

Supplementary Information

CO₂ assimilation strategies in stratified lakes: diversity and distribution patterns of chemolithoautotrophs

A. Alfreider, A. Baumer, T. Bogensperger, T. Posch, M. M. Salcher and M. Summerer

Supplementary Methods

Optimization of the functional gene approach

Flow cytometry

Supporting Tables and Figures

Table S1

Specification of primers applied for the detection of different CO₂ fixation pathways

Table S2

Summary of physico-chemical parameters in the studied lakes

Fig. S1

Neighbor-Joining trees obtained from deduced amino acid sequences of forms IA (**Fig. S1A**), IC (**Fig. S1B**) and II (**Fig. S1C**) RubisCO genes in the CBB cycle

Fig. S2

Coverage of two different primer sets targeting Form II RubisCO sequences

Fig. S3

Coverage of two different primer sets targeting Form IC RubisCO sequences

Fig. S4

Evaluation of different annealing temperatures for thaumarchaeal *hcd* qPCR primers

Fig. S5

Evaluation of the specificity of *hcd* qPCR primers

Fig. S6

Vertical profiles of ammonium and nitrate in the studied lakes

Fig. S7

Vertical profile of autotrophic cell numbers in Lake Egelsee

Supplementary References (not included in main text)

Supplementary Methods

Optimization of the functional gene approach

In this study, a major effort was undertaken to (re)evaluate already published PCR-based marker systems and to design and test new primer pairs before they were applied on a routine basis at our study sites. Table S1 summarizes the specification of primers that were used. Not included in this table are other published primer pairs that were tested (in silico and/or in vitro), but were not appropriate for the investigation of lake water samples from our study sites.

Three sets of oligonucleotide primers were designed and successfully used for PCR and RT-PCR amplification of RubisCO form IA, form IC and form II gene fragments as applied in groundwater samples by Alfreider *et al.* (2003, 2009, 2012) and in exploratory studies in selected lakes (data not shown). Nonetheless, the specificity and coverage of these primers was evaluated based on new RubisCO sequences available in public sequence databases and recently published primer sets. For example, Kato *et al.* (2012) designed a PCR primer set targeting RubisCO form II gene where the coverage (in silico) was reported to be superior compared with the primers designed by Alfreider *et al.* (2003). However, in vivo validation of both primers systems revealed a different picture (Fig. S2) and individual primer sets covered only a part of the *cbbM* diversity observed in our study sites. In order to increase the coverage of *cbbM* genes, selected lake water samples were amplified with both primers sets for sequence analysis. For the same reason, two primers pairs were also used to study the diversity of the form IC RubisCO (Table S1, Fig. S3). Form IA genes were targeted with one primer pair because other primer pairs tested were not specific (resulting in multiple banding of PCR products).

Autotrophy in *Thaumarchaeota* was investigated by the analysis of genes coding for 4-hydroxybutyryl-CoA dehydratase (*hcd*), a key protein in the HP/HB and dicarboxylate/4-hydroxybutyrate (DC/4-HB) cycle (Berg, 2011). Several PCR based protocols were developed, and we selected two primers sets that produced specific amplification products in our samples (Offre *et al.* 2010; Yakimov *et al.* 2011, Table S1). For the quantification of *hcd* genes, a real time PCR primer pair (qPCR_hcd_f/qPCR_hcd_r, Table 1) was designed that specifically covers the sequence diversity of *hcd* genes obtained in this study. This primer was specifically designed on basis of the *hcd* sequences derived from PCR products with primers hcd-465F/hcd-1267R and 4HBD312F/4HBD1360R (Table

S1) of lakes ACH, STA and ZUR (Table S1, Fig. 2), and sequences of representatives within the marine Group I (MGI) *Thaumarchaeota* lineage. Gel electrophoresis and sequence analysis of selected PCR-amplificates produced with the qPCR primer from different samples of lakes ACH, STA and ZUR support the appropriateness of the qPCR assay (Fig. S3 and S4).

CO₂ fixation based on the rTCA cycle is generally detected by targeting genes coding for the alpha or beta subunit of the ATP citrate lyase (*aclAB* genes), and/or the alpha subunit of 2-oxoglutarate:ferredoxin oxidoreductase enzymes (*oorA* genes). Published primer sets, which were tested on samples from our study sites, mostly produced multiple bands or faint amplification products. Furthermore, sequence analysis of randomly selected PCR amplicons of *aclAB* genes produced nucleotide sequences that were mostly not affiliated with the target genes (data not shown). In fact, most published primers were originally designed to study marine hydrothermal environments; therefore, these primers specifically target thermophilic *Aquificales* and *Epsilonproteobacteria* (reviewed by Hügler and Sievert, 2011). Further primer optimization was necessary and we developed a broad range detection system, based on a nested PCR approach, targeting genes coding for ATP citrate lyase alpha subunit (*aclA*) within different clades of *Epsilonproteobacteria*, *Deltaproteobacteria*, *Acidobacteria*, and *Nitrospirae*. A second primer set was developed to specifically target *aclA* genes in *Nitrospirae*. Finally, eleven primer systems were used to amplify different functional genes coding for key enzymes in the Calvin cycle, the HP/HB cycle, and the rTCA cycle in samples from six stratified lakes (Table S1).

Flow cytometry

Total prokaryotic abundances and numbers of photoautotrophic cells in lakes EGE and ZUR were quantified by an inFlux V-GS cell sorter (Becton Dickinson, Franklin Lakes, NJ, USA) equipped with a ultraviolet (UV) laser (Lightwave Electronics, CY-PS, 60 mW, wavelength of 355 nm), a blue laser (Coherent, Sapphire, 200 mW, wavelength of 488 nm), and detectors for two scatter and six fluorescence channels. All samples were stained with 4,6-diamidino-2-phenylindole (DAPI, 1 mg ml⁻¹ final concentration), and scatter plots were analyzed with an in-house software (J Villiger, unpublished).

