Eco-AlpsWater

Innovative Ecological Assessment and Water Management Strategy for the Protection of Ecosystem Services in Alpine Lakes and Rivers

Priority 3: Liveable Alpine Space. SO3.2 - Enhance the protection, the conservation and the ecological connectivity of Alpine Space

Project Eco-AlpsWater
Work Package WPT3
Activity A.T3.3
Deliverable D.T3.3.1.
Version 1.0
Date Sep 2021
Coordination: R. Kurmayer
Compilation: R. Kurmayer
With contributions from Nico Salmaso, Ute Mischke, Tina Elersek, and the entire project consortium

Deliverable D.T3.3.1

Extending the approach to the whole Alpine Space area, involving Observers and stakeholders

Report on results obtained in key lakes and additional lakes

General introduction

In the EAW database we have environmental data for 275 lake samples, 124 from biofilm (from 12 lakes) and 151 from plankton (from 34 lakes). Mixing type of the majority of samples from lakes is monomictic (52%), dimictic (36%), meromictic (8%) and rarely polymictic (4%). Water renewal time is quite diverse; less than 5 years for 52% of lakes, 5-10 for 14%, 10-15 for 10% and more than 20 years for 25% of lakes. The half (51%) of lakes have catchment area in the range 101-1000 km^2, followed by 51-100 km^2 (21%) and 0,1-50 km^2 (16%). During the limnological sampling period, the temperature of water at sampling campaigns was ranging from 3 to 30°C, with conductivity from 16 to 588 µS/cm. The trophic status of key lakes has been assessed by three parameters: total phosphorus, transparency and chlorophyll-a concentration and analyzed with OECD fixed boundary trophic classification system (OECD; 1982). Thus trophic status of lakes sampled in EAW was ranging from ultra-oligotrophic to eutrophic conditions (Figure 1).

For plankton, the euphotic layer was sampled by depth integration. Euphotic layers were deeper than 20m for 20% of samples, ranging between 5-10 m for 29% of samples, or 10-15m for 19% and from 15-20 m for the rest of samples (11%). Sampling volume of water for eDNA analyses was 0,5-1L for half of samples (56%), less than 0,5L for one third of samples (34%) and more than 1L for the rest (11%).

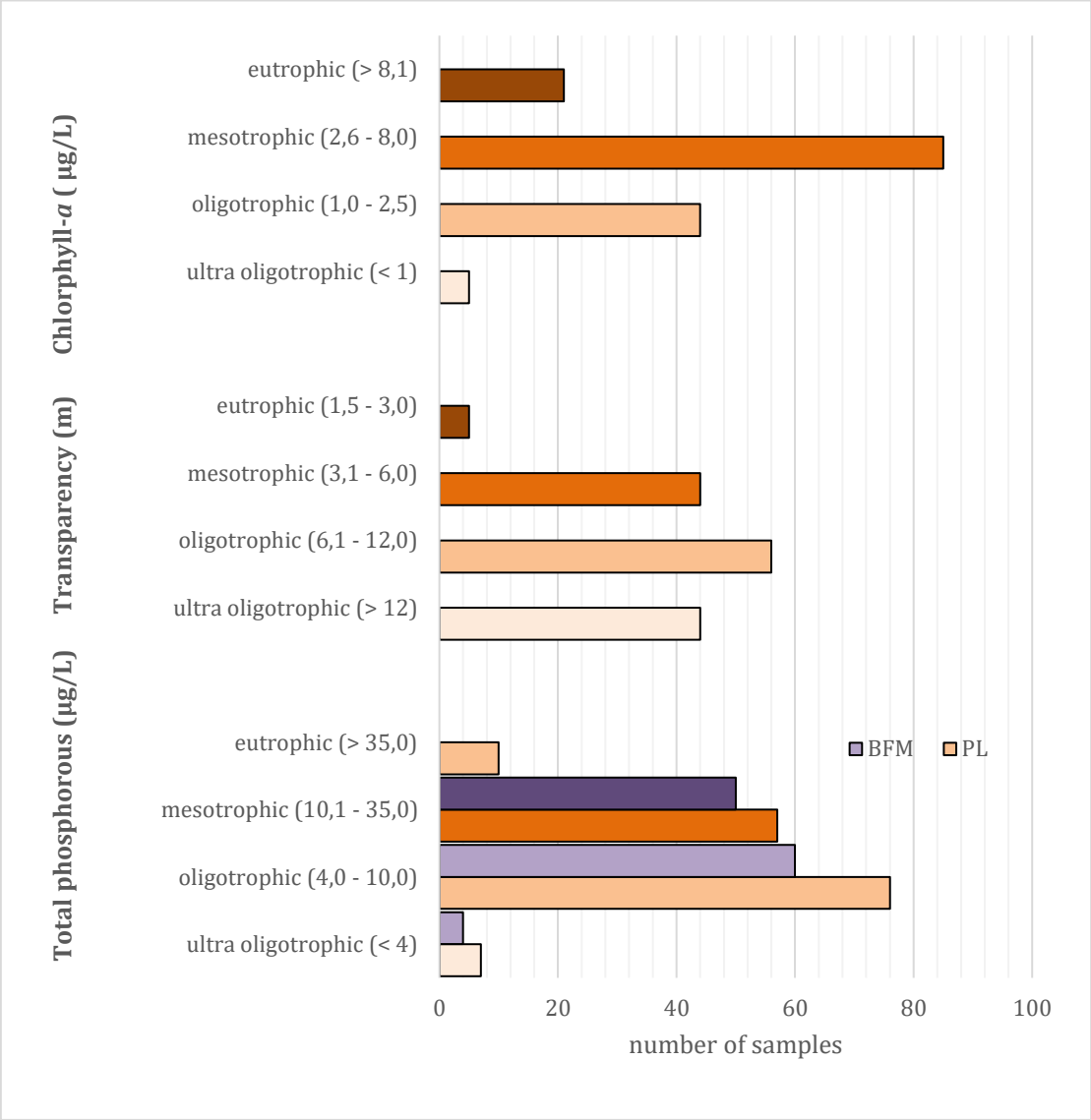


Fig. 1. Number of lake samples from biofilm (BFM) and plankton (PL) with different : (i) total phosphorous concentration, (ii) transparency and (iii) chlorophyll-*a* concentration, in terms of trophic classification system (OECD; 1982) ranging from ultra-oligotrophic to eutrophic conditions.

Part 1: Phytoplankton (incl. cyanobacteria)

Sampling according to national legislative

As for the key lake sampling phytoplankton samples were depth-integrated from 0-20 m corresponding to the euphotic zone at the deepest part of the lake. Sample aliquots were used to determine the chlorophyll-a concentration as well as chemical parameters and nutrients following the national legislative.

The abundance and the total biovolume of the planktonic algae were determined from a subsample under the inverted microscope (quantitative analysis).

The total phytoplankton biovolume was calculated from the sum of the individual taxa. In parallel water chemistry was determined according to the national legislative. Another water volume was filtered for cyanotoxin extraction according to the protocol “Cyanotoxins analyses in lake and biofilm samples” and Cerasino et al. 2017. Briefly among phytoplankton samples microcystins (n=47) and anatoxins (n=18) were found frequently. Overall oncentrations ranged from a few ng/L to 600 ng/L. For microcystin max. concentrations were observed in L. Garda (580 ng/L) and for anatoxin the max. concentration was observed in L. Ledro (628 ng/L). The cyanotoxins nodularin, cylindrospermopsin, and saxitoxins were not detected.

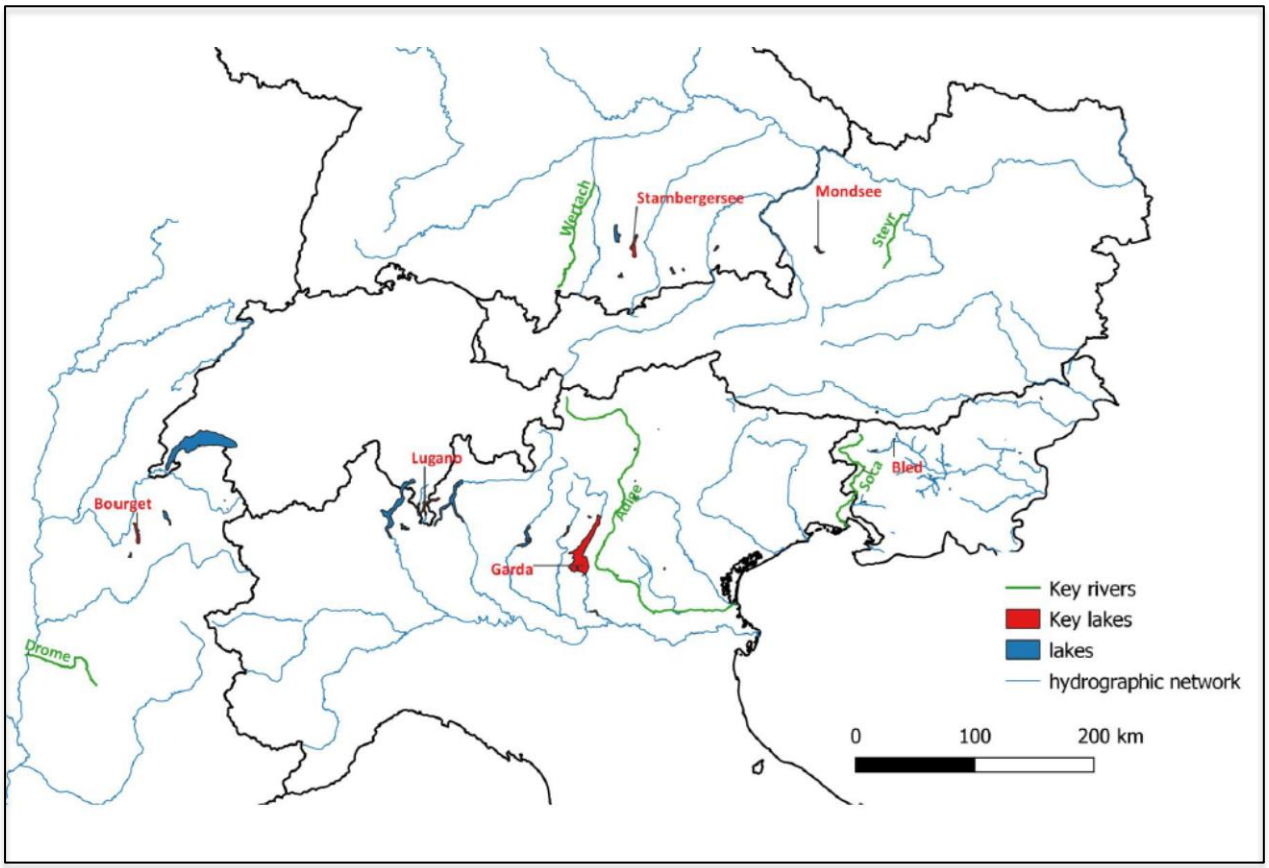


Fig. 2 Study area and sample origin (34 lakes, 6 key lakes and 28 additional lakes)

Table 1. List of 144 plankton samples originating from lakes and of 52 river samples from the six PP countries.

	Pilot lakes	Additional lakes	Pilot rivers	Additional rivers
Austria	14	1	3	3
Germany	9	36	13	1
France	13	5	6	4
Italy	12	14	2	4
Slovenia	12	16	3	8
Switzerland	12	0	0	6
Total	72	72	26	26

For DNA sequencing, the depth-integrated samples were taken transported to the laboratory using cooling boxes. In the laboratory the planktonic samples were filtered through a Sterivex™-GP 0.22 µm filter (Millipore, Billerica, Massachusetts, USA), by pressing water manually through the filter unit with a plastic syringe following the protocol from WP1 (D.T1.1.2 -1 Lake plankton sampling). The syringe was cleaned before with MQ-Water and rinsed once with the sample itself. To estimate the filtered volume, a 1 L Duran bottle was used to capture the filtered water. The filtering was completed until the filter became clogged or when a total volume of 0.5 L was reached.

DNA extraction and sequencing

DNA extraction from Sterivex filters was performed with Mo Bio PowerWater® DNA Isolation Kit (MO BIO Laboratories, a QIAGEN Company, USA) or with modifications according to Vautier et al. (2021). In particular DNA was extracted using the DNeasy® PowerWater® SterivexTM Kit (Qiagen, Hilden, Germany) following the manufacturer`s protocol.

From sample DNA extracts 16S rDNA (V3-V4 region) has been amplified using primers 341Fmod CCTAYGGGRBGCASCAG and 806Rmod GGACTACNNGGGTATCTAAT under the following conditions: 95°C (5 min), 28 cycles including 95°C (30 sec), 55°C (30 sec), 72°C (30 sec), and final 72°C (5 min). For 18S rDNA (V4 region) the primers V4F-18S_ILL and CCAGCASCYGCGGTAATTCC and V4R-18S_ILL ACTTTCGTTCTTGATYRATGA were applied using the cycling conditions from above. Library preparation of purified PCR products for 16S rDNA and 18S rDNA was performed according to EAW Protocols. Bridge amplification and sequencing by synthesis were performed according to standard conditions (FEM, Miseq, Massimo Pindo). One technical replicate was sequenced (DT1.1.2. -10 Library prep 16S marker gene; DT1.1.2. -11, Library prep 18S marker gene).

Bioinformatic processing

Forward and reverse primers in the 16S rRNA and 18S rRNA FASTQ reads were preliminarily removed by using Cutadapt 3.1 (Martin, 2011), implemented in the wrapper rmprim in github (github.com/hts-tools/metatools). Sequences were therefore analysed using DADA2 1.18 (Callahan et al., 2016) in R 4.0.3 (R Core Team, 2020). The DADA2 error model resolves amplicon sequence variants (ASVs, also known as exact sequence variants, ESVs) that differ by as little as one nucleotide (Callahan et al., 2017). The number of ASVs obtained was 52455 (16S rRNA genes) and 21371 (18S rRNA genes). Sequences were assigned to the SILVA 138 (16S rRNA genes) and PR2 v. 4.12 (18SrRNA genes) reference databases for taxonomic classification. The two bioinformatic pipelines used to obtain the bacterial and protists ASVs have been described in detail in Salmaso et al. (2021a, 2021b).

The final ASV tables used in this report were obtained after rarefaction without replacement at the minimum number of sequences per sample (i.e. 6867 and 5552), providing 37530 and 11468 ASVs for the 16S rRNA and 18S rRNA genes, respectively.

For cyanobacteria the automated taxa assignment was improved by using reference sequences from relevant taxonomic literature, using (morphologically described) isolates (strains). For BLAST hits 0-1 mismatch (405 bp alignment) were allowed, which resulted in 51 additional cyanobacterial species or strains. In particular the following taxa defined from strains were used as reference:

+) *Woronichiniana* (Ranjaniemi-Wacklin et al. 2005); *Limnococcus (Chroococcus) limneticus (*Komarkova et al. 2010); *Chroococus minutus* (Komarkova et al. 2010); *Cyanobacterium gracile* PCC6307; *Cyanodictyon* JJCD strain (Jezberova PhD thesis); *Dolichospermum (Anabaena) lemmermanii* (Wacklin et al. 2009, Lyra et al. 2001); +) *Dolichospermum crassum* (Ranjaniemi et al. 2005); *Dolichospermum planctonicum* (Ranjaniemi et al. 2005); *Dolichospermum spiroides*, *Dolichosperum (Anabaena) flos-aquae*; *Aphanocapsa* delicatissima, *A. planktonica*, *A. incerta*; (Komarek et al. 2011), *Eucapsis minor* (Komarek et al. 2016*)*; *Geitlerinema splendidum* (Strunecky et al. 2017); *Gomphosphaeria aponina*; *Komvophoron*; *Limnothrix planktonica*; *Merismopedia tenuissima* (Shen et al. 2018); *Merismopedia punctate*; *Microcystis aeruginosa*; *Phormidium amoenum* (Gaget et al. 2017)

Comparison with traditional microscopy

The microscopical taxa lists have been standardized using the established WFD (EU project WISER) taxa codes, i.e. the REBECCA code for phytoplankton. To facilitate comparison an Excel Access database tool (version 6, May 2021) for all microscopical taxa and REBECCA codes assigned has been prepared (LfU, FEM, LFUI).

Results

HTS Database overview

The 16S rDNA database (extracted on 22 Dec 2020) contained in total 37530 Seq (ASVs). Out of these 2577 Sequences (Seq) or Amplicon Sequence Variants (ASVs) were assigned to cyanobacteria: 1156 Seq were assigned on genus level, 53 Seq were assigned on species level. The 16S rDNA database also contained information from eukaryotic algae through the intracellular chloroplasts phylogenetically originating from cyanobacteria (1523 Seq). Within chloroplasts 503 Seq were assigned to the group of Bacillariophyta (diatoms), 210 Seq to Chlorophyta (coccale and monoid green algae), 54 Seq to Streptophyta (Zygnemales and Desmidiaceae), 89 seq to Cryptomonadavceae, and 11 Seq to Euglenids.

From the 18S rDNA database (extracted on 22 Dec 2020) in total 11468 Seq (ASVs) were obtained. Out of this 5283 Seq were assigned to eukaryotic microalgae (2355 Seq were assigned on genus level, while 1437 Seq were assigned on species level).

Cyanobacteria

Cyanobacterial taxa found by 16S rDNA taxonomy (rarified to signal strength = 4125) included 32 taxa matching with REBECCA code, (raref. Signal strength (read numbers) 1 – 4125), i.e. *Dolichospermum (Anabaena) crassum*, *D. kisseleviana*, *D. lemmermannii*, D. *planktonicum*, *Anabaena sp*., *D. spiroides*, *Aphanizomenon aphanizomenoides*, *A. flos-aquae*, *A. issatschenkoi*, *Aphanizomenon* sp., *Chroococcus limneticus*, *Cuspidothrix* .sp., *Cyanobium* sp., *Cylindrospermopsis raciborskii*, *Geitlerinema splendidum*, *Gloeocapsa* sp., *Limnothrix planctonica*, *Merismopedia* sp., *Microcystis* sp., *Nodosilinea*, N*odularia* sp., *Phormidium* sp., *Planktothrix* sp., *Prochlorothrix* PCC-9006, *Pseudanabaena* sp., *Radiocystis geminata*, *Romeria* sp., *Snowella litoralis*, *Snowella* sp., *Synechococcus* sp., *Tychonema* sp., *Woronichinia* sp.

Cyanobacterial taxa found by 16S rDNA taxonomy (22) but without assigning to the REBECCA code, (raref. Signal strength (read numbers) 1 – 428) included unassigned Genera/Species and thus could be differentiated only at higher taxonomic levels (families, orders), i.e. Candidatus Obscuribacter,

Cyanobiaceae, Leptolyngbyaceae, Microcystaceae, Nodosilineaceae, Nostocaceae, Obscuribacteraceae, Phormidiaceae, Pseudanabaenaceae, Sphaerospermopsis, Caenarcaniphilales, Cyanobacteriales, Gastranaerophilales, Oxyphotobacteria Incertae Sedis, Vampirovibrionales, Obscuribacterales, Synechococcales, Leptolyngbyales, Cyanobacteriales, Phormidesmiales, Obscuribacterales, Pseudanabaenales.

In order to find out whether there is a relationship between signal strength (number of rarified reads) and taxonomic assignment via the REBECCA code, all Cyanobacteria Seqs were compared in read numbers and plotted in classes of read numbers (Suppl. Fig. 1). The vast majority of Seqs (ASVs) occurred with a frequency of 1-80 reads (> 90%). In general 1420 Seqs (80%) that were assigned to REBECCA occurred more frequently (1-80 reads). In contrast 273 Seqs (93%) which were not assigned to REBECCA occurred less often (1-30 reads). Thus, in general the more abundant Seqs (ASVs) also were taxonomically assigned to REBECCA, while about 10% of ASVs could only be assigned to levels above genera/species and thus were considered less characterized.

Via sequencing, 12 planktonic cyanobacteria genera were detected: *Dolichospermum (Anabaena), Aphanizomenon, Cyanobium, Microcystis, Nodosilinea, Planktothrix, Prochlorothrix, Pseudanabaena, Snowella, Synechococcus, Tychonema, Woronichiana*. The genus *Planktothrix* constituted 50% while picocyanobacteria (*Cyanobium, Synechococcus*) constituted 30% (Fig. 3A). Via counting, 15 cyanobacteria genera/species were recorded more frequently, and *Anabaena, Aphanizomenon, Microcystis, Planktothrix, Pseudanabaena, Snowella and Woronichiniana* were recorded by both methods. In contrast, the genera *Aphanocapsa, Aphanothece, Chroococcus, Cylindrospermopsis, Limnothrix*, and *Planktolyngbya* were not confirmed by sequencing. Vice versa, *Prochlorothrix* and *Nodosilinea, Tychonema, Synechococcus* and *Cyanobium* appeared only from sequencing (Fig. 3B).

