Draft genome sequences of *Cylindrospermopsis raciborskii* strains CS-508 and MVCC14, isolated from freshwater bloom events in Australia and Uruguay

Juan J Fuentes-Valdés1,2, Katia Soto-Liebe1, Danilo Pérez-Pantoja3, Javier Tamames4, Lucy Belmar1, Carlos Pedrós-Alió6, Daniel Garrido2 and Mónica Vásquez1*

Abstract

Members of the genus *Cylindrospermopsis* represent an important environmental and health concern. Strains CS-508 and MVCC14 of *C. raciborskii* were isolated from freshwater reservoirs located in Australia and Uruguay, respectively. While CS-508 has been reported as non-toxic, MVCC14 is a saxitoxin (STX) producer. We annotated the draft genomes of these *C. raciborskii* strains using the assembly of reads obtained from Illumina MiSeq sequencing. The final assemblies resulted in genome sizes close to 3.6 Mbp for both strains and included 3202 ORFs for CS-508 (in 163 contigs) and 3560 ORFs for MVCC14 (in 99 contigs). Finally, both the average nucleotide identity (ANI) and the similarity of gene content indicate that these two genomes should be considered as strains of the *C. raciborskii* species.

Keywords: Cylindrospermopsis, Bloom, Cyanobacteria, Environmental toxicity, Non-ribosomal peptide-synthetase, Polyketide synthases

Introduction

Cyanobacterial bloom-forming species are a persistent global problem [1, 2]. *Cylindrospermopsis raciborskii*, is a species responsible for algal blooms that cause serious problems because of the wide variety of toxic compounds that it produces [3, 4]. Animal consumption of contaminated water with toxic metabolites produces symptoms associated with dermal rash, neural disturbance, hepatic and digestive disorder, and in some cases causing death [4, 5]. It was first described in Java (Indonesia) in 1912 [6], and was morphologically characterized in 1972 by Seenayya and Subba-Raju [7] as a Gram-negative-like, cylindrical filament able to fix nitrogen. To date, this species has been characterized as a producer of saxitoxin, a neurotoxin able to block voltage dependent mammalian sodium channels [8]. It also produces cylindrospermopsin, a toxin related with phosphatase metabolic inhibition in hepatocyte cells [9]. Recently, an anti-fungal glycolipopeptide affecting the plasma membrane integrity of *Candida albicans* cells, classified as hassallidin, has also been identified [10–12].

In order to understand the mechanisms responsible for the synthesis of these toxins, representative strains of this species have been characterized both genetically and chromatographically [13]. To date, Brazilian isolates have been characterized as CYL producers (CS-505 and CS-506), HAS producers (CS-505 and CS-509) and as non-toxin producers (CS-508) (unpublished data). In addition, the Uruguayan strain MVCC14 has been described as a STX producer [14]. Moreover, a Brazilian isolate *Raphidiopsis brookii* D9, a species phylogenetically closely related to *C. raciborskii* (Fig. 1), has also been reported as a STX producer [15–17]. The complete genome of *C. raciborskii* CS-505 and draft genomes of strains CS-506, CS-509 and *R. brookii* D9 are currently available [16, 18].

To provide further data to better understand the genomics and physiology of *C. raciborskii*, including its high...
capacity for dispersal, we performed a genome sequence analysis of Australian strain CS-508 and Uruguayan strain MVCC14, including gene annotation using the Clusters of Orthologous Group (COG) database [19]. Moreover, we also conducted a comparative genome analysis on five *C. raciborskii* strains: CS-505, CS-506, CS-508, CS-509 and MVCC14, in addition to *R. brookii* D9 to identify common genes.

**Organism information**

**Classification and features**

*C. raciborskii* is a relevant environmental species causing harmful blooms in freshwater environments, with certain strains synthesizing toxins.

