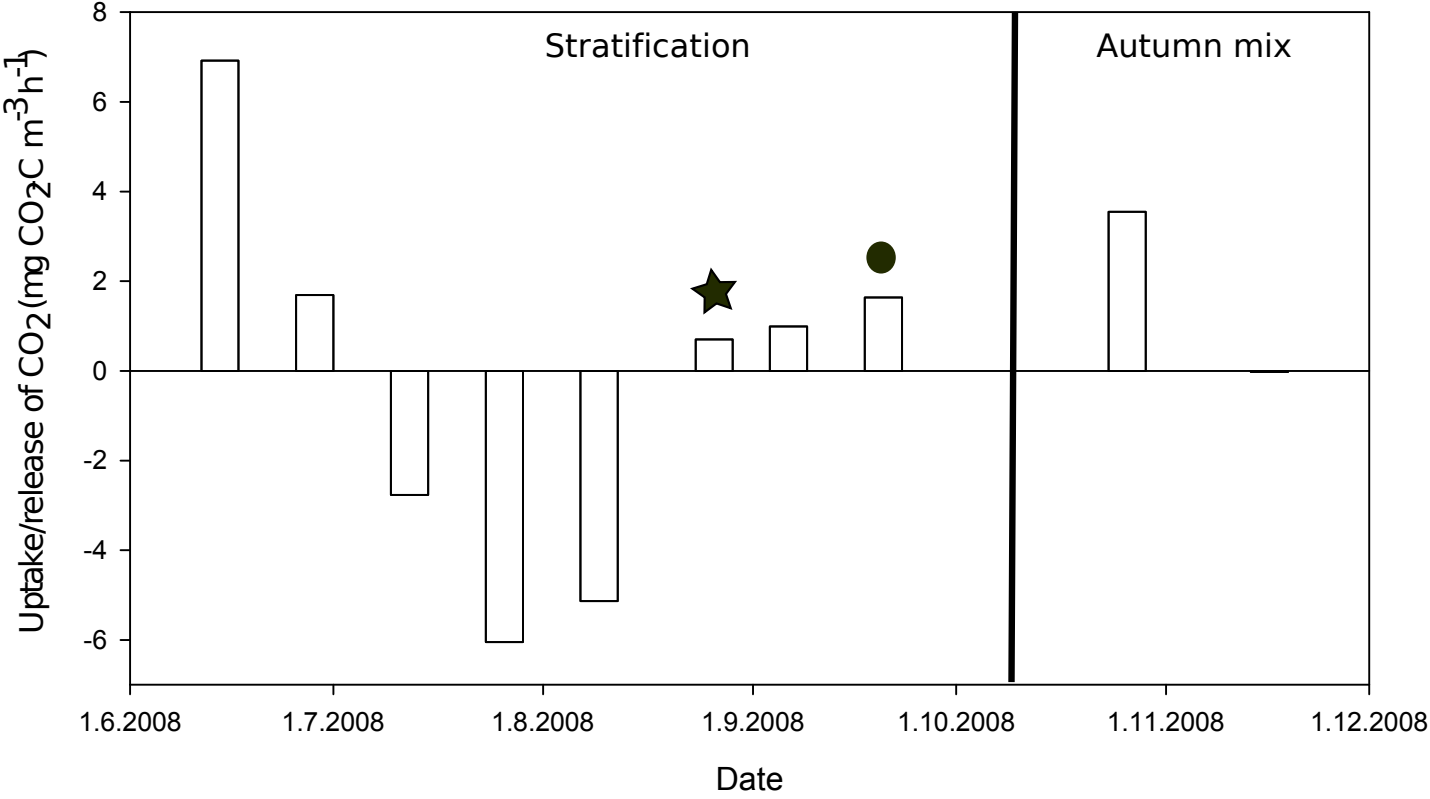
640 **Legends**

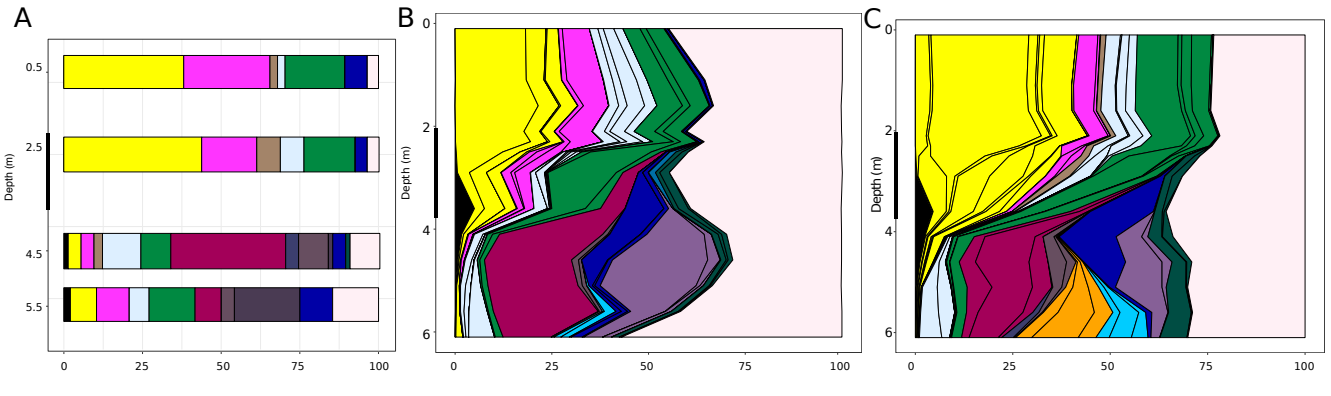
641 **Table 1.** Sampling depth, water layer, data size, assembly coverage in each sample,
642 and inverse Simpson index and Pielou´s evenness of the 16S rRNA $OTU_{0.03}$ data.
643
644 **Figure 1.** Change in CO_2 concentration over 24 h *in situ* dark incubations during
645 open water season 2008 at a depth of 3 m. Positive values indicate respiration
646 dominating over CO_2 assimilation while negative values show net incorporation of
647 inorganic carbon into biomass. Star depicts the time point the clone library was
648 retrieved in 2008 and the circle the time point when the amplicon and shot gun
649 samples were taken in 2013.
650
651 **Figure 2.** Microbial community composition in lake Alinen Mustajärvi based on A)
652 16S rRNA genes in 2008 (a clone library), B) 16S rRNA genes in 2013 (Illumina
653 MiSeq), and C) metagenome assemble genomes from 2013 (Illumina HiSeq). In