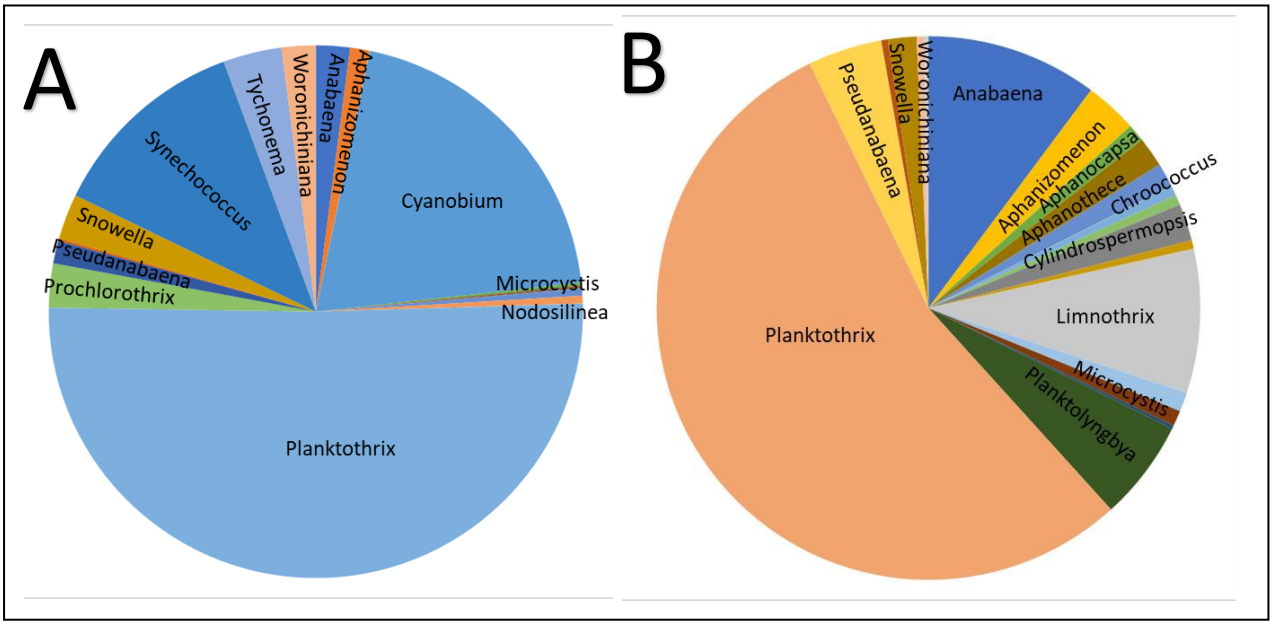


Fig. 3. Overview planktic Cyanobacteria composition (percentage), (A) Read numbers and (B) biovolume determined in Light microscope in phytoplankton from lakes across alpine space (n = 140).

In general for *Anabaena (Dolichospermum)* a good qualitative but little quantitative relationship with biovolume was observed. Overall the number of samples without detection or with detection corresponded although sequencing was found more sensitive, i.e. 33 samples positive via sequencing

vs 24 samples positive via microscopy (Suppl. Fig. 2A, B). From linear regression analysis two outliers (biovolume 12.62, 2.58 mm3/L) were excluded (Suppl. Fig. 2C).

For the genus *Aphanizomenon* a good qualitative and quantitative relationship between both methods was obtained. Overall the number of samples without detection or with detection matched although, as for *Anabaena*, sequencing was found more sensitive, i.e. 63 samples were found positive via sequencing vs 47 samples were found positive via microscopy (Suppl. Fig. 3 A,B). Linear regression analysis showed reasonable quantitative correspondence (R^2 = 0.58).

For pikocyanobacteria, in general, little qualitative and quantitative correspondence between sequencing and biovolume estimated via microscopy could be found. In particular a much higher number of samples (133) were found positive for *Cyanobium* or *Synechococcus* via sequencing than were found to contain picocyanobacteria (i.e. *Aphanothece* or *Aphanocapsa*) via the microscope (67), Suppl. Fig. 4A, B. In particular for the negative samples the correspondence between the two methods was rather low, implying that frequently picocyanobacteria have not been counted using the microscope. Not surprisingly, the quantitative relationship between both methods was poor (Suppl. Fig. 4C).

For the genus Microcystis a good qualitative vs little quantitative relationship between sequencing and biovolume was observed. Similarly, to *Anabaena*/*Aphanizomenon* the genus Microcystis was detected more frequently via sequencing (45) vs microscopy (34). Overall for negative samples, correspondence was reasonable and only eleven samples were found false negative (Suppl. Fig. 5A, B). Nevertheless, the quantitative relationship between sequencing vs biovolume was low (Suppl. Fig. 5C).

Notable, the filamentous cyanobacterium *Nodosilinea* sp. was detected in five lakes located in Italy and Slovenia (Caldaro; Frassino; Garda; Iseo; Pernica). Filaments of *Nodosilinea* sp. typically grow forming a single trichome, but sometimes become multiseriate, with nodules forming under low light conditions (Perkerson III et al. 2011). Although *Nodosilinea* sp occurred in 16 samples, the genus was not recorded in the microscope.

For the genus *Planktothrix* both qualitative and quantitative correspondence between sequencing vs. microscopy has been found. In accordance with other genera the number of positive samples was increased (103) when compared with microscopy (89). Thus sequencing, overall increased detectability of *Planktothrix* among alpine space samples. On the other hand, for negative samples, correspondence between both methods was good, and only eleven samples were found false negative (Suppl. Fig. 6A, B). The quantitative relationship between sequencing and biovolume was found high (R^2 = 0.64), Suppl. Fig. 6C.

Interestingly the closely related cyanobacterium *Tychonema*, which can hardly be discriminated from *Planktothrix* in the microscope showed a rather good qualitative and quantitative relationship when compared with sequencing. Using sequencing 17 samples were found positive (L. Como, Garda, Iseo; Ledro; Maggiore; Staffelsee_Nord) while using the microscope *Tychonema* was found in 11 samples. The quantitative relationship between sequencing and biovolume was found high (R^2 = 0.79), Suppl. Fig. 6C.

Similar to *Nodosilinea*, the filamentous genus cyanobacterium *Prochlorothrix* was never detected by microscopy but through sequencing in 4 lakes (Bled, Schliersee; Tegernsee). So far, the evidence for the presence of the genus *Prochlorothrix* has been restricted to (shallow) eutrophic freshwater or brackish habitats of North Europe (e.g. The Netherlands, Pinevich et al. 2012). In general, co-occurrence with phytoplankton included mostly blue–green filamentous cyanobacteria of the group *Limnothrix/Pseudanabaena* (e.g. Zwart et al., 2005). Since *Prochlorothrix* contains Chlorophyll a and Chlorophyll b it is characterized by a green pigmentation, but might have been overlooked during

microscopical counting. Notably, only in Schliersee higher read numbers were recorded (697-1364 reads).

For the genus *Pseudanabaena* an overall good qualitative and quantitative relationship was observed. However, sequencing overall increased detectability, i.e. 56 samples were found positive while only 14 samples were found to contain *Pseudanabaena* using microscopy (Suppl. Fig. 7A, B). The quantitative relationship between sequencing and biovolume was found high (R^2 = 0.75), Suppl. Fig. 7C.

For the morphologically characterized genera *Snowella* and *Woronichiniana* overall good qualitative and quantitative relationship was found. As for all genera detectability was found increased (Suppl. Fig. 8, 9), nevertheless a good quantitative relationship was observed.

In summary, overall (very) good qualitative relationship between sequencing results (HTS reads) and microscopy (biovolume) was obtained. Thus sequencing results can be used to confirm microscopical counting based on morphological characters, i.e. the differentiation of the genus *Tychonema* from *Planktothrix* based on subtle morphological characters. The mismatch in positive detection between microscopy biovolume vs sequencing was rather low (0-10%), i.e. very little "false positives" were obtained through microscopy. On the other hand, in general for all taxa the sequencing was found more sensitive than the biovolume estimate via microscopy (3 – 30 (50) % higher positive detection rate). Not surprisingly, in the microscope the biovolume of picocyanobacteria was underestimated and in general picocyanobacteria very likely were more reliably detected and quantified using HTS (reads). Nevertheless, two filamentous cyanobacteria (*Nodosilinea, Prochlorothrix*) were detected by sequencing only, implying potential refinement of microscopical analysis. Overall also good quantitative relationship between HTS (reads) and (LM) biovolume were observed for many filamentous or colony-forming genera, i.e. even quantitative conclusions seem possible.

Planktonic cyanobacteria diversity (genotype/strains)

It was interesting to see that populations of planktonic cyanobacteria were composed of variable numbers of partly co-occurring genotypes (oligotypes). The picocyanobacteria (*Cyanobium*) occurred with a maximum number of genotypes (29), while *Tychonema* occurred with one genotype only. The rare genus *Prochlorothrix* occurred with six genotypes (Table 2, Fig. 4). Even harmful and bloom-forming cyanobacteria such as Microcystis or *Planktothrix* characterized phylogenetically intensively using clonal isolates in the laboratory (co-)occurred with several genotypes not described from strain culture collections. Thus in contrast to microscopy, by sequencing a more deep characterization of population genetic structure was achieved.

Various planktonic morphospecies of the former genus Anabaena have been transferred to the new genus *Dolichospermum* (Wacklin et al. 2009). Since the genus *Anabaena* (*Dolichospermum*) has been observed to represent deep phylogenetic diversity (Ranjaniemi et al. 2005), different phylogenetic lineages were assigned with morphospecies well recognized from phytoplankton samples worldwide, i.e. *D. crassum*, *D. circinale*, *D. spiroides*, *D. lemmermannii*, *Sphaerospermopsis*. In the majority of lake samples, the genotype Seq 428 assigned to *D. lemmermannii* was occurring as the only genotype and/or dominating (Fig. 5). Only for a few other smaller and more eutrophic lakes (Serraia, Simsee, Slivnica, Pernica), other *Dolichospermum* taxa were detected: *D. crassum* (Seq 106) showed highest abundance in Serraia and rare occurrence in Simsee and Slivnica. For Pernica a co-occurrence of *D. circinale*, D. *spiroides*, and *Spaerospermopsis* was observed. Overall the absolute dominance of *D. lemmermanii* in alpine lakes is in accordance with its reported prevalence in oligotrophic boreal lakes (i.e. Lepistö et al. 2005). In addition *D. lemmermanni* has been associated with various cyanotoxins (microcystins, anatoxin), and thus is considered relevant for water risk assessment.

Table 2. Number of genotypes (oligotypes) recorded at 16S rDNA for planktonic cyanobacteria from 34 lakes across the alpine space area (n = 140).

	min	median	max
Dolichospermum (Anabaena)	1	1	3
Aphanizomenon	1	1	3
Cyanobium	1	4	29
Microcystis	1	1	3
Nodosilinea	1	1	4
Planktothrix	1	2	8
Prochlorothrix	1	4.5	6
Pseudanabaena	1	1	4
Snowella	1	1	3
Synechococcus	1	2.5	5
Tychonema	1	1	1
Woronichinia	1	1	5

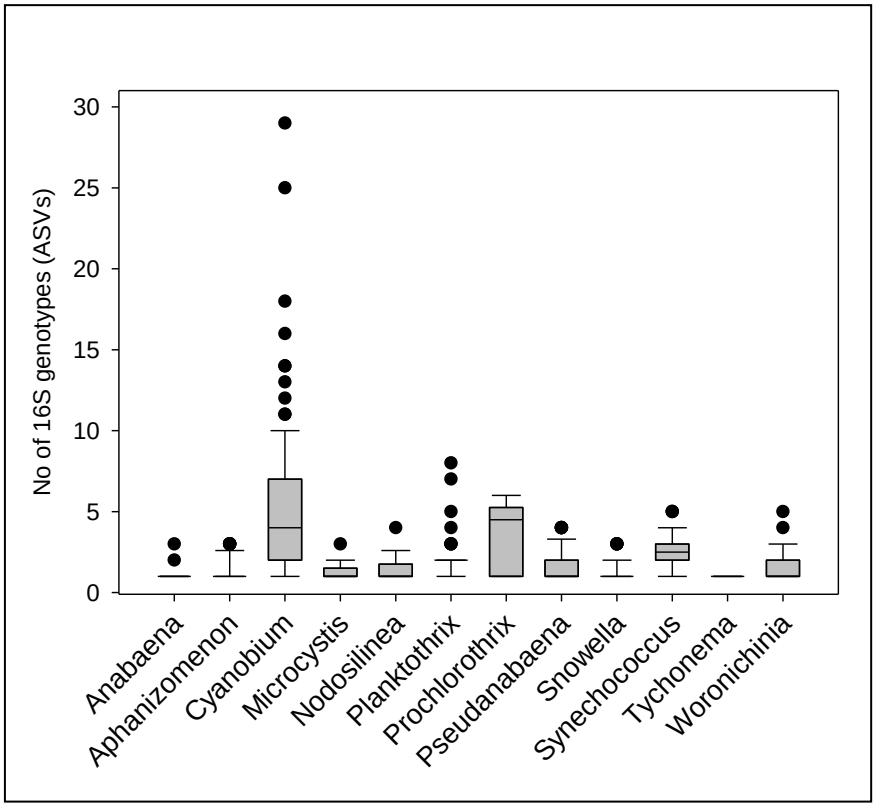


Fig. 4. Frequency distribution of 16S rDNA genotypes occurring in planktonic cyanobacteria populations as recorded from 34 lakes across the alpine space area (n = 140). Box plots show 25% - median – 75% quantiles.

When compared with *Dolichospermum*, for *Aphanizomenon* a higher number (17) of Seq (ASVs) was reported. Seq 189 indicative of *A. flos-aquae* has been found most frequently and also most dominant (Suppl. Fig. 11). Highest read numbers were recorded from L. Lugano (May - Aug 2019) and L. Como and L. Idro. Other ASVs Seq15960, Seq907, Seq9366Seq15960 also indicated A. flos-aquae. Since *A.*

flos-aquae has been linked to saxitoxin and its structural variants it is considered relevant for risk assessment. Other taxa included *Sphaerospermopsis aphanizomenoides* and *Cuspidothrix issatschenkoi* (occurring in L. Pernica).

Notably, intensively studied bloom-forming cyanobacteria such as Microcystis or *Planktothrix* characterized phylogenetically via clonal isolates occurred with genotype diversity so far undescribed. For example *Microcystis* occurred with 5 16S rDNA genotypes, of which either Seq1138 or Seq1196 or Seq1881 were dominating (Suppl. Fig. 12). For the genus *Planktothrix* 2 ASVs (Seq2, Seq7) co-occurred frequently while 3 other genotypes rarely were found. It should be noted that from strain collections typically any genetic diversity within 16S rDNA gene regions is found. In contrast by means of sequencing, new genotypes that have been missed because they are not accessible to isolation under artificial culture conditions might have been found.

In contrast to filamentous cyanobacteria the genera *Cyanobium* and *Synechococcus* are assigned to the order of Synechococcales, constituting a phylogenetic lineage with mostly unicellular morphology of smallest cells (often referred to as pikocyanobacteria) and thriving under nutrient-poor conditions in the water column (Sanchez-Baracaldo et al. 2019). It was interesting to see that the genus *Cyanobium* (differing in its cellular shape from *Synechococcus*) occurred with 29 genotypes while *Synechococcus* occurred with 5 genotypes only. In L. Garda and L. Lugano > 20 *Cyanobium* genotypes occurred also showing seasonal patterns, i.e. with dominance of winter vs. summer genotypes.

In summary, it can be stated that

+) 16S rDNA genotype composition informs about dominant taxa as well as prevailing ecological conditions (e.g. dominance of *D. lemmermannii* within the genus *Dolichospermum* (planktic Anabaena) sp.

+) vice versa 16S rDNA genotype composition can be used for surveillance and control of possibly invasive cyanobacterial taxa (e.g. members of Nostocales in consequence to climate warming), i.e. *Chrysosporum ovalisporum*

+) 16S rDNA genotype composition can inform about ecological forces favoring the growth of numerous 16S genotypes (*Cyanobium* sp.) and thus influencing the development of the total population

+) vice versa a stabilized 16S genotype composition may indicate rather stable (optimized) ecological adaptations (*Synechococcus* sp., *Planktothrix* sp.) and can also be used to indicate relevant environmental change

+) 16S rDNA genotype composition can inform about the true phylogenetic diversity which might be overlooked during laboratory-based isolation and cultivation efforts.

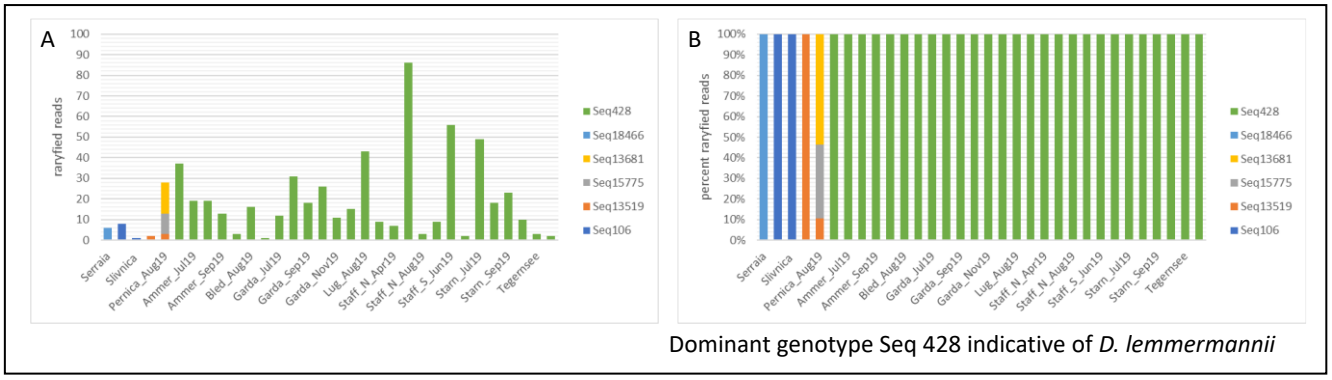


Fig. 5. (A) Absolute and (B) relative biodiversity of the genus *Dolichospermum* (formerly planktic morphospecies of Anabaena transferred to the new genus *Dolichospermum*) from lakes across alpine space (n = 140): Seq 106= *D. crassum*, Seq 428, *D. lemmermannii*, Seq 13519, *D. circinale* strain NIES41, Seq 13681, *Sphaerospermopsis* strain BCCUSP55, Seq 15775, *D. spiroides*, Seq 18466, *D. planctonicum* (1 outlier from Serraia (2541 reads) excluded, Seq106).