*C. raciborskii* species (Tables 1 and 2), were initially described as microorganisms growing in the tropics, however, they have been reported in temperate freshwaters [20]. As previously described [21], the cells belonging to the genus *Cylindrospermopsis* could either be cylindrical filaments with terminal nitrogen fixation structures (heterocysts) (Fig. 1a-e) or resistant cells (akinetes). Both structures could be differentiated under nutrient-deficient culture media. In heterocyst-forming cyanobacteria, heterocysts are distributed in semi-regular intervals along the filament or only in the terminal position. The presence of intercalated heterocysts in *C. raciborskii* has been rarely observed, and has been thus described as a species with terminal heterocysts [22]. However, we observed intercalated heterocysts in strain MVCC14 under nitrogen starvation and under different nitrogen conditions (Fig. 1c-e). The distribution of the heterocysts along the filament has been the subject of research by comparing genetic and physiological traits between *Cylindrospermopsis* and *Anabaena*, as models of differential patterns [23, 24]. *Anabaena* sp. PCC7120 differentiates heterocysts after every 8 to 12 vegetative cells under nitrogen deprivation [23, 24]. We were able to observe heterocysts more frequently in some filaments; regularity between heterocyst cells was approximately of 30 neighboring vegetative cells (SD ± 7, 4). This is the first report showing the transient presence of intercalary heterocyst in this *C. raciborskii* species.
Furthermore, research should help to understand the genetic control that regulates this sporadic distribution of heterocysts in this C. raciborskii strain.

Despite their very similar morphology, C. raciborskii and R. brookii have been classified as different species because the latter is unable of fix nitrogen and does not develop heterocysts (e.g. [25]). Here, the maximum likelihood phylogenetic tree of 16S-rRNA gene sequences shows that R. brookii and C. raciborskii strains constitute a statistically well-supported monophyletic clade (Fig. 2 and Additional file 1: Figure S1). This clade comprises sequences sharing ≥98% of similarity and show low evolutionary rate within the clade. Despite this, it is possible to identify some sub-clusters with a certain coherent phylo-geographical distribution as was previously described [26, 27]. For example, the sub-cluster comprising strains exclusively from South America (R. brookii D9, C. raciborskii MVCC14 and T3) is segregated with a well-supported statistical value (Fig. 2, Additional file 1: Figures. S2 and S4). Phylogenetic analyses from other phylogenetic markers also displayed the monophyletic nature among R. brookii and C. raciborskii strains (Additional file 1: Figures. S2, S3, S4 and S5). This is congruent with a previous study of phylogenetic relationships inferred from several conserved genes, which postulate that Cylindrospermopsis and Raphidiopsis representatives should be congeners [28]. However, to assess the taxonomic classification of these microorganisms further phylogenetic analyses (e.g., global genome comparisons) or more complete physiological descriptions are required.

**Genome sequencing information**

**Genome project history**

Strains CS-508 and MVCC14 were selected for sequencing based on their phylogenetic relationship between strains from South America and Australia. Sequenced
The CS-508 Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession MBQX00000000. The version described here is MBQX01000000. MVCC14 Whole Genome Shotgun Project has been deposited under the accession ID MBQY00000000. The version described in this paper is version MBQY01000000. A summary of the project information is shown in Table 3.

**Growth conditions and genomic DNA preparation**

*C. raciborskii* cultures were grown in MLA medium [30], under 12:12 light:dark cycles at 25 °C. Total DNA extractions were carried out using 100 mL of exponential growth culture, obtaining approximately 1 g of wet cell pellet. DNA purification was conducted by standard CTAB protocol [31]. Total cell pellets were mechanically disrupted and resuspended in 500 μL of CTAB buffer; and incubated at 55 °C for 1 h under constant mixing.

The DNA was purified using 500 μL phenol/chloroform/isoamyl alcohol (25:24:1) and centrifuged at 8000 x g for 7 min. DNA was precipitated using isopropanol/ammonium acetate (0.54 vol cold isopropanol, 0.08 vol ammonium acetate 7.5 M). Finally, DNA was washed with 70% and then with 90% ethanol and resuspended in 50 μL of pure water. DNA extraction was visualized using red gel staining in a 1% agarose gel under UV light.

