



OPEN ACCESS

EDITED BY

Anne Postec,
UMR7294 Institut Méditerranéen
d'océanographie (MIO), France

REVIEWED BY

Lin Xu,
Zhejiang Sci-Tech University, China
Esra Ersoy Omeroglu,
Ege University, Türkiye

*CORRESPONDENCE

Marli F. Fiore
✉ fiore@cena.usp.br

RECEIVED 23 February 2025

ACCEPTED 05 June 2025

PUBLISHED 04 July 2025

CITATION

Machado MJ, Jacinavicius FR, Médice RV,
Dextro RB, Feitosa AMT, Weiss MB,
Pellegrinetti TA, Cotta SR, Crnkovic CM and
Fiore MF (2025) Genetic and biochemical
diversity of terpene biosynthesis
in cyanobacterial strains from tropical soda
lakes.
Front. Microbiol. 16:1582103.
doi: 10.3389/fmicb.2025.1582103

COPYRIGHT

© 2025 Machado, Jacinavicius, Médice,
Dextro, Feitosa, Weiss, Pellegrinetti, Cotta,
Crnkovic and Fiore. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Genetic and biochemical diversity of terpene biosynthesis in cyanobacterial strains from tropical soda lakes

Mauricio J. Machado¹, Fernanda R. Jacinavicius¹,
Rhuana V. Médice², Rafael B. Dextro¹, Anderson M. T. Feitosa¹,
Marcio B. Weiss², Thierry A. Pellegrinetti^{1,3}, Simone R. Cotta⁴,
Camila M. Crnkovic² and Marli F. Fiore^{1*}

¹Center for Nuclear Energy in Agriculture, University of São Paulo, Piracicaba, Brazil, ²School of Pharmaceutical Sciences, University of São Paulo, São Paulo, Brazil, ³Faculty of Agricultural Sciences and Food, Université Laval, Québec City, QB, Canada, ⁴“Luiz de Queiroz” College of Agriculture, University of São Paulo, Piracicaba, Brazil

Introduction: Terpenes and terpenoids are vital components in diverse metabolic pathways, forming the terpenome—the complete spectrum of terpene-related compounds biosynthesized by an organism. Integrating bioinformatic tools has significantly enhanced the ability to assess metabolic potential by combining these computational approaches with experimental biochemical data. Furthermore, gene annotation provides critical insights into specialized targets, facilitating the identification of shared or unique features across different strains.

Aims and methods: This study investigates the presence of terpene compounds in cyanobacterial strains isolated from tropical soda lakes using a combination of gene mining, synteny analysis, phylogenetics, and metabolomics.

Results and discussion: Key enzymes, including phytoene synthase and squalene hopene cyclase, were identified, showing significant similarities and evolutionary links to gene copies in Cyanobacteria from diverse ecological environments. Metabolomic analysis complemented genomic predictions, uncovering a rich diversity of tetraterpene compounds, particularly carotenoids. Notably, triterpene hopanoids were found exclusively in a unicellular strain. These compounds show significant potential for cellular protection, metabolic adaptation, and biotechnological uses. They might support microbial communities in extreme environments, such as the saline-alkaline lakes of the Pantanal Biome in Brazil, developing unique survival and resilience strategies in these harsh conditions.

Conclusion: This study highlights the extensive range of insights that can be obtained by integrating genetics and biochemistry in exploring cyanobacterial diversity, especially from organisms thriving in extreme environments.

KEYWORDS

biosynthetic gene cluster, carotenoids, environmental adaptation, genomic, isoprene, metabolomic

1 Introduction

The tropical soda lakes of Nhocolândia, Mato Grosso do Sul (MS), Brazil, are unique ecosystems of remarkable biological and ecological importance. These lakes in the Brazilian Pantanal biome are characterized by saline-alkaline conditions with a pH gradient varying between 8.62 and 10.26, salinity from 0.41 to 2.42 g L⁻¹, and low water column level (< 2 m) (Pellegrinetti et al., 2023). The unique combination of high pH, salinity, and intense UV radiation in these lakes creates an environment that supports a distinctive microbial community, with cyanobacteria playing a crucial role in sustaining its diversity (Cotta et al., 2022). Specifically, the genera *Limnospira* and *Anabaenopsis* dominate in these lakes, creating blooms that enhance the environment's unique characteristics and support a diversity of heterotrophic organisms that depend on or are modulated by their presence (Andreote et al., 2018; Pellegrinetti et al., 2024). The adaptability of cyanobacteria to these adverse environmental conditions is truly remarkable (Waditee-Sirisattha and Kageyama, 2022; Jacinavicius et al., 2021). The ability of cyanobacteria to adapt and thrive in various ecological niches reflects their metabolic versatility. This phylum possesses a vast diversity of specialized metabolites, such as alkaloids, amino acid derivatives, fatty acid derivatives, polyketides, peptides, and terpenes (Chlipala et al., 2011; Jones et al., 2021; Weiss et al., 2025).

Terpenoid compounds are a diverse group of natural products with more than 80,000 representatives in plants, fungi, marine invertebrates, and bacteria (Christianson, 2017; Rudolf et al., 2021). The terpenome embraces all compounds that contain isoprenoids produced from building blocks of dimethylallyl-pyrophosphate (DMAPP) and isopentenyl-pyrophosphate (IPP) that can be further converted into different molecules by terpene synthases (TPS). Terpene synthases are responsible for the structural diversity found in the terpenoid natural products. Terpenoids are classified based on the number of carbon atoms (C) in their core structure: hemiterpenes (C₅), monoterpenes (C₁₀), sesquiterpenes (C₁₅), diterpenes (C₂₀), triterpenes (C₃₀), and tetraterpenes (C₄₀). Additionally, some meroterpenes have been identified, which are hybrid molecules resulting from mixed biosynthesis (Nazir et al., 2021; Rudolf et al., 2021). Natural products such as steroids, vitamins, plant hormones, and drugs highly utilized in medicine (e.g., taxol and artemisinin) represent terpene compounds (Tholl, 2006; Avalos et al., 2022). Terpenoids are essential in the cellular metabolism of photosynthetic organisms, acting in light conversion, membrane fluidity, and assemblage of photosynthetic reaction centers. In oxygenic phototrophic bacteria, such as cyanobacteria, isoprenoids are synthesized from the methylerythritol-phosphate (MEP) pathway, which relies on glyceraldehyde-3-phosphate (G3P) and pyruvate, both derived from photosynthesis (Pattanaik and Lindberg, 2015).

The carbon metabolism of cyanobacteria, which can convert CO₂ into a wide range of terpenoid compounds, makes these microorganisms an emerging target host for biochemical production, and thus a promising avenue for future research. However, challenges need to be overcome, like carbon partitioning and balancing pigment production (Lin and Pakrasi, 2019). Specifically looking at terpenoids, cyanobacteria have a large diversity of these compounds beyond only pigments, with a considerable variation of functions and structures, which have

been considered important sources of molecules to produce medicines, biofuels, and other applications (Nandagopal et al., 2021). Among the sesquiterpenes that can be engineered in cyanobacteria, farnesene is a compound that has many applications in different areas like biofuels, pest management, cosmetics, flavors, and fragrances (Rautela et al., 2024). The diterpenoid compounds tolypodiol and noscomin, isolated from *Tolypothrix nodosa* and *Nostoc commune*, respectively, exhibit significant anti-inflammatory and antimicrobial activity (Prinsep et al., 1996; Jaki et al., 1999). Additionally, the sesterterpenes cybastacines A and B, derived from *Nostoc* sp., and scytonarol from *Scytonema* sp., also possess noteworthy antimicrobial effects (Mo et al., 2009; Cabanillas et al., 2018). Also, cyanobacterial tetraterpenes can act as an antioxidant agent, principally reducing reactive oxygen species (ROS) and in therapeutic solutions for inflammation-related skin disorders (Hatha and Sumayya, 2023; Morone et al., 2024).

Abiotic stress is known to alter products of the isoprenoid pathway; however, its specific effects on cyanobacterial metabolism remain underexplored. In response to the extreme conditions of soda lakes, such as high pH, salinity, and UV exposure, Cyanobacteria, which dominate in these environments, exhibit specific metabolic pathways that help them adapt and thrive. The increasing availability of genomic data and advancements in metabolomics have significantly boosted the discovery of specialized metabolites, providing new sources for identifying and characterizing cyanobacterial compounds (Winter et al., 2011; Micallef et al., 2015; Weiss et al., 2023). However, further research into cyanobacterial metabolism is essential for a comprehensive understanding of these metabolic adaptations and bioprospecting natural products from organisms isolated from extreme environments.

This study aimed to characterize terpene and terpenoid gene clusters in the genomes of cyanobacterial strains isolated from tropical soda lakes in the Brazilian Pantanal, considered extreme environments. Additionally, it evaluated the biosynthesis of terpene compounds and their potential biological roles, providing insights into the metabolic adaptations of microorganisms to harsh conditions and underscoring the bioprospecting potential of cyanobacterial metabolites.

2 Materials and methods

2.1 Cyanobacteria culture conditions and metabolites extraction

The cyanobacterial strains, *Anabaenopsis elenkinii* CCIBt3563, *Pantanalinema rosaneae* CENA516, *Geminocystis* sp. CENA526, *Alkalinema pantanalense* CENA528, *Limnospira platensis* CENA597, and *Limnospira platensis* CENA650, isolated from soda lakes in the Brazilian Pantanal biome, were used in this study and maintained as uni-cyanobacterial cultures (Supplementary Table 1). These strains are part of the culture collections of the Cell and Molecular Biology Laboratory (CENA/USP) in Piracicaba, São Paulo, Brazil, and the Cyanobacterial Culture Collection (CCIBt) at the Institute of Botany in São Paulo, São Paulo, Brazil. The cyanobacterial strains were cultured in Z8 medium (Kotai, 1972) or a modified Z8 medium containing 7.5 g L⁻¹

of NaCl and adjusted to pH 9.5. Cultures were maintained at $22 \pm 1^\circ\text{C}$ under a 14/10 h light/dark photoperiod with fluorescent illumination ($40 \mu\text{Mol photon m}^{-2}\cdot\text{s}^{-1}$). After 30 days of growth, the culture biomass was lyophilized and stored at -20°C for subsequent extraction.

Freeze-dried cyanobacterial biomass was extracted using three different solvent systems—methanol/water (1:1, v/v), ethyl acetate, and dichloromethane/methanol (1:1, v/v)—applied independently to separate aliquots of the same biomass at a 10:1 (v/w) solvent-to-biomass ratio. Extracts were sonicated (Ultrasonic cleaner, Unique) for 1 min and kept overnight at room temperature for complete extraction. Afterward, extracts were filtered through qualitative paper filters ($80 \mu\text{g}\cdot\text{m}^{-2}$), and all filtrates were evaporated overnight under nitrogen flow. Dried extracts were resuspended in methanol ($10 \text{ mg}\cdot\text{mL}^{-1}$) and subjected to a clean-up step using SPE Columns C18 (Applied Separations).

