Hydrobiologia
https://doi.org/10.1007/s10750-026-06250-8

PHYTOPLANKTON SPATIOTEMPORAL PATTERNS

Shotgun metagenomics uncovers novel features of the cyanobacterial community in a large, shallow lake in Central Italy

Nico Salmaso · Leonardo Cerasino · Margherita Di Brizio · Massimo Pindo · Adriano Boscaini

Received: 5 January 2026 / Revised: 14 May 2026 / Accepted: 23 May 2026

Abstract This study aimed to unravel the taxonomic structure and functional potential of the cyanobacterial community in a large, shallow mesoeutrophic lake in Central Italy (Lake Trasimeno) through shotgun metagenomic analysis. The metagenomic profiling of samples allowed to characterize the structure of the bacterial community, also including the identification of rare species, not detected with the determination of metagenome-assembled genomes (MAGs), in the successive assembling and binning steps. The taxonomic identification of MAGs was based on the computation of genome similarity metrics (average nucleotide and amino acid identity values; ANI, AAI) and phylogenomic analyses, which confirmed the presence of several cyanobacteria previously identified during the monitoring campaigns, and the identification of taxa new for the Mediterranean region, such as *Prochlorothrix hollandica* and an uncharacterized *Prochlorothrix*. The genome annotation of the *Prochlorothrix* MAGs was consistent with that of oxygenic, non-diazotrophic, chlorophyll-*b*-producing, photosynthetic cyanobacteria. The annotation of cyanobacterial MAGs did not reveal the presence of operons encoding cyanotoxins. Based on a conservative analysis, the presence of antimicrobial resistance genes (ARGs) was exclusively identified in *Vibrio cholerae* (Enterobacterales) and in a few unbinned contigs, confirming the global relevance of the ARGs spread in freshwater environments.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10750-026-06250-8.

Guest Editors: Petr Znachor, Pavel Rychtecký, Eliška Kozlíková-Zapomělová, Luigi Naselli-Flores & Judit Padisák / Phytoplankton at Various Temporal and Spatial Scales.

N. Salmaso (✉) · L. Cerasino · M. Pindo · A. Boscaini
Research and Innovation Centre, Fondazione Edmund Mach, Via Edmund Mach, 1, 38098 San Michele all'Adige, Italy
e-mail: nico.salmaso@fmach.it

L. Cerasino
e-mail: leonardo.cerasino@fmach.it

M. Pindo
e-mail: massimo.pindo@fmach.it

A. Boscaini
e-mail: adriano.boscaini@fmach.it

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

M. Di Brizio
UOLM sito Perugia, ARPA Umbria, Via Pievaiola 207/B-3, 06132 Perugia, Italy
e-mail: m.dibrizio@arpa.umbria.it

Published online: 09 June 2026

Keywords Metagenomics · Cyanobacteria · *Prochlorothrix* · Phylogenomics · Pangenomics · Antimicrobial resistance genes

Introduction

The combination of morphology and molecular data, as well as ecological and physiological characteristics, represents a gold standard for a comprehensive and accurate classification of cyanobacteria (Komárek, 2016; Rzymski et al., 2017; Kozlíková-Zapomělová et al., 2025). The polyphasic approach is limited to the study of isolated and cultured specimens, single individuals or colonies, and monospecific environmental samples (Komárek, 2015; Capelli et al., 2017; Zubia et al., 2019; Pokorný et al., 2023). This approach is fully coherent with the requirements of the International Code of Nomenclature of Prokaryotes (ICNP) and the International Code of Nomenclature for algae, fungi, and plants (ICN), which require that the designation of type materials, phylogenetic analyses, and taxonomic identifications is based on isolated strains. An advantage of this approach is the possibility to associate the classified species with morphotypes, allowing their discrimination at the microscope, even if this is restricted to the species possessing a suitable number of phenotypic diacritical characters, and therefore excluding most of the picocyanobacteria (but see Komárek et al., 1999) and many other simple forms. The association of species names to recognizable morphotypes allows the direct and measurable identification of functional traits and physiological functions of ecological relevance, such as shapes and dimensions, maximum growth rates, environmental tolerances, grazer susceptibility, and metabolite production (including cyanotoxins) (Litchman et al., 2007; Salmaso et al., 2015; Liu et al., 2022). The requirement of isolated specimens is an important bottleneck in the use of the polyphasic approach since many unculturable organisms will remain undetected. Similarly, owing to the laboratory time required for the isolation and analysis, many species have not yet been isolated and cultivated (Rosselló-Móra & Whitman, 2019; Hay et al., 2023; Pessi et al., 2023).

In the last two decades, the use of high-throughput sequencing (HTS) technologies has paved the way for the study of environmental DNA using a range of techniques specifically adapted to the characterization of organisms using culture-independent methods. Amplicon sequencing (metabarcoding) approaches have been widely used for the characterization of organisms from microbes to mammals (Ruppert et al., 2019; Compson et al., 2020). Owing to the relatively easy implementation and affordable costs of metabarcoding, several large-scale investigations on the distribution of cyanobacteria have been carried out in a range of aquatic environments (e.g., among others, MacKeigan et al., 2022; Nicolosi Gelis et al., 2024; Salmaso et al., 2024a; Vettorazzo et al., 2024; Moretto et al., 2025). Nevertheless, the short length and resulting low taxonomic resolution of the marker genes (mostly rRNA) after bioinformatic processing (less than ca. 400 bp) do represent important drawbacks, intrinsically limiting taxonomical classification to the genus level and impeding a thorough functional analysis. The parallel introduction of genomic and metagenomic methods based on either short-read assembly (whole-genome shotgun sequencing) or long-read assembly (PacBio and Oxford Nanopore Technologies [ONT]) has enabled the reconstruction of draft or circular cyanobacterial genomes in isolates (e.g., Velichko et al., 2016; Nitnaware et al., 2021) or complex microbial communities (e.g., Cissell & McCoy, 2021; Valadez-Cano et al., 2025), providing information on taxonomy at the strain level and metabolic functions through genome annotation (Leao et al., 2017; Bullerjahn et al., 2023; Cai et al., 2023; Strunecký et al., 2023; Lefler et al., 2025). Besides identification of fundamental metabolic and functional patterns (Ramond et al., 2025), the annotation of cyanobacterial genomes from water bodies has enabled the identification of several biosynthetic gene clusters (BGCs) associated with the production of a wide range of cyanotoxins and other bioactive secondary metabolites, as well as a number of genes and operons linked to antimicrobial resistance acquisition in bacteria (Bashir et al., 2023; Khatiebi et al., 2024; Salmaso et al., 2024b; Stern et al., 2025). Overall, though based on genomic analysis (and not on gene expression), these discoveries opened new perspectives in the assessment of potential risks for human health, which would otherwise be more difficult to appraise using traditional analytical methods or (in the case of cyanotoxins BGCs) less efficient if based on

the analysis of single or only a few gene markers (Timms et al., 2023; Djordjevic et al., 2024; Pereira et al., 2026). In this study, we report the results of a comprehensive shotgun metagenomic analysis conducted on environmental samples collected during the summer of 2023 from two littoral stations of Lake Trasimeno, in Central Italy. The lake is mesoeutrophic and strongly dominated by cyanobacteria, principally *Raphidiopsis raciborskii* (Woloszynska), Aguilera et al., whose presence was described in the lake for the first time in 1995 (Manti et al., 2005; Mugnai et al., 2008). A characterization of the taxonomy and metabolic functions of this species has been recently performed on a cultivated strain isolated during the same sampling campaign of this study from a sample collected at Castiglione del Lago (Salmaso et al., 2025). Besides *R. raciborskii*, many other cyanobacteria have been identified during routine sampling for the control of bathing waters (https://apps.arpa.umbria.it/). Most of these presented a challenge for taxonomic determination based on the use of conventional approaches, while others were overlooked, such as *Prochlorothrix hollandica* Burger-Wiersma, Stal & Mur, which was identified for the first time during this investigation. Among cyanobacteria, and along with a few other genera, this species possesses a unique photosynthetic apparatus, lacking phycobilisomes and including, besides chlorophyll-*a* (Chl-*a*), chlorophyll-*b* (Chl-*b*), which is typical of eukaryotic photosynthetic organisms. The discovery of this species at the northern border of the Italian Mediterranean region was unexpected, because, until now, the known distribution of this species was circumscribed to Northern Europe.

The main objective of this study is the taxonomic and functional characterization of selected cyanobacterial taxa identified in Lake Trasimeno during the summer period. Specific tasks include: (i) the metagenomic characterization of the cyanobacterial community and taxonomical identification of the main bacterial groups; (ii) the taxonomic assessment and functional annotation of secondary metabolites (including BGCs encoding cyanotoxins) of the cyanobacterial metagenome-assembled genomes (MAGs); (iii) the phylogenomic, functional, and pangenomic assessment of two *Prochlorothrix* taxa new for the Mediterranean area; (iv) the identification of antimicrobial resistance (AMR) genes (ARGs) in the bacterial and cyanobacterial MAGs and DNA contigs.

Materials and methods

Study site

Lake Trasimeno is a large (surface, ca. 124 km^2; volume, ca. 586×10^6 m^3) and shallow (z_{max} ca. 6 m) lake located at 258 m a.s.l. in Central Italy. The lake basin falls within the northern inland region of the Mediterranean-type climatic zone (Csa) in Italy (Beck et al., 2023) (https://koppen.earth/). The basin is endorheic and the water level and lake size can vary considerably in response to atmospheric precipitation. Information on the long-term trend of water temperatures and main limnological variables has been reported in Ludovisi & Gaino (2010), Pareeth et al. (2017), Morabito et al. (2018), and Ludovisi et al. (2021).

Sampling stations and field and laboratory measurements

Samplings were carried out on September 12, 2023, at the boat piers located at Castiglione del Lago (CAST) (43°07′23.6″ N 12°03′26.5″ E) and San Feliciano (SFEL) (43°07′02.4″ N 12°09′53.6″ E), which are 8.5 km apart (Supplementary Fig. 1). Samples were collected at the surface using a sterilized bucket; as for cyanotoxins, further samples were collected by concentrating the biomass using 80-µm mesh nets. Subsamples for the cyanotoxins and metagenomic analyses were stored in refrigerated conditions until filtration, performed on the successive day. Subsamples (250 ml) for phytoplankton analyses were fixed after collection with Lugol's solution and successively stored at 4 °C. At the sampling points, water temperatures were measured using a handheld thermometer and transparency using a Secchi disk.

In laboratory, a complete set of physical and water chemical analyses was performed by the local environmental agency (ARPA Umbria) on two pelagic samples collected the day before (September 11, 2023) off the coasts of Passignano sul Trasimeno (TRS30; 43°09′17.1″ N 12°06′46.6″ E) and Magione (TRS35; 43°06′18.0″ N 12°09′34.0″ E) (Supplementary Fig. 1). Results have been reported in Salmaso et al. (2025), Table 1.

The methods used for the microscopical observations in the CAST and SFEL stations, and cyanotoxins analyses, i.e., microcystins (MCs), anatoxins (ATXs), cylindrospermopsins (CYNs), and saxitoxins (STXs), were described in Salmaso et al. (2025) and Rott et al. (2007).

DNA extraction and sequencing

Samples were filtered through 0.2-µm IsoporeTM GTTP polycarbonate membranes (Merck Millipore). After DNA extraction using PowerWater Kit (Qiagen), the library was prepared using the Kapa HyperPlus Kit (Roche) and processed by 150 bp paired-end sequencing on an Illumina NovaSeq 6000 (Illumina Inc., San Diego, CA, USA) (Salmaso et al., 2024b).

Read quality checking and community metagenomic profiling

All the bioinformatic analyses were performed separately on the two individual samples. Reads quality checking, removal of residual adapters, and trimming, keeping only reads with quality scores greater than 20, were performed with FastQC 0.12.1 (github.com/s-andrews/FastQC) and bbduk 39.26 (https://jgi.doe.gov). Removal of human DNA contaminants was performed by mapping reads on the human reference genome GRCh38.p14 using Bowtie 2.5.4 (Langmead & Salzberg, 2012) and successive filtering with SAMtools 1.21 (Danecek et al., 2021).

A read-based metagenomic profiling was performed using a sketch-based *k*-mer containment (Koslicki & Zabeti, 2019) method implemented by Sylph 0.8.1 (Shaw & Yu, 2024; Rozday, 2025). The approach used in Sylph allows for estimating the fraction of sample *k*-mers in the DNA reads that are present in a reference genome; these values can then be used to calculate containment ANI and relative abundances (Sun et al., 2021). From the output of Sylph, the taxonomic profiles were obtained using Sylph-tax 1.5.1 (github.com/bluenote-1577/sylph-tax) and the GlobDB-r226 taxonomic reference database (Speth et al., 2025).

Assembling, binning, and MAGs classification

Assembling and binning followed the procedures described in detail in Salmaso et al. (2025), using the latest versions of the packages and updates, as briefly described below. Paired filtered reads were corrected with metaSPAdes 4.2.0 (–only-error-correction) (Nurk et al., 2017) and assembled with megahit 1.2.9 (–presets meta-sensitive) (Li et al., 2015). Contigs shorter than 1000 bp were discarded using anvi'o 8 (Eren et al., 2021). After binning using CONCOCT 1.1.0 (Alneberg et al., 2014), MetaBAT 2.18 (Kang et al., 2019), and SemiBin2 2.2.0 (Pan et al., 2022), the bins were filtered and combined using DAS Tool 1.1.7 (Sieber et al., 2018). Presence of chimerism and contaminations were checked and removed using GUNC 1.0.6 (Orakov et al., 2021) and MDMcleaner 0.8.7 (Vollmers et al., 2022). MAGs were then classified using GTDB-Tk 2.5.2, using the 10-RS226 GTDB taxonomy (Chaumeil et al., 2022) and skani 0.3.1 (Shaw & Yu, 2023) for ANI screening. Cyanobacterial MAGs were further checked for contaminations using FCS adaptor and FCS-GX 0.5.5 (Astashyn et al., 2024). Taxonomical classification of cyanobacterial MAGs was further assessed by computing ANI values using skani 0.3.1 (Shaw & Yu, 2023) against a complete set of "Cyanobacteriota" genomes downloaded from GTDB RS226. In the two samples, the resulting MAGs were de-replicated, choosing the best representative genome for each group of highly similar genomes using dRep 3.6.2 (last updated 2025) (Olm et al., 2017); in the selection process, dRep retains only dereplicated genomes with a minimum completeness of 75% and maximum contamination of 25% as computed using CheckM v1.2.4 (Parks et al., 2015). Completeness and contamination of MAGs were finally estimated using CheckM2 1.1.0 (Chklovski et al., 2023) and coverages using CoverM 0.7.0 (github.com/wwood/CoverM). The Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the project number PRJNA1395559; besides high-throughput sequencing reads, the project includes high-quality *Prochlorothrix* and *Raphidiopsis* cyanobacterial draft genomes.

Phylogenomic and pangenomic analysis of *Prochlorothrix* MAGs

The MAGs classified under the *Prochlorothrix* genus by GTDB-Tk were further analyzed and compared with a selection of genomes of the genus *Prochlorothrix* obtained from the curated GTDB RS226 database (Parks et al., 2022) and from NCBI (Schoch

et al., 2020) (Supplementary Table 1). Considering the implications of microbial genome completeness for metagenomic functional inference (Eisenhofer et al., 2023), the final list only included genomes with completeness $\geq 80\%$ and contamination $< 6\%$ estimated using Checkm2.

Average nucleotide identity between genomes was computed using ANI_b (blast-based) with pyANI-plus 1.0.0. (Pritchard et al., 2016). Additionally, potential discrimination at the level of genus was also evaluated by computing the average amino acid identity (AAI) using EzAAI 1.2.4 (Kim et al., 2021). When compared with ANI, AAI showed better resolution in describing taxonomic structure beyond the species rank (Konstantinidis & Tiedje, 2005). Phylogenomic analyses were performed using GToTree 1.8.16 based on pre-packaged HMM single-copy genes set specific for Cyanobacteria (251 genes) (Lee, 2019), alignment using muscle 5.1 (Edgar, 2022), and tree building using IQTree, with parameters -m MFP -B 1000 (Nguyen et al., 2015). The tree was rooted using ETE3 (Huerta-Cepas et al., 2016) and *Gloeomargarita lithophora* (Moreira et al.), strain Alchichica-D10 as outgroup. The final tree was annotated with the R package ggtree 3.10.1 (Yu, 2020).