**Genome sequencing and assembly**

Both genomes were obtained by a shotgun strategy using Illumina MiSeq sequencing technology. A total of 8,308,910 paired-end reads were obtained for CS-508 strain and 28,711,437 paired-end reads for MVCC14 strain. Quality control checks were performed on the raw FASTQ data using FastQC (version 0.10.1) [32]. Sequencing reads were trimmed for sequencing adaptors using Trimmomatic (version 0.32) [33] and the quality filtering and trimming was done by Prinseq-lite (version...
Briefly, reads were trimmed for 'N' characters and low quality nucleotides (Phred score cutoff of 24) and then any read with an average Phred score below 29 and shorter than 80 nt was discarded. A de novo assembly strategy involving multiple algorithms and merging of the individual assemblies was performed. Assemblies by IDBA [35], SPADes [36], VELVET [37] and ABYSS [38] algorithms were generated by using the platform MIX software [39] to improve draft assembly by reducing contig fragmentation. Contigs shorter than 1000 bp were discarded. The final assembly resulted in 163 contigs for CS-508 and 99 contigs for MVCC14, accounting for 3,558,956 bp and 3,594,524 bp, respectively. CheckM analysis [40] indicated a genome completeness of 97.57% for CS-508 and 96.29% for MVCC14.

**Genome annotation**

The gene annotation process was conducted using the RAST Server 2.0 [29]. Predicted coding sequences were extracted from RAST platform and homology was evaluated by BLASTp scan, with each predicted ORF as a query against the complete bacterial database.

**Table 3** Project information

| MIGS ID | Property                      | Term (for CS-508)                  | Term (for MVCC14)                  |
|---------|-------------------------------|-----------------------------------|-----------------------------------|
| MIGS 31 | Finishing quality             | High- Quality Draft               | High- Quality Draft               |
| MIGS-28 | Libraries used                | Illumina                         | Illumina                         |
| MIGS 29 | Sequencing platforms          | Illumina HiSeq2000               | Illumina HiSeq2000               |
| MIGS 31.2 | Fold coverage               | 20x                              | 20x                              |
| MIGS 30 | Assemblers                    | IDBA, SPADES, VELVET and ABYSS    | IDBA, SPADES, VELVET and ABYSS    |
| MIGS 32 | Gene calling method           | Rast                             | Rast                             |
|         | Locus Tag                     | CYL_CS508                        | CYL_MVCC14                       |
|         | GenBank ID                    | MBQX000000000                   | MBQY000000000                   |
|         | GenBank Date of Release       | November 01, 2016                | November 01, 2016                |
|         | GOLD ID                       | G0120410                        | G0121371                        |
|         | BIOPROJECT                    | PRJNA327084                      | PRJNA327088                      |
| MIGS 13 | Source Material Identifier    | Freshwater                       | Freshwater                       |
|         | Project relevance             | Environment                      | Environment                      |
### Table 4: Genome statistics of CS-508 (A) and MVCC14 (B)

| Attribute                        | A Value    | A % of Total | B Value    | B % of Total |
|----------------------------------|------------|--------------|------------|--------------|
| Genome size (bp)                 | 3,558,956  | 100          | 3,594,524  | 100          |
| DNA coding (bp)                  | 3,039,246  | 85.34        | 3,074,946  | 85.55        |
| DNA G + C (bp)                   | 1,530,351  | 43           | 1,581,591  | 44           |
| DNA scaffolds                    | 163        | 100          | 99         | 100          |
| Total genes                      | 3344       | 100          | 3616       | 100          |
| Protein coding genes             | 3302       | 98.74        | 3560       | 98.45        |
| RNA genes                        | 42         | 1.26         | 56         | 1.55         |
| Pseudo genes                     | –          | –            | –          | –            |
| Genes in internal clusters       | –          | –            | –          | –            |
| Genes with function prediction   | 2247       | 67.19        | 2337       | 64.63        |
| Genes assigned to COGs           | 1747       | 56.16        | 1796       | 55.55        |
| Genes with Pfam domains          | 2656       | 79.43        | 2800       | 77.43        |
| Genes with signal peptides       | 71         | 2.12         | 63         | 1.74         |
| Genes with transmembrane helices | 255        | 7.63         | 748        | 20.66        |
| CRISPR repeats                   | 7          | –            | 9          | –            |

