Aliterella atlantica gen. nov., sp. nov., and *Aliterella antarctica* sp. nov., novel members of coccoid *Cyanobacteria*

Janaina Rigonato,[1] Watson Arantes Gama,[2] Danillo Oliveira Alvarenga,[1] Luis Henrique Zanini Branco,[3] Frederico Pereira Brandini,[4] Diego Bonaldo Genuário[1] and Marli Fatima Fiore[1]

Correspondence
Marli Fatima Fiore
fiore@cena.usp.br

[1]University of São Paulo, Center for Nuclear Energy in Agriculture, Avenida Centenário 303, 13400-970 Piracicaba, SP, Brazil

[2]Institute of Botany, Nucleus of Phycology, Avenida Miguel Estéfano 3687, 04301-012 São Paulo, SP, Brazil

[3]São Paulo State University UNESP, Institute of Bioscience, Languages and Exact Sciences, São José do Rio Preto, São Paulo, Brazil

[4]University of São Paulo, Oceanography Institute, São Paulo, Brazil

Two *Cyanobacteria* isolated from South Atlantic Ocean continental shelf deep water and from a marine green algae inhabiting the Admiralty Bay, King George Island, Antarctica were investigated based on morphological and ultrastructural traits, phylogeny of 16S rRNA gene sequences, secondary structure of the 16S-23S internal transcribed spacer regions and phylogenomic analyses. The majority of these evaluations demonstrated that both strains differ from the genera of cyanobacteria with validly published names and, therefore, supported the description of the novel genus as *Aliterella* gen. nov. The identity and phylogeny of 16S rRNA gene sequences, together with the secondary structure of D1D1′ and BoxB intergenic regions, further supported the two strains representing distinct species: *Aliterella atlantica* gen. nov., sp. nov. (type SP469036, strain CENA595^T) and *Aliterella antarctica* sp. nov. (type SP469035, strain CENA408^T). The phylogenomic analysis of *A. atlantica* sp. nov. CENA595^T, based on 21 protein sequences, revealed that this genus belongs to the cyanobacterial order *Chroococcidiopsidales*. The isolation and cultivation of two geographically distant unicellular members of a novel cyanobacterial genus and the sequenced genome of the type strain bring new insights into the current classification of the coccoid group, and into the reconstruction of their evolutionary history.

The simplest morphologically coccoid *Cyanobacteria* have often been overlooked and their genetic complexities underestimated. The traditional morphology-based taxonomic system limited the description of coccoid *Cyanobacteria* due to the low degree of morphological variation. Likewise, there was a prevailing notion that larger genomes are needed to support multicellular *Cyanobacteria* than are needed to support single-celled *Cyanobacteria* (Herdman *et al.*, 1979). However, this notion has not always been found to be true, as several unicellular *Cyanobacteria* have a larger genome size than multicellular *Cyanobacteria* (Larsson *et al.*, 2011; Shih *et al.*, 2013). Multicellularity arises as a result of the incomplete separation of daughter cells after cell division, and the behaviour of cells within multicellular structures is coordinated by both shared and unique molecular mechanisms (Rossetti *et al.*, 2010; Claessen *et al.*, 2014). It has been shown that mutations of the cell wall *amiC2* gene in the filamentous heterocytous strain *Nostoc punctiforme* ATCC 29133, encoding an amidase involved in septa formation, induced colonial unicellular morphology in this cyanobacterium, and prevented heterocyte differentiation (Lehner *et al.*, 2011). However, in

Abbreviations: ITS, Internal transcribed spacer; NJ, Neighbour-joining; ML, Maximum-likelihood; BA, Bayesian.

The GenBank/EMBL/DDBJ accession numbers for the genome of *Aliterella atlantica* CENA595^T is JYON01000000 and for the 16S-23S rRNA gene sequence of *Aliterella antarctica* CENA408^T is KU291459.

Two supplementary figures are available with the online Supplementary Material.

the unicellular cyanobacterium, *Synechococcus elongatus* PCC 7942, filament formation was induced by overexpressing the cell division protein FtsZ (Mori & Johnson, 2001). These results indicate that the expression level of cell division proteins are correlated with filament formation, which may be confirmed in the future, as until now no candidate genes involved in cellular morphology have been found in cyanobacterial genomes (Dagan *et al.*, 2013). Even unicellular *Cyanobacteria* frequently adopt transient multicellularity, spending at least part of their lives in temporarily adhered clusters of cells (colonies). Transient multicellularity may occur to promote benefits associated with increased size, to mitigate predation risks or nutrient stressors, and these are associated with distinct evolutionary costs and benefits (Bonner, 1998; Grosberg & Strathmann, 2007). True multicellularity in *Cyanobacteria* has been lost and regained several times during their life history, which has led to the current distribution of both unicellular and multicellular cyanobacterial forms (Tomitani *et al.*, 2006; Schirrmeister *et al.*, 2011, 2015). Therefore, the traditional cyanobacteria classification, based on the most obvious phenotypic features (i.e. cell morphology) does not reflect their evolution and the strength of relationships among taxa, which is the basis of modern systematics (Honda *et al.*, 1999; Robertson *et al.*, 2001; Kim *et al.*, 2014; Komárek *et al.*, 2014). Moreover, advancements in cyanobacterial genome sequencing have enabled large-scale multi-gene phylogenetic analyses that have provided a robust framework to solve deep-branching relationships and help to elucidate the evolutionary history of these phototrophic bacteria (Sánchez-Baracaldo *et al.*, 2005; Blank & Sánchez-Baracaldo, 2010; Larsson *et al.*, 2011; Shih *et al.*, 2013; Bombar *et al.*, 2014; Calteau *et al.*, 2014; Chrismas *et al.*, 2015; Schirrmeister *et al.*, 2015). So far, there are approximately 300 cyanobacterial genomes available in public databases among complete, draft and incomplete genomes, which represent less than 1% of the total bacterial genomes sequenced (https://gold.jgi-psf.org/index, accessed on December 9, 2015). Even so, the phylogenomic approach has been used in a new proposal for cyanobacterial classification (Komárek *et al.*, 2014). Improving cyanobacterial genome sequencing and taxon sampling will contribute to resolving their evolutionary relationships and classification. In this study, investigations were carried out on two underexplored environments, i.e. South Atlantic Ocean water and King George Island in the Antarctic Continent. Two cyanobacterial strains unaffiliated with any of the known genera were recovered. In order to classify these strains under the provisions of the International Code of Nomenclature for algae, fungi and plants, a polyphasic approach was applied, based on morphological and ultrastructural traits, the phylogeny of the 16S rRNA gene, internal transcribed spacer (ITS) secondary structures and phylogenomic analysis.