654 panels B and C different OTUs of the same phylum are separated by black lines. The
655 black highlights on y-axis show the location of oxygen depletion zone.
656
657 **Figure 3.** Environmental conditions in the lake in 2013. A) Concentration of O_2 and
658 CH_4 and water temperature. Concentration of B) NH_4, NO_2, NO_3 and SO_4, C) PO_4 and
659 DOC, and D) Fe^{2+} and Fe^{3+}. E) Bonferroni corrected Pearson correlations of the
660 community composition (measured as Bray-Curtis distances) of the samples
661 according to sampling distance. Black symbols designate significant correlations

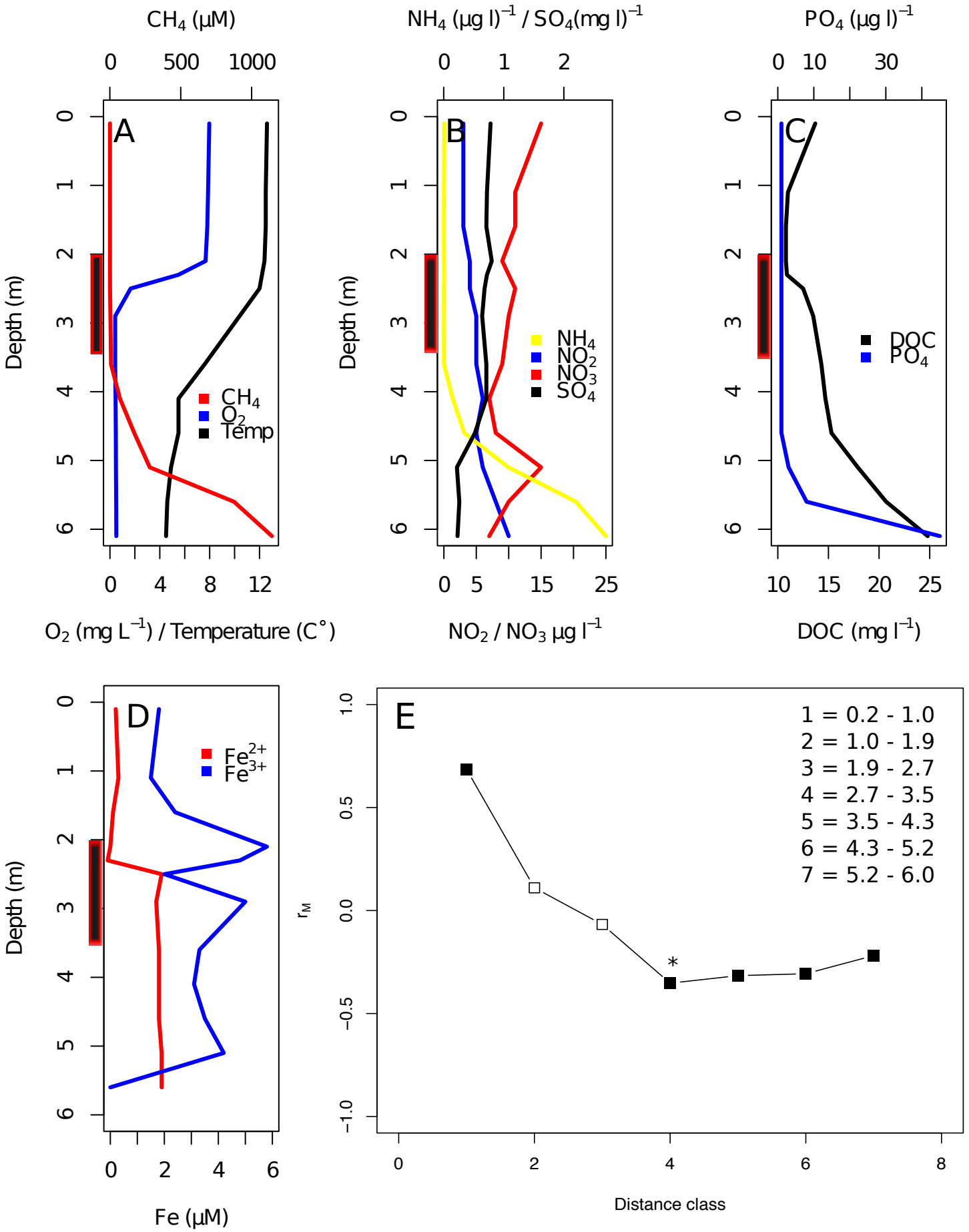
662 with $p < 0.005$ to all except distance 4, for which $p = 0.006$. The black highlights on
663 the y-axis of panels A-D illustrate the oxygen depletion zone.
664
665 **Figure 4.** Heatmap visualizing the abundances of pfam/tigrfam markers related to
666 energy metabolism (as given in z-score standardized per genome equivalent). Only
667 those pfams/tigrfams that had significantly different abundance between different
668 layers of the lake are displayed. Colors at the top of each of the columns reflect the
669 function that the marker represents and colors on the left side of the heatmap
670 illustrate different layers of the lake.
671
672 **Figure 5.** Heatmap visualizing the abundances of pfam/tigrfam markers (as given in
673 z-score standardized per genome equivalent) related to carbon fixation and the total
674 abundance of these genes in the dataset as a sum of the average abundances of the
675 markers for each pathway, respectively. Only those pfams/tigrfams that had
676 significantly different abundance between different layers of the lake are displayed.
677 Colors at the top of each of the columns reflect the function that the marker
678 represents and colors on the left side of the heatmap illustrate different layers of the
679 lake.
