



HAL
open science

Integrating metagenome-scale metabolic modelling and metabolomics to identify biochemical interactions in *Microcystis phycospheres*

Juliette Audemard, Nicolas Creusot, Julie Leloup, Charlotte Duval, Sébastien Halary, Lou Mary, Mélissa Eon, Thomas Forjonel, Mohamed Mouffok, Rémy Puppo, et al.

► To cite this version:

Juliette Audemard, Nicolas Creusot, Julie Leloup, Charlotte Duval, Sébastien Halary, et al.. Integrating metagenome-scale metabolic modelling and metabolomics to identify biochemical interactions in *Microcystis phycospheres*. 2026. <hal-05593509>

HAL Id: hal-05593509

<https://inria.hal.science/hal-05593509v1>

Preprint submitted on 16 Apr 2026












HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

Integrating metagenome-scale metabolic modelling and metabolomics to identify biochemical interactions in *Microcystis* phycospheres

Juliette Audemard^{1,2}, Nicolas Creusot³, Julie Leloup⁴, Charlotte Duval⁵, Sébastien Halary⁵, Lou Mary⁵, Melissa Eon³, Thomas Forjonel³, Mohamed Mouffok², Rémy Puppo⁵, Elodie Belmonte⁶, Veronique Gautier⁶, Jeanne Got⁷, Marie Lefebvre², Gabriel V. Markov⁸, Coralie Muller¹, Benjamin Marie⁵, Binta Diémé^{2,*}, Clémence Frioux^{1,*}

¹Inria, University of Bordeaux, INRAE, F-33400 Talence, France

²Université Clermont Auvergne, INRAE, UNH, PFEM, MetaboHUB Clermont, 63000 Clermont-Ferrand, France

³INRAE, UR1454 EABX, Bordeaux Metabolome, MetaboHub Bordeaux, 33612 Gazinet Cestas, France

⁴Sorbonne Université, Univ Paris-Cité, Univ Paris-Est, CNRS, IRD, INRAE, IEES Paris, F-75005, Paris, France

⁵Muséum National d'Histoire Naturelle, UMR 7245 CNRS-MNHN, MCAM, Paris, France

⁶Plateforme Gentyane, UMR 1095 Génétique Diversité Ecophysiologie des Céréales INRAE, UCA, Clermont-Ferrand, France

⁷Université de Rennes, Inria, CNRS, IRISA, Rennes, France

⁸Sorbonne Université, CNRS, Laboratoire de Biologie Intégrative des Modèles Marins, LBI2M, F-29680 Roscoff, France

* Corresponding authors, equal contributions. binta.dieme@inrae.fr and clemence.frioux@inria.fr

Abstract

Favoured by global changes, freshwater cyanobacterial harmful blooms generate major ecological, economical and public health challenges. *Microcystis*, one of the most widespread cyanobacterial genera, grows within a phycosphere where specialised interactions with its microbiome occur, and are suspected to influence bloom appearance and its potential toxicity. Using a combination of metagenomic, metabolomic and metabolic modelling, we characterised the phycospheres of twelve *Microcystis* strains isolated from a French pond. The distribution of metabolic reactions within *Microcystis* was consistent with their genospecies, whereas the metabolic landscape at the community level diverged from cyanobacterial phylogeny indicating functional decoupling between cyanobacteria and their associated microbiomes. Phycosphere-associated bacteria substantially expand the metabolic repertoire of the system, while maintaining functional redundancy within and across communities. On the other hand, metabolomic profiles were largely driven by cyanobacterial metabolic outputs. Metabolic modelling, together with the identification of toxic specialised metabolites produced by specific biosynthetic gene clusters, further highlighted differences in metabolic potential among phycospheres. Together, these findings deepen the understanding of *Microcystis*' phycosphere functioning, demonstrate the value of multi-omics systems biology approaches, and underscore the ecological relevance of interspecies and inter-phycosphere metabolic interactions as a structuring process in bloom-associated microbiomes.

Keywords systems biology, microbial community, metabolic modelling, freshwater cyanobacteria, metabolomics, phycosphere, *Microcystis*

1 Introduction

Freshwater harmful cyanobacterial blooms (HCBs) have become a globally recurring phenomenon, mainly driven by rising temperatures and increased nutrient loads from anthropogenic activities [65]. These blooms, characterised by the proliferation of cyanobacteria such as *Microcystis*, pose major ecological and societal challenges through the release of large amounts of organic matter and toxic metabolites (*e.g.*

microcystins). Such inputs disrupt ecosystem functioning with severe consequences for biodiversity, public health, and local economies [42, 70]. Despite substantial progress in understanding the environmental drivers and the dynamics of cyanobacterial blooms, these phenomena are still difficult to predict and overcome.

A particularly understudied aspect of these phenomena is the complex network of metabolic interactions within the *phycosphere* - the microscale environment surrounding phytoplankton cells and aggregates - colonised by heterotrophic bacteria. While the taxonomic composition of phycosphere-associated microbiomes varies with cyanobacterial species [56], their global metabolic potential appears conserved, suggesting functional redundancy [68]. These microbial interactions within the phycosphere are thought to support phytoplankton growth, stress tolerance and bloom persistence through nutrient recycling and metabolite exchange, involving commensalism and competition [44, 55, 10, 82]. Yet, precise mechanisms and consequences of this microbial syntrophy remain poorly understood [22].

Addressing these knowledge gaps, a promising approach is to study simplified, cultivated, or synthetic microbial communities [18], as non-axenic cultures of phytoplankton host a specific but reduced microbiome [88]. In this context, genome-scale metabolic networks (GSMNs) are levers to generate hypotheses about functional roles within microbial communities, and connect genes, proteins, and biochemical reactions [14, 4]. However, applications to freshwater cyanobacterial communities remain limited, with most studies focusing on marine systems or isolated cyanobacteria [69, 86, 72]. Metabolomics complements these approaches by detecting main metabolites potentially mediating microbial interactions and refining metabolic model predictions [20, 26, 23, 74, 64].

Here, we present a systems biology-based characterisation of the taxonomic composition and co-metabolism of 12 cultivated phycospheres of *Microcystis* spp., isolated from a eutrophic freshwater pond during a bloom event. By integrating metagenomics, genome-scale metabolic modelling and metabolomics, we investigate the metabolic capabilities of *Microcystis* strains and their associated microbiomes. This combined approach provides a functional view of the phycosphere organisation, and reveals putative metabolic interactions and community-specific functions.

2 Material and Methods

2.1 Phycosphere isolation and culture

Water samples were collected in September 2021, from a suburban freshwater pond near Paris, Ile-de-France (Cergy-Pontoise, GPS coordinates lat. 49.01343 long. 2.02575), experiencing a *Microcystis* bloom, (chlorophyll-content $\approx 10 \mu\text{g.L}^{-1}$) [34].

After pre-filtration on 50 μm porosity, twelve monoclonal *Microcystis* strains and their associated bacterial consortia were isolated by repeated of single-cell or small-colony transfers on solid or liquid media under an inverted microscope (NIKON ECLIPSE TS100, Japan). Viable clones were then cultivated in 25 cm^2 culture flasks (CORNING-FALCON) containing 10 mL of Z8 medium. Cyanobacterial strains were maintained in the Paris Museum Collection (PMC, <https://mcam.mnhn.fr/en/cyanobacteria-and-live-microalgae-470>) at 18°C, using white LED-powered lights providing an irradiance of 8-10 μm photons/ m^2/s , with a photoperiod of 13h light/11h dark. Isolated strains and cultures were all monoclonal and non-axenic. They were then cultivated in BG11 medium, with a photoperiod of 16h light/8h dark and $\sim 14 \mu\text{m}$ photons/ m^2/s . After optimal growth, cultures were centrifuged to separately collect cell pellets and supernatants for genomic and metabolomic analyses. Two complementary extraction protocols were used for metabolomic analyses to maximize coverage (see section 2.8).

2.2 Shotgun metagenomic sequencing

DNA was extracted from lyophilised cultures (50 mg) [30] using the DNeasy PowerLyzer PowerSoil kit (Qiagen) following an adapted protocol (manufacturer's instructions from step 5) without bead beating [76]. Cells were lysed with PowerBead Solution (750 μL) and Solution C1 (60 μL) by two 5-minutes incubation at 70°C with 30s of vortexing before and after each incubation. Long-read shotgun metagenomic sequencing was performed using a PacBio Revio platform (Pacific Biosciences, Menlo Park, CA, USA) at the Gentyane sequencing facility (Clermont-Ferrand, France) with detailed protocols in Supplementary Material and Methods.

2.3 Metagenomic assembly and binning

Reads were assembled using metaMDBG (v1.1) [5], and binned with metaBAT 2 (v2.12.1) [46]. Bin quality was assessed with CheckM2 [17] (v1.1.0). High quality (HQ) metagenome-assembled genomes (MAGs, completeness >90%, contamination <5%) were taxonomically assigned using GTDB-tk (v2.4.1) [16] with the r226 GTDB release as reference [66].

MAG abundances and assembly statistics were calculated with Mapler (v2.0.1) [58]. Contigs from non-HQ bins and unbinned contigs were concatenated to create 12 pseudogenomes (PG) containing the remaining sequences associated to each phycosphere.

2.4 Phylogenomic analysis of *Microcystis*

The phylogenetic placement of the 12 *Microcystis* strains was inferred using a dataset of 349 genomes and method previously described in [34, 57]. Briefly, after predicting the coding sequences of 361 *Microcystis* genomes (including the 12 strains) using Prodigal (v2.6.3), an alignment of 779 concatenated single-copy genes shared by all strains in the analysis (674,919 positions) was generated by Roary (v3.13.0, default parameters except for a protein identity threshold $\geq 90\%$). Spurious and misaligned regions were removed using trimAl (v1.5) before phylogenetic inference was computed by RaxML (v8.2.12) with the following parameters: GTRGAMMA model, 100 bootstraps. In addition, the average nucleotide identity (ANI) was calculated between each pair of strains in the tree using fastANI (v1.33) [43] to define *genospecies* (ANI $\geq 97\%$, [9]), and *genotypes* (ANI $\geq 99\%$). Details about the tree construction are available in Supplementary Material and Methods.

2.5 Gene prediction and functional annotation

Gene prediction of MAGs and PGs was performed with Prodigal (v2.6.3) [41]. Functional annotations were obtained with eggNOG-mapper (v2.1.12) [39], relying on eggNOG database (v5.0.2) [11], and diamond for sequence alignment (v2.1.9) [8].

Biosynthetic Gene Clusters (BGC) were identified using antiSMASH (v7.0.0) [6] with default parameters. Additional targeted annotations were curated using reciprocal blast searches based on known *Microcystis* metabolic capabilities, as explained in Supplementary Material and Methods. Functions related to biogeochemical cycles were assessed using METABOLIC (v.4.0) [87].

2.6 Genome-scale metabolic network reconstruction and comparisons

Genbank files were generated with emapper2gbk (v0.3.1) [4] from gene prediction and functional annotation. Genome-scale metabolic networks (GSMNs) were reconstructed from annotated genomes using Pathway Tools (v27.0) [47] via the mpwt pipeline (v0.8.6) [4]. PADMet library (v5.0.1) [2] was used to manipulate GSMNs and create SBML files [38]. This library was also used to get a description of the GSMNs (*i.e.* size, reaction presence, pathway completeness), and for the curation and integration of missing reactions. For the latter process, we first used the double-blast procedure described in Supplementary Material and Methods for protein annotation, then drafted the reactions catalysed by the identified proteins for integration in corresponding GSMNs with PADMet.

Regarding the pseudogenomes, a single GSMN was reconstructed from concatenated unbinned contigs and low-quality bins per phycosphere, retaining only gene-associated reactions.

Details of GSMNs comparison are described in Supplementary Material and Methods. GSMNs comparisons and metabolome-to-network mapping were performed using MetaNetMap (v1.0.1) [62] (details in Supp. Material and Methods).

2.7 Metabolic Modelling

Community metabolic potential was assessed using the Metage2Metabo (v1.6.1) [4] pipeline integrating MisCoTo (v3.2.0) [33] and MeneTools (v3.4.0) [2], based on network expansion (a Boolean abstraction of metabolic producibility) from seed metabolites corresponding to BG11 medium composition (Table S1) to assess reachability of metabolites [29] (details in Supp. Material and Methods).

Two modelling scenarios were explored: i) intra-phycosphere metabolic complementary under laboratory culture conditions, and ii) inter-phycosphere metabolic complementary simulating pond-like environmental conditions (details in Supp. Material and Methods).