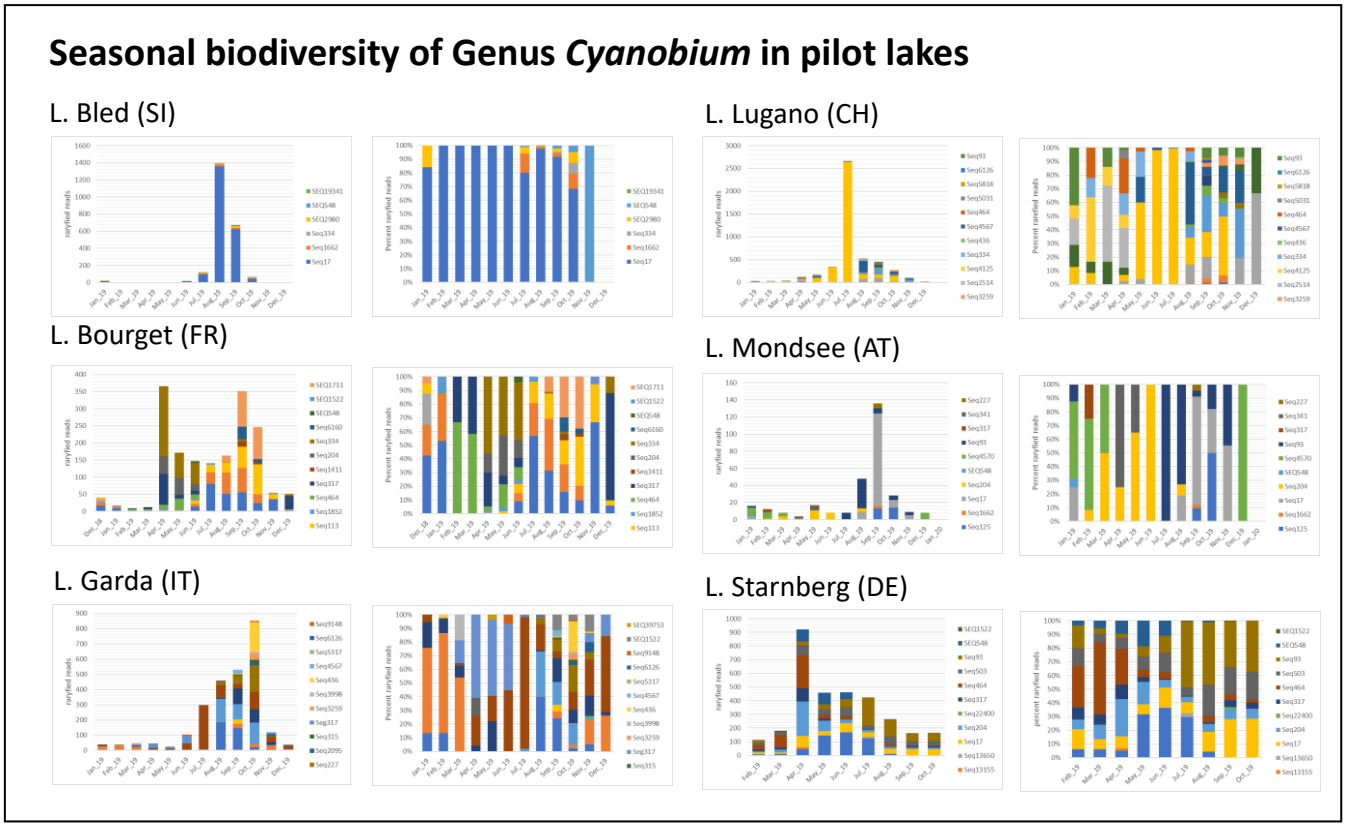


Fig. 6. A) Absolute and (B) relative biodiversity of the genus *Cyanobium* from 6 key lakes across alpine space (n = 140).

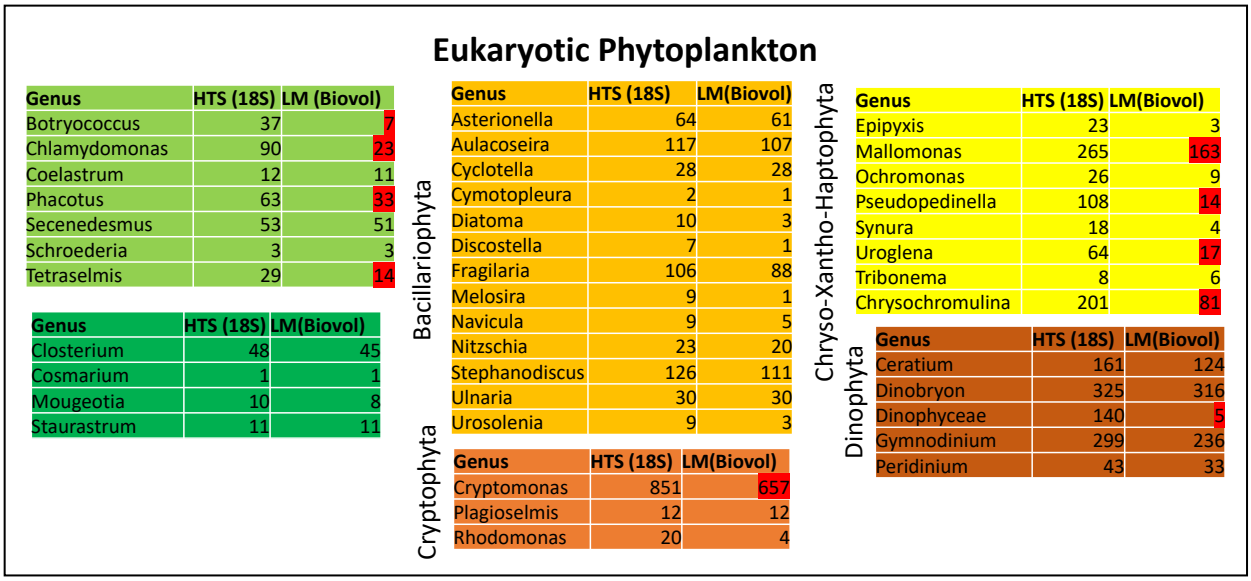
Eukaryotic phytoplankton

Using the database access tool (WP4) the frequency of detection on genus level was compared for all groups of eukaryotic phytoplankton: Chlorophyta (green algae), Streptophyta (including Zygnemales

and desmids), Bacillariophyta (diatoms), Cryptophyta (cryptomonads), Chrysophyta (chrysophytes as well as xanthophytes and haptophytes) and Dinophyta (dinoflagellates). For many genera a good qualitative correspondence on genus level between both methods was observed, including flagellates, coccale, filamentous as well as colony forming organization types.

Using 18S rDNA sequencing, for many genera of chlorophytes and Zygnemales, diatoms, Chrysophytes and dinophytes a good correspondence in detection frequency using both methods was observed. In general these genera are well differentiated in the microscope.

On the other hand a higher sensitivity for detecting flagellates irrespective of taxonomic affiliation was observed, not at least because of increased sampling volume (500 ml of filtered volume vs 25 ml in sedimentation chamber). It is known that flagellates are difficult to preserve using the standard Lugol fixative and thus typically underestimated using microscopy. In addition flagellates show little differentiation in morphology accessible through standard light microscopy and typically are summarized during the counting procedure. Flagellates detected less frequently by LM included the genera *Uroglena*, *Chlamydomonas*, *Pseudopedinella*, *Mallomonas*, *Chrysochromulina*, *Gymnodinium* and *Cryptomonas* (Fig. 8), which were frequently detected by LM counting but the frequency of detection was significantly increased during sequencing.



Genus	HTS (18S)	LM (Biovol)
Botryococcus	37	7
Chlamydomonas	90	23
Coelastrum	12	11
Phacotus	63	33
Secenedesmus	53	51
Schroederia	3	3
Tetraselmis	29	14

Genus	HTS (18S)	LM(Biovol)
Closterium	48	45
Cosmarium	1	1
Mougeotia	10	8
Staurastrum	11	11

Genus	HTS (18S)	LM(Biovol)
Asterionella	64	61
Aulacoseira	117	107
Cyclotella	28	28
Cymotopleura	2	1
Diatoma	10	3
Discostella	7	1
Fragilaria	106	88
Melosira	9	1
Navicula	9	5
Nitzschia	23	20
Stephanodiscus	126	111
Ulnaria	30	30
Urosolenia	9	3

Genus	HTS (18S)	LM(Biovol)
Cryptomonas	851	657
Plagioselmis	12	12
Rhodomonas	20	4

Genus	HTS (18S)	LM(Biovol)
Epipyxis	23	3
Mallomonas	265	163
Ochromonas	26	9
Pseudopedinella	108	14
Synura	18	4
Uroglena	64	17
Tribonema	8	6
Chrysochromulina	201	81

Genus	HTS (18S)	LM(Biovol)
Ceratium	161	124
Dinobryon	325	316
Dinophyceae	140	5
Gymnodinium	299	236
Peridinium	43	33

Fig. 7: Comparing detection frequency using both methods (18S rDNA sequencing & microscopical counting) for eukaryotic phytoplankton. Taxa have been assigned using REBECCA codes on genus level but typically include several species as identified under the microscope. The red marked numbers indicate probably underestimation via LM.

In addition, through sequencing information on certain groups of algae which have not been well recorded before, i.e. eukaryotic flagellates (Chrysophyceae, Dinophyta, Volvocales) but also entire new algal groups (Eustigmatophyta) has been obtained (Fig. 9). The majority of these taxa not detected via LM represent flagellate algae which are not readily detected for reasons as described above. On the other hand many of the new taxa were rare and probably could not have been detected using the lower sampling volume through sedimentation chambers.

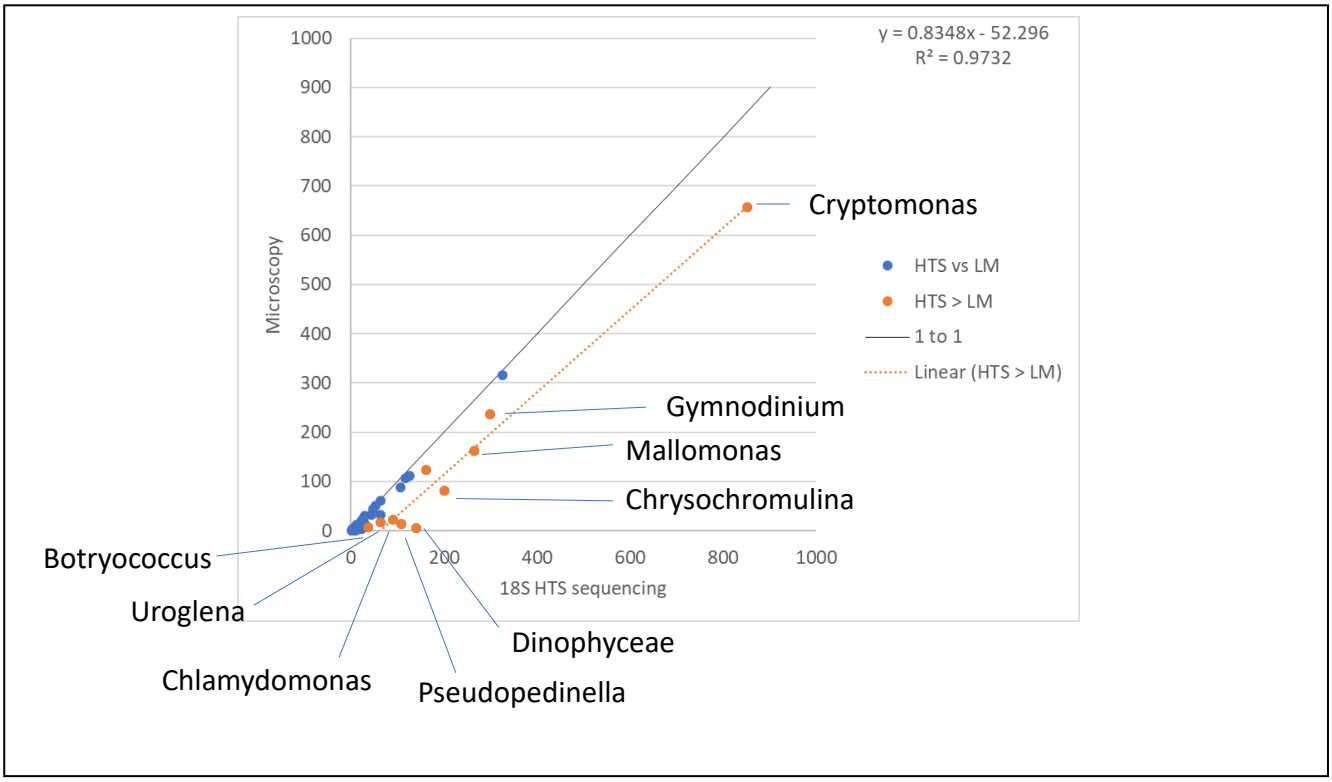


Fig. 8: Comparison of frequency of detection of eukaryotic phytoplankton genera using both methods 18S rDNA sequencing vs. microscopical counting. The straight line indicates the 1:1 relationship. The orange symbols indicate higher frequencies of detection via HTS when compared with microscopical counting. In general phytoplankton of the monadoid organization type were more frequently detected via sequencing.

Eukaryotic Phytoplankton

Chlorophyta

Genus	HTS (18S)
Volvocales sp.	135
Crustomastigaceae	80
Trebouxiophyceae	80
Chlorophyceae sp.	75
Chlorococcales sp.	74
Dolichomastigaceae	59
Choricystis sp.	54
Mychonastes sp.	25
Sphaeropleaceae Gen. sp.	16
Monomastix sp.	14
Hafniomonas reticulata	11
Tetraselmis cordiformis	10
Nephroselmis olivacea	4
Parachlorella kessleri	4
Asterococcus sp.	3
Golenkinia longispicula	3
Tetracystis sp.	3
Chlorococcum sp.	2
Chloromonas sp.	2
Coelastrum reticulatum	2
Ankyra judayi	1
Carteria sp.	1
Chaetophorales	1
Wislouchiella planctonica	1
Zygnematales sp.	26

Bacillariohyta

Genus	HTS (18S)
Bacillariophyceae sp.	140
Synedra sp.	62
Melosira varians	8
Thalassiosira sp	5
Acanthoceras sp.	3
Surirella sp.	3
Cymbella affinis	2
Pseudostaurosiropsis sp.	2
Gyrosigma limosum	1

Cryptophyta

Genus	HTS (18S)
Chroomonas sp	11
Komma caudata	6

Eustigmatophyceae

Genus	HTS (18S)
Nannochloropsis sp.	14
Eustigmatophyceae	13
Pseudotetraëdriella kamillae	2

Chryso-Xantho-Haptophyta

Genus	HTS (18S)
Chrysophyceae sp.	149
Pedinella hexacostata	139
Paraphysomonas sp.	61
Chrysamoeba sp.	27
Monas sp.	20
Prymnesiaceae Gen. sp.	20
Ochromonas sp.	17
Poteriospumella_lacustris	7
Chrysocapsa planctonica	5
Spumella sp.	5
Hydrurus foetidus	4
Chrysosphaerella sp.	2
Poterioochromonas_malhamensis	2
Chromulina nebulosa	1

Dinophyta

Genus	HTS (18S)
Prorocentrum sp.	123
Thoracosphaeraceae	123
Asulcocephalium miricentonis	77
Gymnodinium helveticum	48
Gymnodiniaceae	15
Suessiaceae	6
Tovellia aveirensis	6
Peridinales Gen. sp.	3
Polarella glacialis	3
Baldinia sp.	1
Woloszynskia tenuissima	1

Fig. 9: Eukaryotic phytoplankton taxa detected via 18S rDNA sequencing only. Taxa have been assigned using REBECCA codes on genus level but typically include several species.

In contrast to 16S rDNA quantitative estimates using both methods (18S rDNA sequencing vs microscopical counting) were hardly comparable, probably because gene copies of 18S rDNA in the nucleus have been found highly variable among the different algal groups. Thus in contrast to 16S rDNA the size of a eukaryotic algae cell does not necessarily correlate with 18S sequencing read number, as 18S copies vary several order of magnitude among algal groups. Nevertheless for a limited number of eukaryotic genera quantitative relationships were observed, i.e. *Asterionella* (R = 0.54, n = 61), *Stephanodiscus* (R = 0.5, n = 110), *Phacotus lenticularis* (R = 0.9, n = 33), and *Chrysochromulina* (R = 0.36, n = 81), Suppl. Fig. 13.

Finally, several algal groups reliably detected among lake water samples using microscopical counting were found to occur with high biodiversity in lakes, i.e. chrysophytes, cryptomonads and dinoflagellates. From a theoretical point of view, each genotype (amplicon sequence variant or ASV) might be considered a "cryptic" species. For example *Cryptomonas* occurred with 19 "cryptic" genotypes, the chrysophyte genus *Dinobryon* with 14, or the dinophyte genus *Ceratium* with 12 rDNA (18S) genotypes (Suppl. Fig. 14). It should be noted that the number of ASVs was related to the maximum read number indicating the general abundance (R = 0.58, n = 40). In other words phytoplankton taxa observed with a high read number also occurred with a higher intraspecific diversity.

Furthermore considerable information on interspecific variation among algal groups not recognized by microscopy was obtained (i.e. Chrysophyceae, Volvocales, Bacillariophyceae). Accordingly the number of ASVs was related to the maximum read number indicating their general abundance (R=0.7, n=64). Most importantly the algal groups of Volvocales, Chrysophyceae, Prymnesiaceae, Bacillariophyceae, Gymnodiniacea occurred with 19, 76, 28, 19, 10 genotypes, respectively (Suppl. Fig. 15). In general the algal group of chyrosphytes showed the highest (hidden) biodiversity (76 genotypes).

In summary for eukaryotic phytoplankton we observed the following:

+) overall good qualitative relationship between HTS derived genera and microscopy derived genera, i.e. sequence based confirmation of microscopy and results on genus level (confirmation of microscopical results)

+) overall higher sensitivity for detecting eukaryotic microalgae (flagellates), because of increased sampling volume (500 ml of filtered volume vs 25 ml sedimentation chamber)

+) additional information on certain groups of algae which have not been well recorded before, i.e. eukaryotic flagellates (Chrysophyceae, Dinophyta, Volvocales) but also entire new groups (Eustigmatophyta)

+) information on interspecific genetic variation among populations, i.e. of genotype variability within populations of recognized algal taxa (Chrysophyceae, Cryptomonads, Dinophyceae)

+) information on interspecific variation among algal taxa not recognized by microscopy (Chrysophyceae, Volvocales, Bacillariophyceae).

Part 2: Phytobenthos (incl. diatoms)

Report on results obtained in key rivers and additional rivers

For phytobenthos in total 23 rivers (Fig. 10) were assessed resulting in 53 samples. Two thirds (66%) of river samples had a catchment area > 1000 km2, 23% had catchment area 101-1000 km2, and 9% had a catchment < 50 km^2. During sampling water temperature ranged from 4 to 22°C (10-20°C) and conductivity varied between 25 and 1033 µS/cm (55% of samples had 200-375 µS/cm). To infer overall trophic conditions three nutrient parameters (total phosphorus, phosphate and nitrate concentration) were used. In general river samples were assigned a trophic state ranging from oligotrophic to mesotrophic conditions (Fig. 11), according to the EU commission report (2018).