2.2 DNA extraction and genome sequencing

After 30 days of cultivation under the described conditions, cultures of *P. rosanae* CENA516; *Geminocystis* sp. CENA526; *A. pantanalense* CENA528; *L. platensis* CENA597 and *L. platensis* CENA650 were concentrated by centrifugation at $5,000 \times g$ for 10 min. The concentrated biomass was used for total DNA extraction using the AxyPrep Bacterial Genomic DNA Miniprep Kit (Axygen Biosciences, Union City, CA, United States) according to the manufacturer's requirements. The quality of the DNA was confirmed using 1% (m/w) of agarose gel. DNA quantification was performed with a Qubit 2.0 Fluorometer, using a Qubit dsDNA BR Assay Kit (Life Technologies, Carlsbad, CA, United States). An amount of 1 μg of DNA was used to prepare paired-ends libraries with the Kit Nextera DNA Flex (Illumina, San Diego, CA, United States), which was sequenced in a platform HiSeq 2500 (Illumina) following the manufacturer's instructions. The gDNA was then sent to the Joint Genome Institute (JGI) (Project ID: 504264), where a PacBio SMRTbell library was prepared for circular consensus sequencing (CCS) and sequenced using the PacBio HiFi platform.

2.3 Genome assembly and bioinformatics analysis

The quality of the reads obtained from genomic sequencing on the HiSeq and HiFi platforms was evaluated using FastQC v0.11.8 (FastQC), with analysis graphs generated to assess the quality metrics. Adapter sequences were subsequently removed using NxTrim v0.4.3 (O'Connell et al., 2015). Further filtering of reads was performed with Cutadapt v1.18 (Martin, 2011), which excluded sequences with phred quality scores below 30, lengths shorter than 30 bp, and repetitive elements. The resulting high-quality short reads were individually assembled using SPAdes v3.13.0 (Bankevich et al., 2012) with the metaSPAdes module for metagenomic assembly (Nurk et al., 2017), following the methodology described by Alvarenga et al., 2017. Long-read sequences were assembled using Flye v2.9.1 (Kolmogorov et al., 2019), and hybrid assemblies, which combined both types of

data, were generated using SPAdes in hybrid mode. Taxonomic classification of assemblies was performed using Kraken2 (Wood et al., 2019), while contaminant sequences were identified and removed using KrakenTools (Lu et al., 2022). Pilon (Walker et al., 2014) was employed for genome polishing, correcting assembly errors, and, along with the programs SSPACE (Boetzer et al., 2011) and GAPPadder (Chu et al., 2019), was used for closing sequence gaps, thereby enhancing the completeness and accuracy of the final genomes. The assembled genomes were assessed for quality and completeness using QUAST v5.0.2 (Gurevich et al., 2013). Genome integrity and contamination were estimated using CheckM v1.0.13 (Parks et al., 2015). Additionally, BUSCO v3.0 (Simão et al., 2015) was utilized to evaluate genome quality by identifying conserved single-copy orthologs. The genomic data for *A. elenkinii* can be accessed through the NCBI GenBank database (National Center for Biotechnology Information) under the accession number CP063311. The assembled genomes studied are also deposited in NCBI (BioProject number: PRJNA1226917).

Genome annotation was performed with a rapid prokaryotic sequence annotation algorithm implemented in Prokka v1.12 (Seemann, 2014). The presence of genes associated with the methylerythritol phosphate (MEP) pathway, plus other terpenoid biosynthetic pathways, such as triterpenes, tetraterpenes, and meroterpenoids, were identified through manual curation and using BlastKOALA tool (Kanehisa et al., 2016). Biosynthetic Gene Clusters (BCGs) related to terpene production were predicted by submitting the genome assemblies to AntiSMASH bacterial version 6.0 (Blin et al., 2021) using default parameters. For insights into their evolution and potential differences in the pathways they encode, the predicted BCGs were analyzed with Clinker to visualize and compare gene cluster architectures across the strains using default settings (Gilchrist and Chooi, 2021). Additionally, orthologous gene clusters among the genomes were compared using OrthoVenn2, allowing for the identification of shared and unique genes across the studied strains (Xu et al., 2019).

2.4 Gene phylogeny analysis

Amino acid sequences of the genes encoding phytoene synthase (*crtB*) and terpene synthase (*sqhC*) annotated on the Pantanal strains' genomes were compared to the National Center for Biotechnology Information (NCBI) database via BLAST. Homologs were selected based on coverage ($> 95\%$) and identity scores ($> 70\%$, $e\text{-value} > 1.10 \times 10^{-50}$). Due to the high variability of terpene synthases, genes associated with the biosynthesis of geosmin and germacrene were included in the analysis to enhance resolution and more effectively explore evolutionary relationships. A maximum-likelihood phylogenetic analysis was constructed for both enzymes in MEGA11 (Tamura et al., 2021) using the substitution model Jones-Taylor-Thornton (JTT) based on matrixes of differential substitution rate for each amino acid from the protein sequence. For robustness, 1,000 bootstrap replication steps were used.

2.5 HPLC-MS/MS analysis

The HPLC-MS/MS analyses were carried out on a high-performance liquid chromatography system (Shimadzu®

Prominence Liquid Chromatography) coupled with a high-resolution tandem mass spectrometer (Micro TOF-QII; Bruker Daltonics®, MA, United States) with electrospray ionization source (ESI)—(HPLC-ESI-QTOF-MS/MS). Chromatographic separations were performed over a Kinetex C18 column (2.1 × 50 mm × 1.7 μM, Phenomenex) equipped with pre-column. Sample injection volumes were 2 μL and the mobile phase consisted of 0.1% v/v formic acid in ultrapure water (solvent A) and 0.1% v/v formic acid in acetonitrile (solvent B). The gradient was as follows: 10–100% B (7 min), 100% B (in 1 min), 100–10% B (0.1 min), 10% B (in 1.9 min). Flow rate was 0.5 mL/min. The ESI conditions were capillary potential at 3.5 kV, drying gas (N₂) at 200°C at a flow rate of 9.0 mL/min, and nebulization pressure at 43.0 psi. The mass spectrometer (QTOF) was operated in auto-scan MS/MS mode, and mass spectra were acquired in positive mode, with collision-induced dissociation (CID) energy at 60 eV (isolation mass of 100 *m/z* with width of 3 *m/z*, and isolation of mass 500 *m/z* with width of 6 *m/z*) or 70 eV (isolation mass of 1,000 *m/z* with width of 9 *m/z*, and isolation of mass 1,500 *m/z* with width of 12 *m/z*) and averaged stepping of 65–100% CID. The mass spectra were acquired in the mass range of 150–2,000 Da (MS1) and 50–2,000 Da (MS2).

2.6 Terpenoid profile investigation

LC-MS/MS data were converted to “.mzXML” format in MSConvert® software from ProteoWizard tools. Data was filtered by removing MS/MS fragment ions within ± 17 Da of the precursor *m/z*. Window filtering selected the top 6 fragment ions in a ± 50 Da window. The converted data were processed using the Classical Molecular Networking pipeline of the GNPS platform¹ (Wang et al., 2016). The precursor ion mass tolerance was 0.02 Da, and the MS/MS fragment ion tolerance was 0.02 Da. The minimum cosine score was 0.7, and at least 6 matched peaks. Network spectra were compared against GNPS’ spectral libraries, with matches requiring a score above 0.7 and at least 6 matched peaks. Additional filters included precursor window settings and exclusion of spectra identified as blanks before networking. Cytoscape® software was used to visualize molecular networks. *In silico* structure annotations from the GNPS library were incorporated to enhance the chemical structural information within the molecular network. *In silico* structure annotations were incorporated using GNPS tools such as Dereplicator, Dereplicator+, and the *In silico* Peptidic Natural Product Dereplicator (Gurevich et al., 2018; Mohimani et al., 2018), enhancing identification of peptidic and non-peptidic natural products. The search results were integrated into the GNPS MolNetEnhancer workflow, enabling annotation of a wide range of chemical classes, including lipids and lipid-like molecules, alkaloids, phenolic compounds, terpenoids (monoterpenes, sesquiterpenes, and carotenoids), and peptides (Ernst et al., 2019). Chemical class annotations were performed using the ClassyFire chemical ontology. For putative metabolite identification, HRMS data were also processed using DataAnalysis 4.4 and MetaboScape 4.0 software (Bruker Daltonics, Germany). Accurate masses were used to search against CyanoMetDB

(Jones et al., 2021), PubChem, ChemSpider, METLIN (Smith et al., 2005), NPAtlas (Van Santen et al., 2019), the Dictionary of Natural Products, Bruker’s MetaboBASE Plant Library, MetaboBASE Personal Library 3.0, and in-house databases. Additional database searching was performed using GNPS Theoretical/*In silico* tools, SIRIUS 4, and CANOPUS (Dührkop et al., 2019, 2021). Elementary compositions and deviations from theoretical values (ppm error) were calculated using the SmartFormula algorithm, adopting a 5 ppm threshold. Natural isotopic patterns were considered to refine formula predictions and to aid in matching experimental MS/MS spectra with published data. Furthermore, the ChemCalc web service² (Patinay and Borel, 2013) was used for additional molecular formula predictions and mass error calculation, and the Carotenoids Database³ (Yabuzaki, 2017) was used for prospecting carotenoids.

3 Results

3.1 Genomics analysis

Anabaenopsis elenkinii CCIBt3563, *Pantanalinema rosaneae* CENA516, *Geminocystis* sp. CENA526, *Alkalinema pantanalense* CENA528, *Limnospira platensis* CENA597, and *Limnospira platensis* CENA650 had a complete set of MEP pathway genes responsible for IPP and DMAPP biosynthesis (Supplementary Figure 1). Among these coding regions, genes associated with the MEP pathway of terpene biosynthesis (*dxr*, *dxs*, *ispD*, *ispE*, *ispF*, *ispG*, *ispH*, and *idi*) were orthologous to all evaluated strains from the Pantanal saline-alkaline lakes. The targeted genes are scattered in the genomes rather than organized in clusters. Gene size and identity were conserved among the different sequences from the six Pantanal cyanobacteria that were evaluated. Orthologous gene analysis of the complete genomes highlighted that amongst 9,537 proteins predicted in the six genomes evaluated, 1,507 proteins formed common clusters (Supplementary Figure 2); among them, several genes related to terpene biosynthesis were identified, mainly related to the carotenoid pigments.

