



Genomic insights into adaptations to survival and toxicity of cyanobacteria in hot springs: a study comprising cyanobacterial genomes from Europe, Iceland, and Central Asia

Nataliia Khomutovska^{1,2}, Agnieszka Rudak², Spyros Gkelis³, Mikołaj Kokociński⁴ and Iwona Jasser²

¹ Department of Plant Protection Biology, Swedish University of Agricultural Sciences, Lomma, Sweden

² Faculty of Biology, University of Warsaw, Warsaw, Poland

³ School of Biology, Aristotle University of Thessaloniki, Thessaloniki, Greece

⁴ Faculty of Biology, Adam Mickiewicz University of Poznan, Poznań, Poland

ABSTRACT

Aim of the study. Genomic resources for thermophilic cyanobacteria remain limited, hindering understanding of their ecology and adaptations to extreme high-temperature environments. This study characterizes the genomic features of thermophilic cyanobacteria from thermal springs, focusing on their strategies to cope with energy limitations and the genetic basis of secondary metabolite production.

Material and methods. We analyzed six newly assembled genomes, and compared them with reference genomes from thermophilic and mesophilic cyanobacteria. Functional gene analysis, genome size comparisons, and biosynthetic gene cluster (BGC) mining were performed to assess adaptations.

Results. Genome sizes varied substantially, with *Hillbrichtia pamiria* showing a notably reduced genome, whereas *Calothrix thermalis* and *Thermoleptolyngbya hindakii* were similar to or larger than related mesophiles. Secondary metabolite BGCs were generally reduced in thermophiles, especially in *H. pamiria*, although larger genomes retained diverse BGCs, including cyanotoxins, nonribosomal peptide synthetases (NRPS), ribosomally synthesized and post-translationally modified peptides (RiPPs) and terpenes. Core functions, the Calvin cycle and nitrogen fixation, were conserved, while lineage-specific variation occurred in Tricarboxylic Acid (TCA) cycle, sulfur assimilation, and denitrification genes. *Amphirytois necridicus* harbors features suggesting adaptation to low-energy environments, like alternative TCA cycle enzymes.

Conclusions and relevance. These findings expand the genomic and functional landscape of thermophilic cyanobacteria, reveal lineage-specific adaptations, and underscore the value of learning from these natural strategies for potential biotechnological applications.

Submitted 24 December 2025

Accepted 2 April 2026

Published 3 June 2026

Corresponding author

Agnieszka Rudak,
a.rudak@biol.uw.edu.pl

Academic editor

Mikhail Gelfand

Additional Information and
Declarations can be found on
page 18

DOI 10.7717/peerj.21284

© Copyright

2026 Khomutovska et al.

Distributed under

Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Genomics, Microbiology

Keywords Thermophilic cyanobacteria, Genome sequencing, Extreme environments, Thermal springs

INTRODUCTION

Thermal springs are extreme environments where only highly specialized microorganisms can persist. Among these, thermophilic cyanobacteria stand out as true extremophiles, having evolved unique molecular and genomic adaptations that enable survival and growth at high temperatures conditions that exclude most other life forms. These genomes provide a valuable window into the evolutionary processes that shape life under thermal stress, revealing how extremophiles balance metabolic efficiency, stress resilience, and genomic streamlining. Studying these organisms enhances our understanding of the molecular mechanisms underlying thermotolerance and offers opportunities to harness their unique biochemical and metabolic capacities for biotechnological applications (Ward, Castenholz & Miller, 2012; Kees et al., 2022; Klatt et al., 2011). Adaptation to high temperatures in thermophilic cyanobacteria involves the coordinated action of multiple molecular strategies, including protein stabilization by molecular chaperones, membrane remodeling for thermal resilience, robust DNA repair systems, and enzymatic machinery optimized for heat stress. Together, these adaptations illustrate how extremophilic genomes encode both the flexibility and robustness necessary to thrive under extreme conditions, providing a model for understanding evolution in such environments and for discovering novel biomolecules with industrial and pharmaceutical relevance. Metagenomic analyses of microbial mats have revealed extensive population diversity among thermophilic lineages, as well as unique genomic traits that underpin survival in extreme heat (Pedersen & Miller, 2017; Kees et al., 2022).

Molecular mechanisms of thermal adaptation

Based on previous studies, thermophilic cyanobacteria maintain cellular function under heat stress through coordinated programmes targeting protein stability, genome integrity, and membrane resilience. Protein denaturation represents a major challenge at elevated temperatures, which is counteracted by the expression of heat shock proteins (Hsp20, Hsp60/GroEL, Hsp70/DnaK) and co-chaperones (DnaJ, GrpE) that prevent aggregation and assist in protein refolding (Wang, Cen & Zhao, 2015; Suzuki, Simon & Slabas, 2006). Membrane integrity is preserved through genes such as *fatB* and desaturases (*des*), which adjust fatty acid composition, and by sulfoquinovosyl diacylglycerol synthesis mediated by *sqdX* and *sqdB* (Sato et al., 2017; Eungrasamee et al., 2024). DNA repair systems—including UvrABC, MutSL, RecA, and Dps proteins—maintain genomic stability under heat stress. Thermophiles also rely on structurally stable metabolic enzymes (*gdhA*, *pyk*, *cox*) and antioxidant systems (SOD, KatG, Trx, Prx) to sustain biochemical reactions at high temperatures (Vieira Vaz et al., 2020; Dietz, 2011). Photosynthetic efficiency is maintained via enhanced carbon-concentrating mechanisms (CCM), reflecting the co-evolution of energy capture and thermal adaptation (Brito et al., 2017).

Genomic markers of thermophily

Based on comparative genomic analyses, several gene families and regulatory systems characterise thermophilic cyanobacteria. These include regulatory genes, including sigma (σ) factors, which are proteins that bind RNA polymerase and direct it to specific promoters

to control transcription under heat stress (Wang, Cen & Zhao, 2015), thermostable C-phyocyanins (Kees et al., 2022), and gene patterns from metagenome-assembled genomes (MAGs) that differentiate thermophilic lineages (Pedersen & Miller, 2017). Stress-response genes encoding chaperones and transcriptional regulators further define strategies for heat acclimation (Panyakampol et al., 2015).

Cyanotoxin production in extreme environments

Based on field studies, thermophilic cyanobacteria are capable of producing hepatotoxins (e.g., microcystins) and neurotoxins (e.g., anatoxin, BMAA, DAB, AEG isomers) in hot springs and geothermal lakes (Cirés, Casero & Quesada, 2017; Jablonska et al., 2024). Cyanotoxins in these extreme environments pose ecological risks, including documented wildlife mortality events, such as deaths of Lesser Flamingos in Lake Bogoria, highlighting the environmental and health relevance of thermophilic cyanobacteria.

Study rationale and objectives

Recent genomic research continues to expand understanding of thermophilic cyanobacterial ecology and adaptation (Jasser et al., 2022). However, genomic resources remain limited for many thermophilic genera, restricting detailed evolutionary and functional analyses. By sequencing and analyzing high-quality genomes from multiple thermophilic habitats across Europe, Iceland, and Central Asia, including previously uncharacterized lineages within the Oscillatoriaceae and FACHB-T130 families, this study provides a foundation for understanding the molecular and evolutionary basis of thermotolerance. These genomes enable detailed comparative analyses of metabolic pathways, regulatory networks and stress-response mechanisms, offering insights into how cyanobacteria adapt to extreme thermal environments and providing resources for future ecological, evolutionary and biotechnological studies. To address this gap, the present study aims to characterize the genomic potential of thermophilic cyanobacteria. Specifically, we pursue:

1. Genome mining to identify genes and pathways essential for thermophilic life.
2. Comparative genomic analysis with related mesophilic and thermophilic strains to identify conserved *versus* unique traits.
3. Exploration of biosynthetic gene clusters (BGCs) to assess potential secondary metabolite production with ecological and biotechnological relevance.

MATERIALS AND METHODS

Strains used for genome sequencing

The selected strains originate from locations shown in Fig. 1, with diverse thermal environments, including thermal springs in Greece (*Amphirytyos necridicus*), geothermal ponds in Poland (*Thermoleptolyngbya hindakiae*), hot spring habitats in Tajikistan (*Hillbrichtia pamiria*, Cylx 15.1), and from geothermal baths in Iceland (*Calothrix thermalis*). The strains *A. necridicus*, *T. hindakiae*, *H. pamiria*, and *C. thermalis* were previously isolated, taxonomically described, and characterised with respect to thermotolerance in our earlier study (Jasser et al., 2022). *Anabaena lutea* and *Nostoc*

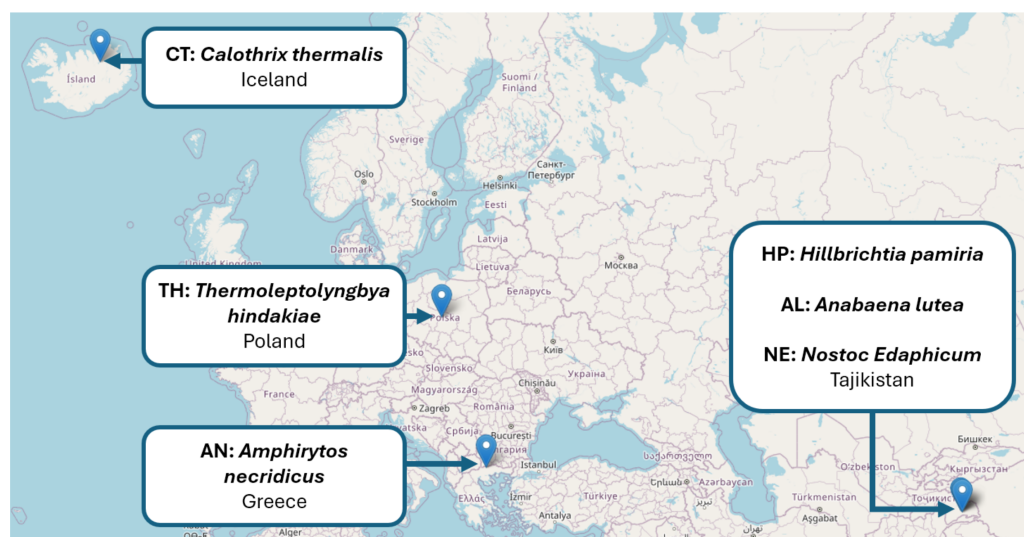


Figure 1 Geographic origin, taxonomic identity and abbreviations of thermophilic cyanobacterial strains which genomes were sequenced and analysed in this study. The analysed strains originate from diverse geothermal environments, including thermal springs in Greece (*Amphiryos necridicus*), geothermal ponds in Poland (*Thermoleptolyngbya hindakiae*), hot springs in Tajikistan (*Hillbrichtia pamiria*, Cylx 15.1), and geothermal baths in Iceland (*Calothrix thermalis*). *A. necridicus*, *T. hindakiae*, *H. pamiria*, and *C. thermalis* were previously isolated, taxonomically described, and characterized with respect to thermo-tolerance (Jasser et al., 2022). *Anabaena lutea* and *Nostoc edaphicum* were previously reported as members of microbial mats E08 and E11, respectively (Khomutovska et al., 2020). Map data ©OpenStreetMap contributors (CC BY-SA 2.0).