2.8 Metabolomic data acquisition

Intracellular metabolites from the 12 phycosphere cultures were collected by centrifugation (3,200 g 15°C, 10 min), lyophilised and extracted using complementary monophasic (MP) [28], mainly dedicated to capture specialised metabolites, and biphasic protocols (BP) [59], to maximize metabolome coverage. Details about the MP and BP protocols are provided in Supplementary Material and Methods.

2.9 Metabolomic annotation and chemometrics

Feature extraction and annotation of metabolomic signals were performed according to two different workflows. At first, for MP extracts, the mass spectrometry (MS) data were processed using MetaScape (v4.0) software (Bruker, Bremen, Germany) in order to generate a data matrix containing semi-quantification results for each analyte in all analysed samples (see details in Supp. Material and Methods). Then, MS-DIAL (v5.25) [79] and MS-cleanR were used to process and clean both MP and BP data and interrogate FragHub spectral libraries [21]. SIRIUS software (v6.1) [27] was used to interrogate structural libraries (SIRIUS database and cyanoMetDB [45]). Output files from MS-DIAL and SIRIUS were further imported into MS-Net [32] in order to suppress analytical redundancy, provide annotation confidence level and classification (NPclassifier, [48]). Annotation workflow parameters are provided in Supplementary Material and Methods. Following MS-Net, spectral library matches were stratified as level 1 (spectral cosine > 0.95 and $\Delta RT < 0.2$ min), level 2a (cosine > 0.85) or 2b ($0.7 \leq \text{cosine} \leq 0.85$). *In silico* candidates were supplemented with taxonomic information by querying COCONUT 2.0 [15] using InChiKey-based matching. Among the candidates per feature, those matching taxonomic criteria (identified in bacterial genus or in *Microcystis*) were classified at level 3a and the remaining *in silico* matches stayed at level 3b. Spectral library matches with cosine < 0.7 or significant precursor mass errors were classified as level 4 (MS/MS analogues).

Prior chemometrics, replicates were combined through averaging as to keep one signal per metabolite for each phycosphere. Feature intensity was normalised, transformed and scaled (*i.e.* sum normalisation, cube root transformation and pareto scaling). A principal component analysis (PCA) and a hierarchical clustering analysis (HCA) were performed and annotated metabolites were summarised according to chemical classes (NPC classification [48]).

2.10 Statistical analysis and data visualisation

Data analysis were handled with python (v.3.12.3), pandas (v.2.2.2), numpy (v.2.1.0), or scikit-learn (v.1.6.1), as well as R (v.4.4.1) package vegan (v.2.6.6.1). Visualisation was generated using python matplotlib package (v.3.9.2), and R package dendextend (v.1.17.1) for tanglegrams.

Phycosphere taxonomic trees were constructed in newick format using GTDB-tk results [67], and displayed with the iTOL tree display and annotation tool [52]. GSMNs similarity across phycospheres were assessed by performing KR Clarke analysis of similarity [19], and visualised with Principal Coordinates Analysis (PCoA). For the latter, we considered the PCoA satisfactory if the condition $stress - 1 < 0.1$ was respected, according to [50].

3 Results

3.1 Phylogenomic and potential metabolism of the 12 *Microcystis* strains

A total of 9,006,345 HiFi metagenomic reads (67.9 Gbp) was obtained, ranging from 516,628 and 1,021,017 per phycosphere community, with median contig lengths between 5,973 to 7,951 bp. One complete *Microcystis* genome (single contig) was reconstructed from each non-axenic culture, with an average size of 5.3 Mbp.

3.1.1 Phylogenomics of *Microcystis* strains

Microcystis phylogeny was inferred from a concatenated alignment of 779 single-copy genes (674,919 aa positions) shared by 361 strains. The twelve strains captured local *Microcystis* diversity and were distributed across the phylogenomic tree, with *Microcystis* sp. PMC1371.22 and *Microcystis* sp. PMC1372.22 closest to the root (Fig. 1A). Using ANI threshold $\geq 97\%$ [9], 21 genospecies (A-U) gathering multiple strains were identified encompassing 352 strains. The twelve strains belonged to six distinct genospecies

(C, D, F, I, S, U). All strains within a genospecies belong to a single genotype (ANI \geq 99%), except for genospecies I, which three strains belong to three distinct genotypes.

3.1.2 Potential functional metabolism of *Microcystis*

GSMNs reconstructed for the twelve *Microcystis* genomes were highly similar (from 1,780 to 1,823 reactions; Table 1), consistent with their phylogenetic proximity (Fig. 1A). The quality of GSMN reconstruction was validated by the congruence between metabolic distances across strains and GSMN contents (Fig. 1B). This suggests that the reconstruction captured clade-specific metabolic features. *Microcystis* GSMNs were identical within genospecies D (PMC1374.22 and PMC1367.22) and U (PMC1375.22 and PMC1356.22), although undetected differences may exist due to annotation limits, and the prevalence of unknown cyanobacterial genes [12].

Specialised metabolism (SM) largely encoded by biosynthetic gene clusters (BGCs) was surveyed across genomes (Fig. 1C, Supp. File 1). BGC content was strictly conserved within genotypes and more loosely conserved within genospecies [40], as illustrated by genospecies I, which exhibits greater genetic divergence ($97\% < \text{ANI} < 99\%$) among its three strains belonging to distinct genotypes. Only five strains from genospecies U, D and I carried microcystin-associated BGCs while carotenoid-associated BGCs were detected in all strains. *Microcystis* PMC1365.22 (genospecies S) lacked most BGCs. Most SM products were detected in intracellular metabolomes, except microviridin, detected only in PMC1364.22, potentially due to sequence variability and annotation limits.

SM-associated reactions were absent from GSMNs and the MetaCyc database [13] highlighting the limited representation of SM in GSMN reconstruction as previously established [73]. To address this, the microcystin-LR biosynthetic pathway was manually curated (Fig. 1D), integrating literature (Sup. Table S8) and BGCs annotations. Overall, *Microcystis* strains shared a conserved core metabolism with phylogenetically structured SM diversity, including variation in pigment and toxin production.

3.2 Taxonomy and functions of the associated microbiome

3.2.1 Taxonomic composition of phycosphere's associated bacteria

Ninety-six high-quality (HQ) MAGs (5-13 per phycosphere) of associated bacteria (AB) were recovered, including 62 MAGs nearly, or fully complete and assembled in a unique contig (Supp. File 1). HQ MAGs accounted for $> 90\%$ of read abundance in most phycospheres, with *Microcystis* genomes showing the highest coverage (82X to 1,115X) consistent with culture conditions (Supp. Fig. S1). Cyanobacterial relative abundance varied among phycospheres, with *Phyco1365* (associated to *Microcystis* PMC1365.22), *Phyco1356* (associated to PMC1356.22) and *Phyco1375* (PMC1375.22) displaying lower dominance, suggesting distinct community structures (Fig. 2A, Supp. Fig. S1). Unbinned contigs and lower quality MAGs were merged into phycosphere-specific pseudogenomes (PGs) to capture residual functional diversity from low-abundant unassembled or unbinned populations.

Taxonomic assignment of MAGs (Fig. 2A, Supp. Fig. S2) revealed that most AB belonged to the *Alphaproteobacteria*, *Gammaproteobacteria*, or *Bacteroidia*. Species-level composition of phycospheres (Fig. 2C) often mirrored *Microcystis* phylogeny, with several MAGs associated with specific genospecies: e.g. *Neoroseomonas sp020249415* with genospecies I, *JANRLA01* (Bacteroidia) and *Burkholderiaceae UBA4615_sp020248665* with genospecies C. In contrast, *Burkholderiaceae sp036405035* and *Sphingomonadaceae 34.65.8* were nearly ubiquitous, indicating stable associations across genospecies. *Fimbrimonadia* were enriched in *Phyco1365* which showed lower cyanobacterial dominance, and also found in smaller proportion in *Phyco1354*.

Phylosymbiosis [72] was assessed by comparing the phylogeny of *Microcystis* with the dissimilarities of the bacterial community (Fig. 2B). The strongest congruence occurred at species and genus levels (Normalised Robinson-Foulds (NRF) metric [75]: 0.55), with clustering by genospecies for several genotypes (D, F, C, and I (2/3 strains)). The dendrogram of genospecies U and its AB was not obviously consistent between the two distinct strains, according to slight AB community variation. Higher taxonomic ranks (family and order) showed weaker strains specificity and congruence (NRF: 0.88 and 0.77 resp. Supp. Fig. S3), indicating that phylosymbiosis may primarily operate at fine taxonomic resolution, that would imply even more pronounced metabolic specificities.

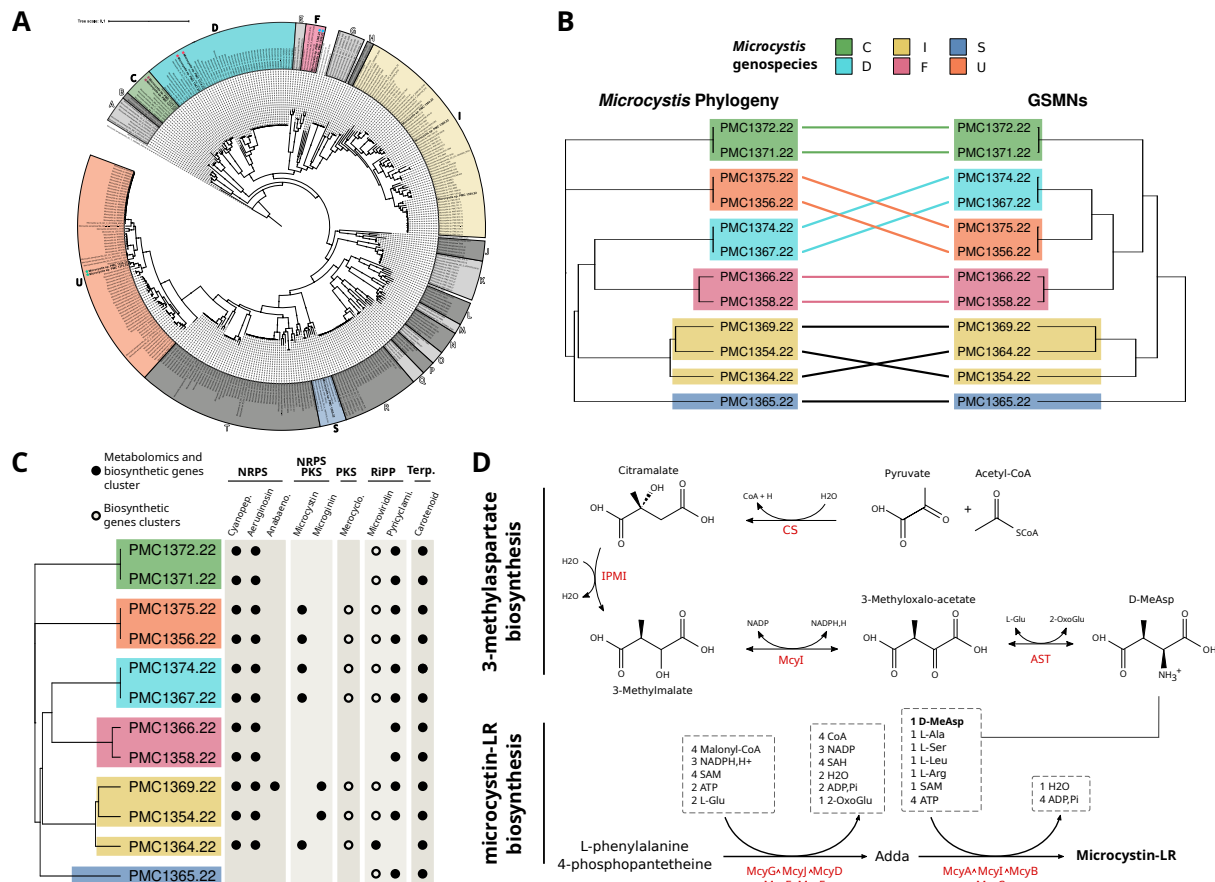


Figure 1: Reconstruction of metabolism for the 12 *Microcystis* strains. **A.** Maximum likelihood phylogenomic tree of *Microcystis* species. The boxes group together strains from the same genospecies, labelled from A to U. The coloured ones (C, D, F, I, S and U) correspond to the genospecies of the surveyed strains (in bold). The stars associated with the labels indicate that the associated strains belong to the same genotype. Tree branches supported by bootstrap values over 60% are indicated by a black circle. The tree is rooted on *M. aeruginosa* Ma_SC.T_19800800_S464, as determined in previous studies [9]. **B.** Tanglegram of *Microcystis* ANI phylogenetic tree on the left, and of the hierarchical clustering of GSMN's reactions presence matrix on the right. Hierarchical clustering was conducted using Jaccard distance and complete-linkage clustering. Species on branches coherent between the two dendrogram are connected by a coloured link. **C.** Specialised metabolism of *Microcystis* strains. On the left, the phylogenetic tree of the 12 *Microcystis* strains. On the right, matrix of the specialised metabolites biosynthetic genes cluster (BGC) presence in genomes, and metabolite identification in the phycosphere according to metabolomics from multiplexed LC-HRMS/MS analyses (Supp. Material). **D.** Microcystin-LR and 3-methylaspartate biosynthesis, according to literature and identified genes in the twelve *Microcystis* genomes. At the bottom, reactions of the MC-LR biosynthesis are presented as integrated in the GSMNs, that is as two lumped reactions. **Abbreviations:** Anabeno., anabenopeptin; Cyanopep., cyanopeptolin; Merocyclo., merocyclophane; NRPS, non-ribosomal peptides synthase; PKS, polyketide synthase; Pyricyclami., pyricyclamide; RiPP, ribosomally synthesized and post-translationally modified peptides; Terp., terpenoids; 2-OxoGlu, 2-oxoglutarate; Adda, (25,35,85,9S)-3-amino-9-methoxy-2,6,8-trimethyl-10-phenyl-4,6-decadienoic, AST: aspartate transaminase; CS, citramalate synthase; D-MeAsp, D-methylaspartate; IPMI, isopropylmalate isomerase acid.