Table S1. Specification of primers applied for the detection of different CO₂ fixation pathways

| Pathway, Target gene | Primer Name | Sequence (5'-3') | bp | Ta ¹ | Reference |
|---|------------------------------|--|------|-----------------|--|
| CBB-cycle | | | | | |
| RubisCO form IA | cbbL_IA_f cbbL_IA_r | CGGCACSTGGACCACSGTSTGGAC GTARTCGTGCATGATGATSGG | 620 | 54 | Alfreider <i>et al.</i> (2003) |
| RubisCO form IC | cbbL_IC_f cbbL_IC_r | GAACATCAAYTCKCAGCCCTT TGGTGCATCTGVCCGGCRTG | 552 | 55 | Alfreider <i>et al.</i> (2009) |
| | cbbLR1F cbbLR1R | AAGGAYGACGAGAACATC TCGGTCGGS GTAGTTGAA | 818 | 57 | Selesi <i>et al.</i> (2009) |
| RubisCO form II | cbbM_f cbbM_r | GGCACCATCATCAAGCCCAAG TCTTGCCGTAGCCCATGGTGC | 505 | 53 | Alfreider <i>et al.</i> (2003) |
| | cbbM343F cbbM1226R | GGYAAAYAACCARGGYATGGG CGYARBGCRTTCATRCRC | 783 | 50 | Kato <i>et al.</i> (2012) |
| HP/HB cycle | | | | | |
| 4-hydroxybutyryl- CoA dehydratase | hcd-465F hcd-1267R | GGHGGTGCWATGACTGA CTCATTCTGTTTTCHACATC | 839 | 53 | Offre <i>et al.</i> (2010) |
| | 4HBD 312F 4HBD 1360R | TTCCAAAGATGTGTGYYGGWATGG CCATGCATTGATTCHGTAA | 1048 | 50 | Yakimov <i>et al.</i> (2011) |
| Quantitative PCR | qPCR_hcd_f qPCR_hcd_r | GACTGATCCWAAAGGDGAYAGAAG CCYTTARCATCTGCWGGGAATTGC | 227 | 56 | This study |
| Acetyl-CoA/propionyl- CoA carboxylase | accA_f accA_r | CCNGCHATGACNGAYTTTG ATRTCWARYGCRCCWGCAAG | 481 | 49 | This study |
| rTCA pathway | | | | | |
| ATP citrate lyase (broad range) | aclA_f680 aclA_r1491 | TNGGHGARRTNGGNGG AYDCKRTGNCCDATNCC | 811 | 50 | This study (nested PCR, primer set 1) ² |
| | aclA_f807 aclA_r1371 | TDMARTTYGGNCAYGCNNGNGC GCNCCDCCRAANCKNNGNCCDAT | 564 | 55 | This study (nested PCR, primer set 2) ² |
| ATP-citrate lyase (<i>Nitrospira</i>) | aclA_Nitro_f aclA_Nitro_r | GGCATGAAGGGTGAAGGATC AACTCTCACGTACCCGACC | 714 | 52 | This study |

¹Ta: Annealing temperature; ² Details for nested PCR see Material and Method section