Full-size DOI: 10.7717/peerj.21284/fig-1

edaphicum were not analysed previously but described as members of microbial mats E08 (*A. lutea*) and E11 (*N. edaphicum*) by Khomutovska et al. (2020). Sampling and isolation steps are presented in Fig. S1.

DNA extraction, genome sequencing, assembly, and annotation

Genomic DNA extraction and sequencing

Genomic DNA was extracted from fresh cyanobacterial biomass using the E.Z.N.A.[®] Soil DNA Kit (Omega Bio-Tek, Norcross, GA, USA), following the manufacturer's instructions. Sequencing was performed exclusively on DNA from pure isolates; no metagenomic datasets or environmental metagenome-assembled genomes (MAGs) were included in this study. Library preparation and whole-genome sequencing were performed commercially by Eurofins Genomics using the Illumina NovaSeq 6000 platform with a 2 × 150 bp paired-end configuration, targeting a minimum of 100 × average coverage for each genome (up to ~10 Mb). Raw sequencing data (FASTQ files) are available via ENA under project PRJEB93759. In addition, assembled genomes for each species, along with annotation outputs generated in KBase, have been deposited in Zenodo (doi: 10.5281/zenodo.17995636), ensuring full reproducibility of downstream analyses: <https://narrative.kbase.us/narrative/222168>.

Genome assembly and quality control

A standardized genome mining protocol for cyanobacteria was applied (Shaffer *et al.*, 2020; Aziz *et al.*, 2008), ensuring reproducible and comprehensive identification of genomic features. Raw reads were processed using Trimmomatic v0.39 to remove adapters and low-quality sequences, and sequence quality was assessed with FastQC v0.11.9 (Andrews, 2010). Contaminant removal was performed at two levels. First, read-level filtering was conducted using Kraken2 v2.1.2 (Wood, Lu & Langmead, 2019), conservatively applied to remove only reads assigned to well-characterized non-cyanobacterial contaminants (e.g., human or other common laboratory contaminants), thereby minimizing the risk of discarding sequences from poorly represented cyanobacterial lineages. Second, post-assembly filtering was applied, where assembled contigs were screened using Kraken2 and ZEUS v1.0 (Alvarenga, Fiore & Varani, 2017) to confirm cyanobacterial origin and remove any residual contamination. Genome assembly was performed using SPAdes v3.15.4 (Prjibelski *et al.*, 2020), followed by scaffolding and gap closure with Platanus v1.2.4 (Kajitani *et al.*, 2019). Assembly quality was evaluated using Assemblathon v1.0 (Bradnam *et al.*, 2013) and QUAST v5.0.2 (Mikheenko *et al.*, 2018). These steps, combined with contaminant filtering and assembly polishing, resulted in high-quality genomes suitable for downstream comparative and functional analyses, although a small number of assemblies with lower completeness or higher contamination were retained and interpreted with appropriate caution.

Functional annotation was carried out with DRAM v1.2.4 (Shaffer *et al.*, 2020) and RAST v2.0 (Aziz *et al.*, 2008), enabling systematic identification of genes involved in primary metabolism, stress adaptation, and secondary metabolite biosynthesis. This genome mining approach allowed us to detect both conserved and lineage-specific genomic traits.

Comparative genomics and phylogenetic analysis

For comparative purposes, additional reference and published genomes were included, selected based on phylogenetic proximity to our isolates, as well as representative thermophilic and mesophilic cyanobacteria. Full list of used genomes is in Table 1. Prior to selection, a comprehensive survey of the literature and public genomic databases was conducted to identify all available thermophilic cyanobacterial genomes, and all relevant genomes available at the time of analysis were included. This allowed robust assessment of conserved *versus* unique genomic features and metabolic capacities. The whole workflow summary is presented in Fig. 2.

To evaluate genome similarity and redundancy, dRep v2.6.2 (Olm *et al.*, 2017) was used to enable accurate genomic comparisons and de-replication. Phylogenetic relationships were inferred using GTDB-Tk v2.12.10 (Chaumeil *et al.*, 2022) and validated with the “Insert Set of Genomes Into SpeciesTree” pipeline implemented in FastTree v2.1.10 (Price, Dehal & Arkin, 2010), with reference genomes from GTDB (Parks *et al.*, 2020).

All analyses were conducted within the KBase platform, using integrated modules for read preprocessing, genome assembly (SPAdes), genome annotation (RAST), and comparative genomics workflows. The KBase narrative provides full reproducibility, including parameter settings and intermediate outputs for each step.

Table 1 List of additional reference genomes included in our study as comparisons to our assembled genomes.

Reason for inclusion	Organism	GenBank access number	Thermophilic	Source publication
Closest relatives to <i>Hillbrichtia pamiria</i> according to phylogenetic tree (Fig. S3)	<i>Coleofasciculus</i> sp.	GCF_014698845.1	no	Chen et al. (2021)
	<i>Cyanosarcina radialis</i>	GCA_019359615.1	no	Ward et al. (2021)
	<i>Rippkaea orientalis</i> / <i>Synechococcus</i> sp.	GCF_000021805.1	no	Cornet et al. (2018)
	<i>Coleofasciculus chthonoplastes</i>	GCA_000155555.1	no	Tang et al. (2024)
Thermophilic species	<i>Thermosynechococcus vestitus</i>	GCA_003555505.2	yes	Liang et al. (2018)
	<i>Moorella thermoacetica</i> ATCC 39073	GCF_000013105.1	yes	Redl et al. (2020)
	<i>Leptodesmis sichuanensis</i>	GCA_021379005.1	yes	Tang et al. (2024)
	<i>Leptolyngbya</i> cf. <i>ectocarp</i> i	GCA_015207065.1	no	Tang et al. (2024)
Non-thermofilic species as a reference group	<i>Leptolyngbya boryana</i>	GCF_002142475.1	no	Hiraide et al. (2014)
	<i>Nostoc punctiforme</i>	GCA_000020025.1	no	Cornet et al. (2018)
	<i>Nostoc</i> sp. PCC 7107	GCA_000316625.1	no	Shih et al. (2013)
	<i>Anabaena</i> sp. PCC 7108	GCA_000332135.1	no	Shih et al. (2013)
	<i>Calothrix</i> sp. PCC 6303	GCF_000317435.1	no	Shih et al. (2013)
	<i>Calothrix</i> sp. PCC 7103	GCF_000331305.1	no	Shih et al. (2013)

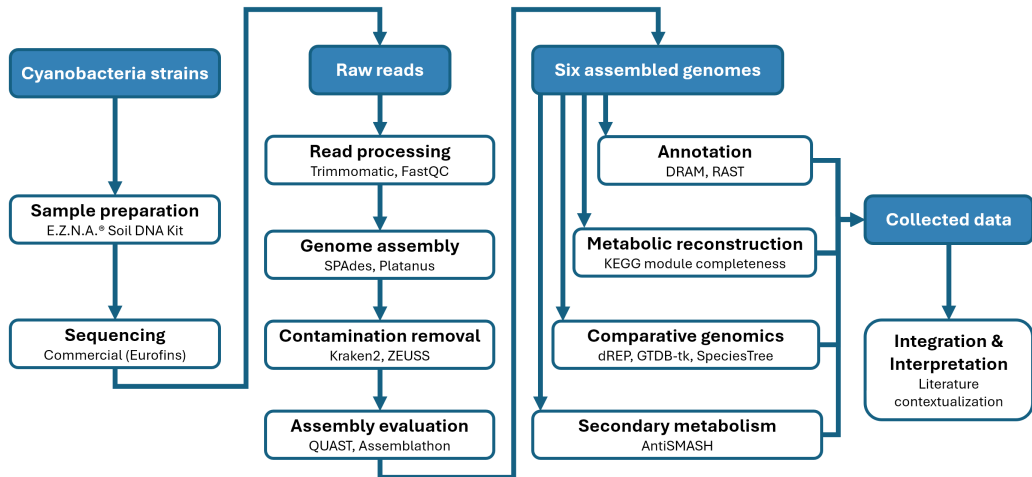


Figure 2 Overview of the genomic analysis workflow and bioinformatic tools used in this study. Workflow summarizing genome assembly, quality assessment, annotation, and comparative genomic analyses. Prior to comparative analysis, a comprehensive survey of the literature and public genomic databases was conducted to identify all available thermophilic cyanobacterial genomes; all relevant genomes were included to enable robust comparison of conserved and unique genomic and metabolic features.