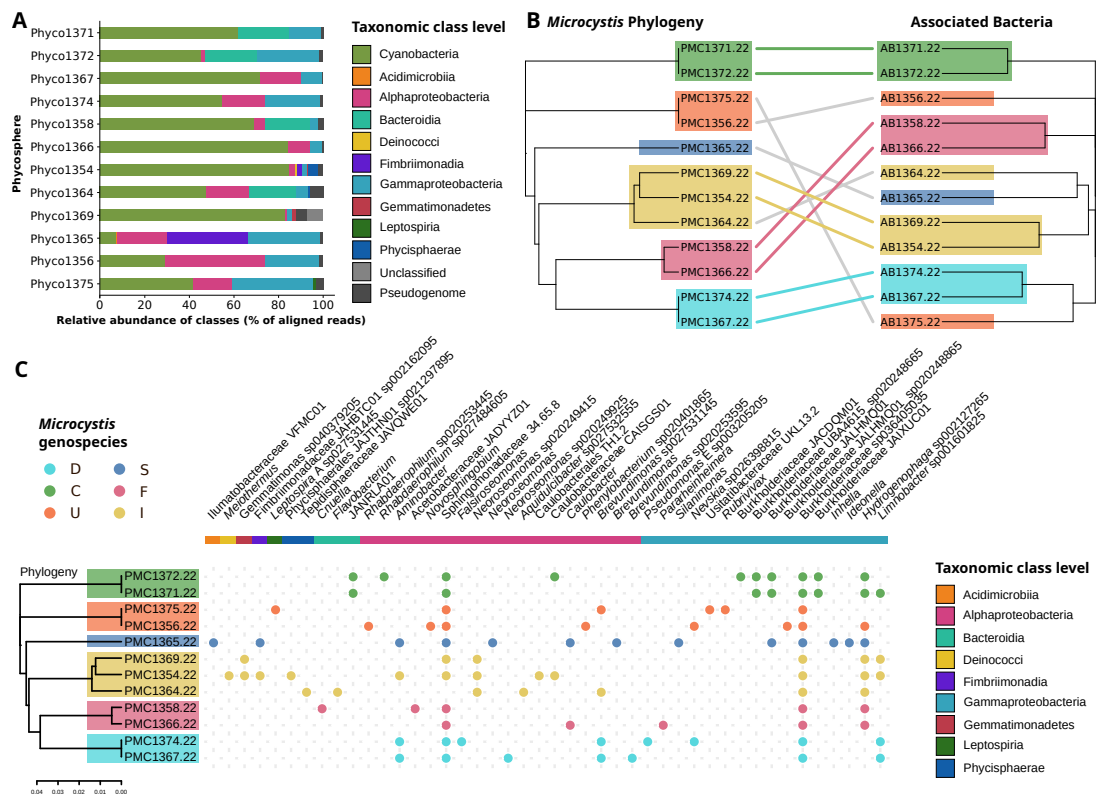


Figure 2: Metagenomic assembly of the 12 phycospheres. **A.** Class taxonomy level and relative abundance of the associated bacteria based on metagenomic read mapping **B.** Relationships between *Microcystis* strain phylogeny and the taxonomic composition of the respective associated bacteria according to hierarchical classification **C.** Heatmap of the bacterial MAGs occurrence within the twelve phycospheres.

3.2.2 Metabolic potential of phycosphere's associated bacteria

GSMNs of AB and PGs were heterogeneous and generally larger than *Microcystis* networks, averaging 2,010 reactions (Table 1). Some PGs yielded small networks (*e.g.* *Phyco1367*) whereas others were larger but largely redundant with the rest of the phycosphere. PGs of *Phyco1366* and *Phyco1375* contained many genes and brought the most additional unique reactions to the community, with respectively 14.5% and 23.5% of unique reactions, suggesting the presence of low-abundance - uncaptured - populations in those phycospheres with distinct metabolic capabilities.

To further investigate functional redundancy and complementarity within phycosphere communities, we performed a rarefaction analysis of the number of enzyme commissions (EC) and other functional annotations in *Microcystis* and AB genomes, excluding PGs (Fig. 3A, Supp. Fig. S4). The absence of a plateau indicates that even in these small cultivated phycospheres, each member likely contributes unique functions. However, the curve inflected beyond five members, suggesting diminishing functional novelty with increasing community size. The union of all reconstructed phycosphere genomes set an upper limit of ~ 300 additional EC numbers, whereas PGs contributed ~ 100 EC numbers absent from other genomes. Analysis of genes related to biogeochemical cycles (Supplementary Material) revealed clear functional partitioning between community members. Cyanobacteria and AB differed in nitrogen metabolism, whereas heterotrophic bacteria encoded pathways for the degradation of complex carbon substrates.

We next compared metabolic reaction repertoires across phycospheres using three configurations: all members including PGs, all members excluding PGs, and AB only (Fig. 3B). Complete phycospheres showed a significant but moderate clustering according to *Microcystis* genospecies (KR Clarke analysis of similarity [19], $R = 0.539$, p-value= 0.007). This pattern persisted when PGs were excluded ($R = 0.554$, p-value= 0.005) but weakened when only AB were considered ($R = 0.317$, p-value= 0.084). In all analyses, phycospheres associated with genospecies I displayed the highest functional diversity.

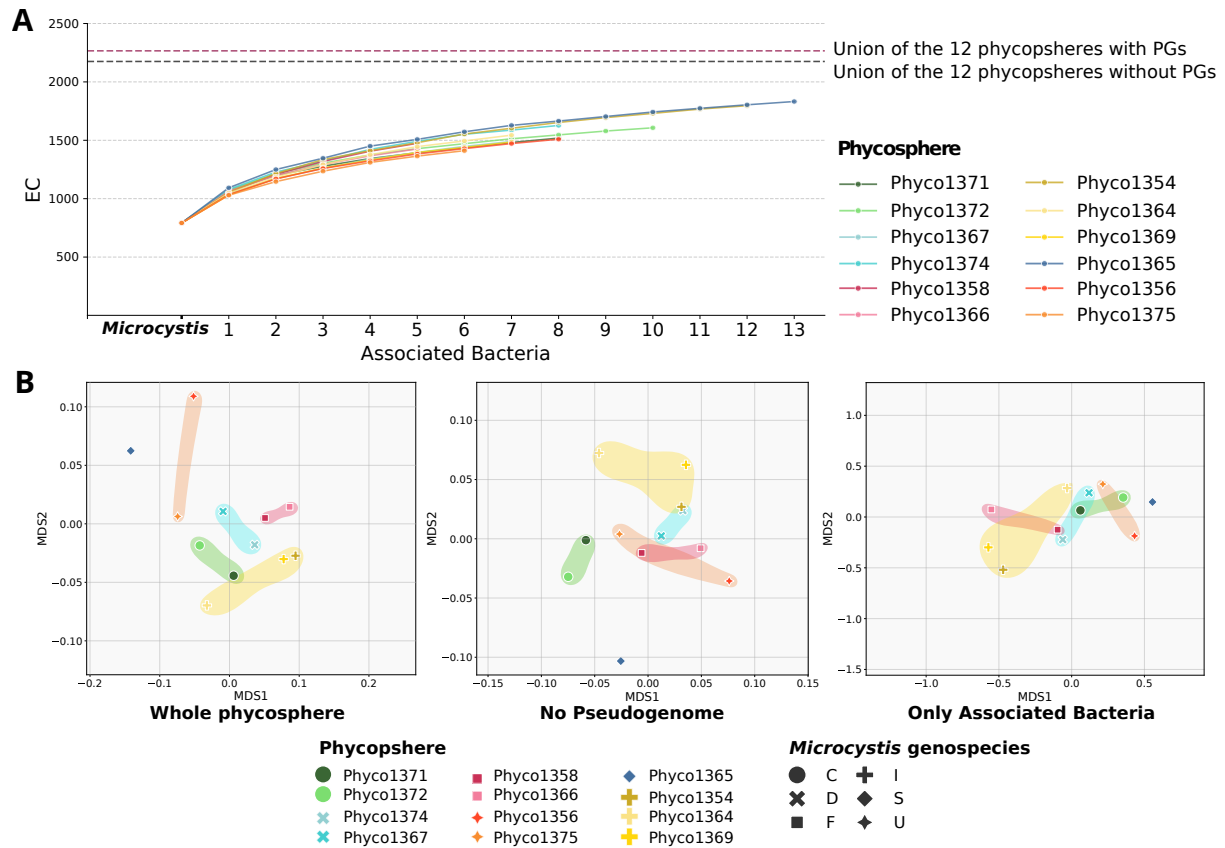


Figure 3: **Metabolic complementarity within phycospheres** **A.** Rarefaction curve of EC numbers annotated in phycosphere metagenomes. Rarefaction curves represent the accumulation of non-redundant EC numbers as a function of the number of sampled associated bacteria, based on 100 sampling iterations. The starting point is *Microcystis* EC numbers. PGs are excluded from the sampling of phycosphere's individual curves. **B.** Principal Coordinates analyses of the GSMNs reactions content, weighted by species abundance and grouped by phycosphere. From left to right ; for the whole phycospheres, without the pseudogenome, and with only the associated bacteria. Shaded areas show phycospheres associated to *Microcystis* strains belonging to the same genospecies. **Abbreviations:** EC, enzyme commission number; PG, pseudogenome

3.3 Meta-metabolomes of the phycospheres

The merged endometabolome features (from both MP and BP extractions; Supp. Fig. S5, Fig. 4A) of the 12 phycospheres comprised 20,931 features: 12,608 with only MS1 and 8,323 with MS2 spectra. Among these, 0.6% (135 features) and 0.3% (63 features) were annotated into levels 2a and 2b, respectively. Including *in silico* candidates (level 3a, 144 features and 3b, 5674 features), up to around 31% of features were putatively annotated. Both HCA (Fig. 4A) and PCA (Supp Fig. S6) revealed clustering of phycospheres according to *Microcystis* genospecies, consistent with a massive cyanobacterial-driven metabolite production and likely influenced by high *Microcystis* dominating biomass within cultures, that is also fully congruent with our metagenomic observations considering the above 1000-times higher biovolume of *Microcystis* cells compared to heterotroph AB (Fig. 2A). Only slight deviations were observed with *Phyco1369* diverging from genospecies I and *Phyco1365* showing a metabolomic profile closer to genospecies U (Fig. 4A). Annotated metabolites (6,016 features with L2 or L3 confidence) were categorised by NPC classification (Supp. Fig. S7) and the 20 most represented superclasses were highly similar across phycospheres (Fig. 4B). Only 144 of the annotated features could be specifically linked to taxa detected by metagenomics, mainly *Microcystis*, *Pseudomonas*, and *Burkholderiaceae* (details, including putative ecological roles and compound identities, in Supp. Material, Supp. Fig. S8).

