MICROALGAL FUNCTIONAL TRAITS

Biodiversity patterns of cyanobacterial oligotypes in lakes and rivers: results of a large-scale metabarcoding survey in the Alpine region

Nico Salmaso · **Serena Bernabei** · **Adriano Boscaini** · **Camilla Capelli** · **Leonardo Cerasino** · **Isabelle Domaizon** · **Tina Elersek** · **Claudia Greco** · **Aleksandra Krivograd Klemenčič** · **Paolo Tomassetti** · **Rainer Kurmayer**

Abstract In this work, we characterised the cyanobacterial communities in the plankton and littoral biofilm of 38 lakes and in the biofilm of 21 rivers in the Alps and surrounding subalpine regions by 16S rRNA gene metabarcoding. We found little overlap in the distribution of amplicon sequence variants (ASVs) between the three habitats and between water bodies. The differences were caused by environmental filtering acting on the selection of the most abundant ASVs and a high contribution of rare oligotypes. The differentiation of community and genotype composition from specific water bodies was explained to a significant extent by environmental variables and morphometry. The taxonomic consistency of ASVs classified under the same genus name was assessed by phylogenetic analyses performed on three representative dominant genera, namely *Cyanobium*, *Tychonema* and *Planktothrix*. The analyses revealed eco-evolutionary adaptations in lakes and rivers, including some evidence for a polyphyletic nature. Monitoring individual genotypes in relation to environmental conditions will be useful to define the ecological amplitude of these taxa. However, the persistence or ephemeral nature of some of the rarest and most unusual ASVs has remained unknown.

N. Salmaso (✉) · A. Boscaini · L. Cerasino
Research and Innovation Centre, Fondazione Edmund Mach, San Michele All'Adige, Italy
e-mail: nico.salmaso@fmach.it

N. Salmaso
NBFC, National Biodiversity Future Center, 90133 Palermo, Italy

S. Bernabei · C. Greco · P. Tomassetti
Italian Institute for Environmental Protection and Research (ISPRA), Rome, Italy

C. Capelli
Institute of Earth Sciences, University of Applied Sciences and Arts of Southern Switzerland (SUPSI), Mendrisio, Switzerland

I. Domaizon
INRAE, UMR Carrtel, Université Savoie Mont Blanc, Pole R&D ECLA, Thonon les Bains, France

T. Elersek
National Institute of Biology, Ljubljana, Slovenia

A. Krivograd Klemenčič
Slovenian Environment Agency, Ljubljana, Slovenia

R. Kurmayer
Research Department for Limnology, Universität Innsbruck, Mondsee, Austria

R. Kurmayer
Universität Innsbruck, Innrain 52, 6020 Innsbruck, Austria

Introduction

The completeness of biogeographical surveys of cyanobacteria based on morphological traits of species is severely limited by several factors. The overlap of morphological characters and the plasticity of phenotypic traits that can vary with environmental conditions are challenging elements in the identification of specimens by light microscopy (LM), especially when the dimensions do not allow taxonomic discrimination of the smallest individuals. Extensive experience and the ability to make use of identification manuals and sparse literature are required, and yet these skills do not always guarantee the comparability of taxonomic identifications by LM between laboratories (Lee et al., 2014). These difficulties are only partially mitigated by the adoption of polyphasic or genetic approaches (Komárek, 2016; Wilmotte et al., 2017), which are generally used to characterise isolated and cultured specimens (Komarek et al., 2013), single individuals or colonies (Mareš et al., 2015; Kurmayer et al., 2018; Pokorný et al., 2023), and environmental monospecific samples (Zubia et al., 2019). Since the description of any new taxa within the International Code of Nomenclature of Prokaryotes (ICNP) and the International Code of Nomenclature for algae, fungi, and plants (ICN) requires the designation of type materials, the phylogenetic analyses and taxonomic identifications are mostly based on strains, whereas many species have not yet been isolated and cultivated (Rosselló-Móra & Whitman, 2019; Zubia et al., 2019) possibly due to metabolic limitations that prevent their cultivation (Brown et al., 2015).

The emergence of culture independent techniques based on the analysis of community and environmental DNA (eDNA; Ruppert et al., 2019) by high-throughput sequencing (HTS) has opened new perspectives in the study of microbial biodiversity and ecology (Pawlowski et al., 2018; Johnson et al., 2019). Although these techniques allow the analysis of virtually all the most abundant taxa, the main metabarcoding techniques used for the determination of bacteria and cyanobacteria suffer from many drawbacks mainly due to the short length and corresponding low taxonomic resolution of the 16S rRNA gene that can be obtained after the application of bioinformatic pipelines (around 400 nucleotides maximum) and the corresponding lack of sensitivity of the reference databases for the selected DNA markers (Salmaso et al., 2022).

The recent widespread adoption of DNA read denoising approaches (Callahan et al., 2016; Nearing et al., 2018) has allowed the fine distribution of specific classified or unclassified cyanobacterial oligotypes across environmental gradients to be studied using exact 16S rRNA oligotypes (exact amplicon sequence variants, ASVs) (Berry et al., 2017; García-García et al., 2019; Salmaso, 2019). However, despite these improvements, the biological significance of the extent of oligotype variability and the inclusion and significance of the range of ASVs classified within the same genus remain largely unexplored.

In the last decade, HTS approaches have been increasingly used for the taxonomic classification of cyanobacteria in a variety of aquatic water bodies (Pushkareva et al., 2015; Lopes dos Santos et al., 2022; Pawlowski et al., 2022; Sandzewicz et al., 2023). Most of the investigations were carried out in selected habitats, contributing to disentangle the taxonomic nature, structure, and temporal dynamics of cyanobacteria (Guellati et al., 2017; Scherer et al., 2017; Nwosu et al., 2021). Due to the difficulty of organising coordinated initiatives with harmonised field and laboratory approaches, large-scale spatial investigations of cyanobacteria across large regions or different countries have rarely been attempted, and in such cases have focused on unique habitats (MacKeigan et al., 2022) or integrative variables, such as cyanotoxins (Mantzouki et al., 2018) and photosynthetic pigments (Donis et al., 2021). The lack of studies over large areas and across ecotones prevented the assessment of the nature, distribution and overlap of cyanobacterial species along spatial gradients within and between different habitats.

In this context, within the framework of the Interreg Alpine Space project Eco-AlpsWater, a large-scale survey was coordinated in 2019 in the Alps and surrounding subalpine regions to characterize the microbial communities (bacteria, cyanobacteria and protists) in the plankton and biofilm of 37 lakes and in the biofilm of 22 rivers (Domaizon et al., 2021; Kurmayer et al., 2021; Salmaso et al., 2022). The

study was based on the adoption of common protocols and involved 12 partners from 6 countries. In this paper, we will specifically focus on the biogeographic distribution of cyanobacterial taxa as inferred from 16S rRNA gene sequences obtained from eDNA extracted from plankton and biofilm samples. Specific aims of this study are: (i) to assess the diversity and overlap of oligotypes in different aquatic habitats and the role of environmental variables in filtering ASVs; (ii) to assess the internal variability of ASVs classified under the same genus name.

Materials and methods

The methods used in the field and in the laboratory, and the repositories (ENA and Zenodo) where the raw molecular data were deposited have already been described (Salmaso et al., 2022). In the following, we will specifically highlight the methodological points that are most relevant or new for this work.

Study sites and sampling

Investigations were carried out in 2019 in lakes and rivers throughout the Alpine region (Supplementary Table 1 and Supplementary Fig. 1); in lakes Mondsee, Aigueblette and Bourget, and in rivers Arve and Drome, the biofilm samples were collected in 2018. Samples were collected from the pelagic zone of the lakes (plankton; lake_PL) and from the biofilm in the littoral zone of the lakes (lake_BFM) as well as from the biofilm in the rivers (river_BFM). In the River Lech, Germany, the biofilm sampling station was located at the Mandichosee, a small artificial lake (1.6 km^2) created after the river was dammed in 1978; in this work, this station was therefore classified as a lake. At Staffelsee, in Germany, plankton samples were collected in the northern and southern basins. Overall, the altitude, surface area, and maximum depth of the 38 lakes analysed ranged between 18 m a.s.l. and 2125 m a.s.l., < 0.01 km^2 and 582 km^2, and 1.3 m and 410 m, respectively. The length of the 21 rivers included in this work ranged between < 1 and 945 km.

Plankton samples (lake_PL) were collected once a month throughout the year or between April and October at 8 key lakes (8 to 13 samples: lakes Bourget, Mondsee, Bled, Garda, Lugano, Ammersee, Starnberger See, Staffelsee) and 1 to 4 times generally during the warmer months at a further 27 additional sites, for a total of 144 samples (Supplementary Table 1). Sampling was performed at the deepest point of the lakes, depth-integrating the epilimnetic or euphotic zones. Water samples were filtered using Sterivex cartridges (0.22 µm, Hydrophilic PVDF Durapore membrane, Sigma Aldrich) within 12 h after sampling and then immediately frozen at − 20 °C.

Biofilm was mainly sampled between July and October in lakes (from 1 to 27 stations, for a total of 125 samples) and between February and October in rivers (from 1 to 6 stations, 45 samples) (Supplementary Table 1). Biofilm was collected by brushing the surface of at least 5 stones identified in the sampling areas as described in Rimet et al. (2020, 2021) (Supplementary Fig. 2). Other substrates were not sampled because the collection from large cobbles and small boulders ensured that diatom subsamples were also collected from both lakes and rivers for analysis according to the Water Framework Directive (Vasselon et al., 2017; Salmaso et al., 2022). Therefore, the periphyton communities considered in this paper are exclusively epilithic cyanobacteria. In the case of more than one sampling station, each lake and river was sampled on the same day/week or, less frequently, within 1 or 2 months. Approximately 10 mL subsamples were collected in 50 mL sterilized Falcon tubes filled with ca. 40 mL of absolute ethanol. Samples were stored at 4 °C in the dark.

Environmental data

Complete measurements of environmental variables were only made in the pelagic areas of lakes. All analyses were based on standard methods fully tested in each laboratory (Wetzel & Likens, 2000; APHA, AWWA & WEF, 2018). In this work, we only considered a set of variables almost fully represented in all lakes, with the exclusion of lakes Anterne and Brevent. In the field, water temperature (Temp) was measured along the water column with probes, and transparency (Secchi) was estimated using a Secchi disk. Algal nutrients were determined by colorimetric methods (total phosphorus, TP; soluble reactive phosphorus, SRP; nitrate nitrogen, NO_3_N; ammonium nitrogen, NH_4_N) and/or ion chromatography (NO_3_N, NH_4_N). Dissolved oxygen (DO)

concentrations were measured with calibrated probes or by titration (Winkler method); percentage DO saturation (DOp) was estimated from dissolved O_2, water temperature and lake elevation (Mortimer, 1981). Chlorophyll-a (Chl_a) was determined spectrophotometrically from acetonic extracts or using hot ethanol (ISO, 1992; APHA, AWWA & WEF, 2018). For the unique samples collected in lakes Maggiore and Mantova Superiore, water temperature and dissolved oxygen, and water temperature, were obtained from technical reports, namely CNR IRSA, sede di Verbania (2020) and Marchesi et al. (2020), respectively. Missing Secchi disk values in lakes Staffelsee (July) and Bourget (October), and Chl_a in Starnberger See (April) were interpolated by calculating the average of the previous and subsequent sampling dates.

DNA extraction, library preparation and sequencing

DNA extraction from Sterivex filters and biofilm was performed with Mo Bio PowerWater® DNA Isolation Kit (MO BIO Laboratories, a QIAGEN Company, USA) (Vautier et al., 2021), and NucleoSpin® Soil kit (Macherey–Nagel) (Vautier et al., 2020), respectively. PCR amplification of each DNA sample was carried out by targeting ~ 460 base fragments of the 16S rRNA gene variable regions V3–V4 using primers 341F (5′CCTACGGGNGGCWGCAG 3′) and 805Rmod (5′GACTACNVGGGTWTCTAATCC 3′). All barcoded libraries were pooled in equimolar concentrations by Real-Time qPCR and checked on a TapeStation 2200 platform (Agilent Technologies, Santa Clara, CA, USA). The library thus obtained was sequenced on an Illumina® MiSeq (PE300) platform (MiSeq Control Software 2.6.2.1 and Real-Time Analysis software 1.18.54) (Salmaso et al., 2022).

Bioinformatic analyses

After primers removal using Cutadapt 3.1 (Martin, 2011), raw sequences were analysed using DADA2 1.18 (Callahan et al., 2016) in R 4.0.3 (R Core Team, 2020) following the pipeline described in Salmaso et al. (2021). Taxonomic assignment of ASVs was performed in DADA2 using the RDP naive Bayesian classifier ("Wang classifier") with a 95% minimum bootstrap confidence threshold, and the reference taxonomic database SILVA 138 (Quast et al., 2013). After discarding singletons and doubletons (single

ASVs present in the whole dataset with one or two sequences, respectively), sequences were further checked for the presence of chimeras and removed if necessary using the uchime2_ref command in usearch v. 11 in the "specific" mode (Edgar, 2016) and the SILVA 138 reference database. The phylum Cyanobacteria was then separated from the remaining bacteria, resulting in 2620 ASVs.

Statistical data analysis

The taxonomic, abundance and environmental data tables, DNA sequences and associated phylogenetic tree were merged into a single dataset and analysed in R using the package phyloseq 1.40.0 (McMurdie & Holmes, 2013). The phylogenetic tree associated with the dataset was computed by Maximum Likelihood (ML) using RAxML 8.2.10 (Stamatakis, 2014) and the GTRCAT model applied to the whole set of cyanobacterial sequences (2620 ASVs) aligned using Muscle v. 5.1 and the super5 command (Edgar, 2022).

Depending on the type of analysis, and as specified in the description of results, ASVs were evaluated without or after rarefaction without replacement to a minimum number of sequences per dataset. Rarefaction was used specifically when the comparison between samples required the use of absolute abundances (i.e., number of reads). In these cases, and considering that the absolute abundances obtained in the planktic and biofilm habitats were not comparable, rarefaction was applied individually to the three habitats (lake_PL, lake_BFM and river_BFM). Differences in the number of ASVs between habitats were estimated using the Kruskal–Wallis rank sum test (KW), whereas pairwise comparisons were performed using the Wilcoxon rank sum test, with the Benjamini and Hochberg correction (stats package in R).

Comparison of communities in the three habitats was assessed by ordination of samples using Principal Coordinates Analysis (PCoA, aka Metric Multidimensional Scaling) of a distance matrix computed using the unweighted UniFrac distance. UniFrac measures the distance between communities based on the lineages they contain i.e., by exploiting the different degrees of similarity between sequences (Lozupone & Knight, 2005; Lozupone et al., 2011). If two habitats share similar ecological characteristics and species, most nodes in the phylogenetic tree will

have descendants from both communities. Under the opposite condition, most nodes will have exclusive descendants, with much of the branch length in the tree not shared between species from the two habitats (Lozupone & Knight, 2005). The distribution of the relative contribution of the cyanobacterial families in the three habitats was analysed by correspondence analysis (CA) (Greenacre & Hastie, 1987).

As the number of samples collected in each lake and river was different, in order to have a comparable number of samples in the water bodies belonging to the three different habitats, subsets of samples were extracted from the main database for a selected number of analyses (e.g., ordinations and Mantel test; see Results and Supplementary Table 1). The first set of samples (lake_PL, dataset 1) included the 8 key lakes sampled monthly in the pelagic zone from April to October, i.e., Ammersee, Bled, Bourget, Garda, Lugano, Mondsee, Starnberger See and Staffelsee (with 2 stations, Staffelsee_Nord and Staffelsee_Sud). Additional analyses in lakes were carried out on a more extended spatial set of data (dataset 2), which included one sample for each water body collected during the summer months (mostly between July and August; 33 stations, including Staffelsee_Nord and Staffelsee_Sud); this dataset did not include lakes Fimon, Anterne, and Brevent, which were excluded from the analyses due to a low number of cyanobacterial reads. A third dataset (lake_BFM, dataset 3) included 9 lakes with 10–12 (6 in Staffelsee) biofilm samples each (Aiguebelette, Ammersee, Bled, Bourget, Garda, Lugano, Mondsee, Staffelsee, and Starnberger See). Dataset 4 (river_BFM) included 6 rivers with a number of biofilm samples ranging from 3 to 6 (Arve, Drome, Salzach, Soča-Isonzo, Steyr, and Wertach). These datasets were analysed by calculating a PCoA on the Bray & Curtis (B&C) dissimilarity matrices after rarefying the samples to a uniform abundance and double square root transformation of ASVs abundances (Legendre & Legendre, 1998), and/or calculating a Canonical Correspondence Analysis (CCA) after log-transformation of the environmental data and ASVs abundance data transformed as in PCoA, respectively.

UniFrac, B&C, and CCA (scaling 2) were calculated using the vegan package in R (Oksanen et al., 2020); CA was calculated using vegan and FactoMineR (Lê et al., 2008), whereas PCoA was calculated using the function cmdscale in the stats package

(Venables & Ripley, 2002). Differences in taxonomic composition and abundance between groups of samples were tested using PERMANOVA with 9999 permutations computed using the adonis2 function in the vegan package applied to the same UniFrac and B&C matrices used in PCoA; pairwise PERMANOVA comparisons were performed using the pairwiseAdonis package.