Pangenome analysis was performed following the anvi'o 8 pangenomic workflow (Delmont et al., 2018; Eren et al., 2021), as implemented in Salmaso et al. (2025); after protein sequence extraction from genomes, proteins were identified using KEGG and COG classifications. In the pangenome, gene clusters were identified using the option –mcl-inflation 5. A subset of single-copy core genes (SCGs) was extracted and used to build a phylogenomic tree using IQTree (-m MFP) with midpoint rooting using the ETE3 package. A second pangenomic analysis was computed by means of PIRATE 1.0.5 (Bayliss et al., 2019), based on annotation gff tables obtained by prokka 1.15.6 (Seemann, 2014) and using mcl = 5. The Heaps law ($n = kN^{\gamma}$) was applied to evaluate if the pangenome was open ($\gamma < 0$) or closed ($\gamma > 0$) (Tettelin et al., 2008). Heaps parameters were estimated using a sample-based rarefaction approach (Gotelli & Colwell, 2001) with 1000 genome random samples using the specaccum function in the R package vegan 2.7.1 and nonlinear least squares fit of the power law to the average values calculated for each distribution; confidence intervals were computed from the standard deviations with the option "polygon" and default "ci" parameter value (2) (Oksanen et al., 2025).

Functional annotations

Functional annotations of the draft genomes were performed using Bakta v1.11.4 with Database: v6.0, full (Schwengers et al., 2021), and the NCBI stand-alone software package Prokaryotic Genome Annotation Pipeline (PGAP) version 2025-05-06.build7983 (Li et al., 2020). Based on the Bakta/PGAP annotation tables, a selection of rRNA and housekeeping genes relevant for taxonomic assessment was extracted from the *Prochlorothrix* draft genomes; besides 5S, 16S, and 23S rRNA, marker genes did include *rbcX* (RuBisCO chaperone), *rplB* (large subunit ribosomal protein L2), *leuS* (leucyl-tRNA synthetase), *fusA* (elongation factor G), and RNA polymerase subunits alpha, beta, and gamma (*rpoA, rpoB*, *rpoC2, rpoC1*). Identification of ribosomal rRNA genes was further performed using Barrnap 0.9 (github.com/tseemann/barrnap). Presence of ARGs was assessed using AMRFinderPlus 4.2.5 with Database version: 2025-12-03.1 (Feldgarden et al., 2021).

Metabolism of the *Prochlorothrix* genomes was evaluated using the metabolism suite of programs in anvi'o 8 (Eren et al., 2021; Watson et al., 2023). Functional annotations used the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al., 2014), matching KEGG Orthology identifiers (KOs) to metabolic pathways (modules) (Kanehisa & Sato, 2020). Selected phenotypic traits were analyzed using KEGG pathway maps (Kanehisa et al., 2022). Considering the peculiar traits characterizing the photosynthetic machinery and antenna pigments in *Prochlorothrix*, specific attention was given to the identification of gene complexes involved in the biosynthesis of chlorophyll-*b* and phycobiliproteins. The essential and discriminant enzyme for Chl-*b* synthesis in *Prochlorothrix* and prochlorophytes is the chlorophyllide *a* oxygenase (CAO), which converts chlorophyllide-*a* to chlorophyllide-*b* (Averina et al., 2019); this enzyme is not present in other non-synthesizing Chl-*b* cyanobacteria. Furthermore, the primary light-harvesting system of *P. hollandica* is associated with the chlorophyll a/b-binding apoproteins encoded by the tightly linked *pcbABC* gene cluster (Nikolaitchik & Bullerjahn, 1998). More broadly, this operon

is also referred to as the CBP (Chl-binding protein) superfamily (Averina et al., 2019).

Secondary metabolite BGC profiles in the cyanobacterial genomes were predicted using antiSMASH 8.0.4 (default mode) (Blin et al., 2025). Identification of a selection of individual genes encoding legacy cyanotoxins (MCs, ATX-a, CYNs, and STXs) and odorous compounds (geosmin), was assessed using ISeqDb 0.0.6 (github.com/hts-tools/iseqdb).

Results

Environmental data and cyanotoxins analysis

In the two sampling sites, surface water temperatures were between 24.6 °C and 25.0 °C, while transparency was 0.3 m. In the two pelagic stations sampled by ARPA Umbria, water temperature and transparency showed equivalent values as those recorded in the two littoral stations, while the high values of soluble reactive phosphorus (17–18 µg P l^{-1}), total phosphorus (70–80 µg P l^{-1}), chlorophyll a (44–50 µg l^{-1}), pH (9), and dissolved oxygen saturation (110–126%) were consistent with the meso-eutrophic/eutrophic status of the lake. Nitrogen compounds always showed concentrations <0.1 mg l^{-1} (NO_3–N and NH_4–N) and <2 mg l^{-1} (total nitrogen) (Salmaso et al., 2025), suggesting potential limitation by N. Conductivity was between 1826 and 1841 µS cm^{-1} at 20 °C.

The analysis by LC–MS did not reveal measurable concentrations of cyanotoxins in the whole set of environmental and net samples collected in the two littoral stations.

Light microscopy

In the two littoral sampling sites analyzed in this work (CAST and SFEL), the phytoplankton community was completely dominated by *R. raciborskii*, with abundances between 22,600 filaments ml^{-1} (CL) and 37,200 filaments ml^{-1}. Other rare cyanobacteria, difficult to observe due to crowding and despite sample dilutions (1:4, 1:10), included species determined only qualitatively and belonging to *Anabaenopsis*, *Pseudanabaena*, *Planktothrix*, *Merismopedia*, as well as undetermined filaments (which, based on current metagenomic analyses, included the *P. hollandica* population), and Nostocales.

Read-based metagenomic profiling and identification of MAGs

In the two stations, the fraction of corrected raw DNA reads mapped on the reference genomes was between 27 and 28%. Average and standard deviations of the containment ANI values of all the taxa identified in the two stations were 97.6 ± 1.4 (CAST) and 97.8 ± 1.4 (SFEL).

Excluding the unmapped fraction, the contribution of the bacterial species identified in the two littoral stations is reported in Fig. 1A. The prokaryotic community in the two stations showed a coincident structure, with the dominant bacterial classes (>5%) mostly represented by a common presence of Nanopelagicaceae (including *Nanopelagicus* spp.), *Aquiluna* spp., and *Pontimonas* spp. (Actinomycetes), Burkholderiaceae, Steroidobacteraceae, *Pseudohongiella* spp., *Vibrio cholerae* Pacini, and *Rubrivivax* spp. (Gammaproteobacteria), *Pollutibacter* spp., Chitinophagaceae, Crocinitomicaceae, and Schleiferiaceae (Bacteroidia), *Fonsibacter* spp., *Candidatus Fonsibacter ubiquis* Henson et al., *Neoroseomonas* spp., Rhodobacteraceae, *Hyphomonas* spp. (Alphaproteobacteria), and Ilumatobacteraceae (Acidimicrobiia). Among the less abundant bacteria, it was worth highlighting the presence, in both stations, of Patescibacteriota.

The largest fraction of mapped reads in photosynthetic cyanobacteria indicated a clear dominance of *Raphidiopsis* / *R. raciborskii*, followed by *Nodosilinea* spp., *Prochlorothrix* (*P. hollandica*), *Pseudanabaena* spp., *Anabaenopsis arnoldii* Aptekarj, *Sphaerospermopsis kisseleviana* (Lemmermann) Zapomelová et al., and *Cyanobium* spp. (Fig. 1B). Besides unclassified species belonging to Microcoleacee, Geitlerinemataceae, and Cyanobiaceae, other rare taxa grouped under "Other_taxa" in Fig. 1B included *Planktothrix* spp. and *Planktothrix agardhii* (Gomont) Anagnostidis & Komárek, *Dolichospermum heterosporum* (Nygaard) Wacklin et al., *Cuspidothrix* sp., *Microcystis* sp., *Vulcanococcus* spp., *Woronichinia* sp., and *Umezakia ovalisporum* (Forti) McGregor et al.. In both stations, non-photosynthetic cyanobacteria were represented by Vampirovibrionia (Fig. 1B; Supplementary Table 2).

Overall, the identification of the cyanobacterial taxa distinguished by the metagenomic profiling was confirmed by the assembly and binning of MAGs, which included the main dominant taxa reported in Fig. 1B (Table 1). The list of MAGs did not include *S. kisseleviana* and Vampirovibrionia JACEUU01, which were discarded by dRep due to their low completeness (< 70%) in the CAST and SFEL stations, respectively. The cyanobacterial MAGs included in Table 1 are represented by both high-quality (completeness ≥ 90% and contamination ≤ 5%) and medium-quality draft genomes (completeness ≥ 50% and contamination ≤ 10%), except for *Nodosilinea* sp014697375, which was characterized by a higher contamination (Table 1). It is worth highlighting the very close correspondence between the relative contribution (sequence abundance) of genera estimated by Sylph and the coverage values of the corresponding draft genomes estimated by CoverM ($r = 0.94$; taking into account the unmapped fractions).

The cyanobacterial taxa common to the two samples collected at CAST and SFEL showed coincident genomic similarity (ANI_b between 0.988 and 0.997). The same high ANI_b values (0.997) were computed between the two draft genomes of the *R. raciborskii* assembled from the two environmental samples and the *R. raciborskii* LT_0923 strain isolated at CAST on the same date (Salmaso et al., 2025). The dominance of *R. raciborskii* documented by the microscopical analyses and metagenomic profiling was confirmed by the high coverage values of the respective MAGs computed in the two stations (Table 1).

In both stations, *Prochlorothrix* was identified with two different organisms, i.e., *P. hollandica* and an undetermined *Prochlorothrix* sp.. The two pairs of genomes in the CAST and SFEL stations were practically coincident ($ANI_b > 0.99$). However, only *P. hollandica* showed clear correspondence and high skani ANI values after comparisons with the corresponding GTDB strains ($ANI_b > 0.98$; Table 1). Conversely, owing to lower skani ANI values, the two undetermined *Prochlorothrix* genomes were classified at the genus level by GTDB-Tk using a classification method based on placement in a class-level tree (pplacer; Table 1) (Chaumeil et al., 2022).

Taxonomy and main functional characteristics of the Lake Trasimeno *Prochlorothrix* assemblies

The *P. hollandica* and *Prochlorothrix* sp. draft genomes assembled from the Lake Trasimeno samples were between 4.3 and 4.4 Mbp and 4.8 and 5.0 Mbp, while GC content values were 0.56 and 0.57, respectively (Table 1). The number of contigs was between 454 and 472 and 321 and 338, while the number of coding genes was between 3704 and 3815 and 3441 and 3522, respectively. Genome completeness and contamination computed with CheckM2 or CheckM1 (marker lineage, Cyanobacteria) were greater than 89-90% and lower than 1%, respectively.

Two identical copies of 5S rRNA were identified in each of the two *Prochlorothrix* sp. assemblies (Supplementary Table 3). By querying the wgs NCBI database, the 5S rRNA sequences corresponded to the *Prochlorothrix* sp. isolate Tbon (USA: Louisiana, Terrebonne Bay off Cocodrie) and to the *Prochlorothrix* sp. isolate 315547 (India: Chilika Lagoon, Odisha) (query length 100%, pident 99.1%). No complete or partial 16S rRNA and 23S rRNA were identified in the four draft genomes. The results obtained from the other marker genes fully confirmed the identity of *P. hollandica*, while the two *Prochlorothrix* assemblies showed only a moderate correspondence (pident around 90%) with other *Prochlorothrix* sampled in North America and Asia (Supplementary Table 3).

The reactions necessary for the basic metabolisms of photosynthetic cyanobacteria in the four draft genomes incorporated modules involved in oxygenic photosynthesis (photosystems II and I; modules M00161 and M00163), including beta-carotene biosynthesis (M00097), the reductive pentose phosphate cycle (Calvin cycle) (M00165), the tricarboxylic acid (Krebs) cycle (M00009), and glycolysis (M00001, M00002). Excluding the modules heavily incomplete (e.g., only represented by 1-block in multi-block modules), other modules were representative of the carbohydrate, energy, lipid, nucleotide, amino acid, and glycan metabolism, metabolism of cofactors and vitamins, and biosynthesis of terpenoids and polyketides.

The presence of the chlorophyllide a oxygenase (CAO) (orthology, K13600; enzyme, EC:1.14.13.122) was confirmed by the anvi'o metabolomic analysis in all four genomes (Supplementary Table 4A). Blast analyses showed full correspondence between the CAO sequences of *P. hollandica* from Lake

Table 1 Cyanobacterial MAGs identified in Lake Trasimeno, in the stations of (A) Castiglione del Lago and (B) San Feliciano

	GTDB-Tk						Skani (GTDB taxonomy)		CheckM2				CoverM
	Class	Order	Family	Genus	Species	Method	Taxon	ANI	Compl	Cont	Size	GC	Cov
(A)	Cyanobacteria	Cyanobacteriales	Microcoleaceae			pplacer			84.1	0.6	3.79	0.38	8
	Cyanobacteria	Cyanobacteriales	Nostocaceae	*Anabaenopsis*	*Anabaenopsis arnoldii*	ANI	*Anabaenopsis arnoldii*	98.7	78.6	4.1	3.14	0.42	9
	Cyanobacteria	Cyanobacteriales	Nostocaceae	*Raphidiopsis*	*Raphidiopsis brookii*	ANI	*Raphidiopsis brookii*[(1)]	99.61	94.1	2.7	3.16	0.4	238
	Cyanobacteria	PCC-9006	Prochlorotrichaceae	*Prochlorothrix*		pplacer			92.9	0.8	5.01	0.57	45
	Cyanobacteria	PCC-9006	Prochlorotrichaceae	*Prochlorothrix*	*Prochlorothrix hollandica*	ANI	*Prochlorothrix hollandica*	99.17	91.4	0.3	4.40	0.56	29
	Cyanobacteria	Phormidesmidales	Phormidesmidaceae	*Nodosilinea*		pplacer	*Nodosilinea* sp023264315	91.62	88.7	1.1	2.91	0.58	23
	Cyanobacteria	Phormidesmidales	Phormidesmidaceae	*Nodosilinea*	*Nodosilinea* sp014697375	ANI	*Nodosilinea* sp014697375	98.49	84.4	11.5	4.95	0.57	17
	Cyanobacteria	Pseudanabaenales	Pseudanabaenaceae	*Pseudanabaena*		pplacer	*Pseudanabaena* sp014696925	94.94	85.0	1.9	4.24	0.42	11
	Vampirovibrionia	LMEP-6097	LMEP-6097	LMEP-6097		pplacer	LMEP-6097 sp018970325	95.03	86.0	4.7	2.07	0.36	19
	Vampirovibrionia	Vampirovibrionales	JAJTHC01	JACEUU01		pplacer			70.6	2.0	1.94	0.43	8
(B)	Cyanobacteria	Cyanobacteriales	Nostocaceae	*Anabaenopsis*	*Anabaenopsis arnoldii*	ANI	*Anabaenopsis arnoldii*	98.7	81.5	7.7	3.72	0.42	10
	Cyanobacteria	Cyanobacteriales	Nostocaceae	*Raphidiopsis*	*Raphidiopsis brookii*	ANI	*Raphidiopsis brookii*[(1)]	99.6	95.4	2.3	3.12	0.4	205
	Cyanobacteria	PCC-9006	Prochlorotrichaceae	*Prochlorothrix*		pplacer			92.3	0.8	4.85	0.57	27
	Cyanobacteria	PCC-9006	Prochlorotrichaceae	*Prochlorothrix*	*Prochlorothrix hollandica*	ANI	*Prochlorothrix hollandica*	99.2	89.0	0.2	4.31	0.56	18
	Cyanobacteria	Pseudanabaenales	Pseudanabaenaceae	*Pseudanabaena*		pplacer	*Pseudanabaena* sp014696925	95.15	87.0	5.6	4.72	0.43	10
	Vampirovibrionia	LMEP-6097	LMEP-6097	LMEP-6097		pplacer	LMEP-6097 sp018970325	94.81	91.9	1.4	2.19	0.36	17