Strain CENA595^T was isolated from the water column, at the level coinciding with the deep chlorophyll maximum layer (ca. 107 m) in a fixed station over the continental shelf of south-eastern Brazil (25° 15.595′ S 45° 07.67′ W) during an expedition of the R/V Alpha Crucis of the University of São Paulo, on November 03, 2012 (Rigonato *et al.*, 2015). Strain CENA408^T was isolated from a sample of a green turf alga growing together with decomposed *Deschampsia* grass on a ornithogenic soil (*Pygoscelis adeliae*), collected close to the sea, on January 21, 2009, in Arctowski (62° 09.76′ S 58° 27.76′ W), Admiralty Bay, King George Island and South Shetlands Archipelago, Antarctica. Subsamples of the material collected were spread out onto agar solidified BG-11 medium (1.2%, w/v) (Allen, 1968) containing cycloheximide (70 mg l^{-1}) (Rippka, 1988). A monospecific culture was obtained by successive streaking onto fresh solid media, followed by optical microscope observations. In all of the isolation steps, the culture was grown with a 14 : 10 h light/dark cycle under white fluorescent irradiance (40 μmol photons $\mu m^{-2} s^{-1}$) at 24 ±1 °C. Morphological observations were conducted using a Zeiss Axioskop 40 optical light microscope equipped with an AxioVision LE 4.6 digital imaging system (Carl Zeiss). Aliquots of lineages CENA408^T and CENA595^T, fixed in 4 % (v/v) formaldehylde, were deposited in Maria Eneyda P. Kauffaman Fidalgo Herbarium at the Institute of Botany (São Paulo, Brazil), under the identification numbers SP469035 and SP469036, respectively. In addition, cultures of the strains are maintained in the Center for Nuclear Energy in Agriculture Collection/University of São Paulo (CENA/USP), Brazil, in BG-11 medium (Allen, 1968) under white fluorescent light (40 μmol photons $m^{-2} s^{-1}$) with a 14 : 10 h light:dark (L : D) cycle at 25 °C±1 °C.

The morphology of the cells was investigated with a light microscope (Olympus; BX53 DIC) equipped with an Olympus DP71 imaging system and the mean cell size of 50 cells was measured for each strain. The morphological features observed were compared to taxa with validly published names in the literature (Komárek & Anagnostidis, 1998; Komárek *et al.*, 2014).

The procedure for ultrastructural examination was carried out by fixing the cells with Karnovsky's fixative solution (Karnovsky, 1965) for 64 h at 4 °C. Subsequently, the cells were washed with 0.05 M sodium cacodylate buffer and post-fixed with 1 % osmium tetroxide 0.1 % (w/v) sodium cacodylate for 1.5 h at room temperature. Afterwards the cells were washed with distilled water and pre-stained with 2.5 % uranyl acetate at 4 °C overnight, followed by sequential dehydrations with 30 %, 50 %, 70 %, 90 % and 100 % (v/v) acetone for 10 min each. Spurr's resin (Electron Microscopy Sciences) was used for the embedding of samples (Spurr, 1969) and polymerization occurred at 65 °C after 72 h. The resulting resin blocks were cut into 70 nm thin sections in a Leica Ultracut UCT ultramicrotome (Leica Microsystems), transferred to 200-mesh copper grids, and stained with uranyl acetate and lead citrate. Visualization and photography were performed under a transmission electron microscope (Carl Zeiss; EM 900) at 50 kV.

Total genomic DNA of the strains was extracted using a PowerSoil DNA extraction kit (MoBio Laboratories). The 16S rRNA gene plus 16S-23S ITS region was obtained by

PCR using the set primer 27F-23S30R (Taton *et al.*, 2003). The PCR products were cloned, sequenced and the reads were assembled into contigs using the Phred/Phrap/Consed package (Ewing & Green, 1998; Ewing *et al.*, 1998; Gordon *et al.*, 1998). The sequences obtained were compared with sequences available at the National Center for Biotechnology Information (NCBI) database using the BLAST tool (http://www.ncbi.nlm.nih.gov/BLAST). Selected cyanobacterial sequences retrieved from the NCBI were used for phylogenetic inferences as additional sampling of taxa. The alignment was performed using the CLUSTAL W (Thompson *et al.*, 1994) tool from MEGA 6 (Tamura *et al.*, 2013). The best-fitting evolutionary model, GTR+G+I, was selected by ModelTest under Akaike information criterion (Posada & Crandall, 1998). The 16S rRNA gene sequence phylogenetic trees were inferred by neighbour-joining (NJ), maximum likelihood (ML) and Bayesian (BA) criteria. NJ and ML analyses were conducted using MEGA 6 with 1000 bootstraps. BA analysis was run in MrBayes 3 (Ronquist & Huelsenbeck, 2003) in 5×10^6 generations. Chains were sampled every 100th cycle and 25 percent of the sampled trees were discarded as burn-in. Convergence of the MCMC algorithm was monitored by the average standard deviation of split frequencies (<0.003). A similarity matrix was built comparing the 16S rRNA sequences using MEGA 6. The secondary structure of the intergenic region, 16S-23S rRNA, was obtained using the Mfold webserver (Zuker, 2003), and the presence of tRNA in the intergenic spacer was verified by tRNAscan-SE 1.21 (Lowe & Eddy, 1997; Schattner *et al.*, 2005).

The genome of strain CENA595^{T}has been previously sequenced (Rigonato *et al.*, 2015). The phylogenomic tree was inferred based on concatenation of 21 protein sequences selected from bacterial markers (Wu & Eisen, 2008; Calteau *et al.*, 2014), aligned to 126 cyanobacterial genomes using Clustal Omega (Sievers *et al.*, 2011). The LG+I+G model of evolution was selected using ProtTest 2.4 (Abascal *et al.*, 2005), and the ML tree was inferred using RAxML (Stamatakis, 2014) with 1000 bootstraps.