Mantel test

Correlations between environmental variables and cyanobacterial ASVs in lakes were calculated by computing Mantel tests (Legendre & Legendre, 1998) using the function mantel in the vegan package (Oksanen et al., 2020). Environmental distance matrices were calculated using a set of log-transformed and standardized basic limnological (Temp, Secchi, DOp, NO3_N, NH4_N, TP, and Chl_a) and physiographic (altitude, surface, maximum depth, catchment area) variables. The cyanobacterial dissimilarity matrix was calculated using the same methods as for the PCoA. The significance of the statistic was evaluated by 999 permutations of the rows and columns of the first (ASVs) dissimilarity matrix.

Blast and phylogenetic analyses

The taxonomic assignments of the most abundant genera in the cyanobacterial families were further verified by submitting the sequences to megablast analyses.

DADA2 distinguishes sequence variants differing by as little as one nucleotide by inferring the biological sequences before introducing errors in the PCR amplification and sequencing steps (Callahan et al., 2017), thus allowing the number and distribution of ASVs (oligotypes) belonging to individual species or genera to be examined. The phylogenetic relatedness of ASVs attributed to individual genera and species by the Wang classifier in DADA2 have been explored in three representative dominant species, i.e., the genus with the highest number of ASVs (*Cyanobium*, 169), and two toxigenic genera (*Tychonema* and *Planktothrix*, 38 and 18 ASVs, respectively). The mutual phylogenetic position of ASVs in the three genera was examined by including in the analyses a selection of homologous sequences imported from GenBank, and using *Gloeobacter violaceus* Rippka,

J.B. Waterbury & Cohen-Bazire, 1974 (strain PCC 7421) as outgroup. Sequences were aligned using Muscle v. 5.1 (Edgar, 2022), resulting in alignments between 385 and 419 nucleotides in length after trimming to the shortest sequence. Maximum likelihood trees were calculated using PhyML 3.1 (Guindon et al., 2010) with models obtained using the function phymltest in the R package ape (Paradis et al., 2004); GTR + G + I was found to be the best fitting evolutionary model for all trees. Branch support was estimated by the SH-like branch support (Anisimova et al., 2011). The Newick rooted trees obtained with PhyML were annotated with the R package ggtree (Yu, 2020).

The presence of a phylogenetic signal in the ASVs of the three selected genera was assessed by testing the distribution of abundances for the three habitats, calculating Pagel's λ and the Abouheif-Moran test (R packages phytools and adephylo); the alignments and phylogenetic trees used in the tests were calculated as described above.

Results

Taxonomic characterization in the pelagic and biofilm habitats

The number of ASVs was much higher in the biofilm samples than in the plankton samples (KW and Wilcoxon tests, $P < 0.001$). The median values of the number of ASVs in lake_BFM, river_BFM, and lake_PL samples were 141, 42, and 14, respectively (Fig. 1A). Moreover, in the biofilm samples, the number of ASVs showed a wider range of variations (11–325) compared to the planktic samples (1–60). These differences were paralleled by the Shannon diversity values (Fig. 1B). While the number of ASVs shared in the three habitats was 0.7%, the number of ASVs shared between any of the two habitats was between 0.04% and 7.8% (Fig. 1C). The high fraction of exclusive ASVs in the three habitats was due to the presence of many sequences occurring at low frequency. The fraction of ASVs occurring once, twice or three times in the whole dataset was 36%, 14%, and 8%, respectively. The frequency of occurrence of ASVs in the three habitats followed a similar pattern, but with a wider range in the number of ASVs detected in one or two lake BFM samples

(Supplementary Fig. 3). After removing the ASVs occurring in only one or two samples, the number of exclusive ASVs in the three habitats decreased only slightly. This small decrease was caused by a persistent number of exclusive ASVs in lake_BFM.

The differences due to the exclusive presence of ASVs in specific habitats were paralleled by significant compositional differences between the three habitat types (Fig. 2). In the PCoA ordination, which was based solely on phylogenetic distance between groups of taxa, the samples showed a clear and significant separation in the three habitats (PERMANOVA, $P < 0.001$) (Fig. 2A). Some exceptions were due to a few samples that were representative of ecotones, such as the lake_BFM samples collected near the mouths of the Cassarate and Cuccio rivers, which are located around the river_BFM cluster, and the lake_BFM sample collected at Mandichosee, which was located near the lake_PL cluster.

The most abundant families in lake_PL samples were Phormidiaceae and Cyanobiaceae (>10% on the total sum of lake_PL samples). Besides Cyanobiaceae, lake_BFM samples showed high relative contributions especially from Nostocaceae (>10%). The river_BFM samples showed a higher (>10%) relative abundance of Leptolyngbyaceae, Xenococcaceae, and an Unknown Family (Oxyphotobacteria Incertae Sedis), which included periphytic filamentous species (Supplementary Fig. 4). The biofilm samples showed a significant contribution (around 20% of total abundances) of ASVs without any classification at the family level; this number decreased to less than 2% in the pelagic samples. Although always present with a lower fraction in the three habitats (average < 0.6%), the non-photosynthetic cyanobacteria (classes Vampirivibrionia and Sericytochromatia) were identified with the orders Caenarcaniphilales, Gastranaerophilales, Obscuribacterales, and Vampirovibrionales (297 ASVs). In this group, the only classified species was *Vampirovibrio chlorellavorus* (ex Gromov and Mamkayeva 1972) Gromov and Mamkayeva 1980, which was identified in Lake Aigueblette, whereas the other taxa were previously identified from metagenome-assembled genomes (MAGs) (Soo et al., 2017). The strict association of most family groups with specific habitats is resumed in the correspondence analysis of Fig. 2B; to better discriminate groups in the CA plane, the analysis included only the most abundant families (>0.2% of total abundance in each habitat,
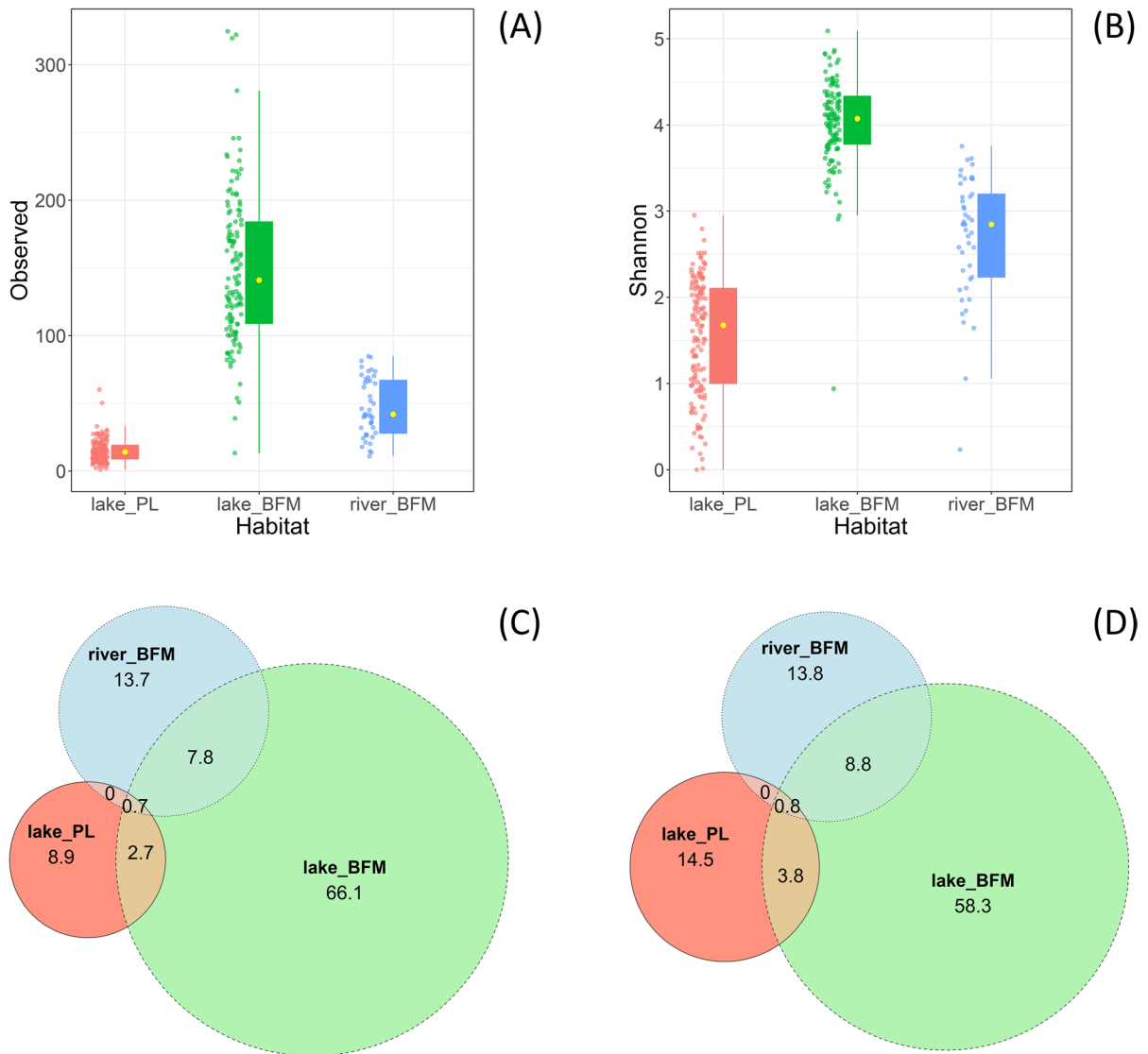
Fig. 1 A Observed number of ASVs and **B** Shannon diversity in the lake plankton (lake_PL), lake biofilm (lake_BFM) and river biofilm (river_BFM) habitats; boxplots show median and hinges as 25th and 75th quartiles, while whiskers extend from the hinges to the largest value no further than $1.5 \times$ interquartile range. Euler diagrams, showing the percentage of shared (C) ASVs ($n = 2620$) and (D) genera ($n = 1187$) in the three habitats

thus excluding 5 families). The different distribution of families in the three habitats was paralleled by a different distribution of prevalent life-habits. The lake_PL samples showed a prevalence of pelagic filamentous and coccoid taxa. In addition to the coccoid taxa in the lake_BFM samples, all lake and river biofilm samples shared common characteristics with a prevalence of periphytic colonies, periphytic filaments, and periphytic heterocytous filaments.

Main genera in the pelagic and biofilm habitats

Overall, the number of genera identified by the Wang classifier in DADA2 was 45% of the total number of ASVs. The fraction of the number of classified genera was higher in the lake_PL samples (70%) than in the lake_BFM and river_BFM samples (42% and 48%, respectively). Nevertheless, considering the abundances, the fraction of ASVs classified at genus
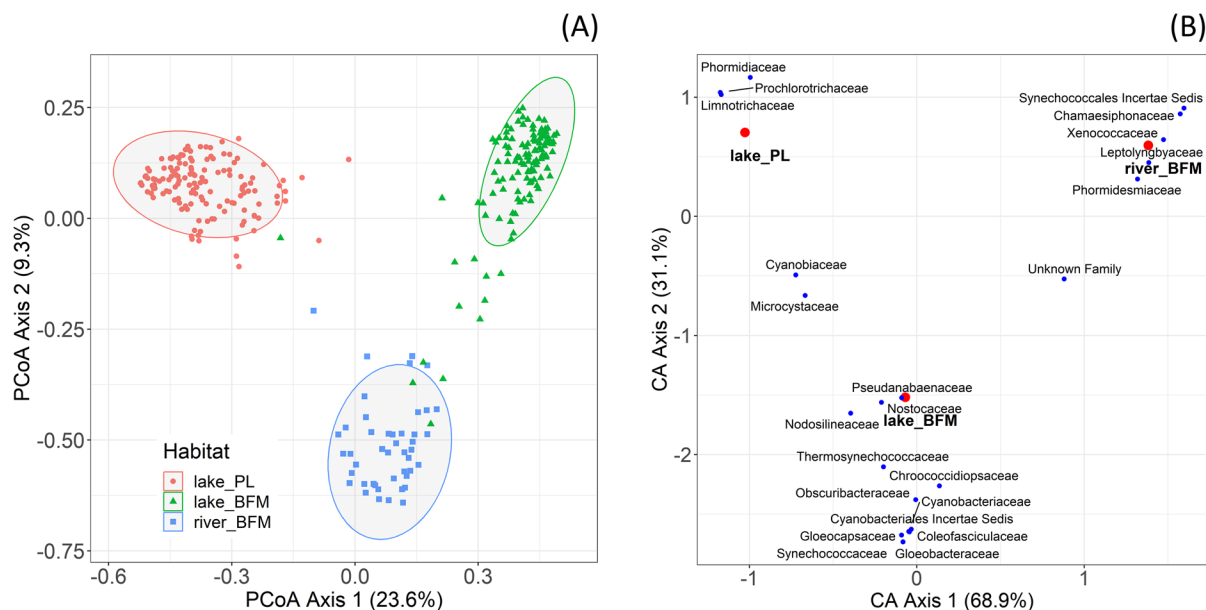
(A)

(B)

Fig. 2 A Principal coordinate analysis (PCoA) of planktic and biofilm samples in the three habitats, calculated using the unweighted UniFrac distance. **B** Correspondence analyses using the mean percentages of cyanobacterial families in the three habitats; the analysis was performed including the families with mean abundances > 0.2% (25 families out of 28) and excluding the unclassified (NA) families

level increased to 96% and 56% in the lake_PL and lake_BFM samples, respectively, while decreasing to 39% in the river_BFM samples. These considerations are confirmed by the tendency for a higher proportion of ASVs to be classified at genus level as their abundance increases (Supplementary Fig. 5); for sequences detected with abundances < 20, the proportion of ASVs classified at genus level was almost always less than 50%. By further extending the analysis to singletons and doubletons only (previously removed and therefore not included in this dataset), the proportion of ASVs classified at genus level dropped to 15% of the total number of ASVs.

Based on the sequences classified at the genus level, the fraction of shared ASVs between the three habitats did not change appreciably (Fig. 1D).

The distribution of the most abundant genera and species contributed to further distinguish the cyanobacterial assemblages in the three habitats (Supplementary Table 2). For each family, the table lists the most abundant ASVs, i.e., those with a relative contribution > 2% on the total abundance for each single habitat. The taxonomic classifications at genus level obtained from SILVA138 were in most cases (67%) confirmed by those obtained from the blast analyses.

Nevertheless, in this group of genera, in the 32% of cases the blast best hits provided more than one genus name, as in the case of *Limnothrix* sp. (with 100% blast best hits: *Limnothrix*, *Anagnostidinema*, *Jaaginema*, and *Planktothrix*) or *Tychonema* CCAP 1459-11B (with 100% blast best hits: *Tychonema*, *Microcoleus*, and *Phormidium*). Among the 33% of taxa classified with different genus names by SILVA and BLAST, most of the identifications by SILVA were however included in the classifications obtained with lower percent identities by the blast analyses. Conversely, a few genera classified by SILVA were not included in the first 50 best hits by blast, namely *Annamia* sp., *Merismopedia* AICB1015, *Acrophormium* PCC-7375, and *Aphanizomenon* NIES81 (Supplementary Table 2). The sequences of the latter taxon were the same as those determined from individuals isolated from the largest southern perialpine lakes and identified as *Dolichospermum lemmermannii* (Richter) P.Wacklin, L.Hoffmann & J.Komárek 2009 (Salmaso et al., 2015b, 2015a; Capelli et al., 2017).

The genera listed in Supplementary Table 2 were those representing the most common ASVs. In most cases (78%), individual genera were represented by

more than one ASV differing by one or more nucleotides (Supplementary Fig. 6). After normalisation of the data by log-transformation, the number of ASVs in each genus was strictly related to the corresponding total abundances ($r^2 = 0.66$, $P < 0.001$; Figure not shown). Considering the genera represented by more than one ASV, the percentage similarity between ASVs calculated for individual genera ranged from 91.5% to 99.8%, with the 25th and 75th quantiles being 96.3% and 98.8%, respectively. These differences were due to oligotypes belonging to different or the same species (strains) within a single genus.

Phylogenetic characterization of ASVs

To clarify the extent of genetic divergence of ASVs within each genus, we examined the phylogenetic relationships of the oligotypes classified by SILVA 138 in three representative genera, namely *Cyanobium* PCC-6307, *Tychonema* CCAP 1459-11B, and *Planktothrix* NIVA-CYA 15.