Table S2. Summary of physico-chemical parameters in the studied lakes

| Depth (m) | Temp* °C | pH | Cond** µS cm ⁻¹ | O ₂ mg L ⁻¹ | Ptot [†] µg L ⁻¹ | NH ₄ ⁺ µg L ⁻¹ | NO ₃ ⁻ µg L ⁻¹ | SO ₄ ²⁻ mg L ⁻¹ | DOC µg L ⁻¹ | HCO ₃ ⁻ mmol L ⁻¹ |
|------------|-------------|------|-------------------------------|--------------------------------------|---|--|--|---|---------------------------|---|
| ZUR | | | | | | | | | | |
| 20 | 7.4 | 7.58 | 296.0 | 4.8 | 10.0 | 5 | 784 | 14.90 | 1250 | 2.62 |
| 60 | 4.4 | 7.76 | 293.0 | 8.4 | 15.0 | b.d. | 764 | 15.00 | n.d. | 2.53 |
| 80 | 4.2 | 7.73 | 295.0 | 7.2 | 26.0 | b.d. | 764 | 15.10 | n.d. | 2.57 |
| 100 | 4.2 | 7.68 | 297.0 | 5.8 | 38.0 | 3 | 777 | 15.10 | n.d. | 2.60 |
| 115 | 4.2 | 7.58 | 301.0 | 2.8 | 50.0 | 3 | 797 | 14.80 | n.d. | 2.62 |
| 120 | 4.2 | 7.49 | 302.0 | 1.0 | 99.0 | 165 | 443 | 14.80 | n.d. | 2.66 |
| 130 | 4.2 | 7.47 | 306.0 | 0.1 | 152.0 | 460 | 210 | 14.10 | n.d. | 2.74 |
| 135 | 4.2 | 7.45 | 307.0 | 0.1 | 155.0 | 562 | 192 | 13.70 | n.d. | 2.75 |
| ACH | | | | | | | | | | |
| 0 | 10.2 | 8.37 | 274 | 9.6 | 1.8 | 1 | 406 | 2.88 | 1136 | 2.82 |
| 10 | 10.4 | 8.42 | 274 | 9.8 | 2.4 | 1 | 407 | 2.88 | 1136 | 2.83 |
| 20 | 9.7 | 8.39 | 277 | 10.1 | 2.4 | 1 | 408 | 2.88 | 1251 | 2.83 |
| 25 | 7.5 | 8.32 | 277 | 10.4 | 2.1 | 1 | 460 | 2.98 | 1149 | 2.82 |
| 30 | 5.9 | 8.29 | 278 | 10.4 | 1.8 | 1 | 484 | 3.00 | 1428 | 2.83 |
| 60 | 5.0 | 8.29 | 279 | 10.5 | 1.5 | 1 | 478 | 3.07 | 1039 | 2.86 |
| 100 | 5.0 | 8.23 | 282 | 9.9 | 1.5 | 1 | 477 | 3.12 | 1011 | 2.88 |
| 129 | 4.9 | 8.19 | 284 | 9.8 | 2.4 | 5 | 467 | 3.16 | 1027 | 2.90 |
| STA | | | | | | | | | | |
| 1 | 20.7 | 8.50 | 292 | 8.9 | 3.8 | 5 | 206 | 8.18 | 405 | 2.60 |
| 9 | 18.3 | 8.58 | 295 | 8.9 | 5.5 | 5 | 217 | 8.22 | 399 | 2.60 |
| 10 | 15.5 | 8.62 | 293 | 9.3 | 4.1 | 5 | 207 | 8.22 | 399 | 2.59 |
| 13 | 10.9 | 8.26 | 317 | 9.7 | 6.1 | 1 | 382 | 8.40 | 512 | 2.81 |
| 19 | 6.9 | 8.09 | 321 | 9.1 | 6.3 | 7 | 451 | 8.55 | 547 | 2.81 |
| 30 | 4.8 | 8.06 | 322 | 9.5 | 4.5 | 1 | 462 | 8.64 | 552 | 2.81 |
| 80 | 4.2 | 8.09 | 320 | 10.1 | 3.6 | 3 | 446 | 8.63 | 553 | 2.82 |
| 100 | 4.2 | 8.07 | 322 | 9.6 | 4.2 | 2 | 442 | 8.59 | 538 | 2.84 |
| HEC | | | | | | | | | | |
| 0 | 25.4 | 8.62 | 310 | 9.3 | 7.4 | 5 | 254 | 4.32 | 3659 | 3.02 |
| 5 | 15.7 | 8.55 | 350 | 13.1 | 12.0 | 7 | 358 | 5.24 | 3628 | 3.43 |
| 10 | 7.2 | 8.17 | 366 | 8.3 | 8.8 | 28 | 445 | 5.88 | 2994 | 3.54 |
| 12 | 6.1 | n.d. | n.d. | 3.4 | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. |
| 14 | 5.7 | n.d. | n.d. | 0.3 | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. |
| 15 | 5.7 | 7.83 | 388 | 0.4 | 11.4 | 29 | 459 | 6.50 | 2752 | 3.76 |
| 25 | 6.3 | 7.39 | 539 | b.d. | 167.5 | 4168 | 7 | 4.02 | 2176 | 5.54 |
| 50 | 6.7 | 7.34 | 578 | b.d. | 320.7 | 5638 | 18 | 2.56 | 2434 | 5.97 |
| PIB | | | | | | | | | | |
| 0 | 17.2 | 7.68 | 67 | 8.8 | 4.4 | 15 | 184 | 5.78 | 2382 | 0.463 |
| 3 | 16.9 | 7.79 | 67 | 8.8 | 5.6 | 15 | 184 | 5.75 | 2368 | 0.462 |
| 6 | 12.2 | 7.50 | 70 | 10.4 | 10.9 | 10 | 168 | 6.14 | 2194 | 0.493 |
| 9 | 7.2 | 7.01 | 73 | 7.2 | 7.6 | 7 | 220 | 6.38 | 1870 | 0.510 |
| 12 | 5.6 | 6.84 | 73 | 4.5 | 8.8 | 15 | 251 | 6.34 | 1754 | 0.512 |
| 15 | 4.9 | 6.71 | 74 | 1.6 | 7.1 | 60 | 251 | 6.29 | 1724 | 0.522 |
| 18 | 4.8 | 6.66 | 75 | 0.2 | 11.8 | 120 | 167 | 6.09 | 1814 | 0.540 |
| 21 | 4.6 | 6.65 | 77 | 0.1 | 12.6 | 240 | 54 | 5.99 | 1823 | 0.566 |
| 24 | 4.6 | 6.65 | 79 | 0.0 | 19.7 | 415 | 8 | 5.66 | 1949 | 0.586 |
| EGE | | | | | | | | | | |
| 5 | 15.7 | n.d. | 393 | 7.01 | 16.5 | 4 | 7 | 68.8 | 3524 | 4.21 |
| 5.5 | 14.3 | n.d. | 418 | 5.94 | 33.5 | 6 | 5 | 53.9 | 3747 | 4.53 |
| 5.75 | 13.4 | n.d. | 421 | 4.21 | 45.8 | 4 | 7 | 49.4 | 3788 | 4.63 |
| 6 | 12.8 | n.d. | 428 | 2.04 | 64.3 | 5 | 7 | 46.6 | 3744 | 4.69 |
| 6.25 | 11.4 | n.d. | 437 | 0.93 | 89.2 | 6 | 11 | 48.0 | 3753 | 4.73 |
| 6.5 | 10.8 | n.d. | 441 | 0.66 | 133.3 | 261 | 10 | 37.0 | 4014 | 4.745 |
| 6.75 | 10.0 | n.d. | 448 | 0.47 | 89.7 | 788 | 0 | 22.7 | 3461 | 4.87 |
| 7 | 9.3 | n.d. | 453 | 0.36 | 82.0 | 1297 | 0 | 16.5 | 3497 | 4.97 |
| 8.5 | 6.5 | n.d. | 505 | b.d. | 79.0 | 3840 | 0 | 15.4 | 5516 | 5.12 |
| 10 | 5.7 | n.d. | 593 | b.d. | 152.5 | 8532 | 0 | 15.8 | 6590 | 5.52 |

* Temperature

** Conductivity

Fig. S1. Neighbor-Joining trees obtained from deduced amino acid sequences of forms IA (Fig. S1A), IC (Fig. S1B) and II (Fig. S1C) RubisCO genes in the CBB cycle derived from this study (shown in blue letters) and representative sequences from NCBI and IMG databases. Bootstrap values are shown as percentages of 1000 replicates and values over 50% are indicated on nodes. Scale bar indicates 5 % changes in the amino acid sequence. Vertical bars with different colors, as identified in the legend inset, indicate phylogenetic affiliation of closest relatives. Where applicable, potential functional traits derived from physiological properties of the cultured closest relatives are designated by differently colored dots.