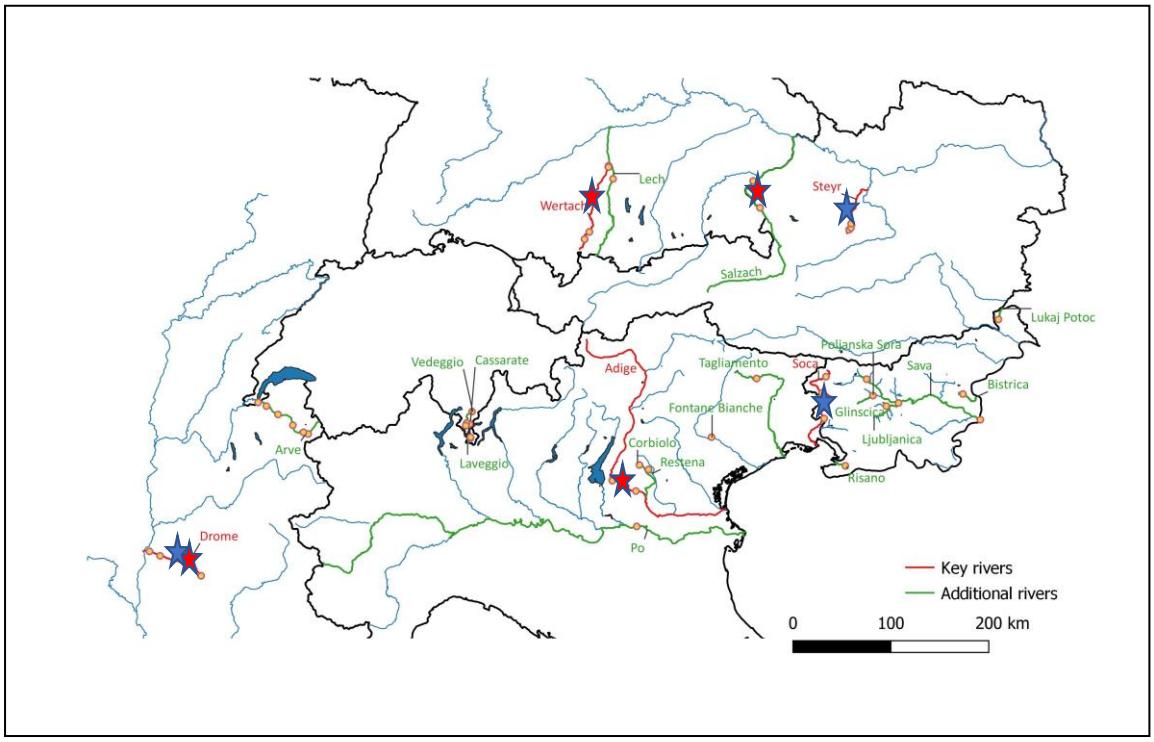


Fig. 10. Study area and sample origin (23 rivers, 5 key rivers and 18 additional rivers), Red star: indicates sites of hydromorphological restoration/ management during the HyMoCARES project (https://www.alpine-space.org/projects/hymocares/en/home), Blue star: indicates sites of strategic Planning for Alpine River Ecosystems during the SPARE project (https://www.alpine-space.org/projects/spare/en/home)

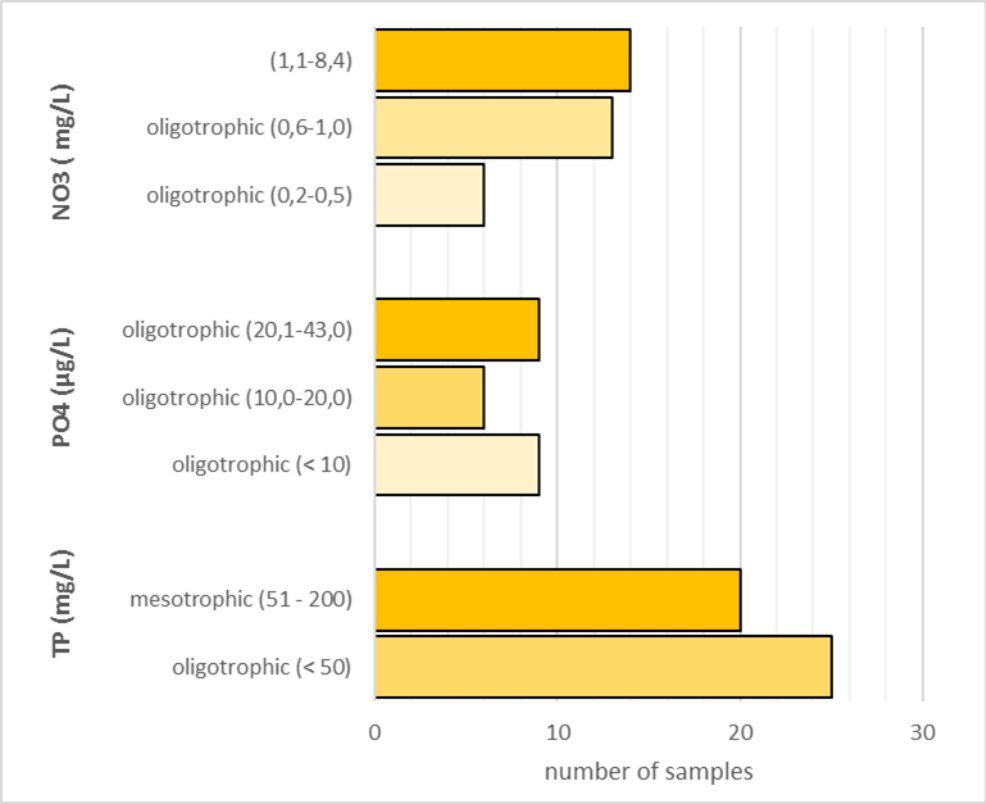


Fig. 11. Characterisation of key river samples from biofilm (BFM) according to (i) TP - total phosphorous concentration exhibit oligotrophic and mesotrophic state, and (ii) PO4 - phosphate and (iii) NO3 - nitrate concentration (trophic classification according to the EU commission working document 2018)

Sampling according to national legislative

In general sampling was performed according to the national legislative for phytobenthic sampling. In brief, five stones were selected by wading into the water and algae growing at the stone surface were brushed off into a tray. Both soft algae and diatoms were sampled using formaldehyde fixation.

In parallel to sampling for microscopy, for DNA extraction from the same stones aliquots were preserved using 80% Ethanol as described in protocol (DT1.1.2. -2, Lake biofilms sampling protocol).

Finally aliquots were scratched directly onto pre-weighed GF/C Filters and the dry-weight was determined from the difference in dried filter (105°C, 24 h) weight before and after filtration. Aliquots without drying but stored at -20°C were then used for cyanotoxin extraction (protocol: Cyanotoxins analyses in lake and biofilm samples and Cerasino et al. 2017.) In general Cyanotoxins were found only in 2 rivers out of 17 rivers sampled (River samples from Germany and France were not analysed for cyanotoxins). Sites from River Salzach (Austria) showed anatoxin (50 ng/g DW) and sites from river Corbiolo (Italy) showed the occurrence of microcystins (300 ng/g DW).

DNA extraction

DNA extraction from biofilm was performed using the NucleoSpin® Soil kit (Macherey-Nagel) according to Vautier et al. (2020).

Sequencing and bioinformatic processing

For the 16 S and 18S rRNA genes, see part 1. For rbcL gene assignment to diatom taxa the curated database R-Syst::diatom (Rimet et al. 2016) was used (INRAE).

Comparison with traditional microscopy

All microscopical taxa lists have been standardized using the established WFD (EU project WISER) taxa codes, i.e. the VALID code system for diatoms in phytobenthos (LfU) and the REBECCA code for non-diatoms (soft algae) An Excel Access database for all microscopical taxa and the VALID codes assigned has been prepared (LfU, FEM, LFUI).

Results on comparison between traditional microscopy and HTS

Using 16S rDNA the biodiversity observed for cyanobacteria in river samples was high and was including 25 genera, such as from unicellular cyanobacteria (*Chamaesiphon*), unicellular cyanobacteria forming nanocytes (*Pleurocapsa*, *Chroococcidiopsis*, *Aliterella*), filamentous cyanobacteria (*Tychonema*, *Phormidium*, *Phormidesmis*, *Leptolyngbyaceae*, *Microseira*, *Wilmottia*) and filamentous cyanobacteria forming heterocytes (*Calothrix*). The genera *Chamaesiphon*, *Pleurocapsa*, *Tychonema* and *Calothrix* occurred with a high number of genotypes (>30). Previous unknown cyanobacteria included (i) the coccale cyanobacterium *Aliterella*, which has been described as a marine deep water or benthic species and (ii) the thin filamentous cyanobacterium genus *Phormidesmis* described from stones in oligotrophic glacial streams or subaerophytic from cold wet rocks (Fig. 12).

Taking all samples together the unicellular genera *Chamaesiphon, Aliterella, Chroococcidiopsis, Pleuorcapsa and the filamentous genera Leptolyngbyaceae, Phormidesmis, Phormidium, Pseudanabaena, Schizothrix, Tychonema* and *Calothrix* showed high read numbers and also up to two dozen of individual genotypes. In general the abundance (max 16S rDNA read numbers) and the number of ASVs was positively related (R = 0.72), Fig. 13.

In summary it can be stated that for cyanobacteria correspondence between microscopy and 16S rDNA sequencing is useful to confirm microscope-based identification of genera. Several previously unknown cyanobacteria have been detected that might require further study (genera *Aliterella*, *Phormidesmis*). Finally the 16S sequencing information can be useful to infer the toxigenic potential at certain sampling suites, e.g the *Tychonema* genotype Seq No34 which has been detected among river samples but has been linked to anatoxin-a production in the phytoplankton in lakes previously (L. Como, L. Garda, L. Iseo, L. Ledro, L. Maggiore, Staffelsee Nord).

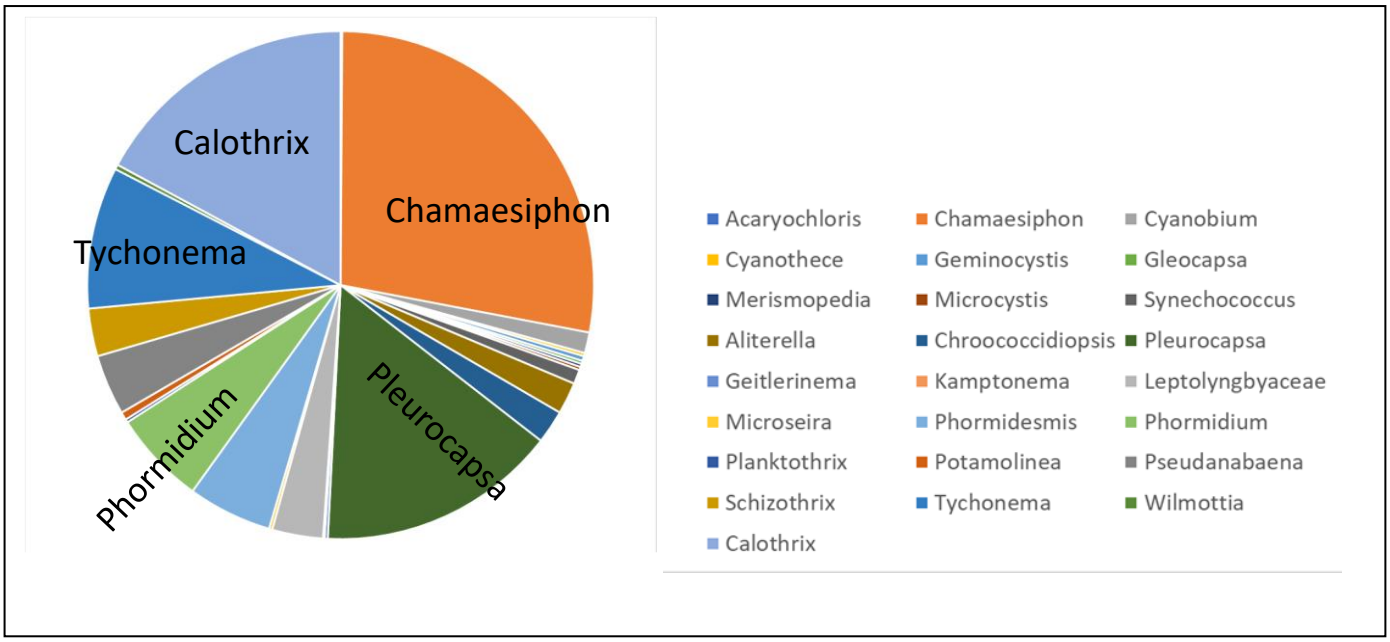


Fig. 12. Biodiversity of benthic cyanobacteria in rivers (n = 53). Section I (unicellular cyanobacteria) include the genera *Acaryochloris, Cyanobium, Cyanothece, Gleocapsa, Merismopedia, Chamaesiphon, Geminocystis, Microcystis, Synechococcus*, section II (unicellular cyanobacteria with nanocytes) include the genera *Aliterella, Chrococcidiopsis, Pleurocapsa*, section III (filamentous cyanobacteria) include the genera *Geiterlinema, Microseira, Planktothrix, Schizothrix, Kamptonema, Leptolyngbyaceae, Phormidium, Phorimdesmis, Potamolinea, Pseudanabaena, Tychonema, Wilmottia*, section IV (filamentous cyanobacteria with heterocytes) include the genus *Calothrix*.

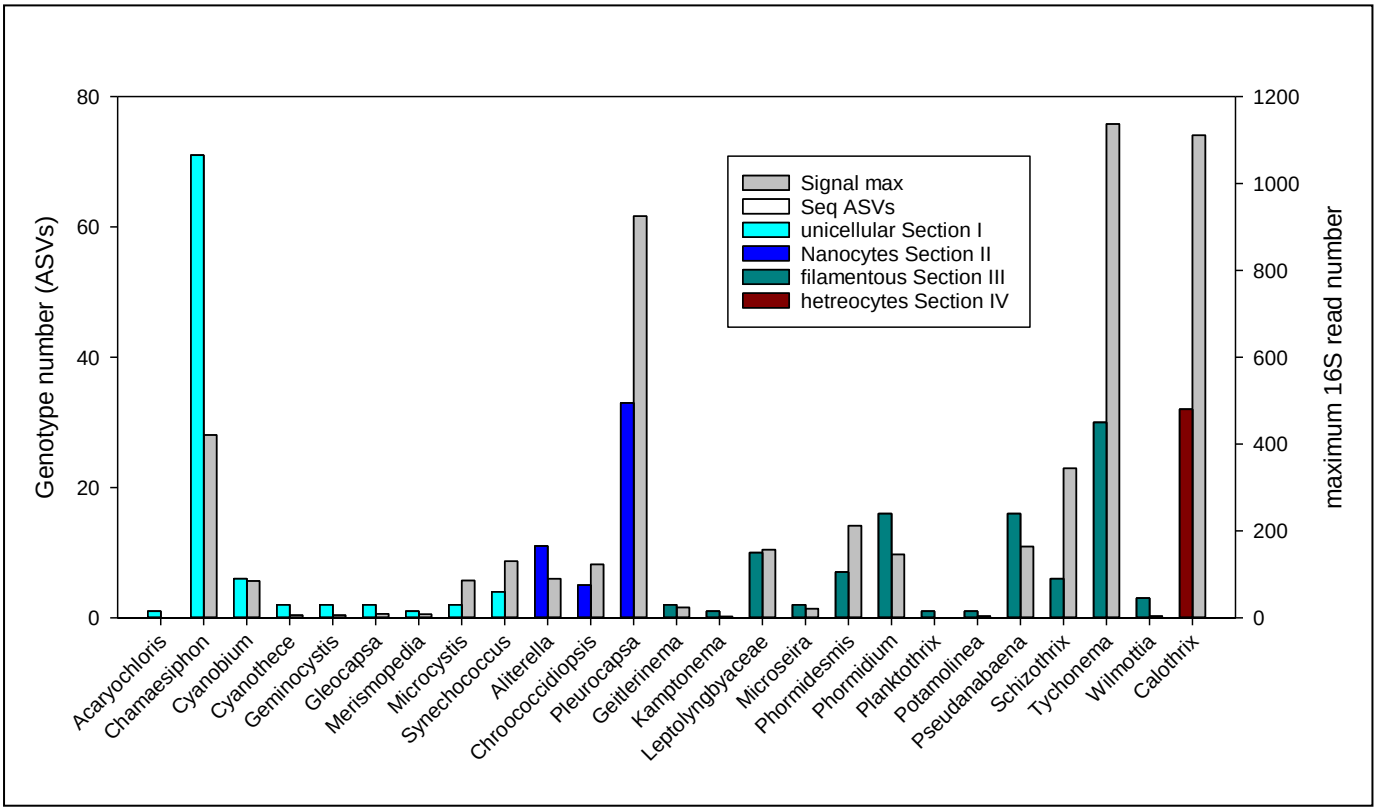


Fig. 13. Biodiversity of benthic cyanobacteria in rivers (n = 53). The cyanobacterial genera are listed according to morphological sections I-IV as described under Fig. 12.

Using 18S rDNA sequencing in parallel 7 divisions of eukaryotic algae were detected: **Chlorophyta** (5 classes Chlorophyceae, Nephrosemidophyceae, Pedinophyceae, Trebouxiophyceae, Ulvophyceae), **Streptophyta** (4 classes Zygnemophyceae, Klebsormidiophyceae, Coleochaetophyceae, Charophyceae), **Cryptophyta (**1 class Cryptophyceae), **Dinophyta** (1 class Dinophyceae), **Haptophyta** (2 classes Pavlovophyceae, Prymnesiophyceae), **Rhodophyta** (4 classes, Bangiophyceae, Compsopogonphyceae, Florideophyceae, Stylonematophyceae), **Ochrophyta** (5 classes Bacillariophyceae, Chrysophyceae, Eustigmatophyceae, Phaeophyceae, Xanthophyceae). With regard to genotype (ASV) and maximum read numbers the Chlorophyta and the Ochrophyta (Bacillariophyceae) were most frequent (Fig. 14), followed by Rhodophyta and Streptophyta. As expected the divisions representing flagellates (Dinophyta, Haptophyta, Cryptophyta) were least abundant.

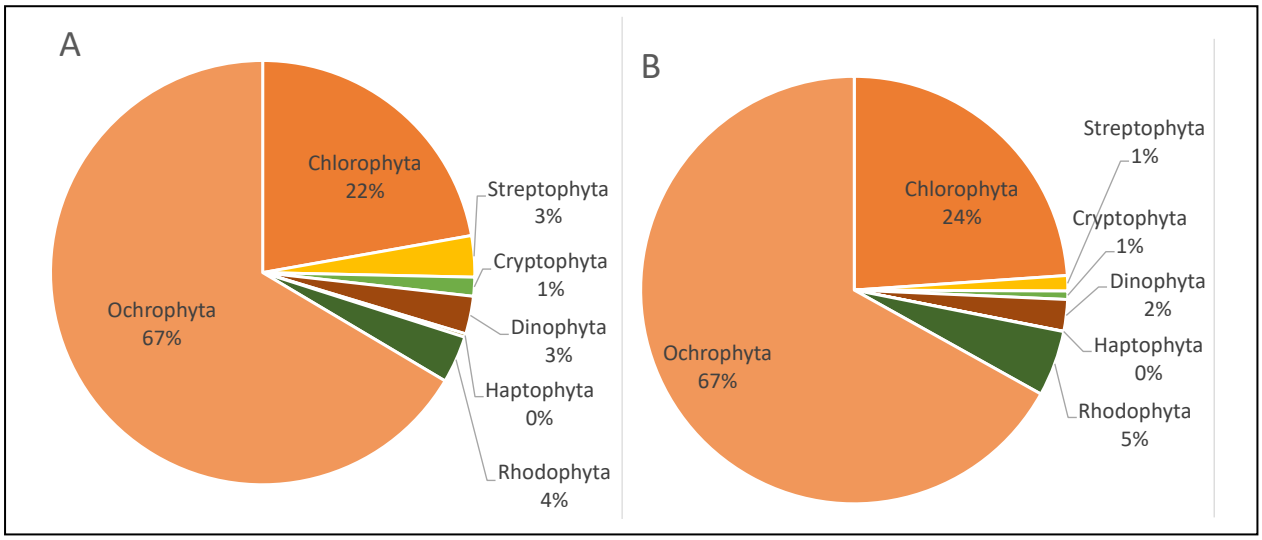


Fig. 14. Divisions of eukaryotic algae in rivers (n = 53). A) Proportion in genotype (ASV) numbers (n= 1527), B) Proportion in maximum read numbers (n= 33149).

Within Chlorophyta the Chlorophyceae were comprised of the filamentous orders Chaetophorales (e.g. *Stigeoclonium*), Oedogoniales (*Oedogonium*), the coccale orders Chlorococcales (e.g. *Scenedesmus* and undetermined), Sphaeropleales (e.g. *Mychonastes*) and Tetrasporales, and the monadoid orders Chlamydomonadales (e.g. *Chlorochytrium*) and Volvocales (e.g. *Chlorogonium* and undetermined). The Trebouxiophyceae contained Chlorellales with coccoid organisation type. Notably Ulvophyceae contained two abundant filamentous orders Ulotrichales (mostly undetermined) and Ulvales relatives (mostly undetermined), (Fig. 15).

Within Ochrophyta the classes Bacillariophyceae and Chrysophyceae had the highest share. Not surprisingly within Bacillariophyceae the Pennales (araphid-pennate or raphid pennate) were found most abundant comprising a large number of genera typically recorded during microscopical counting. Within Chrysophyceae the Hydrurales (the macroalgae *Hydrurus foetidus*) contributed most to read numbers (Fig. 16).

Within Rhodophyta 3 classes were frequently found: Bangiophyceae incl. filamentous Bangiales and coccale Cyanidiales, Florideophyceae incl. filamentous macroscopic Batrachospermales and Hildenbrandiales, and Stylonematophyceae incl. filamentous Porphyridiales (*Chroodactylon*). Taking

all 18S read numbers on eukaryotic algae orders together the ASV (genotype) number and 18S max read number were positively related (R = 0.80, n = 51). In other words, orders occurring more frequently in general also shared higher ASV (genotype) numbers (Suppl. Fig. 16).

In summary the additional information derived from both 16S and 18R rDNA sequencing data for phytobenthos in rivers includes

+) taxonomic confirmation/corrections regarding the expert assessments enabling higher reproducibility (usually including diatoms only)

+) Detection of invasive taxa or species that cannot be easily differentiated in the microscope via an increased resolution of taxa lists

+) information on potentially toxigenic genotypes (i. e. *Tychonema*) of cyanobacteria useful for early warning under conditions of increased phytobenthic growth.

