Indian Journal of Biotechnology
Vol 16, January 2017, pp 84-90

Similarity analysis of *Spirulina/Arthrospira* strains on the basis of phycocyanin operon locus (cpcB-IGS-cpcA) and 16S rRNA gene sequences

Devendra Kumar[1]*, Dolly Wattal Dhar[1], Priyanka Nehra[2] and Neeraj Kumar[2]

[1]Centre for Conservation and Utilisation of Blue Green Algae (CCUBGA), Indian Agricultural Research Institute (IARI) New Delhi 110 012, India

[2]Microbiology Department, Kurukshetra University, Kurukshetra 136 119, India

Received 31 August 2015; revised 7 October 2015; accepted 25 October 2015

Spirulina/Arthrospira is a species of cyanobacteria used in health foods, animal feed, food additives and fine chemicals. The present study conducted a comparison of the 16S rRNA and cpcBA-intergenic spacer (cpcBA-IGS) gene sequences in *Spirulina/Arthrospira* strains from culture collection of CCUBGA, IARI, New Delhi. All the strains of *Spirulina* used in this study had shown nearly 99% similarity amongst them. About fifty sequences (cpcBA-IGS) of *Spirulin*a strains taken from NCBI with ten from the present strains of *Spirulina*, a neighbour-joing (NJ) tree was constructed with the help of MEGA5.0. The tree showed 99% similarity. All the sequences were put to Multiple Sequence Alignment with the help of T-Coffee (version 7.38) and BioEdit (version 7.38) software. Similarity studies undertaken based upon 16S rRNA and cpcBA-IGS genes sequence analysis indicated similarity coefficient of 0.84. *S. platensis* and *Arthrospira* sp. showed 100 percent similarity. Therefore, the current study supports some previous conclusions based on 16S rRNA gene and cpcBA-IGS sequences, which found that *Arthrospira* taxa are monophyletic. However, compared to 16S rRNA sequences, cpcBA-IGS sequences might be better suited to resolve close relationships and interspecies variability.

Keywords: *Spirulina*, cpcBA-IGS, genetic diversity, 16S rRNA

Introduction

Spirulina/Arthrospira is a commercially important filamentous cyanobacterium with an annual production estimated to be over 3,000 tons per year, the largest among microalgae[1,2]. It is found in tropical and subtropical regions, in warm lakes with high carbonate and bicarbonate content, and high pH and salinity. *Arthrospira* is a rich source of proteins, minerals, vitamin B12, ß-carotene, and essential fatty acids, such as, linolenic acid. Owing to its high protein content (up to approx 60-70% on a dry wt basis), the amino acids present in *Arthrospira* match the proportions recommended by the Food and Agriculture Organization (FAO)[3,4]. The cyanobacteria are a morphologically distinct group of oxygenic photosynthetic organisms that inhabit terrestrial and aquatic ecosystems. Their versatility lie in their ability to grow in extreme environments, their significant role in the biogeochemical cycle and the production of bioactive compounds. Initially, cyanobacteria were classified as prokaryotic algae due to their phototrophic nature. Later, the taxonomic assignments were based on morphological and cytological characteristics. Since cyanobacterial morphology is strongly influenced by environmental stimuli at individual sampling locales, it has been difficult to classify cyanobacteria in appropriate taxonomic groups. It has been estimated that as many as 50% of cyanobacterial strains found in culture collections have been misidentified leading to erroneous organism phylogenies[5].

DNA sequences play an essential role in the reconstruction of evolutionary relationships among organisms and have led to new genetic classifications that may confirm or conflict with traditional taxonomy. Application of molecular techniques to amplify some portions of the genome in order to characterize and deduce phylogenetic relationships of cyanobacteria has increased considerably in the recent years[6]. At the molecular level, the rRNA genes are the most widely used markers for the identification of bacteria and cyanobacteria due to their conserved function and universal presence. Several researchers have exploited the conserved regions of the 16S rRNA gene for phylogenetic analysis of cyanobacteria. However, the application of 16S rRNA to identify at the species level and below has been contested[7]. Moreover, the conserved nature of 16S rRNA and the lower