Morphological analysis showed that both strains grown in cyanobacterial culture media presented with a unicellular thallus; cells were solitary or organized in irregularly formatted groups, were blue-green and had a rigid mucilage. The cell shape was spherical or irregular-round and the dimensions ranged from 3.0–4.8 μm long and 2.0–3.3 μm diameter for CENA595^T, and 3.4–5.8 μm long and 2.5–4.6 μm diameter for CENA408^T. Cell division was by binary fission in three or more planes (Figs 1 & S1, available in the online Supplementary Material). By light microscopy the periplasm and nucleoplasm of the cells were noticable, suggesting an accumulation of thylakoids in the peripheral regions (Fig. 1h arrows). However, the ultrastructure (TEM) showed that the thylakoids could be irregularly distributed in the cell cytoplasm (Fig. 2a, c and e) or parallel to the cell wall (Fig. 2b, f and g), and in this case were in agreement with the observations made by light microscopy. Interestingly, the presence of circular dense structures of around 41–55.5 nm in diameter were noted inside the cells of strain CENA595^T (Fig. 2h–l). Comparison of these structural features with others from the literature indicate a resemblance to viruses (Baudoux & Brussaard, 2005; Gao *et al.*, 2012). This assumption is reinforced by the fact that viral sequences were found in the sequenced genome of CENA595^T (data not shown).

Blast analysis of 16S rRNA gene sequences of the two novel strains showed that similarities vary from 89.0 to 94.9 % with sequences of coccoid *Cyanobacteria* available in GenBank. The highest 16S rRNA gene sequence similarities were found with sequences of uncultured bacteria isolated from diverse habitats, such as forest soil from China (KF037769), clean rooms where spacecrafts are assembled (DQ532167; Moissl *et al.*, 2007) and skin microbiome of popliteal fossa of children (JF167907; Kong *et al.*, 2012). Furthermore, phylogenetic analysis grouped the two novel strains with these uncultivated bacteria together with two sequences of uncultured *Cyanobacteria* [FN811217, obtained from a soil sample from Alexander Island in the Antarctic (Chong *et al.*, 2012) and JX127185, recovered from biofilms on building stones in Germany (Hallmann *et al.*, 2013)] with 100 % of bootstrap (Fig. 3). These results indicate that strains CENA595^T and CENA408^T are members of a novel cyanobacterial genus, possibly with a broad distribution. Based on the few sequences so far accessed most of the members of the novel genus occur in terrestrial (soil/mineral) habitats, although CENA595^T was isolated from oceanic waters. Therefore, the novel genus *Aliterella* gen. nov. is proposed under the provisions of the International Code of Nomenclature for algae, fungi and plants. Another point to be highlighted is that Blast analysis revealed low identities ($\leq$94 %) of CENA595^T and CENA408^T 16S rRNA gene sequences with sequences from GenBank. The best hit (94 % identity, 100 % coverage) was observed with *Gloeocapsa* sp. PCC 7428 isolated from the water of a moderately hot spring in Sri Lanka, which was separated from the *Aliterella* group in the phylogenetic analysis. The *Aliterella* sister group encompassed two sequences in the phylogenetic analysis denominated as *Chroococcus* (FR798925; FR798931) and one as *Chroococcidiopsis* CC1 (DQ914863), sharing among 93–95 % similarity with CENA595^T and CENA408^T, based on pairwise nucleotide similarity estimated according to Kimura's 2-parameter model (Kimura, 1980). However, these were probably misidentified, as they were placed far from the *Chroococcus* (Kováčik *et al.*, 2011) and *Chroococcidiopsis* (Waterbury, 1989) type groups in trees. Despite the relationship with these few unicellular taxa with uncertain identification, Blast and phylogenetic analyses indicated that CENA595^T and CENA408^T 16S rRNA gene sequences are closely related to sequences of multicellular strains of nostocalean, especially of the genus *Camptylonemopsis*. An early phylogenetic study showed that strains assigned to the genus *Chroococcidiopis* were the closest coccoid relatives to members of the order *Nostocales* (Fewer *et al.*, 2002). In this study, however, phylogenetic analysis based on 16S rRNA gene