+) information on all eukaryotic algal groups (not only diatoms) which co-occur more frequently and influence the growth of diatoms, i.e. macroscopic algae with epiphytic diatoms.

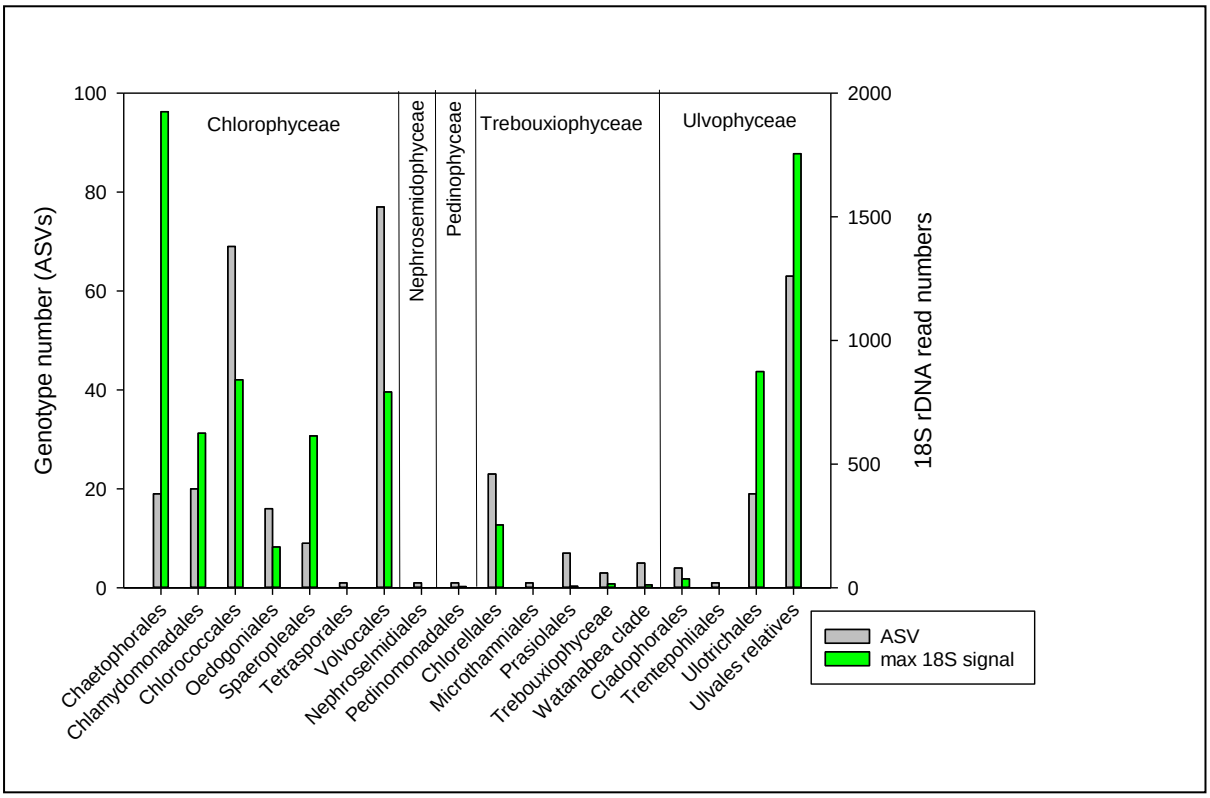


Fig. 15: Biodiversity of Chlorophyta in rivers (n=53): Genotype (ASV) numbers and maximum 18S rDNA read numbers for orders among the 5 classes.

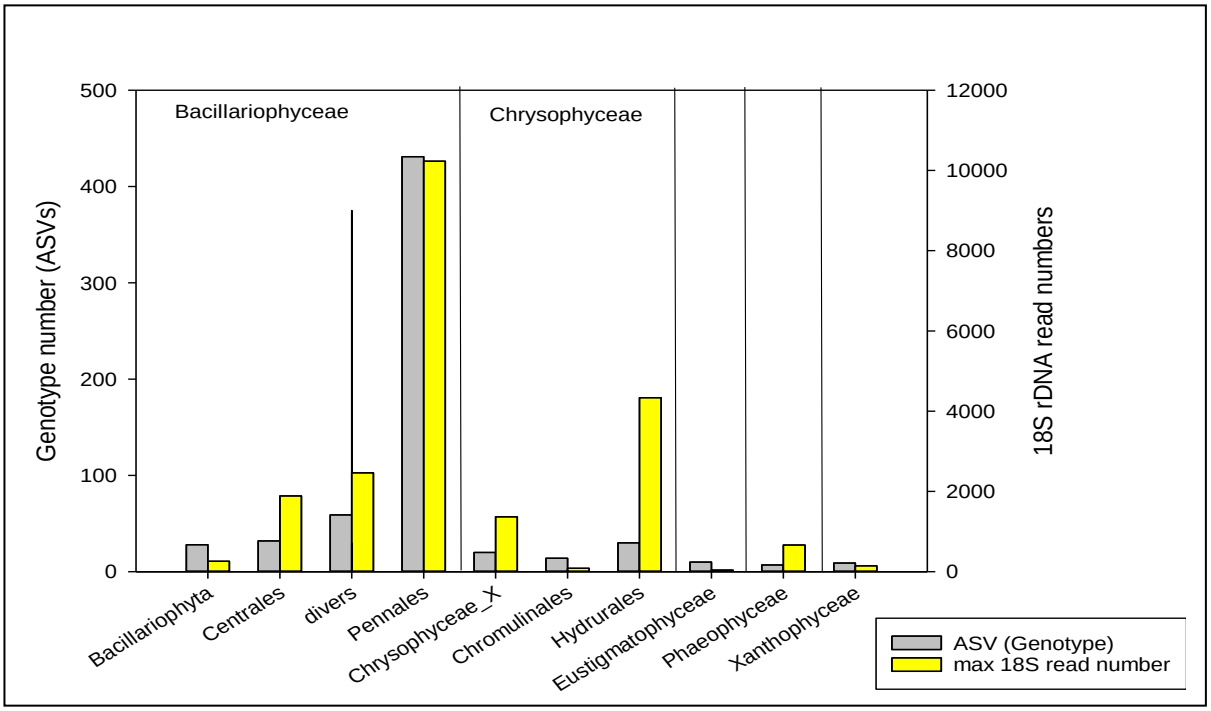


Fig. 16: Biodiversity of Ochrophyta in rivers (n=53): Genotype (ASV) numbers and maximum 18S rDNA read numbers for orders among the 5 classes.

Using rbcL the biodiversity was assessed for benthic diatoms only, taking advantage of the curated database differentiating diatom taxa on the level of morphospecies incl morphological varieties (R-Syst::diatom, Rimet et al. 2016). Since all samples were counted for diatom composition in the microscope (n = 53) in parallel to sequencing (i) a direct comparison as well as (ii) an estimate of the additional information through rbcL sequencing was obtained.

In total 82 diatom taxa were identified through HTS which were detected at least partly also during LM counting. Most frequent occurring taxa (in more than half of the samples) included *Achnanthidium minutissimum, Amphora pediculus, Navicula cryptotenella, Nitzschia dissipata var dissipata, Encyonema minutum, Nitzschia fonticola, Achnanthidium pyrenaicum, Navicula antonii, Nitzschia palea, Reimeria sinuata, Navicula gregaria, Navicula tripunctata* (Fig. 17, taxa names are ordered in decreasing frequency of occurrence).

On the other hand 188 diatom taxa recorded during LM counting were not well supported through HTS, i.e. the intersection between both methods was close to zero. Out of these 188 LM detected taxa the majority (161 taxa) occurred with a maximum proportion < 5% according to microscopical counting and thus were considered of minor relevance. Consequently, in general the diatom taxa not detected through HTS had a low maximum proportion in microscopical counting (< 1 %) and typically occurred in less than 4 samples (Suppl. Fig. 17 A, B).

Nevertheless 27 taxa showed a maximum proportion ranging from 5-54% according to microscopical counting but were not included through HTS. The most frequent taxa (10-54% of maximum proportion in the microscope) included *Cocconeis placentula var euglypta, Gomphonema angustivalva, Cocconeis placentula var lineata, Pinnularia spec, Psammothidium subatomoides, Achnanthidium rivulare, Gomphonema olivaceum var olivaceoum, Achnanthidium minutissima var affinis, Fragilaria capucina var austriaca, Gomphonema lateripunctatum, Achnanthidium atomoides, Gomphonema pumilum var*

elegans (taxa names are ordered with decreasing max. proportion). It is possible that (ongoing) taxonomic revisions are contributing to the exclusion of these more frequent taxa.

In order to estimate the additional information through rbcL sequencing the frequency of occurrence of diatom taxa recorded through HTS and (at least partly) through LM vs. diatom taxa recorded through HTS only were compared. Among the 82 diatom taxa detected by both methods the majority of the taxa (n = 57) were detected more frequently through HTS, ranging from 1.1 - 18 (54)-fold increased frequency of detection (median 3.3 - fold), (Fig. 18). Vice versa, the minor share of taxa (n = 18) were detected less frequently through HTS, i.e. 0.26 – 0.93 – fold lower detection rate (median 0.54-fold).

In particular the diatom taxa detected more frequently through HTS included *Nitzschia dissipata var media, Nitzschia sigmoidea, Nitzschia soratensis, Thalassiosira weissflogii, Gyrosigma sciotense, Cymbella spec., Cocconeis placentula var. placentula, Cymbella compacta, Encyonopsis spec., Didymosphenia geminata var geminate, Encyonopsis subminuta, Nitzschia spec., Amphora spec., Fragilaria gracilis, Caloneis fontinalis, Fragilaria radians, Gomphonema parvulum var parvulum fo saprophilum, Cymatopleura elliptica var elliptica, Cyclotella meneghiniana, Fragilaria spec., Fistulifera saprophila, Cymatopleura solea var solea* (taxa names ordered along a 54 - 5-fold range of higher frequency of detection).

It is anticipated that several factors contribute to the difference between methods, i.e. in particular the higher frequency of detection through HTS. For example in one case the taxon *Nitzschia dissipata var dissipata* (LM) appeared to be replaced by the taxon N*itzschia dissipata var. media* (HTS), (Fig. 18.). For several closely related diatom taxa complexes the identification according to the morphology is difficult (e.g. the *Achnanthidium minutissium* complex). This difficulty may explain that *A. eutrophilum, A. delmontii, A. subatomus* are detected more frequently (1.5 - 4.7 -fold) via HTS, while *A. minutissimum* and *A. pyrenaicum* are detected slightly less frequently (0.8 - 0.9- fold). In addition, in analogy to picosized cyanobacteria and eukaryotic flagellates (part 1) small sized diatom cells (frustules) are overlooked and probably underestimated in the LM counting procedure, such as *Mayamaea atomus var permitis* (2.2-fold higher detection rate).

Finally 97 diatom taxa were not recorded by LM based counting but through HTS. It should be noted that 44 out of this taxa occurred only once or twice, and thus were considered of minor relevance. The other 52 taxa occurred 3 – 38 times among river samples and might be considered more relevant in terms of biodiversity. In particular those taxa included *Caloneis spec, Encyonema spec, Gomphonella olivacea, Surirella spec, Fragilaria spec, Ulnaria spec, Staurosira spec, Discostella woltereckii, Nitzschia pusilla, Cymbella lanceolata var laneolata, Planothidium spec, Gomphonema olivaceum var olivaceoides, Diploneis suboralis, Nitzschia capitellata, Nitzschia draveillensis, Brachysira spec, Cyclotella costei, Iconella sp., Nitzschia supralitorea* (taxa names order with decreasing frequency from 38 – 10 times). Often the respective genotype was not assigned implying the occurrence of cryptic or yet unknown species.

In summary it can be stated that a significant share of diatom taxa was represented by both methods supporting modern estimates of diatom based biodiversity in rivers. The additional information through HTS can be summarized as follows:

+) HTS analysis supports through confirmation of the microscopy based results and supports diatom morphospecies differentiation in field samples

+) The confirmed detection of *Achnanthidium delmontii* in many river samples is of relevance since this is an invasive species and not readily differentiated by light microscopy from the more abundant *A. pyrenaicum*.

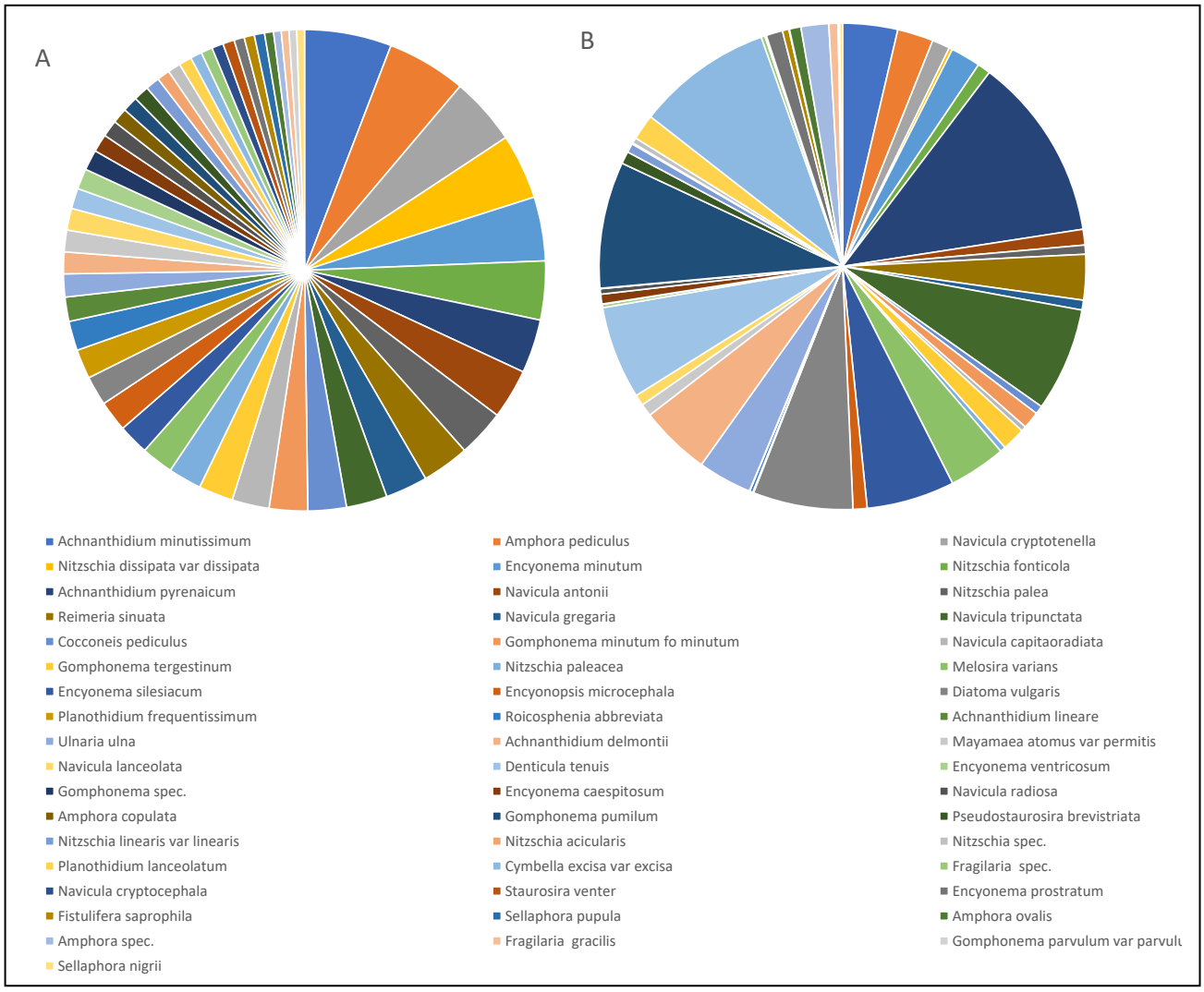


Fig. 17: Diatom taxa recorded through HTS and at least partly through LM on the level of morphospecies incl morphological varieties (R-Syst::diatom, Rimet et al. 2016). In total 82 diatom taxa showed significant overlap. (A) Relative frequency of occurrence (n =53), (B) relative maximal signal intensity (rbcL read number). Diatom taxa names are ordered according to decreasing frequency of occurrence. For reasons of clarity only 52 taxa occurring >10% of all samples are shown (n = 53).

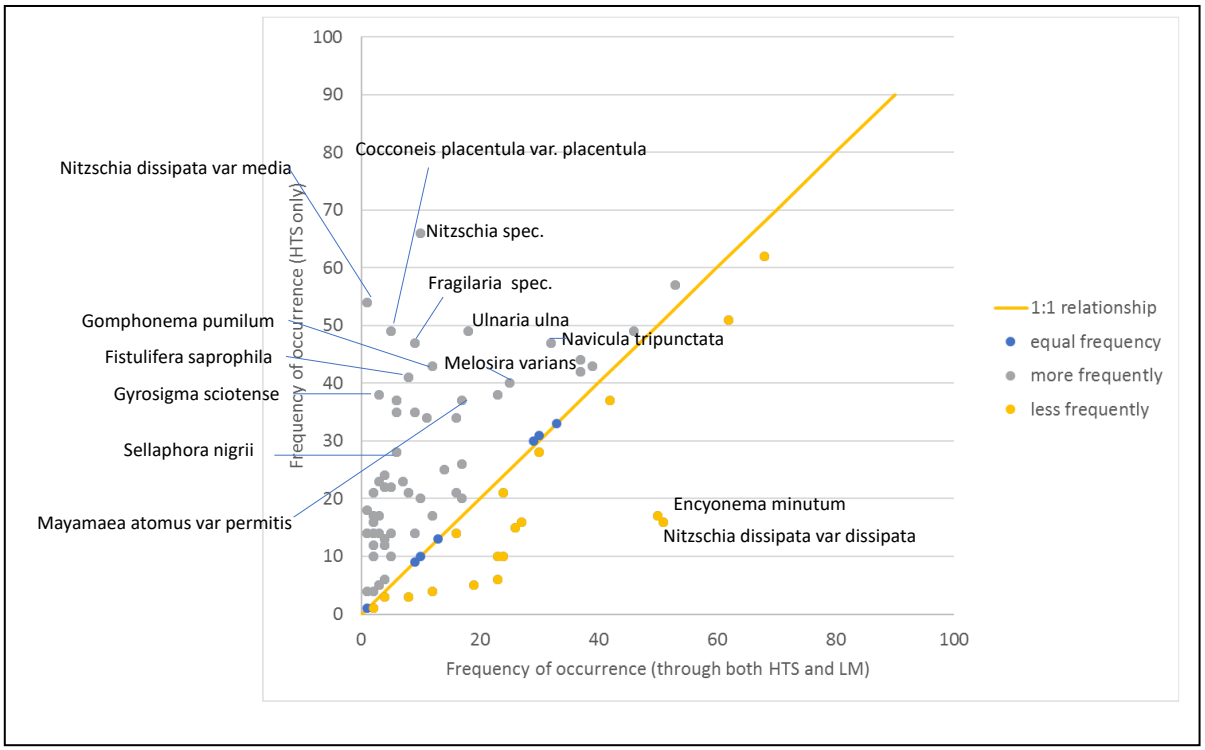


Fig. 18: Frequency of occurrence of diatom taxa recorded through HTS and at least partly through LM on the level of morphospecies incl morphological varieties (x-axis) vs recorded through HTS only (y-axis) from river samples (n=53). The grey symbols above the 1:1 line represent an increased detection frequency through HTS only.

References

EU Commission staff working document, 2018. Report from the Commission to the Council and the European Parliament on the implementation of Council Directive 91/676/EEC concerning the protection of waters against pollution caused from agricultural sources based on Member State reports for the period 2012-2015 (accessed on 16.6.2021 at https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52018SC0246&from=EN)

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., Holmes, S.P., 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nat. Methods 13, 581–583. https://doi.org/10.1038/nmeth.3869

Callahan, B.J., McMurdie, P.J., Holmes, S.P., 2017. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. ISME J. 11, 2639–2643. https://doi.org/10.1038/ismej.2017.119

Cerasino, L., C. Capelli, & N. Salmaso, 2017. A comparative study of the metabolic profiles of common nuisance cyanobacteria in southern perialpine lakes. Advances in Oceanography and Limnology 8: 22–32.