[(1)]Raphidiopsis raciborskii (Salmaso et al., 2025)

Classification was performed using GTDB-Tk 2.5.2 updated to use the RS226 taxonomy. In the "Method" column, ANI and pplacer indicate classification of the query genomes based on either ANI or their placement in the reference tree (Chaumeil et al., 2022). The ANI values were computed by skani using a complete set of "Cyanobacteriota" genomes downloaded from GTDB RS226; for distantly related genomes, computations were skipped. Completeness (Compl.), contamination (Cont.), genome size (Size, Mbp), and GC fraction were computed using CheckM2, while mean coverage (Cov.) was computed using CoverM

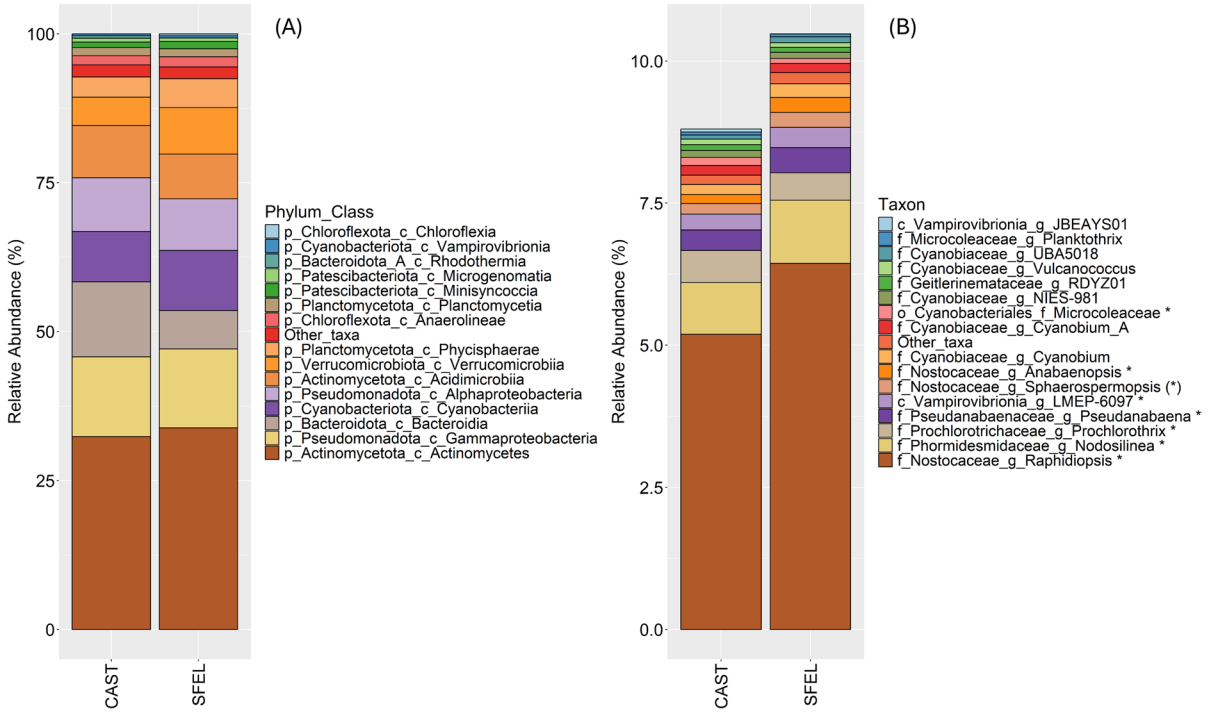

Fig. 1 Relative abundance of (**A**) the bacterial and (**B**) cyanobacterial community in the two sampling stations (CAST, Castiglione del Lago; SFEL, San Feliciano) of Lake Trasimeno. The fraction of unmapped reads (without taxonomic classification) is not included. Prefixes indicate the taxonomic ranks at the level of phylum (p_), class (c_), order (o_), family (f_), and genus (g_). In (B), the cyanobacterial taxa characterized at the genomic level are indicated by asterisks

Trasimeno and those characterized in *P. hollandica* AB126597.1 (Nagata et al., 2004), confirming the inclusion of CAO in the Rieske non-heme iron oxygenase superfamily (pident, 100%). The corresponding pident values computed for the two uncharacterized *Prochlorothrix* were lower (76–89%). As for the *pcb* genes, Bakta did indicate the presence of *Chlorophyll a/b light-harvesting protein* in the four genomes, with the exclusion of *Prochlorothrix*_LT_SFP (Supplementary Table 4B). After blasting the corresponding protein sequences on the UniProt KB reference proteomes, the results confirmed the presence of the PcbABC proteins in the *P. hollandica* genomes from Lake Trasimeno (with a pident value of 98.9–99.7%) and in the *Prochlorothrix* determined in the CAST station, with a lower pident value of 86.6–94.6%.

The complete set of genes involved in phycobilisome biosynthesis was not present in the four *Prochlorothrix* genomes. Indeed, no KOs corresponding to genes encoding allophycocyanin (AP) (*apcA-F*), phycocyanin (PC)/phycoerythrocyanin (PEC) (*cpcA-G*), and phycoerythrin (PE) (*cpeA-E*, *cpeR-S*, *cpeU*, *cpeY-Z*), were identified, with the exception of the presence of an isolated *cpeT* gene in all four genomes.

Nitrogen metabolism was maintained by assimilatory nitrate reduction (M00531), while genes involved in N-fixation were not identified. Nitrogen uptake was inferred based on the presence of specific ATP-binding cassette (ABC) transporter genes, in particular the nitrate/nitrite (*nrtABD*) transporter genes. Besides N, several other ABC transporters involved in the transport of nutrients, microelements, and organic molecules were identified. Among others, these included transporters for Phosphate and Phosphonate, α-Glucoside, Chitobiose and Phospholipids, amino acids and Urea, Iron and Molybdate, Lipo-oligosaccharide and Lipopolysaccharide. The potential uptake of osmoprotectants, which may confer tolerance to elevated osmotic stress in this ion-rich lake, was potentially sustained by an OpuBC-BB-BA ABC transport system.

Following the Bakta annotation, both *P. hollandica* draft genomes included genes encoding functional gas vesicles (*gvpA*, *gvpC, gvpN, gvpJ, gvpK*). A reduced and different set of genes (*gvpA*, *gvpC*, *gvpN*, *gvp*G) was detected in the two undetermined *Prochlorothrix* genomes.

Comparison of the main functional characteristics with other *Prochlorothrix* assemblies analyzed in this work

Overall, the comparison of the KEGG modules in the Lake Trasimeno draft genomes to the other assemblies considered in this study revealed no significant differences between the *P. hollandica* and the *Prochlorothrix* taxa. In particular, the CAO and the chlorophyll *a/b* binding light-harvesting protein were identified in all the assemblies, whereas complete complexes of genes involved in the biosynthesis of phycobilisomes or nitrogen fixations were not present.

Based on the Bakta analysis, all the *P. hollandica* showed a comparable set of *gvp* genes, including *gvpACNJK*, also complemented by *gvpG*. While all the assembled draft genomes from the Louisiana environmental samples (see Supplementary Table 1) showed only the presence of the *gvpG* gene, the other unclassified *Prochlorothrix* taxa contained combinations of *gvpACNG* genes. A broader functional analysis is included in the next two sections.

Phylogenomic and pangenomic analyses of *Prochlorothrix*

A clear separation between the *P. hollandica* and the undetermined *Prochlorothrix* assemblies emerged from the phylogenomic analysis (Fig. 2). The separation of the two clades was confirmed by the low interclade ANI_b (range, 0.76–0.78) and AAI (0.74–0.75) values. Nevertheless, while the intraclade ANI_b and AAI values were always over 0.99 in the *P. hollandica* clade, the corresponding ANI_b and AAI values in the *Prochlorothrix* clade were much lower and dispersed (ranges, 0.79–0.99 and 0.81–0.99, respectively), clearly indicating the existence of different subclades within this group (Fig. 2). It is interesting to observe that the two couples of MAGs assembled from the US samples and the two Indian assemblies were included in different subclades. In the tree, the presence of the CAO and *pcb* genes described in the previous sections was highlighted.

Overall, the *Prochlorothrix* pangenome was represented by a total of 52,502 genes grouped in 8772 gene clusters (Fig. 3). Single-copy core gene clusters (SCGs) and genes present in a single genome (singletons) were 7.7% and 24.4% of the total gene clusters, respectively. Total gene clusters annotated with COG functions were 71%, while SCGs and singletons showed an annotation rate of 87% and 24%, respectively. The phylogenomic analysis computed using the 677 SCGs was fully congruent with the corresponding tree based on the analysis of selected marker genes and with the heatmap of ANI_b values, which confirmed the separation of *P. hollandica* and *Prochlorothrix*, and the recognition of three subclades within the *Prochlorothrix* group (Figs. 2 and 3). Differences between the two main clades were supported by the presence of different groups of gene clusters in the pangenome graph, particularly apparent in the top-left sector of Fig. 3.

Based on and adapted from the pangenome types discussed in Matthews et al. (2024), the principal COG20 categories classified within the Soft Core genome (genes present in the majority of genomes) included genes involved in processes necessary for primary, non-dispensable physiological functions, such as energy production, translation and protein quality control, cytoskeleton (maintenance of cell structural integrity), coenzyme metabolism, and overall general metabolism (Supplementary Fig. 2). Conversely, in the Shell and Cloud pangenome, besides unknown (especially Cloud) and "general function prediction only" genes, the most striking contrast differences with the Soft Core pangenome were represented by Mobilome (prophages, transposons) and Defense Mechanisms, including toxin–antitoxin systems, CRISPR, phage resistance, surface structure variability. The Signal Transduction category was well represented in the three pangenome types (Supplementary Fig. 2), including over 260 COG20 different functions for adapting to various stimuli, such as, among others, light (bacteriophytochrome and light-regulated signal transduction histidine kinase), nitrogen availability (N-regulatory protein PII and GlnK), and redox sensors like thioredoxin, which were all widely represented in the Soft Core or Shell pangenome.

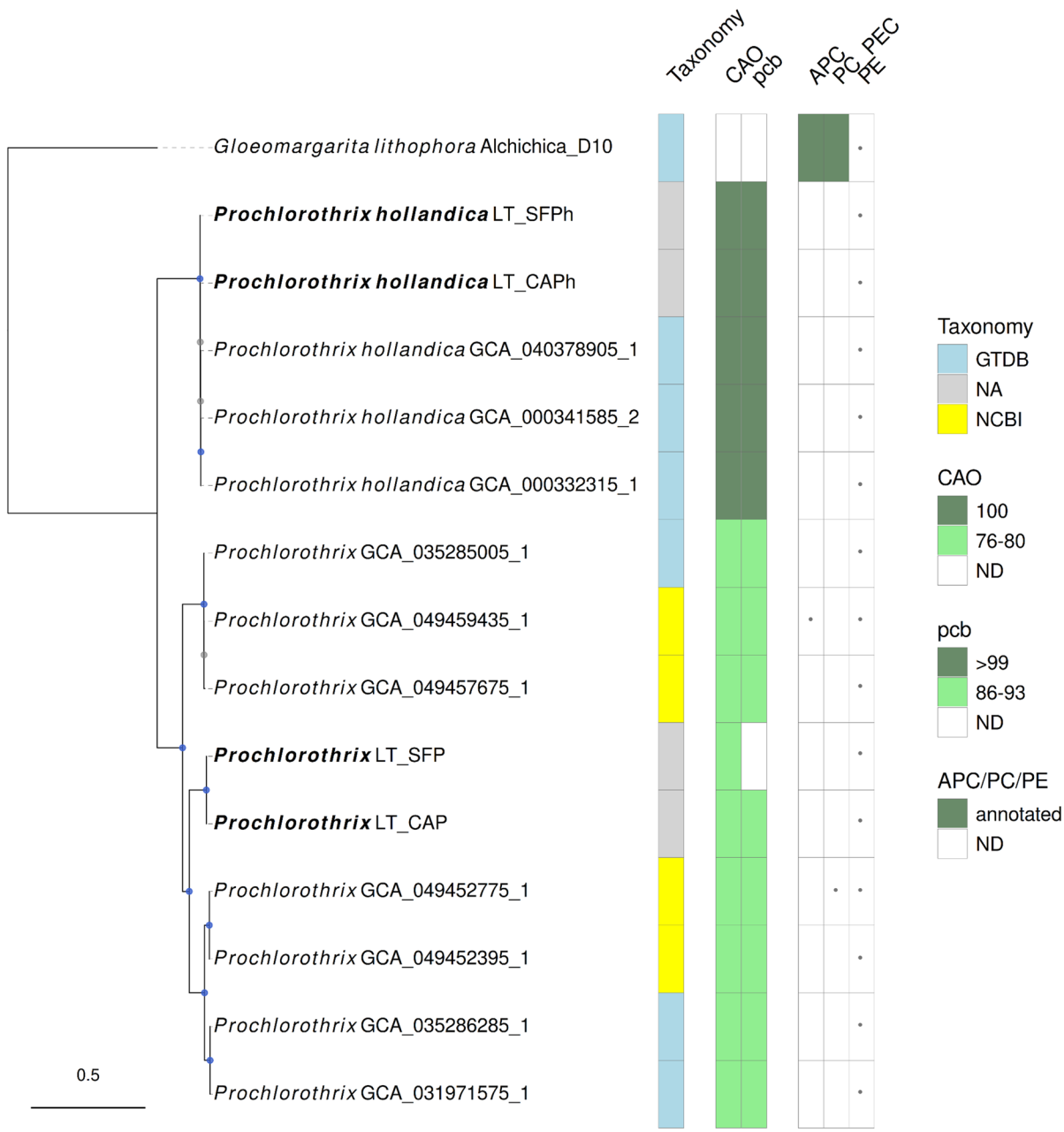

Fig. 2 Phylogenomic tree of *Prochlorothrix*, including a selection of genomes from the Genome Taxonomy Database (GTDB) and NCBI GenBank, and the four genomes assembled from the two stations of Lake Trasimeno, which are highlighted in bold; details on the taxa included in the analysis are reported in Supplementary Table 1; taxa are indicated with the genus or species name, and the NCBI accession codes. The color code columns CAO and pcb indicate the presence of chlorophyllide a oxygenase and *pcbABC* genes encoding chlorophyll a/b-binding apoproteins, respectively; the numbers in the legend resume the pident values reported in Supplementary Table 4 (ND, not detected). The presence of genes involved in the biosynthesis of phycobiliproteins are reported in the columns APC (allophycocyanin), PC_PEC (phycocyanin/phycoerythrocyanin), and PE (phycoerythrin); these complexes of genes were not detected in all the *Prochlorothrix* genomes, with the exclusion of single genes (indicated by small dots; see text). The tree was rooted with *G. lithophora* Alchichica D10 as outgroup. The UFBoot supporting values $\geq 98\%$ and $< 98\%$ are reported in blue and gray, respectively. The scale bar indicates the number of substitutions per site