### Table 5: Number of genes associated with general COG functional categories

| Code | CS-508 Value | %age | MVCC14 Value | %age | Description                                                      |
|------|--------------|------|--------------|------|------------------------------------------------------------------|
| J    | 142          | 4.56 | 143          | 4.37 | Translation, ribosomal structure and biogenesis                  |
| A    | 0            | 0.00 | 0            | 0.00 | RNA processing and modification                                  |
| K    | 69           | 2.22 | 64           | 1.96 | Transcription                                                   |
| L    | 88           | 2.83 | 112          | 3.43 | Replication, recombination and repair                            |
| B    | 0            | 0.00 | 0            | 0.00 | Chromatin structure and dynamics                                 |
| D    | 21           | 0.68 | 19           | 0.58 | Cell cycle control, Cell division, chromosome partitioning      |
| V    | 0            | 0.00 | 0            | 0.00 | Defense mechanisms                                              |
| T    | 49           | 1.58 | 54           | 1.65 | Signal transduction mechanisms                                  |
| M    | 123          | 3.95 | 130          | 3.98 | Cell wall/membrane biogenesis                                   |
| N    | 6            | 0.19 | 5            | 0.15 | Cell motility                                                   |
| U    | 0            | 0.00 | 0            | 0.00 | Intracellular trafficking and secretion                          |
| O    | 111          | 3.57 | 111          | 3.40 | Posttranslational modification, protein turnover, chaperones    |
| C    | 157          | 5.05 | 163          | 4.99 | Energy production and conversion                                |
| G    | 99           | 3.18 | 93           | 2.84 | Carbohydrate transport and metabolism                            |
| E    | 125          | 4.02 | 123          | 3.76 | Amino acid transport and metabolism                              |
| F    | 45           | 1.45 | 44           | 1.35 | Nucleotide transport and metabolism                              |
| H    | 104          | 3.34 | 107          | 3.27 | Coenzyme transport and metabolism                                |
| I    | 32           | 1.03 | 31           | 0.95 | Lipid transport and metabolism                                  |
| P    | 128          | 4.11 | 130          | 3.98 | Inorganic ion transport and metabolism                           |
| Q    | 40           | 1.29 | 36           | 1.10 | Secondary metabolites biosynthesis, transport and catabolism     |
| R    | 252          | 8.10 | 262          | 8.01 | General function prediction only                                 |
| S    | 156          | 5.01 | 169          | 5.17 | Function unknown                                                |
| –    | 1364         | 43.84| 1473         | 45.06| Not in COGs                                                     |

The total is based on the total number of protein coding genes in the genome.
genes related to STX biosynthesis. In the MVCC14 draft genome, we identified a group of according to LC-MS/MS analysis (unpublished results). We detected the presence of hassallidin in CS-508 cultures, the hassallidin cluster. Surprisingly, we were not able to detect the presence of hassallidin in CS-508 cultures, containing 3202 and 3560 ORFs each. Table 5 shows the COG distribution of the corresponding genes. A high number of these encode metabolic proteins (COG codes R, S, M, C, E, P, O, H and T). Interestingly, no genes for the “RNA processing and modification” category were found in any genome. This has been observed in another cyanobacterial genome [41] and could be explained by genetic divergence of these cyanobacteria. Approximately 22% (CS-508) and 26% (MVCC14) of the total coding genes were not classified in any COG category.