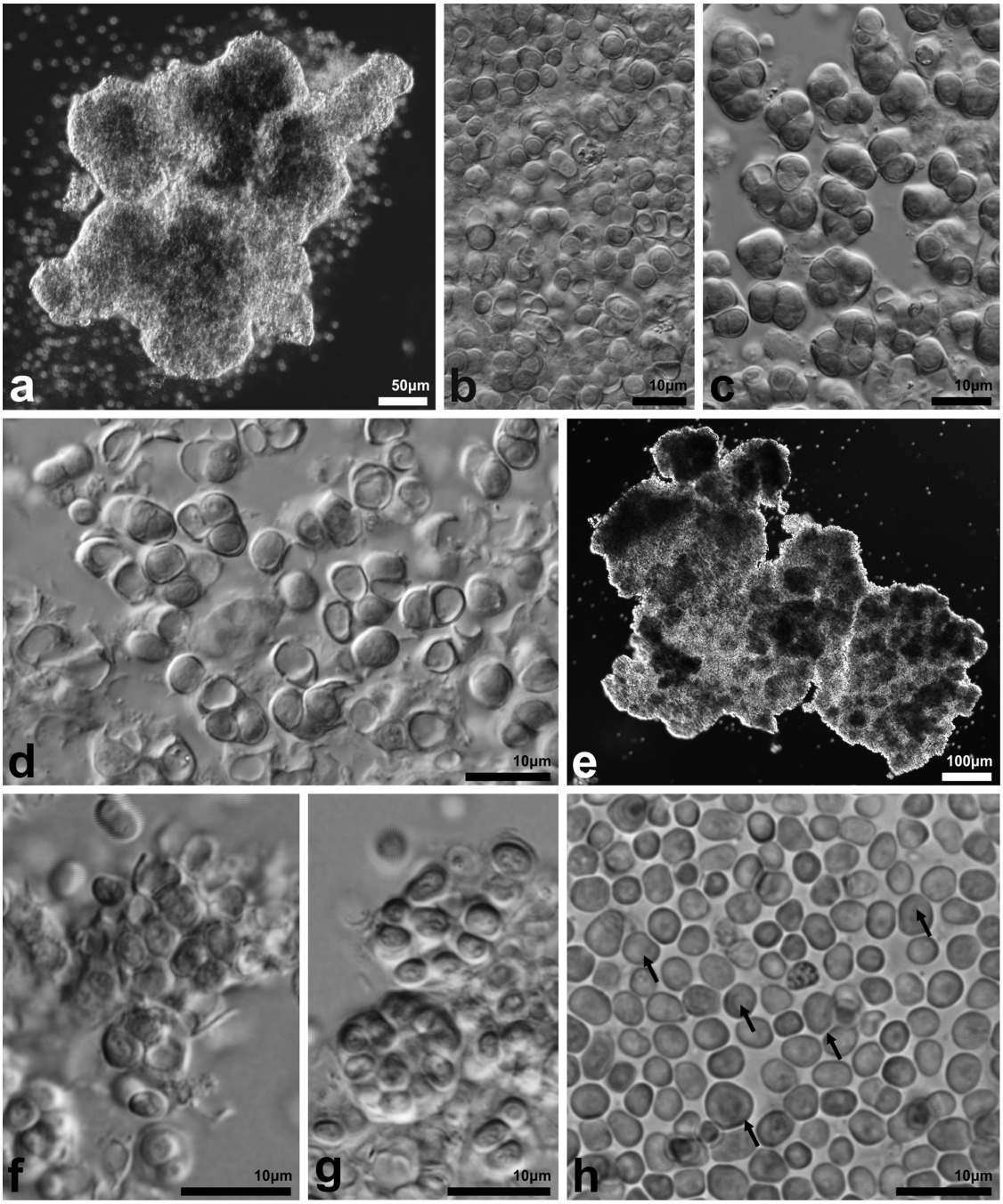

Fig. 1. Photomicrographs of species of *Aliterella*. a–d, *A. atlantica* CENA595^T gen. nov., sp. nov; e–h, *A. antarctica* CENA408^T sp. nov. A coloured version of this figure is available (Fig. S1).

sequences showed that coccoid strains of *Aliterella, Chroogloeocystis* (Brown *et al.*, 2005; Ramos *et al.*, 2010) and *Gloeocapsopsis* (Azua-Bustos *et al.*, 2014) were more closely related to members of the *Nostocales* than of the true *Chroococcidiopis* group (Fewer *et al.*, 2002).

Strains CENA595^T and CENA408^T shared 98.3 % similarity between their 16S rRNA gene sequences and each one is positioned in a different branch inside the clade in the phylogenetic tree, which is evidence of two distinct species. Moreover, despite both strains having 16S-23S rRNA intergenic space sequences for alanine and isoleucine tRNA, configuration regions and nucleotides of 16S-23S rRNA D1-D1′ and BoxB secondary structures (Fig. 4) also suggest that they are two different species. The basal and loop portions of the D1-D1′ secondary structures were equal for both strains, while the central portion of the helix presented

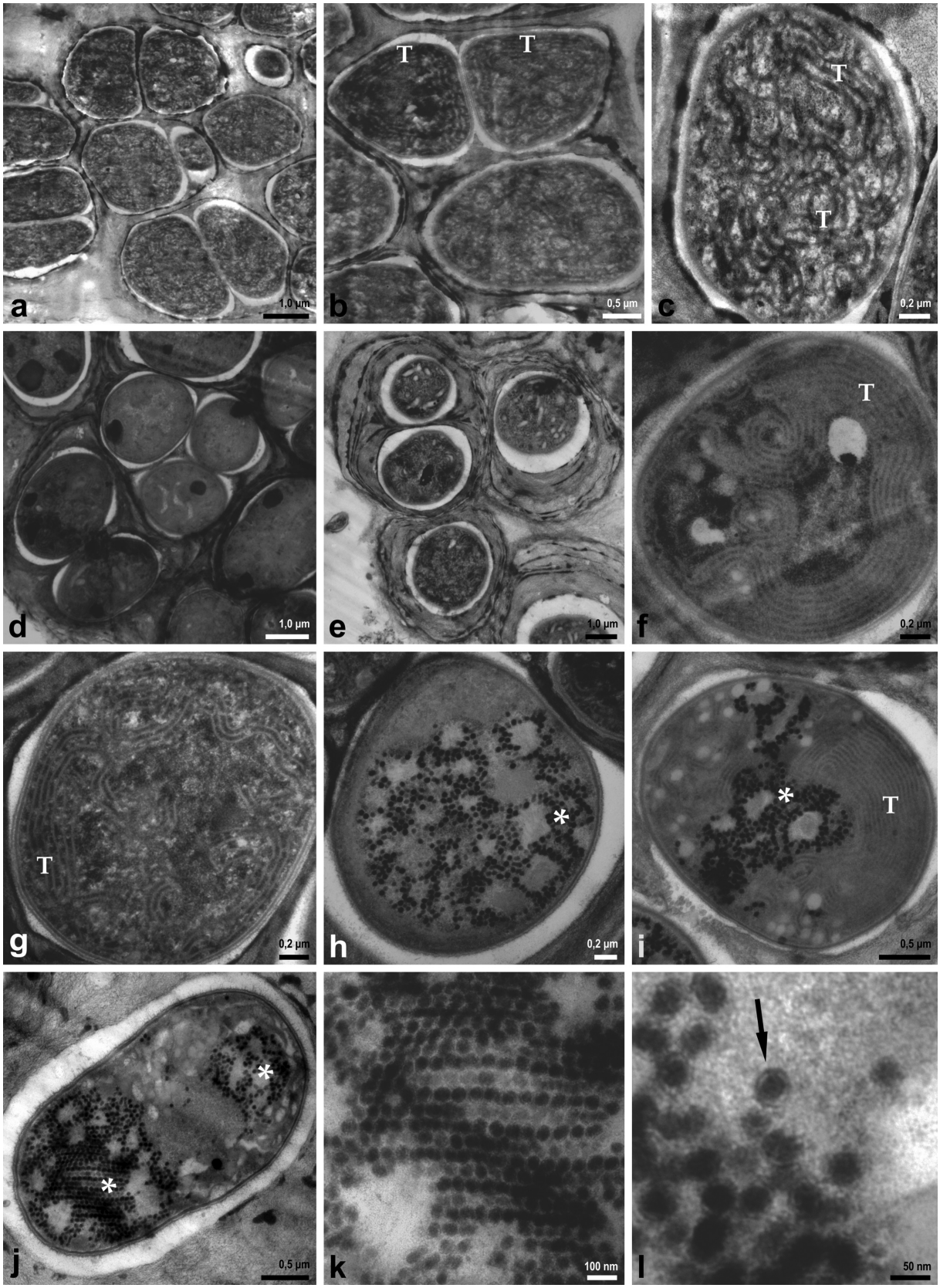

Fig. 2. Ultrastructure of cells of *A. antarctica* CENA408^T gen. nov., sp. nov. (a–c), and *A. atlantica* CENA595^{T}sp. nov. (d–l). H–L highlight the dense structures within the cells that are proposed to be viruses.