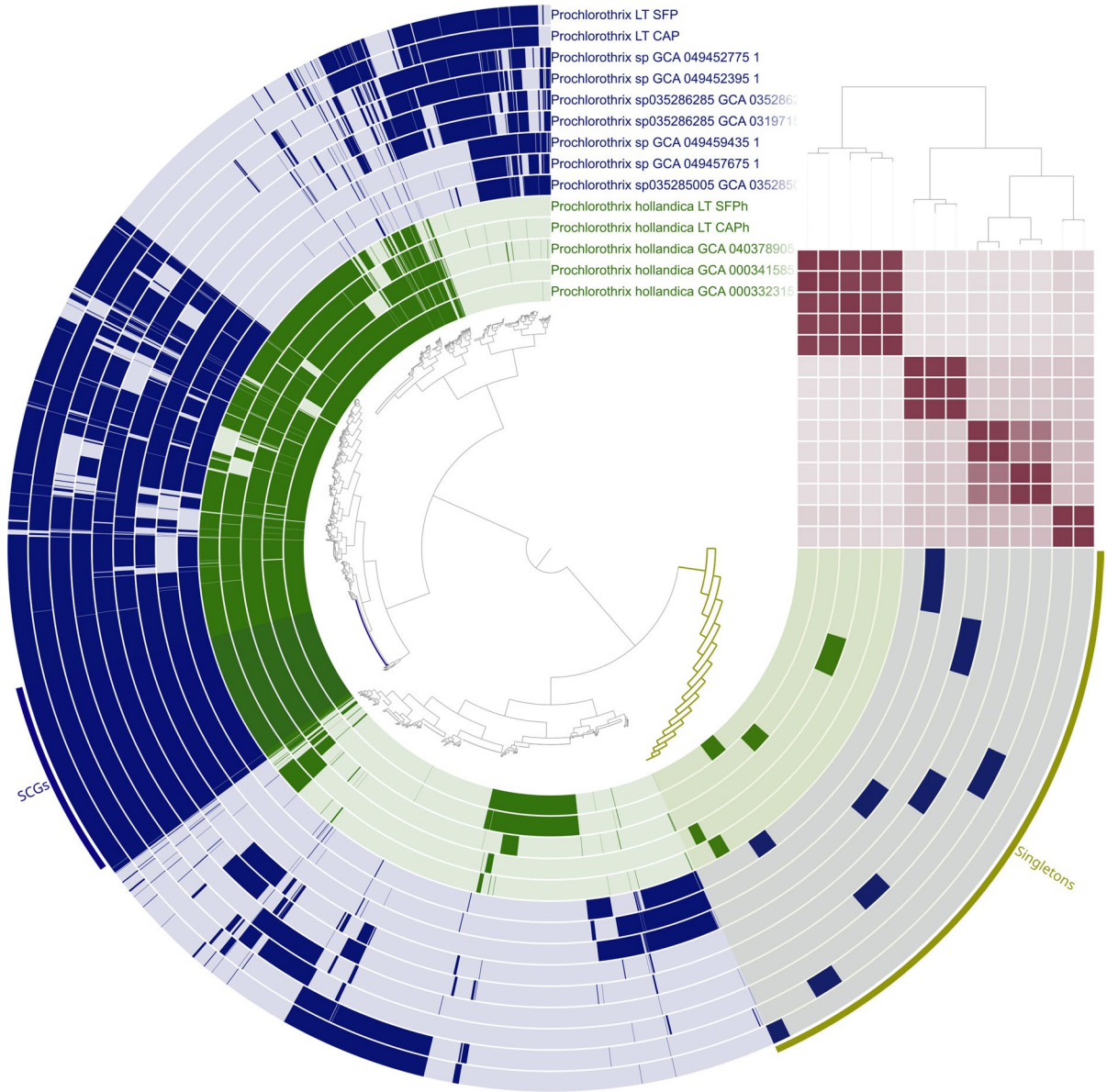

Fig. 3 The pangenome of *Prochlorothrix*. The 8772 gene clusters are represented by radii containing one or more amino acid sequences from one or more genomes, which are represented by the 14 circles. Gene clusters that occur together in the same group of genomes are closer together. Genomes are ordered consistently with the phylogenomic tree on the top right side of the figure, computed from the single-copy core genes (SCGs). The order is also consistent with the symmetrical matrix reporting the ANI_b values computed between all the 14 genomes. The outer circle highlights the two regions of the pangenome including the single-copy gene clusters (SCGs; core genome) and the singletons. Details on the taxa object of the pangenome analysis are reported in Supplementary Table 1

The comparison of the two main clades highlighted the presence of gene clusters involved in specific functions present exclusively in *Prochlorothrix* or *P. hollandica* (Supplementary Table 5). Considering that the enrichment test is not reliable when the number of cases is low, the table reports only the functions exclusively present in one of the two groups. Exclusive functions in *Prochlorothrix* showed a higher representation of signaling domains to sense environmental and intracellular

stimuli, inorganic ion transport and metabolism processes (glna and carbonic anhydrase), membrane functionality and transport (e.g., ABC permeases). Among others, exclusive metabolic functions in *P. hollandica* included osmoprotectant transport systems (OpuBA/OpuBB, OsmF) involved in the adaptation to osmotic fluctuations, proteins linked to cell-wall functionality and processes, intermembrane transporters (PqiABC; Mg_2^+ and Co_2^+ transporters), mobile genetic elements, phage-related proteins, and a peculiar (Csm3) CRISPR, suggesting exclusive interactions with bacteriophages, and other metabolic and regulatory functions contributing to distinguish the two groups of *Prochlorothrix*. Further information is reported in Supplementary Table 5.

The Heaps models obtained from the anvi'o and PIRATE workflows indicated the existence of an open *Prochlorothrix* pangenome, with $k = 3728$ and $\gamma = 0.33$, and $k = 3557$ and $\gamma = 0.32$, respectively (Fig. 4).

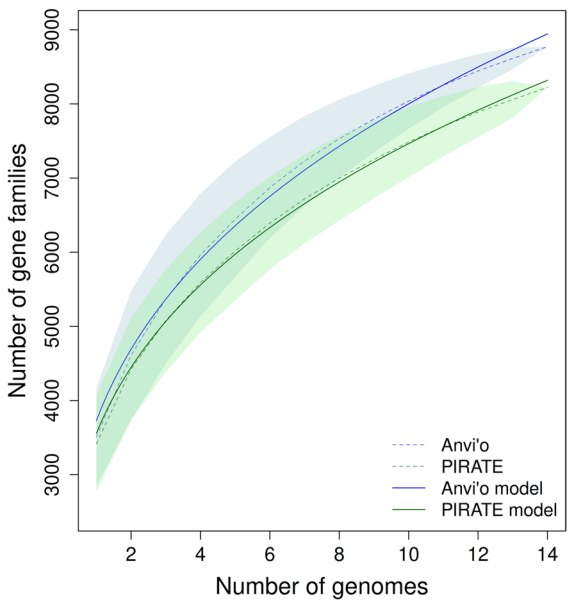

Fig. 4 Accumulation curves calculated from the pangenome analyses of *Prochlorothrix* by Anvi'o and PIRATE. The curves and the parameters of the Heap's model k and γ were calculated from the average values (connected by dashed lines) obtained from 1000 random genome combinations, fitted with nonlinear least squares (solid lines). Confidence intervals were computed from standard deviation values and reported as "polygon."

Secondary metabolites

After annotation of secondary metabolites, the four *Prochlorothrix* draft genomes determined in Lake Trasimeno did show the presence of a few main types, represented by terpene, terpene-precursor, NRPS/NRPS-like, but without showing any significant similarity with specific known clusters. Widening the analysis to the other cyanobacteria detected in Lake Trasimeno, a few other types were added, including hglE-KS, lanthipeptide-class-v, arylpolyene, RiPP-like, T1PKS, T3PKS, mycosporine, T1PKS-NRPS, and cyanobactin. Distinct BGCs were identified in the Nostocales taxa (heterocyst glycolipids; *A. arnoldii* and *R. raciborskii*), *R. raciborskii* (hassallidin C/hassallidin D), *A. arnoldii* (hexose-palythine-serine/hexose-shinorine), and Vampirovibrionia (carotenoid). No operons encoding legacy cyanotoxins were identified in the whole group of cyanobacteria determined in Lake Trasimeno. With the exclusion of a short (70 bp) *cyrJ* fragment gene (pident 98.6% to *R. raciborskii*) in *R. raciborskii* assembled in the CAST station, this was further confirmed by the absence of selected genes encoding legacy cyanotoxins in all the Lake Trasimeno cyanobacterial assemblies, and, excluding two minor hits (*mcyB*, 113–78 bp, pident 86–87% to *Oscillatoria* sp./*M. aeruginosa*, and *mcyD*, 197–201 bp, with MITE insertions, pident 93–94% to *Anabaena* sp.) in the whole set of the two unassembled contigs.

Apart from the addition of a triceptide, the full set of *Prochlorothrix* taxa included in the phylogenomic and pangenomic analyses showed the presence of the same main BGCs identified in the four Lake Trasimeno *Prochlorothrix* genomes, without any further similarity with specific known clusters. Similarly, no genes encoding legacy cyanotoxins were identified in the whole set of *Prochlorothrix* genomes analyzed in this work.

Antimicrobial resistance

No ARGs were identified by AMRFinderPlus in all the cyanobacterial genomes assembled from eDNA collected in Lake Trasimeno. Similarly, no ARGs were identified in all the *Prochlorothrix* genomes analyzed in this work. Conversely, the analyses performed by AMRFinderPlus on the whole set of assembled bacterial genomes of the two stations

identified with a strong confidence in genes identity and coverages (99–100% and 100%, respectively) the presence of a contiguous *almEFG* cluster (core colistin-associated determinants) in the *V. cholerae* MAG in the CAST station. These genes encode the lipid A glycine modification system previously associated with reduced susceptibility to polymyxins, including colistin (Herrera et al., 2014). It is interesting to observe that in this genome, excluding the *toxR* (a major virulence regulatory gene), the genes linked to virulence, i.e., *ctxA*, *ctxB*, and *tcpA* (Miller & Mekalanos, 1985; Taylor et al., 1987; Kaper et al., 1995), were not detected in the current assembly and annotation and should require phenotypic confirmation.

A wider analysis performed on the whole set of contigs, including the unbinned fraction and those < 1000 bp, identified other set of ARGs in the two stations, which included, besides the *almEFG* operon, the lincosamide class (Lnu(H), Lnu(I)), beta-lactam class (*blaVCC-1*, *blaOXA*, *blaM*, *cphA*), and mercury class (*merR*).

Discussion

The metagenomic analyses of two environmental samples collected in the large and shallow Lake Trasimeno allowed to clarify the taxonomic nature of the cyanobacterial community and accompanying bacteria and to investigate the taxonomical and functional characteristics of two groups of *Prochlorothrix* identified for the first time in Central Italy, at the northern border of the Mediterranean region.

Taxonomic nature of cyanobacteria and co-living bacteria

The metagenomic profiling method used in this work proved to be efficient in microbial species identification and microbial abundance estimation. The k-mer approaches, such as that implemented in Sylph, do not rely on sequence alignment or assembly of shotgun-sequenced reads, therefore avoiding computationally expensive analysis steps, and allowing detection of rare genomes present in the samples (Ghylin et al., 2014; Kim et al., 2019). As with all the metagenomic profiling techniques, this approach provides information on taxonomical and not on functional characteristics of organisms.

Metagenomic profiling showed that a large fraction of sequenced reads (around 70%) did not find known taxonomic correspondence. This fraction was comparable with those commonly detected in aquatic environmental samples (e.g., Giordano et al., 2024; Salmaso et al., 2024b; Rodríguez-Gijón et al., 2025), indicating a common presence of several microbial populations still not represented in reference taxonomic databases. Among the principal groups of bacteria found in the two sampling stations, *Nanopelagicus*, part of the acI lineage, and other Actinobacteria are often well represented in freshwater habitats, particularly in oligotrophic lakes (Newton et al., 2011; Garcia et al., 2013). In eutrophic environments, this group may play important roles in the cycling of dissolved organic carbon (DOC) and dissolved organic matter (DOM) produced by phytoplankton (Xie et al., 2024). The presence of taxa adapted to saline environments, such as *Pontimonas* (Actinobacteria) and, among Gammaproteobacteria, *Pseudohongiella* (Gorrasi et al., 2022), is consistent with the large fluctuations in water level and salinity that characterize Lake Trasimeno (Frondini et al., 2019). Gammaproteobacteria were further represented by a well-diversified group. The Burkholderiaceae and Steroidobacteraceae families are able to break complex organic compounds, which is consistent with the presence of high levels of organic matter in Lake Trasimeno; *Rubrivivax* spp. is represented by photoheterotrophic or chemoheterotrophic bacteria, which can exploit light or organic compounds; in aquatic environments, *V. cholerae* strains are generally associated with warm, nutrient-rich, and salty waters (Singleton et al., 1982; Huq et al., 2005). Bacteria within the Bacteroidia group are linked to high DOM recycling, connected with the decomposition of the high phytoplankton biomass. The large group of Alphaproteobacteria comprises a variety of species often linked to DOM degradation, also including photoheterotrophs or facultative degraders (Rhodobacteraceae and *Neoroseomonas*) (Pohlner et al., 2019). Within this group, *Hyphomonas* attaches to surfaces, such as resuspended particles (Quintero et al., 1998). Ilumatobacteraceae are often associated with organic-matter degradation and uptake from sediments (Silva-Solar et al., 2024; Zhang et al., 2024b). We should not disregard the presence of Patescibacteria (Candidate Phyla Radiation), a large but poorly studied group of tiny, mostly uncultivated microbes, characterized by

reduced genomes missing key pathways, and therefore relying on host bacteria for growth (Tian et al., 2020; Wang et al., 2023). The overall picture is that of a lake characterized by elevated internal input of organic matter produced by phytoplankton and sustained recycling by a diversified bacterial community.

Metagenomic profiling identified a large number of cyanobacterial taxa, of which only a few among the most abundant found correspondence in the microscopical examination, which did not detect *Prochlorothrix*, *Nodosilinea,* and *Sphaerospermopsis*, as well as all the tiny picocyanobacteria (*Cyanobium* spp.), the undetermined groups and non-photosynthetic cyanobacteria (Vampirovibrionia) listed in Fig. 1B, and other rarest taxa additionally listed in Supplementary Table 2. As for Vampirovibrionia, excluding *Vampirovibrio chlorellavorus* Gromov & Mamkayeva (Soo et al., 2015), which was not detected in this work, no morphological descriptions for this group are available.

Compared to metagenomic profiling, the number of medium-/high-quality cyanobacterial draft genomes was circumscribed to only the most abundant taxa (Fig. 1B). The number of MAGs that may be potentially characterized is limited by the sequencing depth, while the sequenced reads are partitioned among the bacterial draft genomes. Therefore, though not relevant for direct functional characterization, metagenomic profiling can complement the description of a limited number of draft genomes by providing more extensive taxonomic classifications.

Prochlorothrix MAGs distribution

The species *P. hollandica* was established in 1989 by Burger-Wiersma et al. (1989), based on a strain isolated from a sample from a mixed water column from the eutrophic Lake Loosdrecht (The Netherlands). The isolated strain was successively stored as CALU1027 at the CALU Culture Collection, while other strains have been reported with different collection indexes (Velichko et al., 2016), including an axenic strain (PCC 9006) isolated at the Pasteur Culture Collection (Schyns et al., 1997). Besides *P. hollandica*, a second species, *P. scandica*, originating from Lake Mälaren (Central Sweden), was characterized by Pinevich et al. (1999). The filamentous and freshwater species of the genus *Prochlorothrix* are distinct from the other Chl-*b* synthesizing unicellular and mostly marine counterparts, i.e. *Prochlorococcus, Acaryochloris*, and *Prochloron* (the latter an obligate symbiont of ascidians) (Pinevich et al., 2012; Nielsen et al., 2015; Pinevich & Averina, 2024).

The ANI_b, the AAI values, and the phylogenomic analyses clearly showed the existence of a compact cluster including all the *P. hollandica* assemblies (Richter & Rosselló-Móra, 2009). With the same criteria, though being formally part of a congruent genus, the unclassified *Prochlorothrix* spp. MAGs showed the existence of different geographical subclades originating from diverse geographical areas. Furthermore, the lower AAI values computed between the two main clades *P. hollandica* and *Prochlorothrix* (range 0.74–0.75) may still be considered within the AAI cutoff value applied for the delineation of genera, which was indicated in the range 0.60–0.80 (Konstantinidis & Tiedje, 2005; Park et al., 2022).

