Geminocystis urbisnovae sp. nov. (Chroococcales, Cyanobacteria): polyphasic description complemented with a survey of the family Geminocystaceae

Elena Polyakova¹, Svetlana Averina² and Alexander Pinevich²,*

¹Institute of Fundamental Medicine and Biology, Kazan Federal University, Kazan 420012, Russian Federation
²Department of Microbiology, St. Petersburg State University, St. Petersburg 199178, Russian Federation

Progress in phylogenomic analysis has led to a considerable re-evaluation of former cyanobacterial system, with many new taxa being established at different nomenclatural levels. The family Geminocystaceae is among cyanobacterial taxa recently described on the basis of polyphasic approach. Within this family, there are six genera: Geminocystis, Cyanobacterium, Geminobacterium, Annamia, Picocyanobacterium, and Microcrocis. The genus Geminocystis previously encompassed two species: G. herdmanii and G. papuanica. Herein, a new species G. urbisnovae was proposed under the provisions of the International Code of Nomenclature for algae, fungi, and plants (ICN). Polyphasic analysis was performed for five strains from the CALU culture collection (St. Petersburg State University, Russian Federation), and they were assigned to the genus Geminocystis in accordance with high 16S rRNA gene similarity to existing species, as well as because of proximity to these species on the phylogenetic trees reconstructed with RaxML and Bayes methods. Plausibility of their assignment to a separate species of the genus Geminocystis was substantiated with smaller cell size; stenohaline freshwater ecotype; capability to complementary chromatic adaptation of second type (CA2); distinct 16S rRNA gene clustering; sequences and folding of D1–D1’ and B box domains of the 16S–23S internal transcribed spacer region. The second objective pursued by this communication was to provide a survey of the family Geminocystaceae. The overall assessment was that, despite attention of many researchers, this cyanobacterial family has been understudied and, especially in the case of the crucially important genus Cyanobacterium, taxonomically problematic.

Keywords: complementary chromatic adaptation; Cyanobacterium; Geminocystaceae; Geminocystis; ITS; rbcL gene; rpoC1 gene; 16S rRNA gene phylogeny

INTRODUCTION

The genus Geminocystis (Korelusová et al. 2009) was proposed under provisions of the International Code of Nomenclature for algae, fungi, and plants (ICN) for a group of cyanobacterial strains previously identified as Synechocystis sp. With Geminocystis as the type genus, a new family Geminocystaceae (Chroococcales) was established (Tuji et al. 2021).

At the same time, some of abovementioned strains, in particular the reference strain Synechocystis sp. PCC 6308, were attributed to “Synechocystis Cluster 1” within the taxonomic system of cyanobacteria adopted in Bergey’s Manual of Systematic Bacteriology (Herdman et al. 2001). Unfortunately, this system has been essentially based on morphological distinctions rather than phylo-
genetic relationships; besides, it did not comply with the provisions of the International Code of Nomenclature of Prokaryotes (ICNP) (Oren 2020). In this connection, it was a long-awaited event when the International Committee on Systematics of Prokaryotes (ICSP) decided to emend the ICNP rules so that the names of cyanobacterial taxa validly published under provisions of the ICN are automatically validly published under the ICNP (Oren et al. 2021). Thus, a long-standing problem of illegitimate taxonomic status of cyanobacteria in prokaryotic nomenclature was at last solved.

Meanwhile, current progress in the taxonomy of cyanobacteria has been greatly impelled by polyphasic approach combining in situ and ex situ morphological data with molecular phylogeny records (Komárek et al. 2014, Oren et al. 2022, Strunceký et al. 2023). However, regarding cyanobacteria of the family Geminocystaceae, the analysis has involved a relatively low number of cultured strains. As the result, these understudied cyanobacteria are taxonomically problematic.

Herein, we applied a polyphasic approach to the group of strains previously identified as *Synechocystis* sp. On the basis of obtained data, we proposed a new species *Geminocystis urbisnovae*. Taking into account many researchers’ interest, and accumulating proposals of new Gemonocystaceae genera and species, we additionally reviewed this family especially focusing on the molecular phylogeny criteria used in taxonomy of cyanobacteria.

**MATERIALS AND METHODS**

**Strains and cultivation**

The strains CALU 1334, CALU 1759, CALU 1794, CALU 1801, and CALU 1807 previously identified as *Synechocystis* sp. were from the CALU culture collection (Pinevich et al. 2004) maintained at St. Petersburg State University (St. Petersburg, Russian Federation) Research Park (https://researchpark.spbu.ru/en/collection-ccem-eng/1930-ccem-kollekiya-calu-eng). Despite numerous efforts to ensure culture purity, all these strains have not been purged of satellite bacteria (residual contamination was revealed via plating and light microscopy). Culturing was performed by monthly re-inoculations into modified (1/5 nitrate amount) BG-11 medium–liquid or solidified with 1.2% Difco agar. For analytical purposes, batch cultures were grown for one–two weeks in unshaken 50 mL–Erlenmeyer flasks, or sterile air-bubbled 400 mL–cylindrical vessels, at ambient temperature and continuous illumination of 10 μmol photons m⁻² s⁻¹ cool-white light. Culture density was measured at 750 nm wavelength with a 2802S UV/VIS spectrophotometer (Unico, Franksville, WI, USA); growth rate (μ) was calculated for three independent essays with two biological duplicates and two technical duplicates each.

Salinity responses were examined in the strains cultured in liquid-modified BG-11 medium additionally supplied with 0.3–3.5% NaCl (variants with 0.3, 1.0, 2.0, and 3.5% NaCl, correspondingly) during 10 days; culture density was measured as above.

**Morphological characterization**

Light microscopy was performed with a Leica 2500 microscope equipped with a DFC 500 camera (Leica, Wetzlar, Germany) under differential interference contrast regime.

Mean cell dimensions were measured with a Leica TCS SP5 confocal microscope for ≥30 cells in each of three independent cultivations; resulting variation in dimensions was represented as standard deviation calculated with Microsoft Excell 2010 software. Transmission electron microscope samples were prepared as in Velichko et al. (2012), and viewed at 50–70 kV accelerating voltage in a JEM-1400 electron microscope (Jeol Ltd., Tokyo, Japan) equipped with an Olympus-SIS Veleta camera (Tokyo, Japan). Scanning electron microscope samples obtained by fixation with 4.0% glutaraldehyde in 0.1 M Na-phosphate buffer (pH 7.0) were placed onto 1.0 μm-mesh Millipore membrane filters (Sigma, St. Louis, MO, USA), post-fixed in 1.0% osmium tetroxide, stepwise dehydrated with ethanol and acetone, critical point-dried in a Leica EM CPD 300 apparatus, shadowed with a 10 nm-thick gold layer in a Leica EM SCD 500 apparatus, and viewed in a Tescan MIRA 3 LMU scanning electron microscope (TESCAN, Brno, Czech Republic).