Fig. S1A

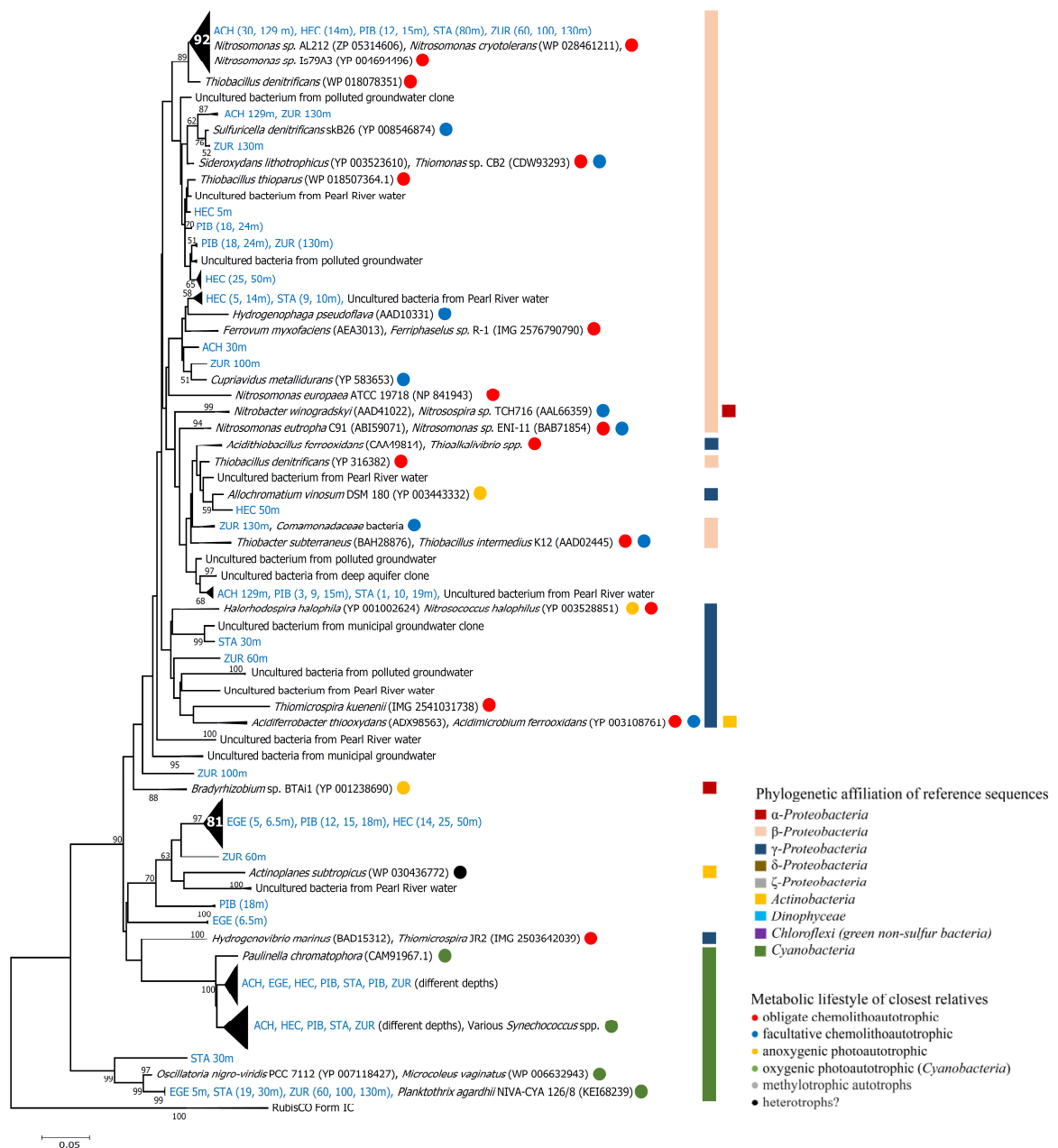


Fig. S1B

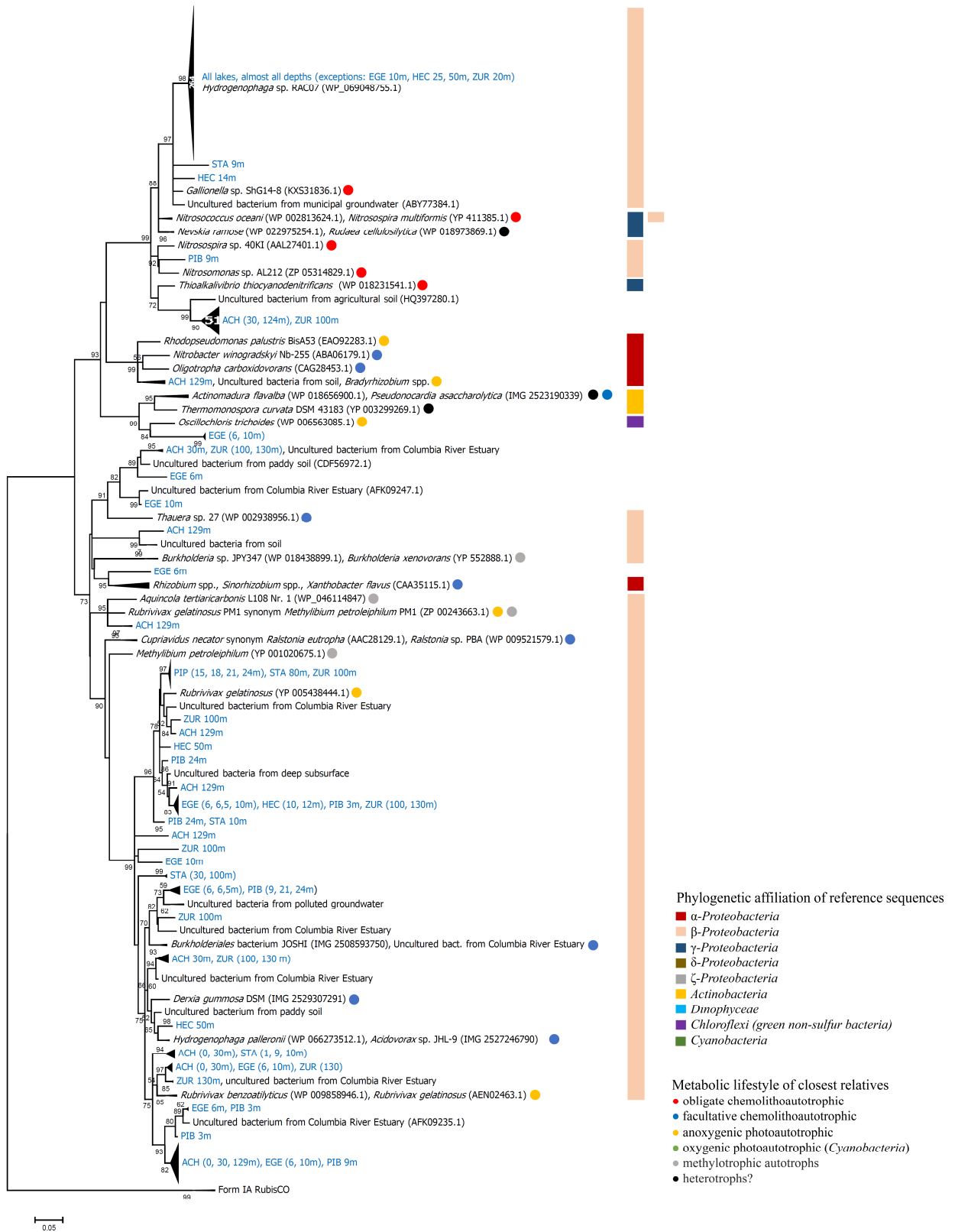


Fig. S1C

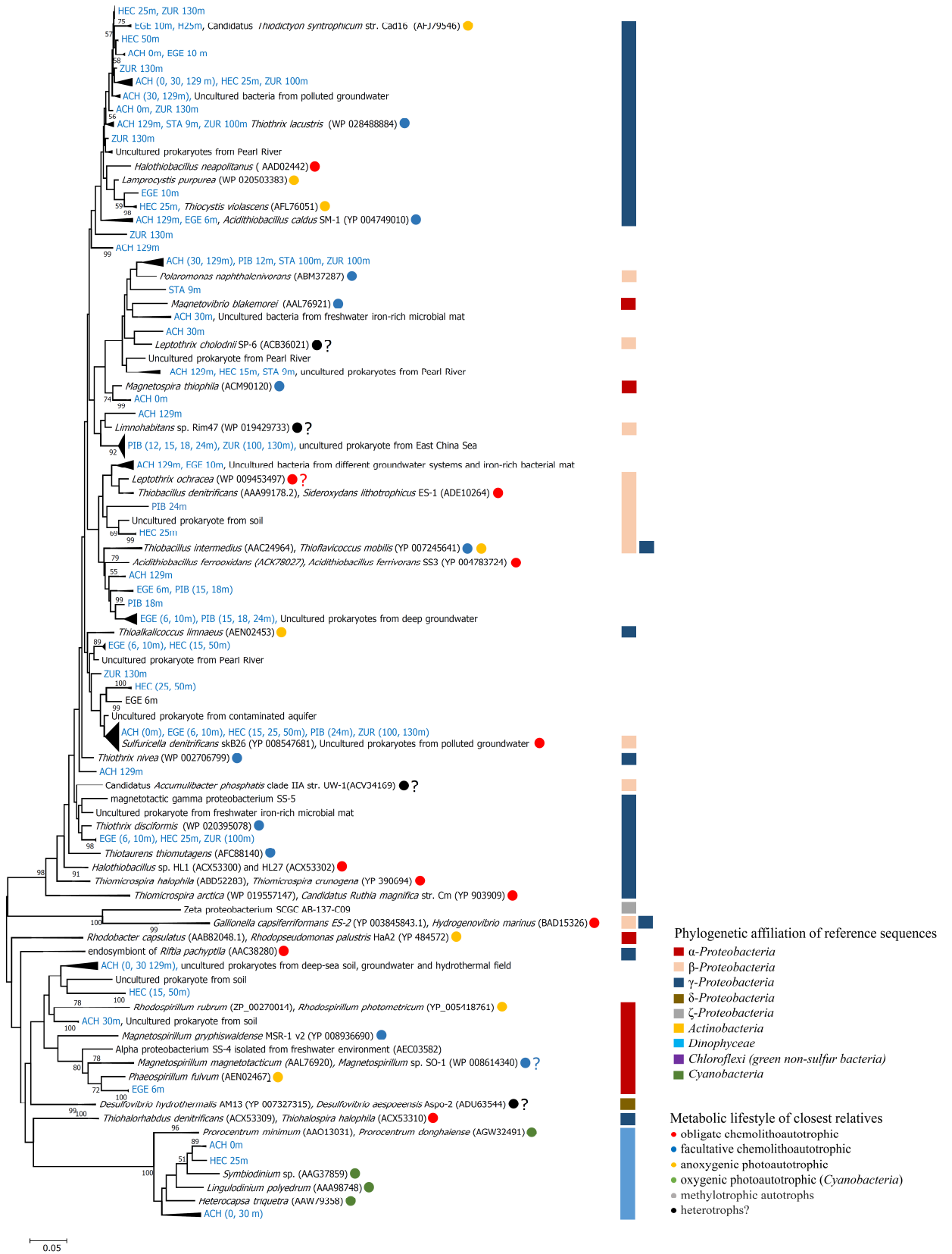


Fig. S2. Coverage of two different primer sets targeting Form II RubisCO sequences (see Table S1) retrieved from different samples of the studied lakes. The green color indicates *cbbM* lineages that are covered by both primer pairs (Alfreider *et al.*, 2003 and Kato *et al.*, 2012).