The oligopeptide NPC_superclass (14% of annotated compounds), showed strong variability among phycospheres, with genospecies-level similarities (Fig. 4C). Actinomycins and linear-lipopeptides were specific to genospecies U and C, respectively, while microcin SF608, likely produced by AB, was enriched in genospecies U. Microcystins (e.g. microcystin-RR, MC-LR) were most abundant in *Phyco1374* and

Phycosphere	Cyanobacteria			Associated Bacteria (average data)			Pseudogenome				
	Genes	Reactions	Reactions with genes as-sociation	High Quality MAGs	Genes	Reactions	Reactions with genes as-sociation	Genes	Reactions	Added value	
C	Phyco1371	5251	1790	68.1 %	8	3498.4 ± 893.7	2029.0 ± 287.8	67.3 % ± 0.8	9677	1937	12.1 %
	Phyco1372	5256	1791	68.1 %	10	3538.6 ± 859.2	2068.1 ± 278.4	67.4 % ± 0.9	19128	2359	9.0 %
D	Phyco1367	5501	1792	68.2 %	7	3622.9 ± 1179.2	1971.6 ± 267.7	67.5 % ± 1.7	567	183	4.4 %
	Phyco1374	5491	1792	68.2 %	8	3968.5 ± 1130.1	2075.4 ± 318.4	67.9 % ± 2.0	15188	2123	5.0 %
F	Phyco1358	4897	1784	68.3 %	5	3949.2 ± 660.2	2014.0 ± 216.5	67.1 % ± 1.4	15773	2121	12.8 %
	Phyco1366	4944	1780	68.2 %	5	3757.2 ± 596.3	2045.2 ± 204.1	67.9 % ± 2.2	11191	1960	14.5 %
I	Phyco1354	4909	1823	67.7 %	12	3718.8 ± 1074.1	1971.9 ± 273.7	67.8 % ± 1.4	15702	2054	4.3 %
	Phyco1364	4864	1796	68.5 %	7	3760.0 ± 948.2	1941.6 ± 258.5	67.5 % ± 0.5	15255	2229	11.8 %
	Phyco1369	5323	1795	68.9 %	7	3620.4 ± 737.1	1960.9 ± 212.8	67.2 % ± 0.8	13385	2212	13.4 %
S	Phyco1365	4627	1795	68.6 %	13	3967.8 ± 1167.2	2083.4 ± 263.7	68.0 % ± 1.6	14651	2265	5.7 %
	Phyco1356	5784	1795	67.8 %	8	3419.4 ± 601.4	1948.4 ± 191.4	67.1 % ± 1.4	11426	2036	9.9 %
U	Phyco1375	5774	1795	67.8 %	6	3483.0 ± 807.4	1943.7 ± 284.1	67.0 % ± 1.1	33414	2839	23.5 %

Table 1: **Description of the twelve phycosphere metagenomes and their reconstructed GSMNs.** For each community, a distinction is made between *Microcystis*, its associated bacteria and the pseudogenome (PG) that is reconstructed from the remaining contigs of the assembly. PG's reactions introduced in the phycosphere is the percentage of reactions of the pseudogenome's metabolic network that are not in any of the other phycosphere's members GSMN. On the very left column and coloured, the *Microcystis* genospecies is indicated. Bold: minimum and maximum values.

Phyco1367 (genospecies D), *Phyco1375* and *Phyco1356* (genospecies U), and especially in *Phyco1364*, matching BGC-based predictions (Fig. 1C). Additionally, a few low-intensity putative microcystins were detected in strains lacking microcystin BGCs. These are likely false positives arising from limitations of automated metabolite annotation, as certain micropeptins/cyanopeptilins share structural features with microcystins but can be manually distinguished based on the presence of specific diagnostic ions.

Saccharides, for their potential role in cross-feeding [61], were examined and varied by genospecies (Fig. 4D); for example, paulomycins were most abundant in genospecies F phycospheres, and some monosaccharides were detected exclusively in *Phyco1365*, the sole representative of genospecies S (Fig. 1C, 2A). No relevant metabolites were detected in the exometabolome beyond culture medium components, suggesting a rapid consumption of extracellular components by heterotroph AB.

Mapping metabolomic features to phycosphere GSMNs yielded only 133 matches (Fig. 4E). Mapped metabolites were mainly amino acids, amino sugars, dipeptides, purines, pyrimidines, and nucleosides, with variable distribution across GSMNs. Notably, carotenoids were mapped only to GSMNs of AB, including PGs, that may correspond to a false negative regarding *Microcystis* which carotenoid biosynthetic genes remain poorly characterised and further automatically annotated.

3.4 Modelling the intra- and inter-phycosphere metabolic complementarity

To investigate metabolic complementarity within and across phycospheres, we simulated metabolite production for each GSMN under BG-11 medium conditions, *i.e.* we identified the metabolites of the GSMNs that could be reached from these nutrients. Predictions distinguished metabolites produced individually from those requiring mutualistic interactions. We tested two scenarios: (i) separate simulations for each phycosphere to assess within-community complementarity, and (ii) a merged network simulation to predict larger-scale interactions between phycospheres (Fig. 5)

Intra-phycosphere metabolic potentials and interactions are shown in Fig. 5A. Under the simulated conditions, *Microcystis* could access a core of over 400 metabolites, largely autonomously. Additional metabolites became accessible only through metabolic interactions with AB and PGs, with variability among phycospheres. Differences among cyanobacterial GSMNs were largely offset by community composition (Fig. 1, Table 1). In contrast, AB metabolic potential varied despite a core of 510 shared metabolites and did not align cyanobacterial genospecies (analysis of similarity [19], $R=0.355$, $p=0.07$). PGs unlocked additional pathways, contributing approximately 25 unique metabolites per phycosphere. *Phyco1364* and *Phyco1375* shared unique predicted cyanobacterial metabolites, mainly terpenoids. Comparisons between model predictions and metabolomics are detailed in Supp. Material.

Inter-phycosphere simulations predicted access to approximately 100 additional metabolites per community (Fig. 5B), including fatty acids, shikimates, phenylpropanoids, carbohydrates, and terpenoids. Notably, terpenoid functions from *Phyco1364* and *Phyco1375* spread to other communities, suggesting transferable metabolic benefits.

Together, these results demonstrate that metabolic complementarity enhances metabolite accessibility both within and between phycospheres and is shaped by phycosphere composition.

4 Discussion

Microcystis metabolism reflects genospecies structure

Microcystis genomes are highly plastic [12], with non-ribosomal peptide synthase and polyketide synthase gene clusters accounting for 5% of the genome [78], and up to 240 natural products classified as specialised metabolites through genomics [51, 25]. Given their high versatility, capturing metabolic variation across genospecies is therefore essential. Our GSMN reconstructions revealed a robust consistency between metabolic reaction content and genospecies phylogeny. BGCs repertoires were similarly conserved within genospecies, except for genospecies I which displayed greater heterogeneity. Metabolomic analyses further confirmed the production of multiple microcystin variants by *Microcystis*. However, their biosynthetic pathways were largely absent from GSMNs due to incomplete database representation. Manual curation of the MC-LR pathway in GSMN underscores the substantial efforts still required to incorporate SM into genome-scale models and could benefit from existing efforts in BGC description [73, 85].

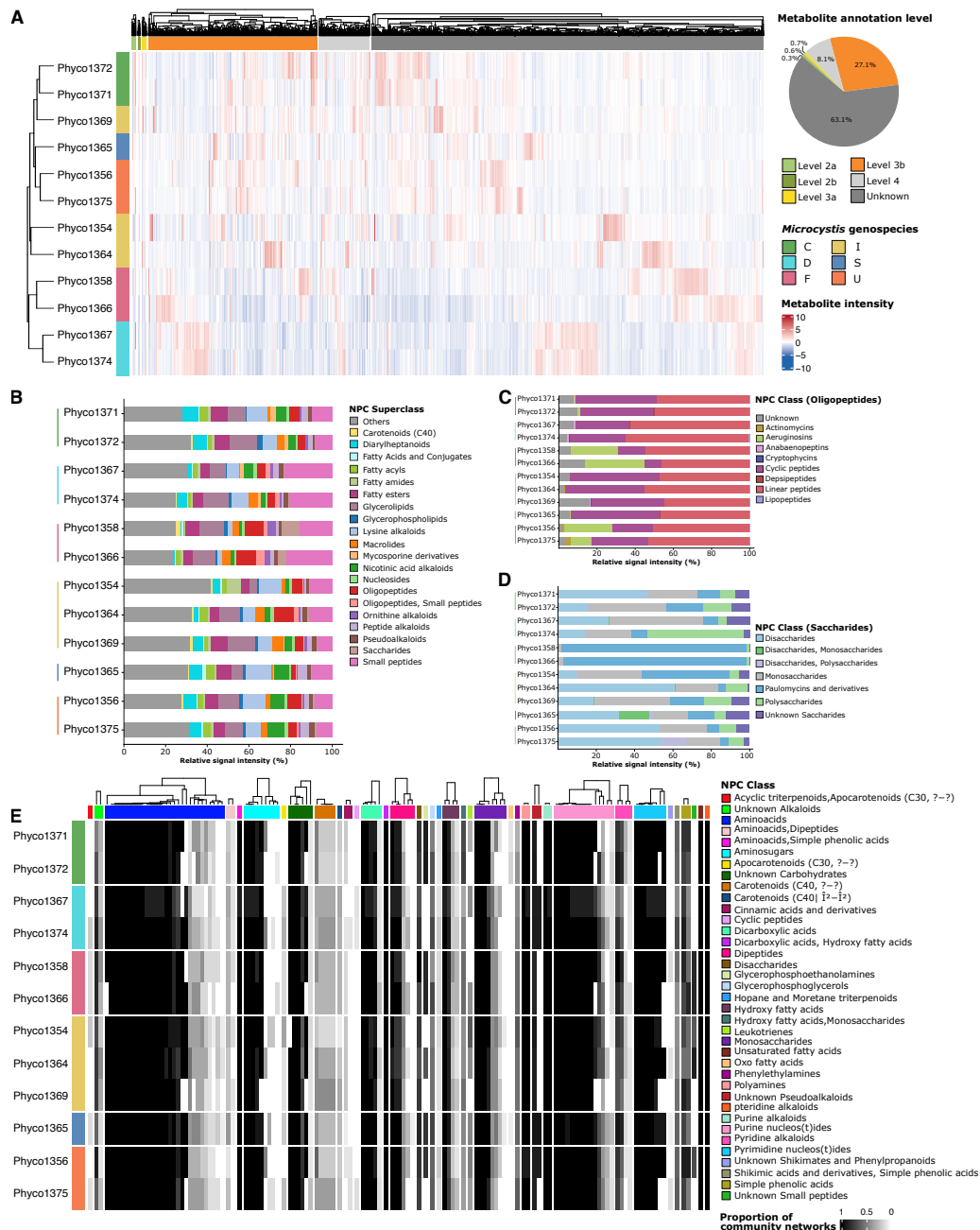


Figure 4: **Metabolomic analysis of the 12 phycospheres.** **A.** Hierarchical clustering analysis on phycospheres transformed endometabolome signals. For each phycosphere, triplicates were averaged to merge the different extractions results and get one set of transformed signals per community. Metabolites annotations was conducted as previously described, and evaluated with levels assigned as described in Methods. Hierarchical clustering was based on euclidean distance and complete linkage method. **B.** Barplot depicting the relative abundance of L2 and L3 annotated metabolomic signals in phycospheres stratified by NP Classifier superclasses. Only the 20 most abundant superclasses are depicted, others being aggregated in *Others*. **C.** Focus on the annotated metabolomic signal classified as "Oligopeptide" superclass (relative abundance). The 20 most abundant classes are depicted, others being aggregated in *Others*. Cyclic peptides include microcystins, additional toxins and other specialised metabolites, such as aeruginosins, or cyanopeptolins. **D.** Focus on the annotated metabolomic signal classified as "Saccharides" superclass. Relative abundance of Saccharides classes in the twelve phycospheres. **E.** Heatmap depicting the mapping of L2-L3 annotated metabolomic signals onto the GSMNs of all phycospheres members. Heatmap values represent the proportion of GSMNs per phycosphere that contain the related metabolite. Identification in GSMN means the metabolite is a substrate or a product of a metabolic reaction. Metabolites are stratified by their NP-Classifier classes. Cyclic peptides correspond to Microcystin-LR. Carotenoids include echinenone, canthaxanthin, zeaxanthin, adonixanthin and dike-tospirilloxanthin. **Abbreviations:** NPC, NP Classifier.