Analysis using AntiSMASH predicted two secondary metabolite biosynthetic gene clusters (BGCs) with significant identity (> 50%) for terpenes among the Pantanal cyanobacterial genomes. The first was for scalene-hopene cyclase, and the second was for phytoene synthase. The scalene hopene cyclase (SCH) enzyme, encoded by the gene *sqhC*, is linked with the cyclization catalysis of hopanoids from the acyclic precursor squalene. This gene is associated with the precursor of triterpenoid compounds (hopene and hopanol), which BlastKOALA predicted in the strains *P. rosaneae* CENA516, *Geminocystis* sp. CENA526, and *A. pantanalense* CENA528 (Supplementary Table 2). The SCH enzyme clusters predicted by AntiSMASH were compared to data available in the platform’s database. Synteny analysis for the scalene hopene cyclase gene found several homologs from different genera and two other strains of *Geminocystis* with significant sequence similarity compared to the Pantanal strains’ genomes

¹ <http://gnps.ucsd.edu>

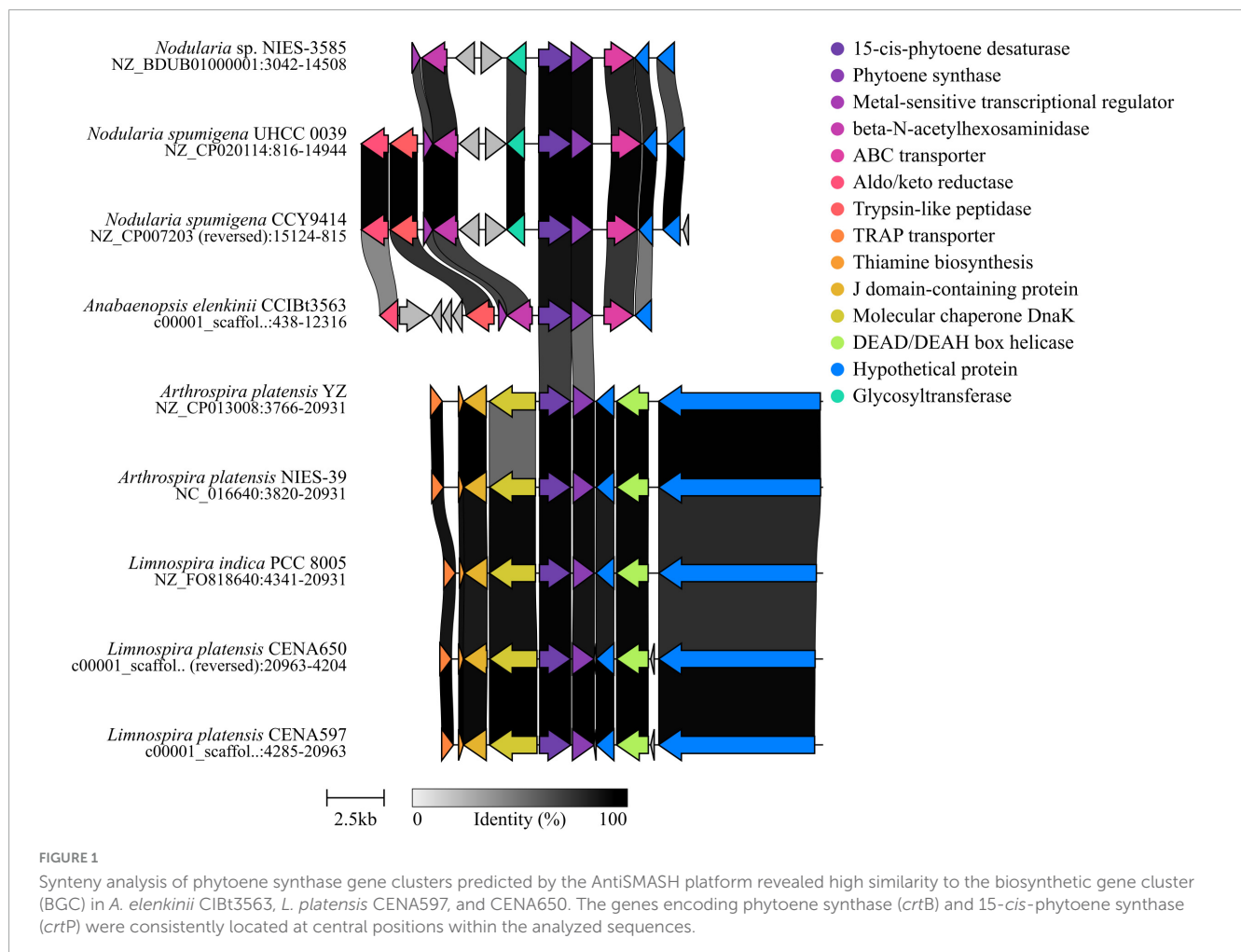
² <http://www.chemcalc.org>

³ <http://carotenoiddb.jp/index.html>

(Supplementary Figure 3). Similarly, genes related to phytoene synthase (*CrtB*), the first enzyme on the carotenoid biosynthetic pathway, were also predicted by AntiSMASH. The studied genomes showed the presence of genes related to phytoene synthase (*crtB*) and 15-sis-phytoene-desaturase (*crtP*). The *crtB* gene clusters found in *A. elenkinii* CCIBt3563 and both strains of *L. platensis* (CENA597 and CENA650) showed the highest similarity to strains of *Nodularia* and the *Arthrospira* genus found in the database that presented this same gene concatenated in the genome (Figure 1). Notably, the strains of *Nodularia* sp. NIES-3585 and *Nodularia spumigena* (UHCC 0039 and CCY 9414), exhibiting significant similarity, are sourced from marine environments, while the strains related to *L. platensis* are from freshwater habitats. The similarity of the *crtB* gene in phylogenetically related strains from distinct aquatic environments suggests that the carotenoid biosynthesis is conserved among the same cyanobacteria species inhabiting different ecological niches.

As for the phylogenetic analysis, the *crtB* gene tree (Supplementary Figure 4) showed that the gene copies found in the strains from saline lakes in Pantanal may exhibit different evolutionary trajectories, presenting common ancestry among the close phylogenetic members. The genes identified in the *L. platensis* strains (CENA597 and CENA650) grouped with homologs retrieved from other *Arthrospira* and *Limnospira* strains, forming a basal group on the tree, with an early ancestry to

the remaining copies of phytoene synthase used in the analysis. Meanwhile, *crtB* annotated in *A. elenkinii* CCIBt3563 clusters with other *Anabaenopsis* strains, sharing a common ancestry with *Nodularia* strains and *crtB* genes from various distinct genera. While the *A. pantanalense* CENA528 gene appears to be ancestral to two others freshwater *Alkalinema* strains, the close branch formed by the MAG of *Alkalinema* sp. CACIAM 70d represents an environmental sample isolated from the Amazon biome in Brazil. Furthermore, the strains *P. rosanae* CENA516 and *Geminocystis* sp. CENA526 clustered together, sharing a common ancestor with the strain *Pantanalinema* sp. GBBB05, which was isolated from a freshwater environment in the Cerrado biome, also in Brazil. As expected, the *sqhC* gene phylogenetic tree (Supplementary Figure 5) displays an external group formed by terpene synthases of *Nostoc* strains associated with the sesquiterpenes compounds geosmin, selinene, and germacrene. The addition of other terpene synthases aimed to capture a broader range of biosynthetic diversity and provide deeper insights into terpene production's evolutionary patterns. This allowed the grouping of all squalene/hopene cyclases with a common ancestral state. *Geminocystis* sp. CENA526 is grouped with other *Geminocystis* sp. and a *Cyanobacterium aponinum* strain, both of which are from freshwater, forming the sister branch to all remaining copies of the *sqhC* gene. Both *P. rosanae* CENA516 and *A. pantanalense* CENA528 shared ancestry with *Synechococcus* and *Thermosynechococcus* strains. The



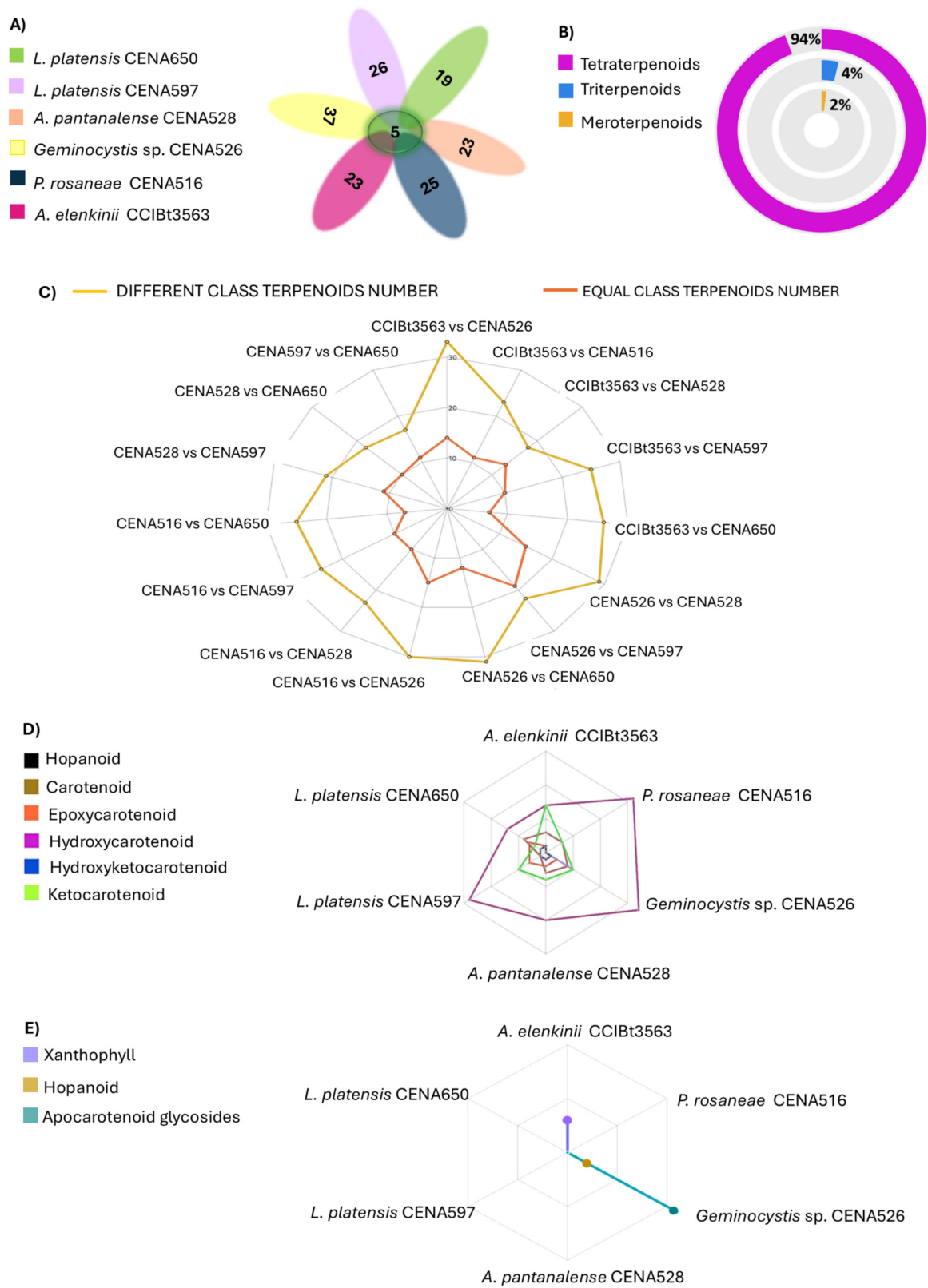


FIGURE 3
(A) Venn Diagram representing the number of nodes annotated as a terpenoid in each strain for terpenoid production and common to all samples. (B) Proportion of terpenoids found in cyanobacteria from tropical soda lakes. (C) Spider plot depicting the distribution of compound classes across different strains. (D) Spider plot showing the distribution of different groups of terpene compounds for each strain. (E) Spider plot presenting the exclusive presence of groups of terpene compounds in strains isolated from Pantanal.

compounds and genes for hopanoids, 35-O- β -6-amino-6-deoxyglucopyranosyl (m/z 708.5330 $[M + H]^+$, $C_{41}H_{73}NO_8$, $\Delta = -0.87$ ppm) and bacteriohopanetetrol (m/z 722.5465 $[M + H]^+$, $C_{42}H_{75}NO_8$, $\Delta = -3.83$ ppm), which are exclusive to this strain (Figure 3 and Supplementary Tables 2, 3).