[Full-size](#) DOI: 10.7717/peerj.21284/fig-2

RESULTS

Assembled genomes

We assembled six genomes of isolated strains with a completeness rate of over 87%. Short contigs with a length of <1,000 bp were removed before the analysis. Assembly statistics are summarized in Table 2, including N50 (73,141–357,711 bp), L50 (7–32), number of

contigs (41–385), and largest contig length (297,150–635,415 bp), indicating variable but overall moderate-to-high assembly continuity among the genomes. Detailed assembly and contig statistics are additionally presented in Figs. S2 and S3. Table 2 and Table S1 provide genome codes used throughout the manuscript, together with basic genome characteristics and initial gene presence screening. Genes *nif* D, *nif* K, *nif* H, and Rubisco S and L have been found in all newly assembled genomes. *A. necridicus* and *H. pamiria* genomes had the highest percentage of GC pairs, having 50% and 47%, respectively, with *H. pamiria* also being the smallest genome among our assemblies, having 4.7 mln bp. Other smaller genomes included *A. lutea* with 5.7mln and *T. hindakiae* and *A. necridicus* with 6.6mln bp. Despite its small genome size, *A. lutea* had the lowest GC percentage. Cytochrome oxidase genes have been found in all genomes, while quinol oxidase is also found in *C. thermalis*, *H. pamiria*, *N. edaphicum*, and *T. hindakiae*. All species differed in the presence of sulphur metabolism genes, and nitrate reductase genes *nir*B and *nir*D were found only in two genomes (*A. lutea* and *T. hindakiae*).

Phylogenomics

Included additional reference genomes for comparative purposes were selected based on quantitative measures of phylogenetic proximity to our isolates, including average nucleotide identity (ANI) and alignment fraction, as well as representation of thermophilic and mesophilic cyanobacteria. For example, *Amphirytos necridicus* (AN) showed 96.17% ANI and 0.824 alignment fraction with its closest GTDB-Tk placement (*Elainella sp000733415*), while *Calothrix thermalis* (CT) shared 94.35% ANI and 0.904 alignment fraction with *Calothrix sp015272395*. Table 3 contains details of all newly assembled genomes.

AN, HP, and TH are all strains of a newly described genera and/or species (Jasser et al., 2022); therefore, their taxonomic assignment could not be performed fully with existing databases. *Hillbrichtia* appears to be the closest relative to members of the *Coleofasciculus* genus; however, not with a high certainty (Fig. S5). *Amphirytos* genus seems to be closely related to *Elainella* (Table 3). TH is a newly described species of *Thermoleptolyngbya hindakiae*, which belongs to Oscillatoriaceae family and is grouped with *Oscillatoria* genus. CT assignment confirmed our initial assessment of this genome to the species *Calothrix thermalis*. AL and NE genomes have been assigned to the family Nostocaceae (Fig. S6). AL appears to be assigned to *Anabaena* genus and is closely related to *A. lutea*, while NE is assigned to *Nostoc*, exhibiting high similarity to *N. edaphicum* (Table 3). Full phylogenetic trees are shown in Figs. S4 and S6. The genomes of *Amphirytos necridicus* (AN) and *Calothrix thermalis* (CT) are draft-quality, with 87.93% completeness and 16.8% contamination (AN) and 29.94% contamination (CT). These genomes were included in analyses for completeness of representation but interpreted cautiously.

Figure S4 shows a phylogenomic tree compatible with NCBI taxonomy (based on <http://algaebase.org> (Guiry & Guiry, 2025)), illustrating the phylogenetic relationships among the cyanobacterial genomes included in this study. Genomes highlighted in blue correspond to the isolates sequenced in this study, as well as thermophilic reference genomes included for comparative purposes. This visualization confirms the placement of

Table 2 Basic statistics of assembled genomes with completeness and contamination percentage, length of genomes, contigs and presence of Oxygen, Sulfur, and Nitrogen metabolism genes.

Genome code	CheckM completeness (%)	CheckM contamination (%)	Genome size (bp)	N50 (bp)	L50 (bp)	Number of contigs	LargestContig (bp)	G+C (%)	BioE: Oxygen	BioE: Sulfur	BioE: Nit_Ox_Red
AN	87.93	16.80	6,669,676	135,684	17	250	297,150	50.70	coxA, coxB	aprA, sdo	nirB, nirD
CT	96.55	29.94	9,123,886	241,165	12	385	573,875	41.30	coxA, coxB, cydA, cydB	cysC, sat, sqr, sdo	–
HP	99.33	0.67	4,788,368	357,711	7	41	635,415	47.60	coxA, coxB, cydA, cydB	cysC, sat, sqr, sdo	–
TH	90.98	4.80	6,634,043	169,088	10	83	455,092	43.30	coxA, coxB, cydA, cydB	cysC, sat, sqr	nirB, nirD
AL	99.67	0.56	5,763,363	184,907	10	59	426,492	39.30	coxA, coxB	sat	–
NE	99.78	0.44	8,207,313	73,141	32	246	330,772	41.50	coxA, coxB, cydA, cydB	cysC, sat, sqr, sdo	–

Table 3 Taxonomic assessment of assembled genomes according to GTDB-Tk. All genomes were assigned as d__Bacteria;p__Cyanobacteriota;c__Cyanobacteriia. According to traditional taxonomy (e.g., classification used within NCBI), these genomes belong to the phylum Cyanobacteria and class Cyanophyceae.

Genome	Taxon	Closest placement taxonomy	Closest placement ANI	Closest placement alignment fraction	Taxonomic classification method	RED value
AN	<i>Amphirytos necridicus</i>	o__Elainellales;f__Elainellaceae; g__Elainella;s__Elainella sp000733415	96.17	0.824	ANI screen	–
CT	<i>Calothrix thermalis</i>	o__Cyanobacteriales;f__Nostocaceae; g__Calothrix;s__Calothrix sp015272395	94.35	0.904	topology and ANI	0.98999
HP	<i>Hillbrichtia pamiria</i>	o__Cyanobacteriales;f__FACHB-T130; g__s__	–	–	topology	0.84051
TH	<i>Thermoleptolyngbya hindakiae</i>	o__Cyanobacteriales;f__Oscillatoriaceae; g__s__	–	–	novelty determined using RED	0.85017
AL	<i>Anabaena lutea</i>	o__Cyanobacteriales;f__Nostocaceae; g__Anabaena;s__Anabaena lutea	94.37	0.834	topology and ANI	0.9955
NE	<i>Nostoc edaphicum</i>	o__Cyanobacteriales;f__Nostocaceae; g__Nostoc;s__Nostoc edaphicum_A	94.94	0.779	topology and ANI	0.99665

our sequenced strains within their respective phylogenetic clades and allows comparison with representative thermophilic cyanobacteria.

Functional genes analysis

CheckM analysis shows that most genomes are highly complete (>96%) with low contamination (<1%). The exceptions are *Amphiryos necridicus* (~88% completeness, 16.8% contamination) and *Calothrix thermalis* (~96.6% completeness, 29.9% contamination). Reduced completeness and higher contamination in these genomes may partly explain some of the observed differences in metabolic modules. All genomes are presented with full quality metrics to allow readers to interpret functional annotations appropriately in comparative analyses.

Figure 3 presents a heatmap summarizing DRAM-based functional annotations across the studied cyanobacterial genomes. Each row represents a genome, and columns correspond to metabolic modules or pathways, including carbon fixation, central carbon metabolism, respiratory complexes, nitrogen and sulfur metabolism, photosynthesis, and carbohydrate-active enzymes (CAZymes). Values indicate the completeness or presence of each pathway/module, allowing a visual comparison of conserved *versus* variable functions among the strains. Only thermophilic strains are included in the heatmap for direct functional comparison.

All studied genomes conserve the Calvin-Benson cycle (reductive pentose phosphate pathway) with near-complete completeness (~1.0), reaffirming its central role in CO₂ fixation across cyanobacteria (Fig. 3). However, key differences emerge in other carbon metabolic routes. The tricarboxylic acid (TCA) cycle, vital for energy generation and biosynthesis, is most complete in *L. boryana*, indicating strong aerobic respiration capacity. In contrast, *A. necridicus* and *Calothrix* sp. PCC 6303 shows reduced TCA completeness (~0.63–0.75), implying alternative metabolic strategies potentially linked to thermophilic adaptation. Glycolysis *via* the Embden-Meyerhof-Parnas pathway is well developed in *Leptodesmis sichuanensis*, accompanied by an active pentose phosphate pathway, which supports biosynthesis through NADPH production.

Alternative carbon fixation pathways show varied representation: the reductive citrate (reverse TCA) cycle is prominent in *Anabaena* sp. PCC 7108 and *Calothrix* sp. PCC 7103 (~0.6–0.7) (Fig. 3), enabling utilization of diverse carbon sources, while *A. necridicus* lacks significant activity here but shows traces of the hydroxypropionate-hydroxybutylate cycle, suggesting limited alternative CO₂ fixation. The Wood-Ljungdahl pathway, an energy-efficient autotrophic route, is notably enriched in *L. boryana* and *L. sichuanensis*.

Photosystems I and II are universally conserved and nearly complete across most genomes, especially in *L. boryana*, *Anabaena* sp. PCC 7108, and *Calothrix* sp. PCC 7103 (Fig. 3), reinforcing their phototrophic lifestyle. Conversely, *A. necridicus* exhibits a reduced presence of photosystem components, indicative of a possible mixotrophic or heterotrophic strategy. The respiratory electron transport chain is well-developed in *L. sichuanensis* and *L. boryana*, with robust complexes I and III, facilitating efficient ATP production under variable environmental conditions.