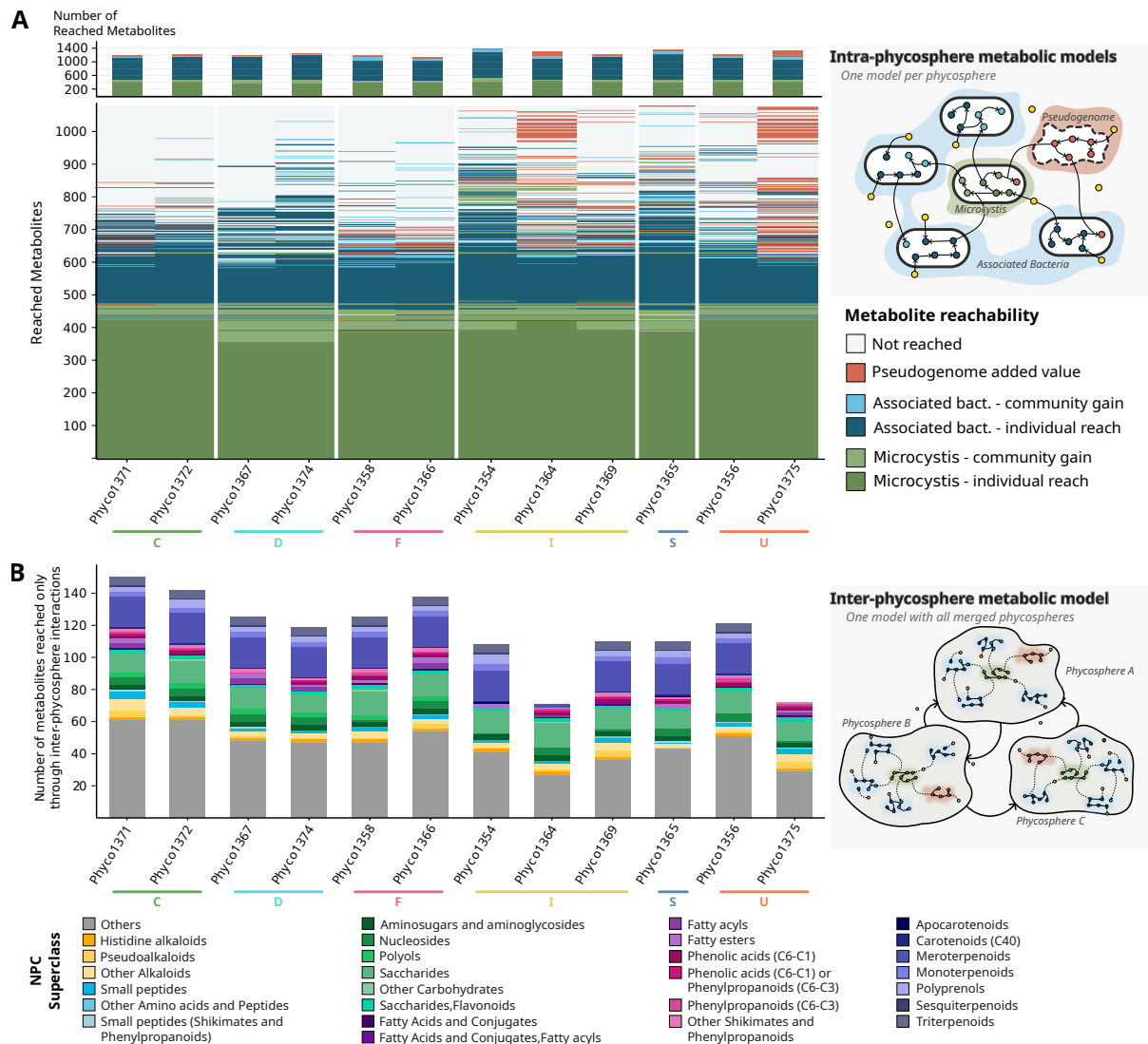


Figure 5: Modelling the metabolic complementarity intra- and inter-phycospheres A. Heatmap of metabolites reach across each phycosphere metabolic capabilities modelling, grouped by genospecies (C, D, F, I, S, and U). Producers of reachable metabolites were grouped to distinguish the roles of *Microcystis*, associated bacteria (AB), and pseudogenomes (PGs), as well as to assess the impact of interactions within and between these groups. Each row is a metabolite, preferably shown as reached by the cyanobacteria, when it is also reached by one of the AB. Non-overlapping reachability is presented in Supplementary Figure 9. "Individual reach", metabolites produced by at least one community member in the phycosphere; "Added value", metabolites reached only through metabolites exchanges; "not reached", metabolites not reached in modelling or absent from GSMNs. The upper barplot enumerates the (possibly overlapping) metabolites reached in each of the observed situation (by *Microcystis*, AB, through the pseudogenome). **B.** Classification of inter-phycosphere added-value. For each phycosphere, metabolites that are reached by one of their member only through phycosphere-phycosphere exchanges are classified according to their NPC Superclass level.

Phycospheres harbour a reduced taxonomic diversity driven by *Microcystis* phylogeny

The reduced taxonomic diversity observed in phycospheres is consistent with culture conditions [88, 49, 37], and reflects selective pressure favouring *Microcystis* growth thus prioritising AB that are essential for it, or at least non-antagonistic [49, 83]. It was previously reported that bacterial communities in the phycosphere are shaped by the genotypic characteristics of *Microcystis* [72]. Despite the limited dataset, this observation implies that bacterial partners conserved through cultivation process may present specific selective added-value that directly depends on *Microcystis* genospecies. The molecular mechanisms behind, potentially mediated by metabolic complementary or selective processes, remain however unknown so far.

Dominant taxa, including Caulobacterales and Burkholderiales, are frequently reported in cyanobacterial phycospheres in both environmental and lab studies. The ubiquitous presence of *Hydrogenophaga* sp. aligns with its recurrent association with *Microcystis*, in natural environments and lab cultures [49, 54, 88], while *Sphingomonadales*, known for microcystin tolerance and degradation [24, 81], were systematically detected. While the metabolomic and BGC analyses confirmed microcystin production within the phycospheres, the absence of known microcystin degradation genes (*mlr*) suggests that tolerance rather than active degradation, may be the main adaptation mechanism of these bacterial partners.

Functional complementarity within the phycosphere, but a cyanobacteria-driven metabolome

Comparative functional analyses and GSMNs revealed substantial diversity and metabolic complementarity among phycosphere members. Notably, PGs contributed up to 23% unique metabolic reactions to phycosphere, highlighting the underestimated role of low-abundance community members. While AB functions alone did not reflect *Microcystis* genospecies, genospecies-specific signals emerged when considering the full phycosphere. *Microcystis* strains likely influence their communities via their own metabolism [72] but the lack of a clear signal at the global AB genomic level suggests that community shaping depends more on specific metabolic functions or environmental responses. Functional complementarity included nitrate and nitrite reduction by AB, potentially supporting redox balance and nutrient recycling [54, 37]. In contrast, thiosulfate oxidation genes were absent from *Microcystis* but present in *Limnobacter* spp. in half of the phycospheres, consistent with prior detection during bloom events [54]. Yet, metabolism captured by metagenomics and annotation likely reflects only a fraction of the total metabolic landscape, as illustrated by the limited fraction of annotated metabolites that could be directly linked to GSMNs, highlighting opportunities for improvement [31].

Metabolomic annotations captured only a small fraction of all detected signals, with fewer than 2% annotated at level 2 and most assigned *in silico*. The absence of annotated exometabolites prevented formal inference of metabolite exchange, possibly due to optimal growth conditions or rapid metabolite turnover. Nonetheless, endometabolome analyses demonstrate the value of combining multiple extraction and acquisition strategies to reveal chemical diversity beyond what metagenomics alone can capture.

Endometabolomic profiles mirrored *Microcystis* phylogeny, reflecting cyanobacterial dominance in biomass and metabolites production. Most annotated compounds belonged to small peptides, glycerolipids, and oligopeptides, which discriminated phycospheres. The tight coupling between *Microcystis* genospecies and chemical profiles aligns with recent reports linking genotypes and chemotypes in environmental phycospheres [40]. Among the phycosphere-specific chemical signatures, we surveyed the oligopeptide and saccharide signals. Cyanobacterial oligopeptides, known for their structural diversity and inhibitory activities, likely contribute to metabolic plasticity and defence [1, 63, 7]. Saccharide analyses revealed the importance of paulomycins, particularly in genospecies F, although their typical producer *Streptomyces* [53] was absent from high-quality MAGs. This may reflect low-abundance or uncharacterised producers, or limitations of *in silico* metabolite annotation.

Phyco1365 deviated from others with distinct BGC content, taxonomic structure, phycosphere GSMNs, and metabolomic profiles. Differences in endometabolome content are likely due to low cyanobacterial abundance, distinct AB, and the near-exclusive presence of *Fimbriimonadia*, the latter being also observed at a lower abundance in *Phyco1364*, the second endometabolome outlier. These differences suggest altered microbial interactions but may also reflect the incompleteness of reference databases.

Ecological insights from cultivated freshwater phycospheres

Recent studies have shown that metabolomic profiles of environmental *Microcystis* strains closely reflect their genotype, likely shaped by local environmental conditions [40]. Although our study relies on cultivated phycospheres, our results support this genotype–chemotype link and its ecological relevance. While Huré et al. [40] identified 13 chemotypes from 65 strains collected from 13 lakes, here we detected six distinct genotypes within a single site, indicating that multiple genotypes can coexist under similar environmental conditions and that AB patterns may persist under lab-cultivation.

Several annotated compounds have reported allelopathic or bactericidal properties, including sorbistin A2, ornibactin C6/C8, and paulomycins [80, 60, 53], and may contribute to shaping phycosphere bacterial communities. However, as these assignments rely solely on *in silico* annotation, they should be interpreted with caution.

Regarding toxin production, the inability of some *Microcystis* strains to produce microcystins did not result in major differences in the overall phycosphere metabolism. How microcystin-producing and non-microcystin-producing strains interact at the bloom scale remains unresolved, but our modelling approach highlights functional complementarity, supporting potential *phycosphere–phycosphere* interactions at fine spatial scales [35]. These interpretations remain limited by incomplete reference databases, meaning GSMN reconstructions provide only a partial view of metabolism [73]. Given the number of analysed genomes, we therefore relied on discrete modelling approaches [29, 4, 3]. Although cultivated phycospheres do not fully recapitulate bloom conditions, they nonetheless reveal substantial diversity and potential interaction networks, offering valuable insights into freshwater microbial ecology.

Intra- and inter-phycosphere metabolic interactions: ecological perspectives from metabolic modelling

This study explored micro-diversity at multiple organisational levels. We first characterised intra-population genomic micro-diversity within the genus *Microcystis* across twelve naturally co-occurring isolates, consistent with patterns previously reported during bloom events [36, 77]. In parallel, micro-diversity was also observed within the AB communities of each phycosphere, with strong congruence between cyanobacterial genotypes and their bacterial partners. Although isolation and laboratory cultivation reduced community complexity, core *Microcystis*–AB interactions appeared to be preserved across genospecies. Together, these results extend the well-described micro-diversity of *Microcystis* to the phycosphere level, where it is thought to support adaptation to fine-scale ecological niches [36].

Metabolic complementarity simulations further indicated that phycospheres can collectively expand their metabolic repertoire beyond individual organismal capacities. Rather than solely reflecting functional redundancy, this complementarity enables access to otherwise unreachable metabolites and aligns with previous observations of cooperative metabolism in cyanobacteria-associated communities [72]. Maintaining phycosphere micro-diversity at small spatial scales may enhance resource use efficiency by AB and support *Microcystis* growth, contributing to bloom persistence and providing a more relevant framework for modelling bloom dynamics [71].

When modelling interactions among the twelve phycospheres, an additional increase in metabolic potential emerged, highlighting that phycosphere metabolism is shaped not only by *Microcystis*–AB interactions but also by complementary processes operating across neighbouring communities. During bloom events, closely co-occurring phycospheres may thus be viewed as neighbour ecotypes represented by distinct genotypes. We therefore hypothesise that environmental adaptation during blooms should be studied at the metacommunity level, and that maintaining phycosphere micro-diversity through interactions among closely related phycospheres represents a potential ecological advantage.

Overall, the systems biology framework adopted here demonstrates the value of integrating metagenomics, metabolomics, and metabolic modelling to investigate the structure and function of *Microcystis* phycospheres. We show that *Microcystis* phylogeny influences both bacterial recruitment and the chemical landscape, while simulations highlight the potential for metabolic cross-feeding within and across phycospheres. These predictions generate testable hypotheses on mutualistic interactions and resource-sharing strategies in *in situ* metacommunities across contrasting ecotypes that now require experimental validation.

Supplementary information

Supplementary files include the following:

- Supplementary Material *supplementary_material.pdf*: additional text, figures and tables related to methods and results
- Supplementary File 1 *Supplementary_tables.xlsx*: tables summarising the characteristics of MAGs reconstructed from metagenomic data, and antiSMASH results.