Gaget, V., A. R. Humpage, Q. Huang, P. Monis & J. D. Brookes, 2017. Benthic cyanobacteria: A source of cylindrospermopsin and microcystin in Australian drinking water reservoirs. Water Research 124:454-464 doi:https://doi.org/10.1016/j.watres.2017.07.073.

Komárek, J., J. Kaštovský & J. Jezberová, 2011. Phylogenetic and taxonomic delimitation of the cyanobacterial genus Aphanothece and description of Anathece gen. nov. European Journal of Phycology 46(3):315-326 doi:10.1080/09670262.2011.606373.

Komarek, J., F. Hindak & J. Jezberova, 2016. Review of the cyanobacterial genus Eucapsis. Nova Hedwigia 103(3-4):441-456 doi:10.1127/nova_hedwigia/2016/0363.

Komárková, J., J. Jezberová, O. Komárek & E. Zapomělová, 2010. Variability of Chroococcus (Cyanobacteria) morphospecies with regard to phylogenetic relationships. Hydrobiologia 639(1):69-83 doi:10.1007/s10750-009-0015-3.

Lepistö, L., J. Rapala, C. Lyra, K. Berg, K. Erkomaa & J. Issakainen, 2005. Occurrence and toxicity of cyanobacterial blooms dominated by Anabaena lemmermannii P. RICHTER and Aphanizomenon spp. in boreal lakes in 2003. Algological Studies 117:315-328 doi:10.1127/1864-1318/2005/0117-0315.

Lyra, C., S. Suomaleinen, M. Gugger, C. Vezie, P. Sundman, L. Paulin & K. Sivonen, 2001. Molecular characterization of planktic cyanobacteria of Anabaena, Aphanizomenon, Microcystis and Planktothrix genera. International Journal of Systematic and Evolutionary Microbiology 51:513-526.

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17, 10–12.

Perkerson III, R. B., J. R. Johansen, L. Kovácik, J. Brand, J. Kaštovský & D. A. Casamatta, 2011. A UNIQUE *PSEUDANABAENA*LEAN (CYANOBACTERIA) GENUS *NODOSILINEA* GEN. NOV. BASED ON MORPHOLOGICAL AND MOLECULAR DATA1. Journal of Phycology 47(6):1397-1412 doi:https://doi.org/10.1111/j.1529-8817.2011.01077.x.

Pinevich, A., N. Velichko & N. Ivanikova, 2012. Cyanobacteria of the Genus *Prochlorothrix*†. Frontiers in Microbiology 3(173) doi:10.3389/fmicb.2012.00173.

Rajaniemi-Wacklin, P., A. Rantala, M. Mugnai, S. Turicchia, S. Ventura, J. Komarkova, L. Lepistö & K. Sivonen, 2005. Correspondence between phylogeny and morphology of Snowella spp. and Woronichinia naegeliana, cyanobacteria commonly occurring in lakes. Journal of Phycology 42:226-232.

Rajaniemi, P., P. Hrouzek, K. Kastovska, R. Willame, A. Rantala, L. Hoffmann, J. Komarek & K. Sivonen, 2005. Phylogenetic and morphological evaluation of the genera Anabaena, Aphanizomenon, Trichormus and Nostoc (Nostocales, Cyanobacteria). International Journal of Systematic and Evolutionary Microbiology 55:11-26 doi:10.1099/ijs.0.63276-0.

R Core Team, 2020. R: A language and environment for statistical computing (v. 4.0.3). R Foundation for Statistical Computing, Vienna.

Salmaso, N., Riccioni, G., Pindo, M., Vasselon, V., Domaizon, I., Kurmayer, R., 2021a. Metabarcoding protocol – Analysis of Bacteria (including Cyanobacteria) using the 16S rRNA gene and a DADA2 pipeline. Zenodo. https://doi.org/https://doi.org/10.5281/zenodo.5232772

Salmaso, N., Riccioni, G., Pindo, M., Kurmayer, R., Vasselon, V., Domaizon, I., 2021b. Metabarcoding protocol – Analysis of protists using the 18S rRNA gene and a DADA2 pipeline. Zenodo. https://doi.org/https://doi.org/10.5281/zenodo.5233527

Sánchez-Baracaldo, P., G. Bianchini, A. Di Cesare, C. Callieri & N. A. M. Chrismas, 2019. Insights Into the Evolution of Picocyanobacteria and Phycoerythrin Genes (mpeBA and cpeBA). Frontiers in Microbiology 10(45) doi:10.3389/fmicb.2019.00045.

Shen, L., S. Ma, F. Cai, G. Yu, S. Li & R. Li, 2018. Polyphasic examination on Merismopedia tenuissima CHAB 7021 from Ganjiang River, China revealed the polyphyly of the genus Merismopedia (Cyanobacteria). Journal of Oceanology and Limnology 36(4):1157-1165 doi:10.1007/s00343-018-7341-0.

Strunecký, O., J. Komárek & J. Šmarda, 2014. Kamptonema (Microcoleaceae, Cyanobacteria), a new genus derived from the polyphyletic Phormidium on the basis of combined molecular and cytomorphological markers. Preslia 86:193-207.

Vautier, M., Vasselon, V., Chardon, C., Rimet, F., Bouchez, A., Domaizon, I., 2020. DNA extraction from environmental biofilm using the NucleoSpin® Soil kit (MACHEREY-NAGEL). protocols.io. https://doi.org/dx.doi.org/10.17504/protocols.io.bd52i88e

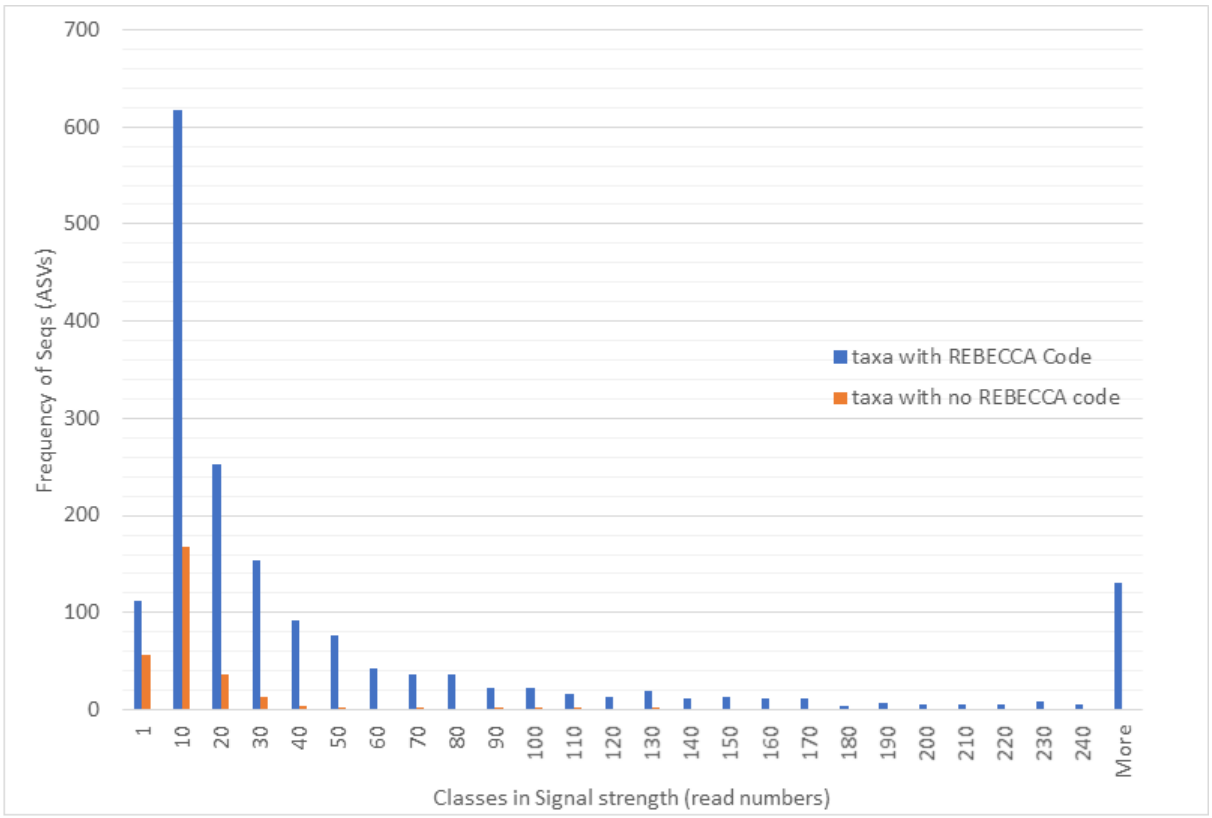
Vautier, M., Chardon, C., Capelli, C., Kurmayer, R., Salmaso, N., Domaizon, I., 2021. Plankton DNA extraction from Sterivex filter units. protocols.io. https://doi.org/dx.doi.org/10.17504/protocols.io.bvgzn3x6

Vollenweider, R. & J. Kerekes, 1982. Eutrophication of waters. Monitoring, assessment and control. OECD Cooperative programme on monitoring of inland waters (Eutrophication control), Environmental Directorate, OECD, Paris:154pp.

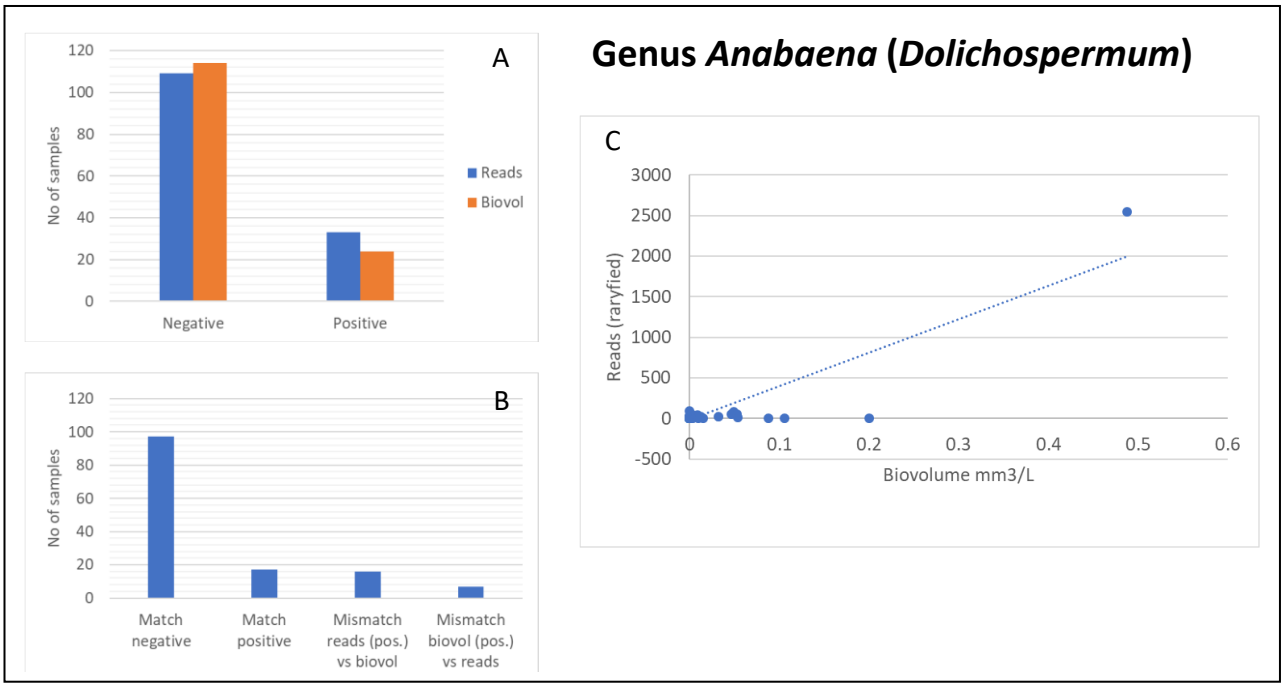
Wacklin, P., L. Hoffmann & J. Komárek, 2009. Nomenclatural validation of the genetically revised cyanobacterial genus Dolichospermum (RALFS ex BORNET et FLAHAULT) comb. nova. Fottea 9(1):59-64 doi:10.5507/fot.2009.005.

Zwart, G., M. P. Kamst-van Agterveld, I. Van Der Werff-Staverman, F. Hagen, H. L. Hoogveld & H. J. Gons, 2005. Molecular characterization of cyanobacterial diversity in a shallow eutrophic lake. Environmental Microbiology 7(3):365-377 doi:https://doi.org/10.1111/j.1462-2920.2005.00715.x.

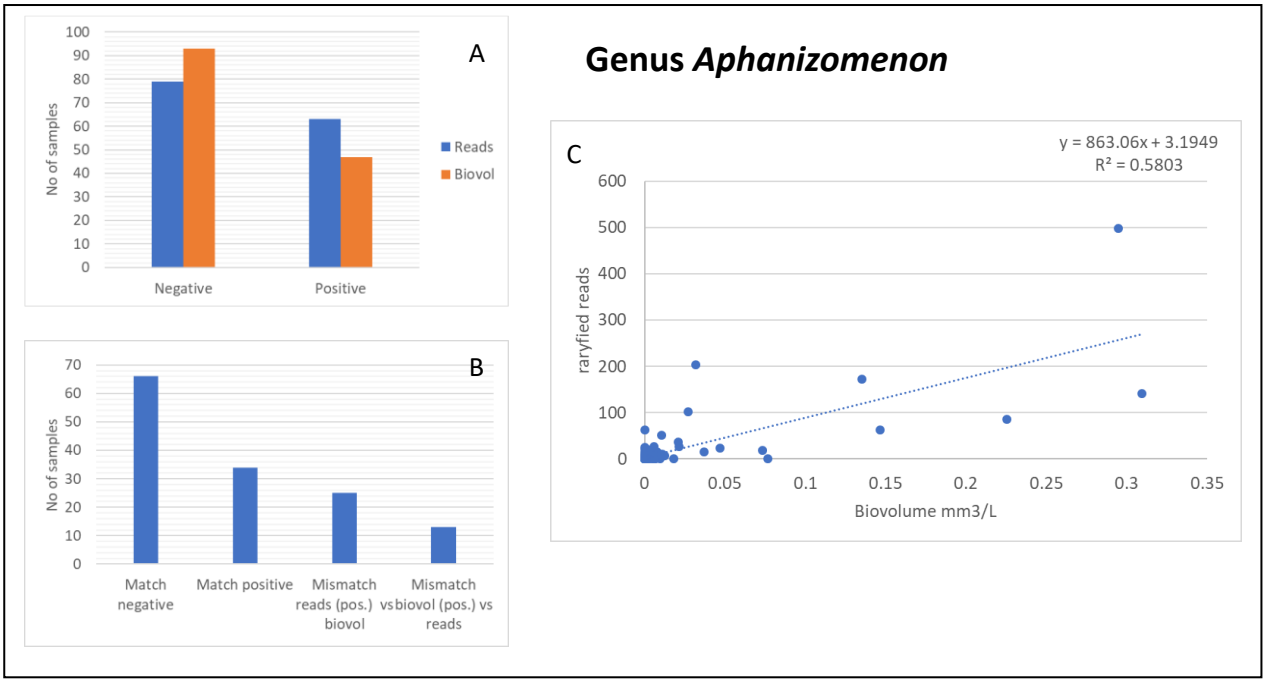
Appendix (Suppl. Figs and Suppl. Tables)



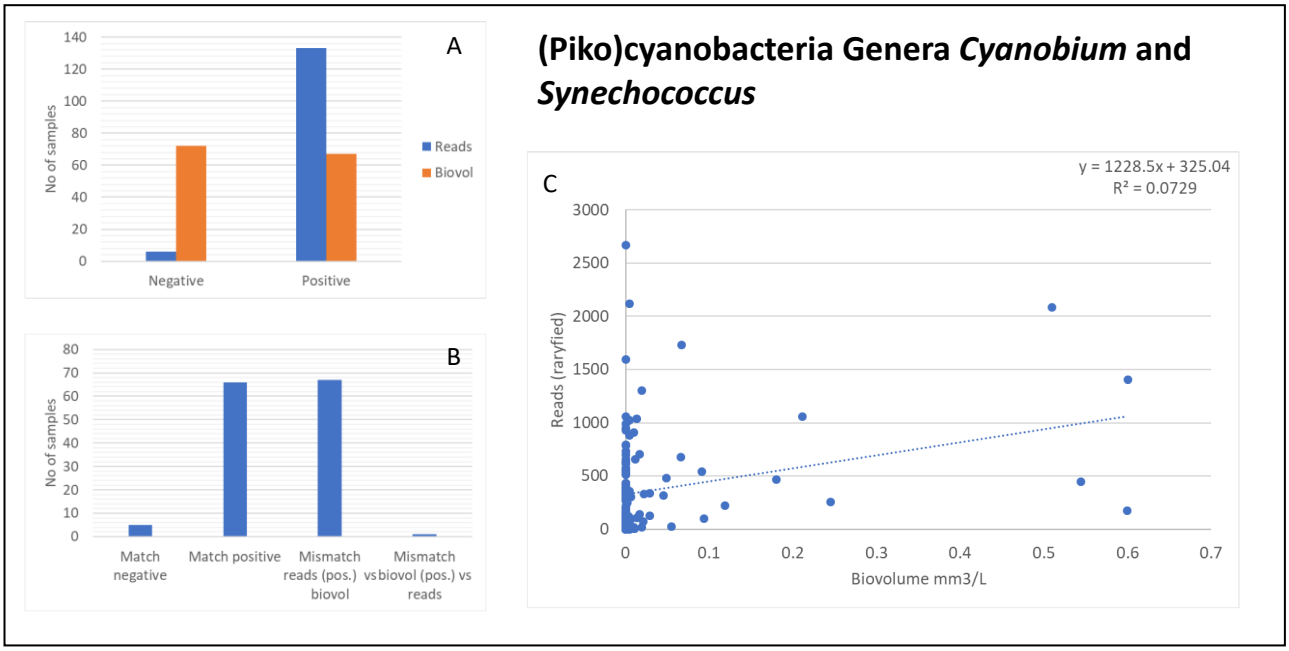
Suppl. Fig. 1. Signal strength (number of rarified reads) in relation to taxonomic assignment via the REBECCA code (blue bars) and Seq (ASVs) which could not be assigned to a REBECCA code (orange bars).



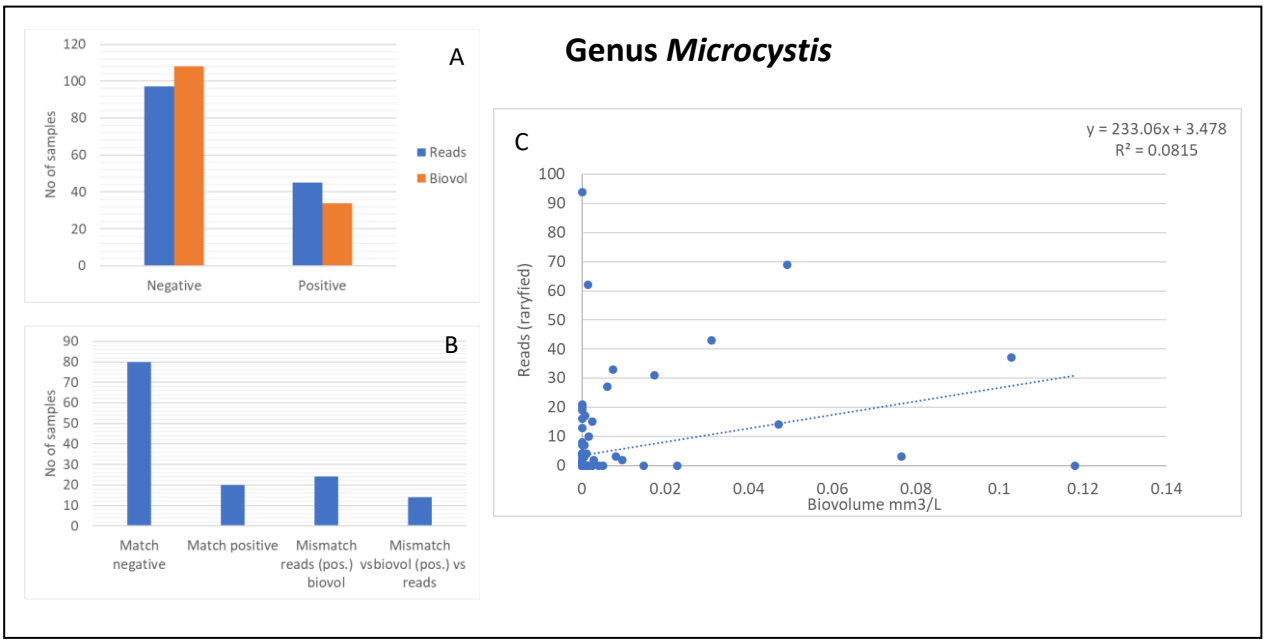
Suppl. Fig. 2. (A, B) Qualitative and (C) quantitative comparison of sequencing vs microscopy for detection of the genus *Dolichospermum* (*Anabaena*) in phytoplankton from lakes across alpine space (n = 140). Two outliers (biovolume 12.62, 2.58 mm^3/L) were excluded (R^2 = 0.23).