**Complementary chromatic adaptation analysis**

Batch cultures were grown for 12–14 days under a Lee 108/124 green or a Lee 106 red filter (Lee Filters, Andover, UK) that correspondingly ensured preferential syntheses of phycocerythrin (PE) or phycocyanin (PC). Harvested cells were rinsed with 50 mM Na-phosphate buffer (pH 7.0), and disrupted in a UD-20 ultrasonic apparatus (Techpan, Warsaw, Poland). Particulate fraction was removed by centrifugation at 8,000 ×g, followed by filtration through 0.22 μm-mesh Millipore membrane. Light absorbance was recorded at 300–800 nm with a 2802S UV/VIS
Gene amplification and sequencing

The cells of two-week-old culture were harvested by centrifugation and rinsed with 10 mM Tris-HCl buffer (pH 8.0) containing 1 mM Na₂EDTA. DNA was isolated according to Koksharova et al. (1998), and treated with RNase A (Thermo Fisher Scientific, Russian Federation). The following gene regions were amplified using primers as follows: 16S rRNA gene with 27F and 1492R (Lane 1991, Turner et al. 1999), 16S–23S internal transcribed spacer (ITS) region with F322 and R340 (Iteman et al. 2000). The rpoCl gene was amplified with the primer pair rpcMF/rpcCR-1 (Mareš et al. 2019a), and the rbcL gene was amplified with the primer pair rbclf/rbclr (Singh et al. 2015). Polymerase chain reaction (PCR) products were visualized and purified with a Nucleospin Gel & PCR Clean-Up Kit (Macherey Nagel, Düren, Germany). Sanger sequencing was performed using a BigDye Terminator version 3.1 Cycle Sequencing Kit and an ABI3000 gene analyzer (Applied Biosystems, Foster City, CA, USA). Sequences were assembled and proofread in Sequencher version 5.4.6 software (Gene Codes Corporation, Ann Arbor, MI, USA).

Phylogenetic analysis

16S rRNA gene sequences alignment was performed with Muscle algorithm and AliView program (1.27 version) using the GenBank NCBI database references. The 16S rRNA gene + rpoCl gene + rbcL gene concatemer was constructed with SequenceMatrix (1.7.8 version) involving 89 OTUs (a total of 2,892 nucleotides). The GTR + G + I model was used for all loci. To reconstruct phylogenies, the RaxML method was employed. 16S rRNA sequences used varied in length, although after cut-off and gap filling they were of 1,300 bp. Bootstrap re-sampling was performed on 1,000 replicates. Bayesian analysis was performed in MrBayes version 3.2.6 (Ronquist et al. 2012). Two runs of four Monte Carlo Markov chains were utilized in 20 million generations, with 1,000 sampling turns, and 25% burn-in. The trees were visualized in FigTree version 1.4.4, and post-edited in Adobe Illustrator CC 2019 v23.1.0.670.

16S–23S ITS analysis

Complete 16S–23S ITS, tRNAs coding regions, D1–D1’ boxes, B boxes, V2 and V3 domains were evaluated as elsewhere (Iteman et al. 2000, Johansen et al. 2011, Müh-
Fig. 1. Differential-interference micrographs of Geminocystis urbisnovae sp. nov. cells. Strain CALU 1334 (A), strain CALU 1759 (B), strain CALU 1794 (C), strain CALU 1801 (D), and strain CALU 1807 (E). Arrows in (E) indicate cell tetrads. Scale bars represent: A–E, 10 µm.
monophyletic species concept (Johansen and Casamatta 2005) relies on phylogeny traced, in the first instance, after distances between 16S rRNA gene sequences (Kim et al. 2014). Another approach which is widely known as polyphasic combines phylogenetic data with conventional phenotypic characters (Komárek et al. 2014). This approach usually includes 16S rRNA gene phylogeny sensu monophyletic species concept.

In the case of 16S rRNA gene marker, a separate species is delimited with ≥98.65% similarity between its members (Kim et al. 2014); in turn, genera have ≤94.5% similarity threshold (Yarza et al. 2014).

Reliance on 16S alone is problematic. First, although this marker is rather conservative, its dissimilarities do not always reflect the differences in speciation rate among evolving species (Dvořák et al. 2023). Second, although 16S rRNA gene marker is generally applied in phylotaxonomy, its role is unequal regarding higher taxa (from phyla downwards to families) and lower taxa (genera and species). In the former case, this marker has been considered the *primas agere* criterion. In the latter case, more reliable proof is brought by the concatemers of 16S rRNA gene with other informational genes (e.g., *rpoC1*) and some of the major housekeeping genes (e.g., *rbcL*). Distinguishing details of the 16S–23S ITS region can be also used as the supporting phylogeny records. Incorporating multiple phylogenetic records into polyphasic approach is preferred from informative viewpoint while discarding it in favor of a single 16S rRNA gene marker delimitation in cyanobacteria. One of them, termed the