Data availability

Raw reads and assembled genomes are available on EBI ENA under accession number PRJEB102163. Metabolomic data is available on MetaboLights [84] with the study identifier MTBLS13909 (reviewer access: <https://www.ebi.ac.uk/metabolights/reviewerdd864ca1-e28b-44b3-9706-532e6ff40fdc>). Processed data related to the analyses are shared on Recherche DataGouv <https://doi.org/10.57745/2P0IKJ>. Scripts to perform data analysis are available on <https://gitlab.inria.fr/comic-project/COMIC-Metabo>.

Competing interests

No competing interest is declared.

Acknowledgements

Experiments presented in this paper were carried out using the PlaFRIM experimental testbed (<https://www.plafrim.fr>), supported by Inria, CNRS (LaBRI and IMB), Université de Bordeaux, Bordeaux INP and Conseil Régional de la Nouvelle Aquitaine. We acknowledge the GenOuest bioinformatics core facility (<https://www.genouest.org>) for providing the computing infrastructure. We acknowledge the METABOHUB research infrastructure (ANR, 11-INBS-0010, <https://www.metabohub.fr/>)

Funding information

This project was funded by INRAE métaprogramme HOLOFLUX project COMIC. CF, CM were supported by the French National Research Agency (ANR) France 2030 PEPR Agroécologie et Numérique MISTIC ANR-22-PEAE-0011. Sample collection, and *Microcystis* phycospheres isolation was funded and part of the project CaCO₃ funded by the Emergence funding from Alliance-Sorbonne Université. LM was funded by the "Fond mécénat environnement et mobilité du Crédit Agricole d'Île-de-France".

Authors contributions statement

Juliette Audemard: methodology, software, formal analysis, investigation, writing - original draft, visualization, writing - review & editing. **Nicolas Creusot**: conceptualization, methodology, supervision, visualization, writing - review & editing, project administration, funding acquisition. **Julie Leloup**: resources, validation, writing - review & editing, supervision, funding acquisition. **Charlotte Duval**: resources. **Sébastien Halary**: resources, data curation, review & editing. **Lou Mary**: formal analysis, visualization, review & editing. **Mélissa Eon**: resources, review & editing. **Thomas Forjonel**: resources. **Mohamed Mouffok**: resources, formal analysis, investigation. **Rémy Puppo**: formal analysis **Elodie Belmonte**: resources, writing, review & editing. **Véronique Gautier**: resources, writing, review & editing. **Jeanne Got**: writing - review & editing. **Marie Lefebvre**: writing - review & editing. **Gabriel V. Markov**: formal analysis, writing - review & editing. **Coralie Muller**: formal analysis, software. **Benjamin Marie**: validation, writing - original draft, writing - review & editing, supervision, funding acquisition. **Binta Diémé**: conceptualization, formal analysis, investigation, methodology, supervision, writing - review & editing, project administration, funding acquisition. **Clémence Frioux**: conceptualization, formal analysis, investigation, methodology, writing - original draft, visualization, supervision, writing - review & editing, funding acquisition.

References

- [1] Ramsy Agha and Antonio Quesada. Oligopeptides as biomarkers of cyanobacterial subpopulations. toward an understanding of their biological role. *Toxins*, 6(6):1929–1950, 2014.
- [2] Méziane Aite, Marie Chevallier, Clémence Frioux, Camille Trottier, Jeanne Got, María Paz Cortés, Sebastián N. Mendoza, Grégory Carrier, Olivier Dameron, Nicolas Guillaudeux, Mauricio Latorre, Nicolás Loira, Gabriel V. Markov, Alejandro Maass, and Anne Siegel. Traceability, reproducibility and wiki-exploration for “à-la-carte” reconstructions of genome-scale metabolic models. *PLOS Computational Biology*, 14(5):e1006146, May 2018. Publisher: Public Library of Science.
- [3] Constanza M. Andreani-Gerard, Natalia E. Jiménez, Ricardo Palma, Coralie Muller, Pauline Hamon-Giraud, Yann Le Cunff, Verónica Cambiazo, Mauricio González, Anne Siegel, Clémence Frioux, and Alejandro Maass. Modeling the emergent metabolic potential of soil microbiomes in Atacama landscapes. *Environmental Microbiome*, 20(1):142, 2025.
- [4] Arnaud Belcour, Clémence Frioux, Méziane Aite, Anthony Bretaudeau, Falk Hildebrand, and Anne Siegel. Metage2Metabo, microbiota-scale metabolic complementarity for the identification of key species. *eLife*, 9:e61968, December 2020. Publisher: eLife Sciences Publications, Ltd.
- [5] Gaëtan Benoit, Sébastien Raguideau, Robert James, Adam M. Phillippy, Rayan Chikhi, and Christopher Quince. High-quality metagenome assembly from long accurate reads with metaMDBG. *Nature Biotechnology*, 42(9):1378–1383, 2024.
- [6] Kai Blin, Simon Shaw, Satria A Kautsar, Marnix H Medema, and Tilmann Weber. The antiSMASH database version 3: increased taxonomic coverage and new query features for modular enzymes. *Nucleic Acids Research*, 49(D1):gkaa978–, 2020.
- [7] Anja Bubik, Bojan Sedmak, Marko Novinec, Brigita Lenarčič, and Tamara T Lah. Cytotoxic and peptidase inhibitory activities of selected non-hepatotoxic cyclic peptides from cyanobacteria. *Biological Chemistry*, 389(10), 2008.
- [8] Benjamin Buchfink, Chao Xie, and Daniel H. Huson. Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, 12(1):59–60, January 2015.
- [9] Haiyuan Cai, Christopher J. McLimans, Jessica E. Beyer, Lee R. Krumholz, and K. David Hambricht. Microcystis pangenome reveals cryptic diversity within and across morphospecies. *Science Advances*, 9(2):eadd3783, 2023.
- [10] Haiyuan Cai, Christopher J. McLimans, Helong Jiang, Feng Chen, Lee R. Krumholz, and K. David Hambricht. Aerobic anoxygenic phototrophs play important roles in nutrient cycling within cyanobacterial Microcystis bloom microbiomes. *Microbiome*, 12(1):88, May 2024.
- [11] Carlos P Cantalapiedra, Ana Hernández-Plaza, Ivica Letunic, Peer Bork, and Jaime Huerta-Cepas. eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. *Molecular Biology and Evolution*, 38(12):5825–5829, December 2021.
- [12] Hengchun Cao, Da Xu, Tiantian Zhang, Qiufang Ren, Li Xiang, Chunhui Ning, Yusen Zhang, and Rui Gao. Comprehensive and functional analyses reveal the genomic diversity and potential toxicity of *Microcystis*. *Harmful Algae*, 113:102186, March 2022.
- [13] Ron Caspi, Richard Billington, Ingrid M Keseler, Anamika Kothari, Markus Krummenacker, Peter E Midford, Wai Kit Ong, Suzanne Paley, Pallavi Subhraveti, and Peter D Karp. The MetaCyc database of metabolic pathways and enzymes—a 2019 update. *Nucleic Acids Research*, 48(D1):D445–D453, 2019.
- [14] Klara Cerk, Pablo Ugalde-Salas, Chabname Ghassemi Nedjad, Maxime Lecomte, Coralie Muller, David J. Sherman, Falk Hildebrand, Simon Labarthe, and Clémence Frioux. Community-scale models of microbiomes: Articulating metabolic modelling and metagenome sequencing. *Microbial Biotechnology*, 17(1):e14396, 2024.

- [15] Venkata Chandrasekhar, Kohulan Rajan, Sri Ram Sagar Kanakam, Nisha Sharma, Viktor Weißenborn, Jonas Schaub, and Christoph Steinbeck. COCONUT 2.0: a comprehensive overhaul and curation of the collection of open natural products database. *Nucleic Acids Research*, 53(D1):D634–D643, 2024.
- [16] Pierre-Alain Chaumeil, Aaron J Mussig, Philip Hugenholtz, and Donovan H Parks. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics*, 2019.
- [17] Alex Chklovski, Donovan H. Parks, Ben J. Woodcroft, and Gene W. Tyson. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. *Nature Methods*, 20(8):1203–1212, 2023.
- [18] John L Chodkowski and Ashley Shade. A synthetic community system for probing microbial interactions driven by exometabolites. *MSystems*, 2(6):10–1128, 2017.
- [19] K. R. Clarke. Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology*, 18(1):117–143, 1993. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1442-9993.1993.tb00438.x>.
- [20] Nicolas Creusot, Floriant Bellvert, Samuel Bertrand, Cédric Bertrand, Simon Colas, Ludovic Cottret, Binta Diémé, Mélissa Eon, Cédric Hubas, Mary-Lorène Goddard, Cyril C. Jousse, Benjamin Marie, Guillaume Marti, Pierre Millard, Pierre Pétriacy, Josep Valls Fonayet, Marie-Virginie Salvia, Mechthild Schmitt-Jansen, and Lin Zi. Unraveling the Microbiome-Metabolome Nexus for Innovative One-Health Solution. August 2025. working paper or preprint.
- [21] Axel Dabanc, Solweig Hennechart, Amélie Perez, Guillaume Cabanac, Yann Guitton, Nils Paulhe, Bernard Lyan, Emilien L. Jamin, Franck Giacomoni, and Guillaume Marti. Fraghub: A mass spectral library data integration workflow. *Analytical Chemistry*, 96(30):12489–12496, 2024. PMID: 39028894.
- [22] Giulia Daly, Veronica Ghini, Alessandra Adessi, Marco Fondi, Alison Buchan, and Carlo Viti. Towards a mechanistic understanding of microalgae–bacteria interactions: integration of metabolomic analysis and computational models. *FEMS Microbiology Reviews*, 46(5):fuac020, 2022.
- [23] Giulia Daly, Veronica Ghini, Alessandra Adessi, Marco Fondi, Alison Buchan, and Carlo Viti. Towards a mechanistic understanding of microalgae–bacteria interactions: integration of metabolomic analysis and computational models. *FEMS Microbiology Reviews*, 46(5), 2022.
- [24] Qin Ding, Xiaolei Song, Mengxuan Yuan, Kai Xu, Jiawei Huang, Rongli Sun, Juan Zhang, Lihong Yin, and Yuepu Pu. Microcystin-LR exposure enhances toxin-degrading capacity and reduces metabolic diversity of sediment microbial communities. *Environmental Pollution*, 311:119947, 2022.
- [25] Elke Dittmann, Muriel Gugger, Kaarina Sivonen, and David P. Fewer. Natural product biosynthetic diversity and comparative genomics of the cyanobacteria. *Trends in Microbiology*, 23(10):642–652, 2015.
- [26] Angela E. Douglas. The microbial exometabolome: ecological resource and architect of microbial communities. *Philosophical Transactions of the Royal Society B*, 375(1798):20190250, 2020.
- [27] Kai Dührkop, Markus Fleischauer, Marcus Ludwig, Alexander A Aksenov, Alexey V Melnik, Marvin Meusel, Pieter C Dorrestein, Juho Rousu, and Sebastian Böcker. Sirius 4: a rapid tool for turning tandem mass spectra into metabolite structure information. *Nature methods*, 16(4):299–302, 2019.
- [28] Sébastien Duperron, Pierre Foucault, Charlotte Duval, Midoli Goto, Alison Gallet, Simon Colas, and Benjamin Marie. Multi-omics analyses from a single sample: prior metabolite extraction does not alter the 16s rRNA-based characterization of prokaryotic community in a diversity of sample types. *FEMS Microbiology Letters*, 370:fnad125, 2023.
- [29] Oliver Ebenhöf, Thomas Handorf, and Reinhart Heinrich. Structural analysis of expanding metabolic networks. *Genome informatics. International Conference on Genome Informatics*, 15(1):35–45, 2004.