680
681 **Figure 6.** Redox reactions potentially driving the autotrophic processes in the lake
682 and the most abundant organisms harboring these pathways. The colors represent
683 the taxonomic annotation of the organisms at the phylum level or, in the case of
684 Proteobacteria, the order. The height of the boxes visualizes the depth of the

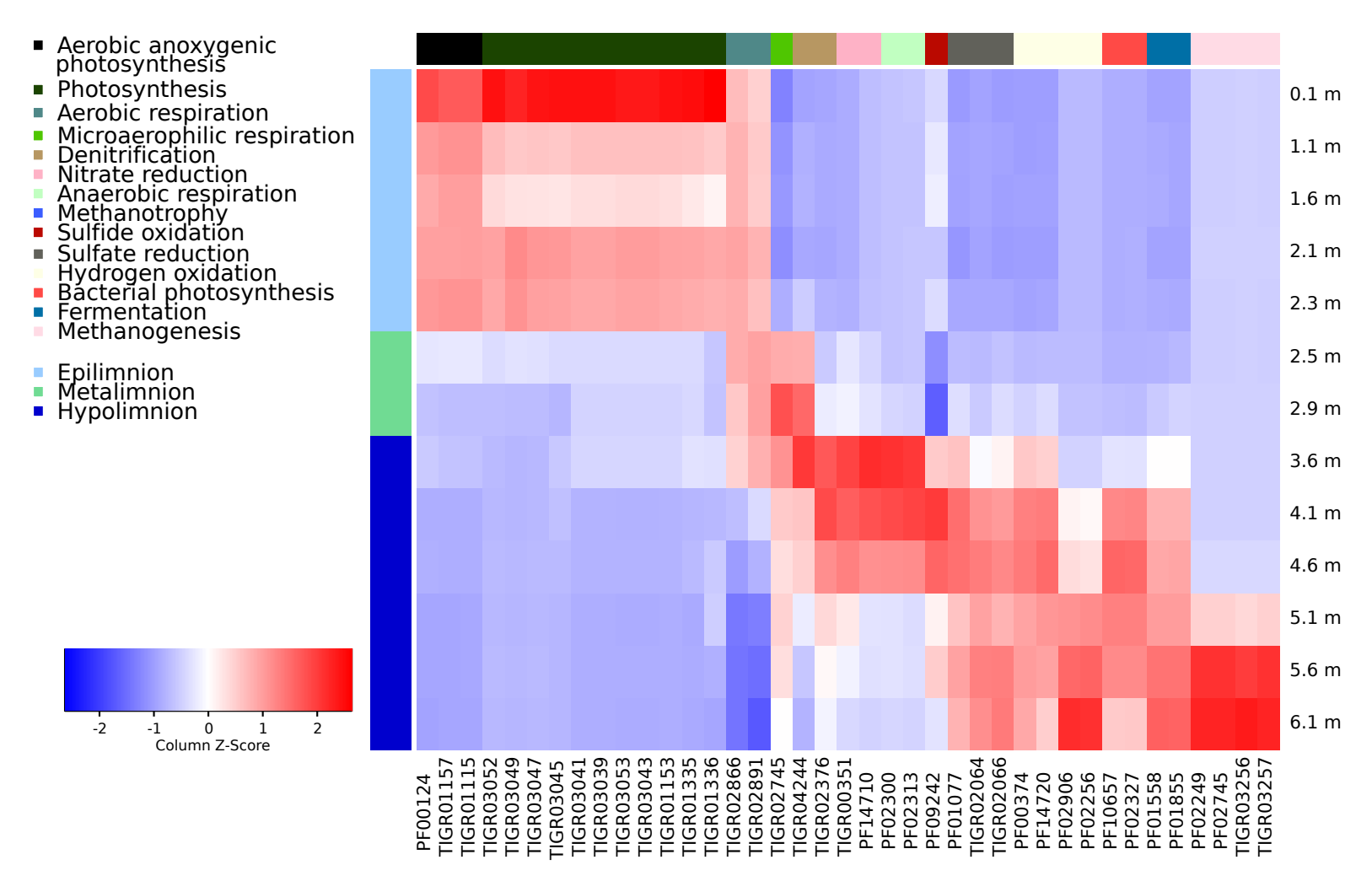
685 maximum abundance and in the case of multiple MAGs of the same taxon, the height
686 of the box is covering the depths where the organisms were most abundant. Also,
687 for the organisms with multiple MAGs with the same taxonomy, dominant pathways
688 are displayed. Only the MAGs with marker genes specific for inorganic carbon
689 fixation pathways are presented.