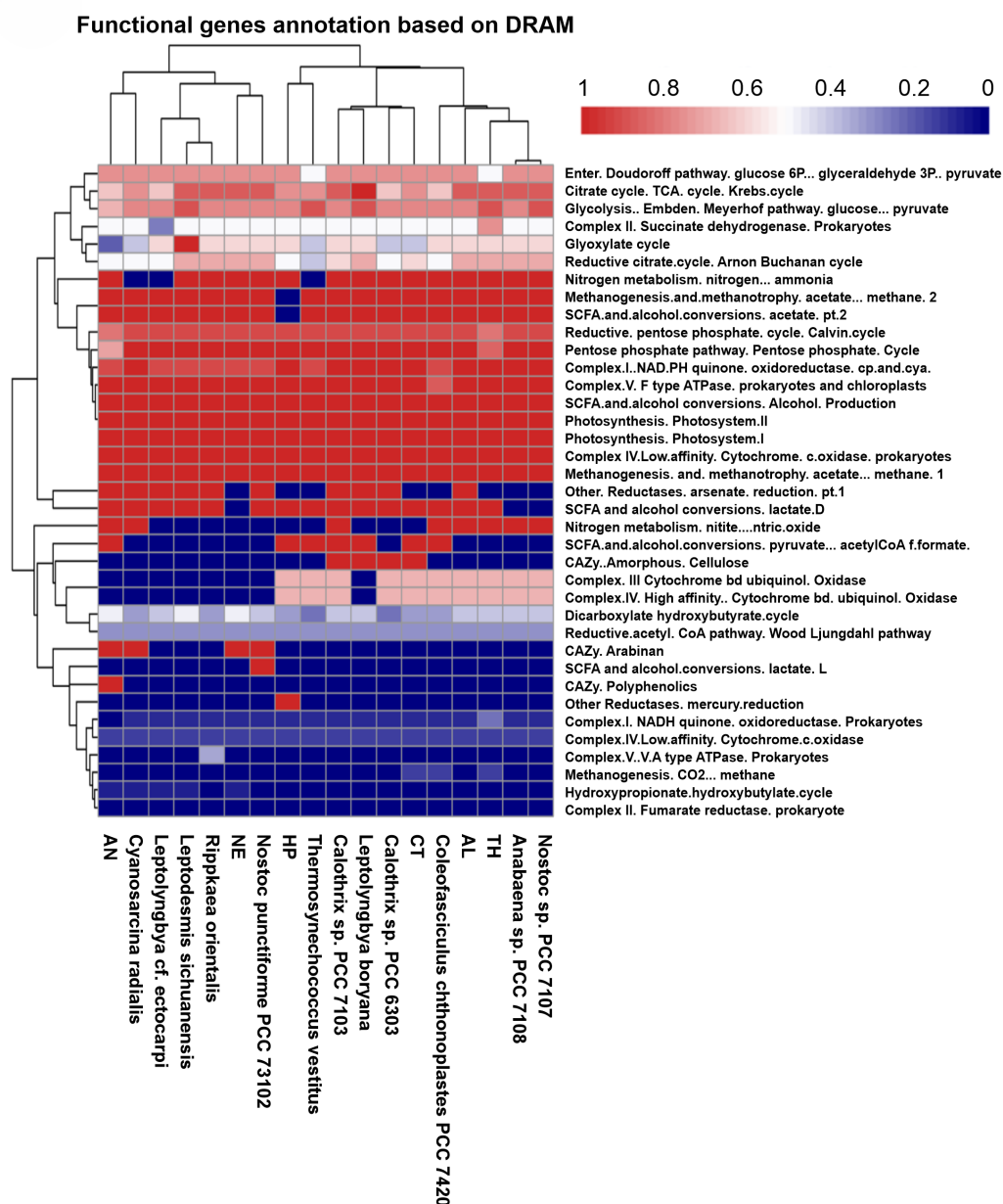


Figure 3 Heatmap showing DRAM-based metabolic annotation and pathway completeness across analysed cyanobacterial genomes. Completeness scores of major metabolic pathways based on DRAM annotation. All genomes show high completeness of the Calvin–Benson cycle, confirming conserved CO₂ fixation capacity. Variability is observed across other pathways, including the tricarboxylic acid (TCA) cycle, glycolysis, the pentose phosphate pathway, and alternative carbon-fixation pathways, indicating differences in metabolic strategies and potential adaptations to thermophilic environments.

[Full-size DOI: 10.7717/peerj.21284/fig-3](https://doi.org/10.7717/peerj.21284/fig-3)

Nitrogen fixation genes are conserved in *Anabaena* sp. PCC 7108 and *Calothrix* sp. PCC 7103 (Fig. 3), supporting their key ecological role in nitrogen cycling. In contrast, *A. necridicus* shows minimal nitrogen fixation potential, possibly relying on external nitrogen

sources. Sulfur metabolism genes for dissimilatory sulfate reduction are uniquely present in *A. necridicus*, suggesting adaptation to sulfur-rich environments and distinguishing it metabolically from other species.

Carbohydrate-active enzymes (CAZymes), essential for polysaccharide degradation, are abundant in *L. sichuanensis* and *L. boryana*, enabling efficient breakdown of cellulose, starch, and pectin, which may contribute to ecological versatility and biotechnological applications. *A. necridicus*, however, lacks key enzymes such as chitinase and xylanase, limiting its polysaccharide degradation capacity. Furthermore, short-chain fatty acid metabolism, including acetate and butyrate fermentation, is well-developed in *L. sichuanensis* and *Calothrix* sp. PCC 7103, supporting survival in fluctuating oxygen conditions (Fig. 3).

To better understand the metabolic potential of the cyanobacterial strains, genome-scale metabolic models (GEMs) were constructed in KBase for six representative genomes: *A. necridicus*, *C. thermalis*, *N. edaphicum*, *A. lutea*, *T. hindakiae*, and *H. pamiria*. The models vary in size, with *C. thermalis*, belonging to Nostocales, showing the most extensive metabolic network (1,141 reactions and 1,195 compounds), indicating broader biosynthetic and catabolic versatility. NE and AL, also belonging to Nostocales, have moderately large models (559 and 570 reactions, respectively), consistent with the typically more complex life strategies of this group, including nitrogen fixation and developmental plasticity.

In contrast, *A. necridicus* and *T. hindakiae*, affiliated with Synechococcales and Oscillatoriales, respectively, have smaller models (~548–549 reactions), reflecting potentially more specialized or streamlined metabolic systems adapted to thermophilic environments. All models share a common structure with two cellular compartments and one biomass objective function. Only minor gap-filling (two reactions per model) was needed to restore model completeness. Full raw data and sequences are deposited under ENA project number PRJEB93759 (Table S2) and DOI: 10.5281/zenodo.17995635.

Biosynthetic and stress-responsive genes

To investigate the biosynthetic potential and stress adaptation mechanisms of thermophilic cyanobacteria, we performed comprehensive gene screening using the antiSMASH bacterial version to identify secondary metabolite biosynthetic gene clusters, and the MultiGene BLAST application within KBase to screen for functional genes associated with DNA repair, stress tolerance, and metabolic adaptation. The BLAST queries specifically targeted genes involved in DNA repair pathways (e.g., *uvrA*, *uvrB*, *recA*, *mutS*), oxidative stress responses (e.g., *sod*, *katG*, *trx*, *prx*), and thermal stress adaptations such as chaperone systems (*dnaK*, *groEL*, *hsp20*) and membrane stability (e.g., *fatB*, *des*, *sqdX*, *sqdB*). However, the identification of these functional genes was limited by the composition of the reference database. Inspection of best matches showed that most homologous sequences corresponded to non-cyanobacterial thermotolerant organisms, whereas thermophilic or thermotolerant cyanobacteria represented only a small fraction (a few percent) of hits. This highlights the strong taxonomic bias in current genomic resources and emphasizes the need for additional cyanobacterial genome assemblies to better evaluate thermotolerance and stress-response capabilities.

A detailed list of secondary metabolite genes found is in [Table S3](#). In the genomes of NE, AL, CT, and TH, we identified more gene clusters involved in secondary metabolite production than in AN and HP. This aligns with genome size, except for AL, which has many secondary metabolite genes and a considerably smaller genome than those of other closely related taxa. We found genes encoding for terpenes and nonribosomal peptide synthetases (NRPS)-like metabolites in all genomes. Heterocyte glycolipids have been found in both Nostocaceae genomes and in AN.

The BLASTn (KBase) analysis identified the Excinuclease ABC subunit C, a key DNA repair enzyme, in *Calothrix thermalis* (CT) with 75.8% identity over 99.6% alignment coverage (1,897 bp), a highly significant *E*-value of 0.0, and a bit score of 931. In *Nostoc edaphicum* (NE), the same gene was detected with 80.7% identity, but lower alignment coverage (51.0%), with an *E*-value of 0.0 and a bit score of 750. No significant hits were obtained in other genomes, suggesting limited conservation or high sequence divergence across the thermophilic strains. Similarly, ATP synthase alpha chain (EC 3.6.3.14) was found only in NE, showing 81.8% identity over 98.6% alignment coverage (1,499 bp). The presence of mismatches and gaps further indicates potential genetic divergence, limiting direct comparisons among the thermophilic isolates. Overall, BLAST-based screening presented in [Table 4](#), for stress-response and cyanotoxin-related genes is constrained by the scarcity of thermophilic cyanobacterial reference genomes. While Excinuclease ABC subunit C was detected in two genomes, its absence in most other strains prevents strong conclusions regarding conservation. Therefore, we interpret these findings cautiously, highlighting only that the gene is present in some thermophilic isolates rather than asserting a broadly conserved role.

DISCUSSION

Comparative genomic analysis of six thermophilic cyanobacterial genomes revealed key adaptive features and metabolic diversity. Most genomes showed completeness levels suitable for broad comparative analyses. For genomes initially below the completeness threshold, quality was improved using standardized bioinformatic procedures, including contaminant removal, scaffolding and gap closure, and assembly polishing, as described in the Methods. To contextualize these genomes, we compared them with closely related reference strains identified through ANI and phylogenomic placement. Several grouped with known taxa such as *Elainella* spp., *Calothrix* spp., *Anabaena* spp., and *Nostoc* spp., while others (TH and HP) appear to represent previously uncharacterized lineages within Oscillatoriaceae and an unclassified FACHB-T130-associated family.