- [30] Pierre Foucault, Sébastien Halary, Charlotte Duval, Midoli Goto, Benjamin Marie, Sahima Hamlaoui, Ludwig Jardillier, Dominique Lamy, Emilie Lance, Emmanuelle Raimbault, Fayçal Allouti, Marc Troussellier, Cécile Bernard, Julie Leloup, and Sébastien Duperron. A summer in the greater Paris: trophic status of peri-urban lakes shapes prokaryotic community structure and functional potential. *Environmental Microbiome*, 20(1):24, February 2025.
- [31] Clément Frainay, Emma L. Schymanski, Steffen Neumann, Benjamin Merlet, Reza M. Salek, Fabien Jourdan, and Oscar Yanes. Mind the Gap: Mapping Mass Spectral Databases in Genome-Scale Metabolic Networks Reveals Poorly Covered Areas. *Metabolites*, 8(3):51, 2018.
- [32] Pereira Francisco, Duthen, Crossay, Alignan, Hennechart, Perez, and Marti. MS-net: Multi-similarity based network annotation for untargeted metabolomics. *Research Square*, 2025.
- [33] Clémence Frioux, Enora Fremy, Camille Trottier, and Anne Siegel. Scalable and exhaustive screening of metabolic functions carried out by microbial consortia. *Bioinformatics*, 34(17):i934–i943, September 2018.
- [34] Juliette Gaëtan, Sébastien Halary, Maxime Millet, Cécile Bernard, Charlotte Duval, Sahima Hamlaoui, Amandine Hecquet, Muriel Gugger, Benjamin Marie, Neha Mehta, David Moreira, Fériel Skouri-Panet, Cynthia Travert, Elodie Duprat, Julie Leloup, and Karim Benzerara. Widespread formation of intracellular calcium carbonates by the bloom-forming cyanobacterium *Microcystis*. *Environmental Microbiology*, 25(3):751–765, March 2023.
- [35] Rebecca Große, Markus Heuser, Jonna E Teikari, Dinesh K Ramakrishnan, Ahmed Abdelfattah, and Elke Dittmann. Microcystin shapes the microcystis phycosphere through community filtering and by influencing cross-feeding interactions. *ISME Communications*, 5(1):ycae170, 2024.
- [36] Sébastien Halary, Sébastien Duperron, Sandra Kim Tiam, Charlotte Duval, Cécile Bernard, and Benjamin Marie. Intra-population genomic diversity of the bloom-forming cyanobacterium, *Aphanizomenon gracile*, at low spatial scale. *ISME Communications*, 3(1):57, 2023.
- [37] Alexa K. Hoke, Guadalupe Reynoso, Morgan R. Smith, Malia I. Gardner, Dominique J. Lockwood, Naomi E. Gilbert, Steven W. Wilhelm, Isabella R. Becker, Grant J. Brennan, Katherine E. Crider, Shannon R. Farnan, Victoria Mendoza, Alison C. Poole, Zachary P. Zimmerman, Lucy K. Utz, Louie L. Wurch, and Morgan M. Steffen. Genomic signatures of lake erie bacteria suggest interaction in the microcystis phycosphere. *PLOS ONE*, 16(9):e0257017, 2021. Publisher: Public Library of Science.
- [38] Michael Hucka, Frank T. Bergmann, Stefan Hoops, Sarah M. Keating, Sven Sahle, James C. Schaff, Lucian P. Smith, and Darren J. Wilkinson. The Systems Biology Markup Language (SBML): Language Specification for Level 3 Version 1 Core. *Journal of Integrative Bioinformatics*, 12(2):266, September 2015.
- [39] Jaime Huerta-Cepas, Damian Szklarczyk, Davide Heller, Ana Hernández-Plaza, Sofia K Forslund, Helen Cook, Daniel R Mende, Ivica Letunic, Thomas Rattei, Lars J Jensen, Christian von Mering, and Peer Bork. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Research*, 47(D1):D309–D314, January 2019.
- [40] Aurore Huré, Maiwenn Le Meur, Charlotte Duval, Jean-Pierre Bouly, Lou Mary, Manon Quiquand, Michella Dawra, Muriel Gugger, Sébastien Halary, and Benjamin Marie. Metabolite diversity of microcystis strains shows tight correspondence to genotype and may contribute to ecotype specificities. *Communications Biology*, 2026.
- [41] Doug Hyatt, Gwo-Liang Chen, Philip F. Locascio, Miriam L. Land, Frank W. Larimer, and Loren J. Hauser. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC bioinformatics*, 11:119, March 2010.
- [42] Bastiaan W. Ibelings, Jutta Fastner, Myriam Bormans, and Petra M. Visser. Cyanobacterial blooms. Ecology, prevention, mitigation and control: Editorial to a CYANOCOST Special Issue. *Aquatic Ecology*, 50(3):327–331, September 2016.

- [43] Chirag Jain, Luis M. Rodriguez-R, Adam M. Phillippy, Konstantinos T. Konstantinidis, and Srinivas Aluru. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nature Communications*, 9(1):5114, 2018.
- [44] Jennifer G. Jankowiak and Christopher J. Gobler. The composition and function of microbiomes within microcystis colonies are significantly different than native bacterial assemblages in two north american lakes. *Frontiers in Microbiology*, 11, 2020. Publisher: Frontiers.
- [45] Martin R. Jones, Ernani Pinto, Mariana A. Torres, Fabiane Dörr, Hanna Mazur-Marzec, Karolina Szubert, Luciana Tartaglione, Carmela Dell’Aversano, Christopher O. Miles, Daniel G. Beach, Pearse McCarron, Kaarina Sivonen, David P. Fewer, Jouni Jokela, and Elisabeth M.-L. Janssen. Cyanometdb, a comprehensive public database of secondary metabolites from cyanobacteria. *Water Research*, 196:117017, 2021.
- [46] Dongwan D. Kang, Feng Li, Edward Kirton, Ashleigh Thomas, Rob Egan, Hong An, and Zhong Wang. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*, 7:e7359, 2019.
- [47] Peter D. Karp, Peter E. Midford, Richard Billington, Anamika Kothari, Markus Krummenacker, Mario Latendresse, Wai Kit Ong, Pallavi Subhraveti, Ron Caspi, Carol Fulcher, Ingrid M. Keseler, and Suzanne M. Paley. Pathway Tools version 23.0 update: software for pathway/genome informatics and systems biology. *Briefings in Bioinformatics*, 22(1):109–126, January 2021.
- [48] Hyun Woo Kim, Mingxun Wang, Christopher A. Leber, Louis-Félix Nothias, Raphael Reher, Kyo Bin Kang, Justin J. J. van der Hooft, Pieter C. Dorrestein, William H. Gerwick, and Garrison W. Cottrell. NPClassifier: A deep neural network-based structural classification tool for natural products. *Journal of Natural Products*, 84(11):2795–2807, 2021.
- [49] Minkyung Kim, Bora Shin, Jaebok Lee, Hye Yoon Park, and Woojun Park. Culture-independent and culture-dependent analyses of the bacterial community in the phycosphere of cyanobloom-forming microcystis aeruginosa. *Scientific Reports*, 9(1):20416, 2019.
- [50] J. B. Kruskal. Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 29(2):115–129, 1964.
- [51] Séverine Le Manach, Charlotte Duval, Arul Marie, Chakib Djediat, Arnaud Catherine, Marc Edery, Cécile Bernard, and Benjamin Marie. Global Metabolomic Characterizations of Microcystis spp. Highlights Clonal Diversity in Natural Bloom-Forming Populations and Expands Metabolite Structural Diversity. *Frontiers in Microbiology*, 10:791, April 2019.
- [52] Ivica Letunic and Peer Bork. Interactive tree of life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Research*, 52:W78–W82, 2024.
- [53] Jine Li, Zhoujie Xie, Min Wang, Guomin Ai, and Yihua Chen. Identification and analysis of the paulomycin biosynthetic gene cluster and titer improvement of the paulomycins in streptomyces paulus NRRL 8115. *PLOS ONE*, 10(3):e0120542, 2015.
- [54] Qi Li, Feibi Lin, Chen Yang, Juanping Wang, Yan Lin, Mengyuan Shen, Min S. Park, Tao Li, and Jindong Zhao. A large-scale comparative metagenomic study reveals the functional interactions in six bloom-forming microcystis-epibiont communities. *Frontiers in Microbiology*, 9:746, 2018.
- [55] Imen Louati, Naoise Nunan, Kevin Tambosco, Cécile Bernard, Jean-François Humbert, and Julie Leloup. The phyto-bacterioplankton couple in a shallow freshwater ecosystem: Who leads the dance? *Harmful Algae*, 126:102436, July 2023.
- [56] Imen Louati, Noémie Pascual, Didier Debros, Cécile Bernard, Jean-François Humbert, and Julie Leloup. Structural Diversity of Bacterial Communities Associated with Bloom-Forming Freshwater Cyanobacteria Differs According to the Cyanobacterial Genus. *PLOS ONE*, 10(11):e0140614, November 2015. Publisher: Public Library of Science.
- [57] Benjamin Marie, Maiwenn Le Meur, Charlotte Duval, Manon Quiquand, Emilie Lance, and Sébastien Duperron. The threat is in the details – critical gap in ecotoxicological assessment of *Microcystis* blooms revealed by critical distinctions of genotype effects induced on medaka fish. *Environmental Pollution*, 387:127344, 2025.

- [58] Nicolas Maurice, Claire Lemaitre, Riccardo Vicedomini, and Clémence Frioux. Mapler: A pipeline for assessing assembly quality in taxonomically rich metagenomes sequenced with HiFi reads. *Bioinformatics*, 41(6):btaf334, 2025.
- [59] Arthur Medina, Melissa Eon, Nicolas Mazzella, Chloé Bonnineau, Débora Millan-Navarro, Aurelie Moreira, Soizic Morin, and Nicolas Creusot. Sensitivity shift of the meta-metabolome and photosynthesis to the chemical stress in periphyton between months along one year and a half period: Case study of a terbuthylazine exposure. *Science of The Total Environment*, 957:177681, 2024.
- [60] Jean-Marie Meyer, Trần Van Van, Alain Stintzi, Odile Berge, and Günther Winkelmann. Ornibactin production and transport properties in strains of burkholderia vietnamiensis and burkholderia cepacia (formerly pseudomonas cepacia). *Biometals*, 8(4):309–317, 1995.
- [61] Marco Mühlenbruch, Hans-Peter Grossart, Falk Eigemann, and Maren Voss. Mini-review: Phytoplankton-derived polysaccharides in the marine environment and their interactions with heterotrophic bacteria. *Environmental microbiology*, 20(8):2671–2685, 2018.
- [62] Coralie Muller, Juliette Audemard, Sylvain Prigent, and Clémence Frioux. Metanetmap: automatic mapping of metabolomic data onto metabolic networks. *bioRxiv*, 2026.
- [63] M Namikoshi and KL Rinehart. Bioactive compounds produced by cyanobacteria. *Journal of Industrial Microbiology*, 17(5):373–384, 1996.
- [64] Cecilia Noecker, Alexander Eng, Sujatha Srinivasan, Casey M. Theriot, Vincent B. Young, Janet K. Jansson, David N. Fredricks, and Elhanan Borenstein. Metabolic Model-Based Integration of Microbiome Taxonomic and Metabolomic Profiles Elucidates Mechanistic Links between Ecological and Metabolic Variation. *mSystems*, 1(1):10.1128/msystems.00013–15, 2016.
- [65] Hans W. Paerl and Jef Huisman. Climate change: a catalyst for global expansion of harmful cyanobacterial blooms. *Environmental Microbiology Reports*, 1(1):27–37, February 2009.
- [66] Donovan H Parks, Maria Chuvochina, Christian Rinke, Aaron J Mussig, Pierre-Alain Chaumeil, and Philip Hugenholtz. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Research*, 50(D1):D785–D794, 2021.
- [67] Donovan H Parks, Maria Chuvochina, Christian Rinke, Aaron J Mussig, Pierre-Alain Chaumeil, and Philip Hugenholtz. 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, 2022.
- [68] Noémie Pascault, Olivier Rué, Valentin Loux, Jacques Pédrón, Véronique Martin, Jennifer Tambosco, Cécile Bernard, Jean-François Humbert, and Julie Leloup. Insights into the cyanosphere: capturing the respective metabolisms of cyanobacteria and chemotrophic bacteria in natural conditions? *Environmental Microbiology Reports*, 13(3):364–374, 2021.
- [69] Ruchi Pathania, Amit Srivastava, Shireesh Srivastava, and Pratyosh Shukla. Metabolic systems biology and multi-omics of cyanobacteria: Perspectives and future directions. *Bioresource Technology*, 343:126007, January 2022.
- [70] Leanne Pearson, Troco Mihali, Michelle Moffitt, Ralf Kellmann, and Brett Neilan. On the Chemistry, Toxicology and Genetics of the Cyanobacterial Toxins, Microcystin, Nodularin, Saxitoxin and Cylindrospermopsin. *Marine Drugs*, 8(5):1650–1680, May 2010.
- [71] Helena L. Pound, Robbie M. Martin, Cody S. Sheik, Morgan M. Steffen, Silvia E. Newell, Gregory J. Dick, R. Michael L. McKay, George S. Bullerjahn, and Steven W. Wilhelm. Environmental Studies of Cyanobacterial Harmful Algal Blooms Should Include Interactions with the Dynamic Microbiome. *Environmental Science & Technology*, 55(19):12776–12779, 2021.
- [72] Olga M. Pérez-Carrascal, Nicolas Tromas, Yves Terrat, Elisa Moreno, Alessandra Giani, Laís Corrêa Braga Marques, Nathalie Fortin, and B. Jesse Shapiro. Single-colony sequencing reveals microbe-by-microbiome phyllosymbiosis between the cyanobacterium *Microcystis* and its associated bacteria. *Microbiome*, 9(1):194, September 2021.