690
691 **Supplementary Tables**
692 **Supplementary Table S1.** A table listing HMMs used to compare the abundance of
693 different energy and carbon assimilation pathways between different depths layers
694 and p-values of these comparisons.
695
696 **Supplementary Table S2.** Characteristics of the high-quality MAGs. Bin Id,
697 taxonomic annotation by i) manual annotation from a phylogenetic tree from
698 PhyloPhlan and ii) PhyloPhlan automatic annotation, size of the MAG, average
699 proportion of the MAG among all samples, depth of the maximum abundance and
700 the proportion in this depth, completeness, contamination and strain heterogeneity
701 of the MAG and number of marker genes for different energy and carbon
702 assimilation pathways in the MAG. Presence of key genes of each of the pathway is
703 shown with letters. AAP: M = pufM; methanogenesis: M = mcrA; CBB: R = RuBisCo;
704 PPP: Z = glucose-6-phosphate dehydrogenase, P = 6-phosphogluconolactonase, G =
705 6-phosphogluconate dehydrogenase; rTCA: C = citrate lyase, O = oxoglutarate
706 synthase, P = pyruvate synthase; WL: C = CO dehydrogenase; 3HP: A = acetyl-CoA
707 carboxylase, P = propionyl-CoA carboxylase. For hydrogenases sub groups are

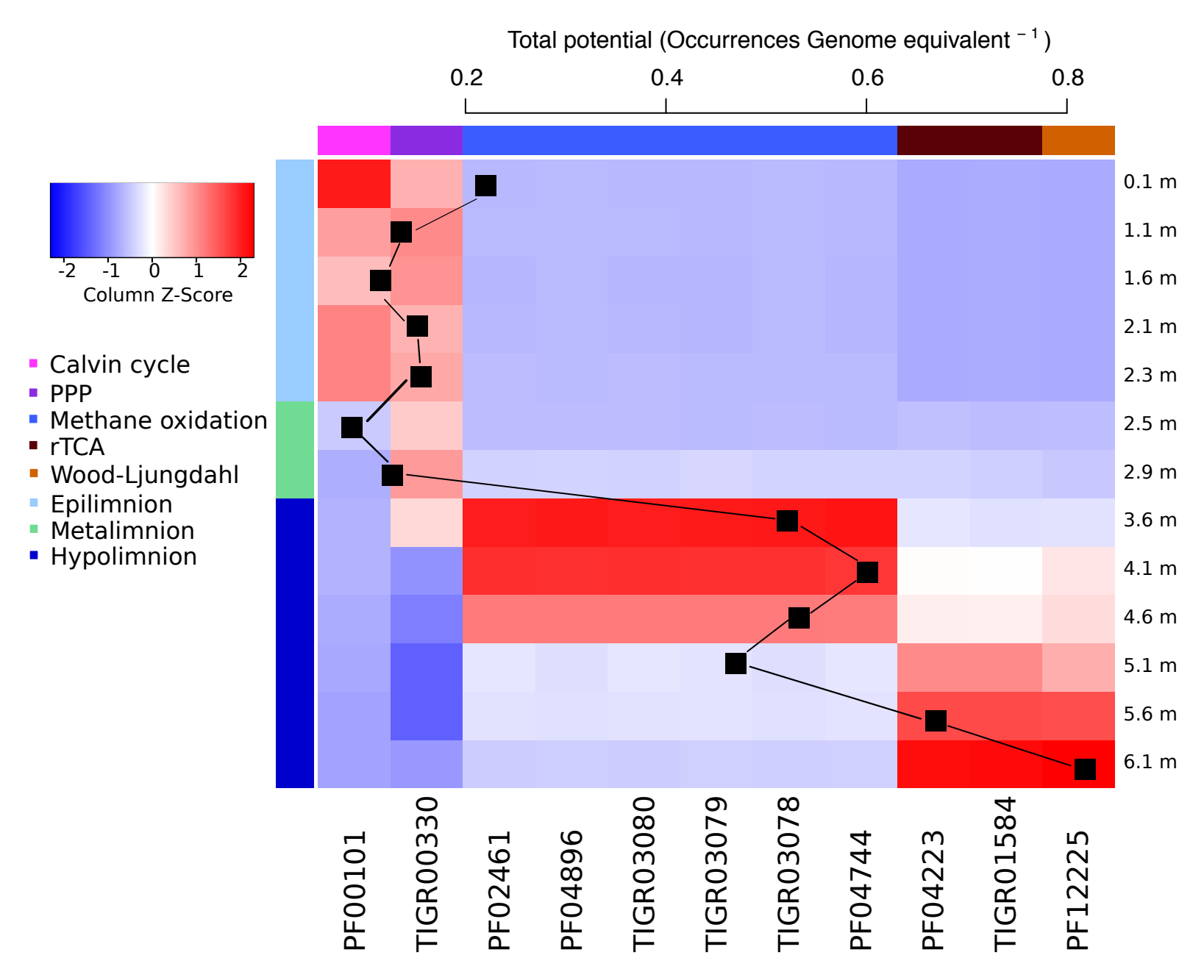
708 presented as in D. Søndergaard, C.N. Pedersen, and C Greening. Sci Rep 6:34212,
709 2016.
710
711 **Supplementary Figures**