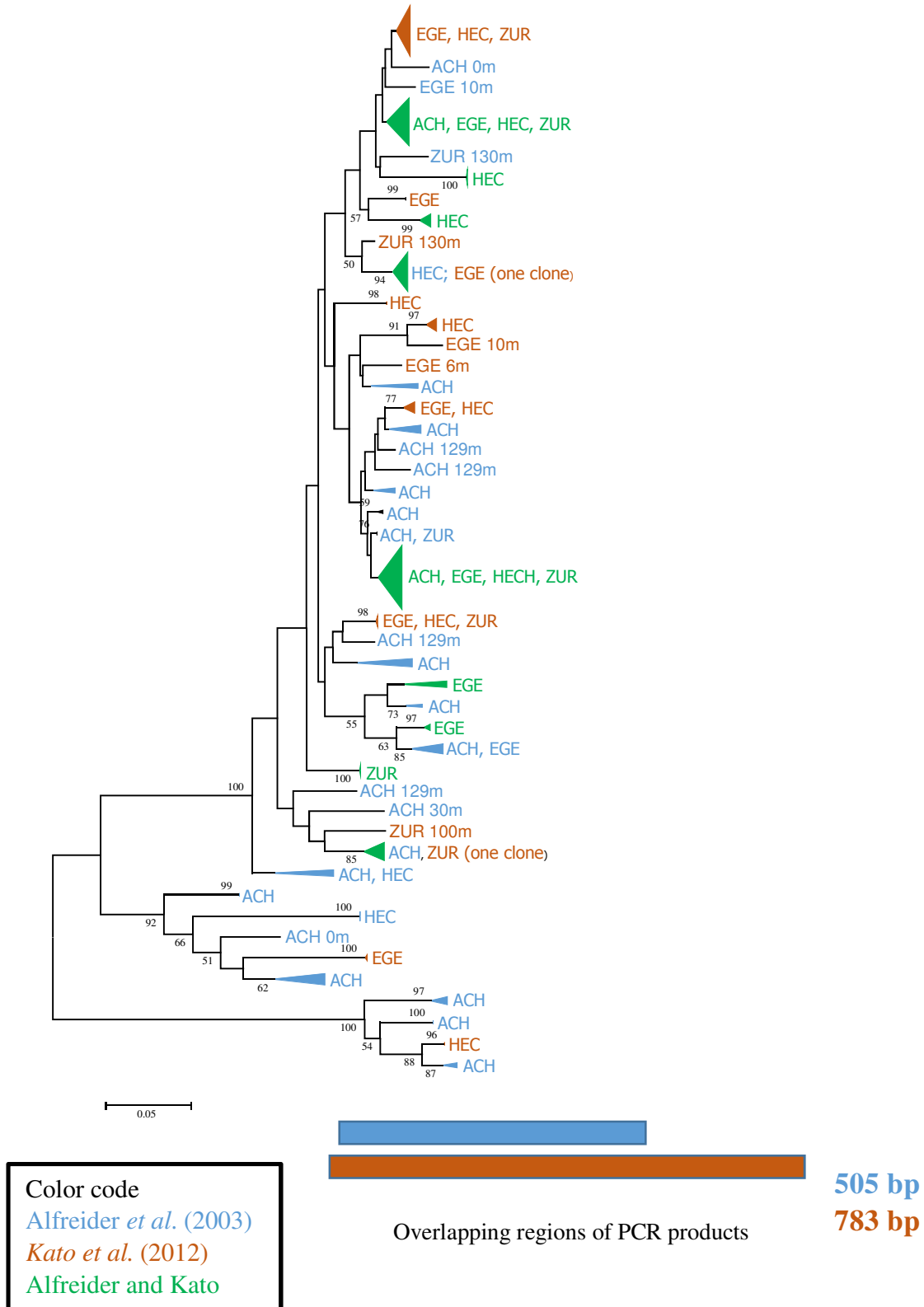


Fig. S3. Coverage of two different primer sets targeting Form 1C RubisCO sequences (see Table S1) retrieved from different samples of the studied lakes. The green color indicates *cbbL* IC lineages that are covered by both primer pairs (Alfreider *et al.*, 2009 and Selesi *et al.*, 2009).

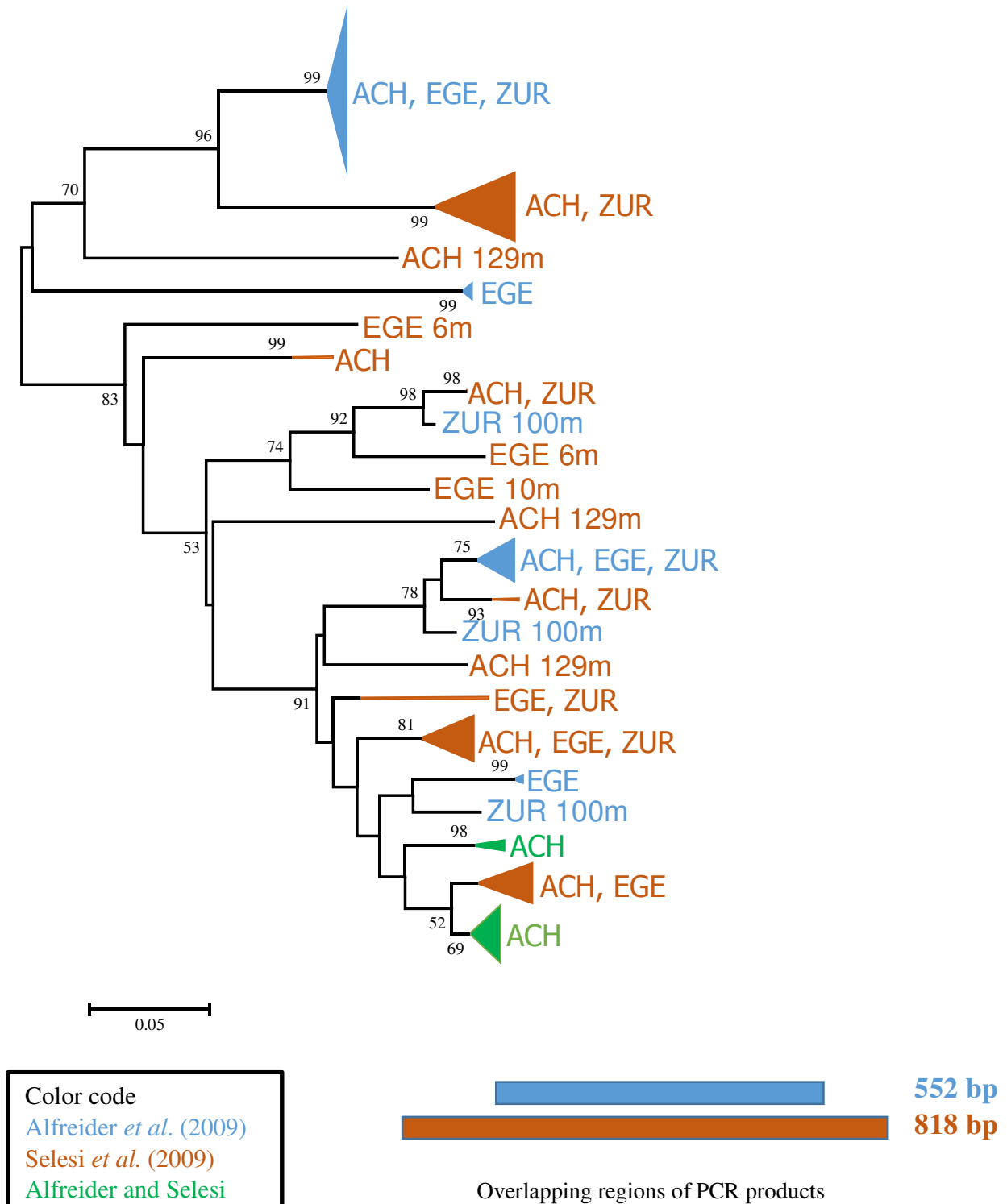


Fig. S4. Evaluation of different annealing temperatures for thaumarchaeal *hcd* qPCR primers with different samples of lakes ACH (0m, 129m), STA (30m) and ZUR (130). A sample from lake PIB (18m) was used as negative control.