Genome size, GC content, and general features

Genome sizes ranged from 4.7 Mbp (HP) to 9.1 Mbp (CT), reflecting variation in genome complexity. Larger genomes, such as those of CT and NE (>8.5 Mbp), are consistent with previously described Nostocaceae genomes, which are known for their expansive gene content and secondary metabolite pathways ([Huo et al., 2021](#); [Moten et al., 2020](#)). GC content varied from ~39% (AL) to ~50% (HP and AN), potentially contributing to thermal genome stability and aligning with ranges reported in thermophilic cyanobacteria ([Tang](#)

Table 4 List of toxin-related and thermophilic-associated genes identified in assembled genomes using BLAST.

Taxon	Gene	Related product	NE	AL	HP	CT	TH	AN	Source publication
<i>Anabaena</i> sp.	<i>mcyA-J</i>	microcystin	*						<i>Dreher et al. (2019)</i>
<i>Phormidium</i>	<i>mcy</i> cluster	microcystin	*						<i>Shishido et al. (2019)</i>
<i>Microcystis aeruginosa</i>	<i>uma1-6</i> cluster	microcystin	*	*					<i>Tillett et al. (2000)</i>
<i>Leptolyngbya</i> sp.	<i>mcyE</i>	microcystin							<i>Saran & Mishra (2015)</i>
<i>Prochlorococcus marinus</i>	<i>psbA</i>	photosystem component		*		*	*		<i>Lindell et al. (2004)</i>
<i>Moorea producents</i>		NRPS/PKS cluster		*		*			<i>Boudreau et al. (2015)</i>
<i>Nostoc</i> sp.		res B-like protein, ABC transporter, nostophycin	*	*					<i>Fewer et al. (2011)</i>
<i>Nostoc punctiforme</i>	<i>aerA</i>	aerolysin							<i>Tikhonova et al. (2017)</i>
<i>Oscillatoria</i> sp.	<i>cyr</i> cluster	cylindrospermopsin	*			*	*		<i>Mazmouz et al. (2010)</i>
<i>Desmonostoc muscorum</i>	<i>nosA-G</i>	nostopeptolide	*	*					<i>Mares et al. (2020)</i>
<i>Desmonostoc muscorum</i>	<i>mcxA-C</i>	muscotoxin	*	*					<i>Mares et al. (2020)</i>
<i>Nostoc flagelliforme</i>		C-phycocyanin beta and alpha subunit		*		*		*	<i>Jiao et al. (2010);</i>
<i>Kamptomena laetevirens</i>									<i>Shakena Fathima et al. (2023)</i>
<i>Fisherella thermalis</i>	<i>uvrA</i>	excinuclease A subunit	*			*			<i>Dagan et al. (2013)</i>

Notes. The “Gene” column reports standard gene names, while the “Related product” column lists the corresponding protein or functional product. In case of whole clusters, the columns are merged by the name of final product. Marks (*) indicate the presence of a gene or cluster in each genome. Source publications are provided for reference.

et al., 2022; Vieira Vaz et al., 2020). Two of the assembled genomes, *Amphiryos necridicus* (AN) and *Calothrix thermalis* (CT), are draft-quality, with 87.93% completeness and 16.8% contamination (AN) and 29.94% contamination (CT). These genomes were included in comparative analyses to provide a broader representation of thermophilic cyanobacterial diversity. However, due to incompleteness and contamination, any pathway absences or variations observed in these genomes may partly reflect assembly gaps rather than true biological differences. Consequently, conclusions regarding metabolic adaptations and lineage-specific features are drawn primarily from the more complete, lower-contamination genomes, while observations from AN and CT should be interpreted with caution. This transparent reporting ensures acknowledging both the potential insights and limitations of these draft genomes.

Respiration and central carbon metabolism

Genes encoding components of the electron transport chain, including cytochrome oxidases (coxA, coxB) and cytochrome bd oxidase (cydA, cydB), were widely distributed. However, AL and AN strains lack cyd genes, indicating potential functional divergence in aerobic respiration (*Winstedt et al., 1998*). Sulfur assimilation genes (cysC, sat, sqr, sdo) showed heterogeneous presence, reflecting adaptation to local sulfur availability (*Yarbrough, 2020*). The Calvin cycle (RuBisCO-mediated carbon fixation) was nearly complete (~1.0) across all genomes, consistent with core carbon fixation pathways conserved in cyanobacteria (*Shi & Falkowski, 2008; Shih et al., 2013*). The tricarboxylic acid (TCA) cycle exhibited lineage-specific variation: it was incomplete in *A. necridicus* (~0.63) but largely complete in *Leptolyngbya* and *Nostoc*-related genomes (0.88–1.0). This partial TCA cycle in some

thermophiles reflects a known prokaryotic adaptation: cyanobacteria lack 2-oxoglutarate dehydrogenase, but the cycle can be functionally completed *via* alternative enzymes (Shih *et al.*, 2013; Alcorta *et al.*, 2020). Such streamlining likely supports energy-efficient respiration under thermal stress while maintaining essential metabolite fluxes. Alternative carbon fixation pathways, including the Wood-Ljungdahl and hydroxybutyrate cycles, were absent or incomplete, in agreement with prior observations that these pathways are rare in cyanobacteria (Ragsdale & Pierce, 2008). DRAM-based pathway completeness revealed that some draft genomes, including AN and CT, show reduced completeness for the TCA cycle (~ 0.63 – 0.75). Because these genomes are incomplete or have moderate contamination, these observations may partly reflect assembly gaps rather than true biological differences. Consequently, interpretations regarding alternative metabolic strategies are presented with caution, focusing on high-quality genomes for robust conclusions.

Nitrogen metabolism

All genomes retained complete nitrogen fixation gene sets (*nifHDK*), supporting $N_2 \rightarrow NH_3$ conversion even under thermophilic conditions. The presence of additional *nif* genes (*nifENB*) was verified where detectable, ensuring coverage of the minimal set proposed for functional nitrogenase activity (Dos Santos *et al.*, 2012). The nitrite-to-nitric oxide module (*nirB/nirD*) exhibited variable presence: it was absent in *Hillbrichtia* and *Thermoleptolyngbya*, suggesting differential denitrification capacities among thermophiles, consistent with freshwater cyanobacteria studies (Meeks & Elhai, 2002; Sivonen & Jones, 1999). Nostocaceae genomes (NE and AL) contained additional genes for heterocyte glycolipid biosynthesis, indicating the presence of functional heterocyte-like structures that facilitate oxygen-tolerant nitrogen fixation (Bauersachs *et al.*, 2013).

Photosynthesis and carbon-concentrating mechanisms

Photosystems I and II were universally present and nearly complete, even in genomes with streamlined metabolic pathways, emphasizing their indispensable role in light-driven energy capture. Some genomes, such as those of *Leptolyngbya* spp., exhibited an expanded repertoire of photosystem II components and carbon-concentrating mechanisms (CCMs), consistent with adaptations to high-light and carbon-rich environments (Brito *et al.*, 2017). Conversely, *A. necridicus* exhibited reduced photosystem complexity alongside genes supporting heterotrophic carbon utilization, reflecting adaptation to low-energy niches (Pierpont *et al.*, 2024; Tang *et al.*, 2024).

Secondary metabolism and biosynthetic gene clusters

BGC abundance is generally correlated with genome size. NE, AL, CT, and TH contained more BGCs, whereas AN and HP exhibited reduced secondary metabolite potential. Notably, AL retained numerous BGCs despite a smaller genome, illustrating genome streamlining with selective retention of biosynthetic genes, a phenomenon documented in cyanobacteria from extreme environments (Weber *et al.*, 2015; Shih *et al.*, 2013; Konstantinou *et al.*, 2021). All genomes contained terpene and NRPS-like pathways, reflecting the widespread biosynthetic potential of cyanobacteria (Kehr, Gatte Picchi & Dittmann, 2011).

Toxin biosynthesis

Microcystin (mcy) clusters were restricted to Nostocaceae genomes, consistent with prior studies (Sivonen & Jones, 1999; Rantala et al., 2006). Cylindrospermopsin (cyr) clusters were detected in *Nostoc edaphicum*, *Calothrix thermalis*, and *Thermoleptolyngbya hindakii*, in line with reports of thermophilic and freshwater cyanobacteria producing this toxin (Mihali et al., 2008). *A. necridicus* and *H. pamiria* lacked toxin genes, correlating with lower secondary metabolite content. Additional ribosomally synthesized and post-translationally modified peptides (RiPPs) and lanthipeptides in TH and AL suggest antimicrobial potential, consistent with the known pharmaceutical relevance of cyanobacterial RiPPs (Arnison et al., 2013). Cyanobactins and microviridins identified in NE and TH highlight further bioactive compound potential (Sivonen et al., 2010) (Table 4).

Some apparent absences in metabolic pathways in AN and CT may reflect incomplete assembly or contamination rather than true biological loss. Therefore, interpretations of metabolic streamlining or alternative strategies are drawn primarily from more complete, low-contamination genomes, whereas draft genomes are included to represent thermophilic diversity.

Importantly, recent population genomics of the thermophilic *Thermosynechococcus taiwanensis* revealed within-species divergence into distinct populations and selective sweeps targeting different sets of genes related to photosynthesis, motility, and ion transport, providing direct evidence of fine-scale adaptation processes in hot spring cyanobacteria. These findings illustrate that even within a single thermophilic species, evolutionary pressures can drive genomic divergence and candidate adaptation at subspecies levels, emphasizing the role of selection in structuring genetic diversity in extreme environments (Chang et al., 2024).