712 **Supplementary Figure S1.** Change in dissolved inorganic carbon concentration
713 during 24 h dark incubations at 3 m depth in 2008-2010. In 2008 measurements
714 were done biweekly, in 2009 monthly and in 2010 three times during the open
715 water season.
716
717 **Supplementary Figure S2.** Photosynthetically active radiation in the lake water
718 column.
719
720 **Supplementary Figure S3.** The average abundance of marker genes for different
721 pathways related to energy generation.
722
723 **Supplementary Figure S4.** Phylogenetic tree of the Metagenome Assembled
724 Genomes depicting their taxonomic placement in the tree of life.


Stratification
Autumn mix
8
6
4
2
0
-2
-4
-6
Uptake/release of CO2 (mg CO2C m-3 h-1)
1.6.2008
1.7.2008
1.8.2008
1.9.2008
1.10.2008
1.11.2008
1.12.2008
Date


A
B
C
Depth (m)
0.5
2.5
4.5
5.5
0
2
4
6
0
25
50
75
100
Acidobacteria
Actinobacteria
Alphaproteobacteria
Armatimonadetes
Bacteroidetes
Betaproteobacteria
Chlorobia
Chloroflexi
Deltaproteobacteria
Elusimicrobia
Euryarchaeota
Firmicutes
Gammaproteobacteria
Parcubacteria
Verrucomicrobia
Others


A
B
C
D
E
Depth (m)
CH4 (µM)
O2 (mg L−1) / Temperature (C°)
CH4
O2
Temp
NH4 (µg l)−1 / SO4(mg l)−1
NO2 / NO3 µg l−1
NH4
NO2
NO3
SO4
PO4 (µg l)−1
DOC (mg l−1)
DOC
PO4
Fe (µM)
Fe2+
Fe3+
rM
Distance class
1 = 0.2 - 1.0
2 = 1.0 - 1.9
3 = 1.9 - 2.7
4 = 2.7 - 3.5
5 = 3.5 - 4.3
6 = 4.3 - 5.2
7 = 5.2 - 6.0


Aerobic anoxygenic photosynthesis