*Author for correspondence:
devendra2228@gmail.com

evolutionary rate variation compared to the protein-encoding genes makes it less useful for phylogenetic studies of closely related organisms. As a result, researchers have targeted other variable regions, such as, the 16S-23S internal transcribed spacer region (ITS) and the intergenic spacer region (IGS) of the phycocyanin (PC) locus[8]. In cyanobacteria, the 16S-23S ITS region has variable tRNA genes both in terms of length and nucleotides. Therefore, the sequence information of this region can be a valuable tool for use in phylogenetic investigations. However, the existence of multiple rRNA operons in some cyanobacteria has to be considered in those organisms during application in phylogenetic studies[9]. Yet another phylogenetic marker commonly employed by phycologists is the phycocyanin operon, which includes the genes responsible for coding of two phycobiliprotein subunits (cpcB & cpcA) and three linker polypeptides. The substitution rate of the nucleotides in this region is higher than that of 16S rRNA and, thus, can be used as a potential genetic marker for phylogenetic studies. The coding regions show little sequence divergence among closely related species, whereas the spacer regions may exhibit perceptible variability. Therefore, the conserved coding regions of the phycocyanin locus can be used for comparison among cyanobacterial species, while the highly variable spacer region may be used to discriminate strains. Several studies have indicated intragenic recombinations and possible exchanges of genetic material between cyanobacterial strains within the phycocyanin operon[10]. These observations call into question the use of the PC-IGS region alone for phylogenetic analyses, especially for phyletically related strains susceptible to homologous recombination. To avoid the complex evolutionary pattern of the PC-IGS locus within Cyanobacteria, inclusion of additional gene loci during phylogenetic analyses has been suggested[11]. At the National Facility for Marine Cyanobacteria (NFMC), Bharathidasan University, Tiruchirappalli, India, the cyanobacterial cultures collected across different geographical locations within the Indian subcontinent were characterized based on their morphological features.

The aim of the present study was to clarify the diversity of cultivated strains of *Spirulina*/*Arthrospira* using sequence data from a highly variable DNA fragment, including a comparison of the phylogeny of *Spirulina*/*Arthrospira* strains based on the 16S rRNA gene and cpcBA-IGS. Besides, to define and delimit the genus *Spirulina*/*Arthrospira*, other cyanobacteria were also included. DNA divergence studies showed significant values of divergence with greater gene conversion tracts in the unbranched (Nostocales) compared to the branched (Stigonematales) strains. The present study hence represents a unique blend of molecular phylogeny with evogenomic sequence analyses for understanding the genetic diversity, phylogeny, and evolutionary pace within the heterocystous Cyanobacteria.

Materials and Methods

Growth and Maintenance of Culture

Axenic *Spirulina/Arthrospira* strains were procured from the culture collection of Centre for Conservation and Utilization of Blue Green Algae (CCUBGA), IARI, New Delhi, India (Fig. 1, Table 1). Cultures

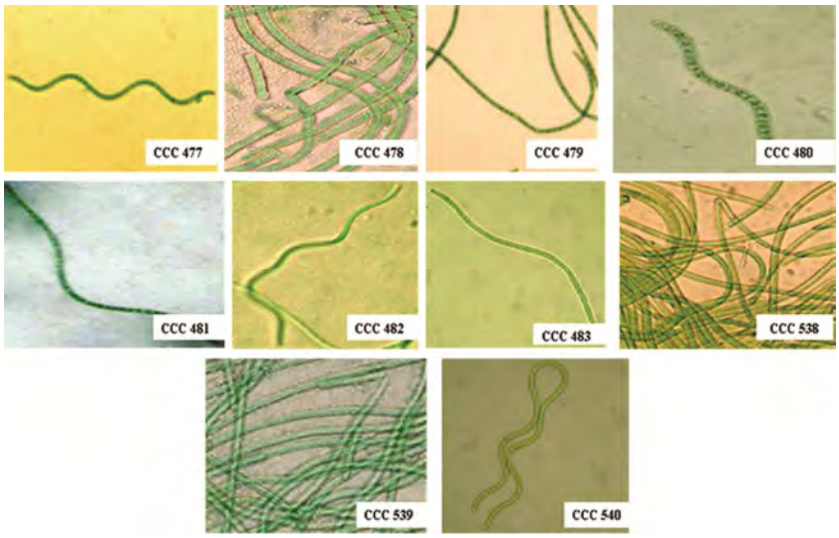

Fig. 1 — Photomicrograph of *Spirulina/Arthrospira* strains.

Table 1 — Geographical origin and cultural characteristics of *Spirulina* with *Arthrospira*

Strain no.*	Taxonomic description	Origin/Source	Growth pattern		
			Liquid medium	Solid medium	Colour of thallus
CCC 477	*S. platensis*	West Germany	Uniform suspension	Spreading	Blue green
CCC 478	*S. platensis*	Israel	Uniform suspension	Spreading	Dark green
CCC 479	*S. platensis*	Vietnam	Uniform suspension	Spreading	Green
CCC 480	*S. platensis*	CFTRI, Mysore, India	Uniform suspension	Spreading	Blue green
CCC 481	*S. maxima*	China	Planktonic	Spreading	Blue green
CCC 482	*S. lonar*	Israel	Planktonic	Spreading	Blue green
CCC 483	*S. platensis* (mutant)	CCUBGA, IARI, New Delhi, India	Uniformb suspension	Spreading	Blue green
CCC 538	*Arthrospira* sp.	Canary Islands, Spain	Planktonic	Spreading	Pale green
CCC 539	*Arthrospira* sp.	Canary islands, Spain	Planktonic	Spreading	Blue green
CCC 540	*Spirulina* sp.	Loktak Lake, Manipur, India	Uniform suspension	Spreading	Blue green