Collectively, these observations demonstrate that thermophilic cyanobacteria balance genomic streamlining with selective retention of adaptive pathways, shaping metabolic, respiratory, and stress-response traits that support survival under thermal, energetic, and ecological constraints. Understanding the mechanisms and evolutionary trajectories that enable cyanobacteria to thrive at high temperatures is crucial for reconstructing how life expands ecological tolerance limits and for uncovering the rules of genome evolution under extreme stress. However, the limited availability of high-quality thermophilic cyanobacterial genomes continues to constrain comprehensive evolutionary and functional analyses. Expanding these genomic resources and dissecting the molecular basis of thermophily will not only advance our understanding of extremophile evolution and ecology but also reveal natural adaptations with tangible biotechnological applications.

These observations are consistent with broader comparative genomic studies showing that late-branching cyanobacterial lineages typically exhibit expanded biosynthetic gene cluster (BGC) repertoires, particularly NRPS and polyketide synthase (PKS) clusters (Shih et al., 2013). For example, Shih et al. (2013) identified over 450 NRPS and PKS gene clusters across 89 cyanobacterial genomes, and Konstantinou et al. (2021) reported numerous BGCs across more than 70 marine Oscillatoriales genomes, underscoring the phylum's biosynthetic potential. NRPS-like and Type I PKS clusters in CT and NE support

the likelihood of bioactive compound synthesis in these strains (Kehr, Gatte Picchi & Dittmann, 2011).

Furthermore, lanthipeptide and ribosomally synthesized and post-translationally modified peptide (RiPP)-like clusters in TH and AL genomes suggest antimicrobial capabilities, which aligns with previous findings on the pharmaceutical relevance of cyanobacterial RiPPs (Arnison et al., 2013). The detection of cyanobactins and microviridins in NE and TH, known for their bioactive properties, further highlights the pharmaceutical potential of these thermophilic strains (Sivonen et al., 2010).

Our results emphasize that, due to the very limited number of publicly available thermophilic cyanobacterial genomes, genome mining and detailed analysis of specific adaptations to extreme environments remain constrained (Shih et al., 2013; Pierpont et al., 2024). Despite these limitations, the present study provides valuable insights into the metabolic and biosynthetic potential of thermophilic cyanobacteria, offering important genomic resources that will facilitate future comparative and functional studies in this underexplored group (Meeks & Elhai, 2002; Weber et al., 2015).

CONCLUSIONS

This comparative genomic study of six thermophilic and mesophilic cyanobacteria reveals substantial diversity in genome size, gene content, and metabolic capabilities that reflect adaptive evolution to extreme thermal environments. Despite this variation, all genomes preserve core functions—including the Calvin cycle and nitrogen fixation—underscoring their essential roles in primary productivity and nitrogen cycling.

Lineage-specific differences in TCA cycle completeness, sulfur assimilation, and denitrification genes illustrate metabolic streamlining consistent with thermophilic strategies for energy conservation and stress resilience. Incomplete TCA cycles in some strains may be compensated by alternative enzymatic routes that enable maintenance of redox balance and energy production under thermal stress. Larger genomes, such as those of *Nostoc*- and *Calothrix*-related strains, encode expanded biosynthetic gene clusters (BGCs) for toxins, NRPS, RiPPs and terpenes, highlighting an enriched secondary metabolism that may contribute to stress mitigation in extreme habitats and offer biotechnological potential. The identification of novel lineages within Oscillatoriaceae and FACHB-T130 families further expands known thermophilic cyanobacterial diversity (Table S4). These newly sequenced genomes complement prior characterizations of hot-spring genera and species, providing a foundation for understanding the genetic basis of thermotolerance, which involves heat shock proteins, regulatory networks, and secondary metabolite pathways shaped by both vertical inheritance and horizontal gene transfer. Recent population genomic studies demonstrate that even within a single species, selective sweeps and localized genomic divergence contribute to adaptation to thermal niches, highlighting ongoing evolutionary processes. By increasing the number of high-quality thermophilic cyanobacterial genomes, this work not only enriches reference datasets for comparative analyses but also advances our understanding of extremophile evolution, ecological adaptation, and the molecular mechanisms underlying thermophily. Furthermore, these

insights may guide the discovery of natural adaptations with potential biotechnological applications, reinforcing the broader relevance of expanding genomic resources in this field.

ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to David P. Fewer and Rafael Popin from the University of Helsinki for generously sharing their insights and protocols for genome analysis, which greatly contributed to the success of this study.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was supported by the Polish National Centre (Grant 2015/19/B/NZ9/00473) and co-funded by the University of Warsaw through the “Mentoring Program” (Action I.1.1/IV.1.1; Grant BOB-IDUB-622-258/2022) of the “Excellence Initiative –Research University”. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

Polish National Centre: 2015/19/B/NZ9/00473.

University of Warsaw through the Mentoring Program: Action I.1.1/IV.1.1.

Excellence Initiative –Research University: BOB-IDUB-622-258/2022.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Nataliia Khomutovska conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Agnieszka Rudak analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Spyros Gkelis conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Mikołaj Kokociński conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Iwona Jasser conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.

DNA Deposition

The following information was supplied regarding the deposition of DNA sequences:

The sequence data is available in the [Supplementary File](#) and at ENA: [PRJEB93759](#).

Data Availability

The following information was supplied regarding data availability:

The raw data and sequences are available at Zenodo:

- Khomutovska, N., Ruda, A., Gkelis, S., Kokociński, M., & Jasser, I. (2025). Genomic insights into hot spring cyanobacteria [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.17995636>.

The sequence data is available in the [Supplemental File](#) and at ENA: [PRJEB93759](https://ena.ebi.ac.uk/ena/record/PRJEB93759).

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.21284#supplemental-information>.

REFERENCES

- Alcorta J, Alarcón-Schumacher T, Salgado O, Díez B. 2020. Taxonomic novelty and distinctive genomic features of hot spring cyanobacteria. *Frontiers in Genetics* 11:568223 DOI [10.3389/fgene.2020.568223](https://doi.org/10.3389/fgene.2020.568223).
- Alvarenga DO, Fiore MF, Varani AM. 2017. A metagenomic approach to cyanobacterial genomics. *Frontiers in Microbiology* 8:809 DOI [10.3389/fmicb.2017.00809](https://doi.org/10.3389/fmicb.2017.00809).
- Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. Available at <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- Arnison PG, Bibb MJ, Bierbaum G, Bowers AA, Bugni TS, Bulaj G, Camarero JA, Campopiano DJ, Challis GL, Clardy J, Cotter PD, Craik DJ, Dawson , Dittman E, Donadio S, Dorrestein PC, Entian K-D, Fischbach MA, Garavelli JS, Göransson U, Gruber CW, Haft DH, Hemscheidt TK, Hertweck C, Hill C, Horswill AR, Jaspars M, Kelly WL, Klinman JP, Kuipers OP, Link AJ, Liu W, Marahiel MA, Mitchell DA, Moll GN, Moore BS, Müller R, Nair SK, Nes IF, Norris GE, Olivera BM, Onaka H, Patchett ML, Piel J, Reaney MJT, Rebuffat S, Ross RP, Sahl H-G, Schmidt EW, Selsted ME, Severinov K, Shen B, Sivonen K, Smith L, Stein T, Süßmuth RD, Tagg JR, Tang G-L, Truman AW, Vederas JC, Walsh CT, Walton JD, Wenzel SC, Willey J, vander Donk W. 2013. Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Natural Product Reports* 30(1):108–160 DOI [10.1039/c2np20085f](https://doi.org/10.1039/c2np20085f).
- Aziz RK, Bartels D, Best AA, De Jongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75 DOI [10.1186/1471-2164-9-75](https://doi.org/10.1186/1471-2164-9-75).
- Bauersachs T, Miller SR, vander Meer MTJ, Hopmans EC, Schouten S, Sinninghe Damsté JS. 2013. Distribution of long chain heterocyst glycolipids in cultures of the thermophilic cyanobacterium *Mastigocladus laminosus* and a hot spring microbial mat. *Organic Geochemistry* 56:19–29 DOI [10.1016/j.orggeochem.2012.11.013](https://doi.org/10.1016/j.orggeochem.2012.11.013).