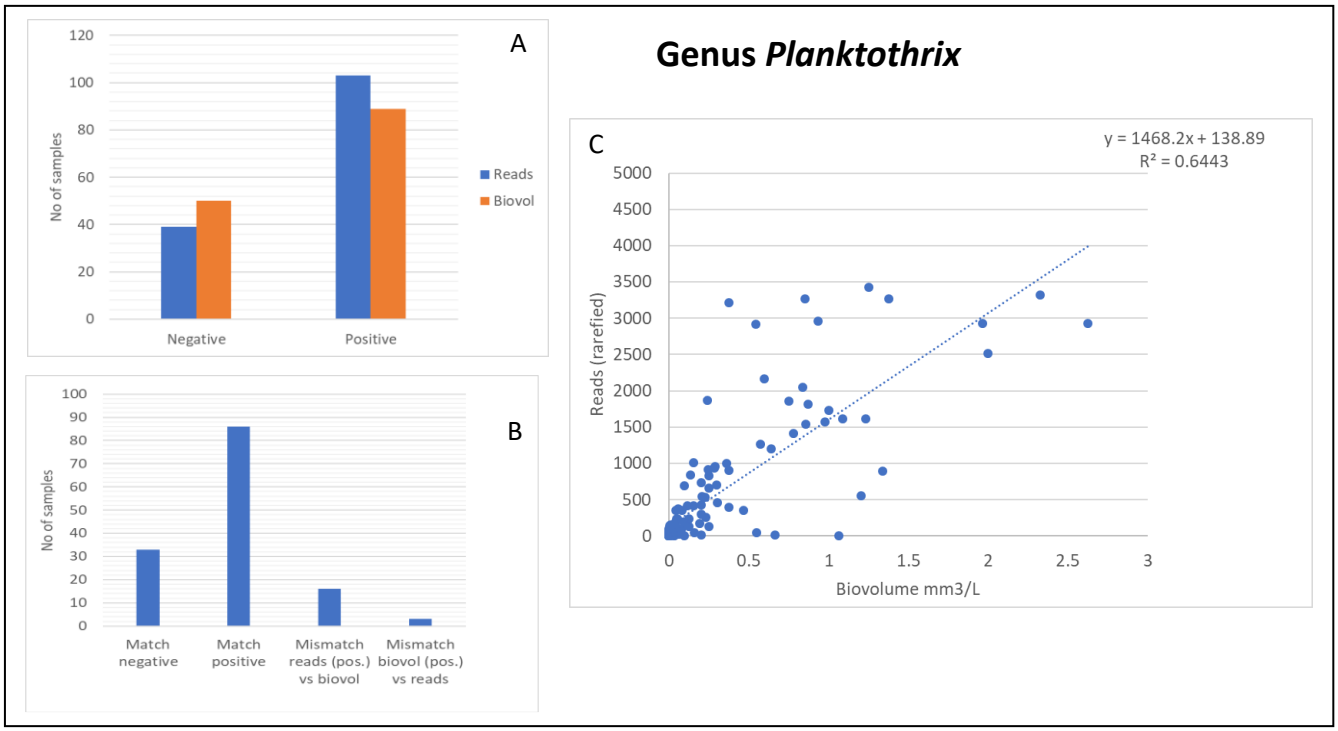
Suppl. Fig. 3. (A, B) Qualitative and (C) quantitative comparison of sequencing vs microscopy for detection of the genus *Aphanizomenon* in phytoplankton from lakes across alpine space (n = 140). Linear regression (R^2 = 0.58).



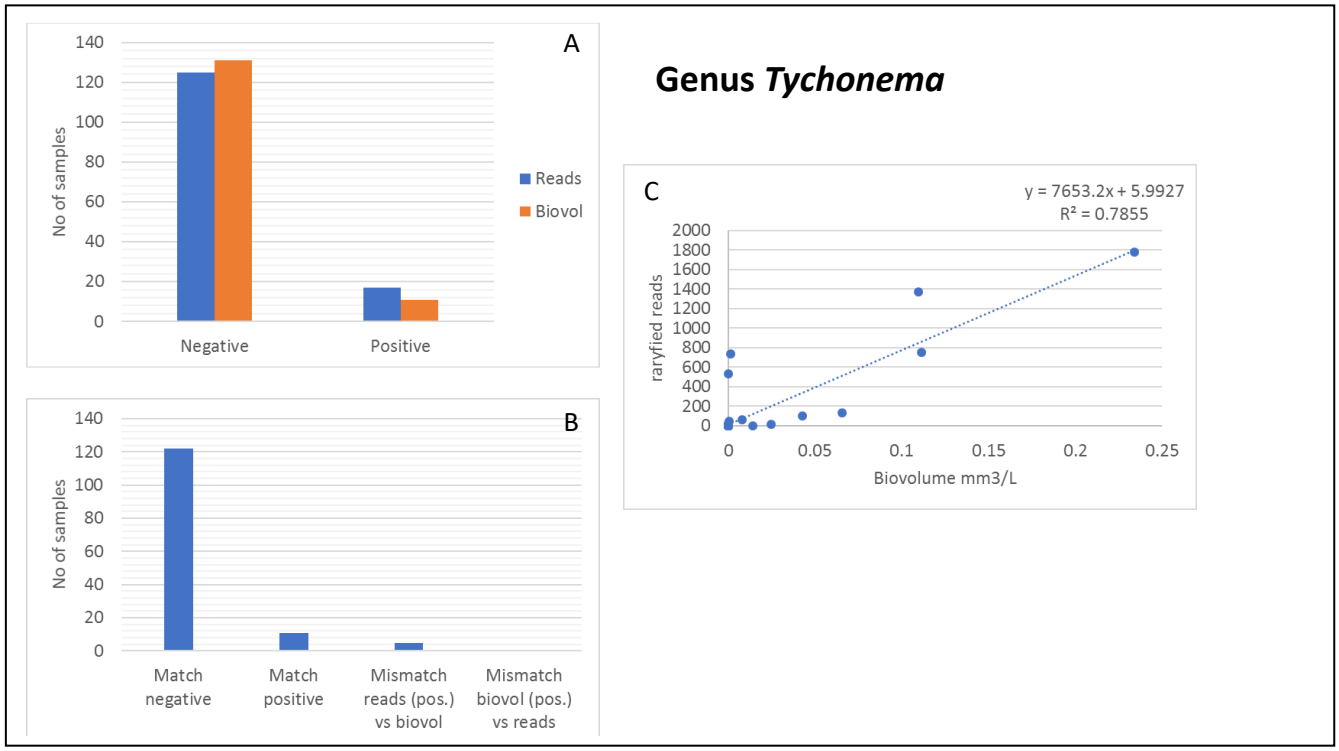
Suppl. Fig. 4. (A, B) Qualitative and (C) quantitative comparison of sequencing vs microscopy for detection of pikocyanobacteria (*Synechococcus, Coenobium, Aphanocapsa, Aphanothece*) in phytoplankton from lakes across alpine space (n = 140). Linear regression (R^2 = 0.07).



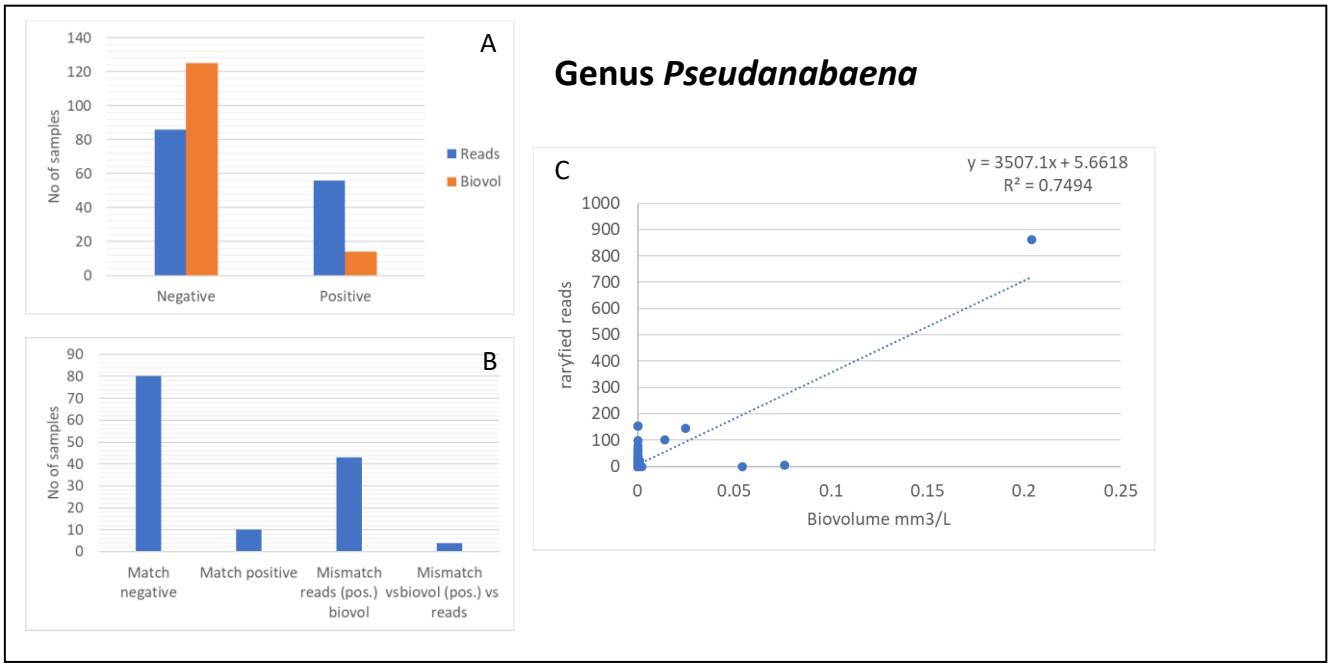
Suppl. Fig. 5. (A, B) Qualitative and (C) quantitative comparison of sequencing vs microscopy for detection of the genus *Microcystis* in phytoplankton from lakes across alpine space (n = 140). Linear regression (R^2 = 0.08), 2 outliers (biovolume 0.405, 0.32 mm3/L) excluded.



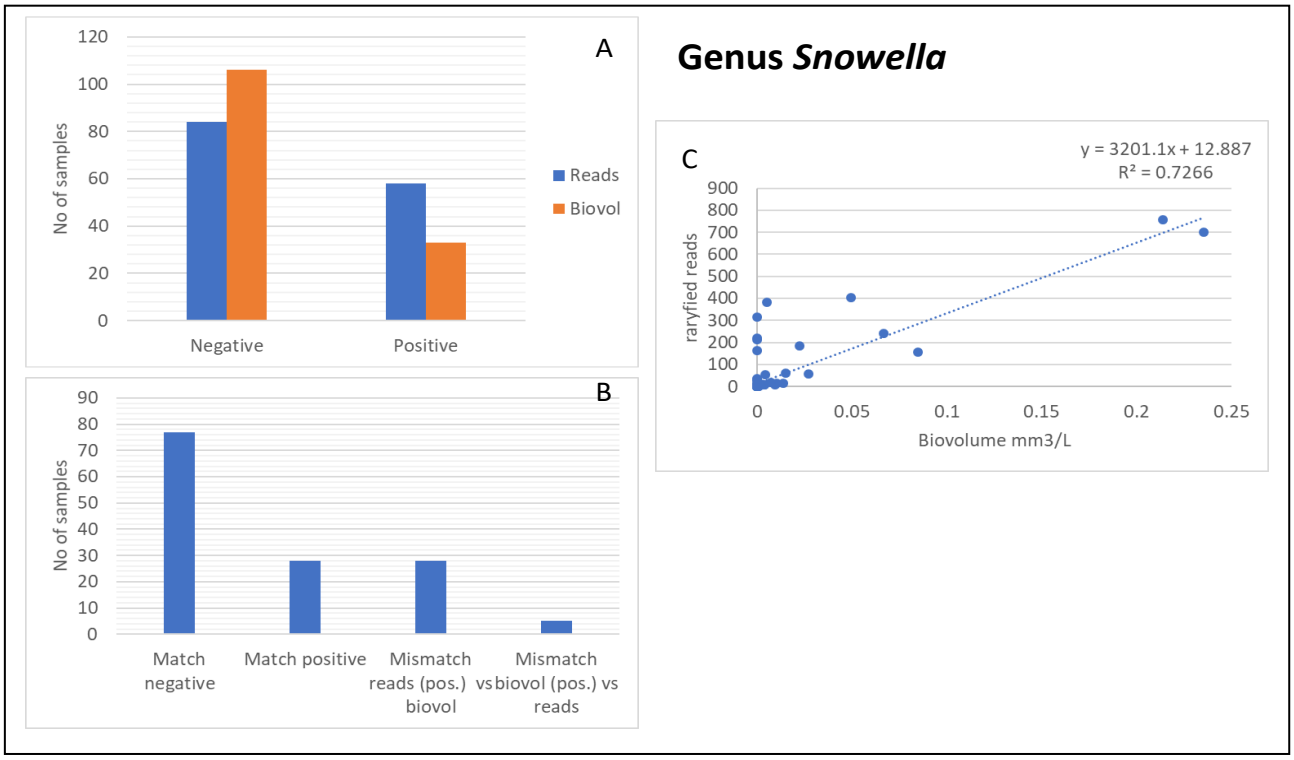
Suppl. Fig. 6. (A, B) Qualitative and (C) quantitative comparison of sequencing vs microscopy for detection of the genus *Planktothrix* in phytoplankton from lakes across alpine space (n = 140). Linear regression (R^2 = 0.64), 2 outliers (biovolume 14.08, 21.92mm^3/L) excluded.



Suppl. Fig. 7. (A, B) Qualitative and (C) quantitative comparison of sequencing vs microscopy for detection of the genus *Tychonema* in phytoplankton from lakes across alpine space (n = 140). Linear regression (R^2 = 0.79).

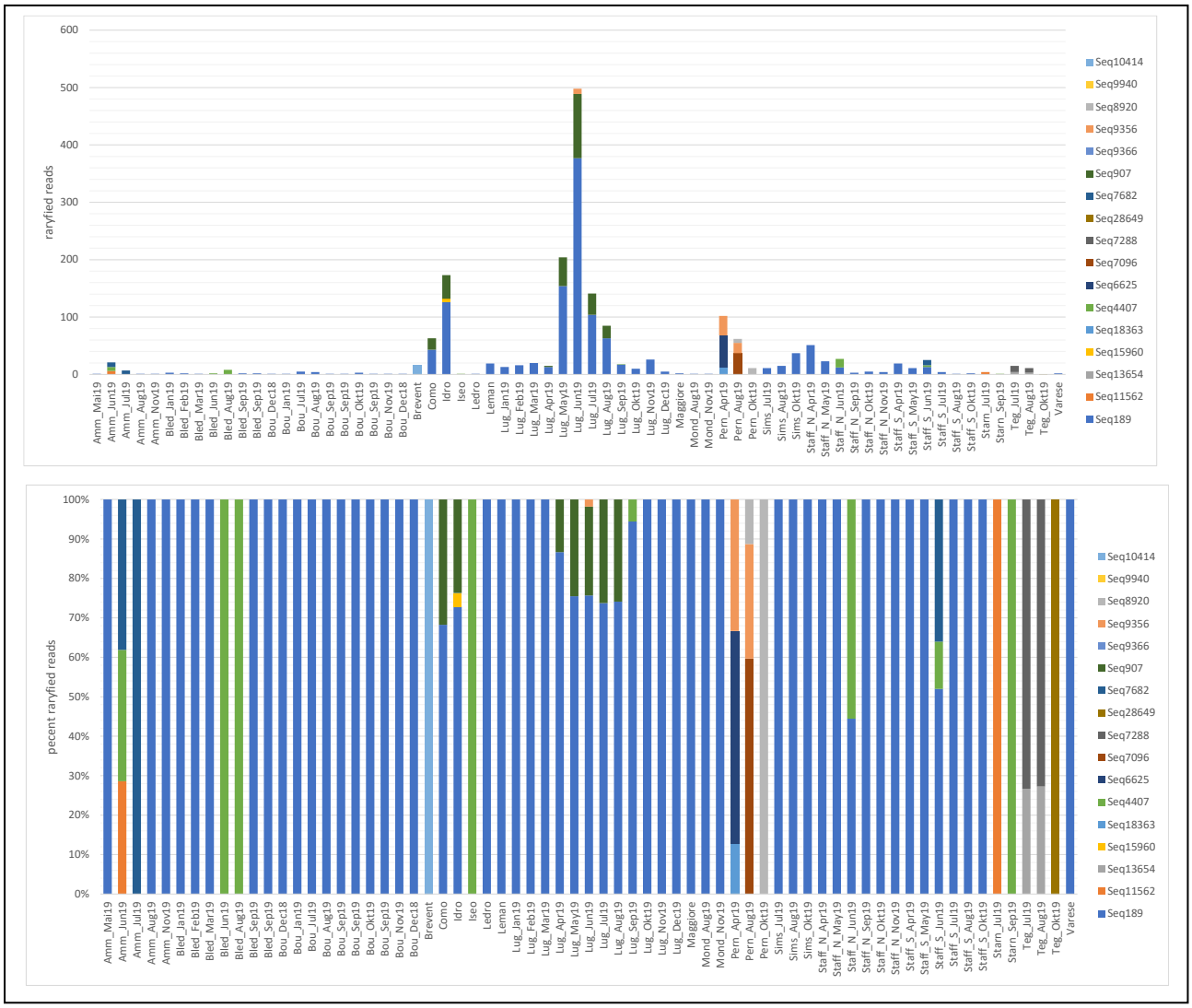


Suppl. Fig. 8. (A, B) Qualitative and (C) quantitative comparison of sequencing vs microscopy for detection of the genus *Pseudanabaena* in phytoplankton from lakes across alpine space (n = 140). Linear regression (R^2 = 0.75), 2 outliers (biovolume 1.74, 3.77 mm^3/L) excluded.