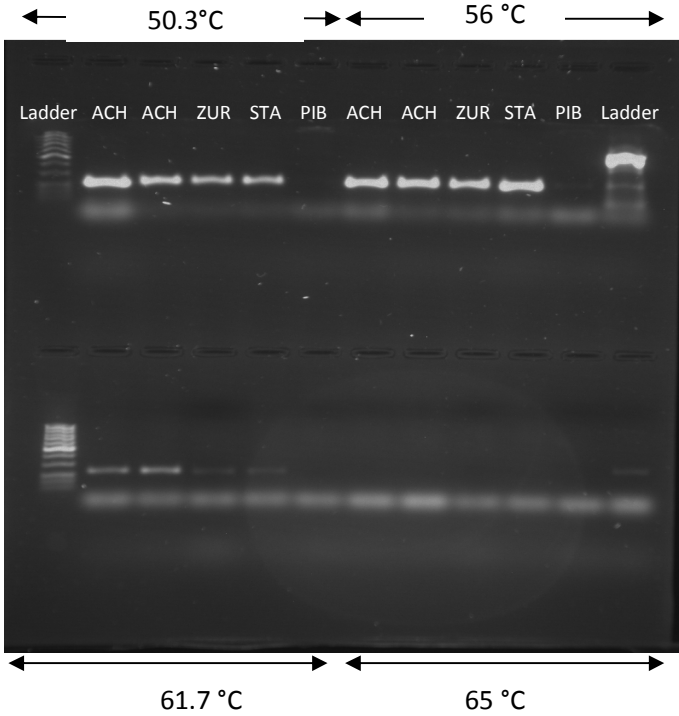


Fig. S5. Evaluation of the specificity of *hcd* qPCR primers (annealing temperature 56°C) by sequencing of cloned PCR products derived from samples of lakes ACH, STA and ZUR.

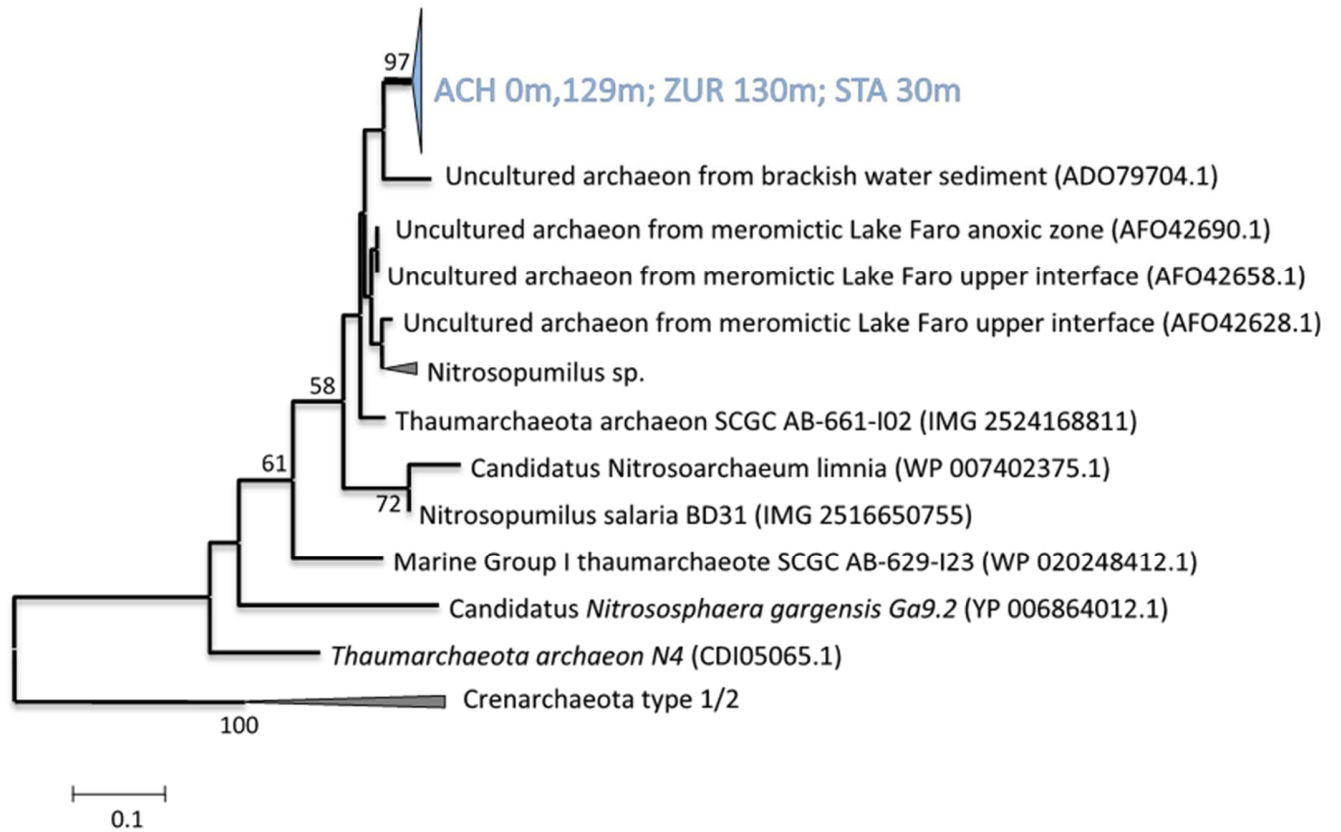


Fig. S6. Vertical profiles of ammonium, nitrate and oxygen in the studied lakes at the time of sampling.