**Insights from the genome sequence**

Photoautotrophic metabolic pathways were reconstructed in CS-508 and MVCC14 draft genomes, based on the predicted metabolic pathways in previous sequenced genomes of C. raciborskii [16, 18]. Nitrogen metabolic systems related to ammonium, nitrate and nitrite acquisition, as well as heterocyst differentiation and nitrogen fixation, were identified in both genome drafts.

Sequenced genomes were compared to previously published C. raciborskii and R. brookii genomes. We determined the average nucleotide identity in these genomes by a two-way comparison analysis (Table 6), using the inference tool ANI calculator [20]. The percentage of shared genes between strains ranged from 93.23 to 99.77%. Finding the ANI value, the complete group, C. raciborskii and R. brookiicould be considered as members of the same species, considering a threshold value of 95% [42].

We identified four genes encoding a non-ribosomal peptide synthetase complex in the CS-508 genome related to the hassallidin biosynthesis. We found in CS-508 the same gene cluster as in the hassallidin producers CS-509, CS-505 and Anabaena SYKE748A [10, 16, 18], with no evidence of mutations in the hassallidin cluster. Surprisingly, we were not able to detect the presence of hassallidin in CS-508 cultures, according to LC-MS/MS analysis (unpublished results). In the MVCC14 draft genome, we identified a group of genes related to STX biosynthesis. STX is a paralytic biotoxin produced by marine dinoflagellates and freshwater cyanobacteria [14]. The sxt gene cluster found in MVCC14 has a similar distribution and toxin profile to R. brookii D9 [16]. We did not find NRPS sequences in the MVCC14 genome.

**Conclusions**

In order to understand the genomics of the toxin producing, bloom forming C. raciborskii, this work presents two drafts of sequenced genomes from the non-toxic Australian strain CS-508 and the Uruguayan neurotoxin-producer strain MVCC14. An NRPS gene cluster related with hassallidin production was identified in CS-508 and PKS-like set of genes related with STX production was identified in the genome of the MVCC14 strain. Considering the 16S rRNA gene phylogenetic analysis and genome level comparison, we identified a phyleogeographical segregation of the C. raciborskii and R. brookii strains retrieved from South America. Disregarding nitrogen fixation ability, these results suggest R. brookii D9 and C. raciborskii mvcc14 are closely related at genome level, which could lead to new research to corroborate the Cylindrospermopsis /Raphidiopsis clade as one comprised by two genera or by a single genus with different species.

**Additional file**

Additional file 1: Figure S1. Cyanobacterial ML phylogenetic tree based on 16S rRNA gene sequences. Figure S2. ML phylogenetic tree based on rbcL gene sequences from relatives cyanobacteria. Figure S3. ML phylogenetic tree based on ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) proteins from relatives cyanobacteria. Figure S4. ML phylogenetic tree based on psbA gene sequences from relatives cyanobacteria. Figure S5. ML phylogenetic tree based on Photosystem II D1 (PsbA) proteins from relatives cyanobacteria.

**Acknowledgements**

This work was financed by the following grants: Fondecyt regular 1131037, 1161232, Fondecyt de Iniciación 11103518 and JF PhD Conicyt Fellowship 21120837, CTM2016-80095-C2-1-R from the Spanish Ministry of Economy and Competitiveness; KSL was financed by postdoctoral Fondecyt N° 3130681, LB was funded by Postdoctoral Fondecyt N° 3140330. K. del Rio for strain cultivation and DNA extraction and Dr. Sylvia Bonilla for kindly providing the MVCC14 C. raciborskii strain.
Authors’ contributions
JFF, KSL conceived the molecular genetic studies, participated in the sequence alignment and drafted the manuscript. DG, DPP, JT participated in designing and performing the bioinformatics analysis. MV, CPA conceived the study, and participated in the draft of the manuscript. LB conducted the 16S rRNA gene phylogenetic analysis. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
1Department of Molecular Genetics and Microbiology, Pontificia Universidad Católica de Chile, 8331150 Santiago, Chile. 2Department of Chemical and Bioprocess Engineering, Pontificia Universidad Católica de Chile, 7820436 Santiago, Chile. 3Programa Institucional de Fomento a la Investigación, Desarrollo e Innovación, Universidad Tecnológica Metropolitana, 8940577 Santiago, Chile. 4Systems Biology Program, CNB, CSIC, Calle Darwin 3, 28049 Madrid, Spain.