*CCC= Cyanobacterial Culture

were maintained in chemically defined Z-medium[12] at 28±2°C under the light intensity of 52-55 µmol photon m^{-2} s^{-1} and L:D cycles of 16:8 h. Protocol was optimized for higher mass production of *Spirulina/Arthrospira* strains[13].

DNA Extraction, Amplification and Sequencing

Extraction of DNA was performed on the cell pellet obtained from 2.0 mL of exponential growth phase cultures centrifuged at 14,000 rpm for 10 min. DNeasy Tissue Kit manufacturer's protocol (Qiagen, Cat. No. 69504) was used and the protocol was slightly modified at specific steps for efficient recovery of DNA. The extracted DNA preparation was quantified by taking absorbance at 260 nm. The 16S rRNA gene was then amplified using the universal primers FD1 (5′-AGAGTTTGATCCTGGCTCAG-3′) and RP2 (5′-ACGGCTACCTTGTTACGACTT-3′)[14] by modified PCR protocol (Fig. 2). For the 16S, the thermal cycling was performed with an initial denaturation at 94°C for 5 min, followed by 35 cycles of 94°C for 30 sec, 64°C for 45 sec and 72°C for 2 min, and a final extension of 72°C for 5 min. The cpcBA-IGS was amplified using primers located within the coding region of cpcB and cpcA, respectively, where the forward primer was cpc_arF-TCGAAGATCGTTGCTTGAACG and the reverse primer was cpc_arR- TTAGGTCCCTGCATTTGGGTG[15] (Fig. 3). The PCR reaction was performed using Ready-to-Go PCR Beads (Amersham Biosciences, Sweden) containing 10 mM Tris-HCl, pH 9.0, 50 mM KCl, 1.5 mM MgCl2, 200 µM of each dNTP, 0.30 µM of each primer, 1.25 units of Taq DNA polymerase and 100 ng of template DNA in a 25.0 µL

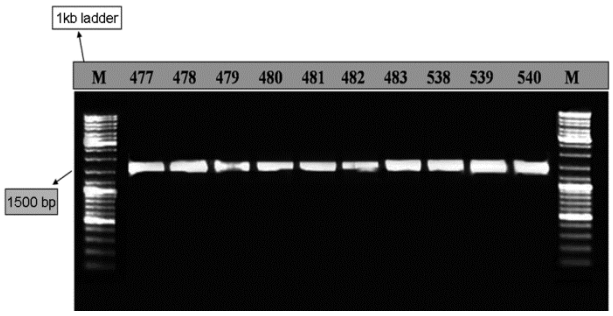

Fig. 2 — PCR amplification of 16S rRNA gene in *Spirulina/Arthrospira* strains.

reaction using a Progene thermal cycler (Techne Cambridge Ltd., UK). The cycling profile for the cpcBA-IGS region included an initial denaturation at 95°C for 5 min, followed by 35 cycles of 95°C for 25 sec, 60°C for 15 sec and 72°C for 30 sec, and a final extension at 72°C for 5 min. About 5 µL of the amplified products was subjected to 1.5% agarose gel electrophoresis, stained with ethidium bromide and visualized using a UV transilluminator. PCR products were purified and then directly sequenced at a commercial facility (DNA Sequencing Facility, Department of Biochemistry, University of Delhi South Campus, New Delhi, India). DNA sequencing was performed with ABI PRISM Big Dye Terminator Kit (Perkin Elmer) and an ABI PRISM (Model-3730, Version-3.0) genetic analyzer (Perkin Elmer) according to manufacturer's instruction. Sequence readings were performed giving partial 16S rRNA sequence of about 1500 bp and cpcBA-IGS of 600 bp. The nucleotide sequences described in this study were submitted to the NCBI under GenBank accession numbers JQ926188 to JQ926196 for cpcBA-IGS and JX014313 (CCC478) for 16S rRNA.