The first genus had the highest number of ASVs in the dataset (169) (Supplementary Fig. 6). The sequence similarity between the ASVs classified within *Cyanobium* PCC-6307 ranged between 93.1% and 99.8%, with the 25th and 75th quantiles being 96.3% and 98.0%, respectively. The most abundant oligotypes of *Cyanobium*, with a relative contribution of > 1% to the total abundance in each habitat (35 ASVs), were included in the phylogenetic analysis with a selection of homologous species retrieved from GenBank (Fig. 3). All *Cyanobium* ASVs were intermingled with several other sequences from GenBank classified under the genera *Cyanobium* and

Fig. 3 Maximum likelihood rooted topology of ASVs classified in the genus *Cyanobium* by the naïve Bayesian classifier and the taxonomic reference database SILVA 138 based on the alignment of 16S rRNA gene fragments (385 bp); the tree is rooted by *Gloeobacter violaceous*. The tips of the tree have been annotated with different colours to indicate the prevalence of ASVs in lake plankton and biofilm (> 0.5% of the total habitat, or exclusively present in one habitat). The annotation does not include *Cyanobium* ASVs detected in rivers, as they were always very rare (total abundances < 0.2% on the whole dataset). The size of the symbols is scaled according to the total abundance in the dataset. Numbers at nodes indicate SH-like branch supports
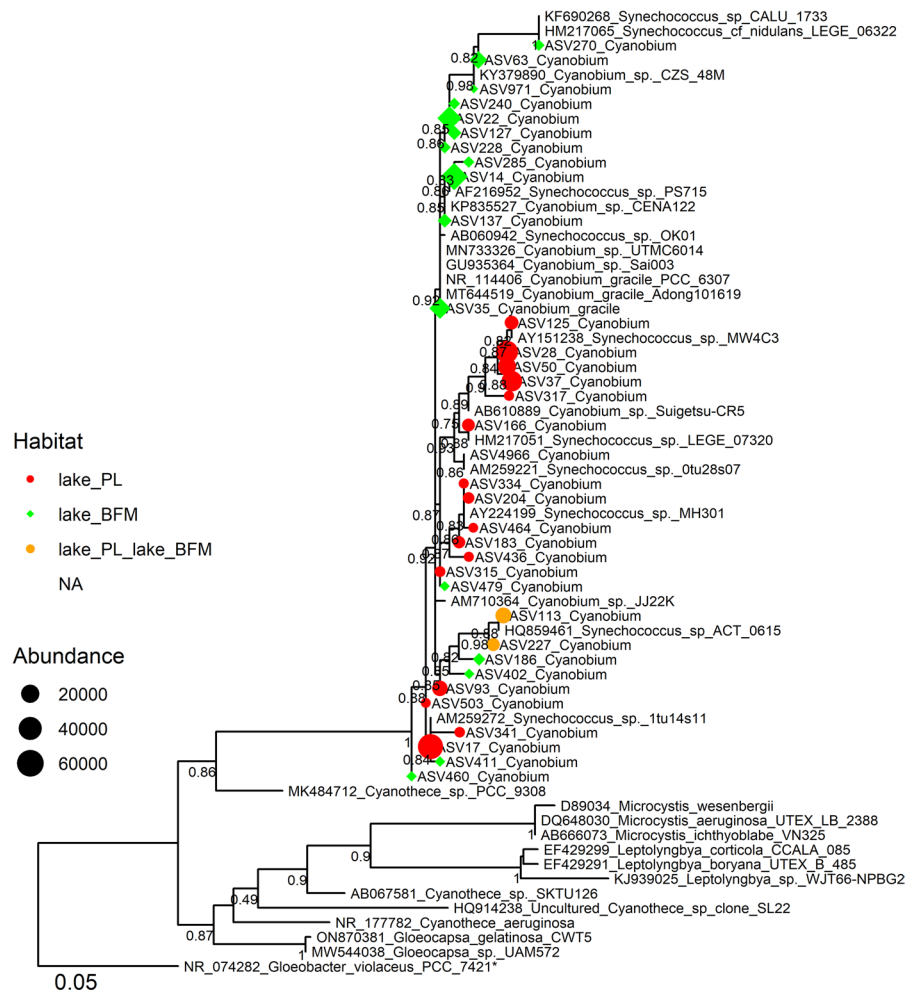
Synechococcus, whereas all other genera from Gen-Bank were part of other distinct branches. The tips of the ASVs were annotated with different colours to indicate the exclusive or almost exclusive prevalence ($> 0.5\%$ of the total habitat) of the individual sequences in each habitat. Most *Cyanobium* ASVs were prevalent almost exclusively in the plankton and biofilm of lakes, with a very small number (7) of low abundance sequences identified in the biofilm of rivers. Furthermore, the position of the ASVs in the phylogenetic tree suggested the existence of two main clusters grouping the lake_PL and lake_BFM ASVs. The PCoA analysis computed on the whole set of *Cyanobium* ASVs showed the existence of two well distinct large groups of lake_PL and lake_BFM samples, with the few river_BFM samples dispersed between these two groups (Supplementary Fig. 7A; PERMANOVA, $P < 0.001$). The habitat phylogenetic structure was further confirmed by the Pagel's and Abouheif-Moran tests calculated for the lake_PL and

lake_BFM abundances ($P < 0.001$) on the tree built using all 119 *Cyanobium* ASVs (Figure not shown).

The genus *Tychonema* was represented by 38 ASVs, with sequence similarities between 94.2% and 99.8%, and with 25th and 75th quantiles of 96.2% and 98.8%. All 38 ASVs classified under this genus were included in the phylogenetic analysis (Fig. 4). ASVs showed little consistent habitat specific distribution in three major clades. In the first clade, ASVs were mainly represented in the biofilm of rivers, or rivers and lakes. Excluding a taxon provisionally attributed to cf. *Tychonema*, in this cluster the taxa from GenBank were represented by *Microcoleus* and *Phormidium* species. Together with two other species from GenBank (*Phormidium uncinatum* Gomont and *Microcoleus vaginatus* Gomont), five of the ASVs in this group shared a common 11-nucleotide insertion (5'-GTTGTGAAAGC-3'), whereas in the sixth ASV (10,916) the insertion had two different nucleotides (5'-GTTACGAAAGT-3'). In the second main

Fig. 4 Maximum likelihood rooted topology of ASVs classified in the genus *Tychonema* by the naïve Bayesian classifier and SILVA 138 taxonomic reference database based on the alignment of 16S rRNA gene fragments (419 bp); the tree is rooted by *Gloeobacter violaceus*. Other features as in Fig. 3
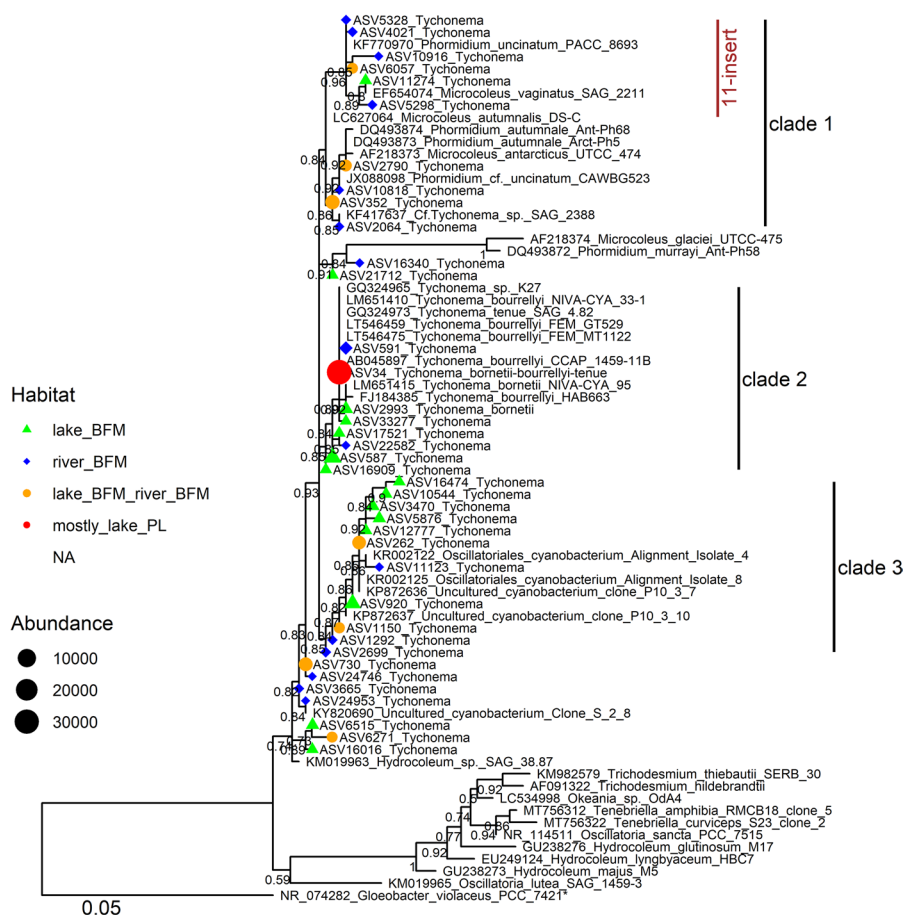
clade, all ASVs were detected mainly in lake biofilm and secondarily in river biofilm and lake plankton; these ASVs clustered together with other *Tychonema* species from the GenBank that are typical for benthic/periphytic (*Tychonema bornetii* (Zukal) Anagnostidis & Komárek and *Tychonema tenue* (Skuja) K.Anagnostidis & J.Komárek) and planktic environments (*Tychonema bourrellyi* (J.W.G.Lund) Anagnostidis & Komárek) (Fig. 4). In the case of ASV 34, the individuals were mainly recorded in the pelagic habitats, but with a measurable presence (about 6% on the whole total) also in the benthic samples. The filaments detected in the lake_PL samples had a very circumscribed distribution (Supplementary Fig. 8A) and can be attributed to the species *T. bourrellyi* (Shams et al., 2015; Salmaso et al., 2016). The third clade was characterized by ASVs clustering exclusively with taxa of uncertain classification identified

in rivers, whereas the remaining ASVs clustered together with other unclassified clones from GenBank isolated from river water mats, and *Hydrocoleum* sp.. The annotation of ASVs according to habitat in the tree in Fig. 4 did not show any pattern. This was confirmed by Pagel's λ and the Abouheif-Moran tests calculated for both lake_BFM and river_BFM abundances ($P > 0.2$). Nevertheless, the PCoA analysis computed on the whole set of *Tychonema* ASVs suggested a gradual transition between lake_BFM and river_BFM samples (Supplementary Fig. 7B; PERMANOVA, $P < 0.001$).

The genus *Planktothrix* included 18 ASVs, with sequence similarities between 84.2% and 99.8%, and with 25th and 75th quantiles of 91.6% and 98.0%. Phylogenetic analyses showed that at least half of the ASVs clustered with other *Planktothrix* species from GenBank (Fig. 5). Most ASVs (13 out of 18)

Fig. 5 Maximum likelihood rooted topology of ASVs classified in the genus *Planktothrix* by the naïve Bayesian classifier and taxonomic reference database SILVA 138 based on alignment of 16S rRNA gene fragments (408 bp); the tree is rooted by *Gloeobacter violaceous*. Other features as in Fig. 3. Abbreviations in red refer to exclusive ASVs detected in lakes Pernica (P), Pernica and Bled (P&B), Pernica and Mondsee (P&M) and Pernica and Mantova (P&Ma)
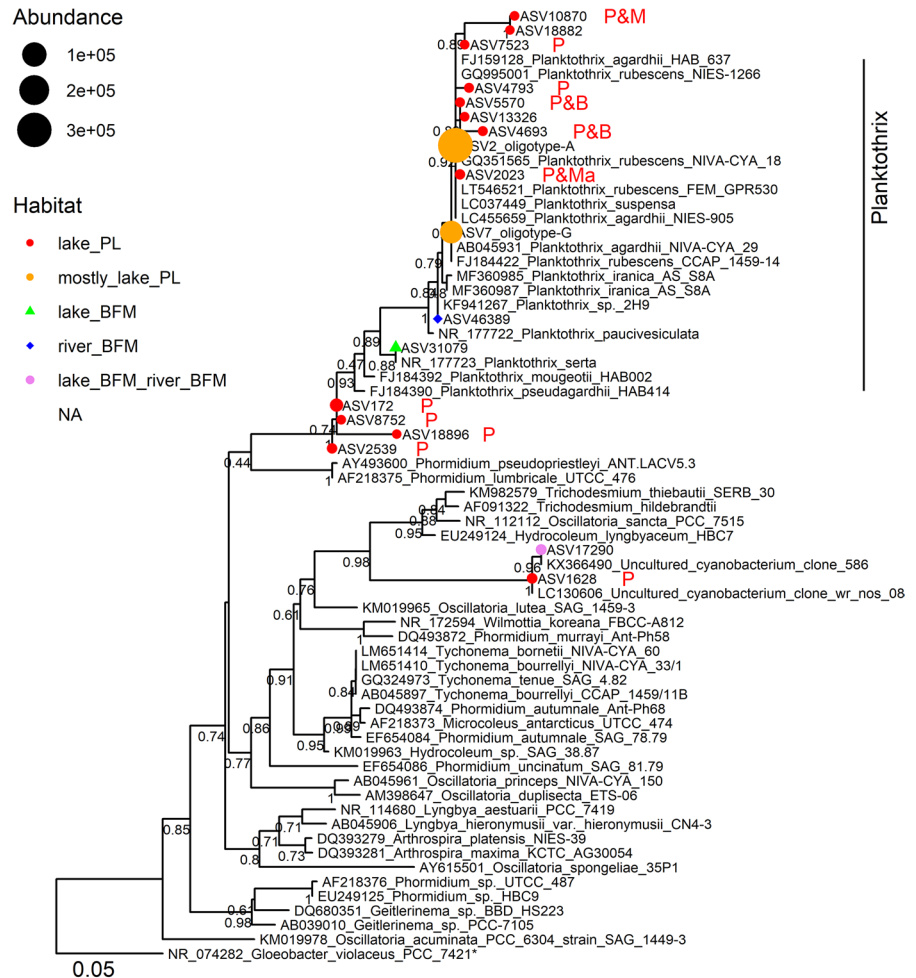
were identified exclusively in the lake_PL samples, although they were also found shared between planktic and benthic habitats. The two most abundant ASVs (oligotypes-A and -G; over 97% of total abundance) were almost exclusively identified in the pelagic samples, with only a few individuals recorded in the biofilm. These ASVs, which differed by only one nucleotide, were widely distributed throughout the Alpine region and corresponded to the species *Planktothrix rubescens* (De Candolle ex Gomont) Anagnostidis & Komárek (Supplementary Fig. 8B). Unexpectedly ASVs 1628 (Lake Pernica) and 17,290 (rivers Adige and Soca, and lakes Aigueblette and Ammersee) were included in a separate cluster, together with two unidentified taxa from GenBank detected from eDNA in Lake Taihu or isolated from a Japanese river. The latter ASV was also detected, using the same laboratory and bioinformatic methods, with 100% sequence similarity in the small Alpine Lake Valagola (NE Italian Alps) and in a small pond nearby Lake Garda (project AcquaViva, MAB-UNESCO; *unpublished data*). The remaining ASVs showed > 99% sequence similarities with different *Planktothrix* species, with the exclusion of ASVs 10,870, 18,882 and 18,896, which are possibly artefacts because the alignment shows higher variability at the 5' end in comparison with *Planktothrix* spp. complete genomes (Entfellner et al., 2022). Notably, 11 of 18 ASVs were identified exclusively or almost exclusively in Lake Pernica (Fig. 5).

Biogeographical distribution in the pelagic and biofilm habitats

Lake plankton, seasonal samples

In dataset 1, 43% of ASVs, representing 2% of the total abundance, occurred exclusively in one lake. The PCoA ordination showed a compact clustering of samples originating from the pelagic habitat of one lake, and an overall significant separation of these clusters (PERMANOVA, $P < 0.001$) (Fig. 6A). Conversely, the two stations of Lake Staffelsee did not show compositional differences (PERMANOVA, pairwise adonis, $P > 0.70$). The latter two stations, together with Lake Garda, were completely separated from the other lakes. Oligotypes A and G of *P. rubescens* were associated with all lakes, with the exception of Lake Staffelsee, where this species was identified with a very low number of sequences (Fig. 6B).

T. bourrellyi (ASV34) showed a strong association with Lake Garda, whereas the two Staffelsee stations were characterized by the presence of several *Snowella* ASVs and a sequence variant of *Cyanobium* (ASV14), which was dominant in the biofilm samples. Lake Lugano showed a high abundance of *Aphanizomenon* MDT14A. Several other *Cyanobium* ASVs showed a specific association with one or more lakes.

The CCA calculated for this dataset showed a distribution of lakes and species consistent with the PCoA (Fig. 6C–D). The importance of the first two components (constrained eigenvalues) was 36% and 24%, respectively; the significance of the CCA ordination (999 permutation tests, $P < 0.01$) was confirmed by the Mantel test computed on the same set of environmental variables associated to the CCA diagram (Table 1A). Their interpretation was not straightforward, mainly because the main trophic response variables (water transparency and Chl_a) showed no significant correlations with TP ($P > 0.2$). The main differences between lakes were due to the higher temperatures in lakes Garda, Lugano and Bourget; the higher values of Chl_a, Secchi and DOp in the large lakes (left and upper left quadrant); the higher concentrations of TP in Lake Staffelsee; the lower and higher concentrations of NO3_N in lakes Garda, and Ammersee and Bourget, respectively. Overall, based on the mean lake values of the biotic and abiotic variables, the ASVs showed a significant association with the whole set of environmental and morphometric variables used in the CCA (Table 1B).