**DISCUSSION**

There are two currently employed means of species delimitation in cyanobacteria. One of them, termed the
<table>
<thead>
<tr>
<th>Strain</th>
<th>Origin</th>
<th>Cell shape</th>
<th>Change of division plane by 90°</th>
<th>Cell dimensions (µm)</th>
<th>Response to salinity</th>
<th>Color; chromatic adaptation (CA)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Geminocystis</em> sp.</td>
<td>Alkaline lake, Nhecolândia, Brazil (Andreote et al. 2014)</td>
<td>Spherical; ellipsoidal or hemispheric after division</td>
<td>Absent</td>
<td>4.08−7.15</td>
<td>Haloalkalophilic</td>
<td>No data</td>
</tr>
<tr>
<td>CHAB 6541</td>
<td>Alkaline Lake Qinghai, China (Ma et al. 2019)</td>
<td>Spherical or ellipsoidal</td>
<td>Present</td>
<td>2.71−4.22</td>
<td>Haloalkalophilic</td>
<td>Blue-green; no data on CA</td>
</tr>
<tr>
<td><em>Geminocystis</em> sp.</td>
<td>A lake in Wisconsin, USA (Rippka et al. 1979)</td>
<td>Spherical or hemispheric after division</td>
<td>Present</td>
<td>3.0−3.5 × 4.0−4.5</td>
<td>Freshwater</td>
<td>Blue-green; no data on CA</td>
</tr>
<tr>
<td>PCC 6308</td>
<td>Thermal Spring Llás, Costa Rica (Brenes-Guillén et al. 2021)</td>
<td>Spherical; ellipsoidal or tandemly hemispheric after division</td>
<td>Absent</td>
<td>4.0</td>
<td>Strongly halotolerant (up to 15% NaCl)</td>
<td>No data</td>
</tr>
<tr>
<td><em>Geminocystis</em> sp.</td>
<td>A stream in Mimomi Hongo Park, Japan (Hirose et al. 2015a)</td>
<td>Spherical</td>
<td>No data</td>
<td>−5</td>
<td>Freshwater</td>
<td>Blue-green to brown; CA2</td>
</tr>
<tr>
<td>NIES-3708</td>
<td>A stream in Mimomi Hongo Park, Japan (Hirose et al. 2015b)</td>
<td>Ellipsoidal</td>
<td>No data</td>
<td>−5</td>
<td>Freshwater</td>
<td>Blue-green to reddish brown; CA2</td>
</tr>
<tr>
<td><em>G. papuanica</em> PAP1</td>
<td>Tropical rain forest, Papua-New Guinea (Korelusová et al. 2009)</td>
<td>Spherical; ellipsoidal or hemispheric after division</td>
<td>Present</td>
<td>3.5−4.0 × 5.0−6.2</td>
<td>No data</td>
<td>Blue-green to dark grey; no data on CA</td>
</tr>
<tr>
<td><em>Geminocystis</em> sp.</td>
<td>Platistomo Spring, Greece (Bravakos et al. 2016)</td>
<td>Spherical or ellipsoidal</td>
<td>Absent</td>
<td>5−7</td>
<td>Freshwater</td>
<td>Brown; no data on CA</td>
</tr>
<tr>
<td><em>G. urbisnovae</em> CALU 1334</td>
<td>Lake Ladoga, Russia (this study)</td>
<td>Spherical; ellipsoidal or hemispheric after division</td>
<td>Present</td>
<td>1.5−2.1 × 2.7−3.2</td>
<td>Freshwater; weakly halotolerant</td>
<td>Dark green, light green or olive green; CA2</td>
</tr>
<tr>
<td><em>G. urbisnovae</em> CALU 1759</td>
<td>Korotskoj Lake, Russia (this study)</td>
<td>Spherical; ellipsoidal or hemispheric after division</td>
<td>Present</td>
<td>2.4−2.7 × 3.3−3.7</td>
<td>Freshwater; weakly halotolerant</td>
<td>Dark green, light green or olive green; CA2</td>
</tr>
<tr>
<td><em>G. urbisnovae</em> CALU 1794</td>
<td>Korotskoj Lake, Russia (this study)</td>
<td>Spherical; ellipsoidal or hemispheric after division</td>
<td>Present</td>
<td>2.5−3.1 × 3.7−4.0</td>
<td>Freshwater; weakly halotolerant</td>
<td>Dark green, light green or olive green; CA2</td>
</tr>
<tr>
<td><em>G. urbisnovae</em> CALU 1801</td>
<td>Korotskoj Lake, Russia (this study)</td>
<td>Spherical; ellipsoidal or hemispheric after division</td>
<td>Present</td>
<td>2.2−2.7 × 3.3−3.6</td>
<td>Freshwater; weakly halotolerant</td>
<td>Dark green, light green or olive green; CA2</td>
</tr>
<tr>
<td><em>G. urbisnovae</em> CALU 1807</td>
<td>Korotskoj Lake, Russia (this study)</td>
<td>Spherical; ellipsoidal or hemispheric after division</td>
<td>Present</td>
<td>2.5−2.8 × 3.4−3.5</td>
<td>Freshwater; weakly halotolerant</td>
<td>Dark green, light green or olive green; CA2</td>
</tr>
</tbody>
</table>
Fig. 3. RaxML/Bayes phylogenetic tree for 16S rRNA gene sequences. Node support indicated by bootstrap values from the maximum likelihood / Bayesian analysis posterior probabilities. Scale bar represents: 0.06 substitutions per nucleotide position.
approach is rejected (Johansen et al. 2021).

Taking into account that 16S rRNA gene similarity of CALU strains to existing Geminocystis spp. was 94.5–97.9% (Supplementary Table S2), they do not belong to any other genus except Geminocystis. This can also be concluded from a close clustering of these strains with the strains Geminocystis sp. 1-1, Geminocystis sp. CHAB-6541, Geminocystis herdmani PCC 6308, Geminocystis sp. NIES-3708, Geminocystis sp. NIES-3709, Geminocystis sp. PltCNmm3, Geminocystis sp. PMC-995-17, and Geminocystis papuanica PAP1 in the phylogenetic trees (Figs 3 & 4, Supplementary Fig. S4) as well as from the peculiarities of the ITS region (Fig. 5).

To conclude, on the basis of morphological and ultrastructural characters of CALU strains (in particular, lamellar system type shared with other Geminocystis spp.) (Fig. 2), and taking into account smaller cell size (Table 1) as well as distinguishing physiological characters such as stenohaline freshwater ecotype and capability of complementary CA of the second type, these cyanobacteria could be ascribed to a new species of the genus Geminocystis. Corresponding phylogeny-based support issued from the 16S rDNA sequencing data that showed a distinct location of CALU strains on evolutionary trees (Figs 3 & 4). An accessory evidence of new species were D1−D1′ and B box structures (Fig. 5) dissimilar from those in existing members of the genus.

**Fig. 4.** Bayes phylogenetic tree for 16S rRNA gene + rbcL gene + rpoC gene concatemers. Node support indicated as posterior probabilities from Bayesian analysis. Scale bar represents: 0.5 substitutions per nucleotide position.
Fig. 5. Schematic representation of D1–D1’ (A) and B box (B) domain folding in 16S-23S internal transcribed spacer of Geminocystaceae. In rectangles, conservative nucleotides within D1–D1’ and B box loops.
Formal diagnosis of *Geminocystis urbisnovae*

On the basis of a polyphasic analysis, the following diagnosis for new species of the genus *Geminocystis* (Chroococcales, Cyanobacteria) is proposed under the provisions of the ICN (Turland et al. 2018).