Information on the geographical distribution of *Prochlorothrix* is quite scanty. Available observations identified *Prochlorothrix* species in eutrophic freshwater or brackish habitats of North Europe, mostly dominated by cyanobacteria, while there was no indication of their presence in oligotrophic lakes (Pinevich et al., 2012). Nevertheless, a reliable assessment of the geographical distribution range was largely impeded by the difficulty to firmly discriminate by light microscopy *Prochlorothrix* species from other filamentous cyanobacteria and particularly from morphologically undistinguishable species belonging to *Pseudanabaena*, *Planktothrix*, and *Limnothrix* (Geiss et al., 2003; Pinevich et al., 2012). Additional information was obtained from the analysis of environmental DNA, with amplification of *pcb* genes, which documented the presence of *Prochlorothrix* in the eutrophic Darss-Zingst estuary (southern Baltic Sea) (Geiss et al., 2003). The discovery was further supported by the detection of autofluorescent *Prochlorothrix*-like filaments consistent with the presence of Chl-*b*. Potential presence of *Prochlorothrix* was also suggested by large-scale metabarcoding surveys in the Alpine region, where several short 16S rRNA variants were identified in the plankton and biofilm in Germany, France, Austria, Switzerland, and Slovenia, with pident values, after blastn, from 100% (uncultured *Prochlorothrix* sp.) to 92.8% (*Prochlorothrix hollandica* PCC-9006 = CALU-1027) (zenodo.org/records/5822484). Furthermore, metabarcoding based

on the same SSU rRNA gene target detected the potential presence of *Prochlorothrix* sp. in the New Orleans region (USA) (Amaral-Zettler et al., 2008). Rare filaments of *Prochlorothrix* sp. were also listed at the Messolonghi saltworks (Western Greece), with a salinity range between 50 and 80 ppt (Hotos, 2021), but no other supporting evidences were reported. More recently, a few draft genomes of *Prochlorothrix* sp. assembled from freshwater samples collected in India, Singapore, and the USA have been deposited in public repositories (Supplementary Table 1). Despite scarce, these observations suggest that *Prochlorothrix* may be more widely distributed, confirming the note by Bullerjahn & Post (2014).

Overall, the genomic and phylogenomic characterization of *P. hollandica* and other undetermined *Prochlorothrix* assemblies does represent the first documented detection of these taxa in the Mediterranean region. More in general, considering the difficulties intrinsic in the microscopical discrimination of species possessing a simple set of diacritical morphological characters (such as *Prochlorothrix* and several other cyanobacteria), the application of metagenomic profiling and characterization of MAGs from environmental samples does represent a straightforward, robust, and objective approach to unveil taxa distribution and functional characteristics of cyanobacteria. The downside of this approach is that the genomes so determined do not have a corresponding culture to be used for morphological, morphometric, and physiological assessment. In the two sets of taxa included in the phylogenomic (Fig. 2) and pangenomic analyses (Fig. 3), only two genomes originated from an axenic culture (GCA_000332315.1) and a unialgal, non-axenic culture (GCA_000341585.2), both derived from strain PCC 9006 (Supplementary Table 1). While the morphological characteristics of *P. hollandica* have been described, which allow tentative classifications under the microscope, the phenotypic characteristics of the *Prochlorothrix* spp. taxa remain incomplete and elusive (next section).

Main functional characteristics of the *Prochlorothrix* MAGs

The basic annotation of the genomes of *P. hollandica* and *Prochlorothrix* was congruent with that expected in oxygenic, non-diazotrophic photosynthetic cyanobacteria, showing no significant differences in the KEGG modules between the two clades. The annotation highlighted the presence of CAO and *pcb* genes, i.e., the main diacritical traits distinguishing the Chl-*b* containing cyanobacteria from the other members of this phylum. Nevertheless, the evaluation of the CAO and *pcb* sequences was heavily influenced by the scarcity of corresponding reference sequences in public databases, which are essentially still limited to those determined in *P. hollandica*. More specifically, regarding the filaments ascribed to *Prochlorothrix* spp., the findings relating to Chl-*b* associated sequences will need to be validated on a larger scale. This will require examination of the full extent of variability found in orthologous genes across a wider range of species within this genus to ensure a more comprehensive understanding of their genetic diversity and functionality.

Other basic potential structural differences regarded the set of *gvp* genes detected in the two clades. The set of *gvp* genes in *P. hollandica* was indicative of an expected capacity to produce gas vesicles, as confirmed by their direct observation at the cell poles of the PCC 9006 strain by transmission electron microscopy and further confirmed by pressure nephelometry that showed a critical pressure of gas vesicles of around 9 bars (Burger-Wiersma et al., 1989). Conversely, excluding incomplete annotations (Eisenhofer et al., 2023), the presence of a reduced set of gas vesicle encoding genes in *Prochlorothrix* spp. (*gvpACNG*) may still include a minimal functional set of genes (possibly forming incomplete or less robust gas vesicles), while the presence of only the *gvpG* gene, lacking the structural *gvpA* gene (Pfeifer, 2012), cannot sustain the formation of the gas vesicles walls. This could lead to the conclusion about the existence of localized populations ascribed under the genus *Prochlorothrix* with adaptations to illuminated shallow and turbulent environments or to periphytic lifestyle.

Core and cloud pangenome

Considering the low number of genomes, the number of gene clusters obtained from the pangenome analysis was relatively high, especially when compared with species analyzed with a much more numerous number of assemblies and using the same computational approach (Salmaso et al., 2025). The two main clades and, partly, the *Prochlorothrix* subclades

defined by the phylogenomic analyses showed correspondence with the subgroups in the shell and cloud gene clusters, which further consolidated the differences at the level of dispensable genes. Contrary to the core genome, which generally includes genes vertically inherited necessary for primary metabolic activities, the shell and cloud pangenome include genes related to the adaptation to specific environments (Riley & Lizotte-Waniewski, 2009; Matthews et al., 2024). For example, besides Soft Core, the high representation of signal transduction mechanisms in the Shell pangenome (and partly Cloud) was indicative of an important role of this category in the regulation of the response to specific environmental signals, allowing rapid adaptation to fluctuating nutrient availability (Watzer et al., 2019; Zhang et al., 2024a), light (Bhaya, 2004; Montgomery, 2007), temperature (Suzuki et al., 2001; Llop et al., 2025), and other stress-related stimuli (Rachedi et al., 2020; Rai et al., 2021). Overall, environmental sensing functions are linked to a wide range of both universal signals (e.g., basic nutrient status) and varying niche-specific signals (e.g., chemotaxis, pathogen–host interactions, competitors), allowing species and populations to adapt to a different range of environmental conditions (Alm et al., 2006; Los et al., 2010; Matilla & Krell, 2018; Matthews et al., 2024). Conversely, defense mechanisms and mobilome were increasingly or exclusively (the latter) present in the Shell and Cloud pangenome, a highly dynamic region involved in the development of new CRISPR-Cas systems and restriction–modification systems (Tock & Dryden, 2005; Pattharaprachayakul et al., 2020; Jungblut et al., 2021; Papoulis et al., 2021) for protection against local bacteriophages and foreign genetic elements. The abundance of mobile genetic elements underscores the dynamic nature of the cloud pangenome, a genomic region characterized by horizontal gene transfer (HGT) and extensive genomic plasticity, enabling rapid adaptation and diversification through recently acquired genetic material (Tokuda & Shintani, 2024; Bhaya et al., 2025; Stark et al., 2025). This aspect could contribute to explaining the higher fraction of unclassified gene families in the Cloud pangenome compared to the Soft Core and Shell pangenomes.

Differences in the exclusive presence of COGs in the two *P. hollandica* and *Prochlorothrix* spp. clades indicated peculiar functional adaptations and lifestyle strategies, which are commented on in Supplementary Table 5. It is important to note that the presence of unique gene clusters in one clade does not imply the presence of unique life traits. Differences may be mitigated considering the relatively small number of exclusive gene clusters, suggesting that many traits are not clade-specific and that genes with closely related functions may be scattered across both clades. For example, though not present in all assemblies, an ABC-type proline/glycine betaine transport system (ProW) (Kempf & Bremer, 1995) with osmoprotection functions and similar to that exclusively detected in *P. hollandica* (OpuBA/BB) was also detected in a few *Prochlorothrix* MAGs.

The *Prochlorothrix* pangenome has several characteristics conforming to open pangenomes, which have a small fraction of core genes and high rates of gene acquisition by HGT (Brockhurst et al., 2019). These characteristics originated from the presence of two different clades in the analysis, which acquired genes from other species or developed different adaptations to survive in different habitats. In this regard, it cannot be overlooked that rarefaction analysis is fundamentally limited by the number and characteristics of genomes that are sampled. Estimating if a pangenome is open or closed depends on the number of genomes analyzed, the biases present in the dataset, and the bioinformatic tools used for the pangenome estimation (Tonkin-Hill et al., 2022, 2023; Salmaso et al., 2025).

Secondary metabolites and ARGs

Excluding the sporadic presence of very short single genes encoding *cyrJ* and *mcy* on the full set of contigs, the absence of operons encoding cyanotoxins resulting from genome annotation was fully congruent with results obtained with the analytical determinations performed on the water and net samples by LC–MS, which showed the complete absence of a wide range of legacy cyanotoxins. Considering the dominance of non-toxigenic strains of *R. raciborskii* in European countries and Lake Trasimeno (Austoni et al., 2025; Kokociński et al., 2025; Salmaso et al., 2025), the absence of toxicity was not unexpected but still remarkable, given the presence of a wide range of other cyanobacterial species. In this regard, the mesoeutrophic Lake Trasimeno stands out as an exception in Central Italy, where many other lakes did show the

presence of measurable concentrations of MCs and CYN, associated to the development of other dominant species, mostly *Planktothrix rubescens* (De Candolle ex Gomont) Anagnostidis & Komárek, *Aphanizomenon* sp., and *Microcystis* sp. (Manganelli et al., 2014).

The only ARG detected in the Lake Trasimeno MAGs was found in the *V. cholerae* assembly. Nevertheless, the presence of other ARGs belonging to the lincosamide, beta-lactam, and mercury classes detected after the analysis of the whole set of contigs highlighted the presence of a wide range of ARG classes in the bacterial community. It is worth emphasizing that in this work, the detection of ARGs was performed using a conservative approach implemented in AMRFinderPlus, based on homologies with a curated NCBI database, and stringent identity and coverage thresholds. The adoption of less stringent criteria would have identified additional distant homologs (data not shown), although these results may be associated with an increased risk of false positives assignments. Furthermore, while discussing AMR, we should consider that freshwater environments host bacteria that contain both inherent and acquired antimicrobial-resistant genes (Martinez, 2009), and that potential AMR connections should be verified by phenotypic data (van Belkum et al., 2020; Banerjee & Patel, 2023). As for the *V. cholerae* strain, if further confirmed, the annotation results would substantiate the frequent absence of cholera toxin genes in many environmental strains (Faruque et al., 1998), even though other virulence factors may be retained in this species (Gherlan et al., 2025).

Applicability of results and methodological considerations

The metagenomic assessment of water samples provides an objective, robust, and replicable method to evaluate in detail the taxonomic and functional characteristics of cyanobacterial species developing in aquatic ecosystems. In water bodies used for recreational purposes, the exact recognition of potentially toxic strains has important implications for the management of risks connected with human health. The resulting taxonomic and functional profiles provide robust guidance to the regular monitoring activities. While this study is based on two summer samples, we are confident that the main cyanobacterial species representative of the warmest months are included in this work. Nevertheless, since the analysis is based on a single set of samples collected in summer, these results cannot be considered exhaustive of the bacterial and cyanobacterial communities and functions developing during a whole annual cycle in Lake Trasimeno and would require analysis of samples representative of other seasons.

Conclusions

While confirming the presence of several known species already identified during the monitoring campaigns in Lake Trasimeno, the results of the metagenomic analysis distinguished a few new species never detected before in the Mediterranean area, such as *P. hollandica* and an uncharacterized *Prochlorothrix*. The evolutionary relationships between *P. hollandica* and the more heterogeneous and geographically dispersed group of uncharacterized *Prochlorothrix* assemblies remain unclear and open to updates, requiring the analysis and, possibly, phenotypic characterization of more strains than those analyzed in the present study. Except for two *P. hollandica* genomes, all the MAGs included in the present analyses were of metagenomic origin, i.e., exclusively obtained from environmental samples. The annotation of the complete set of cyanobacterial MAGs did not reveal the presence of operons encoding cyanotoxins, providing an important background for environmental monitoring. Based on a conservative analysis, the absence of ARGs in all the analyzed cyanobacterial genomes was equally significant. Conversely, the identification of different classes of ARGs in *V. cholerae* and unbinned contigs underscores the global relevance of antimicrobial resistance and the need for accurate evaluation and interpretation of AMR patterns also in freshwater environments. Finally, it is important to note that this work is open to updates also considering the analysis of samples representative of other seasons.

Acknowledgements We thank the personnel of ARPA Umbria involved in the environmental monitoring of Lake Trasimeno. We thank A. Zampedri, A. Paoli, S. Larger, and E. Stefani for technical assistance in the laboratory analysis. We are grateful to two anonymous reviewers for valuable comments and suggestions on an earlier version of the manuscript.

Author contributions Nico Salmaso helped in conceptualization, software, investigation, formal analysis, writing—original draft, review and editing. Leonardo Cerasino, Massimo Pindo, and Margherita Di Brizio contributed to methodology, investigation, writing—review & editing. Adriano Boscaini helped in methodology, investigation, data curation, writing—review & editing.

Funding Open access funding provided by Fondazione Edmund Mach - Istituto Agrario di San Michele all'Adige within the CRUI-CARE Agreement. This work was supported by the National Biodiversity Future Center and received funding from the European Union Next-GenerationEU—Piano Nazionale di Ripresa e Resilienza (PNRR) Missione 4 Componente 2, Investimento 1.4-D.D. 1034 17/06/2022, CN00000033, CUPD43C22001280006. This manuscript reflects only the authors' views and opinions; neither the European Union nor the European Commission can be considered responsible for them.

Data availability This Whole-Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the project number PRJNA1395559, Biosamples SAMN54366539-SAMN54366544, accessions JBTLXB000000000, JBTLXA000000000, JBTLWZ000000000, JBTLWY000000000, JBTLWX000000000, JBTLWW000000000. Environmental data have been reported in Salmaso et al. (2025), Table 1.

Declarations

Competing interests The authors have no relevant financial or non-financial interests to disclose.

Ethical approval No approval of research ethics committees was required to accomplish the goals of this study because experimental work was conducted with unregulated invertebrate species.

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

References

Alm, E., K. Huang & A. Arkin, 2006. The evolution of two-component systems in bacteria reveals different strategies for niche adaptation. PLoS Computational Biology 2: e143. https://doi.org/10.1371/journal.pcbi.0020143.

Alneberg, J., B. S. Bjarnason, I. de Bruijn, M. Schirmer, J. Quick, U. Z. Ijaz, L. Lahti, N. J. Loman, A. F. Andersson & C. Quince, 2014. Binning metagenomic contigs by coverage and composition. Nature Methods 11: 1144–1146. https://doi.org/10.1038/nmeth.3103.

Amaral-Zettler, L. A., J. D. Rocca, M. G. Lamontagne, M. R. Dennett & R. J. Gast, 2008. Changes in microbial community structure in the wake of Hurricanes Katrina and Rita. American Chemical Society. Environmental Science & Technology 42: 9072–9078. https://doi.org/10.1021/es801904z.

Astashyn, A., E. S. Tvedte, D. Sweeney, V. Sapojnikov, N. Bouk, V. Joukov, E. Mozes, P. K. Strope, P. M. Sylla, L. Wagner, S. L. Bidwell, L. C. Brown, K. Clark, E. W. Davis, B. Smith-White, W. Hlavina, K. D. Pruitt, V. A. Schneider & T. D. Murphy, 2024. Rapid and sensitive detection of genome contamination at scale with FCS-GX. Genome Biology 25: 60. https://doi.org/10.1186/s13059-024-03198-7.

Austoni, M., A. Boscaini, F. Buzzi, L. Cerasino, G. Franzini, F. Giacomazzi, M. Marchesi, C. Zampieri & N. Salmaso, 2025. Recent spread of *Raphidiopsis raciborskii* in the lake district south of the Alps. Advances in Oceanography and Limnology 16: 13225. https://doi.org/10.4081/aiol.2025.13225.

Averina, S. G., N. V. Velichko, A. A. Pinevich, E. V. Senatskaya & A. V. Pinevich, 2019. Non-*a* chlorophylls in cyanobacteria. Photosynthetica 57: 15109–1118. https://doi.org/10.32615/ps.2019.130.

Banerjee, R. & R. Patel, 2023. Molecular diagnostics for genotypic detection of antibiotic resistance: current landscape and future directions. JAC-Antimicrobial Resistance 5: dlad018. https://doi.org/10.1093/jacamr/dlad018.