- Boudreau PD, Monroe EA, Mehrotra S, Desfor S, Korobeynikov A, Sherman DH, Murray TF, Gerwick L, Dorrestein PC, Gerwick WH. 2015.** Expanding the described metabolome of the marine cyanobacterium *Moorea producens* JHB through orthogonal natural products workflows. *PLOS ONE* **29**;10(7):e0133297 DOI [10.1371/journal.pone.0133297](https://doi.org/10.1371/journal.pone.0133297).
- Bradnam KR, Fass JN, Alexandrov A, Baranay P, Bechner M, Birol I, Boisvert S, Chapman JA, Chapuis G, Chikhi R. 2013.** Assemblathon 2: evaluating de novo methods of genome assembly in three vertebrate species. *Gigascience* **2**(1):2047–217.
- Brito FF, Ochoade Alda JA, Vasconcelos V, Pinto E. 2017.** Comparative genomics of *Leptolyngbya* species reveals adaptive strategies to extreme environments. *Frontiers in Microbiology* **8**:1066 DOI [10.1038/s41396-020-00775-z](https://doi.org/10.1038/s41396-020-00775-z).
- Chang HY, Yen HC, Chu HA, Kuo CH. 2024.** Population genomics of a thermophilic cyanobacterium revealed divergence at subspecies level and possible adaptation genes. *Botanical Studies* **65**:35 DOI [10.1186/s40529-024-00442-y](https://doi.org/10.1186/s40529-024-00442-y).
- Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. 2022.** GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics* **38**(23):5315–5316 DOI [10.1093/bioinformatics/btac672](https://doi.org/10.1093/bioinformatics/btac672).
- Chen MY, Teng WK, Zhao L, Hu CX, Zhou YK, Han BP, Song LR, Shu WS. 2021.** Comparative genomics reveals insights into cyanobacterial evolution and habitat adaptation. *ISME Journal* **15**(1):211–227 DOI [10.1038/s41396-020-00775-z](https://doi.org/10.1038/s41396-020-00775-z).
- Cirés S, Casero MC, Quesada A. 2017.** Toxicity at the edge of life: a review on cyanobacterial toxins from extreme environments. *Marine Drugs* **15**(7):233 DOI [10.3390/md15070233](https://doi.org/10.3390/md15070233).
- Cornet L, Bertrand AR, Hanikenne M, Javaux EJ, Wilmotte A, Baurain D. 2018.** Metagenomic assembly of new (sub)polar Cyanobacteria and their associated microbiome from non-axenic cultures. *Microbial Genomics* **4**(9):e000212 DOI [10.1099/mgen.0.000212](https://doi.org/10.1099/mgen.0.000212).
- Dagan T, Roettger M, Stucken K, Landan G, Koch R, Major P, Gould SB, Goremykin VV, Rippka R, Tandeau de Marsac N, Gugger M, Lockhart PJ, Allen JF, Brune I, Maus I, Pühler A, Martin WF. 2013.** Genomes of Stigonematalean cyanobacteria (subsection V) and the evolution of oxygenic photosynthesis from prokaryotes to plastids. *Genome Biology and Evolution* **5**(1):31–44 DOI [10.1093/gbe/evs117](https://doi.org/10.1093/gbe/evs117).
- Dietz KJ. 2011.** Peroxiredoxins in plants and cyanobacteria. *Antioxidants & Redox Signaling* **15**(4):1129–1159 DOI [10.1089/ars.2010.3657](https://doi.org/10.1089/ars.2010.3657).
- Dos Santos PC, Fang Z, Mason SW, Setubal JC, Dixon R. 2012.** Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. *BMC Genomics* **13**:162 DOI [10.1186/1471-2164-13-162](https://doi.org/10.1186/1471-2164-13-162).
- Dreher TW, Collart LP, Mueller RS, Halsey KH, Bildfell RJ, Schreder P, Sobhakumari A, Ferry R. 2019.** Anabaena/Dolichospermum as the source of lethal microcystin levels responsible for a large cattle toxicosis event. *Toxicon: X* **1**:100003 DOI [10.1016/j.toxcx.2018.100003](https://doi.org/10.1016/j.toxcx.2018.100003).
- Eungrasamee K, Zhu Z, Liu X, Jantaro S, Lindblad P. 2024.** Lipid metabolism in cyanobacteria: biosynthesis and utilization. In: cyanobacteria: metabolisms to

molecules. *Progress in Biochemistry and Biotechnology* 2:85–116
DOI 10.1016/B978-0-443-13231-5.00017-9.

- Fewer DP, Osterholm J, Rouhiainen L, Jokela J, Wahlsten M, Sivonen K. 2011.** Nostophycin biosynthesis is directed by a hybrid polyketide synthase-nonribosomal peptide synthetase in the toxic cyanobacterium *Nostoc* sp. strain 152. *Applied and Environmental Microbiology* 77(22):8034–8040 DOI 10.1128/AEM.05993-11.
- Guiry MD, Guiry GM. 2025.** AlgaeBase. World-wide electronic publication, University of Galway. Available at <https://www.algaebase.org>.
- Hiraide Y, Oshima K, Fujisawa T, Uesaka K, Hirose Y, Tsujimoto R, Yamamoto H, Okamoto S, Nakamura Y, Terauchi K, Omata T, Ihara K, Hattori M, Fujita Y. 2014.** Loss of cytochrome cM stimulates cyanobacterial heterotrophic growth in the dark. *Plant and Cell Physiology* 56(2):334–345 DOI 10.1093/pcp/pcu165.
- Huo D, Li H, Cai F, Guo X, Qiao Z, Wang W, Yu G, Li R. 2021.** Genome evolution of filamentous cyanobacterium *Nostoc* species: from facultative symbiosis to free living. *Microorganisms* 9(10):2015 DOI 10.3390/microorganisms9102015.
- Jablonska M, Eleršek T, Kogovšek P, Skok S, Oarga-Mulec A, Mulec J. 2024.** Molecular screening for cyanobacteria and their cyanotoxin potential in diverse habitats. *Toxins* 16(8):333 DOI 10.3390/toxins16080333.
- Jasser I, Panou M, Khomutovska N, Sandzewicz M, Panteris E, Niyatbekov T, Łach Ł, Kwiatkowski J, Kokociński M, Gkelis S. 2022.** Cyanobacteria in hot pursuit: characterization of cyanobacteria strains, including novel taxa, isolated from geothermal habitats from different ecoregions of the world. *Molecular Phylogenetics and Evolution* 170:107454 DOI 10.1016/j.ympev.2022.107454.
- Jiao GF, Liang WY, Chen , You XR, Xie DL, Xu HC. 2011.** Cloning and expression of genes which encode subunit alpha and beta of C-phycoyanin of *Nostoc flagelliforme*. *Plant Science Journal* 29(3):331–339. Available at <https://plantscience.cn/en/article/id/793>.
- Kajitani R, Yoshimura D, Okuno M, Minakuchi Y, Kagoshima H, Fujiyama A, Kubokawa K, Kohara Y, Toyoda A, Itoh T. 2019.** Platanus-allee is a de novo haplotype assembler enabling a comprehensive access to divergent heterozygous regions. *Nature Communications* 10:1702 DOI 10.1038/s41467-019-09575-2.
- Kees ED, Murugapiran SK, Bennett AC, Hamilton TL. 2022.** Distribution and genomic variation of thermophilic cyanobacteria in diverse microbial mats at the upper temperature limits of photosynthesis. *mSystems* 7:e00317-22 DOI 10.1128/msystems.00317-22.
- Kehr JC, Gatte Picchi D, Dittmann E. 2011.** Natural product biosyntheses in cyanobacteria: a treasure trove of unique enzymes. Beilstein. *The Journal of Organic Chemistry* 7:1622–1635 DOI 10.3762/bjoc.7.191.
- Khomutovska N, Sandzewicz M, Łach Ł, Suska-Malawska M, Chmielewska M, Mazur-Marzec H, Ceglowska M, Niyatbekov T, Wood SA, Puddick J, Kwiatowski J, Jassers I. 2020.** Limited microcystin, anatoxin and cylindrospermopsin production by cyanobacteria from microbial mats in cold deserts. *Toxins* 12(4):244 DOI 10.3390/toxins12040244.

- Klatt CG, Wood JM, Rusch DB, Bateson MM, Hamamura N, Heidelberg JF, Grossman AR, Bhaya D, Cohan FM, K hl M, Bryant DA, Ward DM. 2011. Community ecology of hot spring cyanobacterial mats: predominant populations and their functional potential. *ISME Journal* 5(8):1262–1278 DOI 10.1038/ismej.2011.73.
- Konstantinou D, Popin RV, Fewer DP, Sivonen K, Gkelis S. 2021. Genome reduction and secondary metabolism of the marine sponge-associated cyanobacterium *Leptothoe*. *Marine Drugs* 19.6:298 DOI 10.3390/md19060298.
- Liang Y, Kaczmarek MB, Kasprzak AK, Tang J, Shah MMR, Jin P, Klepacz-Sm łka A, Cheng JJ, Ledakowicz S, Daroch M. 2018. Thermosynechococcaceae as a source of thermostable C-phycocyanins: properties and molecular insights. *Algal Research* 35:223–235 DOI 10.1016/j.algal.2018.08.037.
- Lindell D, Sullivan MB, Johnson ZI, Tolonen AC, Rohwer F, Chisholm SW. 2004. Transfer of photosynthesis genes to and from prochlorococcus viruses. *Proceedings of the National Academy of Sciences of the United States of America* 101(30):11013–11018 DOI 10.1073/pnas.0401526101.
- Mares J, Chmelik D, Galica T, Urajova P, Hajek J, Rehakova K, Hrouzek P. 2020. A newly discovered NRPS/PKS pathway of the cyanobacterial polipeptide muscotoxin utilizes 4-methyl proline generated by the nostopeptolide biosynthetic machinery. Available at <https://www.ncbi.nlm.nih.gov/nucore/2099398330>.
- Mazmouz R, Chapuis-Hugon F, Mann S, Pichon V, M jean A, Ploux O. 2010. Biosynthesis of cylindrospermopsin and 7-epicylindrospermopsin in *Oscillatoria* sp. strain PCC 6506: identification of the *cyr* gene cluster and toxin analysis. *Applied and Environmental Microbiology* 76(15):4943–4949 DOI 10.1128/AEM.00717-10.
- Meeks JC, Elhai J. 2002. Regulation of cellular differentiation in filamentous cyanobacteria in free-living and plant-associated symbiotic growth states. *Microbiology and Molecular Biology Reviews* 66(1):94–121 DOI 10.1128/mmbr.66.1.94-121.2002.
- Mihali TK, Muenchhoff R, Kellmann J, Barrow KD, Neilan BA. 2008. Characterization of the gene cluster responsible for cylindrospermopsin biosynthesis. *Applied and Environmental Microbiology* 74:716–722 DOI 10.1128/AEM.01988-07.
- Mikheenko A, Prjibelski A, Saveliev V, Antipov D, Gurevich A. 2018. Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics* 34(13):i142–i150 DOI 10.1093/bioinformatics/bty266.
- Moten D, Batsalova T, Dzhambazov B, Teneva I. 2020. *Comparative genome analysis of some representatives of genus Nostoc. vol. 20.5.1*. Sofia: Surveying Geology & Mining Ecology Management (SGEM), 167–173 DOI 10.5593/sgem2020/5.1/s20.021.
- Olm MR, Brown CT, Brooks B, Banfield JF. 2017. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *The ISME Journal* 11(12):2864–2868 DOI 10.1038/ismej.2017.126.
- Panyakampol J, Cheevadhanarak S, Sutheeworapong S, Chaijaruwanich J, Senachak J, Siangdung W, Jeamton W, Tanticharoen M, Paithoonrangsarid K. 2015. Physiological and transcriptional responses to high temperature in *Arthrospira* (*Spirulina*) *platensis* C1. *Plant and Cell Physiology* 56(3):481–496 DOI 10.1093/pcp/pcu192.