Received: 3 August 2017 Accepted: 24 September 2018
Published online: 12 October 2018

References
1. Carmichael WW. Health Effects of Toxin-Producing Cyanobacteria. “The CyanohaBis”. Hum Ecol Risk Assess. An Int J. 2001;7:393–407.
2. Parf HW, Paul VJ. Climate change: Links to global expansion of harmful cyanobacteria. Water Res. 2012;46:1340–63.
3. Hawkins PR, Runnegar AR, Jackson AR, Falconer IR. Severe Hepatotoxicity Caused by the Tropical Cyanobacterium Supply Reservoir. Appl Env Microbiol. 1985;50:1296–9.
4. Hawkins PR, Chandrasena NR, Jones GJ, Humpage AR, Falconer IR. Isolation and toxicity of Cylindrospermopsis raciborskii from an ornamental lake. Toxicon. 1997;35(3):341–6.
5. Bitand JF, Jacquet S, Bernard C, Humbert J-F. Health hazards for terrestrial vertebrates from toxic cyanobacteria in surface water ecosystems. Vet Res. 2003;34:361–77.
6. Woloszyńska J. Das Phytoplankton einiger javanischer Seen, mit Berücksichtigung des Sawa-Planktons: Imprimerie de l’Université, 1912.
7. Seenayya G, Raju NS. On the ecology and systematic position of the alga known as Anabaenopsis raciborskii (Wolosz) Elek. and a critical evaluation of the forms described under the genus Anabaenopsis. Int. Symp. Taxon. Biol. Biog Donna Algae. 1st, Madras, 1970. Pap. 1972.
8. Lagos N, Onodera H, Zagatto PA, Andrinolo D, Azevedo SMF, Oshima Y. The diversity of Cylindrospermopsis raciborskii isolated from Brazil. Toxicon. 1999;37:1359–73.
9. Griffiths DJ, Saker ML. The Palm Island mystery disease 20 years on: A review of research on the cyanotoxin cylindrospermopsin. Environ Toxicol. 2003;18:76–93.
10. Vestola J, Shishido TK, Jokela J, Fewer DP, Arto O, Permi P, et al. Hassallidins, antifungal glycoplipopeptides, are widespread among cyanobacteria and are the end-product of a nonbosalom pathway. Proc Natl Acad Sci U S A. 2014;111:E1909–17.
11. Neuhof T, Schmieder P, Preussel K, Dieckmann R, Pham H, Bartl F, et al. Hassallidin A and a glycosylated lipopeptide with antifungal activity from the cyanobacterium Hassallia sp. J Nat Prod. 2005;68:685–700.
12. Neuhof T, Schmieder P, Seibold M, Preussel K, von Döhlen H. Hassallidin B- Second antifungal member of the Hassallidin family. Bioorganic Med Chem Lett. 2006;16:4220–2.
13. Dahlmann J, Budakowski WR, Luckas B. Liquid chromatography-electrospray ionisation-mass spectrometry based method for the simultaneous determination of algal and cyanobacterial toxins in phytoplankton from marine waters and lakes followed by tentative structural elucidation of microcystins. J Chromatogr A. 2003;994:45–57.
14. Piccini C, Aubriot L, Fabre A, Amara V, González-Pérez M, Giani A, et al. Genetic and eco-physiological differences of South American Cylindrospermopsis raciborskii isolates support the hypothesis of multiple ecotypes. Harmful Algae. 2011;10:644–53.
15. Stucken K, Murillo AA, Soto-Lebre K, Fuentes-Valdés JJ, Méndez MA, Vásquez M. Toxicity phenotype does not correlate with phylogeny of Cylindrospermopsis raciborskii strains. Syst Appl Microbiol. 2009;32:37–48.
16. Stucken K, John LJ, Gembella A, Murillo AA, Soto-Lebre K, Fuentes-Valdés JJ, et al. The smallest known genomes of multicellular and toxic cyanobacteria: comparison, minimal gene sets for linked traits and the evolutionary implications. PLoS One. 2010;5:e9235.
17. Fuenzalida L. Genetic and physiologic studies in the cyanobacterium Cylindrospermopsis raciborskii, a PSP-toxin producer. Santiago: Universidad de Chile; 2005.