Lake plankton, summer samples

In dataset 2, 48% of ASVs, representing 18% of total abundance, occurred exclusively in a single lake. The importance of the first two components (constrained eigenvalues) in the CCA (Fig. 7A–B) was 27% and 21%, respectively; the CCA ordination was significant at $P < 0.01$ (999 permutation tests). In the CCA plane, the full set of lakes showed a distribution organized along a trophic and climatic gradient (temperature, TP, Chl_a), and a composite gradient pointing towards large and transparent waterbodies with high DOp and NO3_N content. After linearisation by logarithmic transformation, TP and Chl_a were highly and positively correlated ($r = 0.81$, $P < 0.001$); in turn, these two variables were negatively correlated
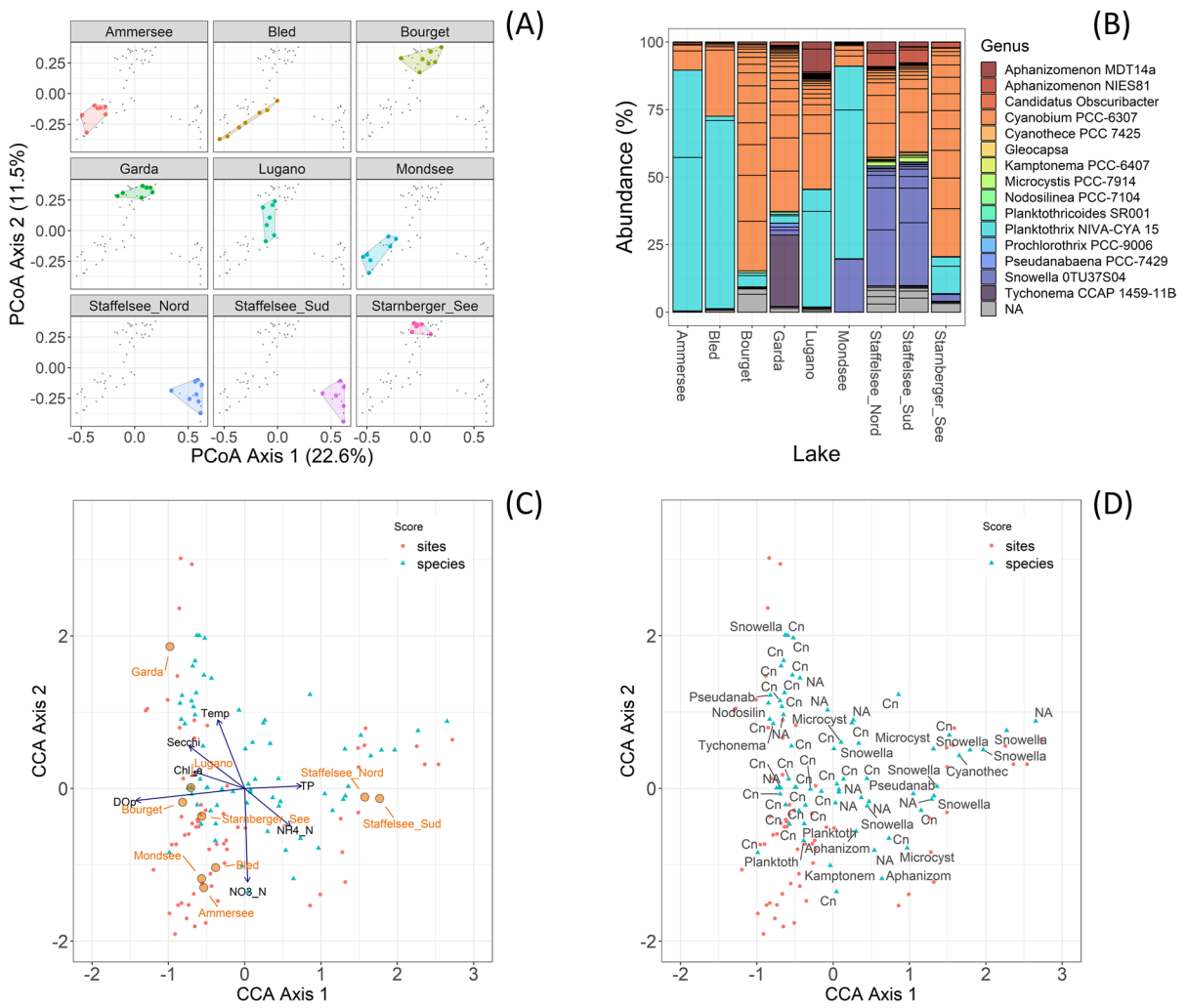
Fig. 6 **A** PCoA of monthly plankton samples collected between April and October in 9 lake stations (8 lakes). **B** Percentage contribution of each ASVs to the cyanobacterial genera calculated from average abundances in each lake station; for each genus, horizontal lines in the bars indicate different oligotypes; genus names follow the SILVA 138 taxonomy. *Aphanizomenon* NIES81 corresponds to *Dolichospermum lemmermannii*. **C** Canonical Correspondence Analysis (CCA) of the same sample set; ASVs with abundances < 15 reads were not included in the analysis. For each individual lake, the position in the graph is indicated by the centroid calculated on the corresponding sample set; Temp, water temperature; TP, total phosphorus; NH4_N, ammonium nitrogen; NO3_N, nitrate nitrogen; DOp, dissolved oxygen saturation (%); Chl_a, chlorophyll-a; Secchi, Secchi disk transparency. **D** Same as (**C**), with labels showing genus names. For ease of visualisation, species names are given in 9 characters; Cn and NA indicate the genus *Cyanobium* and unclassified ASVs, respectively

to the Secchi disk depth ($r = -0.77$ and $r = -0.70$, respectively; $P < 0.001$). The results obtained by the CCA were confirmed by the Mantel test calculated on the same set of environmental and morphometric variables (Table 1C). As expected, the majority of ASVs were represented by different oligotypes of *Cyanobium*, scattered around the CCA plane (Fig. 7B); the most eutrophic water bodies were lakes

Pernica, Frassino, Ragogna, Fiè allo Sciliar, Slivnica, Mantova Superiore, and Serraia. These lakes were associated with different ASVs of *Microcystis*, *Limnothrix*, *Pseudanabaena*, *Snowella,* and an oligotype of *Planktothrix* identified exclusively in lakes Pernica and Mantova Superiore (Figure 5). On the opposite side of the gradient, the larger lakes showed a different group of taxa, represented by *Planktothrix*

Table 1 Association of cyanobacterial community structure with environmental factors based on (A) monthly plankton samples collected between April and October in 9 lake stations (8 lakes); (B) lake averages of monthly plankton samples collected between April and October in 9 lake stations (8 lakes); (C) plankton samples collected during the summer months in 33 lake stations (32 lakes); (D) averages of biofilm samples collected in 9 lake stations (9 lakes)

	Dataset	Habitats	N. stations	N. cases	Factor	Mantel r	P
(A)	1	Lake plankton—April–October	9	64	Env. (7 var)	0.35	<0.01
(B)	1	Lake plankton—April–October	9	9	Env. (7 var)	0.44	<0.01
		(Averages)	9	9	Morph. (4 var)	0.42	<0.01
(C)	2	Lake plankton—Summer samples	33	33	Env. (6 var)	0.49	<0.01
			33	33	Morph. (4 var)	0.37	<0.01
(D)	3	Lake biofilm—Averages	9	9	Env. (7 var)	0.48	<0.01
			9	9	Morph. (4 var)	0.45	<0.05

The different groups of samples refer to datasets 1–3 described in the "Statistical data analysis" section. Analyses based on the six environmental variables (Env.) include Temp, Secchi, DOp, NO_3_N, TP, and Chl_a; the seven variables additionally include NH4_N. Lake physiographic variables (Morph.) include lake elevation, surface area, maximum depth, and catchment area

(mostly *P. rubescens*), *Dolichospermum*, *Tychonema* and, partly, *Pseudanabaena* and *Snowella*. Notably, in a few of the eutrophic lakes (Pernica, Frassino and Mantova Superiore), 1 or 2 oligotypes of *Cylindrospermopsis* (*Raphidiopsis*) were detected. Looking at the first 150 most abundant genera, the proportion of unclassified genera was always low, with the exception of lakes Frassino and Serraia, which showed a high contribution of unclassified taxa at order level (Oxyphotobacteria Incertae Sedis) and unclassified Nostocaceae, respectively (Fig. 7C). In Lake Serraia, a blast analysis of the unclassified ASVs identified several species of *Dolichospermum* and *Anabaena* sharing the same sequence with 100% identity.

Lake biofilm

The 47% of ASVs, corresponding to 4% of the total abundance, were exclusively present in a single lake (dataset 3). The lake_BFM samples showed a compact and significant clustering in the PCoA ordination within each lake (PERMANOVA, Adonis and pairwise Adonis, $P < 0.001$) (Fig. 8A). In addition to ASVs identified at the genus level, the differences between lakes were also due to a large (25–50%) contribution of ASVs classified at taxonomic levels above the genus (Fig. 8B; "NA"); most of the unclassified ASVs showed no classification even at the order level. The average composition and distribution of ASVs in the individual lakes were significantly related to environmental and lake morphometric variables (Table 1D).

River biofilm

The 59% of ASVs, representing 15% of the total abundance, were found exclusively in a single river (dataset 4). In PCoA ordination, BFM samples showed a compact and significant clustering within each river (PERMANOVA, adonis, $P < 0.001$) (Fig. 8C). This was consistent with the high variation in the proportions of several genera between lakes (Fig. 8D). Differences were also associated to a large proportion (30–80%) of unclassified ASVs at the genus level. The majority of unclassified ASVs belonged to the Leptolyngbyaceae family.

Discussion

Investigations in a large number of lakes and rivers in the Alpine region allowed us to assess a high diversity and very limited overlap of ASVs and genera of cyanobacteria in the pelagic areas of lakes and in the epilithic biofilm of lakes and rivers. In addition, the analysis of the composition of oligotypes classified under the same genus names in three selected taxa revealed the presence of known sequence variants among the most abundant reads, but also several sequences with undetermined or uncertain classification.
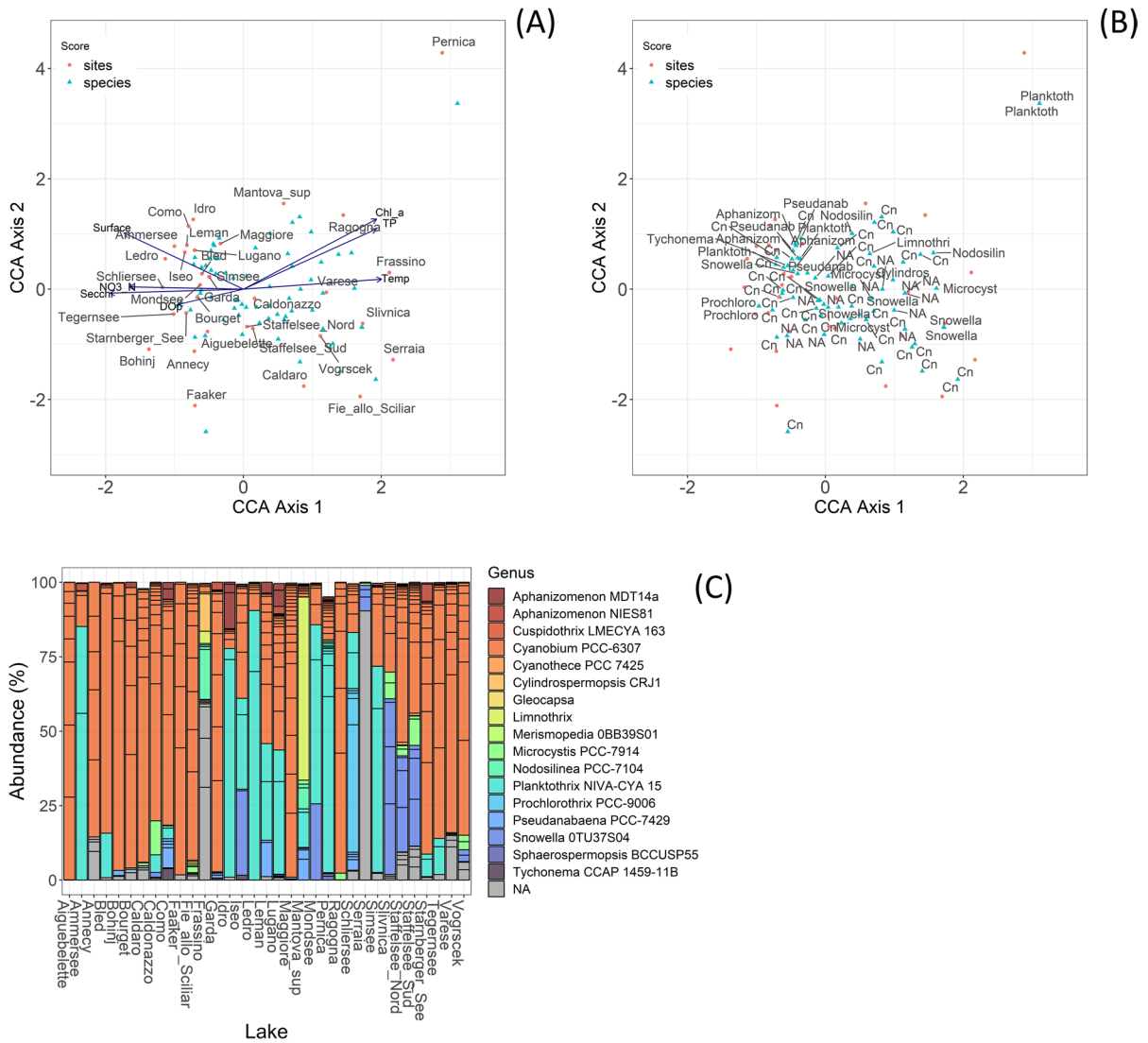
Fig. 7 A Canonical correspondence analysis of plankton samples collected during the summer months in 33 stations (32 lakes); labels show lake names; ASVs with abundances < 25 reads were not included in the analysis. **B** Same as (**A**), with labels showing genus names. For ease of visualisation, species names are given in 9 characters; Cn and NA indicate the genus *Cyanobium* and unclassified ASVs, respectively. **C** Percentage contribution of each ASVs to the cyanobacterial genera calculated from abundances in each lake; the graph shows the first 150 most abundant ASVs; for each genus, horizontal lines in the bars indicate individual oligotypes (e.g., several oligotypes attributed to *Planktothrix* can be observed in Lake Pernica)

Biodiversity of ASVs in the three aquatic habitats

The number of ASVs identified in lake biofilms was an order of magnitude higher and three times higher than the corresponding values determined in plankton and river biofilms, respectively. These differences can be explained by the high diversity and concentration of bacterial and cyanobacterial communities that develop in benthic substrates and the corresponding complexity of trophic interactions (Besemer, 2015; Zancarini et al., 2017; Farkas et al., 2020). However, excluding cyanobacteria and keeping only the other bacterial phyla (over 39,000 ASVs), the number of ASVs in lake_BFM samples was about 3 and 2.5 times higher than the corresponding values in lake_PL and river_BFM, respectively, highlighting a strong
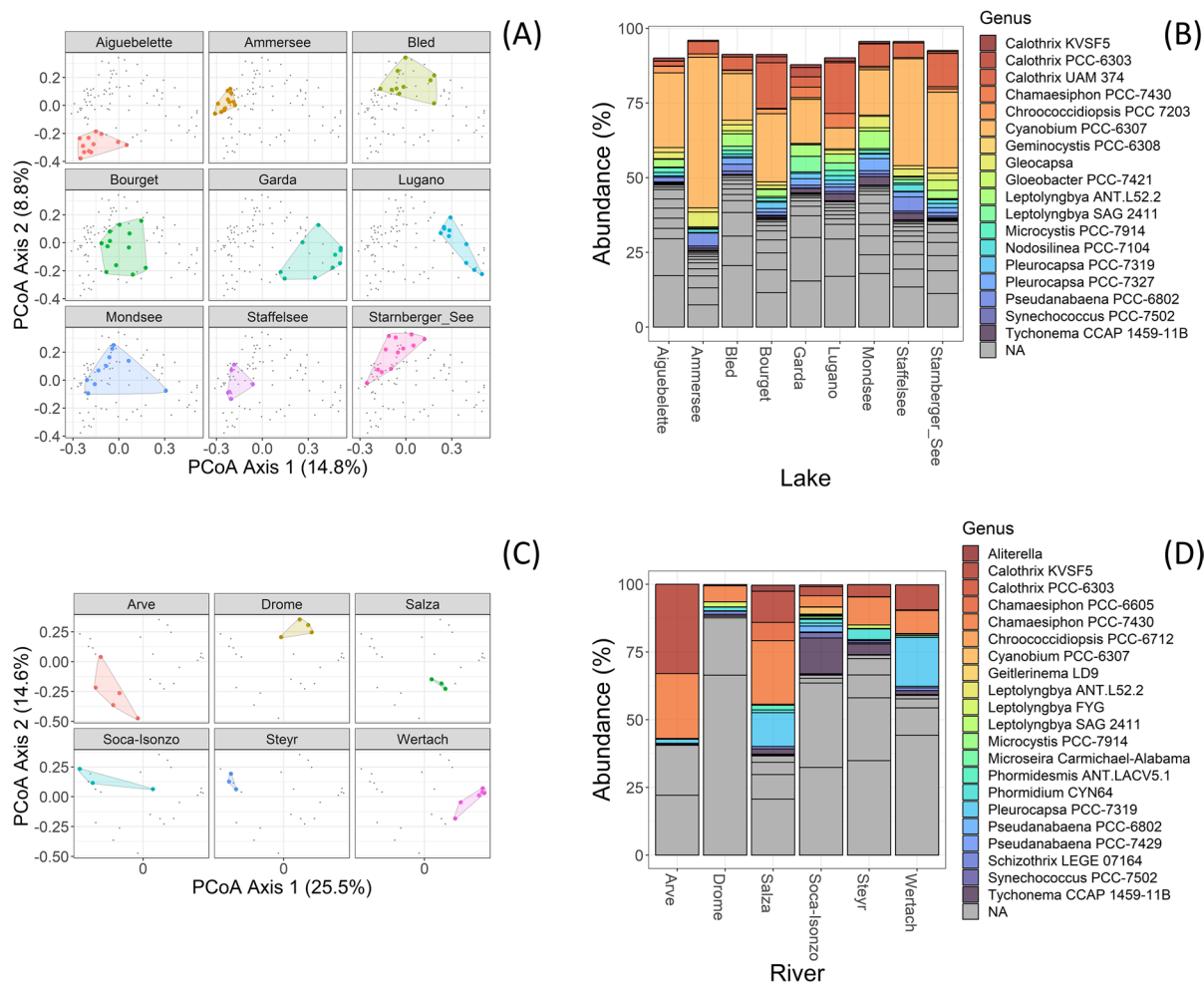
Fig. 8 A PCoA of biofilm samples collected in 9 lakes and **B** mean percentage contribution of the most abundant genera calculated for each individual lake. **C, D** Same as in (**A, B**), but referring to biofilm samples collected in rivers

environmental filtering of cyanobacteria in lake_PL samples compared to bacteria. Furthermore, the flow regime was a negative factor affecting the biodiversity of cyanobacteria in the river biofilm compared to the lentic environments. These considerations implicitly assume the ecological adaptation of cyanobacteria to different habitats and lifestyles. At the functional level of organisation, the plankton and biofilm communities were characterised by the prevalence of known planktic (mostly coccoid and filamentous) and periphytic taxa (at least those classified at the genus and/or family level), respectively, with a clear dominance of different families in the pelagic samples and in the biofilm of lakes and rivers (Stevenson et al., 1996; Wehr & Sheath, 2003). These differences were

further supported by the extremely low number of ASVs shared between the three habitats. These results demonstrate the existence of strong environmental filtering not only at the morpho-functional level, but also at higher taxonomic levels (Fig. 2B), followed by differences also in the selection of oligotypes in the three individual habitats. The main drivers are differences in water and substrate matrices, which require adaptations to the diluted planktic lifestyle or the development of organisms in crowded substrates with high organic and nutrient content.