Bashir, F., A. Bashir, N. Bouaïcha, L. Chen, G. A. Codd, B. Neilan, W.-L. Xu, L. Ziko, V. D. Rajput, T. Minkina, R. S. Arruda & B. A. Ganai, 2023. Cyanotoxins, biosynthetic gene clusters, and factors modulating cyanotoxin biosynthesis. World Journal of Microbiology and Biotechnology 39: 241. https://doi.org/10.1007/s11274-023-03652-x.

Bayliss, S. C., H. A. Thorpe, N. M. Coyle, S. K. Sheppard & E. J. Feil, 2019. PIRATE: a fast and scalable pangenomics toolbox for clustering diverged orthologues in bacteria. GigaScience 8: giz119. https://doi.org/10.1093/gigascience/giz119.

Beck, H. E., T. R. McVicar, N. Vergopolan, A. Berg, N. J. Lutsko, A. Dufour, Z. Zeng, X. Jiang, A. I. J. M. van Dijk & D. G. Miralles, 2023. High-resolution (1 km) Köppen-Geiger maps for 1901–2099 based on constrained CMIP6 projections. Scientific Data 10: 724. https://doi.org/10.1038/s41597-023-02549-6.

Bhaya, D., 2004. Light matters: phototaxis and signal transduction in unicellular cyanobacteria. Molecular

Microbiology 53: 745–754. https://doi.org/10.1111/j.1365-2958.2004.04160.x.

Bhaya, D., G. Birzu & E. P. C. Rocha, 2025. Horizontal gene transfer and recombination in Cyanobacteriota. Annual Review of Microbiology 79: 685–711. https://doi.org/10.1146/annurev-micro-041522-100420.

Blin, K., S. Shaw, L. Vader, J. Szenei, Z. L. Reitz, H. E. Augustijn, J. D. D. Cediel-Becerra, V. de Crécy-Lagard, R. A. Koetsier, S. E. Williams, P. Cruz-Morales, S. Wongwas, A. E. Segurado Luchsinger, F. Biermann, A. Korenskaia, M. M. Zdouc, D. Meijer, B. R. Terlouw, J. J. J. van der Hooft, N. Ziemert, E. J. N. Helfrich, J. Masschelein, C. Corre, M. G. Chevrette, G. P. van Wezel, M. H. Medema & T. Weber, 2025. antiSMASH 8.0: extended gene cluster detection capabilities and analyses of chemistry, enzymology, and regulation. Nucleic Acids Research 53: W32–W38. https://doi.org/10.1093/nar/gkaf334.

Brockhurst, M. A., E. Harrison, J. P. J. Hall, T. Richards, A. McNally & C. MacLean, 2019. The ecology and evolution of pangenomes. Current Biology: CB 29: R1094–R1103. https://doi.org/10.1016/j.cub.2019.08.012.

Bullerjahn, G. S. & A. F. Post, 2014. Physiology and molecular biology of aquatic cyanobacteria. Frontiers in Microbiology 5: 359. https://doi.org/10.3389/fmicb.2014.00359.

Bullerjahn, G. S., R. M. L. McKay & P. M. Visser, 2023. Editorial: Ecology and molecular biology of bloom-forming cyanobacteria. Frontiers in Microbiology 14: 134658. https://doi.org/10.3389/fmicb.2023.1346581.

Burger-Wiersma, T., L. J. Stal & L. R. Mur, 1989. *Prochlorothrix hollandica* gen. nov., sp. nov., a filamentous oxygenic photoautotrophic procaryote containing chlorophylls *a* and *b*: assignment to Prochlorotrichaceae fam. nov. and Order Prochlorales Florenzano, Balloni, and Materassi 1986, with emendation of the ordinal description. International Journal of Systematic and Evolutionary Microbiology 39: 250–257. https://doi.org/10.1099/00207713-39-3-250.

Cai, H., C. J. McLimans, J. E. Beyer, L. R. Krumholz & K. D. Hambright, 2023. *Microcystis* pangenome reveals cryptic diversity within and across morphospecies. Science Advances 9: 3783. https://doi.org/10.1126/sciadv.add3783.

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. https://doi.org/10.1016/j.hal.2017.05.004.

Chaumeil, P.-A., A. J. Mussig, P. Hugenholtz & D. H. Parks, 2022. GTDB-Tk v2: memory friendly classification with the genome taxonomy database. Bioinformatics 38: 5315–5316. https://doi.org/10.1093/bioinformatics/btac672.

Chklovski, A., D. H. Parks, B. J. Woodcroft & G. W. Tyson, 2023. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. Nature Methods 20: 1203–1212. https://doi.org/10.1038/s41592-023-01940-w.

Cissell, E. C. & S. J. McCoy, 2021. Shotgun metagenomic sequencing reveals the full taxonomic, trophic, and functional diversity of a coral reef benthic cyanobacterial mat from Bonaire, Caribbean Netherlands. Science of the Total Environment 755: 142719. https://doi.org/10.1016/j.scitotenv.2020.142719.

Compson, Z. G., B. McClenaghan, G. A. C. Singer, N. A. Fahner & M. Hajibabaei, 2020. Metabarcoding from microbes to mammals: comprehensive bioassessment on a global scale. Frontiers in Ecology and Evolution 8: 379. https://doi.org/10.3389/fevo.2020.581835.

Danecek, P., J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M. O. Pollard, A. Whitwham, T. Keane, S. A. McCarthy, R. M. Davies & H. Li, 2021. Twelve years of SAMtools and BCFtools. GigaScience 10: giab008. https://doi.org/10.1093/gigascience/giab008.

Delmont, T. O., C. Quince, A. Shaiber, Ö. C. Esen, S. T. Lee, M. S. Rappé, S. L. McLellan, S. Lücker & A. M. Eren, 2018. Nitrogen-fixing populations of Planctomycetes and Proteobacteria are abundant in surface ocean metagenomes. Nature Microbiology 3: 804–813. https://doi.org/10.1038/s41564-018-0176-9.

Djordjevic, S. P., V. M. Jarocki, T. Seemann, M. L. Cummins, A. E. Watt, B. Drigo, E. R. Wyrsch, C. J. Reid, E. Donner & B. P. Howden, 2024. Genomic surveillance for antimicrobial resistance – a one health perspective. Nature Reviews Genetics 25: 142–157. https://doi.org/10.1038/s41576-023-00649-y.

Edgar, R. C., 2022. Muscle5: high-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. Nature Communications 13: 6968. https://doi.org/10.1038/s41467-022-34630-w.

Eisenhofer, R., I. Odriozola & A. Alberdi, 2023. Impact of microbial genome completeness on metagenomic functional inference. ISME Communications 3: 1–5. https://doi.org/10.1038/s43705-023-00221-z.

Eren, A. M., E. Kiefl, A. Shaiber, I. Veseli, S. E. Miller, M. S. Schechter, I. Fink, J. N. Pan, M. Yousef, E. C. Fogarty, F. Trigodet, A. R. Watson, Ö. C. Esen, R. M. Moore, Q. Clayssen, M. D. Lee, V. Kivenson, E. D. Graham, B. D. Merrill, A. Karkman, D. Blankenberg, J. M. Eppley, A. Sjödin, J. J. Scott, X. Vázquez-Campos, L. J. McKay, E. A. McDaniel, S. L. R. Stevens, R. E. Anderson, J. Fuessel, A. Fernandez-Guerra, L. Maignien, T. O. Delmont & A. D. Willis, 2021. Community-led, integrated, reproducible multi-omics with anvi'o. Nature Microbiology 6: 3–6. https://doi.org/10.1038/s41564-020-00834-3.

Faruque, S. M., M. J. Albert & J. J. Mekalanos, 1998. Epidemiology, genetics, and ecology of toxigenic *Vibrio cholerae*. Microbiology and Molecular Biology Reviews: MMBR 62: 1301–1314. https://doi.org/10.1128/MMBR.62.4.1301-1314.1998.

Feldgarden, M., V. Brover, N. Gonzalez-Escalona, J. G. Frye, J. Haendiges, D. H. Haft, M. Hoffmann, J. B. Pettengill, A. B. Prasad, G. E. Tillman, G. H. Tyson & W. Klimke, 2021. AMRFinderPlus and the Reference Gene Catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. Scientific Reports 11: 12728. https://doi.org/10.1038/s41598-021-91456-0.

Frondini, F., W. Dragoni, N. Morgantini, M. Donnini, C. Cardellini, S. Caliro, M. Melillo & G. Chiodini, 2019. An endorheic lake in a changing climate: geochemical

investigations at Lake Trasimeno (Italy). Multidisciplinary Digital Publishing Institute. Water 11: 1319. https://doi.org/10.3390/w11071319.

Garcia, S. L., K. D. McMahon, M. Martinez-Garcia, A. Srivastava, A. Sczyrba, R. Stepanauskas, H.-P. Grossart, T. Woyke & F. Warnecke, 2013. Metabolic potential of a single cell belonging to one of the most abundant lineages in freshwater bacterioplankton. The ISME Journal 7: 137–147. https://doi.org/10.1038/ismej.2012.86.

Geiss, U., I. Bergmann, M. Blank, R. Schumann, M. Hagemann & A. Schoor, 2003. Detection of Prochlorothrix in brackish waters by specific amplification of pcb genes. Applied and Environmental Microbiology 69: 6243–6249. https://doi.org/10.1128/AEM.69.10.6243-6249.2003.

Gherlan, G. S., D. S. Lazar, S. A. Florescu, R. M. Dirtu, D. R. Codreanu, S. Lupascu & M. Nica, 2025. Non-toxigenic *Vibrio cholerae* – just another cause of vibriosis or a potential new pandemic? Archive of Clinical Cases 12: 5–16. https://doi.org/10.22551/2025.46.1201.10305.

Ghylin, T. W., S. L. Garcia, F. Moya, B. O. Oyserman, P. Schwientek, K. T. Forest, J. Mutschler, J. Dwulit-Smith, L.-K. Chan, M. Martinez-Garcia, A. Sczyrba, R. Stepanauskas, H.-P. Grossart, T. Woyke, F. Warnecke, R. Malmstrom, S. Bertilsson & K. D. McMahon, 2014. Comparative single-cell genomics reveals potential ecological niches for the freshwater acI Actinobacteria lineage. The ISME Journal 8: 2503–2516. https://doi.org/10.1038/ismej.2014.135.

Giordano, N., M. Gaudin, C. Trottier, E. Delage, C. Nef, C. Bowler & S. Chaffron, 2024. Genome-scale community modelling reveals conserved metabolic cross-feedings in epipelagic bacterioplankton communities. Nature Communications 15: 2721. https://doi.org/10.1038/s41467-024-46374-w.

Gorrasi, S., M. Pasqualetti, M. Braconcini, B. Muñoz-Palazon & M. Fenice, 2022. Could *Pontimonas* harbour halophilic members able to withstand very broad salinity variations? Multidisciplinary Digital Publishing Institute. Microorganisms 10: 790. https://doi.org/10.3390/microorganisms10040790.

Gotelli, N. J. & R. K. Colwell, 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecology Letters 4: 379–391. https://doi.org/10.1046/j.1461-0248.2001.00230.x.

Hay, M. C., A. C. Mitchell, A. R. Soares, A. R. Debbonaire, D. C. Mogrovejo, N. Els & A. Edwards, 2023. Metagenome-assembled genomes from high Arctic glaciers highlight the vulnerability of glacier-associated microbiota and their activities to habitat loss. Microbiology Society, Microbial Genomics 9: 001131. https://doi.org/10.1099/mgen.0.001131.

Herrera, C. M., A. A. Crofts, J. C. Henderson, S. C. Pingali, B. W. Davies & M. S. Trent, 2014. The Vibrio cholerae VprA-VprB two-component system controls virulence through endotoxin modification. mBio. https://doi.org/10.1128/mBio.02283-14.

Hotos, G. N., 2021. A preliminary survey on the planktonic biota in a hypersaline pond of Messolonghi Saltworks (W. Greece). Diversity 13: 270. https://doi.org/10.3390/d13060270.

Huerta-Cepas, J., F. Serra & P. Bork, 2016. ETE 3: reconstruction, analysis, and visualization of phylogenomic data. Molecular Biology and Evolution 33: 1635–1638. https://doi.org/10.1093/molbev/msw046.

Huq, A., R. B. Sack, A. Nizam, I. M. Longini, G. B. Nair, A. Ali, J. G. Morris, M. N. H. Khan, A. K. Siddique, M. Yunus, M. J. Albert, D. A. Sack & R. R. Colwell, 2005. Critical factors influencing the occurrence of *Vibrio cholerae* in the environment of Bangladesh. Applied and Environmental Microbiology 71: 4645–4654. https://doi.org/10.1128/AEM.71.8.4645-4654.2005.

Jungblut, A. D., F. Raymond, M. B. Dion, S. Moineau, V. Mohit, G. Q. Nguyen, M. Déraspe, É. Francovic-Fontaine, C. Lovejoy, A. I. Culley, J. Corbeil & W. F. Vincent, 2021. Genomic diversity and CRISPR-Cas systems in the cyanobacterium *Nostoc* in the High Arctic. Environmental Microbiology 23: 2955–2968. https://doi.org/10.1111/1462-2920.15481.

Kanehisa, M. & Y. Sato, 2020. KEGG Mapper for inferring cellular functions from protein sequences. Protein Science: A Publication of the Protein Society 29: 28–35. https://doi.org/10.1002/pro.3711.

Kanehisa, M., S. Goto, Y. Sato, M. Kawashima, M. Furumichi & M. Tanabe, 2014. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Research 42: D199-205. https://doi.org/10.1093/nar/gkt1076.

Kanehisa, M., Y. Sato & M. Kawashima, 2022. KEGG mapping tools for uncovering hidden features in biological data. Protein Science: A Publication of the Protein Society 31: 47–53. https://doi.org/10.1002/pro.4172.

Kang, D. D., F. Li, E. Kirton, A. Thomas, R. Egan, H. An & Z. Wang, 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ 7: e7359. https://doi.org/10.7717/peerj.7359.

Kaper, J. B., J. G. Morris & M. M. Levine, 1995. Cholera. Clinical Microbiology Reviews 8: 48–86. https://doi.org/10.1128/cmr.8.1.48.

Kempf, B. & E. Bremer, 1995. OpuA, an osmotically regulated binding protein-dependent transport system for the osmoprotectant glycine betaine in *Bacillus subtilis*. Journal of Biological Chemistry 270: 16701–16713. https://doi.org/10.1074/jbc.270.28.16701.

Khatiebi, S., K. Kiprotich, Z. Onyando, J. Mwaura, C. Wekesa, C. N. Chi, C. Mulambalah & P. Okoth, 2024. High-throughput shotgun emtagenomics of microbial footprints uncovers a cocktail of Noxious antibiotic resistance genes in the Winam Gulf of Lake Victoria. Journal of Tropical Medicine 2024: 7857069. https://doi.org/10.1155/jotm/7857069.

Kim, S., I. Kang, J.-H. Seo & J.-C. Cho, 2019. Culturing the ubiquitous freshwater actinobacterial acI lineage by supplying a biochemical "helper" catalase. The ISME Journal 13: 2252–2263. https://doi.org/10.1038/s41396-019-0432-x.

Kim, D., S. Park & J. Chun, 2021. Introducing EzAAI: a pipeline for high throughput calculations of prokaryotic average amino acid identity. Journal of Microbiology 59: 476–480. https://doi.org/10.1007/s12275-021-1154-0.

Kokociński, M., C. Graco-Roza, I. Jasser, J. Karosienė, J. Kasperovičienė, J. Kobos, J. Koreivienė, J. Mankiewicz-Boczek, J. Soininen & A. Szczurowska, 2025. Environmental factors determining the distribution patterns of invasive *Raphidiopsis raciborskii* and *R. mediterranea* in central east Europe. Frontiers in Microbiology 16: 1533716.

Komárek, J., 2015. Review of the cyanobacterial genera implying planktic species after recent taxonomic revisions according to polyphasic methods: state as of 2014. Hydrobiologia 764: 259–270. https://doi.org/10.1007/s10750-015-2242-0.