*Geminocystis urbisnovae* E. Polyakova, Averina & Pinevich sp. nov. (Fig. 1)

**Description.** Dark green, light green, or olive green coci 2.7 ± 0.3–3.6 ± 0.2 µm long (min–max, 2.1–4.5 µm) and 2.4 ± 0.3–3.2 ± 0.2 µm wide (min–max, 1.5–3.7 µm). Binary fission in one plane (preferentially), or in two perpendicular planes (occasionally); single, in pairs, or in tetrads. Sheath and capsule absent, mucilage produced. Parallel rows of thylakoids stretched from cell pole to pole. Loose cell precipitate in unstirred culture; homogeneous growth in stirred culture. Colonies on 1.2% agar, with slightly waveform edge and smooth shiny surface, moderately granular, bordered with slime. Freshwater, stenohaline: growth inhibited at ≥0.3% NaCl addition to modified BG-11 medium (1/5 nitrate amount; 0.05% total salt). PE produced; complementary CA2. Pairwise identity of 16S rRNA gene sequences with published strains of the genus *Geminocystis* 94.5–97.9%; species-specific nucleotide sequence and folding of D1–D1’ and B box domains of the 16S–23S ITS region.

**Etymology.** The species epithet *urbisnovae* (ur. bis. novae. Lat. fem. nom. urbis a city. L. adj. novus new, a relation to new city) is after Latinized name of Novgorod city (Nov’go.rod; literally, a new city) in north-west Russia and adjoining administrative region from which the type strain was isolated.

**Habitat.** Northern Hemisphere middle latitude freshwater pools; in plankton, benthos, or periphyton.

**Type locality.** Lake Korotskoye, Novgorod administrative region, Russia (57°56’ N, 33°15’ E).

**Holotype here designated.** Herbarium material of the strain CALU 1759, dense culture exsiccatum LE A0001561 stored at Komarov Botanical Institute of Russian Academy of Sciences, St. Petersburg, Russia; https://www.bin-ran.ru/collections/.

**Isotype.** LE A0001562.

**Paratypes.** LE A0001563, LE A0001565, LE A0001567, and LE A0001569 herbarium materials of the strains CALU 1334, CALU 1794, CALU 1801, and CALU 1807, correspondingly; preserved and stored as the holotype.

**Reference strain.** CALU 1759.

**Additional strains.** CALU 1334, CALU 1794, CALU 1801, and CALU 1807.

**Gene sequences.** The following gene sequences deposited at the GenBank NCBI database under the accession numbers MT463169–MT463173 (16S rRNA gene), MT467232–MT467236 (16S–23S ITS), MT411485–MT411489 (rpoCl gene), and MT411494–MT411498 (rbcL gene).

**A survey of the family Geminocystaceae**

The second objective pursued by this communication was to provide *status quo* survey of the family *Geminocystaceae*.

Description of new objects and progress in phylogenetic analysis have led to a re-evaluation of former cyanobacterial system and the establishment of numerous taxa at different nomenclatural levels. It has been constantly questioned whether (or not) the traditional morphology-based system mirrored phylogenetic relationships (Dvořák et al. 2023, Strunecký et al. 2023). In this connection, the understudied family *Geminocystaceae* is among the cyanobacterial taxa of indisputable interest.

**Representatives.** At present, the family *Geminocystaceae* encompasses six genera characterized to varying degrees. For example, although the genus *Microcrocis* (see below) encompasses many phenotypically described representatives, only one species has been genotyped, and cultured strains have not been reported (Pokorny et al. 2023).

The genus *Geminocystis* is represented by the species *G. herdmannii* and *G. papuanica* (Korelusová et al. 2009) as well as by a new species *G. urbisnovae* (this study)—a total number of 18 strains (Table 1).

The genus *Cyanobacterium* (10 phenotypically characterized strains) (Table 2) contains the species *C. stanieri* initially proposed by Ripppka and Cohen-Bazire (1983) and the species *C. aponinum* (Morro et al. 2007). Only few cultured strains were genotyped (Andreote et al. 2014, Deepa 2018, Mobberley et al. 2018), and several draft genomes were deposited at the GenBank database.

The genus *Geminobacterium* contains a single species *G. atlanticum* (Brito et al. 2017).

The genus *Annamia*, in contrast to abovementioned unicellular genera, is represented by filamentous species *A. toxica* (Nguyen et al. 2013) and *A. dubia* (Tuji et al. 2021). Although great distinction between unicellularity and producing trichome was against belonging to one and the same taxon in the outdated phenotypic system (Herdman et al. 2001), modern taxonomy of cyanobacteria issues from the supremacy of phylogeny-based approach. Thus, coexistence of unicellular and filamentous
representatives within the family Geminocystaceae (also, e.g., in the family Cymatolegaceae) (Strunecký et al. 2023) is non-paradoxical in case of shared recent common ancestry. The genus Picosynechococcus was established and described by Komárek et al. (2020) basing on the analysis of several unicellular strains previously attributed to combined genus Synechococcus, and grouped as a separate clade within the order Chroococcales. Although the genus Picosynechococcus was initially attributed to the family Aphanotheceae (Komárek et al. 2020), extended phylogenomic analysis empowered to re-attribute it to the family Geminocystaceae (Strunecký et al. 2023). The genus Microcrocis has been for a long time familiar from field observations. As many as ten species were described at the level of morphology (Pokorný et al. 2023). Recent genotyping of environmental sample phenotypically identified as M. geminata demonstrated corresponding similarities of 16S rRNA gene and the ITS region to those in the family Geminocystaceae (Pokorný et al. 2023, Strunecký et al. 2023). Importantly, the morphology of M. geminata has been described using only environmental samples. The absence of cultured strains has made it impossible to observe the development of colonies.

Morphology. Unicellular Geminocystaceae are assigned to separate genera and species under relatively small morphological distinctions. Cyanobacterium spp. cells were from spherical to rod-like (Moro et al. 2007, Andreote et al. 2014, Deepa 2018, Brenes-Guillén et al. 2021) (Table 2). Cell form and dimensions depended on growth conditions, temperature, and salinity, in particular (Moro et al. 2007, Bravakos et al. 2016).