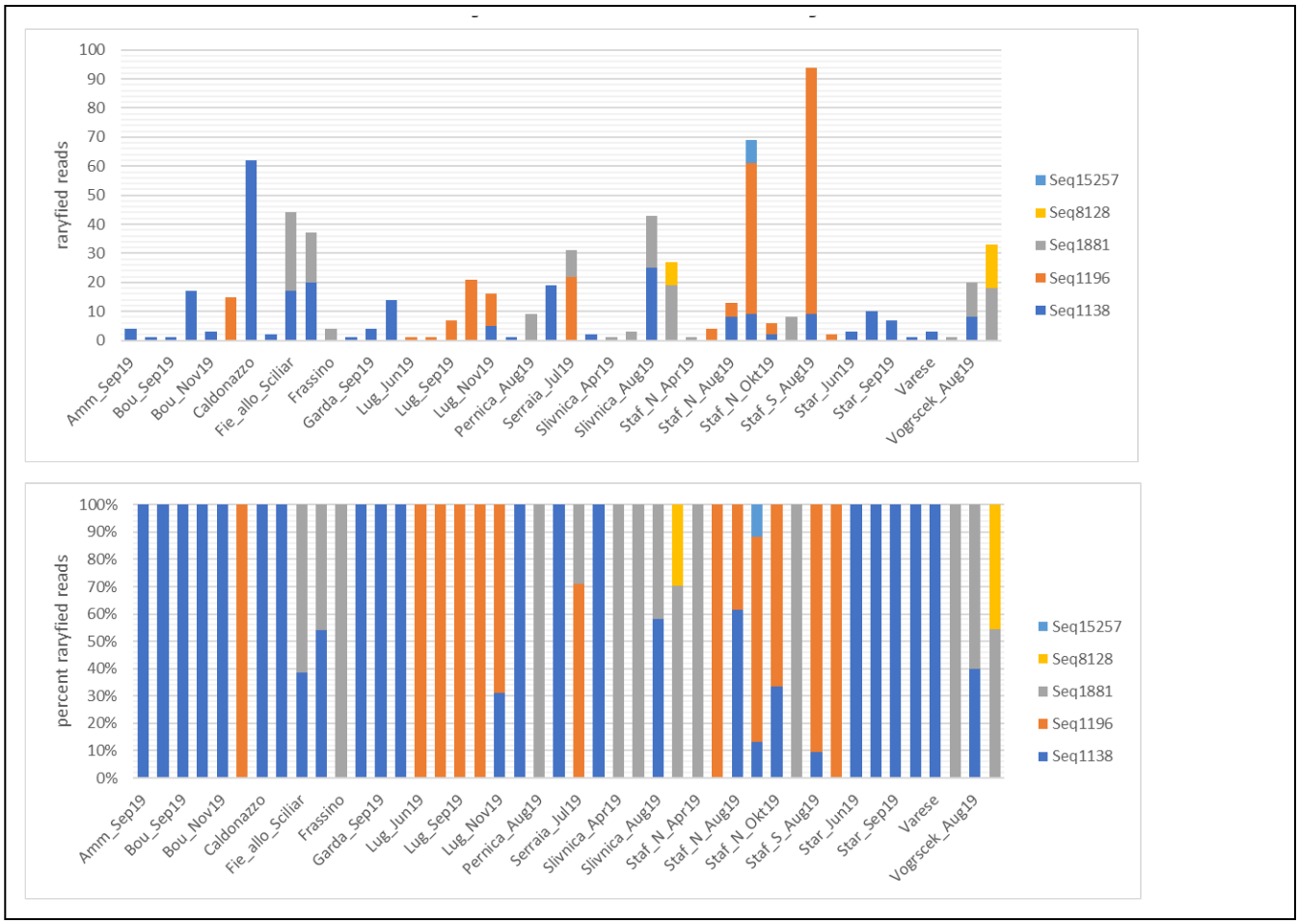


Suppl. Fig. 9. (A, B) Qualitative and (C) quantitative comparison of sequencing vs microscopy for detection of the genus *Snowella* in phytoplankton from lakes across alpine space (n = 140). Linear regression (R^2 = 0.73), 1 outlier (biovolume 1.5 mm^3/L) excluded.

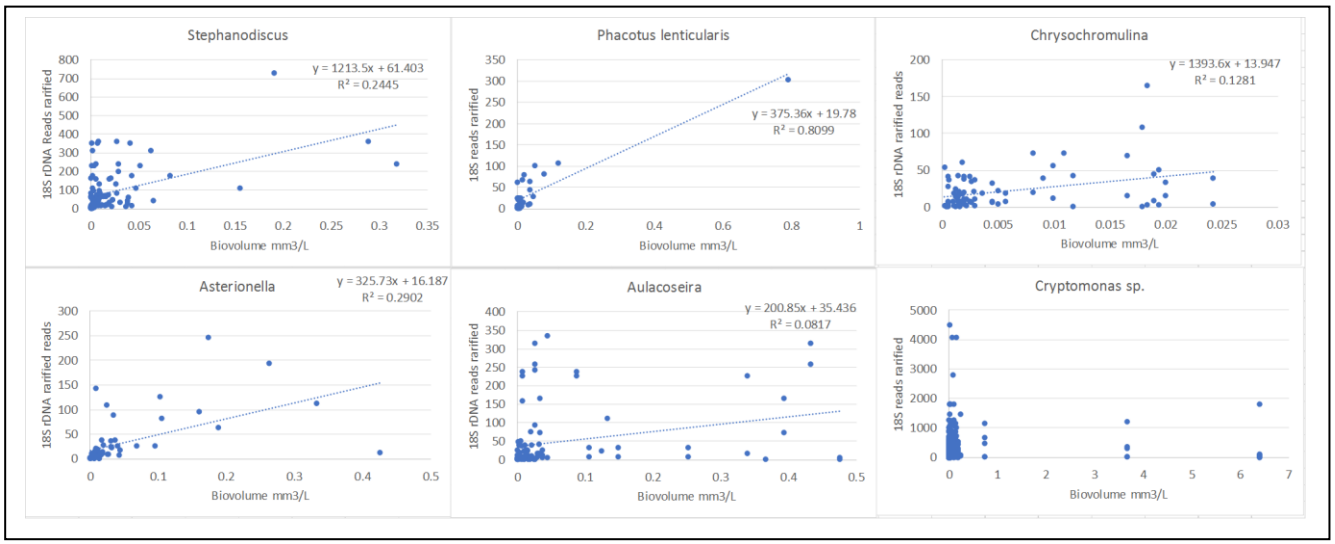
Suppl. Fig. 10. (A, B) Qualitative and (C) quantitative comparison of sequencing vs microscopy for detection of the genus *Woronichiniana* in phytoplankton from lakes across alpine space (n = 140). Linear regression (R^2 = 0.64), 2 outliers (biovolume 14.08, 21.92mm^3/L) excluded.



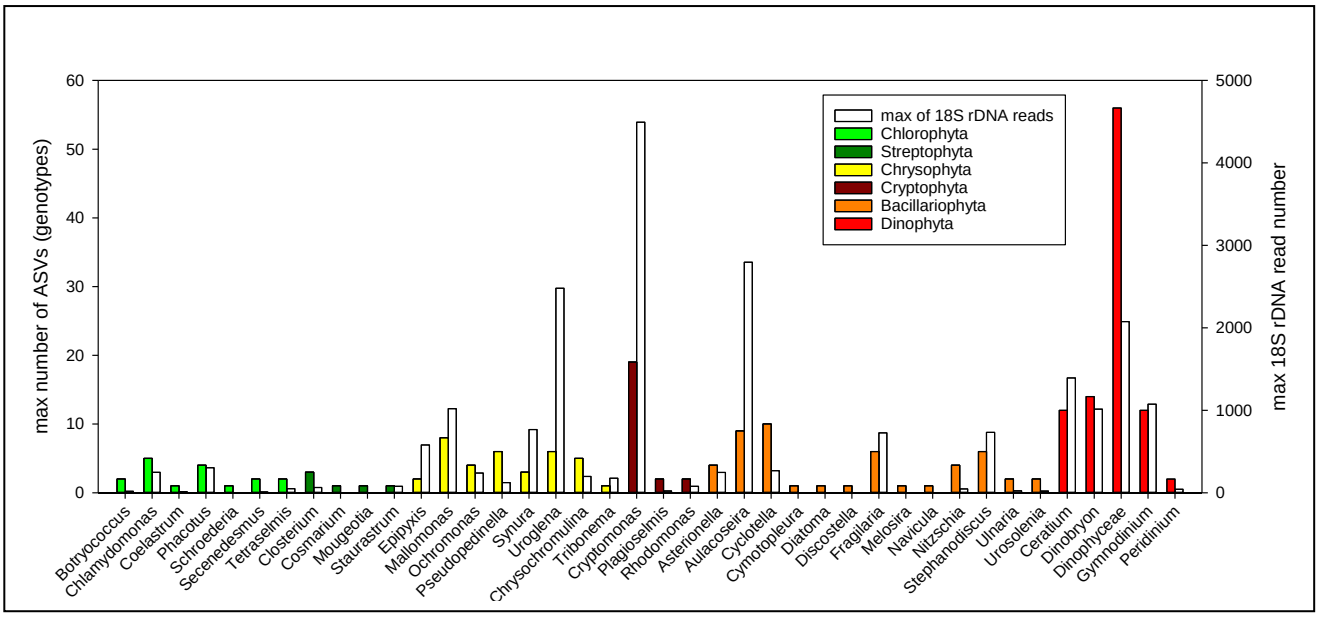
Suppl. Fig. 11. A) Absolute and (B) relative biodiversity of the genus *Aphanizomenon* from lakes across alpine space (n = 140): Seq189=*A. flos-aquae* PH218, Seq11562=A. sp, Seq13654=A. sp, Seq15960=*A. flos-aquae*, Seq18363=A. sp, Seq4407=A. sp, Seq6625=A. sp, Seq7096=*Sphaerospermopsis aphanizomenoides*, Seq7288=A. sp, Seq28649=A. sp, Seq7682=A. sp, Seq907=*A. flos-aquae* PH218, Seq9366=*A. flos-aquae*, Seq9356=A. sp, Seq8920=*C. issatschenkoi*, Seq9940=A. sp, Seq10414=A. sp



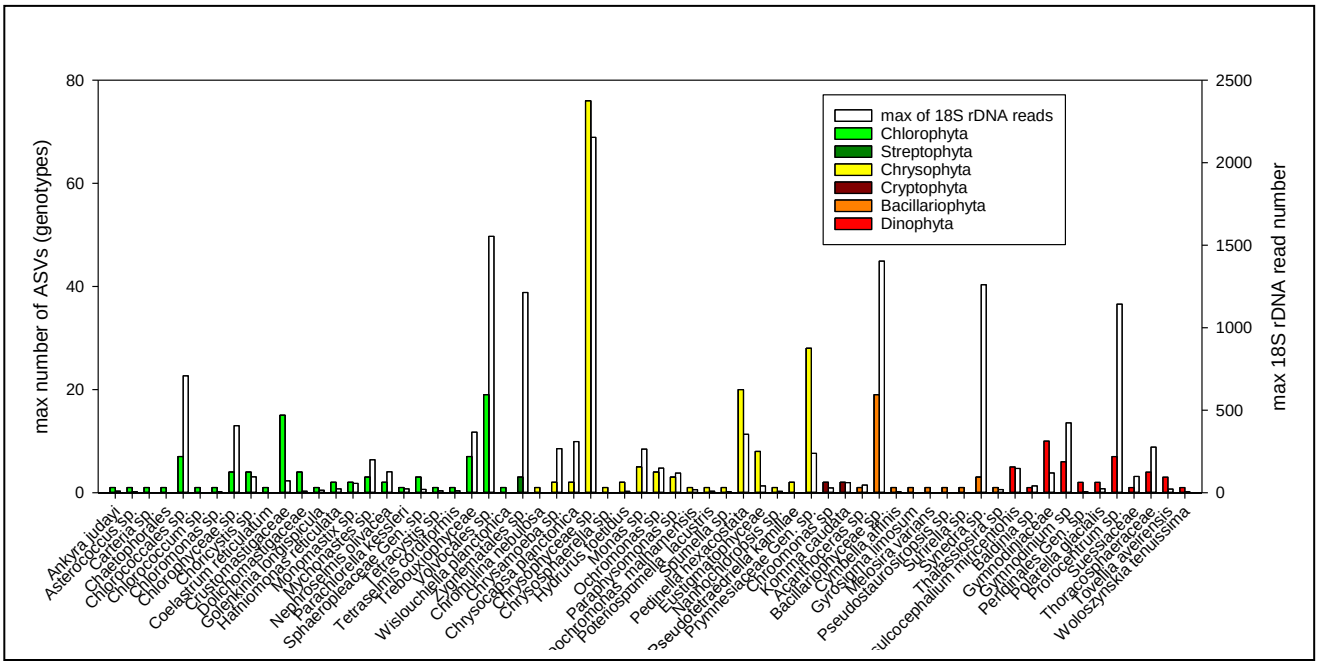
Suppl. Fig. 12. A) Absolute and (B) relative biodiversity of the genus *Microcystis* from lakes across alpine space (n = 140).



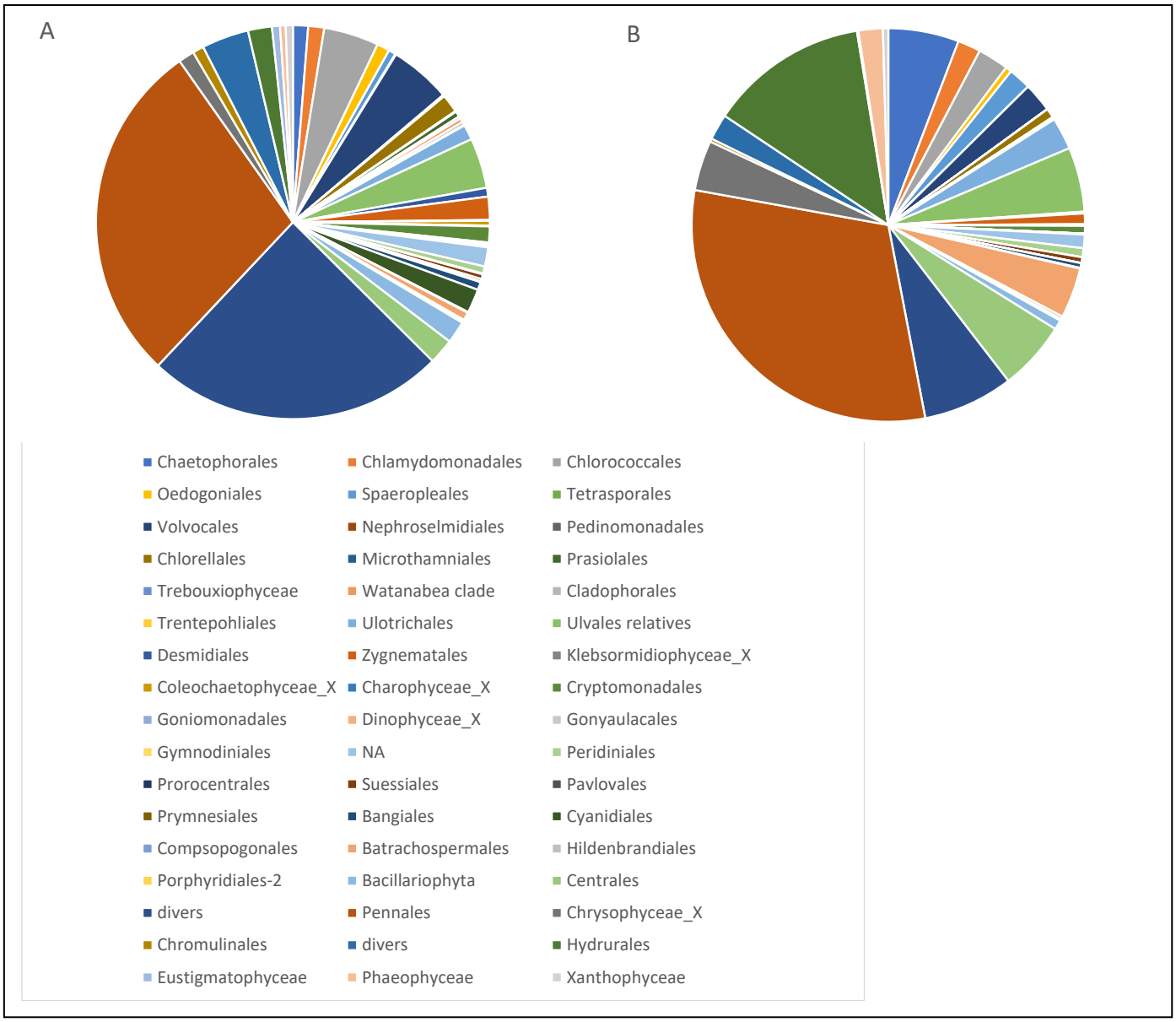
Suppl. Fig. 13. Comparing quantitative estimates for eukaryotic phytoplankton using both methods (18S rDNA sequencing read number vs biovolume from microscopical counting)



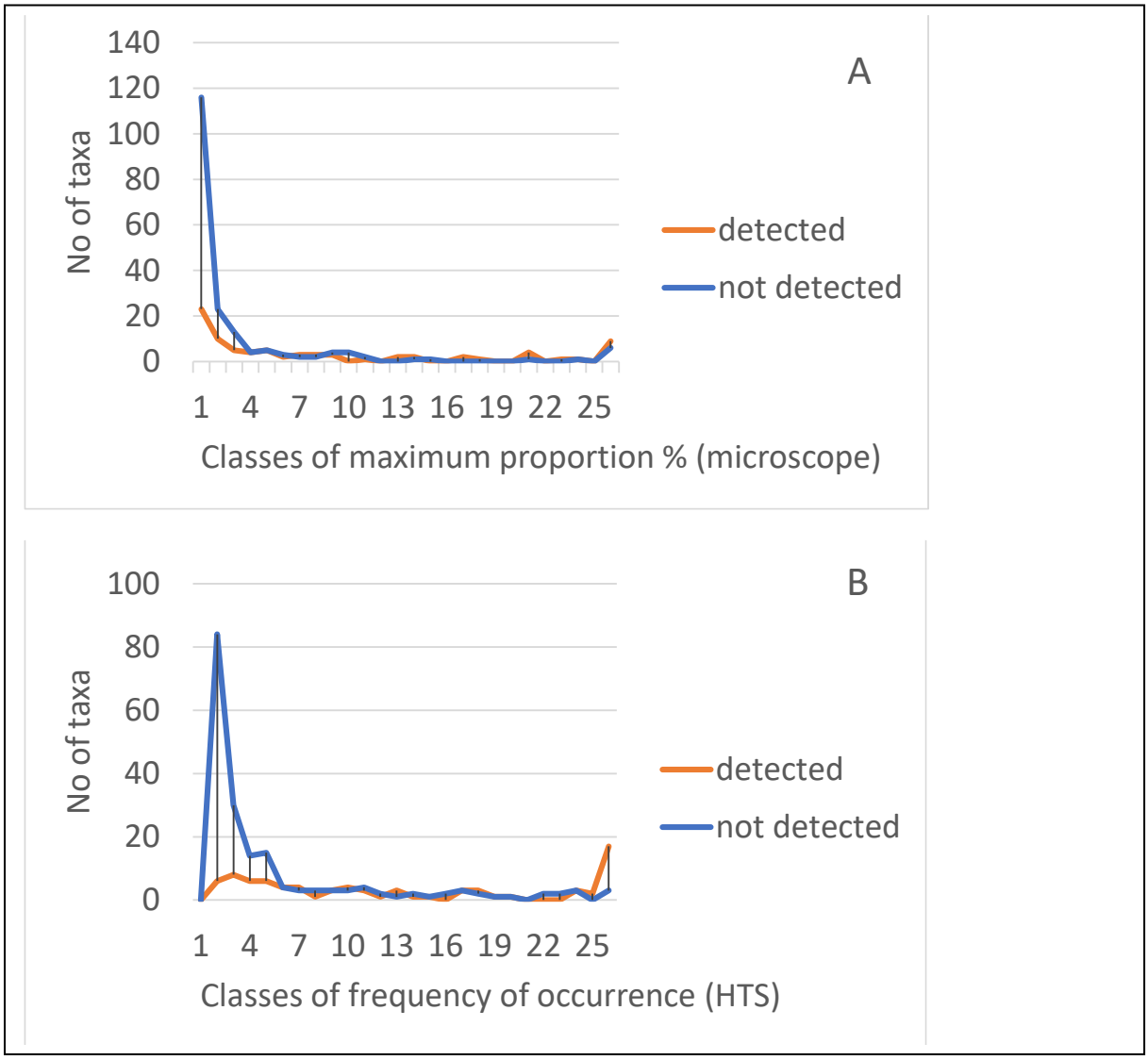
Suppl. Fig. 14: Interspecific/Intraspecific biodiversity of phytoplankton as revealed through 18S rDNA sequencing within taxa counted in the microscope. Theoretically each genotype (amplicon sequence variant) might be considered a “cryptic” species.



Suppl. Fig. 15: Interspecific/Intraspecific biodiversity of phytoplankton as revealed through 18S rDNA sequencing within taxa not counted in the microscope. Theoretically each genotype (amplicon sequence variant) might be considered a “cryptic” species.



Suppl. Fig. 16. Orders of eukaryotic algae in rivers (n = 53). A) Proportion in genotype (ASV) numbers (n= 1527), B) Proportion in maximum read numbers (n= 33149).



Suppl. Fig. 17: Number of diatom taxa recorded through HTS (rbcL) and at least partly through LM (detected) vs taxa recorded through LM counting only (not detected). (A) Comparing the number of detected taxa vs. the number of not detected taxa plotted against the maximum proportion as revealed through LM counting, (B) Number of taxa according to the frequency of occurrence in river samples (n = 53).