The differences between habitats were also contributed to by a high number of ASVs occurring in one or a few samples exclusive to one of the three habitats. These ASVs cannot be considered to be

of high discriminatory value, as they occur at very low frequencies and are generally low in abundance (Zhang et al., 2016; Lee et al., 2021).

Diversity estimates based on the number of ASVs must be interpreted correctly, taking into account the multicopy nature and intragenomic variability of the 16S rRNA gene (Větrovský & Baldrian, 2013). As reported by Schirrmeister et al. (2012) and Stoddard et al. (2015), the number of 16S rRNA copies in cyanobacteria is generally between 1 and 5, i.e. much lower than the 18S ribosomal gene copies in protists (> 500,000 in ciliates; Wang et al., 2017). Furthermore, Espejo & Plaza (2018) showed that polymorphic sites in intragenomic 16S rRNA genes are rare and occur at much lower frequencies than those found in different species. Therefore, ASVs do not represent species or clones, but rather different amplicon variants (oligotypes) of the same or different species and clones (Eren et al., 2013). Their variability contributes to the differences in sequences attributed to individual species or genera.

Cyanobacterial biodiversity in lakes and rivers

A complete assessment of biodiversity in ecosystems should consider all their habitat components. In lakes, these include the limnetic zone, and the littoral and wetland zone with their constitutive variety of substrates, benthic communities and vegetation (Wetzel, 2001). Similarly, river substrate diversity and zonation are influential in controlling microbial biomass and diversity (Tett et al., 1978; Nowicka-Krawczyk & Zelazna-Wieczorek, 2013). In this work, the periphytic communities were only sampled in the biofilm developing on stones (epilithon; Supplementary Fig. 2), thus excluding the communities developing on other substrates (i.e. epiphyton, epipelon, epipsammon, and epizoon) or other metaphytic communities composed of microbial/cyanobacterial organisms originating from true floating populations that aggregate among macrophytes and debris of the littoral zone (Stevenson et al., 1996; Wetzel, 2001; Timoshkin et al., 2016). Each of these components can be considered as habitats hosting specific and adapted microbial communities, including cyanobacteria (Stevenson et al., 1996; Zębek et al., 2021). For example, Levi et al. (2017) and Wijewardene et al. (2022) showed that epiphytic communities have unique and different structures and functions compared to other periphyton biofilms in freshwater habitats, therefore contributing to the overall microbial diversity of benthic areas of lakes and rivers. The development of epiphytic cyanobacteria can become massive, as in the case of the recent colonization of reed stems by *Gloeotrichia pisum* Thuret ex Bornet & Flahault after a large rise in the water level of Lake Kinneret flooded the macrophyte stands (Lang-Yona et al., 2023).

The importance of studying cyanobacteria growing on other substrates or in metaphytic mats is also dictated by the potential development of toxigenic and geosmin-producing cyanobacteria not only in open waters, but also in the littoral zone of lakes and rivers (van Breemen et al., 1991; Quiblier et al., 2013; Harland et al., 2014). In this work, many of the most abundant genera identified in the epilithic samples are recognised as toxin-producing (e.g., *Aphanocapsa*, *Cyanobium*, *Geitlerinema*, *Limnothrix*, *Merismopedia*, *Microcystis*, *Oscillatoria*, *Phormidium*, *Pseudanabaena*, *Synechococcus*) (Jakubowska & Szeląg-Wasielewska, 2015; Bernard et al., 2017) and/or geosmin-producing species (e.g., *Calothrix*, *Phormidium*, *Oscillatoria*, *Nodosilinea*, *Geitlerinema*, *Pseudanabaena*) (Suurnäkki et al., 2015; Senavirathna & Jayasekara, 2023). The development of littoral mats or periphyton populations of toxic cyanobacteria is a common cause of fatal poisoning in dogs (Wood et al., 2007, 2010; Backer et al., 2013; Fastner et al., 2018; Bauer et al., 2020), livestock and other animals that drink contaminated water (Huisman et al., 2005; Stewart et al., 2008), whereas in benthic mats cyanobacteria have been found to be the main source of geosmin (Gaget et al., 2022). In the future, the development of benthic cyanobacteria could possibly be intensified due to an increase in the transparency of the littoral zones following restoration and reoligotrophication (Chorus et al., 2021). This process could be further sustained by climate change, which favours an intensification of extreme events, including a higher frequency and persistence of flow reduction episodes in streams (van Vliet et al., 2013; Robichon et al., 2023) and water temperatures and stability in lakes (Paerl & Huisman, 2009; Burford et al., 2019).

Despite their importance and toxigenic potential, benthic cyanobacteria have received much less attention than pelagic species (Burford et al., 2019; Salmaso et al., 2022). This represents a gap in our knowledge, which becomes even more important

when considering the much higher cyanobacterial biodiversity found in this work in a limited fraction of littoral substrata compared to pelagic cyanobacterial populations.

Taxonomic classification of ASVs

The higher proportion of ASVs classified in the lake_PL samples compared to the biofilm habitats can be interpreted considering the greater completeness of the taxonomic databases with information on the planktic cyanobacteria (Salmaso et al., 2022). This bias was contributed by the greater number of studies on toxigenic and non-toxigenic cyanobacteria in open lake waters in the context of eutrophication (Meriluoto et al., 2017) and the main focus on biomonitoring of diatoms in river and lake biofilm (Kelly et al., 2014; Levi et al., 2017). Based on abundance values, the proportion of ASVs classified to genus level in the lake planktic samples was extremely high. Conversely, the number of unclassified ASVs remained high in the lake_BFM and river_BFM samples, even when considering abundances. In the case of the lake_BFM samples, several taxa were also unclassified at order rank, highlighting the existence of unknown groups.

Wang's classification of genera, based on the SILVA database, gave broadly comparable results to the blast analyses, and fully equivalent results at the family taxonomic level. Even where discrepancies occurred, the genera were almost always listed under the best blast hit results. Although providing results that can be used for ecological assessment, these divergences are not always satisfactorily for a complete taxonomic survey of ecosystems. Compared to blast, which provides complete information even for sequences shared by different species and even genera, Wang's classifier reports the classification at the immediately higher rank in such cases. Together with the incompleteness of taxonomic reference archives, this may help to explain the high proportion of unclassified ASVs at genus and species level. A paradigmatic example is Lake Serraia (Fig. 7C), where more than 85% of the unclassified ASVs were represented by a unique sequence attributed to the family Nostocaceae and shared by several genera of *Dolichospermum* and *Anabaena*.

Phylogenetic assessment of ASVs taxonomy

Most of the ASVs classified at the genus level were represented by a variable number of oligotypes with different abundances and sequence identity. The phylogenetic analysis of ASVs within the genera *Cyanobium*, *Tychonema* and *Planktothrix* allowed to clarify several taxonomic and ecological aspects of the nature and significance of the observed variability, but also raised some questions about the nature and biological significance of some ASVs.

The *Cyanobium* sequences were interspersed with several other taxa classified under the *Cyanobium* and *Synechococcus* genera in GenBank. This is in full agreement with several other studies that showed no comprehensive and coherent phylogenetic relationships (Lopes et al., 2012) and identical sequences shared between these two picocyanobacterial genera (Bukowska et al., 2014), possibly due to difficulties in annotating and discriminating the isolated morphotypes deposited in the taxonomic databases. On the other hand, several studies based on isolates and physiological assessment demonstrated the existence of physiologically distinct groups of strains within *Cyanobium* and *Synechococcus* taxa, with different pigment composition, salt tolerance and ecological niches (Callieri & Stockner, 2002; Ernst et al., 2003; Jezberová & Komárková, 2007; Callieri et al., 2022), as well as specific biosynthetic traits in marine environments (Doré et al., 2023). These results are consistent with the differences in the distribution of two different clusters of ASVs in the lake_PL and lake_BFM samples, indicating the existence of different groups of strains adapted to different ecological niches in the pelagic and benthic littoral zones of the lakes, respectively.

The *Tychonema* ASVs were in different major branches of the phylogenetic tree, together with well-known or poorly classified homologous sequences from GenBank. Excluding one doubtful genus (cf. *Tychonema*), one of the major clades (1) included different sequences from GenBank attributed to *Microcoleus* and *Phormidium*. These two genera were the subject of many recent taxonomic revisions (Palinska et al., 2011; Strunecký et al., 2013; Niiyama & Tuji, 2019), although some questions remain (Komárek et al., 2014). As a further element indicating a close relationship of clade-1 with *Microcoleus* and *Phormidium*, six *Tychonema* taxa identified in the biofilm

samples showed the presence of two 11-nucleotide insertions, the most common of which was identical except for a one nucleotide shift in alignment to a previously described 11-insert, e.g. by Taton et al. (2003) and Jungblut et al. (2016). The one nucleotide shift is due to slight differences in the alignments provided by the major rewrite of Muscle5 based on new algorithms (Edgar, 2022) compared to previous versions and other alignment software used in former works. This insert appears to be widespread, although its presence has not always been emphasised (e.g., accession numbers KF770970 and EF654074; Fig. 4). The heterogeneity observed in clade-1 is consistent with the results of a phylogenomic analysis by Strunecký et al. (2023), who found that strains previously designated as *Tychonema* are monophyletic with *Microcoleus*, and *Tychonema* may become a *Microcoleus* species. This is confirmed by the Genome Taxonomy Database (Parks et al., 2022), where several former *Tychonema* strains are now classified into the genus *Microcoleus*. The second major clade contained only *Tychonema* strains, both planktic (*T. bourrellyi*) and benthic (*T. bornetii* and *T. tenue*). When comparisons are restricted to the 16S rRNA region used for metabarcoding (V3-V4), different strains of *Tychonema* spp. may share identical sequences. In such cases, species discrimination should integrate other distinctive features, such as habitat preferences and morphometric characters (Salmaso et al., 2016). In the planktic samples, *Tychonema* was essentially represented by one unique ASV, which could be attributed to *T. bourrellyi*. Its distribution still seems to be mostly restricted to the southern perialpine regions (Supplementary Fig. 8A). On the other hand, the benthic *Tychonema* populations showed a wider distribution, covering areas throughout the Alpine region and evolving, as the planktic ecotypes, with toxigenic strains producing anatoxin-a, as demonstrated in some of the recent cases of animal poisoning caused by toxic clumps and mats of *Tychonema* sp. in the littoral area of Lake Tegel (Fastner et al., 2018) and in the reservoir Mandichosee (Lech River) (Bauer et al., 2020, 2022). Compared to the previous two clades, the remaining ASVs clustered closely with unclassified cyanobacterial clones from GenBank, raising doubts about their taxonomic assignment and nature.

The ASVs designated as *Planktothrix* were almost all represented by the two planktic oligotypes "A" (ASV2) and "G" (ASV7). Both genotypes constitute

the 16S rRNA genes in a strain (4 gene copies per chromosome), e.g., strain PCC7821 or strain no758 (both originally isolated from Scandinavia) contain these two oligotypes (pos. 97, A vs G) (Entfellner et al., 2022). A few rare and quite unusual oligotypes (ASVs 1628 and 17,290) coincided with unclassified sequences, both deposited in GenBank and recently determined in other alpine or perialpine waters. These oligotypes most likely do not belong to the genus *Planktothrix*, and their taxonomic position needs to be clarified. In contrast, some other unusual ASVs showed no correspondence with classified or unclassified sequences in the reference databases. Further investigation is needed in the basins where these sequences were identified to verify the consistency and non-transitory nature of these ASVs. The largest number of unusual sequences was found in Lake Pernica. This lake is highly eutrophic (Ambrožič et al., 2008) due to its location at the edge of the Alpine chain in northeastern Slovenia, which is relatively flat and intensively farmed, and therefore has different characteristics from the other Alpine or perialpine lakes included in this study. Accordingly, *P. rubescens*, which is typical for Alpine lakes and found in the majority of lakes in this study, was never detected in Lake Pernica, whereas *P. agardhii* is commonly observed during regular monitoring (data from Slovenian Environment Agency). In particular, ASV172, ASV2539, ASV8752, ASV31079 were most closely related to phylogenetic lineage 3 including tychoplanktic genera such as *P. tepida* or *P. pseudagardhii* (Entfellner et al., 2022).

Commonness and rarity of ASVs: implications for taxonomic classification

Of the three genera analysed in the previous section, *Tychonema* and *Planktothrix* have been the subject of several studies that have allowed their ecological, taxonomic, and genetic characterisation, especially in planktic environments. Compared to other less known taxa, this can help in the interpretation of the observed diversity and in the assessment of the uniqueness of some of the rarest ASVs found in this work. In both genera, the oligotypes that showed a clear taxonomic classification were by far the most abundant, whereas the ASVs that showed discrepancies in their classification were among the rarest oligotypes. This was consistent with the analysis of all

2620 ASVs included in this work, which showed a tendency for the proportion of taxonomic assignments at the genus level to increase in the most abundant species (Supplementary Fig. 5). In the same samples, similar results were found using a slightly different approach (blast analysis on 16S rRNA and 18S rRNA genes) in both cyanobacteria and phytoplankton (Salmaso et al., 2022). These observations can be explained in two complementary ways. On the one hand, the most abundant species are also those that were more likely to be isolated and genetically characterised. On the other hand, we cannot exclude that some of the rarest species could be ephemeral artefacts that survived the most stringent quality filtering and/or oligotypes that evolved separately in specific water bodies (e.g., Lake Pernica). In the case of some rather rare ASVs, their presence was also fully confirmed in other water bodies in the Alps or in Asia, suggesting that an increase in the range of environments analysed could help to resolve the assessment of abundance, rarity, and uniqueness of ASVs.

Compositional patterns and environmental drivers in lakes and rivers

Analyses of samples collected in lake plankton and littoral epilithic biofilm from lakes and rivers revealed individual compositional patterns unique to each water body. Looking at the most abundant ASVs, the monthly lake_PL samples differed due to several ASVs that were classified within a limited number of taxa restricted to well-known genera such as *Cyanobium*, *Planktothrix*, *Snowella*, *Aphanizomenon*, *Dolichospermum* and, restricted to Lake Garda, *Tychonema*. Based on the whole set of ASVs, the compositional patterns were strongly linked to a number of physical and trophic factors. In this dataset, the lack of correlation between trophic variables was unexpected at first sight but was actually explained by their small range of variation (e.g., mean values of TP between 7 and 15 µg P L^{-1}). Furthermore, the higher values of transparency and Chl_a were found in the largest and deeper lakes, where the contribution of mineral turbidity is generally lower than in the smaller water bodies (Havens & James, 1999; Jones et al., 2008; Nõges, 2009). Conversely, the analysis of lake_PL samples collected annually during the summer months in a larger number of lakes allowed to identify a distribution of lakes and taxa along a trophic, physical, and physiographic gradient. Overall, genera that developed predominantly in larger, deeper lakes and smaller, shallower lakes were recognised as belonging to typical functional groups living in oligo- and mesotrophic waters and in more eutrophic and/or turbid waters, respectively (see Reynolds et al., 2002).

This work has allowed, for the first time, the identification of *Cylindrospermopsis* (*Raphidiopsis*) sequences in some eutrophic lakes in Italy and Slovenia (Salmaso et al., 2022). Within this genus, the toxigenic species *Raphidiopsis* (*Cylindrospermopsis*) *raciborskii* (Woloszynska) Aguilera & al. (Supplementary Table 2) has been recognised as an invasive species of tropical origin (Sukenik et al., 2012; Kokociński et al., 2017). In Italy, this species was previously identified in lakes in Lombardy by microscopic analysis (M. Austoni, pers. comm.). Its identification highlights the potential of eDNA analysis as an effective tool for the early detection of invasive and toxigenic cyanobacteria during geographical surveys.

The existence of compositional patterns unique to each water body was also confirmed in the biofilms collected at different lake and river stations. Analogous to the lake_PL samples, the compositional differences in the lake_BFM samples were correlated with a number of trophic, physical, and physiographic variables, further highlighting the strong environmental filtering at the level of individual lakes. However, analysis of biofilm community structure in lakes and rivers was severely hampered by the high number of unclassified ASVs, which limited comparisons at the genus level to about 50% of the ASVs.

Conclusion

Our work revealed a distinct biodiversity and low overlap of amplicon sequence variants in individual lakes and rivers, with the development of specific families and broad morpho-functional types. Within each pelagic and benthic habitat, the individual character of the water bodies indicated a strong role of environmental filtering in the selection of major genera and ASVs. All habitats and water bodies were characterised by a high proportion of low abundance ASVs occurring in one or a few samples. Due to their stochastic nature, these taxa only partially contributed to the characterisation and functioning of the

community, despite their contribution to increasing biodiversity.