Komárek, J., J. Kopecký & V. Cepák, 1999. Generic characters of the simplest cyanoprokaryotes *Cyanobium, Cyanobacterium* and *Synechococcus*. Cryptogamie Algologie 20: 209–222. https://doi.org/10.1016/S0181-1568(99)80015-4.

Komárek, J., 2016. A polyphasic approach for the taxonomy of cyanobacteria: principles and applications. European Journal of Phycology 51: 346–353. https://doi.org/10.1080/09670262.2016.1163738

Konstantinidis, K. T. & J. M. Tiedje, 2005. Towards a genome-based taxonomy for prokaryotes. Journal of Bacteriology 187: 6258–6264. https://doi.org/10.1128/JB.187.18.6258-6264.2005.

Koslicki, D. & H. Zabeti, 2019. Improving MinHash via the containment index with applications to metagenomic analysis. Applied Mathematics and Computation 354: 206–215. https://doi.org/10.1016/j.amc.2019.02.018.

Kozlíková-Zapomělová, E., P. Urajová & J. Mareš, 2025. Phylogenetic heterogeneity within the *Cuspidothrix* clade (Nostocales, Cyanobacteria): *Dolichospermum* longicellulare as a black sheep of its genus. Taylor & Francis. Phycologia 64: 346–354. https://doi.org/10.1080/00318884.2025.2545165.

Langmead, B. & S. L. Salzberg, 2012. Fast gapped-read alignment with Bowtie 2. Nature Methods 9: 357–359. https://doi.org/10.1038/nmeth.1923.

Leao, T., G. Castelão, A. Korobeynikov, E. A. Monroe, S. Podell, E. Glukhov, E. E. Allen, W. H. Gerwick, & L. Gerwick, 2017. Comparative genomics uncovers the prolific and distinctive metabolic potential of the cyanobacterial genus *Moorea*. Proceedings of the National Academy of Sciences 114: 3198–3203. https://doi.org/10.1073/pnas.1618556114.

Lee, M. D., 2019. GToTree: a user-friendly workflow for phylogenomics. Bioinformatics 35: 4162–4164. https://doi.org/10.1093/bioinformatics/btz188.

Lefler, F. W., D. E. Berthold, M. Barbosa, J. Hu, A. Taylor, J. Moretto, J. D. Chaffin, H. Raymond & H. D. Laughinghouse, 2025. Metagenomes from cyanobacterial harmful algal blooms from lakes in Ohio (USA). Microbiology Resource Announcements 14: e00400. https://doi.org/10.1128/mra.00400-25.

Li, D., C.-M. Liu, R. Luo, K. Sadakane & T.-W. Lam, 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics (Oxford, England) 31: 1674–1676. https://doi.org/10.1093/bioinformatics/btv033.

Li, W., K. R. O'Neill, D. H. Haft, M. DiCuccio, V. Chetvernin, A. Badretdin, G. Coulouris, F. Chitsaz, M. K. Derbyshire, A. S. Durkin, N. R. Gonzales, M. Gwadz, C. J. Lanczycki, J. S. Song, N. Thanki, J. Wang, R. A. Yamashita, M. Yang, C. Zheng, A. Marchler-Bauer & F. Thibaud-Nissen, 2020. RefSeq: expanding the Prokaryotic genome annotation pipeline reach with protein family model curation. Nucleic Acids Research 49: D1020–D1028. https://doi.org/10.1093/nar/gkaa1105.

Litchman, E., C. A. Klausmeier, O. M. Schofield & P. G. Falkowski, 2007. The role of functional traits and trade-offs in structuring phytoplankton communities: scaling from cellular to ecosystem level. Ecology Letters 10: 1170–1181. https://doi.org/10.1111/j.1461-0248.2007.01117.x.

Liu, S., C. D. Moon, N. Zheng, S. Huws, S. Zhao & J. Wang, 2022. Opportunities and challenges of using metagenomic data to bring uncultured microbes into cultivation. Microbiome 10: 76. https://doi.org/10.1186/s40168-022-01272-5.

Llop, A., S. Bibak, T. Mata-Balaguer, L. Tremiño, L. Fuertes-García, J. L. Neira, R. Dixon, & A. Contreras, 2025. Temperature changes are signaled in cyanobacteria through the PipX interaction network. bioRxiv. 2025.05.19.654826. https://doi.org/10.1101/2025.05.19.654826.

Los, D. A., A. Zorina, M. Sinetova, S. Kryazhov, K. Mironov & V. V. Zinchenko, 2010. Stress sensors and signal transducers in cyanobacteria. Sensors (Basel, Switzerland) 10: 2386–2415. https://doi.org/10.3390/s100302386.

Ludovisi, A., D. Cappelletti, A. C. Elia, D. Gigante, E. Goretti, M. Lorenzoni, G. La Porta, & R. Venanzoni, 2021. IT21-A Lago Trasimeno. In Capotondi L. et al. La Rete Italiana per la Ricerca Ecologica di Lungo Termine. Lo studio della biodiversità e dei cambiamenti. CNR-Edizioni, Rome. pp. 663–673.

Ludovisi, A. & E. Gaino, 2010. Meteorological and water quality changes in Lake Trasimeno (Umbria, Italy) during the last fifty years. Journal of Limnology 69: 174–188. https://doi.org/10.4081/jlimnol.2010.174.

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://doi.org/10.1016/j.hal.2022.102187.

Manganelli, M., E. Viaggiu, R. Barone, F. Buzzi, F. Caviglia, R. Congestri, D. Copetti, R. De Angelis, F. Godeas, L. Guzzella, E. Masala, L. Naselli-Flores, N. Salmaso, & S. Scardala, 2014. Situazione nazionale: corpi idrici interessati da cianobatteri tossici. In Cianobatteri: linee guida per la gestione delle fioriture nelle acque di balneazione. Istituto Superiore di Sanità, Roma. pp. 116–143.

Manti, G., D. Mattei, V. Messineo, S. Melchiorre, S. Bogialli, N. Sechi, P. Casiddu, A. Luglié, M. Di Brizio & M. Bruno, 2005. First report of *Cylindrospermopsis raciborskii* in Italy. Harmful Algae News 28: 8–9.

Martinez, J. L., 2009. Environmental pollution by antibiotics and by antibiotic resistance determinants. Environmental

Pollution 157: 2893–2902. https://doi.org/10.1016/j.envpol.2009.05.051.

Matilla, M. A. & T. Krell, 2018. The effect of bacterial chemotaxis on host infection and pathogenicity. FEMS Microbiology Reviews 42: fux052. https://doi.org/10.1093/femsre/fux052.

Matthews, C. A., N. S. Watson-Haigh, R. A. Burton & A. E. Sheppard, 2024. A gentle introduction to pangenomics. Briefings in Bioinformatics 25: bbae588. https://doi.org/10.1093/bib/bbae588.

Miller, V. L. & J. J. Mekalanos, 1985. Genetic analysis of the cholera toxin-positive regulatory gene toxR. Journal of Bacteriology 163: 580–585. https://doi.org/10.1128/jb.163.2.580-585.1985.

Montgomery, B. L., 2007. Sensing the light: photoreceptive systems and signal transduction in cyanobacteria. Molecular Microbiology 64: 16–27. https://doi.org/10.1111/j.1365-2958.2007.05622.x.

Morabito, G., M. G. Mazzocchi, N. Salmaso, A. Zingone, C. Bergami, G. Flaim, S. Accoroni, A. Basset, M. Bastianini, G. Belmonte, F. Bernardi Aubry, I. Bertani, M. Bresciani, F. Buzzi, M. Cabrini, E. Camatti, C. Caroppo, B. Cataletto, M. Castellano, P. Del Negro, A. de Olazabal, I. Di Capua, A. C. Elia, D. Fornasaro, M. Giallain, F. Grilli, B. Leoni, M. Lipizer, L. Longobardi, A. Ludovisi, A. Lugliè, M. Manca, F. Margiotta, M. A. Mariani, M. Marini, M. Marzocchi, U. Obertegger, A. Oggioni, B. M. Padedda, M. Pansera, R. Piscia, P. Povero, S. Pulina, T. Romagnoli, I. Rosati, G. Rossetti, F. Rubino, D. Sarno, C. T. Satta, N. Sechi, E. Stanca, V. Tirelli, C. Totti & A. Pugnetti, 2018. Plankton dynamics across the freshwater, transitional and marine research sites of the LTER-Italy network. Patterns, fluctuations, drivers. Science of the Total Environment 627: 373–387. https://doi.org/10.1016/j.scitotenv.2018.01.153.

Moretto, J. A. S., F. W. Lefler, M. B. Lin, D. E. Berthold, M. Barbosa, A. R. Smyth & H. D. Laughinghouse, 2025. 16S rRNA amplicon sequencing of bacterial communities from stormwater ponds in South Florida. Microbiology Resource Announcements 14: e00690-e725. https://doi.org/10.1128/mra.00690-25.

Mugnai, M., M. Angela Margheri, C. Cristina Sili, S. Turicchia, E. Soldati, E. Maffettone, E. Funari, S. Scardala, M. Di Brizio & S. Ventura, 2008. The cyanobacterial community of Lake Trasimeno. Algological Studies. https://doi.org/10.1127/1864-1318/2008/0128-0037.

Nagata, N., S. Satoh, R. Tanaka & A. Tanaka, 2004. Domain structures of chlorophyllide a oxygenase of green plants and *Prochlorothrix hollandica* in relation to catalytic functions. Planta 218: 1019–1025. https://doi.org/10.1007/s00425-003-1181-6.

Newton, R. J., S. E. Jones, A. Eiler, K. D. McMahon & S. Bertilsson, 2011. A Guide to the natural history of freshwater lake bacteria. Microbiology and Molecular Biology Reviews 75: 14–49. https://doi.org/10.1128/mmbr.00028-10.

Nguyen, L.-T., H. A. Schmidt, A. von Haeseler & B. Q. Minh, 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Molecular Biology and Evolution 32: 268–274. https://doi.org/10.1093/molbev/msu300.

Nicolosi Gelis, M. M., A. Canino, A. Bouchez, I. Domaizon, C. Laplace-Treyture, F. Rimet & B. Alric, 2024. Assessing the relevance of DNA metabarcoding compared to morphological identification for lake phytoplankton monitoring. Science of the Total Environment 914: 169774. https://doi.org/10.1016/j.scitotenv.2023.169774.

Nielsen, D. A., M. Pernice, M. Schliep, G. Sablok, T. C. Jeffries, M. Kühl, D. Wangpraseurt, P. J. Ralph & A. W. D. Larkum, 2015. Microenvironment and phylogenetic diversity of *Prochloron* inhabiting the surface of crustose didemnid ascidians. Environmental Microbiology 17: 4121–4132. https://doi.org/10.1111/1462-2920.12983.

Nikolaitchik, O. A. & G. S. Bullerjahn, 1998. Transcript analysis of the *pcbABC* genes encoding the antenna apoproteins in the photosynthetic prokaryote, *Prochlorothrix hollandica*. FEMS Microbiology Letters 168: 187–194. https://doi.org/10.1111/j.1574-6968.1998.tb13272.x.

Nitnaware, K. M., K. B. Raskar, G. Agarwal, R. A. Chávez Montes, R. Chopra, D. L. López-Arredondo, T. D. Nikam & G. B. Patil, 2021. Whole-genome characterization and comparative genomics of a novel freshwater cyanobacteria species: *Pseudanabaena punensis*. Molecular Phylogenetics and Evolution 164: 107272. https://doi.org/10.1016/j.ympev.2021.107272.

Nurk, S., D. Meleshko, A. Korobeynikov & P. A. Pevzner, 2017. metaSPAdes: a new versatile metagenomic assembler. Genome Research 27: 824–834. https://doi.org/10.1101/gr.213959.116.

Orakov, A., A. Fullam, L. P. Coelho, S. Khedkar, D. Szklarczyk, D. R. Mende, T. S. B. Schmidt, & P. Bork, 2021. GUNC: detection of chimerism and contamination in prokaryotic genomes. Genome Biology 22: 178. https://doi.org/10.1186/s13059-021-02393-0

Oksanen, J., G. L. Simpson, F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'Hara, P. Solymos, M. H. H. Stevens, E. Szoecs, H. Wagner, M. Barbour, M. Bedward, B. Bolker, D. Borcard, T. Borman, G. Carvalho, M. Chirico, M. De Caceres, S. Durand, H. B. A. Evangelista, R. FitzJohn, M. Friendly, B. Furneaux, G. Hannigan, M. O. Hill, L. Lahti, C. Martino, D. McGlinn, M.-H. Ouellette, E. Ribeiro Cunha, T. Smith, A. Stier, C. J. F. Ter Braak, & J. Weedon, 2025. vegan: Community ecology package. 2.7–2.

Olm, M. R., C. T. Brown, B. Brooks & J. F. Banfield, 2017. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. The ISME Journal 11: 2864–2868. https://doi.org/10.1038/ismej.2017.126.

Pan, S., C. Zhu, X.-M. Zhao & L. P. Coelho, 2022. A deep siamese neural network improves metagenome-assembled genomes in microbiome datasets across different environments. Nature Communications 13: 2326. https://doi.org/10.1038/s41467-022-29843-y.

Papoulis, S. E., S. W. Wilhelm, D. Talmy & E. R. Zinser, 2021. Nutrient loading and viral memory drive accumulation of restriction modification systems in bloom-forming cyanobacteria. mBio. https://doi.org/10.1128/mbio.00873-21.

Pareeth, S., M. Bresciani, F. Buzzi, B. Leoni, F. Lepori, A. Ludovisi, G. Morabito, R. Adrian, M. Neteler & N. Salmaso, 2017. Warming trends of perialpine lakes from

homogenised time series of historical satellite and in-situ data. Science of the Total Environment 578: 417–426. https://doi.org/10.1016/j.scitotenv.2016.10.199.

Park, M.-J., Y. J. Kim, M. Park, J. Yu, T. Namirimu, Y.-R. Roh & K. K. Kwon, 2022. Establishment of genome based criteria for classification of the Family Desulfovibrionaceae and proposal of two novel genera, *Alkalidesulfovibrio* gen. nov. and *Salidesulfovibrio* gen. nov. Frontiers in Microbiology 13: 738205. https://doi.org/10.3389/fmicb.2022.738205.

Parks, D. H., M. Imelfort, C. T. Skennerton, P. Hugenholtz & G. W. Tyson, 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Research 25: 1043–1055. https://doi.org/10.1101/gr.186072.114.

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.

Pattharaprachayakul, N., M. Lee, A. Incharoensakdi & H. M. Woo, 2020. Current understanding of the cyanobacterial CRISPR-Cas systems and development of the synthetic CRISPR-Cas systems for cyanobacteria. Enzyme and Microbial Technology 140: 109619. https://doi.org/10.1016/j.enzmictec.2020.109619.

Pereira, A., F. Martínez-Jerónimo, D. P. Fewer, D. F. Simon, M. Hernández-Zamora, L. Martínez-Jerónimo, P. Antuna-González, G. Munoz, S. Sauvé, B. J. Shapiro, & N. Tromas, 2026. Linking cyanobacterial genomes to toxin dynamics through genome-resolved metagenomics. bioRxiv. 2026.02.24.707707.

Pessi, I. S., R. V. Popin, B. Durieu, Y. Lara, B. Tytgat, V. Savaglia, B. Roncero-Ramos, J. Hultman, E. Verleyen, W. Vyverman & A. Wilmotte, 2023. Novel diversity of polar Cyanobacteria revealed by genome-resolved metagenomics. Microbiology Society, Microbial Genomics 9: 001056. https://doi.org/10.1099/mgen.0.001056.

Pfeifer, F., 2012. Distribution, formation and regulation of gas vesicles. Nature Reviews Microbiology 10: 705–715. https://doi.org/10.1038/nrmicro2834.

Pinevich, A. V. & S. G. Averina, 2024. Taxonomy of cyanobacteria: The era of change. Microbiology 93: 521–536. https://doi.org/10.1134/S0026261724605724.

Pinevich, A., O. Skulberg, H. Matthijs, H. Schubert, E. Wil, O. Gavrilova & N. Velichko, 1999. Characterization of a novel chlorophyll *b*-containing *Prochlorothrix* species (Prochlorophyta) and its photosynthetic apparatus. Microbios 100: 159–174.