18. Sinha R, Pearson LA, Davis TW, Muenchhoff J, Piatrana R, Ives A, et al. Comparative genomics of Cylindrospermopsis raciborskii strains with differential toxicities. BMC Genomics. 2014;15:83.
19. Tatusov RL, Koonin EV, Lipman DJ. A Genomic Perspective on Protein Families. Science. 1997;278:63–7.
20. Sukenik A, Hadas O, Kaplan A, Quesada A. Invasion of Nostocales (cyanobacteria) to subtropical and temperate freshwater lakes - physiological, regional, and global driving forces. Front Microbiol. 2012;3:1–9.
21. Padisák J. Cylindrospermopsis raciborskii (Woloszyńska) Seenayya et Subba Raju, an expanding, highly adaptive cyanobacterium: worldwide distribution and review of its ecology. Arch. Für Hydrobiol. Suppl. Monogr. Beitrage. 1997:563–93.
22. Chonudomkul D, Yongmanitchai W, Theeragool G, Kawachi M, Kasai F, Kaya K, et al. Morphology, genetic diversity, temperature tolerance and toxicity of Cylindrospermopsis raciborskii (Nostocales, Cyanobacteria) strains from Thailand and Japan. FEMS Microbiol Ecol. 2004;48:345–55.
23. Plominsky AM, Larsson J, Bengman B, Delherbe N, Oüssi I, Vásquez M. Dinitrogen fixation is restricted to the terminal heterocysts in the invasive cyanobacterium Cylindrospermopsis raciborskii C5-505. PLoS One. 2013;8:e51682.
24. Muñoz-García J, Ares S. Formation and maintenance of nitrogen-fixing cell patterns in filamentous cyanobacteria. Proc Natl Acad Sci. 2016;113:201524383.
25. Mohamed ZA. First report of toxic Cylindrospermopsis raciborskii and Raphidiopsis mediterranea (Cyanoprokaryota) in Egyptian fresh waters. FEMS Microbiol Ecol. 2007;59:749–61.
26. Gugger M, Mollica R, Le Berre B, Dufour P, Bernard C, Humbert J-F. Genetic Diversity of Cylindrospermopsis Strains (Cyanobacteria) Isolated from Four Continents. Appl Environ Microbiol. 2005;71:1097–100.
27. Haande S, Rohrflack T, Ballot A, Raberg K, Skullberg R, Beck M, et al. Genetic characterisation of Cylindrospermopsis raciborskii (Nostocales, Cyanobacteria) isolates from Africa and Europe. Harmful Algae. 2008;7:692–701.
28. Wu W, Zhu Z, Fu L, Niu B, Li W, WebMGA: a customizable web server for fast metagenomic sequence analysis. BMC Genomics. 2011;12:444.
29. Aziz RK, Bartels D, Best AA, Dejongh M, Dizis T, Edwards RA, et al. The RAST Server: rapid annotations using subsystems technology. BMC Genomics. 2008;9:75.
30. Castro D, Vera D, Lagos N, García C, Vásquez M. The effect of temperature on growth and production of paralytic shellfish poisoning toxins by the cyanobacterium Cylindrospermopsis raciborskii C10. Toxicon. 2004;44:483–9.
31. Richards E, Reichardt M, Rogers S. Preparation of Genomic DNA from Plant Tissue.Curr. Protoc. Mol. Biol. John Wiley & Sons, Inc.; 2001. p. 12.3.2;1–2.3.7.
32. Andrews S. FastQC: A quality control tool for high throughput sequence data [Internet]. Babraham Bioinforma. 2010 [cited 2016 Dec 2]. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
33. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:211–20.
34. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. Bioinformatics. 2011;27:963–4.
35. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:211–20.
36. Andrews S. FastQC: A quality control tool for high throughput sequence data [Internet]. Babraham Bioinforma. 2010 [cited 2016 Dec 2]. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
37. Zerbino DR, Birney E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 2008;18:821–9.
38. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJM, Birol I. ABySS: A parallel assembler for short read sequence data. Genome Res. 2009;19:1117–23.