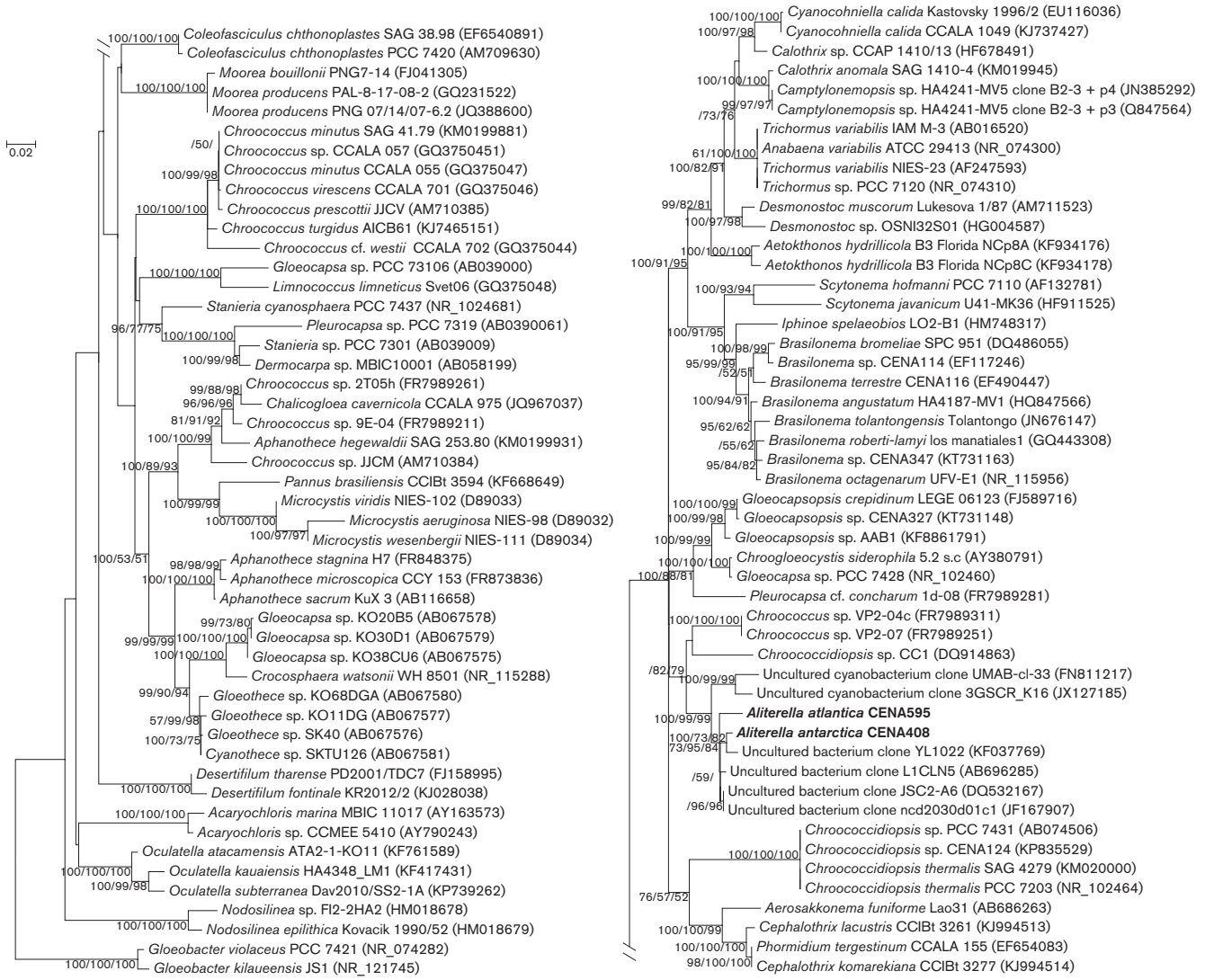

Fig. 3. 16S rRNA phylogenetic tree, inferred by the maximum likelihood method, GTR+I+G evolutive model. Numbers in nodes represent the NJ, ML bootstrap and BA posterior probability, respectively. Bar, 0.01 substitutions per nucleotide position.

two bubbles that differed in position and in nucleotide numbers (Fig 4a, b). With respect to BoxB, its basal portion in the CENA595^T helix is two nucleotide pairs longer than in CENA408^T, and at its central portion, an irregular bulge occurred in CENA595^T, but was not observed in CENA408^T (Fig. 4c, d). These differences in the 16S rRNA sequences and 16S-23S rRNA ITS, and also the distinct ecological origin of strains CENA595^T and CENA408^T, support the description of two novel species, *Aliterella atlantica* gen. nov., sp. nov. and *Aliterella antarctica* sp. nov.