### Table 2. General characters of cultured Cyanobacterium spp. strains

<table>
<thead>
<tr>
<th>Strain</th>
<th>Origin</th>
<th>Cell shape</th>
<th>Cell width (µm)</th>
<th>Salinity response</th>
<th>Color; phycobiliproteins; chromatic adaptation (CA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyanobacterium aponinum PCC 10605</td>
<td>Thermal springs, Padua, Italy (Moro et al. 2007)</td>
<td>Ellipsoidal; single or in pairs after division</td>
<td>2–3</td>
<td>Tolerates 34 ppt NaCl</td>
<td>Blue-green; phycoerythrin containing; no data on CA</td>
</tr>
<tr>
<td>C. stanieri PCC 7202</td>
<td>Alkaline pond in Chad Republik (Rippka and Cohen-Bazire 1983)</td>
<td>Spherical or rod-like</td>
<td>1.9–4.4</td>
<td>Euryhaline</td>
<td>No data on color; phycoerythrin absent; no data on CA</td>
</tr>
<tr>
<td>C. stanieri LEGE 03274</td>
<td>Freshwater lagoon Lagoa dos Salgados, Albufeira, Portugal (Ramos et al. 2018)</td>
<td>Rod-like to wide oval</td>
<td>1.4–2.2</td>
<td>No data</td>
<td>Blue-green; no data on CA</td>
</tr>
<tr>
<td>C. stanieri (former Synechococcus sp. PCC 8806)</td>
<td>Freshwater lagoon Port Gentil, Carbon, France (<a href="https://catalogue-crbip.pasteur.fr/resultatRecherche.xhtml">https://catalogue-crbip.pasteur.fr/resultatRecherche.xhtml</a>)</td>
<td>Rod-like to wide oval</td>
<td>1–3</td>
<td>Euryhaline</td>
<td>PE absent; no data on CA</td>
</tr>
<tr>
<td>Cyanobacterium sp. CENA 526</td>
<td>Alkaline Lake Grande, Brazil (Andreote et al. 2014)</td>
<td>Rod-like; single or in pairs after division</td>
<td>2.10–2.33</td>
<td>Haloalkalophilic</td>
<td>No data</td>
</tr>
<tr>
<td>Cyanobacterium sp. IKISCC 30</td>
<td>Thermal spring in Lefkada, Greece (Bravakos et al. 2016)</td>
<td>Spherical or rod-like</td>
<td>2–3</td>
<td>Tolerates 35.7 ppt NaCl</td>
<td>Dark green; no data on phycobiliproteins and CA</td>
</tr>
<tr>
<td>Cyanobacterium sp. MBIC 10216</td>
<td>Sakamoto hot spring, Japan (Moro et al. 2007)</td>
<td>Spherical or rod-like</td>
<td>2–3</td>
<td>Tolerates 34 ppt NaCl</td>
<td>Blue-green; low phycoerythrin content; no data on CA</td>
</tr>
<tr>
<td>Cyanobacterium sp. MCCB 114, MCCB 115, and MCCB 238</td>
<td>Vypeen-Periyar river, Cochin, India (Deepa 2018)</td>
<td>Spherical or rod-like</td>
<td>2–4</td>
<td>Tolerates 30–40% NaCl</td>
<td>No data on color; phycoerythrin and allopheocyanin absent; no data on CA</td>
</tr>
<tr>
<td>Cyanobacterium sp. ThrSCSp4</td>
<td>Thermal Spring Thermopiles, Greece (Bravakos et al. 2016)</td>
<td>Spherical or rod-like</td>
<td>2–4</td>
<td>Tolerates 5.5 ppt NaCl</td>
<td>Green; no data on phycobiliproteins and CA</td>
</tr>
<tr>
<td>Cyanobacterium sp. IPPAS B-1200</td>
<td>Bay Kurta, Lake Balkhash, Republic of Kazakhstan (Sarsekeyeva et al. 2014)</td>
<td>Oval or rod-like</td>
<td>2.0–4.5</td>
<td>Haloalkalophilic; tolerates 1 g·L⁻¹ NaCl/16.8 g·L⁻¹ NaHCO₃</td>
<td>Blue-green; no data on phycobiliproteins and CA</td>
</tr>
</tbody>
</table>

http://e-algae.org
The genus *Geminobacterium* is represented by the species *G. atlanticum* has ~2.5 μm cells. The same size was reported for an unidentified nitrogen-fixing strain CLg1 closely similar in 16S rRNA gene sequence to *Geminobacterium atlanticum* (Falcón et al. 2004, Deschamps et al. 2008).

*Geminocystis* spp. had spherical or ellipsoidal 1.5−7.0 μm cells (Andreote et al. 2014, Bravakos et al. 2016, Hirase et al. 2017, Ma et al. 2019, Brenes-Guillén et al. 2021) (Table 1). The majority of described strains (Table 1) can reorient division plane by 90° producing double pairs, or tetrads (Korelusová et al. 2009, Andreote et al. 2014, Bravakos et al. 2016, Brito et al. 2017, Brenes-Guillén et al. 2021). Tetrads have also been observed in *Synechocystis* spp. (Korelusová et al. 2009) and *Altericista* spp. (Averina et al. 2021) where such division mode was considered a diagnostic trait. Although a change of division plane has been sometimes hard to observe, this character was included into formal diagnoses of *G. herdmanii*, *G. papuanaica*, and *G. urbisnovae* (this study).

The species *Anamnia toxica* produces trichome: (1) without necridia, akinetes, and heterocysts; (2) uniseriate; (3) straight; (4) with circular furrows at cross-walls; (5) sheath-less; (6) incapable of false branching; (7) 20–46 μm long in environmental samples, and up to 229 μm long in culture. Intercalary cells were barrel-shaped or cylindrical (1.5−3.5 × 6.3−12.0 μm) while terminal cells were flattened or hemispheric (Nguyen et al. 2013). The morphology of *A. dubia* moderately differs from that of *A. toxica*: straight or waveform trichomes are 30–40 μm long, in environmental samples, and up to several hundred μm, in culture; furrows at cross-walls are absent. Although intercalary cells have approximately the same width as in *A. toxica* (1.2−3.5 μm), their length varies over a broader range (4.6−16.2 μm, occasionally 26.7 μm) (Tuji et al. 2021).