Ecological interpretation of the data was hampered by the low proportion of ASVs classified at least to genus level, particularly in biofilm samples and for many of the rarest ASVs. This was mainly due to the incompleteness of taxonomic reference databases for species living in poorly studied environments and the limited genetic information provided by the short 16S rRNA gene currently used for metabarcoding. The incompleteness of taxonomic data for benthic environments was also highlighted by a consistent proportion of ASVs that showed no classification even at taxonomic levels higher than family.

In the lake biofilm, the number of individual ASVs was ten times higher than the number of ASVs detected in the pelagic zone. This large imbalance in the biodiversity of these two habitats is even more significant when considering that in our work the biofilm was only representative of the epilithic fraction and therefore only provided a rough estimate of the biodiversity complexity of the littoral zone. Despite their importance in terms of biodiversity and as a potential substrate for the colonisation of toxigenic populations, benthic and periphytic habitats in lakes and rivers have been much less studied. Besides diatoms, which are the main target of biomonitoring studies, a consistent inclusion of other components of the microbial community would allow a better characterisation of the functionality of lakes and rivers.

The phylogenetic analysis of three selected genera, namely *Cyanobium*, *Tychonema* and *Planktothrix,* allowed the identification of heterogeneous phylogenetic clades, indicating both the polyphyletic nature of the taxa and the specificity of the results obtained with the taxonomic classifiers and reference databases used. In the case of the two filamentous cyanobacteria, the ASVs were placed either in clades along very well characterised species (e.g., planktic and benthic *Tychonema* or *Planktothrix*) or in branches that were poorly or not taxonomically characterised. In the eutrophic Lake Pernica, Slovenia, some unusual sequences were identified and classified as "*Planktothrix*". However, the non-ephemeral nature of these and many other rare unclassified ASVs remains unknown.

Although limited to three representative genera, the phylogenetic analyses allowed clarification of the taxonomic nature and positions of ASVs that would otherwise be classified in a single genus. In general, this approach should be used whenever there is interest in delving deeper into the taxonomy and differential distribution of ASVs in specific cyanobacterial taxa, including ecotypes. A revealing example was the small picocyanobacteria, which are usually classified into broad functional groups, but in our work were separated into at least two major ecotypes restricted to pelagic habitats and lake biofilms.

Declarations

Competing interest The authors declare that they have no competing interests.

Ethical approval Not applicable.

Informed consent All authors have consented to participate in this work and approve the submission to Hydrobiologia.

References

Ambrožič, Š., I. Cvitanič, M. Dobnikar Tehovnik, M. Gacin, J. Grbović, B. Jesenovec, Š. Kozak Legiša, M. Krajnc, P. Mihorko, M. Poje, Š. Remec Rekar, B. Rotar, & E. Sodja, 2008. Kakovost voda v Sloveniji. Ljubljana: Agencija RS za okolje, Ljubljana.

Anisimova, M., M. Gil, J.-F. Dufayard, C. Dessimoz, & O. Gascuel, 2011. Survey of Branch Support Methods Demonstrates Accuracy, Power, and Robustness of Fast Likelihood-based Approximation Schemes. Systematic Biology 60: 685–699, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3158332/.

APHA, AWWA & WEF, 2018. Standard Methods for the Examination of Water and Wastewater. APHA Press, Washington DC, https://www.wef.org/publications/publications/books/StandardMethods/.

Backer, L. C., J. H. Landsberg, M. Miller, K. Keel, & T. K. Taylor, 2013. Canine Cyanotoxin Poisonings in the United States (1920s–2012): Review of Suspected and Confirmed Cases from Three Data Sources. Toxins 5: 1597–1628, https://www.mdpi.com/2072-6651/5/9/1597.

Bauer, F., J. Fastner, B. Bartha-Dima, W. Breuer, A. Falkenau, C. Mayer, & U. Raeder, 2020. Mass Occurrence of Anatoxin-a- and Dihydroanatoxin-a-Producing *Tychonema* sp. in Mesotrophic Reservoir Mandichosee (River Lech, Germany) as a Cause of Neurotoxicosis in Dogs. Toxins 12: 726, https://www.mdpi.com/2072-6651/12/11/726.

Bauer, F., M. Stix, B. Bartha-Dima, J. Geist, & U. Raeder, 2022. Spatio-Temporal Monitoring of Benthic Anatoxin-a-Producing *Tychonema* sp. in the River Lech, Germany. Toxins 14: 357, https://www.mdpi.com/2072-6651/14/5/357.

Bernard, C., A. Ballot, S. Thomazeau, S. Maloufi, A. Furey, J. Mankiewicz-Boczek, B. Pawlik-Skowronska, C. Capelli, & N. Salmaso, 2017. Appendix 2. Cyanobacteria associated with the production of cyanotoxins In Meriluoto, J., L. Spoof, & G. A. Codd (eds), Handbook on Cyanobacterial Monitoring and Cyanotoxin Analysis. Wiley, Chichester: 501–525.

Berry, M. A., J. D. White, T. W. Davis, S. Jain, T. H. Johengen, G. J. Dick, O. Sarnelle & V. J. Denef, 2017. Are oligotypes meaningful ecological and phylogenetic units? A case study of *Microcystis* in freshwater lakes. Frontiers in Microbiology 08: 365. https://doi.org/10.3389/fmicb.2017.00365/full.

Besemer, K., 2015. Biodiversity, community structure and function of biofilms in stream ecosystems. Research in Microbiology 166: 774–781, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4601154/.

Brown, C. T., L. A. Hug, B. C. Thomas, I. Sharon, C. J. Castelle, A. Singh, M. J. Wilkins, K. C. Wrighton, K. H. Williams, & J. F. Banfield, 2015. Unusual biology across a group comprising more than 15% of domain Bacteria. Nature 523: 208–211, https://www.nature.com/articles/nature14486.

Bukowska, A., A. Bielczyńska, A. Karnkowska, R. J. Chróst & I. Jasser, 2014. Molecular (PCR-DGGE) versus morphological approach: analysis of taxonomic composition of potentially toxic cyanobacteria in freshwater lakes. Aquatic Biosystems 10: 2. https://doi.org/10.1186/2046-9063-10-2.

Burford, M. A., C. C. Carey, D. P. Hamilton, J. Huisman, H. W. Paerl, S. A. Wood, & A. Wulff, 2019. Perspective: Advancing the research agenda for improving understanding of cyanobacteria in a future of global change. Harmful Algae 91: 101601, https://www.sciencedirect.com/science/article/pii/S1568988319300514.

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

Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, & S. P. Holmes, 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods 13: 581–583, http://www.nature.com/articles/nmeth.3869.

Callieri, C., R. Sabatino, A. Di Cesare & R. Bertoni, 2022. Spatial-temporal study of cluster 5 picocyanobacteria and exopolymeric microgels in Lake Maggiore. Advances in Oceanography and Limnology 13: 1–10. https://doi.org/10.4081/aiol.2022.11043.

Callieri, C. & J. G. Stockner, 2002. Freshwater autotrophic picoplankton: a review. Journal of Limnology 61: 1–14. https://doi.org/10.4081/jlimnol.2002.1.

Capelli, C., A. Ballot, L. Cerasino, A. Papini, & N. Salmaso, 2017. Biogeography of bloom-forming microcystin producing and non-toxigenic populations of *Dolichospermum lemmermannii* (Cyanobacteria). Harmful Algae 67: 1–12, http://linkinghub.elsevier.com/retrieve/pii/S1568988316303663.

Chorus, I., J. Fastner, & M. Welker, 2021. Cyanobacteria and Cyanotoxins in a Changing Environment: Concepts, Controversies, Challenges. Water 13: 2463, https://www.mdpi.com/2073-4441/13/18/2463/htm.

CNR IRSA, sede di Verbania, 2020. Ricerche sull'evoluzione del Lago Maggiore. Aspetti limnologici. Programma triennale 2019–2021. Campagna 2019. Commissione Internazionale per la protezione delle acque italo-svizzere (Ed.), Verbania: 1–110.

Domaizon, I., G. Riccioni, M. Pindo, V. Vasselon, R. Kurmayer, A. Boscaini, C. Capelli, A. Bouchez, F. Rimet, M. Vautier, C. Chardon, M. Logez, J.-M. Baudoin, J. Wanzenböck, H. Rund, S. Dobrovolny, P. Hufnagl, A. Gandolfi, J. Bylemans, U. Mischke, T. Elersek, & N. Salmaso, 2021. Technical guidelines for eDNA monitoring in Alpine waters. National Institute of Biology, Ljubljana, https://www.alpine-space.eu/wp-content/uploads/2022/06/2-en-technical-guidelines-1.pdf.

Donis, D., E. Mantzouki, D. F. McGinnis, D. Vachon, I. Gallego, H. P. Grossart, L. N. de Senerpont Domis, S.

Teurlincx, L. Seelen, M. Lürling, Y. Verstijnen, V. Maliaka, J. Fonvielle, P. M. Visser, J. Verspagen, M. van Herk, M. G. Antoniou, N. Tsiarta, V. McCarthy, V. C. Perello, D. Machado-Vieira, A. G. de Oliveira, D. Š Maronić, F. Stević, T. Ž Pfeiffer, I. B. Vucelić, P. Žutinić, M. G. Udovič, A. Plenković-Moraj, L. Bláha, R. Geriš, M. Fránková, K. S. Christoffersen, T. P. Warming, T. Feldmann, A. Laas, K. Panksep, L. Tuvikene, K. Kangro, J. Koreivienė, J. Karosienė, J. Kasperovičienė, K. Savadova-Ratkus, I. Vitonytė, K. Häggqvist, P. Salmi, L. Arvola, K. Rothhaupt, C. Avagianos, T. Kaloudis, S. Gkelis, M. Panou, T. Triantis, S. K. Zervou, A. Hiskia, U. Obertegger, A. Boscaini, G. Flaim, N. Salmaso, L. Cerasino, S. Haande, B. Skjelbred, M. Grabowska, M. Karpowicz, D. Chmura, L. Nawrocka, J. Kobos, H. Mazur-Marzec, P. Alcaraz-Párraga, E. Wilk-Woźniak, W. Krztoń, E. Walusiak, I. Gagala-Borowska, J. Mankiewicz-Boczek, M. Toporowska, B. Pawlik-Skowronska, M. Niedźwiecki, W. Pęczuła, A. Napiórkowska-Krzebietke, J. Dunalska, J. Sieńska, D. Szymański, M. Kruk, A. Budzyńska, R. Goldyn, A. Kozak, J. Rosińska, E. Szeląg-Wasielewska, P. Domek, N. Jakubowska-Krepska, K. Kwasizur, B. Messyasz, A. Pełechata, M. Pełechaty, M. Kokocinski, B. Madrecka-Witkowska, I. Kostrzewska-Szlakowska, M. Frąk, A. Bańkowska-Sobczak, M. Wasilewicz, A. Ochocka, A. Pasztaleniec, I. Jasser, A. M. Antão-Geraldes, M. Leira, V. Vasconcelos, J. Morais, M. Vale, P. M. Raposeiro, V. Gonçalves, B. Aleksovski, S. Krstić, H. Nemova, I. Drastichova, L. Chomova, S. Remec-Rekar, T. Elersek, L. A. Hansson, P. Urrutia-Cordero, A. G. Bravo, M. Buck, W. Colom-Montero, K. Mustonen, D. Pierson, Y. Yang, J. Richardson, C. Edwards, H. Cromie, J. Delgado-Martín, D. García, J. L. Cereijo, J. Gomà, M. C. Trapote, T. Vegas-Vilarrúbia, B. Obrador, A. García-Murcia, M. Real, E. Romans, J. Noguero-Ribes, D. P. Duque, E. Fernández-Morán, B. Úbeda, J. Á. Gálvez, N. Catalán, C. Pérez-Martínez, E. Ramos-Rodríguez, C. Cillero-Castro, E. Moreno-Ostos, J. M. Blanco, V. Rodríguez, J. J. Montes-Pérez, R. L. Palomino, E. Rodríguez-Pérez, A. Hernández, R. Carballeira, A. Camacho, A. Picazo, C. Rochera, A. C. Santamans, C. Ferriol, S. Romo, J. M. Soria, A. Özen, T. Karan, N. Demir, M. Beklioğlu, N. Filiz, E. Levi, U. Iskin, G. Bezirci, Ü. N. Tavşanoğlu, K. Çelik, K. Ozhan, N. Karakaya, M. A. T. Koçer, M. Yilmaz, F. Maraşlıoğlu, Ö. Fakioglu, E. N. Soylu, M. A. Yağcı, Ş Çınar, K. Çapkın, A. Yağcı, M. Cesur, F. Bilgin, C. Bulut, R. Uysal, K. Latife, R. Akçaalan, M. Albay, M. T. Alp, K. Özkan, T. O. Sevindik, H. Tunca, H. Önem, H. Paerl, C. C. Carey & B. W. Ibelings, 2021. Stratification strength and light climate explain variation in chlorophyll a at the continental scale in a European multilake survey in a heatwave summer. Limnology and Oceanography 66: 4314–4333. https://doi.org/10.1002/lno.11963.

Doré, H., U. Guyet, J. Leconte, G. K. Farrant, B. Alric, M. Ratin, M. Ostrowski, M. Ferrieux, L. Brillet-Guéguen, M. Hoebeke, J. Siltanen, G. Le Corguillé, E. Corre, P. Wincker, D. J. Scanlan, D. Eveillard, F. Partensky, & L. Garczarek, 2023. Differential global distribution of marine picocyanobacteria gene clusters reveals distinct niche-related adaptive strategies. The ISME Journal 17: 720–732, https://www.nature.com/articles/s41396-023-01386-0.

Edgar, R. C., 2016. UCHIME2: improved chimera prediction for amplicon sequencing. bioRxiv, 074252, https://doi.org/10.1101/074252v1.

Edgar, R. C., 2022. High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. bioRxiv, 2021.06.20.449169, https://doi.org/10.1101/2021.06.20.449169v2.

Entfellner, E., R. Li, Y. Jiang, J. Ru, J. Blom, L. Deng & R. Kurmayer, 2022. Toxic/Bioactive peptide synthesis genes rearranged by insertion sequence elements among the bloom-forming cyanobacteria *Planktothrix*. Frontiers in Microbiology 13: 1–20. https://doi.org/10.3389/fmicb.2022.901762.

Eren, A. M., L. Maignien, W. J. Sul, L. G. Murphy, S. L. Grim, H. G. Morrison, & M. L. Sogin, 2013. Oligotyping: Differentiating between closely related microbial taxa using 16S rRNA gene data. Methods in Ecology and Evolution 4: 1111–1119, http://www.ncbi.nlm.nih.gov/pubmed/24358444.

Ernst, A., S. Becker, U. I. A. Wollenzien & C. Postius, 2003. Ecosystem-dependent adaptive radiations of picocyanobacteria inferred from 16S rRNA and ITS-1 sequence analysis. Microbiology 149: 217–228. https://doi.org/10.1099/mic.0.25475-0.

Espejo, R. T. & N. Plaza, 2018. Multiple Ribosomal RNA operons in bacteria; their concerted evolution and potential consequences on the rate of evolution of their 16S rRNA. Frontiers in Microbiology 9: 1232. https://doi.org/10.3389/fmicb.2018.01232/full.

Farkas, M., E. Kaszab, J. Radó, J. Hähn, G. Tóth, P. Harkai, Á. Ferincz, Z. Lovász, A. Táncsics, L. Vörös, B. Kriszt & S. Szoboszlay, 2020. Planktonic and benthic bacterial communities of the largest central European shallow lake, Lake Balaton and its main inflow Zala River. Current Microbiology 77: 4016–4028. https://doi.org/10.1007/s00284-020-02241-7.

Fastner, J., C. Beulker, B. Geiser, A. Hoffmann, R. Kröger, K. Teske, J. Hoppe, L. Mundhenk, H. Neurath, D. Sagebiel & I. Chorus, 2018. Fatal neurotoxicosis in dogs associated with tychoplanktic, anatoxin-a producing *Tychonema* sp. in mesotrophic Lake Tegel, Berlin. Toxins 10: 1–11. https://doi.org/10.3390/toxins10020060.

Gaget, V., H. Almuhtaram, F. Kibuye, P. Hobson, A. Zamyadi, E. Wert, & J. D. Brookes, 2022. Benthic cyanobacteria: A utility-centred field study. Harmful Algae 113: 102185, https://www.sciencedirect.com/science/article/pii/S1568988322000142.

García-García, N., J. Tamames, A. M. Linz, C. Pedrós-Alió, & F. Puente-Sánchez, 2019. Microdiversity ensures the maintenance of functional microbial communities under changing environmental conditions. The ISME Journal 13: 2969–2983, https://www.nature.com/articles/s41396-019-0487-8.

Greenacre, M. & T. Hastie, 1987. The geometric interpretation of correspondence analysis. Journal of the American Statistical Association 82: 437–447. https://doi.org/10.1080/01621459.1987.10478446.

Guellati, F. Z., H. Touati, K. Tambosco, C. Quiblier, J.-F. Humbert & M. Bensouilah, 2017. Unusual cohabitation

and competition between *Planktothrix rubescens* and *Microcystis* sp. (cyanobacteria) in a subtropical reservoir (Hammam Debagh) located in Algeria. PLoS ONE 12: e0183540. https://doi.org/10.1371/journal.pone.0183540.