Photosynthesis
Aerobic respiration
Microaerophilic respiration
Denitrification
Nitrate reduction
Anaerobic respiration
Methanotrophy
Sulfide oxidation
Sulfate reduction
Hydrogen oxidation
Bacterial photosynthesis
Fermentation
Methanogenesis
Epilimnion
Metalimnion
Hypolimnion
-2
-1
0
1
2
Column Z-Score
0.1 m
1.1 m
1.6 m
2.1 m
2.3 m
2.5 m
2.9 m
3.6 m
4.1 m
4.6 m
5.1 m
5.6 m
6.1 m
PF00124
TIGR01157
TIGR01115
TIGR03052
TIGR03049
TIGR03047
TIGR03045
TIGR03041
TIGR03039
TIGR03053
TIGR03043
TIGR01153
TIGR01335
TIGR01336
TIGR02866
TIGR02891
TIGR02745
TIGR04244
TIGR02376
TIGR00351
PF14710
PF02300
PF02313
PF09242
PF01077
TIGR02064
TIGR02066
PF00374
PF14720
PF02906
PF02256
PF10657
PF02327
PF01558
PF01855
PF02249
PF02745
TIGR03256
TIGR03257


Total potential (Occurrences Genome equivalent $^{-1}$)
0.2
0.4
0.6
0.8
-2
-1
0
1
2
Column Z-Score
Calvin cycle
PPP
Methane oxidation
rTCA
Wood-Ljungdahl
Epilimnion
Metalimnion
Hypolimnion
0.1 m
1.1 m
1.6 m
2.1 m
2.3 m
2.5 m
2.9 m
3.6 m
4.1 m
4.6 m
5.1 m
5.6 m
6.1 m
PF00101
TIGR00330
PF02461
PF04896
TIGR03080
TIGR03079
TIGR03078
PF04744
PF04223
TIGR01584
PF12225

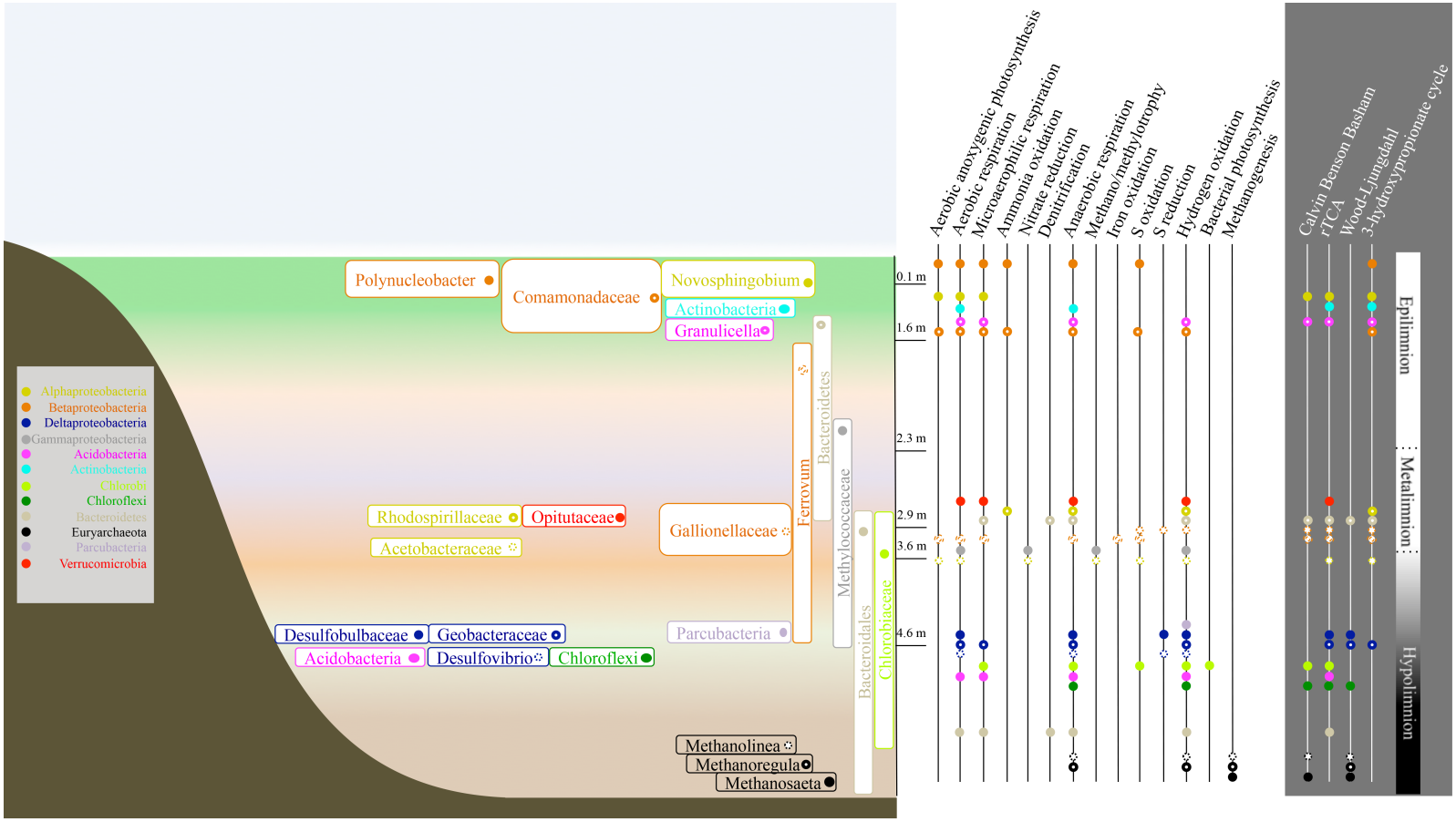

Polynucleobacter
Comamonadaceae
Novosphingobium
Actinobacteria
Granulicella
Ferrovum
Bacteroidetes
Methylococcaceae
Bacteroidales
Chlorobiaceae
Rhodospirillaceae
Opitutaceae
Acetobacteraceae
Gallionellaceae
Desulfobulbaceae
Geobacteraceae
Acidobacteria