The comparative analysis of terpenoid profiles among the Pantanal cyanobacterial strains revealed a significant insight, such as the observation of a greater dissimilarity of compounds rather than a homogenous profile amongst the studied cyanobacteria (Figure 3C). Strains CENA597 and CENA650 (*L. platensis*), belonging to the same species, showed a similarity of 39.29%. The highest similarity, at 46.34%, was between *Geminocystis* sp. CENA526 and *L. platensis* CENA597, despite these strains belonging to different families and having different cell morphologies (coccoid unicellular and filamentous, respectively). The lowest similarity was observed between *A. elenkinii* CCIBt3563 and *Geminocystis* sp. CENA526. Regarding chemical diversity and the predominance of specific terpenoids, it is noteworthy that hydroxycarotenoids were more prevalent in the strains *P. rosaneae* CENA516, *Geminocystis* sp. CENA526, *A. pantanalense* CENA528, and *L. platensis* CENA597 (Figure 3D). On the other hand, strain *A. elenkinii* CCIBt3563 showed more diversity across the categories (except hopanoids) and some exclusive xanthophyll compounds (Figure 3E and Supplementary Table 3). The terpenoid profiles varied among the studied strains. *P. rosaneae* CENA516, *Geminocystis* sp. CENA526 and *L. platensis* CENA597 stood out for their high concentration of hydroxycarotenoids, while *A. elenkinii* CCIBt3563 exhibited a relatively high diversity of ketocarotenoids (Figure 3C and Supplementary Table 2). *Geminocystis* sp. CENA526 exhibited a higher abundance and diverse terpenoid categories, including the exclusive presence of apocarotenoid glycosides and hopanoids (Figure 3E and Supplementary Table 2).

The analysis of carotenoids in cyanobacteria revealed that these compounds are predominantly represented by C40 cyclic species, which can be categorized into two major classes: carotenes (e.g., β -carotene) and xanthophylls (oxygenated derivatives of carotenes such as zeaxanthin and echinenone) (Supplementary Table 2). Based on the spectral similarity between samples and the GNPS database, the feature with m/z 565.4033 $[M + H]^+$ was annotated as the terpene canthaxanthin with a calculated error of $\Delta = -2.29$ ppm. This annotation was further supported by manual analysis of the isotopic and fragmentation patterns, with characteristic fragments (m/z 203.1, m/z 105.06, and m/z 119.08) of canthaxanthin. The SIRIUS platform also proposed the molecular formula $C_{40}H_{52}O_2$ and classified the feature as a terpene. Similarly, the feature detected at m/z 551.4186 $[M + H]^+$ was identified as echinenone ($C_{40}H_{54}O$, $\Delta = 2.05$ ppm) based on MS data, which showed characteristic fragments (m/z 203.1, m/z 93.0, and m/z 69.0). The feature acquired with m/z 585.4231 $[M + H]^+$ was suggested to be antheraxanthin ($C_{40}H_{56}O_3$, $\Delta = 0.26$ ppm), with fragments (m/z 95 and m/z 105) consistent with this carotenoid. These features were used as reference points for the annotation propagation of the other compounds in the network. Three compounds were consistently annotated among the analyzed strains: β -carotene, echinenone, and zeaxanthin (Figure 3 and Supplementary Table 4).

The biosynthesis of carotenoids begins with phytoene, catalyzed by the enzyme phytoene synthase (Supplementary Table 2). Phytoene synthase gene (*crtB*), predicted

in the genomics analysis, performs the condensation of two GGPP molecules, which are then converted into lycopene through desaturation and isomerization steps catalyzed by phytoene desaturase, zeta-carotene desaturase, and cis-carotene isomerase (Supplementary Table 2). Lycopene serves as a precursor for β -carotene (m/z 536.4392 $[M]^+$, $C_{40}H_{56}[M]^+$, $\Delta = 1.86$ ppm) converted by the lycopene cyclase (CruA/P) (Supplementary Table 2). β -carotene is the starting point for the biosynthesis of other carotenoids like zeaxanthin (m/z 568.4246 $[M]^+$, $C_{40}H_{56}O_2$, $\Delta = -6.03$ ppm), mediated by β -carotene hydroxylase (*crtR*) (Supplementary Table 2).

In *A. elenkinii* CCIBt3563, zeaxanthin is then converted into nostoxanthin (m/z 600.4154 $[M]^+$, $C_{40}H_{56}O_4$, $\Delta = -4.09$ ppm). In these Pantanal strains, antheraxanthin (m/z 585.4231 $[M + H]^+$, $C_{40}H_{56}O_3$, $\Delta = 0.26$ ppm), β -cryptoxanthin (m/z 553.4355 $[M + H]^+$, $C_{40}H_{56}O$, $\Delta = 4.40$ ppm), violaxanthin (m/z 601.4191 $[M + H]^+$, $C_{40}H_{56}O_4$, $\Delta = 2.10$ ppm), and other compounds can be formed through the reversible activity of epoxidase and de-epoxidase enzymes in a process known as the violaxanthin cycle. β -cryptoxanthin is converted back into zeaxanthin by hydrogenase, which removes a hydrogen atom from the C3' position of β -cryptoxanthin. β -carotene is also a precursor for the biosynthesis of echinenone and canthaxanthin, both produced by a β -carotene ketolase (*crtS*) and a hydroxylase (*CrtR*), which can convert echinenone to zeaxanthin (Supplementary Table 2). Echinenone can be hydroxylated to hydroxyechinenone by a hydroxylase enzyme. *A. elenkinii* CCIBt3563 produced β -carotene, canthaxanthin, and echinenone (Supplementary Table 4). Curiously, canthaxanthin was detected exclusively in this strain. Additionally, exclusive compounds detected by LC-MS/MS include myxoxanthophylls, myxol 2'-glucoside similar (m/z 746.4761, $C_{46}H_{66}O_8$, $\Delta = 0.49$ ppm), and β -apo-2'-carotenal (m/z 509.3717 $[M + H]^+$, $C_{37}H_{48}O$, $\Delta = 2.32$ ppm), which are directly related to enzymes involved in the xanthophyll cycle (Figure 3 and Supplementary Table 2).

Distinct carotenoid profiles, including glycosylated apocarotenoids, were identified in *A. pantanalense* CENA528, *Geminocystis* sp. CENA526, and *P. rosaneae* CENA516 (Figure 3D). Enzymes like lycopene cyclase (*cruA/cruP*) and carotenoid desaturases add complexity to the carotenoid structures, and carotenoid cleavage dioxygenases (*crtS*, *CCD1*, *CCD4*) form apocarotenoids (Supplementary Table 2). In *A. pantanalense* CENA528, there was a greater production of zeaxanthin, β -cryptoxanthin, and β -carotene (Supplementary Table 4). Putative compounds such as ketohydroxylycopene (m/z 566.4576 $[M]^+$, $C_{40}H_{54}O_2$), glycosyl-4,4'-diaponeurosporenoate (m/z 594.5125, $C_{36}H_{50}O_7$) and 1,2-dihydrolycopene (m/z 538.4392 $[M]^+$, $C_{40}H_{58}$) were also produced by this strain. *CruA/CruP* catalyzes lycopene cyclization, leading to 1,2-dihydrolycopene formation, while *CCD1* and *CCD4* cleave carotenoid molecules to form apocarotenoids. In *P. rosaneae* CENA516, β -carotene, zeaxanthin, and echinenone were the most abundant carotenes (Supplementary Table 4). Exclusively detected compounds include tetrahydro-2,2-diketo- β -carotene (m/z 561.3655 $[M + H]^+$, $C_{40}H_{48}O_2$, $\Delta = 0.12$ ppm) and an apocarotenoid (m/z 608.2908). However, *Geminocystis* sp. CENA526 was the only strain where β -carotene was not the most abundant carotenoid; instead, α -cryptoxanthin, zeaxanthin, and echinenone dominated. Several apocarotenoid glycosides (m/z 580.4880, m/z 585.4225, m/z

594.2759, m/z 620.5712, m/z 622.1915, m/z 624.5204, m/z 638.5671, m/z 664.4858, m/z 668.5379) and a putative phenicoxanthin (m/z 581.4175 $[M + H]^+$, $C_{40}H_{52}O_3$) were detected exclusively in this strain (Figure 3 and Supplementary Table 4).

The strain *L. platensis* CENA597 is notable for also producing β -carotene, β -cryptoxanthin and antheraxanthin. The adonixanthin (m/z 582.4039 $[M]^+$, $C_{40}H_{54}O_3$, $\Delta = -3.61$ ppm) and myxoxanthophyll compound (m/z 703.4564, $C_{44}H_{63}O_7$, $\Delta = 0.60$ ppm) were detected exclusively in this strain, likely related to lycopene cyclase (CruA/CruP) and β -carotenoid hydroxylase involved in xanthin myxol formation (Figure 3 and Supplementary Table 2). Another *L. platensis* strain isolated from Pantanal, CENA650 exclusively produced a xanthophyll compound (m/z 542.4881, $C_{40}H_{62}$, $\Delta = 5.43$ ppm), and a carotene group compound (m/z 540.4738 $[M]^+$, $C_{40}H_{60}$, $\Delta = -0.09$ ppm) suggesting that lycopene cyclase (CruA/CruP) might be involved in carotenoid cyclization, forming tetrahydrolycopene. Additionally, lycopene derivatives are also present, likely as complex derivatives of the lycopene molecule. Both *Limnospira* strains CENA597 and CENA650, also produced in high abundance the compound annotated as 3,3'-ditetrahydropyranosyloxysoreniaterene (m/z 729.4725 $[M + H]^+$, $C_{46}H_{66}O_7$, $\Delta = -1.92$ ppm), which is associated with the isoreniaterene pigment biosynthesis pathway and the *crtU* gene (carotenoid phi/chi-ring synthase) predicted in both genomes (Supplementary Tables 2, 4).

4 Discussion

Terpenoids are generated from a complex process that involves several steps with enzymes in specific sequences of reactions and intermediates in the metabolic pathway (Liang et al., 2006). The genes and enzymes involved in the MEP pathway have already been found in several organisms and in cyanobacterial strains (Lichtenthaler, 2000; Pattanaik and Lindberg, 2015). This pathway is the main metabolic route for the biosynthesis of isoprenoids in prokaryotes. The genes related to the MEP pathway of terpene biosynthesis are considered conserved across cyanobacteria, with high similarity even between evolutionarily distant genera and simultaneously share homology with the mevalonate pathway in eukaryotes (Lange et al., 2000). As an essential metabolic route for cyanobacterial biology, inhibiting enzymes in this pathway causes lethal effects, mainly affecting photosynthesis and respiration and inhibiting oxidative activity (Pattanaik and Lindberg, 2015; Rudolf et al., 2021).