Previous studies have reported a total genome size of 5.2 kb for *A. atlantica* CENA595^T (Rigonato *et al.*, 2015), which exceeds the genome sizes of some multicellular *Cyanobacteria* and indicates a higher degree of organizational complexity than in other marine unicellular *Cyanobacteria*. The genome sequencing of *A. atlantica* CENA595^T sp. nov., enabled phylogenetic inference based on 21 protein sequences. Expanding the minimal data set of 16S ribosomal RNA sequences (around 1400 nucleotide sites) by using a large data set (3172 amino acid sites) helped to overcome the limitations of single-gene analyses, such as loss of phylogenetic information due to multiple substitutions at the same nucleotide site (saturation) and low statistical support across cyanobacterial phylogeny. However, the amount of data for the cyanobacterial genome currently available is much lower than the data for 16S rRNA gene sequences. However, as more cyanobacterial genomes are sequenced, a more phylogenomic approach has been used to study the deep branching relationships and the evolutionary history

of these organisms (Sánchez-Baracaldo *et al*., 2005; Blank & Sánchez-Baracaldo, 2010; Larsson *et al*., 2011; Shih *et al*., 2013; Bombar *et al*., 2014; Komárek *et al*., 2014; Calteau *et al*., 2014; Chrismas *et al*., 2015; Schirrmeister *et al*., 2015). Phylogenomic tree analysis (Fig. S2) fitted the *A. atlantica* CENA595^{T}sp. nov., amino acid sequence within a clade containing sequences of *Synechocystis* sp. PCC7509, *C. thermalis* PCC7203 and *Gloecocapsa* sp. PCC7428. This cyanobacterial group, according to the more recently proposed cyanobacteria classification (Komárek *et al*., 2014), form the new order *Chroococcidiopsidales*, which encompasses one family (*Chroococcidiopsidaceae*) and one genus (*Chroococcidiopsis*), whose members mostly live in extreme habitats. Previously, strain CENA595^T was classified only at the order level (*Chroococcales*) based on morphological traits (Rigonato *et al*., 2015). The data generated in this study allowed a more robust taxonomic analysis of *A. atlantica* CENA595^T sp. nov. Therefore, this strain should be classified in the order *Chroococcidiopsidales*, but in a new monotypic family, *Aliterellaceae* fam. nov., due to its phylogenetic distinctiveness and low similarity with members of the genus *Chroococcidiopsis* (Fewer *et al*., 2002; Komárek *et al*., 2014).

The examples of the novel genus described here are able to grow in habitats with distinct conditions of salinity and naturally occupy different niches. The description of these species has also brought new insights into the identities of several uncultured bacteria and cyanobacterial sequences available in the GenBank database, which likely belong to the genus *Aliterella* gen. nov. Hence, *Aliterella* gen. nov., colonizes a variety of environments (associated with Antarctic green algae, the deep ocean, forest soil, human skin, Antarctic soil, building stone) suggesting a broad distribution and an important ecological role.

Description of *Aliterella* gen. nov.

(A.li.ter.el'la. L. adv. aliter otherwise, differently; N.L. fem. dim. n. *Aliterella* a cyanobacterium with variable cell shape).

In culture medium, cells are solitary or are grouped in microscopic or macroscopic irregular clusters, but form mucilaginous colonies; the mucilage is rigid. Spherical, oval, irregular or elongated, without aerotopes and blue-green. Divide by binary fission in multiple planes in successive generations. Thylakoids are organized randomly in the cytoplasm, with a higher concentration in the cell periphery, with a clear distinction of periplasmic and nucleoplasmic zones in light microscopy.

The type species is *Aliterella atlantica* sp. nov.

Description of *Aliterella atlantica* sp. nov.

(at.lan'ti.ca L. fem. adj. *atlantica* pertaining to the Atlantic ocean, the geographical origin of the type strain).

Thallus is extended, compact, irregular in shape, composed of numerous colonies or isolated ensheathed cells. Colonies are mostly irregular, sometimes rounded, variable in dimensions and in the number of cells (up to 32–64 cells or more). The mucilage is firm surrounding cells and colonies, colourless. With pressure cells are easily removed from colonies and the mucilaginous envelope remains and is the shape of the components. Cells are cylindrical with rounded ends, sometimes irregular, 3.0–4.8 μm long (mean, 3.7 μm), 2.0–3.3 μm diameter (mean, 2.7 μm), 1.1 to 2× longer than wide

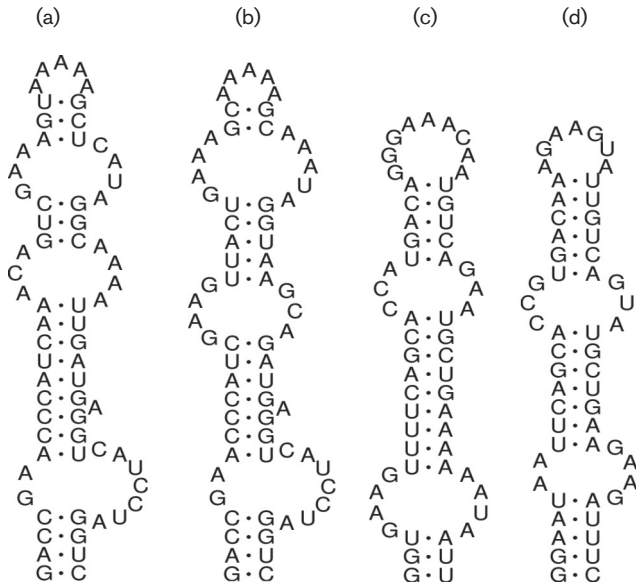

Fig. 4. Secondary structure for D1D1′ and BoxB 16S-23S rRNA intergenic spacer. (a) D1D1′ CENA595^T; (b) D1D1′ CENA408^T; (c) BoxB CENA595^T; (d) BoxB CENA408^T.

(mean, 1.4×). The cell contents are granulated, sometimes homogeneous, with the chromatoplasm and centroplasm usually recognizable. Reproduction is by simple binary cell division in three or more planes. Baeocytes and nanocytes are not observed.

The type strain, CENA595^T, was isolated from waters of the South Atlantic Ocean continental shelf, Brazil.

Description of *Aliterella antarctica* sp. nov.

(ant.arc'ti.ca. L. fem. adj. *antarctica* southern, and by extension, pertaining to Antarctica, the geographical origin of the type strain).

Thallus is extended, compact, irregular in shape, composed of numerous colonies or isolated ensheathed cells. Colonies are mostly irregular, sometimes rounded, variable in dimensions and number of cells (up to 32–64 cells). Mucilage is apparently firm surrounding the cells and colonies, homogeneous or occasionally lamellate, colourless to brownish. With pressure cells are easily removed from the colonies and the mucilaginous envelope remains the shape of their components. Cells are cylindrical with rounded ends or, less frequently, near spherical or irregular, 3.4–5.8 μm long (mean, 4.5 μm), 2.5–4.6 μm in diameter (mean, 3.5 μm), 1.0 to 1.7× longer than broad (mean 1.3×). Cell contents are granulated, sometimes homogeneous. The chromatoplasm and centroplasm are usually recognizable. Reproduction is by simple binary cell division in three or more planes. Baeocytes and nanocytes are not observed.