39. Soueidan H, Maurier F, Groppi A, Sirand-Pugnet P, Tardy F, Citti C, et al. Finishing bacterial genome assemblies with Mix. BMC Bioinformatics. 2013;14:516.

40. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. S. Genome Res. 2015;25:1043–55.

41. Cheevadhanarak S, Paithoonrangsarid K, Prommeenate P, Tardy F, Citti C, et al. Draft genome sequence of Arthrospira platensis C1 (PCC9438). Stand. Genom. Sci. 2012;6:43–53.

42. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int. J Syst Evol. Microbiol. 2007;57:81–91.

43. Edgar RC. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32:1792–7.

44. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 2010;59:307–21.

45. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, et al. The minimum information about a genome sequences (MIGS) specification. Nat. Biotechnol. 2008;26:541–7.

46. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci. 1990;87:4576–9.

47. Castenholz RW. General characteristics of the cyanobacteria. In: Boone DR, Castenholz RW, editors. Bergey’s manual of systematic bacteriology. 2nd ed. New York: Springer; 2001. p. 474–87.

48. Saker ML, Neilan BA. Varied diazotrophies, morphologies, and toxicities of genetically similar isolates of Cylindrospermopsis raciborskii (Nostocales, Cyanophyceae) from northern Australia. Appl. Environ. Microbiol. 2001;67:1839–45.

49. Caumette P, Brochier-Armanet C, Normand P. Taxonomy and Phylogeny of Prokaryotes. Environ. Microbiol. Fundam. Appl. Dordrecht: Springer Netherlands; 2015. p. 145–90.

50. Meeks JC, Elhai J. Regulation of cellular differentiation in filamentous cyanobacteria in free-living and plant-associated symbiotic growth states. Microbiol Mol Biol Rev. 2002;66:94–121 table of contents.

51. Saker ML, Griffiths DJ. The effect of temperature on growth and cylindrospermopsin content of seven isolates of Cylindrospermopsis raciborskii (Nostocales, Cyanophyceae) from water bodies in northern Australia. Phycologia. 2000;39:349–54.

52. Moisander PH, Mclintain E, Paerl HW. Salinity effects on growth, photosynthetic parameters, and nitrogenase activity in estuarine planktonic cyanobacteria. Microb Ecol. 2002;43:432–42.

53. Saker ML. Cyanobacterial blooms in tropical north Queensland water bodies. Townsville: James Cook University; 2000.

54. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: tool for the unification of biology. Nat. Genet. 2000;25:25–9.

55. Vital D, Kruk C. Cylindrospermopsis raciborskii (Cyanobacteria) extends its distribution to Latitude 34°35’S: taxonomical and ecological features in Uruguayan eutrophic lakes. PANAMJAS. 2008;3:142–51.