Carotenoid pigments play a crucial role in photosynthesis, capturing light and protecting against photo-induced damage (Hashimoto et al., 2018; Jacinavicius et al., 2019). In cyanobacteria, β -carotene is the primary carotenoid, along with zeaxanthin, canthaxanthin, β -cryptoxanthin, and echinenone (Takaichi, 2011). These strains share many enzymatic pathways related to terpenoid biosynthesis; however, the comparative analysis of terpenoid profiles among cyanobacterial strains performed in this study revealed significant biochemical diversity. Unique compounds were annotated in each extract from different Pantanal strains (Supplementary Table 2). The metabolomic analysis showed that the most crucial similarity in terpenoid production is not necessarily found among strains of the same species

but was detected among strains from different families, such as *Geminocystis* sp. CENA526 and *L. platensis* CENA597, a unicellular and a filamentous cyanobacteria, respectively. The strains *Geminocystis* sp. CENA526 and *L. platensis* CENA597, isolated from Salina Grande and Salina Centenário, respectively, showed a consistent presence of hydroxycarotenoids and other shared terpenoid classes (Figures 3C, D), which are metabolites commonly linked to photoprotective functions. At the same time, the results highlight metabolic variations within the same species, as seen by the higher number of different instead of equal terpenoid classes in CENA597 and CENA650, two strains of *Limnospira platensis* (Figure 3C). This observation demonstrates significant genetic and biochemical diversity in terpene production among cyanobacteria from tropical soda lakes, which appear to exhibit lineage-specific metabolomic adaptations that likely reflect genetic factors, niche specialization, and evolutionary responses to the extreme physicochemical parameters of their habitats. These adaptations can be observed in the *Synechococcus* PCC7002 and *Synechococcus elongatus* PCC7942, which present differences in sesquiterpene and monoterpene production due to variations in the terpene synthase enzymes (Chenebault et al., 2023). In addition, the differences in carotenoid composition relate to the presence or absence of specific genes and variations in the catalytic properties of key enzymes (Hirschberg and Chamovitz, 1994; Mochimaru et al., 2005; Takaichi and Mochimaru, 2007). Environmental conditions such as growth stage, light intensity, nitrogen source, and nitrogen concentration, as well as strain-specific characteristics within a species, also contribute to these phenomena (Olaizola and Duerr, 1990; Kłodawska et al., 2019; Llewellyn et al., 2020).

Terpenoid composition variation can be attributed to the presence or absence of specific carotenoid biosynthetic pathways and genes, as well as the distinct characteristics of the enzymes involved. As observed, all genomes share common genes for carotenoid production (*crtB*, *crtP*, *crtQ*, *crtH*, and *crtR*), and the presence of respective compounds appears conserved. This conservation might represent a core metabolome that depends on differential gene expression, regulatory mechanisms, or interactions with other genes, which are influenced by specific environmental conditions in Pantanal, such as salinity, solar radiation, or nutrient availability. Such correlations are observed when genes linked to specific pigment groups (*crtS*, *crtW*, *crtX*, and *crtU*) are distributed among different species (Supplementary Table 2), suggesting adaptive mechanisms, as they may relate to increased regulatory demands (e.g., light regulation) and other environmental characteristics (Liang et al., 2006; Schopf et al., 2013; Ma and Cui, 2022). Another condition that may influence the composition of different specific carotenoid compounds that these strains produce would be their responses to oxidative stress. Carotenoids can neutralize reactive oxygen species and scavenge free radicals (Montenegro et al., 2002). The carotenoid pigments found in this study stood out as more efficient in light capture under conditions of excess or scarcity of light (Stamatakis et al., 2014). *A. elenkinii* CCIBt3563, isolated from Salina da Reserva, exhibited a distinct xanthophyll profile not observed in the other strains. The accumulation of xanthophylls like zeaxanthin has been observed in the cytoplasmic membranes of *Synechococcus* cells grown under high irradiance conditions (Masamoto et al., 1999) and is likely associated with enhanced photoprotection and oxidative stress mitigation, crucial for survival

under conditions of high irradiance. In contrast, *A. pantanalense* CENA528, from Salina Preta, displayed a comparatively lower diversity of terpenoids. This may suggest a narrower ecological amplitude or adaptation to different habitats, where a reduced repertoire of secondary metabolites is sufficient for persistence. Additionally, it might correlate with cyanobacterial blooms, given that each species adapts differently to light availability in the presence of dominant cyanobacteria, such as *A. elenkinii* and *L. platensis*, the reported main bloom-forming species in the Pantanal soda lakes (Andreote et al., 2018; Pellegrinetti et al., 2023).

The *crtB* gene encodes phytoene synthase, an essential enzyme in the carotenoid biosynthesis pathway. The enzyme phytoene synthase (CrtB) initiates the biosynthetic pathway of carotenoids, which is essential for producing various carotenoids. It catalyzes the head-to-head condensation of two molecules of geranylgeranyl pyrophosphate (GGPP), forming phytoene—the first committed step in carotenoid biosynthesis. Subsequent desaturation steps are catalyzed by enzymes such as phytoene desaturase (*crtI*, or its cyanobacterial equivalents *crtP* and *crtQ*), which convert phytoene into carotenoids like neurosporene or lycopene (Sandmann, 2002; Kato et al., 2016). Phytoene has increasingly been associated with photoprotective activities. While phytoene has an absorbance maximum of around 286 nm, effectively protecting cells from UV-B radiation, phytofluene absorbs at 348 nm in the UV-A region (Meléndez-Martínez et al., 2019). Combining these carotenoids in a bioproduct might represent an eco-friendly option for chemical sunscreens related to commercial options using mycosporine-like amino acids for photoprotection (Sen and Mallick, 2021). The phytoene synthase gene found in the Pantanal strains showed evolutionary similarity with strains from brackish water on the Baltic Sea, hypersaline lakes in Tanzania, and freshwater lakes in Brazil and China. These diverse environmental conditions might pose different abiotic stresses, but UV exposure is obligatory for photosynthetic organisms. Recent data show that both *L. platensis* CENA597 and CENA650, as well as *A. elenkinii* CCIBt3563, lack the genetic arrangement for mycosporine-like amino acid production, failing to biosynthesize these compounds even when exposed to UV radiation (Dextro et al., 2023). Therefore, it is possible to assume that other photoprotective strategies, such as using carotenoids, are employed by these strains to survive and even form blooms in these highly UV-exposed habitats.

Strains *L. platensis* CENA597, *L. platensis* CENA650, and *A. elenkinii* CCIBt3563 likely share the same adaptive mechanisms that grant endurance in habitats highly exposed to UV radiation, which affects the photosynthetic efficiency of cyanobacteria (Jacinavicius et al., 2021). The xanthophyll cycle was highly active in these strains, with the identification of unique compounds present in this pathway (xanthophyll, adonixanthin, myxoxanthophyll, canthaxanthin, and 2'-fucosyl-4-hydroxy-myxol-2'-fucoside). Xanthophylls are oxygen-containing carotenoids that include functional groups such as aldehydes, carboxyl, and epoxides (Maoka, 2020). Xanthophylls have potent antioxidant activity that contributes to the integrity of the thylakoid membrane against excess light, which makes these molecules significant photoprotective compounds (Srivastava et al., 2022). Exposure to UV-B led to changes in carotenoid transcription regulation, affecting

carotenoid synthesis, photoprotection, and cleavage, with increased expression of ketolase-related genes resulting in higher concentrations of echinenone, canthaxanthin, and myxoxanthophyll that reduced DNA damage and oxidative stress and absorbed radiation, potentially promoting DNA repair (Ehling-Schulz et al., 1997; Llewellyn et al., 2020). The pigments β -cryptoxanthin, echinenone, canthaxanthin, astaxanthin, and zeaxanthin have antioxidant properties, acting against reactive oxygen species within the cell, protecting molecules from oxidative stress (Gruszecki and Strzalka, 2005; Sedoud et al., 2014; Pereira et al., 2021). Additionally, xanthophylls have anti-inflammatory and antitumor activities due to their antioxidant properties and double bonds in their chemical structure, highlighting their commercial appeal for biotechnology in the cosmetics and food industries (Álvarez et al., 2014; Pereira et al., 2021).

Genes involved in the carotenogenesis pathway, from geranylgeranyl-diphosphate to lycopene, are also upregulated following exposure to UV-B light radiation (Llewellyn et al., 2020). Lycopene acts as a precursor to cyclic β -carotenes, as well as to xanthophyll carotenoids, with lycopene cyclase being responsible for catalyzing the formation of cyclic end groups of these carotenoids (Domonkos et al., 2013). Various types of lycopene cyclase, including the first identified type, *crtL* (*crtL*-b, *LCY*-b), have been confirmed in cyanobacteria (Sugiyama and Takaichi, 2020). Cyanobacteria whose genome does not contain genes for lycopene beta cyclase (*crtL*) have a gene like *cruA* (ortholog), which may perform a similar or compensatory function to lycopene cyclase (Maresca et al., 2007). *cruA/cruP* was detected in all strains of the Pantanal genome. The homologous genes of *cruA* and *cruP* are widely distributed in some cyanobacterial species' genomes; however, few reports demonstrate lycopene cyclase activity in products (Sugiyama and Takaichi, 2020). Modifications to carotenes can produce glycosylated carotenoids, which are also very common in cyanobacteria (Hirschberg and Chamovitz, 1994; Takaichi et al., 2005). Specific glycosylated carotenoids, likely as complex derivatives of the lycopene molecule, are present in *Geminocystis* sp. CENA526 and *A. pantanalense* CENA528. Cyanobacteria synthesize these apocarotenoids through a complex biosynthetic pathway involving carotenogenic enzymes via the activity of carotenoid cleavage dioxygenases (CCDs) (Liang et al., 2017; Hambly et al., 2021). These activities reflect the complexity and diversity in carotenoid biosynthesis among these strains. In cyanobacteria, apocarotenoids play crucial roles in protecting photosynthetic machinery from oxidative damage caused by excessive light, capturing light, and transferring energy to the photosynthetic complexes; still, many apocarotenoids have biological functions that remain challenging to determine and are not yet fully understood (Liang et al., 2017).

The hopanoid bacteriohopanetetrol was exclusively annotated in the *Geminocystis* sp. CENA526 strain, suggesting a variation in lipid composition and osmotolerance between the strains (Supplementary Table 2). Hopene and other hopanoids occur in a wide range of Gram-positive and Gram-negative bacteria, supporting small bacterial adaptation to extreme conditions, like high temperatures and pH gradients (Carvalho and Fernandes, 2010). The presence of detectable hopanoids likely protects the membrane in *Geminocystis* sp. CENA26, suggesting an adaptive strategy to the Pantanal lakes, where high salinity and intense

UV radiation occur. *sqhC* is a primitive condition gene found in several modern bacterial genera, but it is estimated to occur in only 5–10% of all bacterial species (Kharbush et al., 2013). The synteny analysis for the predicted SHC enzyme gene showed an identity score of around 60% between the Pantanal strains and the sequences of other cyanobacteria. This analysis corroborates with the *sqhC* tree (Supplementary Figure 5) and the works of Pearson et al. (2007) and Kharbush et al. (2013), where SHCs have a similarity between 45 and 55% within cyanobacterial groups, with an average of 60% similarity between the closest taxonomical subgroups and about 90% for species within the same genus. A study with environmental and culturable samples suggests that possessing a squalene hopene cyclase (SHC) protein implies that cyanobacterial strains in culturable conditions will have detectable levels of bacteriohopanepolyols (BHPs) (Talbot et al., 2008, 2016). Additionally, structural variants of 2-methyl-BHPs, similar to those found for *Geminocystis* sp. CENA526 is present in different cyanobacterial mats from hypersaline marine lagoons in Australia and Mexico (Talbot et al., 2008). This difference reflects the structural and genetic diversity caused by environmental gradients that these organisms face. The predicted clusters showed similarity only for the SHC enzyme since the evaluated genera are not phylogenetically close to each other, thus indicating that this gene might be conserved among the entire phylum Cyanobacteria. In the phylogenetic analysis of *sqhC*, taxonomical similarity was observed, especially in grouping the gene copies from the *Geminocystis* genus, with sequence similarity of over 80%, forming a branch with common ancestry (Supplementary Figure 5). For the homologs genes annotated in *A. pantanalense* CENA528 and *P. rosaneae* CENA516, which does not present detectable production of hopanoids, that grouped with strains isolated from hot springs and freshwater samples, it can be argued that a better evolutionary resolution might only be achieved with further description of cyanobacterial diversity from wide environmental backgrounds, improving possible inferences between shared gene history in cyanobacteria (Dextro et al., 2021).