Guindon, S., J.-F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, & O. Gascuel, 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Systematic Biology 59: 307–321, http://sysbio.oxfordjournals.org/content/59/3/307.long.

Harland, F. M. J., S. A. Wood, P. A. Broady, S. Gaw & W. M. Williamson, 2014. Polyphasic studies of cyanobacterial strains isolated from benthic freshwater mats in Canterbury, New Zealand. New Zealand Journal of Botany 52: 116–135.

Havens, K. E. & R. T. James, 1999. Localized Changes in Transparency Linked to Mud Sediment Expansion in Lake Okeechobee, Florida: Ecological and Management Implications. Lake and Reservoir Management 15: 54–69. https://doi.org/10.1080/07438149909353952.

Huisman, J., H. C. P. Matthijs, & P. M. Visser (eds), 2005. Harmful Cyanobacteria. Springer, Berlin, https://doi.org/10.1007/1-4020-3022-3.

ISO, 1992. Water quality - measurement of biochemical parameters - spectrometric determination of the chlorophyll-a concentration. International Organization for Standardization, Geneve, Switzerland.

Jakubowska, N., & E. Szeląg-Wasielewska, 2015. Toxic Picoplanktonic Cyanobacteria - Review. Marine Drugs 13: 1497–1518, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4377996/.

Jezberová, J. & J. Komárková, 2007. Morphometry and growth of three *Synechococcus*-like picoplanktic cyanobacteria at different culture conditions. Hydrobiologia 578: 17–27. https://doi.org/10.1007/s10750-006-0429-0.

Johnson, J. S., D. J. Spakowicz, B. Y. Hong, L. M. Petersen, P. Demkowicz, L. Chen, S. R. Leopold, B. M. Hanson, H. O. Agresta, M. Gerstein, E. Sodergren, & G. M. Weinstock, 2019. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. Nature Communications 10: 1–11, https://www.nature.com/articles/s41467-019-13036-1.

Jones, J. R., D. V. Obrecht, B. D. Perkins, M. F. Knowlton, A. P. Thorpe, S. Watanabe & R. R. Bacon, 2008. Nutrients, seston, and transparency of Missouri reservoirs and oxbow lakes: an analysis of regional limnology. Lake and Reservoir Management 24: 155–180. https://doi.org/10.1080/07438140809354058.

Jungblut, A. D., I. Hawes, T. J. Mackey, M. Krusor, P. T. Doran, D. Y. Sumner, J. A. Eisen, C. Hillman, & A. K. Goroncy, 2016. Microbial mat communities along an oxygen gradient in a perennially ice-covered antarctic lake. Applied and Environmental Microbiology 82: 620–630, https://europepmc.org/articles/PMC4711149/.

Kelly, M., G. Urbanic, E. Acs, H. Bennion, V. Bertrin, A. Burgess, L. Denys, S. Gottschalk, M. Kahlert, S. M. Karjalainen, B. Kennedy, G. Kosi, A. Marchetto, S. Morin, J. Picinska-Fałtynowicz, S. Poikane, J. Rosebery, I. Schoenfelder, J. Schoenfelder & G. Varbiro, 2014. Comparing aspirations: intercalibration of ecological status concepts across European lakes for littoral diatoms.

Hydrobiologia 734: 125–141. https://doi.org/10.1007/s10750-014-1874-9.

Kokociński, M., R. Akçaalan, N. Salmaso, M. P. Stoyneva-Gärtner & A. Sukenik, 2017. Expansion of alien and invasive cyanobacteria. In Meriluoto, J., L. Spoof & G. A. Codd (eds), Handbook on Cyanobacterial Monitoring and Cyanotoxin Analysis Wiley, Chichester: 28–39.

Komárek, J., 2016. A polyphasic approach for the taxonomy of cyanobacteria: principles and applications. European Journal of Phycology 1–8, https://doi.org/10.1080/09670262.2016.1163738.

Komárek, J., J. Kaštovský, J. Mareš & J. R. Johansen, 2014. Taxonomic classification of cyanoprokaryotes (cyanobacterial genera) 2014, using a polyphasic approach. Preslia 86: 295–335.

Komarek, J., E. Zapomelova, J. Smarda, J. Kopecky, E. Rejmankova, J. Woodhouse, B. A. Neilan, & J. Komarkova, 2013. Polyphasic evaluation of *Limnoraphis robusta*, a water-bloom forming cyanobacterium from Lake Atitlan, Guatemala, with a description of *Limnoraphis* gen. nov. Fottea 13: 39–52, https://doi.org/10.5507/fot.2013.004.html.

Kurmayer, R., G. Christiansen, A. Holzinger & E. Rott, 2018. Single colony genetic analysis of epilithic stream algae of the genus *Chamaesiphon* spp. Hydrobiologia 811: 61–75. https://doi.org/10.1007/s10750-017-3295-z.

Kurmayer, R., T. Elersek, I. Domaizon, C. Capelli, J. Schaumburg, U. Mischke, G. Riccioni, L. Cerasino, S. Bernabei, C. Greco, C. Martone, M. Insolvibile, P. Tomassetti, F. Grassi, A. Giorgio, & N. Salmaso, 2021. Perspectives in eDNA monitoring in alpine waters - for stakeholders and decision-makers. Eco-AlpsWater project. National Institute of Biology, Ljubljana, https://www.alpine-space.org/projects/eco-alpswater/.

Lang-Yona, N., A. Alster, D. Cummings, Z. Freiman, R. Kaplan-Levy, A. Lupu, N. Malinsky-Rushansky, S. Ninio, A. Sukenik, Y. Viner-Mozzini & T. Zohary, 2023. *Gloeotrichia pisum* in Lake Kinneret: a successful epiphytic cyanobacterium. Journal of Phycology 59: 97–110. https://doi.org/10.1111/jpy.13301.

Lê, S., J. Josse, & F. Husson, 2008. FactoMineR: An R Package for Multivariate Analysis. Journal of Statistical Software 25: 1–18, https://doi.org/10.18637/jss.v025.i01.

Lee, E., U. M. Ryan, P. Monis, G. B. McGregor, A. Bath, C. Gordon, & A. Paparini, 2014. Polyphasic identification of cyanobacterial isolates from Australia. Water Research 59: 248–261, http://www.ncbi.nlm.nih.gov/pubmed/24810741.

Lee, S.-H., T.-S. Kim, & H.-D. Park, 2021. Transient-rare bacterial taxa are assembled neutrally across temporal scales. Microbes and Environments 36: ME20110, https://doi.org/10.1264/jsme2.ME20110.

Legendre, P. & L. Legendre, 1998. Numerical Ecology, Elsevier, Amsterdam:

Levi, P. S., P. Starnawski, B. Poulsen, A. Baattrup-Pedersen, A. Schramm & T. Riis, 2017. Microbial community diversity and composition varies with habitat characteristics and biofilm function in macrophyte-rich streams. Oikos 126: 398–409. https://doi.org/10.1111/oik.03400.

Lopes dos Santos, A., C. Gérikas Ribeiro, D. Ong, L. Garczarek, X. L. Shi, S. D. Nodder, D. Vaulot, & A.

Gutiérrez-Rodríguez, 2022. Phytoplankton diversity and ecology through the lens of high throughput sequencing technologies. Advances in Phytoplankton Ecology 353–413.

Lopes, V. R., V. Ramos, A. Martins, M. Sousa, M. Welker, A. Antunes, & V. M. Vasconcelos, 2012. Phylogenetic, chemical and morphological diversity of cyanobacteria from Portuguese temperate estuaries. Marine Environmental Research 73: 7–16, http://www.sciencedirect.com/science/article/pii/S0141113611001383.

Lozupone, C., & R. Knight, 2005. UniFrac: a new phylogenetic method for comparing microbial communities. Applied Environmental Microbiology 71: 8228–8235, http://www.ncbi.nlm.nih.gov/pubmed/16332807.

Lozupone, C., M. E. Lladser, D. Knights, J. Stombaugh, & R. Knight, 2011. UniFrac: an effective distance metric for microbial community comparison. The ISME Journal 5: 169–172, https://www.nature.com/articles/ismej201013.

MacKeigan, P. W., R. E. Garner, M.-È. Monchamp, D. A. Walsh, V. E. Onana, S. A. Kraemer, F. R. Pick, B. E. Beisner, M. D. Agbeti, N. B. da Costa, B. J. Shapiro, & I. Gregory-Eaves, 2022. Comparing microscopy and DNA metabarcoding techniques for identifying cyanobacteria assemblages across hundreds of lakes. Harmful Algae 113: 102187, https://www.sciencedirect.com/science/article/pii/S1568988322000166.

Mantzouki, E., M. Lürling, J. Fastner, L. de Senerpont Domis, E. Wilk-Woˊzniak, E. Wilk-Woźniak, J. Koreivienė, L. Seelen, S. Teurlincx, Y. Verstijnen, W. Krztoń, E. Walusiak, J. Karosienė, J. Kasperovičienė, K. Savadova, I. Vitonytė, C. Cillero-Castro, A. Budzynska, R. Goldyn, A. Kozak, J. Rosińska, E. Szeląg-Wasielewska, P. Domek, N. Jakubowska-Krepska, K. Kwasizur, B. Messyasz, A. Pełechata, M. Pełechaty, M. Kokocinski, A. García-Murcia, M. Real, E. Romans, J. Noguero-Ribes, D. P. Duque, E. Fernández-Morán, N. Karakaya, K. Häggqvist, N. Demir, M. Beklioğlu, N. Filiz, E. E. Levi, U. Iskin, G. Bezirci, Ü. N. Tavşanoğlu, K. Özhan, S. Gkelis, M. Panou, Ö. Fakioglu, C. Avagianos, T. Kaloudis, K. Çelik, M. Yilmaz, R. Marcé, N. Catalán, A. G. Bravo, M. Buck, W. Colom-Montero, K. Mustonen, D. Pierson, Y. Yang, P. M. Raposeiro, V. Gonçalves, M. G. Antoniou, N. Tsiarta, V. McCarthy, V. C. Perello, T. Feldmann, A. Laas, K. Panksep, L. Tuvikene, I. Gagala, J. Mankiewicz-Boczek, M. A. Yağcı, Ş. Çınar, K. Çapkın, A. Yağcı, M. Cesur, F. Bilgin, C. Bulut, R. Uysal, U. Obertegger, A. Boscaini, G. Flaim, N. Salmaso, L. Cerasino, J. Richardson, P. M. Visser, J. M. H. Verspagen, T. Karan, E. N. Soylu, F. Maraşlıoğlu, A. Napiórkowska-Krzebietke, A. Ochocka, A. Pasztaleniec, A. M. Antão-Geraldes, V. Vasconcelos, J. Morais, M. Vale, L. Köker, R. Akçaalan, M. Albay, D. Špoljarić Maronić, F. Stević, T. Žuna Pfeiffer, J. Fonvielle, D. Straile, K. O. Rothhaupt, L. A. Hansson, P. Urrutia-Cordero, L. Bláha, R. Geriš, M. Fránková, M. A. T. Koçer, M. T. Alp, S. Remec-Rekar, T. Elersek, T. Triantis, S. K. Zervou, A. Hiskia, S. Haande, B. Skjelbred, B. Madrecka, H. Nemova, I. Drastichova, L. Chomova, C. Edwards, T. O. Sevindik, H. Tunca, B. Önem, B. Aleksovski, S. Krstić, I. B. Vucelić, L. Nawrocka, P. Salmi, D. Machado-Vieira, A. G. De Oliveira, J. Delgado-Martín, D. García, J. L. Cereijo, J. Gomà, M. C. Trapote, T. Vegas-Vilarrúbia, B. Obrador, M. Grabowska, M. Karpowicz, D. Chmura, B. Úbeda, J. Á. Gálvez, A. Özen, K. S. Christoffersen, T. P. Warming, J. Kobos, H. Mazur-Marzec, C. Pérez-Martínez, E. Ramos-Rodríguez, L. Arvola, P. Alcaraz-Párraga, M. Toporowska, B. Pawlik-Skowronska, M. Niedźwiecki, W. Pęczuła, M. Leira, A. Hernández, E. Moreno-Ostos, J. M. Blanco, V. Rodríguez, J. J. Montes-Pérez, R. L. Palomino, E. Rodríguez-Pérez, R. Carballeira, A. Camacho, A. Picazo, C. Rochera, A. C. Santamans, C. Ferriol, S. Romo, J. M. Soria, J. Dunalska, J. Sieńska, D. Szymański, M. Kruk, I. Kostrzewska-Szlakowska, I. Jasser, P. Žutinić, M. Gligora Udovič, A. Plenković-Moraj, M. Frąk, A. Bańkowska-Sobczak, M. Wasilewicz, K. Özkan, V. Maliaka, K. Kangro, H. P. Grossart, H. W. Paerl, C. C. Carey, & B. W. Ibelings, 2018. Temperature effects explain continental scale distribution of cyanobacterial toxins. Toxins 10: 156, http://www.ncbi.nlm.nih.gov/pubmed/29652856.

Marchesi, M., P. Montanari, E. Bettoni, F. Buzzi, R. M. Di Piazza, & P. Genoni, 2020. Stato delle acque superficiali in Lombardia. Laghi di Mantova. Aggiornamento 2014–2019. ARPA Lombardia: 1–53.

Mareš, J., Y. Lara, I. Dadáková, T. Hauer, B. Uher, A. Wilmotte & J. Kaštovský, 2015. Phylogenetic analysis of cultivation-resistant terrestrial cyanobacteria with massive sheaths (*Stigonema* spp. and *Petalonema alatum*, Nostocales, Cyanobacteria) using single-cell and filament sequencing of environmental samples. Journal of Phycology 51: 288–297. https://doi.org/10.1111/jpy.12273.

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17: 10–12, https://doi.org/10.14806/ej.17.1.200.

McMurdie, P. J. & S. Holmes, 2013. Phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. PLoS ONE Public Library of Science 8: e61217. https://doi.org/10.1371/journal.pone.0061217.

Meriluoto, J., L. Blaha, G. Bojadzija, M. Bormans, L. Brient, G. A. Codd, D. Drobac, E. J. Faassen, J. Fastner, A. Hiskia, B. W. Ibelings, T. Kaloudis, M. Kokocinski, R. Kurmayer, D. Pantelić, A. Quesada, N. Salmaso, N. Tokodi, T. M. Triantis, P. M. Visser, & Z. Svirčev, 2017. Toxic cyanobacteria and cyanotoxins in European waters – recent progress achieved through the CYANOCOST Action and challenges for further research. Advances in Oceanography and Limnology 8: 161–178, http://pagepressjournals.org/index.php/aiol/article/view/6429.

Mortimer, C. H., 1981. The oxygen content of air-saturated fresh waters over ranges of temperature and atmospheric pressure of limnological interest. SIL Communications 1953–1996(22): 1–23. https://doi.org/10.1080/05384680.1981.11904000.

Nearing, J. T., G. M. Douglas, A. M. Comeau, & M. G. I. Langille, 2018. Denoising the Denoisers: an independent evaluation of microbiome sequence error-correction approaches. PeerJ 6: e5364, https://peerj.com/articles/5364.

Niiyama, Y., & A. Tuji, 2019. *Microcoleus pseudautumnalis* sp. nov. (Cyanobacteria, Oscillatoriales) producing

2-methylisoborneol. Bulletin of the National Museum of Nature and Science. Series B, Botany 45: 93–101, https://cir.nii.ac.jp/crid/1523388080253362688.

Nõges, T., 2009. Relationships between morphometry, geographic location and water quality parameters of European lakes. Hydrobiologia 633: 33–43. https://doi.org/10.1007/s10750-009-9874-x.

Nowicka-Krawczyk, P. B., & J. Zelazna-Wieczorek, 2013. Cyanobacteria microflora in a limestone spring (Troniny spring, Central Poland). Acta Societatis Botanicorum Poloniae 82: 219–224, http://agro.icm.edu.pl/agro/element/bwmeta1.element.agro-a9f4dd93-462d-4fa1-84bb-f92797459f19.

Nwosu, E. C., P. Roeser, S. Yang, S. Pinkerneil, L. Ganzert, E. Dittmann, A. Brauer, D. Wagner & S. Liebner, 2021. Species-level spatio-temporal dynamics of cyanobacteria in a hard-water temperate lake in the Southern Baltics. Frontiers in Microbiology 12: 1–17. https://doi.org/10.3389/fmicb.2021.761259.

Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, E. Szoecs, & H. Wagner, 2020. vegan: Community Ecology Package, https://cran.r-project.org/package=vegan.

Paerl, H. W., & J. Huisman, 2009. Climate change: a catalyst for global expansion of harmful cyanobacterial blooms. Environmental Microbiology Reports 1: 27–37, http://www.ncbi.nlm.nih.gov/pubmed/23765717.

Palinska, K. A., B. Deventer, K. Hariri & M. Lotocka, 2011. A taxonomic study on *Phormidium*-group (cyanobacteria) based on morphology, pigments, RAPD molecular markers and RFLP analysis of the 16S rRNA gene fragment. Fottea 11: 41–55. https://doi.org/10.5507/fot.2011.006.html.

Paradis, E., J. Claude & K. Strimmer, 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20: 289–290.

Parks, D. H., M. Chuvochina, C. Rinke, A. J. Mussig, P.-A. Chaumeil & P. Hugenholtz, 2022. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. Nucleic Acids Research 50: D785–D794. https://doi.org/10.1093/nar/gkab776.