This genus *Picosynechococcus* encompasses small ellipsoidal or rod-like cyanobacteria (1.2−3.0 × 0.8−2.0 μm) occasionally aggregated in short chains; capsule or sheath is not produced.

*Microcrocis* spp. have ellipsoidal or rod-like cells as assembled in flat colonies (like those in *Merismopedia*) consisting of numerous entities encapsulated within a slime layer. Cells in young colonies often align in regular rows; in elder colonies, regularity is lost. Of interest, *Microcrocis* cells divide longitudinally rather than along the transversal plane (Pokorný et al. 2023).

**Ultrastructure.** Unicellular *Geminocystaceae* stay apart from many other unicellular cyanobacteria (Mareš et al. 2019b). Namely, in *G. herdmanii* PCC 6308, *G. papuanaica* PAP1 (Korelusová et al. 2009), and *Geminocystis* sp. CHAB 6541 (Ma et al. 2019) parallel thylakoids stretch from cell pole to pole. Similar pattern has been observed in *G. atlanticum* (Brito et al. 2017), *C. stanieri* (Oren et al. 2022), *M. geminata* (Pokorný et al. 2023), and *G. urbisnovae* (this study).

In contrast, *C. aponinum* (Moro et al. 2007) as well as *Cyanobacterium* sp. MCCB 114, MCCB 115, and MCCB 238 (Deepa 2018) demonstrated a fascicular arrangement of thylakoids assembled in bundles or ribbons interspaced with parallel segments and hemicircular loops. Of note, similar pattern has been observed in many of heterocystous cyanobacteria (Mareš et al. 2019b).

Filamentous systems in filamentous species *A. toxica* and *A. dubia* are similar to those in many filamentous non-heterocystous cyanobacteria as well as in some unicellular strains (Hoffmann et al. 2005). Namely, thylakoids orient perpendicularly to plasma membrane in cross-section micrographs while in longitudinal projection they lie in parallel to it (Nguyen et al. 2013, Tuji et al. 2021).

**Habitats and environmental responses.** Despite environmental niches occupied by *Geminocystaceae* are diverse (Tables 1 & 2), these cyanobacteria have been preferentially isolated from aquatic habitats (except *G. papuanaica* PAP1 originating from soil in a tropical rainy forest) (Korelusová et al. 2009). Whilst many of the *Geminocystaceae* have been encountered in continental freshwater pools, *G. atlanticum* LEGE 07459 (Brito et al. 2017) and closely related strain CLg1 (Falcón et al. 2004) are sea inhabitants.

Although *Geminocystaceae* considerably differ in salinity responses, they preferentially adapt themselves to niches of high salt content (Tables 1 & 2). For example, *Geminocystis* sp. CENA526 (Andreote et al. 2014) and *Geminocystis* sp. CHAB 6541 (Ma et al. 2019) were been isolated from salt lakes. Two closely related strains, *Cyanobacterium* sp. CENA527 (Andreote et al. 2014) and *Cyanobacterium* sp. HL-69 (Mobberley et al. 2018) were isolated from a salt lake and a hypersaline lake, correspondingly. The strains *Cyanobacterium* sp. MCCB 114, MCCB 115, and MCCB 238 originating from brackish water tolerated as much as 40% salt (Deepa 2018). At the same time, *C. stanieri* (Castenholz et al. 2001) and *C. aponinum* (Moro et al. 2007) represent euryhaline cyanobacteria. Also, although *Cyanobacterium* sp. IklSCC30 was isolated from a highly mineralized hot spring, it also grew in BG-11 medium containing little salt (Bravakos et al. 2016). Within the genus *Picosynechococcus* there were not only continental plankton or periphyton inhabitants, but also those which grew in highly mineralized springs.
(Komárek et al. 2020). Finally, species of the genus *Microcrocis* are preferentially freshwater inhabitants, although they have been also found in marine or brackish sediments (Pokorný et al. 2023).

Regarding temperature responses, *Geminocystaceae* often belong to a thermophilic ecotype (Tables 1 & 2). Thus, *Geminocystis* sp. L8, L17, L18, L34, and L29 were isolated from a hot spring (Brenes-Guillén et al. 2021). Other hot spring strains encompass *C. aponinum* PCC 10605 (Moro et al. 2007), *Cyanobacterium* sp. MBIC10216 (Blank and Hinman 2016), and *Cyanobacterium* sp. CR 23M (Brenes-Guillén et al. 2021). Less frequently it has been mesophilic ecotype, e.g., in *G. herdmanii* PCC 6308 (Korelusová et al. 2009), *Geminocystis* sp. NIES-3708, *Geminocystis* sp. NIES-3709 (Hirose et al. 2015a, 2015b), and *G. urbisonvae* (this study). A rare exception is several strains characterized as the psychrotrophic or psychrophilic cyanobacteria. For instance, the strain *Geminocystis* sp. KT25 isolated from an Alaskan tundra lake (Strunecký et al. 2020) most possibly represented a psychrotrophic cyanobacterium. According to 16S rRNA gene similarity, it could be assigned to a new species *G. urbisonvae* (this study), but, unfortunately, it was lost prior to a detailed description.

**Phycobiliprotein pigments and complementary CA.** PE and PC composition in *Geminocystaceae* has been insufficiently studied. It also concerned the complementary CA, i.e., structural rearrangement of the phycobilisome in response to preferential illumination with either green or red light (Tandeau de Marsac 1977). Very often a capability of CA has not been tested at all (Tables 1 & 2). In those cases, when a corresponding analysis was performed, CA of the second type (CA2) (Hirose et al. 2019), i.e., regulation of PE but not PC, has been observed.

**Molecular phylogeny.** In agreement with polyphasic approach (Vandamme et al. 1996), *Geminocystaceae* have been analyzed using phylogenetic methods. With 16S rRNA gene sequence as prime record, this family always produced a separate cluster in evolutionary trees reconstructed with different methods (Tuji et al. 2021). Monophyly of *Geminocystaceae* was confirmed in our extended analysis (Figs 3 & 4, Supplementary Fig. S4) also showing that within the same name cluster there were three clades designated as: (1) the genus *Geminocystis*; (2) the “*Cyanobacterium stanieri*” subcluster; and (3) the “*C. aponinum*” subcluster. This cluster additionally incorporated two distantly diverged small clades. One of them was represented by the species *A. dubia* and *A. tosica* while the other contained the species *G. atlanticum* and the strain CLg1. Taking into account the established minimal 98.7% similarity of 16S rRNA gene within a prokaryotic species (Kim et al. 2014), and considering 95.5% similarity between *G. atlanticum* and the cyanobacterium CLg1 (Supplementary Tables S2–S4), the latter strain represented a separate species or even genus.