In addition to tetraterpenoids (carotenoids) and triterpenoids (hopanoids), meroterpenoids (ubiquinones) were also detected. The evolutionary trajectory of ubiquinone pathways suggests that the first set of ubiquinone biosynthetic enzymes were present in cyanobacteria that produced phylloquinone, so the current genes of these metabolic pathways found in other prokaryotes and eukaryotes are homologs inherited from cyanobacteria (Degli Esposti, 2017). Tocopherol was detected in all strains and has antioxidant activity, protecting membranes against the auto-oxidation of polyunsaturated fatty acids (Inoue et al., 2011). Meanwhile, menaquinone and phylloquinone (vitamin K) participate in photosystem I electron transfer (Gross et al., 2006; Sadre et al., 2012). These substances produced by cyanobacteria have the potential to be incorporated into nutraceuticals and pharmaceuticals, fortifying food for human or animal supplementation and acting as natural antioxidants (Kawamukai, 2018).

As observed, the intensity and quality of solar radiation might influence the production of terpenoids, which play a crucial role in protection against UV damage and oxidation. Likewise, salinity and nutrient changes cause carotenoid composition variations (Paliwal et al., 2015). Extreme conditions, such as those found in the Pantanal biome, with high salt concentrations, pH

variations, or low water availability, can stimulate the production of specific carotenoids that help organisms survive in these adverse conditions. This biochemical diversity is the main factor in Cyanobacteria's success in colonizing distinct habitats. Also, these carbon pathways play a crucial role in the tropical soda lakes as cyanobacterial blooms mediate carbon and energy fluxes, where Cyanobacteria represent a leading agent in supporting the heterotrophic bacterial community, particularly during the dry season (Cotta et al., 2022).

5 Conclusion

The genomes of cyanobacterial strains isolated from the soda lakes of the Pantanal biome were mined for genes associated with terpene and terpenoid biosynthesis, identifying several homologs. The studied strains shared many enzymatic pathways related to terpenoid biosynthesis. Still, they mainly differed in their responses to the production of tetraterpenoids, results corroborated by the chemical profile of the samples. Synteny and phylogenetic analysis of critical enzymes such as scalene hopene cyclase and phytoene synthase highlighted sequence conservation and shared evolutionary histories with cyanobacteria from diverse environments. Metabolomic analysis revealed significant diversity and specificity in carotenoid profiles, with unique features contributing to the strains' adaptation to extreme conditions in the soda lakes. This study links genome mining with metabolomics, which serves as a starting point for research focused on biotechnological endpoints. The detection of bioactive compounds, including carotenoids, hopanoids, and tocopherols, further highlights the biotechnological potential of these strains. This potential extends to applications in pharmaceuticals, nutraceuticals, and cosmetics, opening new avenues for drug development and health supplements. The unique metabolic adaptations and biosynthetic pathways identified in these cyanobacteria provide a valuable foundation for further exploring their metabolites' ecological roles and bioprospecting potential, especially under extreme environmental conditions.

Data availability statement

The genomic sequences for this study have been deposited and are available in the NCBI database under the following accession numbers: *Alkalinema pantanalense* CENA528 (JBLZFX000000000), *Anabaenopsis elenkinii* CCIBt3563 (CP063311), *Geminocystis* sp. CENA526 (JBLZFY000000000), *Limnospira platensis* CENA597 (CP185278), *Limnospira platensis* CENA650 (JBLZFW000000000), and *Pantanalinema rosaneae* CENA516 (JBLZFZ000000000).

Author contributions

MM: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Writing – original draft, Writing – review & editing. FJ: Conceptualization, Data curation,

Formal Analysis, Investigation, Methodology, Software, Writing – original draft, Writing – review & editing. RM: Data curation, Formal Analysis, Investigation, Methodology, Writing – review & editing. RD: Data curation, Formal Analysis, Investigation, Methodology, Writing – review & editing. AF: Formal Analysis, Software, Writing – review & editing. MW: Data curation, Writing – review & editing, Formal Analysis. TP: Writing – review & editing. SC: Writing – review & editing. CC: Conceptualization, Methodology, Writing – review & editing. MF: Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the State of São Paulo Research Foundation (FAPESP, 2016/14227-5 and 2022/02872-4). Joint Genome Institute funded long-read sequencing (JGI Project ID: 504264). MM, MW, and RM received scholarships from the Brazilian Federal Agency for the Support and Evaluation of Graduate Education (CAPES, Finance code 001). TP, SC, RD, and FJ were supported by FAPESP (2017/12644-0, 2018/24049-2, 2023/08974-6, 2024/02448-3). RD and MM received a doctoral scholarship (140892/2020-6, 140991/2022-0) from the Brazilian National Council for Scientific and Technological Development (CNPq). MF received a research fellowship (309476/2022-4) from CNPq. The authors thank Fundação de Estudos Agrários Luiz de Queiroz (FEALQ) for the contribution to the payment of APC.

References

- Alvarenga, D. O., Fiore, M. F., and Varani, A. M. (2017). A metagenomic approach to cyanobacterial genomics. *Front. Microbiol.* 8:809. doi: 10.3389/fmicb.2017.00809
- Álvarez, R., Vaz, B., Gronemeyer, H., and de Lera, Á. R. (2014). Functions, therapeutic applications, and synthesis of retinoids and carotenoids. *Chem. Rev.* 114, 1–125. doi: 10.1021/cr400126u
- Andreote, A. P., Dini-Andreote, F., Rigonato, J., Machineski, G. S., Souza, B. C., Barbiero, L., et al. (2018). Contrasting the genetic patterns of microbial communities in soda lakes with and without cyanobacterial bloom. *Front. Microbiol.* 9:244. doi: 10.3389/fmicb.2018.00244
- Avalos, M., Garbeva, P., Vader, L., van Wezel, G. P., Dickschat, J. S., and Ulanova, D. (2022). Biosynthesis, evolution and ecology of microbial terpenoids. *Nat. Prod. Rep.* 39, 249–272. doi: 10.1039/D1NP00047K
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: A New genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021
- Blin, K., Shaw, S., Kloosterman, A. M., Charlop-Powers, Z., Van Wezel, G. P., Medema, M. H., et al. (2021). AntiSMASH 6.0: Improving cluster detection and comparison capabilities. *Nucleic Acids Res.* 49, W29–W35. doi: 10.1093/nar/gkab335
- Boetzer, M., Henkel, C. V., Jansen, H. J., Butler, D., and Pirovano, W. (2011). Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27, 578–579. doi: 10.1093/bioinformatics/btq683
- Cabanillas, A. H., Tena Pérez, V., Maderuelo Corral, S., Rosero Valencia, D. F., Martel Quintana, A., Ortega Doménech, M., et al. (2018). Cybastacines A and B: Antibiotic sesterterpenes from a *Nostoc* sp. cyanobacterium. *J. Nat. Prod.* 81, 410–413. doi: 10.1021/acs.jnatprod.7b00638
- Carvalho, C. C., and Fernandes, P. (2010). Production of metabolites as bacterial responses to the marine environment. *Mar. Drugs* 8, 705–727. doi: 10.3390/md8030705
- Chenebault, C., Blanc-Garin, V., Vincent, M., Diaz-Santos, E., Goudet, A., Cassier-Chauvat, C., et al. (2023). Exploring the potential of the model cyanobacteria *Synechococcus* PCC 7002 and PCC 7942 for the photoproduction of high-value terpenes: A comparison with *Synechocystis* PCC 6803. *Biomolecules* 13:504. doi: 10.3390/biom13030504
- Chlipala, G. E., Mo, S., and Orjala, J. (2011). Chemodiversity in freshwater and terrestrial cyanobacteria - a source for drug discovery. *Curr. Drug Targets* 12, 1654–1673. doi: 10.2174/138945011798109455
- Christianson, D. W. (2017). Structural and chemical biology of terpenoid cyclases. *Chem. Rev.* 117, 11570–11648. doi: 10.1021/acs.chemrev.7b00287
- Chu, C., Li, X., and Wu, Y. (2019). GAPPadder: A sensitive approach for closing gaps on draft genomes with short sequence reads. *BMC Genom.* 20:426. doi: 10.1186/s12864-019-5800-2
- Cotta, S. R., Pellegrinetti, T. A., Andreote, A. P. D., Costa, J. S., Sarmiento, H., and Fiore, M. F. (2022). Disentangling the lifestyle of bacterial communities in tropical soda Lakes. *Sci. Rep.* 12, 1–10. doi: 10.1038/s41598-022-12046-2
- Degli Esposti, M. (2017). A journey across genomes uncovers the origin of ubiquinone in cyanobacteria. *Genom. Biol. Evol.* 9, 3039–3053. doi: 10.1093/gbe/evx225
- Dextro, R. B., Delbaje, E., Cotta, S. R., Zehr, J. P., and Fiore, M. F. (2021). Trends in free-access genomic data accelerate advances in cyanobacteria taxonomy. *J. Physiol.* 57, 1392–1402. doi: 10.1111/jpy.13200
- Dextro, R. B., Delbaje, E., Galdes, V., Pinto, E., Long, P. F., and Fiore, M. F. (2023). Exploring the relationship between biosynthetic gene clusters and constitutive production of mycosporine-like amino acids in Brazilian cyanobacteria. *Molecules* 28:1420. doi: 10.3390/molecules28031420
- Domonkos, I., Kis, M., Gombos, Z., and Ughy, B. (2013). Carotenoids, versatile components of oxygenic photosynthesis. *Prog. Lipid Res.* 52, 539–561. doi: 10.1016/j.plipres.2013.07.001

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2025.1582103/full#supplementary-material>