Pawlowski, J., M. Kelly-Quinn, F. Altermatt, L. Apothéloz-Perret-Gentil, P. Beja, A. Boggero, A. Borja, A. Bouchez, T. Cordier, I. Domaizon, M. J. Feio, A. F. Filipe, R. Fornaroli, W. Graf, J. Herder, B. van der Hoorn, J. Iwan Jones, M. Sagova-Mareckova, C. Moritz, J. Barquín, J. J. Piggott, M. Pinna, F. Rimet, B. Rinkevich, C. Sousa-Santos, V. Specchia, R. Trobajo, V. Vasselon, S. Vitecek, J. Zimmerman, A. Weigand, F. Leese & M. Kahlert, 2018. The future of biotic indices in the ecogenomic era: Integrating (e)DNA metabarcoding in biological assessment of aquatic ecosystems. Science of the Total Environment 637–638: 1295–1310. https://doi.org/10.1016/j.scitotenv.2018.05.002.

Pawlowski, J., K. Bruce, K. Panksep, F. I. Aguirre, S. Amalfitano, L. Apothéloz-Perret-Gentil, T. Baussant, A. Bouchez, L. Carugati, K. Cermakova, T. Cordier, C. Corinaldesi, F. O. Costa, R. Danovaro, A. Dell'Anno, S. Duarte, U. Eisendle, B. J. D. Ferrari, F. Frontalini, L.

Frühe, A. Haegerbaeumer, V. Kisand, A. Krolicka, A. Lanzén, F. Leese, F. Lejzerowicz, E. Lyautey, I. Maček, M. Sagova-Marečková, J. K. Pearman, X. Pochon, T. Stoeck, R. Vivien, A. Weigand, & S. Fazi, 2022. Environmental DNA metabarcoding for benthic monitoring: A review of sediment sampling and DNA extraction methods. Science of the Total Environment 818: 151783, https://www.sciencedirect.com/science/article/pii/S0048969721068595.

Pokorný, J., L. Štenclová & J. Kaštovský, 2023. Unsuspected findings about phylogeny and ultrastructure of the enigmatic cyanobacterium *Microcrocis geminata* resulted in its epitypification and novel placement in Geminocystaceae. Fottea 23: 110–121. https://doi.org/10.5507/fot.2022.016.html.

Pushkareva, E., I. S. Pessi, A. Wilmotte & J. Elster, 2015. Cyanobacterial community composition in Arctic soil crusts at different stages of development. FEMS Microbiology Ecology 91: fiv143. https://doi.org/10.1093/femsec/fiv143.

Quast, C., E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies, & F. O. Glöckner, 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research 41: D590–6, http://www.ncbi.nlm.nih.gov/pubmed/23193283.

Quiblier, C., S. Wood, I. Echenique-Subiabre, M. Heath, A. Villeneuve, & J.-F. Humbert, 2013. A review of current knowledge on toxic benthic freshwater cyanobacteria - ecology, toxin production and risk management. Water research 47: 5464–5479, http://www.ncbi.nlm.nih.gov/pubmed/23891539.

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

Reynolds, C. S., V. Huszar, C. Kruk, L. Naselli-Flores & S. Melo, 2002. Towards a functional classification of the freshwater phytoplankton. Journal of Plankton Research 24: 417–428.

Rimet, F., M. Vautier, R. Kurmayer, N. Salmaso, C. Capelli, A. Bouchez, P. Hufnagl, & I. Domaizon, 2020. River biofilms sampling for both downstream DNA analysis and microscopic counts V.1. Protocols.io , https://www.protocols.io/view/river-biofilms-sampling-for-both-downstream-dna-an-ben6jdhe.

Rimet, F., M. Vautier, R. Kurmayer, N. Salmaso, C. Capelli, A. Bouchez, P. Hufnagl, I. Domaizon, C. Chardon, A. Bouchez, J. Gueguen, M. Vautier, I. Domaizon, J. Gueguen, A. Bouchez, I. Domaizon, J. Gueguen, M. Vautier, I. Domaizon, R. Kurmayer, N. Salmaso, C. Capelli, A. Bouchez, P. Hufnagl, & I. Domaizon, 2021. updated version- Lake biofilms sampling for both downstream DNA analysis and microscopic counts. protocols.io, https://www.protocols.io/view/lake-biofilms-sampling-for-both-downstream-dna-ana-bfk4jkyw.html.

Robichon, C., J. Robin, & S. Dolédec, 2023. Relative effect of hydraulics, physico-chemistry and other biofilm algae on benthic cyanobacteria assemblages in a regulated river. Science of The Total Environment 872: 162142, https://www.sciencedirect.com/science/article/pii/S0048969723007581.

Rosselló-Móra, R., & W. B. Whitman, 2019. Dialogue on the nomenclature and classification of prokaryotes. Systematic and Applied Microbiology 42: 5–14, https://www.sciencedirect.com/science/article/pii/S07232020183016 81?via%3Dihub.

Ruppert, K. M., R. J. Kline & M. S. Rahman, 2019. Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. Global Ecology and Conservation 17: e00547. https://doi.org/10.1016/j.gecco.2019.e00547.

Salmaso, N., 2019. Effects of habitat partitioning on the distribution of bacterioplankton in deep lakes. Frontiers in Microbiology 10: 2257. https://doi.org/10.3389/fmicb.2019.02257.

Salmaso, N., G. Riccioni, M. Pindo, V. Vasselon, I. Domaizon & R. Kurmayer, 2021. Metabarcoding protocol – analysis of bacteria (including cyanobacteria) using the 16S rRNA gene and a DADA2 pipeline (Version 1), *Zenodo*, https://doi.org/10.5281/zenodo.5232772.

Salmaso, N., V. Vasselon, F. Rimet, M. Vautier, T. Elersek, A. Boscaini, C. Donati, M. Moretto, M. Pindo, G. Riccioni, E. Stefani, C. Capelli, F. Lepori, R. Kurmayer, U. Mischke, A. K. Klemenčič, K. Novak, C. Greco, G. Franzini, G. Fusato, F. Giacomazzi, A. Lea, S. Menegon, C. Zampieri, A. Macor, D. Virgilio, E. Zanut, R. Zorza, F. Buzzi & I. Domaizon, 2022. DNA sequence and taxonomic gap analyses to quantify the coverage of aquatic cyanobacteria and eukaryotic microalgae in reference databases: Results of a survey in the Alpine region. Science of the Total Environment 834: 155175. https://doi.org/10.1016/j.scitotenv.2022.155175.

Salmaso, N., A. Boscaini, C. Capelli, L. Cerasino, M. Milan, S. Putelli, & M. Tolotti, 2015a. Historical colonization patterns of *Dolichospermum lemmermannii* (Cyanobacteria) in a deep lake south of the Alps. Advances in Oceanography and Limnology 6: 33–45, http://pagepressjournals.org/index.php/aiol/article/view/5456.

Salmaso, N., C. Capelli, S. Shams, & L. Cerasino, 2015b. Expansion of bloom-forming *Dolichospermum lemmermannii* (Nostocales, Cyanobacteria) to the deep lakes south of the Alps: Colonization patterns, driving forces and implications for water use. Harmful Algae 50: 76–87, http://www.sciencedirect.com/science/article/pii/S1568988315300561.

Salmaso, N., L. Cerasino, A. Boscaini, & C. Capelli, 2016. Planktic *Tychonema* (Cyanobacteria) in the large lakes south of the Alps: Phylogenetic assessment and toxigenic potential. FEMS Microbiology Ecology 92: fiw155, http://www.ncbi.nlm.nih.gov/pubmed/27402712.

Sandzewicz, M., N. Khomutovska, Ł. Łach, J. Kwiatowski, T. Niyatbekov, M. Suska-Malawska, & I. Jasser, 2023. Salinity matters the most: How environmental factors shape the diversity and structure of cyanobacterial mat communities in high altitude arid ecosystems. Frontiers in Microbiology 14, https://doi.org/10.3389/fmicb.2023.1108694.

Scherer, P. I., A. D. Millard, A. Miller, R. Schoen, U. Raeder, J. Geist & K. Zwirglmaier, 2017. Temporal dynamics of the microbial community composition with a focus on toxic cyanobacteria and toxin presence during harmful algal blooms in two south German lakes. Frontiers in Microbiology 8: 2387. https://doi.org/10.3389/fmicb.2017.02387.

Schirrmeister, B. E., D. A. Dalquen, M. Anisimova, & H. C. Bagheri, 2012. Gene copy number variation and its significance in cyanobacterial phylogeny. BMC Microbiology 12: 177, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3552681/.

Senavirathna, M. D. H. J. & M. A. D. D. Jayasekara, 2023. Temporal variation of 2-MIB and geosmin production by *Pseudanabaena galeata* and *Phormidium ambiguum* exposed to high-intensity light. Water Environment Research 95: e10834. https://doi.org/10.1002/wer.10834.

Shams, S., C. Capelli, L. Cerasino, A. Ballot, D. R. Dietrich, K. Sivonen & N. Salmaso, 2015. Anatoxin-a producing *Tychonema* (Cyanobacteria) in European waterbodies. Water Research 69: 68–79.

Soo, R. M., J. Hemp, D. H. Parks, W. W. Fischer, & P. Hugenholtz, 2017. On the origins of oxygenic photosynthesis and aerobic respiration in Cyanobacteria. Science 355: 1436–1440, http://www.ncbi.nlm.nih.gov/pubmed/28360330.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30: 1312–1313. https://doi.org/10.1093/bioinformatics/btu033.

Stevenson, J. R., M. I. Bothwell, & R. L. Lowe, 1996. Algal Ecology: freshwater benthic ecosystems. Academic Press, Elsevier, https://www.elsevier.com/books/algal-ecology/stevenson/978-0-12-668450-6.

Stewart, I., A. A. Seawright, & G. R. Shaw, 2008. Cyanobacterial poisoning in livestock, wild mammals and birds--an overview. Advances in Experimental Medicine and Biology 619: 613–637, http://www.ncbi.nlm.nih.gov/pubmed/18461786.

Stoddard, S. F., B. J. Smith, R. Hein, B. R. K. Roller, & T. M. Schmidt, 2015. rrnDB: Improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. Nucleic Acids Research 43: D593–D598, http://www.ncbi.nlm.nih.gov/pubmed/25414355.

Strunecký, O., A. P. Ivanova & J. Mareš, 2023. An updated classification of cyanobacterial orders and families based on phylogenomic and polyphasic analysis. Journal of Phycology 59: 12–51. https://doi.org/10.1111/jpy.13304.

Strunecký, O., J. Komárek, J. Johansen, A. Lukešová, & J. Elster, 2013. Molecular and morphological criteria for revision of the genus *Microcoleus* (Oscillatoriales, Cyanobacteria). Journal of Phycology 49: 1167–1180, http://www.ncbi.nlm.nih.gov/pubmed/27007635.

Sukenik, A., O. Hadas, A. Kaplan & A. Quesada, 2012. Invasion of nostocales (cyanobacteria) to subtropical and temperate freshwater lakes – physiological, regional, and global driving forces. Frontiers in Microbiology 3: 86. https://doi.org/10.3389/fmicb.2012.00086.

Suurnäkki, S., G. V. Gomez-Saez, A. Rantala-Ylinen, J. Jokela, D. P. Fewer, & K. Sivonen, 2015. Identification of geosmin and 2-methylisoborneol in cyanobacteria and molecular detection methods for the producers of these compounds. Water Research 68: 56–66, https://www.sciencedirect.com/science/article/pii/S0043135414006678.

Taton, A., S. Grubisic, E. Brambilla, R. De Wit, & A. Wilmotte, 2003. Cyanobacterial diversity in natural and

artificial microbial mats of Lake Fryxell (McMurdo Dry Valleys, Antarctica): a morphological and molecular approach. Applied and Environmental Microbiology 69: 5157–5169, http://aem.asm.org/content/69/9/5157.short.

Tett, P., C. Gallegos, M. G. Kelly, G. M. Hornberger & B. J. Cosby, 1978. Relationships among substrate, flow, and benthic microalgal pigment density in the Mechums River, Virginia. Limnology and Oceanography 23: 785–797. https://doi.org/10.4319/lo.1978.23.4.0785.

Timoshkin, O. A., D. P. Samsonov, M. Yamamuro, M. V. Moore, O. I. Belykh, V. V. Malnik, M. V. Sakirko, A. A. Shirokaya, N. A. Bondarenko, V. M. Domysheva, G. A. Fedorova, A. I. Kochetkov, A. V. Kuzmin, A. G. Lukhnev, O. V. Medvezhonkova, A. V. Nepokrytykh, E. M. Pasynkova, A. E. Poberezhnaya, N. V. Potapskaya, N. A. Rozhkova, N. G. Sheveleva, I. V. Tikhonova, E. M. Timoshkina, I. V. Tomberg, E. A. Volkova, E. P. Zaitseva, Yu. M. Zvereva, A. B. Kupchinsky, & N. A. Bukshuk, 2016. Rapid ecological change in the coastal zone of Lake Baikal (East Siberia): Is the site of the world's greatest freshwater biodiversity in danger? Journal of Great Lakes Research 42: 487–497, https://linkinghub.elsevier.com/retrieve/pii/S0380133016000393.

van Breemen, L. W. C. A., H. A. M. Ketelaars & P. VisserEbbeng, 1991. A new method to control growth of geosmin producing benthic cyanobacteria. SIL Proceedings 1922–2010(24): 2168–2173. https://doi.org/10.1080/03680770.1989.11899918.

Vasselon, V., F. Rimet, K. Tapolczai, & A. Bouchez, 2017. Assessing ecological status with diatoms DNA metabarcoding: Scaling-up on a WFD monitoring network (Mayotte island, France). Ecological Indicators Elsevier B.V. 82: 1–12, https://doi.org/10.1016/j.ecolind.2017.06.024.

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

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

Venables, W. N., & B. D. Ripley, 2002. Modern Applied Statistics with S. Springer, New York, NY, https://doi.org/10.1007/978-0-387-21706-2.

Větrovský, T. & P. Baldrian, 2013. The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. PLoS ONE Public Library of Science 8: e57923. https://doi.org/10.1371/journal.pone.0057923.

van Vliet, M. T. H., W. H. P. Franssen, J. R. Yearsley, F. Ludwig, I. Haddeland, D. P. Lettenmaier, & P. Kabat, 2013. Global river discharge and water temperature under climate change. Global Environmental Change 23: 450–464, https://www.sciencedirect.com/science/article/pii/S0959378012001331.

Wang, C., T. Zhang, Y. Wang, L. A. Katz, F. Gao & W. Song, 2017. Disentangling sources of variation in SSU rDNA sequences from single cell analyses of ciliates: Impact of copy number variation and experimental error.

Proceedings of the Royal Society B: Biological Sciences 284: 20170425. https://doi.org/10.1098/rspb.2017.0425.

Wehr, J. D. & R. G. Sheath, 2003. Freshwater Algae of North America - Ecology and Classification, Academic Press, Elsevier:

Wetzel, R. G., 2001. Limnology. Lake and river ecosystems, Academic Press, San Diego:

Wetzel, R. G. & G. Likens, 2000. Limnological Analyses, Springer, New York:

Wijewardene, L., N. Wu, N. Fohrer, & T. Riis, 2022. Epiphytic biofilms in freshwater and interactions with macrophytes: Current understanding and future directions. Aquatic Botany 176: 103467, https://www.sciencedirect.com/science/article/pii/S0304377021001169.

Wilmotte, A., H. D. I. Laughinghouse, C. Capelli, R. Rippka, & N. Salmaso, 2017. Taxonomic identification of cyanobacteria by a polyphasic approach In Kurmayer, R., K. Sivonen, A. Wilmotte, & N. Salmaso (eds), Molecular Tools for the Detection and Quantification of Toxigenic Cyanobacteria. John Wiley: 79–119.

Wood, S. A., A. I. Selwood, A. Rueckert, P. T. Holland, J. R. Milne, K. F. Smith, B. Smits, L. F. Watts & C. S. Cary, 2007. First report of homoanatoxin-a and associated dog neurotoxicosis in New Zealand. Toxicon 50: 292–301.

Wood, S. A., M. W. Heath, P. T. Holland, R. Munday, G. B. McGregor & K. G. Ryan, 2010. Identification of a benthic microcystin-producing filamentous cyanobacterium (Oscillatoriales) associated with a dog poisoning in New Zealand. Toxicon 55: 897–903.

Yu, G., 2020. Using ggtree to visualize data on tree-like structures. Current Protocols in Bioinformatics 69: e96. https://doi.org/10.1002/cpbi.96.

Zancarini, A., I. Echenique-Subiabre, D. Debroas, N. Taïb, C. Quiblier, & J.-F. Humbert, 2017. Deciphering biodiversity and interactions between bacteria and microeukaryotes within epilithic biofilms from the Loue River, France. Scientific Reports 7: 4344, https://www.nature.com/articles/s41598-017-04016-w.

Zębek, E., A. Napiórkowska-Krzebietke, A. Świątecki & D. Górniak, 2021. Biodiversity of periphytic cyanobacteria and algae assemblages in polar region: a case study of the vicinity of Arctowski Polish Antarctic Station (King George Island, Antarctica). Biodiversity and Conservation 30: 2751–2771. https://doi.org/10.1007/s10531-021-02219-2.

Zhang, X., Z. Pu, Y. Li & X. Han, 2016. Stochastic processes play more important roles in driving the dynamics of rarer species. Journal of Plant Ecology 9: 328–332. https://doi.org/10.1093/jpe/rtv058.

Zubia, M., C. Vieira, K. A. Palinska, M. Roué, J.-C. Gaertner, I. Zloch, M. Grellier & S. Golubic, 2019. Benthic cyanobacteria on coral reefs of Moorea Island (French Polynesia): diversity response to habitat quality. Hydrobiologia 843: 61–78. https://doi.org/10.1007/s10750-019-04029-8.