Desulfovibrio
Chloroflexi
Parcubacteria
Methanolinea
Methanoregula
Methanosaeta
Alphaproteobacteria
Betaproteobacteria
Deltaproteobacteria
Gammaproteobacteria
Acidobacteria
Actinobacteria
Chlorobi
Chloroflexi
Bacteroidetes
Euryarchaeota
Parcubacteria
Verrucomicrobia
0.1 m
1.6 m
2.3 m
2.9 m
3.6 m
4.6 m
Aerobic anoxygenic photosynthesis
Aerobic respiration
Microaerophilic respiration
Ammonia oxidation
Nitrate reduction
Denitrification
Anaerobic respiration
Methano/methylotrophy
Iron oxidation
S oxidation
S reduction
Hydrogen oxidation
Bacterial photosynthesis
Methanogenesis
Calvin Benson Basham
rTCA
Wood-Ljungdahl
3-hydroxypropionate cycle
Epilimnion
Metalimnion
Hypolimnion

Table 1. Sampling depth, water layer, data size, assembly coverage in each sample, and inverse Simpson index and Pielou's evenness of the 16S rRNA $OTU_{0.03}$ data.

Depth (m)	Layer	Raw data (Gb)	Coverage (%)	Inverse Simpson index	Pielou's evenness
0.1	Epilimnion	8.06	61.86	9.90	0.49
1.1	Epilimnion	10.75	73.54	13.39	0.52
1.6	Epilimnion	9.54	75.49	12.01	0.50
2.1	Epilimnion	10.55	72.24	11.52	0.50
2.3	Epilimnion	10.24	73.33	12.50	0.51
2.5	Metalimnion	9.75	67.29	19.76	0.61
2.9	Metalimnion	9.68	61.09	22.28	0.62
3.6	Hypolimnion	10.52	65.57	28.52	0.62
4.1	Hypolimnion	10.51	63.62	9.69	0.53
4.6	Hypolimnion	9.26	60.39	8.41	0.51
5.1	Hypolimnion	8.31	48.72	8.39	0.53
5.6	Hypolimnion	8.05	43.77	11.50	0.60
6.1	Hypolimnion	5.24	40.21	35.97	0.69