16S rRNA gene sequences in the genus *Geminocystis* were distributed between two clades (Fig. 3). One of them encompassed the strains CALU 1334, CALU 1759, CALU 1794, CALU 1801, and CALU 1807 (all of them being attributed to a new species *G. urbisonvae*; this study) as well as the strain *Geminocystis* sp. KT25. The other clade contained the species *G. herdmanii*, *G. papuanica*, and *Geminocystis* sp. strain 1.1 as well as *C. aponinum* (96.3, 96.0, and 95.5% similarity, correspondingly) (Supplementary Tables S2–S4). Hence, these strains could be reassigned to the genus *Geminocystis*.

A most interesting case is *Cyanobacterium* 16S rRNA gene sequences showing high heterogeneity of this genus which radiated into the subclusters “*Cyanobacterium stanieri*” and “*C. aponinum*” (Fig. 3). The former subcluster included *C. stanieri* PCC 7202 (type strain) (Oren et al. 2022), the strains AICB1016 and PCC 8806 (deposited in the databases as *Synechococcus* sp.) as well as *Cyanobacterium* sp. strains CR_23M, UMPCCC 1230, CENA527, and HL-69. The latter subcluster encompassed the strain *C. aponinum* PCC 10605 as well as *Cyanobacterium* sp. MCCB 114, MCCB 115, MCCB 238, and MBIC10216.

Importantly, the intra-species 16S rDNA similarity criterion of 98.7% (Kim et al. 2014) empowered a taxonomic reattribution of some strains from the “*Cyanobacterium stanieri*” subcluster. Thus, *Cyanobacterium* sp. CENA527, *Synechococcus* sp. AICB1016, and *Synechococcus* sp. PCC 8806 (correspondingly, 99.2, 98.7, and 98.8% similarity to *C. stanieri* PCC 7202) (Supplementary Tables S2–S4) should in fact be considered *C. stanieri* strains. At the same time, *Cyanobacterium* sp. HL-69 (97.9% similarity to *C. stanieri*) and *Cyanobacterium* sp. UMPCCC 1230 (96.3% similarity to *C. stanieri*) (Supplementary Tables S2–S4) would belong to a separate species. In turn, the strain *Cyanobacterium* sp. CR_23M occupied a position outside the “*Cyanobacterium stanieri*” subcluster (92.8% similarity) and preferentially associated with the genus...
Geminocystis (94.2% similarity to G. herdmani PCC 6308, and 95.1% similarity to Geminocystis sp. NIES-3708) (Fig. 3, Supplementary Tables S2–S4). Furthermore, despite taxonomic status of Cyanobacterium sp. CENA169 was considered indefinite (Silva et al. 2014), this strain does not at any rate belong to the genus Geminocystis (16S rRNA gene similarity to the strains of this genus was <92%) (Supplementary Tables S2–S4).

Phylogenetic relationships within the genus Cyanobacterium could be additionally evaluated using Bayesian and RaxML reconstructions with 16S rRNA + rbcL + rpoC1 genes concateners (Fig. 4, Supplementary Fig. S4). Although a relatively small set of sequences was analyzed (Supplementary Table S5), reliability of two Cyanobacterium subclusters as well as C. stanieri and C. aponinum tree locations were confirmed.

Diversity of Geminocystaceae could be also traced using accessory phylogeny markers such as nucleotide sequence and folding of the 16S–23S ITS of the rrn operon. For this purpose, total spacer length, V2 domain length, D1–D1′ and B box domains folding and the built-in RNA genes were examined (Fig. 4, Supplementary Table S1), as well as the V3 region length.

Additionally, ITS sequences of Geminocystaceae were compared with those in sister branch Picosynechococcus spp. Total ITS length was the lowest in C. aponinum PCC 10605 (312 bp) and the highest in A. dubia NIES-4383 (434 bp). Besides, ITS top length in two Cyanobacterium subclusters was different, namely 312–313 bp in “Cyanobacterium aponinum” and 340–344 bp in “Cyanobacterium stanieri.” In the genus Geminocystis it ranged from 330 to 368 bp. In all the strains of the species G. urbisnovae (this study) it was 360 bp taking close value with Geminocystis sp. NIES-3708 (368 bp) (Fig. 5). D1–D1′ domains were closely similar in all Geminocystaceae except A. dubia NIES-4383. Importantly, in Geminocystis spp. this domain was distinguished by the first four conserved nucleotides GACU/GAUU (instead of GACC in other cyanobacteria); additionally, D1–D1′ domains in Geminocystis spp. were shorter (16–19 bp) than in other cyanobacteria (Itteman et al. 2000). D1–D1′ domains in two Cyanobacterium subclusters had the same length as in Geminocystis spp., and their sequence varied within the “Cyanobacterium stanieri” subcluster (Fig. 5B). B box in Geminocystis spp. showed identical folding (Fig. 4B) except G. herdmannii PCC 6308 where this domain had a strain-specific sequence. In the case of Cyanobacterium spp., B box folding was uniform notwithstanding a certain distinction in conserved GGCA and AGCA sites. In the genera Annamia and Geminobacterium, B box sequence and folding were more specific (Fig. 5B). Finally, within the ITS region of all Geminocystaceae, tRNA\" and tRNA\# were encoded. For comparison, the ITS region in Picosynechococcus spp. contained an extended D1–D1′ domain which was identical in the strains PCC 7002, PCC 8007, and NIES-970 (Fig. 5B) while B box domains in these strains differed (Fig. 4B).

All strains of G. urbisnovae possessed an identical, very short V3 region. At the same time, V3 region was absent in Geminocystis papuanica (EF555569), Geminobacterium atlanticum (KR676352), and Microcrocis geminata (ON426371–ON426380). In other genera, V3 region was too short to be folded into secondary structure (Korelsová et al. 2009, Brito et al. 2017). Because comparative analysis of V3 region in aforementioned cases was impossible, we did not add corresponding data to Supplementary Table S1.