- Parks DH, Chuvochina M, Chaumeil PA, Rinka C, Mussig AJ, Hugenholtz P. 2020. A complete domain-to-species taxonomy for Bacteria and Archaea. *Nature Biotechnology* 38:1079–1086 DOI 10.1038/s41587-020-0501-8.
- Pedersen D, Miller S. 2017. Photosynthetic temperature adaptation during niche diversification of the thermophilic cyanobacterium *Synechococcus* A/B clade. *ISME Journal* 11:1053–1057 DOI 10.1038/ismej.2016.173.
- Pierpont CJ, Baroch JJ, Church MJ, Miller SR. 2024. Idiosyncratic genome evolution of the thermophilic cyanobacterium *Synechococcus* at the limits of phototrophy. *ISME Journal* 18(1) DOI 10.1093/ismejo/wrae184.
- Price MN, Dehal PS, Arkin AP. 2010. FastTree 2 approximately maximum-likelihood trees for large alignments. *PLOS ONE* 5(3):e9490 DOI 10.1371/journal.pone.0009490.
- Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. 2020. Using SPAdes de novo assembler. *Current Protocols in Bioinformatics* 70(1):e102 DOI 10.1002/cpbi.102.
- Ragsdale SW, Pierce . 2008. The wood–Ljungdahl pathway of acetogenesis. *Biochimica et Biophysica Acta (BBA)–Proteins and Proteomics* 1784(12):1873–1898 DOI 10.1016/j.bbapap.2008.08.012.
- Rantala A, Rajaniemi-Wacklin P, Lyra C, Lepistö L, Rintala J, Mankiewicz-Boczek J, Sivonen K. 2006. Detection of microcystin-producing cyanobacteria in Finnish lakes with genus-specific microcystin synthetase gene E (mcyE) PCR and associations with environmental factors. *Applied and Environmental Microbiology* 72(9):6101–6110 DOI 10.1128/AEM.01058-06.
- Redl S, Poehlein A, Esser C, Bengelsdorf FR, Jensen TØ, Jendresen CB, Tindall BJ, Daniel R, Dürre P, Nielsen AT. 2020. Genome-based comparison of all species of the genus *Moorella*, and status of the species *Moorella thermoacetica* and *Moorella thermoautotrophica*. *Frontiers in Microbiology* 10:3070 DOI 10.3389/fmicb.2019.03070.
- Saran N, Mishra AK. 2015. Assessment of genetic diversity and molecular phylogeny of toxin producing cyanobacteria using metagenomic approaches. Available at <https://www.ncbi.nlm.nih.gov/nuccore/1028039980>.
- Sato N, Kamimura R, Kaneta K, Yoshikawa M, Tsuzuki M. 2017. Species-specific roles of sulfolipid metabolism in acclimation of photosynthetic microbes to sulfur-starvation stress. *PLOS ONE* 12(10):e0186154 DOI 10.1371/journal.pone.0186154.
- Shaffer M, Borton MA, McGivern BB, Zayed AA, Leanti La Rosa S, Solden LM, Liu P, Narrowe AB, Rodríguez-Ramos J, Bolduc B, Gazitúa MC, Daly RA, Smith GJ, Vik DR, Pope PB, Sullivan MB, Roux S, Wrighton KC. 2020. DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Research* 48(16):8883–8900 DOI 10.1093/nar/gkaa621.
- Shakena Fathima T, Fayaazuddin T, Dhanveer Ahamed B, Prakash P, Akilan E, Asraf Sithikka MR, Muralitharan G, Thajuddin N. 2023. Suitability of phycocyanin gene sequences for the identification of Cyanobacterial taxa belonging to Oscillatoriales. *Biological Forum –An International Journal* 15(6):491–498. Available at <https://www.researchtrend.net/bfij/pdf/Suitability-of-Phycocyanin-gene-Sequences-for-the->

[Identification-of-Cyanobacterial-taxa-belonging-to-Oscillatoriales-Majoj-Kumar-73.pdf](#).

- Shi T, Falkowski PG. 2008.** Genome evolution in cyanobacteria: the stable core and the variable shell. *Proceedings of the National Academy of Sciences of the United States of America* **105**(7):2510–2515 DOI [10.1073/pnas.0711165105](#).
- Shih PM, Wu D, Latifi A, Axen SD, Fewer DP, Talla E, Calteau A, Cai F, Tandeau de Marsac N, Rippka R, Herdman M, Sivonen K, Coursin T, Laurent T, Goodwin L, Nolan M, Davenport K, Han CS, Rubin EM, Eisen JA, Woyke T, Gugger M, Kerfeld CA. 2013.** Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proceedings of the National Academy of Sciences of the United States of America* **110**:1053–1058 DOI [10.1073/pnas.1217107110](#).
- Shishido TK, Jokela J, Humisto A, Suurnäkki S, Wahlsten M, Alvarenga DO, Sivonen K, Fewer DP. 2019.** The biosynthesis of rare homo-amino acid containing variants of microcystin by a benthic cyanobacterium. *Marine Drugs* **17**(5):271 DOI [10.3390/md17050271](#).
- Sivonen K, Jones G. 1999.** Cyanobacterial toxins. In: Chorus I, Bartram J, eds. *Toxic Cyanobacteria in water*. Boca Raton: CRC Press, 41–111.
- Sivonen K, Leikoski N, Fewer DP, Jokela J. 2010.** Cyanobactins—ribosomal cyclic peptides produced by cyanobacteria. *Applied Microbiology and Biotechnology* **86**:1213–1225 DOI [10.1007/s00253-010-2482-x](#).
- Suzuki I, Simon WJ, Slabas AR. 2006.** The heat shock response of *Synechocystis* sp. PCC 6803 analysed by transcriptomics and proteomics. *Journal of Experimental Botany* **57**(7):1573–1578 DOI [10.1093/jxb/erj148](#).
- Tang J, Hu Z, Zhang J, Daroch M. 2024.** Genome-scale identification and comparative analysis of transcription factors in thermophilic cyanobacteria. *BMC Genomics* **25**:44 DOI [10.1186/s12864-024-09969-7](#).
- Tang J, Zhou H, Yao D, Riaz S, You D, Klepacz-Smółka A, Daroch M. 2022.** Comparative genomic analysis revealed distinct molecular components and organization of CO₂-concentrating mechanism in thermophilic cyanobacteria. *Frontiers in Microbiology* **13**:876272 DOI [10.3389/fmicb.2022.876272](#).
- Tikhonova I, Kuzmin A, Deeva D, Sorokovikova E, Potapov S, Lomakina A, Belykh O. 2017.** Cyanobacteria *Nostoc punctiforme* from abyssal benthos of Lake Baikal: unique ecology and metabolic potential. *Indian Journal of Microbiology* **57**(4):422–426 DOI [10.1007/s12088-017-0673-1](#).
- Tillett D, Dittmann E, Erhard M, von Döhren H, Börner T, Neilan BA. 2000.** Structural organization of microcystin biosynthesis in *Microcystis aeruginosa* PCC7806: an integrated peptide-polyketide synthetase system. *Chemistry & Biology* **7**:753–764 DOI [10.1016/s1074-5521\(00\)00021-1](#).
- Vieira Vaz MGM, Almeida AVM, de Castro NV, Nunes-Nesi A, Araújo WL. 2020.** Carbon and nitrogen metabolism in cyanobacteria: basic traits, regulation and biotechnological application. In: *Handbook of algal science, technology and medicine*. Cambridge MA: Academic Press, 245–254 DOI [10.1016/B978-0-12-818305-2.00015-2](#).

- Wang Q, Cen Z, Zhao J. 2015.** The survival mechanisms of thermophiles at high temperatures: an angle of omics. *Physiology* **30**(2):97–106 DOI [10.1152/physiol.00066.2013](https://doi.org/10.1152/physiol.00066.2013).
- Ward DM, Castenholz RW, Miller SR. 2012.** Cyanobacteria in geothermal habitats. In: Whitton BA, ed. *Ecology of Cyanobacteria II*. Dordrecht: Springer Netherlands, 39–63 DOI [10.1007/978-94-007-3855-3_3](https://doi.org/10.1007/978-94-007-3855-3_3).
- Ward RD, Stajich JE, Johansen JR, Huntemann M, Clum A, Foster B, Foster B, Roux S, Palaniappan K, Varghese N, Mukherjee S, Reddy TBK, Daum C, Copeland A, Chen I-MA, Ivanova NN, Kyrpides NC, Shapiro N, Elie-Fadrosh EA, Pietrasiak N. 2021.** Metagenome sequencing to explore phylogenomics of terrestrial cyanobacteria. *Microbiology Resource Announcements* **10**(22):e0025821 DOI [10.1128/MRA.00258-21](https://doi.org/10.1128/MRA.00258-21).
- Weber T, Blin K, Duddela S, Krug D, Kim HU, Brucoleri R, Lee SY, Fischbach MA, Müller R, Wohlleben W, Breitling R, Takano E, Medema MH. 2015.** antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Research* **43**(W1):W237–W243 DOI [10.1093/nar/gkv437](https://doi.org/10.1093/nar/gkv437).
- Winstedt L, Yoshida K, Fujita Y, von Wachenfeldt C. 1998.** Cytochrome bd Biosynthesis in *Bacillus subtilis*: characterization of the cydABCD Operon. *Journal of Bacteriology* **180** DOI [10.1128/jb.180.24.6571-6580.1998](https://doi.org/10.1128/jb.180.24.6571-6580.1998).
- Wood DE, Lu J, Langmead B. 2019.** Improved metagenomic analysis with Kraken 2. *Genome Biology* **20**:257 DOI [10.1186/s13059-019-1891-0](https://doi.org/10.1186/s13059-019-1891-0).
- Yarbrough KA. 2020.** Conservatism of chemosynthetic sulfur metabolism functional potential varies across prokaryotes. Master’s thesis, Long Beach, California: California State University.