The type strain, CENA408^T, was isloated from green turf alga growing together with decomposed Deschampsia grass on a ornithogenic soil (Pygoscelis adeliae), collected close to the sea in Admiralty Bay, King George Island

Acknowledgements

This study was supported by grants from the State of São Paulo Research Foundation (FAPESP) to JR (2012/06850-3), to MFF (2013/50425-8) and from the National Council for Scientific and Technological Development (CNPq) to FPB (565060/2010-4). JR received a postdoctoral fellowship from the Brazilian Federal Agency for the Support and Evaluation of Graduate Education (CAPES-PNPD20131744 USP/CENA program). DOA., W.A.G and D.B.G. were supported by FAPESP graduate scholarship 2011/08092-6, 2012/16430-1, 2010/00321-3, respectively. MFF would also like to thank CNPq for a research fellowship (310244/2015-3). We thank Prof. Francisco André Ossamu Tanaka for permission to use the transmission electron microscope and assistance with the equipment at NAP-MEPA/ESALQ-USP and his team for helping with the sample preparation. We are also grateful to the R/V 'Alpha Crucis' crew, in name of Captain José Helvécio Moraes de Rezende, and all the Oceanographic Institute team that helped in the sampling.

References

Abascal, F., Zardoya, R. & Posada, D. (2005). ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* **21**, 2104–2105.

Allen, M. M. (1968). Simple conditions for growth of unicellular blue-green algae on plates. *J Phycol* **4**, 1–4.

Azua-Bustos, A., Zúñiga, J., Arenas-Fajardo, C., Orellana, M., Salas, L. & Rafael, V. (2014). *Gloeocapsopsis* AAB1, an extremely desiccation-tolerant cyanobacterium isolated from the Atacama Desert. *Extremophiles* **18**, 61–74.

Baudoux, A. C. & Brussaard, C. P. (2005). Characterization of different viruses infecting the marine harmful algal bloom species *Phaeocystis globosa*. *Virology* **341**, 80–90.

Blank, C. E. & Sánchez-Baracaldo, P. (2010). Timing of morphological and ecological innovations in the cyanobacteria–a key to understanding the rise in atmospheric oxygen. *Geobiology* **8**, 1–23.

Bombar, D., Heller, P., Sanchez-Baracaldo, P., Carter, B. J. & Zehr, J. P. (2014). Comparative genomics reveals surprising divergence of two closely related strains of uncultivated UCYN-A cyanobacteria. *ISME J* **8**, 2530–2542.

Bonner, J. T. (1998). The origins of multicellularity. *Integr Biol* **1**, 27–36.

Brown, I. I., Mummey, D. & Cooksey, K. E. (2005). A novel cyanobacterium exhibiting an elevated tolerance for iron. *FEMS Microbiol Ecol* **52**, 307–314.

Calteau, A., Fewer, D. P., Latifi, A., Coursin, T., Laurent, T., Jokela, J., Kerfeld, C. A., Sivonen, K., Piel, J. & Gugger, M. (2014). Phylum-wide comparative genomics unravel the diversity of secondary metabolism in Cyanobacteria. *BMC Genomics* **15**, 977.

Chong, C. W., Convey, P., Pearce, D. A. & Tan, I. K. P. (2012). Assessment of soil bacterial communities on Alexander Island (in the maritime and continental Antarctic transitional zone). *Polar Biology* **35**, 387–399.

Chrismas, N. A., Anesio, A. M. & Sánchez-Baracaldo, P. (2015). Multiple adaptations to polar and alpine environments within cyanobacteria: a phylogenomic and Bayesian approach. *Front Microbiol* **6**, 1070.

Claessen, D., Rozen, D. E., Kuipers, O. P., Søgaard-Andersen, L. & van Wezel, G. P. (2014). Bacterial solutions to multicellularity: a tale of biofilms, filaments and fruiting bodies. *Nat Rev Microbiol* **12**, 115–124.

Dagan, T., Roettger, M., Stucken, K., Landan, G., Koch, R., Major, P., Gould, S. B., Goremykin, V. V., Rippka, R. & other authors (2013). Genomes of Stigonematalean cyanobacteria (subsection V) and the evolution of oxygenic photosynthesis from prokaryotes to plastids. *Genome Biol Evol* **5**, 31–44.

Ewing, B. & Green, P. (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* **8**, 186–194.

Ewing, B., Hillier, L., Wendl, M. C. & Green, P. (1998). Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* **8**, 175–185.

Fewer, D., Friedl, T. & Büdel, B. (2002). *Chroococcidiopsis* and heterocyst-differentiating cyanobacteria are each other's closest living relatives. *Mol Phylogenet Evol* **23**, 82–90.

Gao, E. B., Gui, J. F. & Zhang, Q. Y. (2012). A novel cyanophage with a cyanobacterial nonbleaching protein A gene in the genome. *J Virol* **86**, 236–245.

Gordon, D., Abajian, C. & Green, P. (1998). Consed: a graphical tool for sequence finishing. *Genome Res* **8**, 195–202.

Grosberg, R. K. & Strathmann, R. R. (2007). The evolution of eulticellularity: a minor major transition? *Annu Rev Ecol Evol Syst* **38**, 621–654.

Hallmann, C., Stannek, L., Fritzlar, D., Hause-Reitner, D., Friedl, T. & Hoppert, M. (2013). Molecular diversity of phototrophic biofilms on building stone. *FEMS Microbiol Ecol* **84**, 355–372.

Herdman, M., Janvier, M., Rippka, R. & Stanier, R. Y. (1979). Genome Size of Cyanobacteria. *J Gen Microbiol* **111**, 73–85.

Honda, D., Yokota, A. & Sugiyama, J. (1999). Detection of seven major evolutionary lineages in cyanobacteria based on the 16S rRNA gene

sequence analysis with new sequences of five marine *Synechococcus* strains. *J Mol Evol* **48**, 723–739.