**Taxonomic items.** A prime challenge has been posed by the genus Cyanobacterium. It should be noted that initial proposal of this genus, as well as of the type species C. stanieri (Rippka and Cohen-Bazire 1983) was invalid under the ICN because of absent Latin diagnosis and reference to herbarium holotype. Next proposal of the family Cyanobacteriaceae on the basis of a polyphasic approach (Hoffmann et al. 2005, Komárek et al. 2014) was emended (Tuji et al. 2021), and the family Geminocystaceae including in particular the validly published genus Cyanobacterium was established.

Valid proposal of the genus Cyanobacterium gen. nov., and the type species C. stanieri sp. nov. under the provisions of the ICN (Oren et al. 2022) was reinforced with ample information about morphology, ultrastructure, and chromosomal / genomic characters; detailed supplementary materials encompassed the issue of genome mining for genes involved in the biosynthesis of carotenoids and phycobiliproteins.

In our opinion, the genus Cyanobacterium represents a heterogeneous assemblage because of high branching of the “Cyanobacterium stanieri” and the “Cyanobacterium aponinum” subclusters (Fig. 3). Their divergence was supported by relatively low 16S rRNA gene similarity between these clades (~93%) (Supplementary Tables S2–S4). Additionally, they differed in the structure of the 16S–23S ITS region (Fig. 5). Unlike dissimilar strains within the “Cyanobacterium stanieri” subcluster, high similarity between the “Cyanobacterium aponinum” subcluster strains (99.2–99.9%) empowers to ascribe them to one and the same species C. aponinum (Supplementary Tables S2–S4). In view of distance from the “Cyanobacterium stanieri” subcluster, the “Cyanobacterium aponinum...
"num" subcluster could claim the range of an independent genus. In contrast to the 98.7% consensus value for species delimitation on the basis of 16S phylogeny (Kim et al. 2014), a cut-off for genera has for a long time been obscure (see Kämpfer and Glaser 2012) until the 94.5% threshold was finally established (Yarza et al. 2014). Anyway, new genus could be proposed if a strain (or a group of strains) is located outside the cluster of validly published genus, and separate taxonomic position is substantiated with phenotypic characters missing in neighboring genera (Stackebrandt 2006, Johansen et al. 2021). Therefore, proposal of a separate genus on the basis of the "Cyanobacterium aponinum" subcluster is premature because of poor phenotypic data.

Regarding Picosynechococcus, 16S (Fig. 3) and concatenated Fig. (Fig. 4, Supplementary Fig. S4) dendrograms demonstrated that this genus was a sister clade to the cluster Geminocystaceae. Additionally, 16S rRNA gene similarity between the analyzed Picosynechococcus strains (NIES 970, PCC 7002, PCC 8807, UH7, and HOG) and other genera of Geminocystaceae was relatively low (89.4–91.2% in the case of Geminocystis spp., 89.6–91.2% in the case of the subcluster "Cyanobacterium aponinum", and 87.8–89.9% in the case of the subcluster "Cyanobacterium stanieri"). Noteworthy, comparable similarity values with aforementioned subclusters had the genera Annamia and Geminobacterium. Moreover, total length of the ITS region in Picosynechococcus was ~500 bp, i.e., a top value in Geminocystaceae except Annamia dubia NIES-4383 (Supplementary Table S1). Additionally, the strains PCC 7002, PCC 8007, and NIES 970 of the Picosynechococcus clade had a long, specific, and mutually identical D1–D1’ domain. Thus, in our opinion, the genus Picosynechococcus could be included in the family Geminocystaceae at least provisionally, i.e., for the same reason with the genera Annamia and Geminobacterium.

To conclude, application of polyphasic approach to freshly isolated strains and / or the strains stored in microbial collections should bring a deeper insight into the environmental diversity of cyanobacteria. Obviously, the comparative genomics research on different Geminocystaceae would be required to promote and improve the taxonomy of this family. Unfortunately, such efforts have been limited. Despite a prime reliance on genotypic characters, detailed phenotypic description is urgently required that is exemplified by a moderate progress in research of the family Geminocystaceae.

ACKNOWLEDGEMENTS

The authors thank O. Gavrilova for long-term collaboration, S. Smirnova for a supply of plankton samples, and N. Velichko for a help with transmission electron microscope. We thank St. Petersburg University Research Centers “Cultivation of Microorganisms”, “Molecular and Cell Technologies”, and “Chromas” for strains maintenance, analytical assistance, and genes sequencing. We especially thank the anonymous reviewers for their criticism and valuable suggestions that greatly helped us in upgrading and improvement of the manuscript.

CONFLICTS OF INTEREST

The authors declare that they have no potential conflicts of interest.

SUPPLEMENTARY MATERIALS

Supplementary Table S1. 16S–23S internal transcribed spacer details in Geminocystaceae (https://www.e-algae.org).

Supplementary Table S2. Pairwise identity (%) of 16S rRNA gene sequences in the Geminocystaceae strains (https://www.e-algae.org).

Supplementary Table S3. Pairwise identity (%) of 16S rRNA gene sequences in the “Cyanobacterium stanieri” subcluster (https://www.e-algae.org).

Supplementary Table S4. Pairwise identity (%) of 16S rRNA gene sequences in the “Cyanobacterium aponinum” subcluster (https://www.e-algae.org).

Supplementary Table S5. GenBank accession numbers for the genes used in phylogenetic analysis (https://www.e-algae.org).

Supplementary Fig. S1. General view of Geminocystis urbisnovae sp. nov. colonies grown for 3–4 weeks on modified BG-11 medium solidified with 1.2% Difco agar (https://www.e-algae.org).

Supplementary Fig. S2. Culture growth graphs for Geminocystis urbisnovae sp. nov. Strain CALU 1334 (A), strain CALU 1759 (B), strain CALU 1794 (C), strain CALU 1801 (D), and strain CALU 1807 (E) (https://www.e-algae.org).

Supplementary Fig. S3. Light absorbance spectra for Geminocystis urbisnovae sp. nov. cultured under green (green line) or red (red line) filter (https://www.e-algae.org).
Supplementary Fig. S4. RaxML phylogenetic tree for the 16S rRNA gene + rbcL gene + rpO C gene concatenator of 89 sequences (https://www.e-algae.org).

REFERENCES


Polyakova et al. Geminocystis urbisnovae sp. nov.