Pinevich, A., N. Velichko & N. Ivanikova, 2012. Cyanobacteria of the Genus *Prochlorothrix*. Frontiers in Microbiology. https://doi.org/10.3389/fmicb.2012.00173.

Pohlner, M., L. Dlugosch, B. Wemheuer, H. Mills, B. Engelen & B. K. Reese, 2019. The majority of active Rhodobacteraceae in marine sediments belong to uncultured genera: a molecular approach to link their distribution to environmental conditions. Frontiers in Microbiology 10: 659. https://doi.org/10.3389/fmicb.2019.00659.

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.

Pritchard, L., R. H. Glover, S. Humphris, J. G. Elphinstone & I. K. Toth, 2016. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. Analytical Methods 8: 12–24. https://doi.org/10.1039/C5AY02550H.

Quintero, E. J., K. Busch & R. M. Weiner, 1998. Spatial and temporal deposition of adhesive extracellular polysaccharide capsule and fimbriae by *Hyphomonas* strain MHS-3. Applied and Environmental Microbiology 64: 1246–1255. https://doi.org/10.1128/AEM.64.4.1246-1255.1998.

Rachedi, R., M. Foglino & A. Latifi, 2020. Stress signaling in Cyanobacteria: a mechanistic overview. Life 10: 312. https://doi.org/10.3390/life10120312.

Rai, R., K. K. Rai, S. Singh, A. Raj & L. C. Rai, 2021. Stress proteins and signal transduction in Cyanobacteria. In Rastogi, R. P. (ed), Ecophysiology and Biochemistry of Cyanobacteria Springer Nature, Singapore: 155–180.

Ramond, P., P. E. Galand & R. Logares, 2025. Microbial functional diversity and redundancy: moving forward. FEMS Microbiology Reviews 49: fuae031. https://doi.org/10.1093/femsre/fuae031.

Richter, M., & R. Rosselló-Móra, 2009. Shifting the genomic gold standard for the prokaryotic species definition. Proceedings of the National Academy of Sciences 106: 19126–19131. https://doi.org/10.1073/pnas.0906412106.

Riley, M. A. & M. Lizotte-Waniewski, 2009. Population genomics and the bacterial species concept. Methods in molecular biology (Clifton, N.J.) 532: 367–377. https://doi.org/10.1007/978-1-60327-853-9_21.

Rodríguez-Gijón, A., A. Pacheco-Valenciana, F. Milke, J. E. Dharamshi, J. J. Hampel, J. Damashek, G. Wienhausen, L. M. Rodriguez-R & S. L. Garcia, 2025. Widely-distributed freshwater microorganisms with streamlined genomes co-occur in cohorts with high abundance. Scientific Reports 15: 34482. https://doi.org/10.1038/s41598-025-22383-7.

Rosselló-Móra, R. & W. B. Whitman, 2019. Dialogue on the nomenclature and classification of prokaryotes. Systematic and Applied Microbiology 42: 5–14. https://doi.org/10.1016/j.syapm.2018.07.002.

Rott, E., N. Salmaso & E. Hoehn, 2007. Quality control of Utermöhl-based phytoplankton counting and biovolume estimates – an easy task or a Gordian knot? Hydrobiologia 578: 141–146. https://doi.org/10.1007/s10750-006-0440-5.

Rozday, T. J., 2025. Metagenomics distilled: new *k*-mer-based methods. Nature Reviews Microbiology 23: 473–473. https://doi.org/10.1038/s41579-025-01192-9.

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.

Rzymski, P., B. Poniedziałek, J. Mankiewicz-Boczek, E. J. Faassen, T. Jurczak, I. Gągała-Borowska, A. Ballot, M. Lürling & M. Kokociński, 2017. Polyphasic toxicological screening of *Cylindrospermopsis raciborskii* and *Aphanizomenon gracile* isolated in Poland. Algal Research 24: 72–80. https://doi.org/10.1016/j.algal.2017.02.011.

Salmaso, N., L. Naselli-Flores & J. Padisák, 2015. Functional classifications and their application in phytoplankton ecology. Freshwater Biology 60: 603–619. https://doi.org/10.1111/fwb.12520.

Salmaso, N., S. Bernabei, A. Boscaini, C. Capelli, L. Cerasino, I. Domaizon, T. Elersek, C. Greco, A. Krivograd Klemenčič, P. Tomassetti & R. Kurmayer, 2024a. Biodiversity patterns of cyanobacterial oligotypes in lakes and rivers: results of a large-scale metabarcoding survey in the Alpine region. Hydrobiologia 851: 1035–1062. https://doi.org/10.1007/s10750-023-05423-z.

Salmaso, N., L. Cerasino, M. Pindo & A. Boscaini, 2024b. Taxonomic and functional metagenomic assessment of a *Dolichospermum* bloom in a large and deep lake south of the Alps. FEMS Microbiology Ecology 100: fiae117. https://doi.org/10.1093/femsec/fiae117.

Salmaso, N., L. Cerasino, M. Di Brizio, M. Pindo & A. Boscaini, 2025. Phylogenomic and pangenomic assessment of a Mediterranean strain of *Raphidiopsis raciborskii* extends knowledge of the global distribution and characteristics of a potentially toxigenic cyanobacterium. Environmental Microbiology Reports 17: e70098. https://doi.org/10.1111/1758-2229.70098.

Schoch, C. L., S. Ciufo, M. Domrachev, C. L. Hotton, S. Kannan, R. Khovanskaya, D. Leipe, R. McVeigh, K. O'Neill, B. Robbertse, S. Sharma, V. Soussov, J. P. Sullivan, L. Sun, S. Turner, & I. Karsch-Mizrachi, 2020. NCBI Taxonomy: a comprehensive update on curation, resources and tools. Database 2020: baaa062. https://doi.org/10.1093/database/baaa062.

Schwengers, O., L. Jelonek, M. A. Dieckmann, S. Beyvers, J. Blom & A. Goesmann, 2021. Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. Microbial Genomics 7: 000685. https://doi.org/10.1099/mgen.0.000685.

Schyns, G., R. Rippka, A. Namane, D. Campbell, M. Herdman & J. Houmard, 1997. *Prochlorothrix hollandica* PCC 9006: genomic properties of an axenic representative of the chlorophyll a/b-containing oxyphotobacteria. Research in Microbiology 148: 345–354. https://doi.org/10.1016/S0923-2508(97)81590-2.

Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics (Oxford, England) 30: 2068–2069. https://doi.org/10.1093/bioinformatics/btu153.

Shaw, J. & Y. W. Yu, 2023. Fast and robust metagenomic sequence comparison through sparse chaining with skani. Nature Methods 20: 1661–1665. https://doi.org/10.1038/s41592-023-02018-3.

Shaw, J. & Y. W. Yu, 2024. Rapid species-level metagenome profiling and containment estimation with sylph. Nature Biotechnology 43: 1348–1359. https://doi.org/10.1038/s41587-024-02412-y.

Sieber, C. M. K., A. J. Probst, A. Sharrar, B. C. Thomas, M. Hess, S. G. Tringe & J. F. Banfield, 2018. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. Nature Microbiology 3: 836–843. https://doi.org/10.1038/s41564-018-0171-1.

Silva-Solar, S., T. Viver, Y. Wang, L. H. Orellana, K. Knittel & R. Amann, 2024. Acidimicrobiia, the actinomycetota of coastal marine sediments: abundance, taxonomy and genomic potential. Systematic and Applied Microbiology 47: 126555. https://doi.org/10.1016/j.syapm.2024.126555.

Singleton, F. L., R. Attwell, S. Jangi & R. R. Colwell, 1982. Effects of temperature and salinity on *Vibrio cholerae* growth. Applied and Environmental Microbiology 44: 1047–1058. https://doi.org/10.1128/aem.44.5.1047-1058.1982.

Soo, R. M., B. J. Woodcroft, D. H. Parks, G. W. Tyson & P. Hugenholtz, 2015. Back from the dead; the curious tale of the predatory cyanobacterium *Vampirovibrio chlorellavorus*. PeerJ 3: e968. https://doi.org/10.7717/peerj.968.

Speth, D. R., N. Pullen, S. T. N. Aroney, B. L. Coltman, J. T. Osvatic, B. J. Woodcroft, T. Rattei, & M. Wagner, 2025. GlobDB: a comprehensive species-dereplicated microbial genome resource. arXiv.

Stark, G. F., L. E. Smith, A. R. Truchon, R. M. Martin, E. R. Denison & S. W. Wilhelm, 2025. *Microcystis* plasmids: the unexplored portion of the mobilome and the presence of potential phage-like plasmids. ISME Communications 5: ycaf154. https://doi.org/10.1093/ismeco/ycaf154.

Stern, D. B., R. T. Raborn, S. P. Lovett, N. R. Boise, L. Carrasquilla, S. Enke, D. Radune, D. L. Woodruff, K. L. Wahl & M. J. Rosovitz, 2025. Novel toxin biosynthetic gene cluster in harmful algal bloom-causing *Heteroscytonema crispum*: insights into the origins of paralytic shellfish toxins. Genome Biology and Evolution 17: evae248. https://doi.org/10.1093/gbe/evae248.

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.

Sun, Z., S. Huang, M. Zhang, Q. Zhu, N. Haiminen, A. P. Carrieri, Y. Vázquez-Baeza, L. Parida, H.-C. Kim, R. Knight & Y.-Y. Liu, 2021. Challenges in benchmarking metagenomic profilers. Nature Methods 18: 618–626. https://doi.org/10.1038/s41592-021-01141-3.

Suzuki, I., Y. Kanesaki, K. Mikami, M. Kanehisa & N. Murata, 2001. Cold-regulated genes under control of the cold sensor Hik33 in *Synechocystis*. Molecular Microbiology 40: 235–244. https://doi.org/10.1046/j.1365-2958.2001.02379.x.

Taylor, R. K., V. L. Miller, D. B. Furlong & J. J. Mekalanos, 1987. Use of *phoA* gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. Proceedings of the National Academy of Sciences of the United States of America 84: 2833–2837. https://doi.org/10.1073/pnas.84.9.2833.

Tettelin, H., D. Riley, C. Cattuto & D. Medini, 2008. Comparative genomics: the bacterial pan-genome. Current Opinion in Microbiology 11: 472–477. https://doi.org/10.1016/j.mib.2008.09.006.

Tian, R., D. Ning, Z. He, P. Zhang, S. J. Spencer, S. Gao, W. Shi, L. Wu, Y. Zhang, Y. Yang, B. G. Adams, A. M. Rocha, B. L. Detienne, K. A. Lowe, D. C. Joyner, D. M. Klingeman, A. P. Arkin, M. W. Fields, T. C. Hazen, D.

A. Stahl, E. J. Alm & J. Zhou, 2020. Small and mighty: adaptation of superphylum Patescibacteria to groundwater environment drives their genome simplicity. BioMed Central Microbiome 8: 1–15. https://doi.org/10.1186/s40168-020-00825-w.

Timms, V. J., K. A. Hassan, L. A. Pearson & B. A. Neilan, 2023. Cyanobacteria as a critical reservoir of the environmental antimicrobial resistome. Environmental Microbiology 25: 2266–2276. https://doi.org/10.1111/1462-2920.16453.

Tock, M. R. & D. T. Dryden, 2005. The biology of restriction and anti-restriction. Current Opinion in Microbiology 8: 466–472. https://doi.org/10.1016/j.mib.2005.06.003.

Tokuda, M. & M. Shintani, 2024. Microbial evolution through horizontal gene transfer by mobile genetic elements. Microbial Biotechnology 17: e14408. https://doi.org/10.1111/1751-7915.14408.

Tonkin-Hill, G., J. Corander & J. Parkhill, 2023. Challenges in prokaryote pangenomics. Microbial Genomics 9: 001021. https://doi.org/10.1099/mgen.0.001021.

Tonkin-Hill, G., R. A. Gladstone, A. K. Pöntinen, S. Arredondo-Alonso, S. D. Bentley, & J. Corander, 2022. Robust analysis of prokaryotic pangenome gene gain and loss rates with Panstripe. bioRxiv. 2022.04.23.489244.

Valadez-Cano, C., N. Tromas, A. Reyes-Prieto, L. Johnston, Y. Huang, H. Morris, L. Zamlynny, D. G. Beach, R. C. Jamieson & J. Lawrence, 2025. Genetic diversity and anatoxin profiles of freshwater benthic cyanobacteria from Nova Scotia (Canada). Environmental Microbiology 27: e70067. https://doi.org/10.1111/1462-2920.70067.

van Belkum, A., C.-A.D. Burnham, J. W. A. Rossen, F. Mallard, O. Rochas & W. M. Dunne, 2020. Innovative and rapid antimicrobial susceptibility testing systems. Nature Reviews Microbiology 18: 299–311. https://doi.org/10.1038/s41579-020-0327-x.

Velichko, N., M. Rayko, E. Chernyaeva, A. Lapidus & A. Pinevich, 2016. Draft genome of *Prochlorothrix hollandica* CCAP 1490/1T (CALU1027), the chlorophyll *a*/*b*-containing filamentous cyanobacterium. Standards in Genomic Sciences 11: 82. https://doi.org/10.1186/s40793-016-0204-4.

Vettorazzo, S., A. Boscaini, L. Cerasino & N. Salmaso, 2024. From small water bodies to lakes: exploring the diversity of freshwater bacteria in an Alpine biosphere reserve. Science of the Total Environment 954: 176495. https://doi.org/10.1016/j.scitotenv.2024.176495.

Vollmers, J., S. Wiegand, F. Lenk & A.-K. Kaster, 2022. How clear is our current view on microbial dark matter? (Re-) assessing public MAG & SAG datasets with MDMcleaner. Nucleic Acids Research 50: e76. https://doi.org/10.1093/nar/gkac294.

Wang, Y., L. A. Gallagher, P. A. Andrade, A. Liu, I. R. Humphreys, S. Turkarslan, K. J. Cutler, M. L. Arrieta-Ortiz, Y. Li, M. C. Radey, J. S. McLean, Q. Cong, D. Baker, N. S. Baliga, S. B. Peterson & J. D. Mougous, 2023. Genetic manipulation of Patescibacteria provides mechanistic insights into microbial dark matter and the epibiotic lifestyle. Elsevier. Cell 186: 4803-4817.e13. https://doi.org/10.1016/j.cell.2023.08.017.

Watson, A. R., J. Füssel, I. Veseli, J. Z. DeLongchamp, M. Silva, F. Trigodet, K. Lolans, A. Shaiber, E. Fogarty, J. M. Runde, C. Quince, M. K. Yu, A. Söylev, H. G. Morrison, S. T. M. Lee, D. Kao, D. T. Rubin, B. Jabri, T. Louie & A. M. Eren, 2023. Metabolic independence drives gut microbial colonization and resilience in health and disease. Genome Biology 24: 78. https://doi.org/10.1186/s13059-023-02924-x.

Watzer, B., P. Spät, N. Neumann, M. Koch, R. Sobotka, B. Macek, O. Hennrich & K. Forchhammer, 2019. The signal transduction protein PII controls ammonium, nitrate and urea uptake in cyanobacteria. Frontiers in Microbiology 10: 1428. https://doi.org/10.3389/fmicb.2019.01428.

Xie, G., C. Sun, Y. Gong, W. Luo & X. Tang, 2024. Beyond the bloom: unraveling the diversity, overlap, and stability of free-living and particle-attached bacterial communities in cyanobacteria-dominated hypereutrophic lake. Microbial Ecology 87: 96. https://doi.org/10.1007/s00248-024-02410-2.

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

Zhang, Q., L. Jia, Y. Chen, H. Yan, Q. Chen, J. Zhang & H. Sun, 2024a. Molecular mechanisms of the cyanobacterial response to different phosphorus sources. Sustainability 16: 5642. https://doi.org/10.3390/su16135642.

Zhang, Y., T. Li, G. Li, T. Yuan, Y. Zhang & L. Jin, 2024b. Profiling sediment bacterial communities and the response to pattern-driven variations of total nitrogen and phosphorus in long-term polyculture ponds. Frontiers in Marine Science 11: 1403909. https://doi.org/10.3389/fmars.2024.1403909.

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.