- [73] Sizhe Qiu, Aidong Yang, and Hong Zeng. Flux balance analysis-based metabolic modeling of microbial secondary metabolism: Current status and outlook. *PLOS Computational Biology*, 19(8):e1011391, 2023.
- [74] Rishiram Ramanan, Byung-Hyuk Kim, Dae-Hyun Cho, Hee-Mock Oh, and Hee-Sik Kim. Algae–bacteria interactions: Evolution, ecology and emerging applications. *Biotechnology Advances*, 34(1):14–29, 2016.
- [75] D. F. Robinson and L. R. Foulds. Comparison of phylogenetic trees. *Mathematical Biosciences*, 53(1):131–147, 1981.
- [76] Théotime Roussel, Sébastien Halary, Charlotte Duval, Bérénice Piquet, Jean-Paul Cadoret, Léa Vernès, Cécile Bernard, and Benjamin Marie. Monospecific renaming within the cyanobacterial genus *limnospira* (spirulina) and consequences for food authorization. *Journal of Applied Microbiology*, 134(8):lxad159, 2023.
- [77] Orna Schweitzer-Natan, Maya Ofek-Lalzar, Daniel Sher, and Assaf Sukenik. The microbial community spatially varies during a microcystis bloom event in lake kinneret. *Freshwater Biology*, 68(2):349–363, 2023.
- [78] Patrick M. Shih, Dongying Wu, Amel Latifi, Seth D. Axen, David P. Fewer, Emmanuel Talla, Alexandra Calteau, Fei Cai, Nicole Tandeau de Marsac, Rosmarie Rippka, Michael Herdman, Kaarina Sivonen, Therese Coursin, Thierry Laurent, Lynne Goodwin, Matt Nolan, Karen W. Davenport, Cliff S. Han, Edward M. Rubin, Jonathan A. Eisen, Tanja Woyke, Muriel Gugger, and Cheryl A. Kerfeld. Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proceedings of the National Academy of Sciences*, 110(3):1053–1058, 2013. Publisher: Proceedings of the National Academy of Sciences.
- [79] Hiroshi Tsugawa, Tomas Cajka, Tobias Kind, Yan Ma, Brendan Higgins, Kazutaka Ikeda, Mitsuhiro Kanazawa, Jean VanderGheynst, Oliver Fiehn, and Masanori Arita. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nature Methods*, 12(6):523–526, 2015.
- [80] Hiroshi Tsukiura, Minoru Hanada, Kyoichiro Saito, Kei-Ichi Fujisawa, Takeo Miyaki, Hideo Koshiyama, and Hiroshi Kawagushi. Sorbistin, a new aminoglycoside antibiotic complex of bacterial origin i. production, isolation and properties. *The Journal of Antibiotics*, 29(11):1137–1146, 1976.
- [81] Amé María Valeria, Echenique José Ricardo, Pflugmacher Stephan, and Wunderlin Daniel Alberto. Degradation of microcystin-RR by *sphingomonas* sp. CBA4 isolated from san roque reservoir (córdoba - argentina). *Biodegradation*, 17(5):447–455, 2006.
- [82] Jason N. Woodhouse, Jennifer Ziegler, Hans-Peter Grossart, and Brett A. Neilan. Cyanobacterial community composition and bacteria–bacteria interactions promote the stable occurrence of particle-associated bacteria. *Frontiers in Microbiology*, Volume 9 - 2018, 2018.
- [83] Qiong Wu, Chunyan Yu, Yanru Liu, Peng Xing, Huabing Li, Biao Li, Shiqiang Wan, and Qinglong L. Wu. Microcystis blooms caused the decreasing richness of and interactions between free-living microbial functional genes in lake taihu, china. *FEMS microbiology ecology*, 100(1):fiad166, 2024.
- [84] Ozgur Yurekten, Thomas Payne, Noemi Tejera, Felix Xavier Amaladoss, Callum Martin, Mark Williams, and Claire O’Donovan. MetaboLights: open data repository for metabolomics. *Nucleic Acids Research*, 52(D1):D640–D646, 2024.
- [85] Mitja M Zdouc, Kai Blin, Nico L L Louwen, Jorge Navarro, Catarina Loureiro, Chantal D Bader, Constance B Bailey, Lena Barra, Thomas J Booth, Kenan A J Bozhüyük, José D D Cediell-Becerra, Zachary Charlop-Powers, Marc G Chevrette, Yit Heng Chooi, Paul M D’Agostino, Tristan de Rond, Elena Del Pup, Katherine R Duncan, Wenjia Gu, Novriyandi Hanif, Eric J N Helfrich, Matthew Jenner, Yohei Katsuyama, Aleksandra Korenskaia, Daniel Krug, Vincent Libis, George A Lund, Shrikant Mantri, Kalindi D Morgan, Charlotte Owen, Chin-Soon Phan, Benjamin Philmus, Zachary L Reitz, Serina L Robinson, Kumar Saurabh Singh, Robin Teufel, Yaojun Tong, Fidele Tugizimana, Dana Ulanova, Jaclyn M Winter, César Aguilar, Daniel Y Akiyama,

- Suhad A A Al-Salihi, Mohammad Alanjary, Fabrizio Alberti, Gajender Aleti, Shumukh A Alharthi, Mariela Y Arias Rojo, Amr A Arishi, Hannah E Augustijn, Nicole E Avalon, J Abraham Avelar-Rivas, Kyle K Axt, Hellen B Barbieri, Julio Cesar J Barbosa, Lucas Gabriel Barboza Segato, Susanna E Barrett, Martin Baunach, Christine Beemelmans, Dardan Beqaj, Tim Berger, Jordan Bernaldo-Agüero, Sandra M Bettenbühl, Vincent A Bielinski, Friederike Biermann, Ricardo M Borges, Rainer Borriss, Milena Breitenbach, Kevin M Bretscher, Michael W Brigham, Larissa Buedenbender, Brodie W Bulcock, Carolina Cano-Prieto, João Capela, Victor J Carrion, Riley S Carter, Raquel Castelo-Branco, Gabriel Castro-Falcón, Fernanda O Chagas, Esteban Charria-Girón, Ayesha Ahmed Chaudhri, Vasvi Chaudhry, Hyukjae Choi, Yukyung Choi, Roya Choupannejad, Jakub Chromy, Melinda S Chue Donahey, Jérôme Collemare, Jack A Connolly, Kaitlin E Creamer, Max Crüsemann, Andres Arredondo Cruz, Andres Cumsille, Jean-Felix Dallery, Luis Caleb Damas-Ramos, Tito Damiani, Martinus de Kruijff, Belén Delgado Martín, Gerardo Della Sala, Jelle Dillen, Drew T Doering, Shraavan R Dommaraju, Suhan Durusu, Susan Egbert, Mark Ellerhorst, Baptiste Faussurier, Artem Fetter, Marc Feuermann, David P Fewer, Jonathan Foldi, Andri Frediansyah, Erin A Garza, Athina Gavriilidou, Andrea Gentile, Jennifer Gerke, Hans Gerstmans, Juan Pablo Gomez-Escribano, Luz A González-Salazar, Natalie E Grayson, Claudio Greco, Juan E Gris Gomez, Sebastian Guerra, Shaday Guerrero Flores, Alexey Gurevich, Karina Gutiérrez-García, Lauren Hart, Kristina Haslinger, Beibei He, Teo Hebra, Jethro L Hemmann, Hindra Hindra, Lars Höing, Darren C Holland, Jonathan E Holme, Therese Horch, Pavlo Hrab, Jie Hu, Thanh-Hau Huynh, Ji-Yeon Hwang, Riccardo Iacovelli, Dumitrita Iftime, Marianna Iorio, Sidharth Jayachandran, Eunah Jeong, Jiayi Jing, Jung J Jung, Yuya Kakumu, Edward Kalkreuter, Kyo Bin Kang, Sangwook Kang, Wonyong Kim, Geum Jin Kim, Hyunwoo Kim, Hyun Uk Kim, Martin Klapper, Robert A Koetsier, Cassandra Kollten, Ákos T Kovács, Yelyzaveta Kriukova, Noel Kubach, Aditya M Kujapur, Aleksandra K Kushnareva, Andreja Kust, Jessica Lamber, Martin Larralde, Niels J Larsen, Adrien P Launay, Ngoc-Thao-Hien Le, Sarah Lebeer, Byung Tae Lee, Kyungha Lee, Katherine L Lev, Shu-Ming Li, Yong-Xin Li, Cuauhtémoc Licona-Cassani, Annette Lien, Jing Liu, Julius Adam V Lopez, Nataliia V Machushynets, Marla I Macias, Taifo Mahmud, Matiss Maleckis, Añadir Maharai Martinez-Martinez, Yvonne Mast, Marina F Maximo, Christina M McBride, Rose M McLellan, Khyati Mehta Bhatt, Chrats Melkonian, Aske Merrild, Mikko Metsä-Ketelä, Douglas A Mitchell, Alison V Müller, Giang-Son Nguyen, Hera T Nguyen, Timo H J Niedermeyer, Julia H O'Hare, Adam Ossowicki, Bohdan O Ostash, Hiroshi Otani, Leo Padva, Sunaina Paliyal, Xinya Pan, Mohit Panghal, Dana S Parade, Jiyeon Park, Jonathan Parra, Marcos Pedraza Rubio, Huong T Pham, Sacha J Pidot, Jörn Piel, Bitu Pourmohsenin, Malik Rakhmanov, Sangeetha Ramesh, Michelle H Rasmussen, Adriana Rego, Raphael Reher, Andrew J Rice, Augustin Rigolet, Adriana Romero-Otero, Luis Rodrigo Rosas-Becerra, Pablo Y Rosiles, Adriano Rutz, Byeol Ryu, Libby-Ann Sahadeo, Murrel Saldanha, Luca Salvi, Eduardo Sánchez-Carvajal, Christian Santos-Medellin, Nicola Sbaraini, Sydney M Schoellhorn, Clemens Schumm, Ludek Sehnal, Nelly Selem, Anjali D Shah, Tania K Shishido, Simon Sieber, Velina Silviani, Garima Singh, Hemant Singh, Nika Sokolova, Eva C Sonnenschein, Margherita Sosio, Sven T Sowa, Karin Steffen, Evi Stegmann, Alena B Streiff, Alena Strüder, Frank Surup, Tiziana Svenningsen, Douglas Sweeney, Judit Szenei, Azat Tagirdzhanov, Bin Tan, Matthew J Tarnowski, Barbara R Terlouw, Thomas Rey, Nicola U Thome, Laura Rosina Torres Ortega, Thomas Tørring, Marla Trindade, Andrew W Truman, Marie Tvilum, Daniel W Udway, Christoph Ulbricht, Lisa Vader, Gilles P van Wezel, Max Walmsley, Randika Warnasinghe, Heiner G Weddelling, Angus N M Weir, Katherine Williams, Sam E Williams, Thomas E Witte, Steffaney M Wood Rocca, Keith Yamada, Dong Yang, Dongsoo Yang, Jingwei Yu, Zhenyi Zhou, Nadine Ziemert, Lukas Zimmer, Alina Zimmermann, Christian Zimmermann, Justin J J van der Hooff, Roger G Lington, Tilmann Weber, and Marnix H Medema. MIBiG 4.0: advancing biosynthetic gene cluster curation through global collaboration. *Nucleic Acids Research*, page gkae1115, 2024.
- [86] Peng Zhou, Li Wang, Hai Liu, Chunyan Li, Zhimin Li, Jinxiang Wang, and Xiaoming Tan. CyanoOmicsDB: an integrated omics database for functional genomic analysis of cyanobacteria. *Nucleic Acids Research*, 50(D1):D758–D764, January 2022.
- [87] Zhichao Zhou, Patricia Q. Tran, Adam M. Breister, Yang Liu, Kristopher Kieft, Elise S. Cowley, Ulas Karaoz, and Karthik Anantharaman. METABOLIC: high-throughput profiling of microbial genomes for functional traits, metabolism, biogeochemistry, and community-scale functional networks. *Microbiome*, 10(1):33, 2022.
- [88] Lin Zhu, Anouk Zancarini, Imen Louati, Silvia De Cesare, Charlotte Duval, Kevin Tambosco, Cécile Bernard, Didier Debroas, Lirong Song, Julie Leloup, and Jean-François Humbert. Bacterial

communities associated with four cyanobacterial genera display structural and functional differences: Evidence from an experimental approach. *Frontiers in Microbiology*, Volume 7 - 2016, 2016.