- Duhrkop, K., Fleischauer, M., Ludwig, M., Aksenov, A. A., Melnik, A. V., Meusel, M., et al. (2019). SIRIUS 4: A rapid tool for turning tandem mass spectra into metabolite structure information. *Nat. Methods* 16, 299–302. doi: 10.1038/s41592-019-0344-8
- Duhrkop, K., Nothias, L. F., Fleischauer, M., Reher, R., Ludwig, M., Hoffmann, M. A., et al. (2021). Systematic classification of unknown metabolites using high-resolution fragmentation mass spectra. *Nat. Biotechnol.* 39, 462–471. doi: 10.1038/s41587-020-0740-8
- Ehling-Schulz, M., Bilger, W., and Scherer, S. (1997). UV-B-induced synthesis of photoprotective pigments and extracellular polysaccharides in the terrestrial cyanobacterium *Nostoc commune*. *J. Bacteriol.* 179, 1940–1945. doi: 10.1128/jb.179.6.1940-1945.1997
- Ernst, M., Kang, K. B., Caraballo-Rodríguez, A. M., Nothias, L. F., Wandy, J., Chen, C., et al. (2019). MolNetEnhancer: Enhanced molecular networks by integrating metabolome mining and annotation tools. *Metabolites* 9:144. doi: 10.3390/metabo9070144
- Gilchrist, C. L., and Chooi, Y. H. (2021). Clinker & clustermap.js: Automatic generation of gene cluster comparison figures. *Bioinformatics* 37, 2473–2475. doi: 10.1093/bioinformatics/btab007
- Gross, J., Cho, W. K., Lezhneva, L., Falk, J., Krupinska, K., Shinozaki, K., et al. (2006). A plant locus essential for phyloquinone (vitamin K1) biosynthesis originated from a fusion of four eubacterial genes. *J. Biol. Chem.* 281, 17189–17196. doi: 10.1074/jbc.M601754200
- Gruszecki, W. I., and Strzalka, K. (2005). Carotenoids as modulators of lipid membrane physical properties. *Biochim. Biophys. Acta* 1740, 108–115. doi: 10.1016/j.bbadis.2004.11.015
- Gurevich, A., Mikheenko, A., Shlemov, A., Korobeynikov, A., Mohimani, H., and Pevzner, P. A. (2018). Increased diversity of peptidic natural products revealed by modification-tolerant database search of mass spectra. *Nat. Microbiol.* 3, 319–327. doi: 10.1038/s41564-017-0094-2
- Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: Quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072–1075. doi: 10.1093/bioinformatics/btt086
- Hambly, A. J., Van Duijneveldt, J. S., and Gates, P. J. (2021). Identification of β -carotene oxidation products produced by bleaching clay using UPLC-ESI-MS/MS. *Food Chem.* 353, 129455. doi: 10.1016/j.foodchem.2021.129455
- Hashimoto, H., Uragami, C., Yukihira, N., Gardiner, A. T., and Cogdell, R. J. (2018). Understanding/unravelling carotenoid excited singlet states. *J. R. Soc. Interface* 15:20180026. doi: 10.1098/rsif.2018.0026
- Hatha, A. M., and Sumayya, N. S. (2023). “Antioxidants from marine cyanobacteria,” in *Marine antioxidants*, eds S. K. Kim, K. H. Shin, and J. Venkatesan (Cambridge, MA: Academic Press), 119–131. doi: 10.1016/B978-0-323-95086-2.00010-2
- Hirschberg, J., and Chamovitz, D. (1994). “Carotenoids in cyanobacteria,” in *The molecular biology of cyanobacteria*, ed. D. A. Bryant Dordrecht (Dordrecht: Springer Netherlands), 559–579. doi: 10.1007/0-306-48205-3_18
- Inoue, S., Ejima, K., Iwai, E., Hayashi, H., Appel, J., Tyystjärvi, E., et al. (2011). Protection by α -tocopherol of the repair of photosystem II during photoinhibition in *Synechocystis* sp. PCC 6803. *Biochim. Biophys. Acta Bioenerg.* 1807, 236–241. doi: 10.1016/j.bbabi.2010.11.003
- Jacinavicius, F. R., Gerales, V., Crnkovic, C. M., Delbaje, E., Fiore, M. F., and Pinto, E. (2021). Effect of ultraviolet radiation on the metabolomic profiles of potentially toxic cyanobacteria. *FEMS Microbiol. Ecol.* 97:fiab243. doi: 10.1093/femsec/fiab243
- Jacinavicius, F. R., Pacheco, A. B. F., Chow, F., da Costa, G. C. V., Kalume, D. E., Rigonato, J., et al. (2019). Different ecophysiological and structural strategies of toxic and non-toxic *Microcystis aeruginosa* (cyanobacteria) strains assessed under culture conditions. *Algal Res.* 41:101548. doi: 10.1016/j.algal.2019.101548
- Jaki, B., Orjala, J., and Sticher, O. (1999). A novel extracellular diterpenoid with antibacterial activity from the cyanobacterium *Nostoc commune*. *J. Nat. Prod.* 62, 502–503. doi: 10.1021/np980444x
- Jones, M. R., Pinto, E., Torres, M. A., Dörr, F., Mazur-Marzec, H., Szubert, K., et al. (2021). CyanoMetDB, a comprehensive public database of secondary metabolites from cyanobacteria. *Water Res.* 196:117017. doi: 10.1016/j.watres.2021.117017
- Kanehisa, M., Sato, Y., and Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG Tools for functional characterization of genome and metagenome sequences. *J. Mol. Biol.* 428, 726–731. doi: 10.1016/j.jmb.2015.11.006
- Kato, S., Takaichi, S., Ishikawa, T., Asahina, M., Takahashi, S., and Shinomura, T. (2016). Identification and functional analysis of the geranylgeranyl pyrophosphate synthase gene (*crtE*) and phytoene synthase gene (*crtB*) for carotenoid biosynthesis in *Euglena gracilis*. *BMC Plant Biol.* 16:1–12. doi: 10.1186/s12870-015-0698-8
- Kawamukai, M. (2018). Biosynthesis and applications of prenylquinones. *Biosci. Biotechnol. Biochem.* 82, 963–977. doi: 10.1080/09168451.2018.1433020
- Kharbush, J. J., Ugalde, J. A., Hogle, S. L., Allen, E. E., and Aluwihare, L. I. (2013). Composite bacterial hopanoids and their microbial producers across oxygen gradients in the water column of the California current. *Appl. Environ. Microbiol.* 79, 7491–7501. doi: 10.1128/AEM.02367-13
- Kłodawska, K., Bujas, A., Turos-Cabal, M., Żbik, P., Fu, P., and Malec, P. (2019). Effect of growth temperature on biosynthesis and accumulation of carotenoids in cyanobacterium *Anabaena* sp. PCC 7120 under diazotrophic conditions. *Microbiol. Res.* 226, 34–40. doi: 10.1016/j.micres.2019.05.003
- Kolmogorov, M., Yuan, J., Lin, Y., and Pevzner, P. A. (2019). Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* 37, 540–546. doi: 10.1038/s41587-019-0072-8
- Kotai, J. (1972). Instructions for preparation of modified nutrient solution Z8 for algae. *Norwegian Inst. Water Res. Oslo* 11:5. doi: 10.1127/0029-5035/2004/0079-0099
- Lange, B. M., Rujan, T., Martin, W., and Croteau, R. (2000). Isoprenoid biosynthesis: The evolution of two ancient and distinct pathways across genomes. *Proc. Natl. Acad. Sci. U.S.A.* 97, 13172–13177. doi: 10.1073/pnas.24045479
- Liang, C., Zhao, F., Wei, W., Wen, Z., and Qin, S. (2006). Carotenoid biosynthesis in cyanobacteria: Structural and evolutionary scenarios based on comparative genomics. *Int. J. Biol. Sci.* 2:197. doi: 10.7150/ijbs.2.197
- Liang, M. H., Zhu, J., and Jiang, J. G. (2017). Carotenoids biosynthesis and cleavage related genes from bacteria to plants. *Crit. Rev. Food Sci. Nutr.* 58, 2314–2333. doi: 10.1080/10408398.2017.1322552
- Lichtenthaler, H. K. (2000). Non-mevalonate isoprenoid biosynthesis: Enzymes, genes and inhibitors. *Biochem. Soc. Trans.* 28, 785–789. doi: 10.1042/bst0280785
- Lin, P.-C., and Pakrasi, H. B. (2019). Engineering cyanobacteria for production of terpenoids. *Planta* 249, 145–154. doi: 10.1007/s00425-018-3047-y
- Llewellyn, C. A., Ains, R. L., Farnham, G., and Greig, C. (2020). Synthesis, regulation and degradation of carotenoids under low-level UV-B radiation in the filamentous cyanobacterium *Chlorogloeopsis fritschii* PCC 6912. *Front. Microbiol.* 11:163. doi: 10.3389/fmicb.2020.00163
- Lu, J., Rincon, N., and Wood, D. E. (2022). Metagenome analysis using the Kraken software suite. *Nat. Protoc.* 17, 2815–2839. doi: 10.1038/s41596-022-00738-y
- Ma, J., and Cui, X. (2022). Aromatic carotenoids: Biological sources and geological implications. *Geosyst. Geoenviron.* 1:100045. doi: 10.1016/j.jgeo.2022.100045
- Maoka, T. (2020). Carotenoids as natural functional pigments. *J. Nat. Med.* 74, 1–16. doi: 10.1007/s11418-019-01364-x
- Maresca, J. A., Graham, J. E., Wu, M., Eisen, J. A., and Bryant, D. A. (2007). Identification of a fourth family of lycopene cyclases in photosynthetic bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 104, 11784–11789. doi: 10.1073/pnas.0702984104
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17, 10–12. doi: 10.14806/ej.17.1.200
- Masamoto, K., Zsiros, O., and Gombos, Z. (1999). Accumulation of zeaxanthin in cytoplasmic membranes of the cyanobacterium *Synechococcus* sp. strain PCC7942 grown under high light condition. *J. Plant Physiol.* 155, 136–138. doi: 10.1016/S0176-1617(99)80155-2
- Meléndez-Martínez, A. J., Stingo, C. M., and Mapelli-Brahm, P. (2019). Skin carotenoids in public health and nutraceuticals: The emerging roles and applications of the UV radiation-absorbing colourless carotenoids phytoene and phytofluene. *Nutrients* 11:1093. doi: 10.3390/nu11051093
- Micallef, M. L., D’Agostino, P. M., Sharma, D., Viswanathan, R., and Moffitt, M. C. (2015). Genome mining for natural product biosynthetic gene clusters in the Subsection V cyanobacteria. *BMC Genom.* 16:1–20. doi: 10.1186/s12864-015-1855-z
- Mo, S., Krunic, A., Pegan, S. D., Franzblau, S. G., and Orjala, J. (2009). An antimicrobial guanidine-bearing sesterterpene from the cultured cyanobacterium *Scytonema* sp. *J. Nat. Prod.* 72, 2043–2045. doi: 10.1021/np900288x
- Mochimaru, M., Masukawa, H., and Takaichi, S. (2005). The cyanobacterium *Anabaena* sp. PCC 7120 has two distinct β -carotene ketolases: *CrtO* for echinenone and *CrtW* for ketomoxol synthesis. *FEBS Lett.* 579, 6111–6114. doi: 10.1016/j.febslet.2005.09.081
- Mohimani, H., Gurevich, A., Shlemov, A., Mikheenko, A., Korobeynikov, A., Cao, L., et al. (2018). Dereplication of microbial metabolites through database search of mass spectra. *Nat. Commun.* 9:14035. doi: 10.1038/s41467-018-06082-8
- Montenegro, M. A., Nazareno, M. A., Durantini, E. N., and Borsarelli, C. D. (2002). Singlet molecular oxygen quenching ability of carotenoids in a reverse-micelle membrane mimetic system. *Photochem. Photobiol.* 75, 353–361. doi: 10.1562/0031-865520020750353SMOQAO2.0.CO2
- Morone, J., Hentschke, G., Pinto, E., Morais, J., Cruz, P., Vasconcelos, V., et al. (2024). Carotenoids from cyanobacteria modulate iNOS and inhibit the production of inflammatory mediators: Promising agents for the treatment of inflammatory conditions. *Algal Res.* 83:103729. doi: 10.1016/j.algal.2024.103729
- Nandagopal, P., Steven, A. N., Chan, L. W., Rahmat, Z., Jamaluddin, H., and Mohd Noh, N. I. (2021). Bioactive metabolites produced by cyanobacteria for growth adaptation and their pharmacological properties. *Biology* 10:1061. doi: 10.3390/biology10101061
- Nazir, M., Saleem, M., Tousif, M. I., Anwar, M. A., Surup, F., Ali, I., et al. (2021). Meroterpenoids: A comprehensive update insight on structural diversity and biology. *Molecules* 11:957. doi: 10.3390/molecules11070957
- Nurk, S., Meleshko, D., Korobeynikov, A., and Pevzner, P. A. (2017). metaSPAdes: A new versatile metagenomic assembler. *Genome Res.* 27, 824–834. doi: 10.1101/gr.213959.116