Karnovsky, M. J. A. (1965). Formaldehyde-glutaraldehyde fixative of high osmolarity for use in electron microscopy. *J Cell Biol* **27**, 137.

Kim, M., Oh, H. S., Park, S. C. & Chun, J. (2014). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol* **64**, 346–351.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* **16**, 111–120.

Komárek, J. & Anagnostidis, K. (1998). Cyanoprokaryota, 1. Teil: Chroococcales. In *Süsswasserflora Von Mitteleuropa* ,19, pp. 1–548. Edited by H. Ettl. Stuttgart: Gustave Fisher.

Komárek, J., Kaštovský, J., Mareš, J. & Johansen, J. R. (2014). Taxonomic classification of cyanoprokaryotes (cyanobacterial genera) using a polyphasic approach. *Preslia* **86**, 295–335.

Kong, H. H., Oh, J., Deming, C., Conlan, S., Grice, E. A., Beatson, M. A., Nomicos, E., Polley, E. C., Komarow, H. D. & other authors (2012). Temporal shifts in the skin microbiome associated with disease flares and treatment in children with atopic dermatitis. *Genome Res* **22**, 850–859.

Kováčik, L., Jezberová, J., Komárková, J., Kopecký, J. & Komárek, J. (2011). Ecological characteristics and polyphasic taxonomic classification of stable pigment-types of the genus *Chroococcus* (Cyanobacteria). *Preslia* **83**, 145–166.

Larsson, J., Nylander, J. A. & Bergman, B. (2011). Genome fluctuations in cyanobacteria reflect evolutionary, developmental and adaptive traits. *BMC Evol Biol* **11**, 187.

Lehner, J., Zhang, Y., Berendt, S., Rasse, T. M., Forchhammer, K. & Maldener, I. (2011). The morphogene AmiC2 is pivotal for multicellular development in the cyanobacterium *Nostoc punctiforme*. *Mol Microbiol* **79**, 1655–1669.

Lowe, T. M. & Eddy, S. R. (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* **25**, 955–964.

Moissl, C., Osman, S., La Duc, M. T., Dekas, A., Brodie, E., DeSantis, T., Desantis, T. & Venkateswaran, K. (2007). Molecular bacterial community analysis of clean rooms where spacecraft are assembled. *FEMS Microbiol Ecol* **61**, 509–521.

Mori, T. & Johnson, C. H. (2001). Independence of circadian timing from cell division in cyanobacteria. *J Bacteriol* **183**, 2439–2444.

Posada, D. & Crandall, K. A. (1998). Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818.

Ramos, V., Seabra, R., Brito, Ângela., Santos, A., Santos, C. L., Lopo, M., Moradas-Ferreira, P., Vasconcelos, V. M. & Tamagnini, P. (2010). Characterization of an intertidal cyanobacterium that constitutes a separate clade together with thermophilic strains. *Eur J Phycol* **45**, 394–403.

Rigonato, J., Alvarenga, D. O., Branco, L. H., Varani, A. M., Brandini, F. P. & Fiore, M. F. (2015). Draft genome sequence of a novel culturable marine chroococcalean cyanobacterium from the South atlantic ocean. *Genome Announc* **3**, e00384.

Rippka, R. (1988). Recognition and identification of cyanobacteria. *Meth Enzymol* **167**, 28–76.

Robertson, B. R., Tezuka, N. & Watanabe, M. M. (2001). Phylogenetic analyses of *Synechococcus* strains (cyanobacteria) using sequences of 16S rDNA and part of the phycocyanin operon reveal multiple evolutionary lines and reflect phycobilin content. *Int J Syst Evol Microbiol* **51**, 861–871.

Ronquist, F. & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574.

Rossetti, V., Schirrmeister, B. E., Bernasconi, M. V. & Bagheri, H. C. (2010). The evolutionary path to terminal differentiation and division of labor in cyanobacteria. *J Theor Biol* **262**, 23–34.

Sanchez-baracaldo, P., Hayes, P. K. & Blank, C. E. (2005). Morphological and habitat evolution in the Cyanobacteria using a compartmentalization approach. *Geobiology* **3**, 145–165.

Schattner, P., Brooks, A. N. & Lowe, T. M. (2005). The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res* **33**, W686–W689.

Schirrmeister, B. E., Antonelli, A. & Bagheri, H. C. (2011). The origin of multicellularity in cyanobacteria. *BMC Evol Biol* **11**, 45.

Schirrmeister, B. E., Gugger, M., Philip, C. J. & Donoghue, P. C. J. (2015). Cyanobacteria and the Great Oxidation Event: evidence from genes and fossils. *Palaeontology* **58**, 769–785.

Shih, P. M., Wu, D., Latifi, A., Axen, S. D., Fewer, D. P., Talla, E., Calteau, A., Cai, F., Tandeau de Marsac, N. & other authors (2013). Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proc Natl Acad Sci U S A* **110**, 1053–1058.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M. & other authors (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539.

Spurr, A. R. (1969). A low-viscosity epoxy resin embedding medium for electron microscopy. *J Ultrastruct Res* **26**, 31–43.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* **30**, 2725–2729.

Taton, A., Grubisic, S., Brambilla, E., De Wit, R. & Wilmotte, A. (2003). Cyanobacterial diversity in natural and artificial microbial mats of Lake Fryxell (McMurdo Dry Valleys, Antarctica): a morphological and molecular approach. *Appl Environ Microbiol* **69**, 5157–5169.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673–4690.

Tomitani, A., Knoll, A. H., Cavanaugh, C. M. & Ohno, T. (2006). The evolutionary diversification of cyanobacteria: molecular-phylogenetic and paleontological perspectives. *Proc Natl Acad Sci U S A* **103**, 5442–5447.

Waterbury, J. B. (1989). Subsection II Order Pleurocapsales. In *Bergeys Manual of Systematic Bacteriology*, Vol. 3, pp. 1746–1770. Edited by R. E. Buchanan & N. E. Gibbons. Baltimore: Williams & Wilkins.

Wu, M. & Eisen, J. A. (2008). A simple, fast, and accurate method of phylogenomic inference. *Genome Biol* **9**, R151.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31**, 3406–3415.