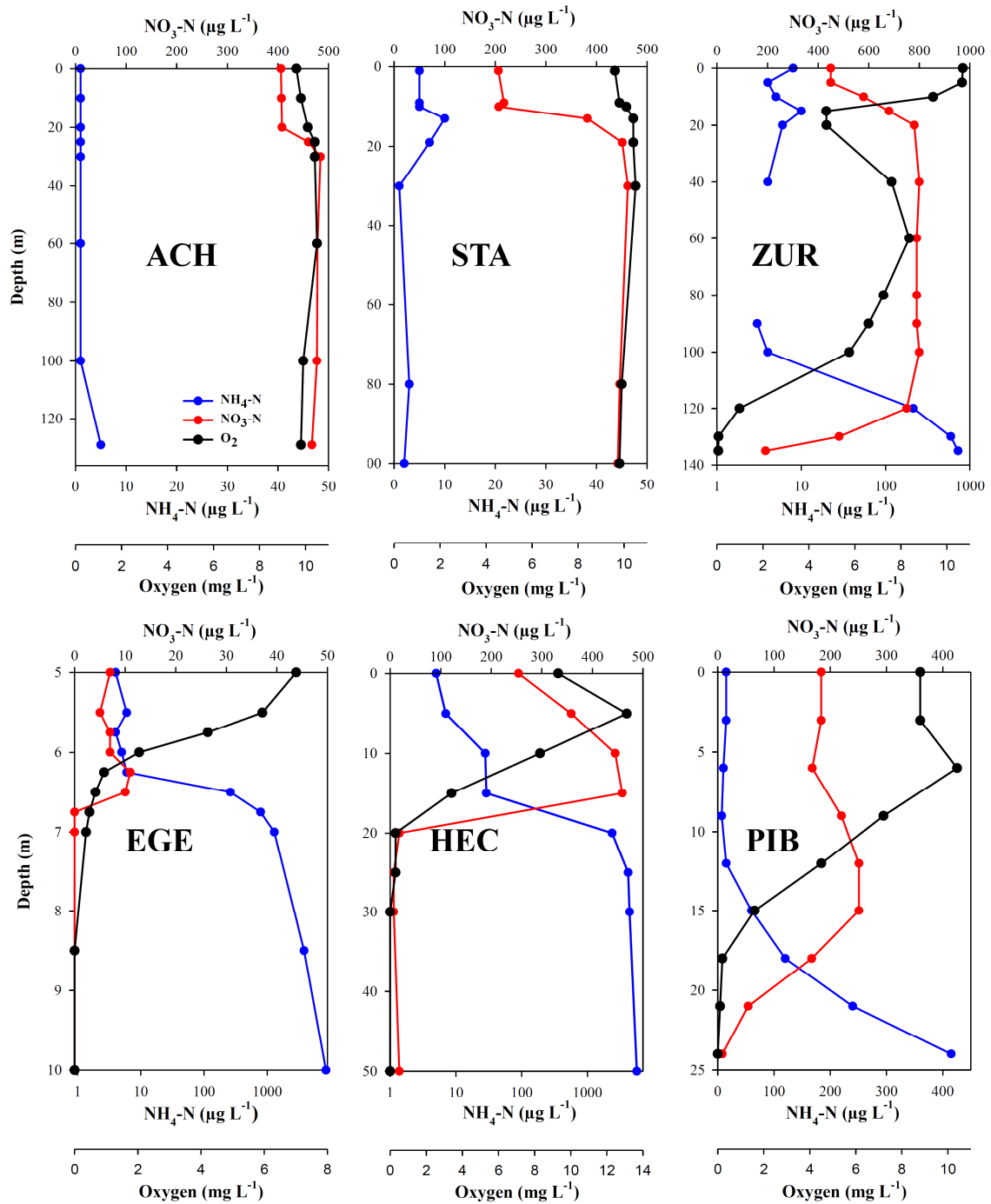
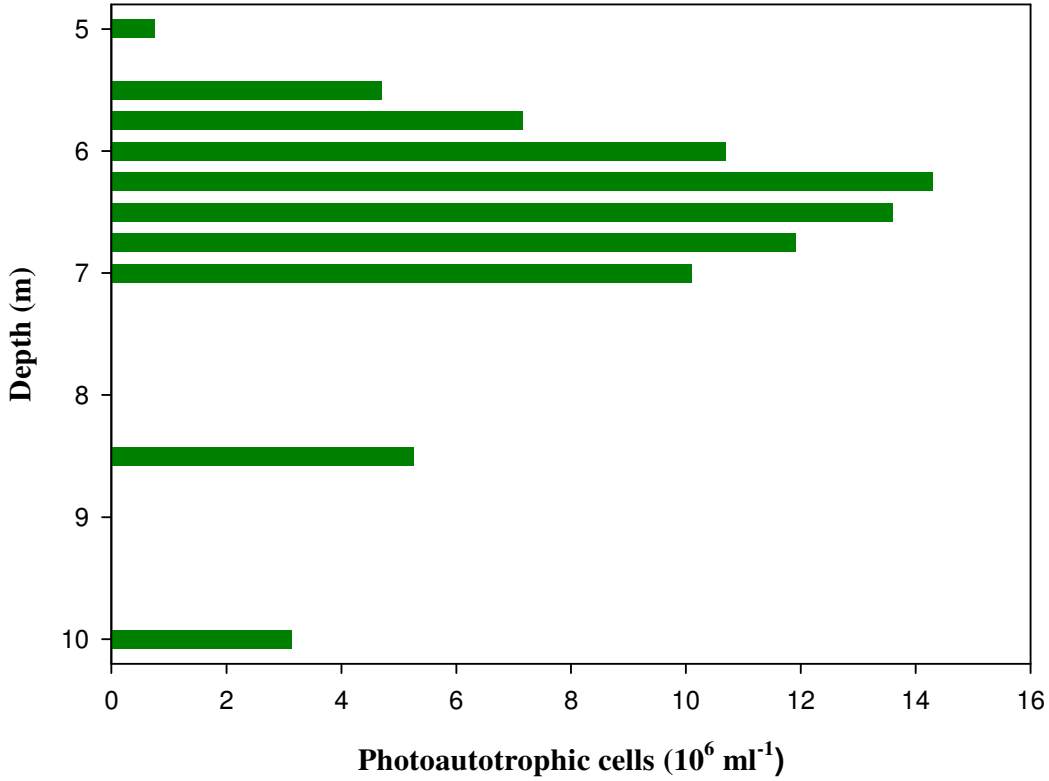


Fig. S7. Vertical profile of photoautotrophic cell numbers in Lake EGE



Supplementary References (not included in main text)

- Alfreider, A., Vogt, C., Hoffmann, D., and Babel, W. (2003) Diversity of ribulose-1,5-bisphosphate carboxylase/oxygenase large-subunit genes from groundwater and aquifer microorganisms. *Microb Ecol* 45: 317-328.
- Offre, P.R., Nicol, G.W., and Prosser, J.I. (2010) Autotrophic community profiling and quantification of putative autotrophic thaumarchaeal communities in environmental samples. *Environ Microbiol Reports* 3: 245-253.
- Kato, S., Nakawake, M., Ohkuma, M., and Yamagishi, A. (2012) Distribution and phylogenetic diversity of *cbbM* genes encoding RubisCO form II in a deep-sea hydrothermal field revealed by newly designed PCR primers. *Extremophiles* 16: 277-283.
- Selesi, D., Schmid, M., and Hartmann, A. (2005) Diversity of green-like and red-like ribulose-1,5-bisphosphate carboxylase/oxygenase large-subunit genes (*cbbL*) in differently managed agricultural soils. *Appl Environ Microbiol* 77: 175–184