- O'Connell, J., Schulz-Trieglaff, O., Carlson, E., Hims, M. M., Gormley, N. A., and Cox, A. J. (2015). NxTrim: Optimized trimming of Illumina mate-pair reads. *Bioinformatics* 31, 2035–2037. doi: 10.1093/bioinformatics/btv057
- Olaizola, M., and Duerr, E. O. (1990). Effects of light intensity and quality on the growth rate and photosynthetic pigment content of *Spirulina platensis*. *J. Appl. Phycol.* 2, 97–104. doi: 10.1007/BF00023370
- Paliwal, C., Pancha, I., Ghosh, T., Maurya, R., Chokshi, K., Bharadwaj, S. V., et al. (2015). Selective carotenoid accumulation by varying nutrient media and salinity in *Synechocystis* sp. CCNM 2501. *Bioresour. Technol.* 197, 363–368. doi: 10.1016/j.biortech.2015.08.122
- Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. doi: 10.1101/gr.186072.114
- Patiny, L., and Borel, A. (2013). ChemCalc: A building block for tomorrow's chemical infrastructure. *J. Chem. Inf. Model.* 53, 1223–1228. doi: 10.1021/ci300563h
- Pattanaik, B., and Lindberg, P. (2015). Terpenoids and their biosynthesis in cyanobacteria. *Life* 5, 269–293. doi: 10.3390/life5010269
- Pearson, A., Flood Page, S. R., Jorgenson, T. L., Fischer, W. W., and Higgins, M. B. (2007). Novel hopanoid cyclases from the environment. *Environ. Microbiol.* 9, 2175–2188. doi: 10.1111/j.1462-2920.2007.01331.x
- Pellegrinetti, T. A., Cotta, S. R., Feitosa, Y. B., Melo, P. L., Bieluczyk, W., Silva, A. M., et al. (2024). The role of microbial communities in biogeochemical cycles and greenhouse gas emissions within tropical soda lakes. *Sci. Tot. Environ.* 947:174646. doi: 10.1016/j.scitotenv.2024.174646
- Pellegrinetti, T. A., Cotta, S. R., Sarmiento, H., Costa, J. S., Delbaje, E., Montes, C. R., et al. (2023). Bacterial communities along environmental gradients in tropical soda lakes. *Microb. Ecol.* 85, 892–903. doi: 10.1007/s00248-022-02086-6
- Pereira, A. G., Otero, P., Echave, J., Carreira-Casais, A., Chamorro, F., Collazo, N., et al. (2021). Xanthophylls from the sea: Algae as source of bioactive carotenoids. *Mar. Drugs* 19:188. doi: 10.3390/md19040188
- Prinsep, M. R., Thomson, R. A., West, M. L., and Wylie, B. L. (1996). Tolypodiol, an antiinflammatory diterpenoid from the cyanobacterium *Tolypothrix nodosa*. *J. Nat. Prod.* 59, 786–788. doi: 10.1021/np9602574
- Rautela, A., Yadav, I., Gangwar, A., Chatterjee, R., and Kumar, S. (2024). Photosynthetic production of α -farnesene by engineered *Synechococcus elongatus* UTEX 2973 from carbon dioxide. *Bioresour. Technol.* 396:130432. doi: 10.1016/j.biortech.2024.130432
- Rudolf, J. D., Alsup, T. A., Xu, B., and Li, Z. (2021). Bacterial terpenome. *Nat. Prod. Rep.* 38, 905–980. doi: 10.1039/D0NP00066C
- Sadre, R., Pfaff, C., and Buchkremer, S. (2012). Plastoquinone-9 biosynthesis in cyanobacteria differs from that in plants and involves a novel 4-hydroxybenzoate solanesyltransferase. *Biochem. J.* 442, 621–629. doi: 10.1042/BJ20111796
- Sandmann, G. (2002). Molecular evolution of carotenoid biosynthesis from bacteria to plants. *Physiol. Plant* 116, 431–440. doi: 10.1034/j.1399-3054.2002.1160401.x
- Schopf, L., Mautz, J., and Sandmann, G. (2013). Multiple ketolases involved in light regulation of canthaxanthin biosynthesis in *Nostoc punctiforme* PCC 73102. *Planta* 237, 1279–1285. doi: 10.1007/s00425-013-1846-8
- Sedoud, A., López-Igual, R., Rehman, A., Wilson, A., Perreau, F., Boulay, C., et al. (2014). The cyanobacterial photoactive orange carotenoid protein is an excellent singlet oxygen quencher. *Plant Cell* 26, 1781–1791. doi: 10.1105/tpc.114.123802
- Seemann, T. (2014). Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- Sen, S., and Mallick, N. (2021). Mycosporine-like amino acids: Algal metabolites shaping the safety and sustainability profiles of commercial sunscreens. *Algal Res.* 58:102425. doi: 10.1016/j.algal.2021.102425
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212. doi: 10.1093/bioinformatics/btv351
- Smith, C. A., O'Maille, G., Want, E. J., Qin, C., Trauger, S. A., Brandon, T. R., et al. (2005). METLIN: A metabolite mass spectral database. *Ther. Drug Monit.* 27, 747–751. doi: 10.1097/01.fid.0000179845.53213.39
- Srivastava, A., Thapa, S., Chakdar, H., Babele, P. K., and Shukla, P. (2022). Cyanobacterial myxoxanthophylls: Biotechnological interventions and biological implications. *Crit. Rev. Biotechnol.* 44, 63–77. doi: 10.1080/07388551.2022.2117682
- Stamatakis, K., Tsimilli-Michael, M., and Papageorgiou, G. C. (2014). On the question of the light-harvesting role of β -carotene in photosystem II and photosystem I core complexes. *Plant Physiol. Biochem.* 81, 121–127. doi: 10.1016/j.plaphy.2014.01.014
- Sugiyama, K., and Takaichi, S. (2020). Carotenogenesis in cyanobacteria: CruA/CruP-type and CrtL-type lycopene cyclases. *J. Gen. Appl. Microbiol.* 66, 53–58. doi: 10.2323/jgam.2020.01.005
- Takaichi, S. (2011). Carotenoids in algae: Distributions, biosyntheses and functions. *Mar. Drugs* 9, 1101–1118. doi: 10.3390/md9061101
- Takaichi, S., and Mochimaru, M. (2007). Carotenoids and carotenogenesis in cyanobacteria: Unique ketocarotenoids and carotenoid glycosides. *Cell. Mol. Life Sci.* 64, 2607–2619. doi: 10.1007/s00018-007-7190-z
- Takaichi, S., Mochimaru, M., Maoka, T., and Katoh, H. (2005). Myxol and 4-ketomyxol 2'-fucosides, not rhamnosides, from *Anabaena* sp. PCC 7120 and *Nostoc punctiforme* PCC 73102, and proposal for the biosynthetic pathway of carotenoids. *Plant Cell Physiol.* 46, 497–504. doi: 10.1093/pcp/pci049
- Talbot, H. M., Sidgwick, F. R., Bischoff, J., Osborne, K. A., Rush, D., Sherry, A., et al. (2016). Analysis of non-derivatised bacteriohopanepolyols by ultrahigh-performance liquid chromatography/tandem mass spectrometry. *Rapid Commun. Mass Spectr.* 30, 2087–2098. doi: 10.1002/rcm.7696
- Talbot, H. M., Summons, R. E., Jahnke, L. L., Cockell, C. S., Rohmer, M., and Farrimond, P. (2008). Cyanobacterial bacteriohopanepolyol signatures from cultures and natural environmental settings. *Org. Geochem.* 39, 232–263. doi: 10.1016/j.orggeochem.2007.08.006
- Tamura, K., Stecher, G., and Kumar, S. (2021). MEGA11: Molecular evolutionary genetics analysis version 11. *Mol. Bio. Evol.* 38, 3022–3027. doi: 10.1093/molbev/msab120
- Tholl, D. (2006). Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. *Curr. Opin. Plant Biol.* 9, 297–304. doi: 10.1016/j.pbi.2006.03.014
- Van Santen, J. A., Jacob, G., Singh, A. L., Aniebok, V., Balunas, M. J., Bunsco, D., et al. (2019). The natural products atlas: An open access knowledge base for microbial natural products discovery. *ACS Cent. Sci.* 5, 1824–1833. doi: 10.1021/acscentsci.9b00806
- Waditee-Sirisattha, R., and Kageyama, H. (2022). "Extremophilic cyanobacteria," in *Cyanobacterial physiology*, eds H. Kageyama and R. Waditee-Sirisattha (Cambridge, MA: Academic Press), 85–99. doi: 10.1016/B978-0-323-96106-6.00012-5
- Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., et al. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS one* 9:p.e112963. doi: 10.1371/journal.pone.0112963
- Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., et al. (2016). Sharing and community curation of mass spectrometry data with global natural products social molecular networking. *Nat. Biotechnol.* 34, 828–837. doi: 10.1038/nbt.3597
- Weiss, M. B., Borges, R. M., Sullivan, P., Domingues, J. P., da Silva, F. H., Trindade, V. G., et al. (2025). Chemical diversity of cyanobacterial natural products. *Nat. Prod. Rep.* 42, 6–49. doi: 10.1039/D4NP00040D
- Weiss, M. B., Médice, R. V., Jacinavicius, F. R., Pinto, E., and Crnkovic, C. M. (2023). "Metabolomics applied to cyanobacterial toxins and natural products," in *Microbial natural products chemistry: A metabolomics approach*, ed. T. P. Fill (Berlin: Springer Nature), 21–49. doi: 10.1007/978-3-031-41741-2_2
- Winter, J. M., Behnken, S., and Hertweck, C. (2011). Genomics-inspired discovery of natural products. *Curr. Opin. Chem. Biol.* 15, 22–31. doi: 10.1016/j.cbpa.2010.10.020
- Wood, D. E., Lu, J., and Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genom. Biol.* 20:257. doi: 10.1186/s13059-019-1891-0
- Xu, L., Dong, Z., Fang, L., Luo, Y., Wei, Z., Guo, H., et al. (2019). OrthoVenn2: A web server for whole-genome comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Res.* 47, W52–W58. doi: 10.1093/nar/gkz333
- Yabuzaki, J. (2017). Carotenoids database: Structures, chemical fingerprints and distribution among organisms. *Database* 2017:bax004. doi: 10.1093/database/bax004