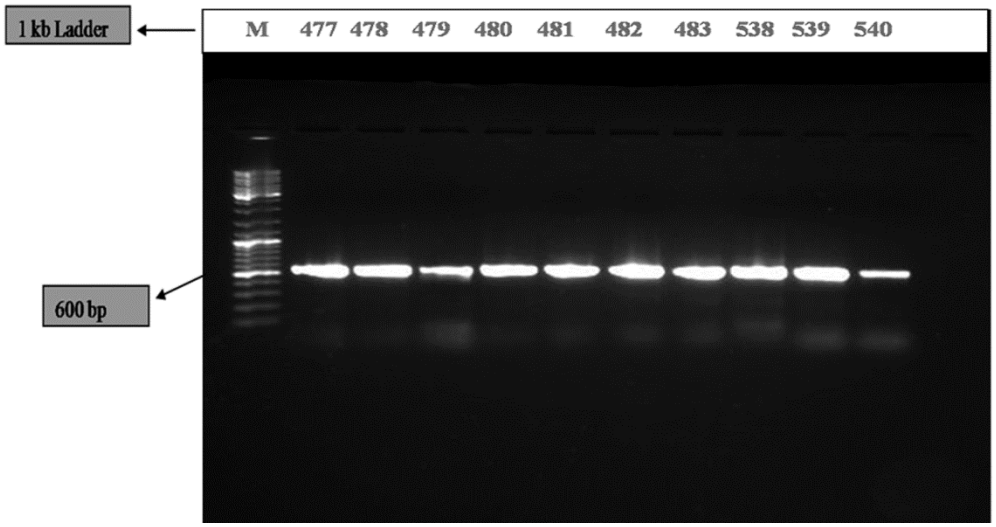

Fig. 3 — PCR amplification of cpcB-IGS-cpcA gene in *Spirulina/Arthrospira* strains.

Phylogenetic Analysis

The 16S rRNA gene and cpcBA-IGS sequences obtained from the *Spirulina/Arthrospira* strains were initially compared with sequences available in the National Center for Biotechnology Information (NCBI) database using BLAST network services (http://www.ncbi.nlm.nih.gov/BLAST) to determine their approx phylogenetic affiliations[16]. The sequences were aligned using PHYDIC 3.0, and unambiguously aligned nucleotide positions then used for phylogenetic analyses using MEGA.5. The similarity values between the sequences were calculated from distance matrices by reversing the Jukes-Cantor distance formula[17]. Phylogenetic trees were then inferred by Neighbour Joining (NJ) using the Kimura two-parameter model. The resulting NJ tree was evaluated by bootstrap analyses based on 1000 resembling. Due to the spacer variability, a phylogenetic analysis of the matrix was also performed using just the two coding regions. Finally, an overview of the phylogenetic position of *Spirulina* in cyanobacteria was created by comparing the 16S rRNA gene and cpcBA-IGS sequences to corresponding cyanobacterial sequences available in databases and the sequences obtained in this study for *S. laxissima* SAG 256.80 and *Oscillatoria sancta* NIER 10027.

Results and Discussion

Near-complete 16S rRNA gene sequences of *Spirulina* strains received from culture collections across the world were determined. A phylogenetic tree was then reconstructed using a NJ analysis based on aligning the all sequences with *E. coli* K-12 as an outgroup and Nostocales (*Anabaena*) as other cyanobacterial group[18] (Fig. 4). The corrected sequence alignment, providing the basis of the phylogenetic analyses, corresponded to positions 8-1512 according to the *E. coli* numbering system and was 1,383 nucleotides (nt) in length after removing all gaps and ambiguous positions. Clade I and clade II are apart from each other because they are belonging to different genus.

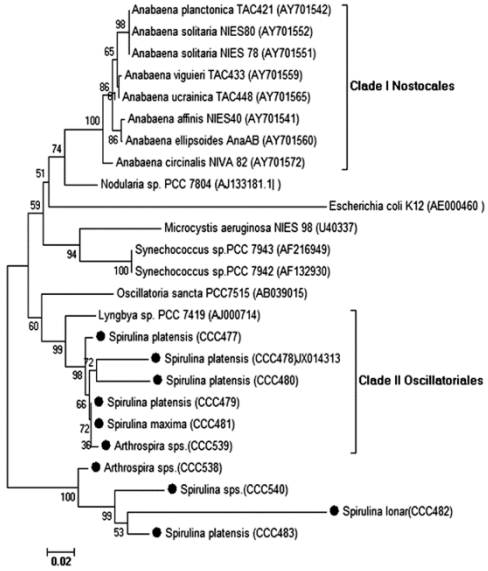

Fig. 4 — Phylogenetic tree showing relationship of 16S rRNA gene sequences of *Spirulina/Arthrospira* strains with other cyanobacteria using Neighbor Joining Program in MEGA 5.0. [Bootstrap values (≥50%) are indicated on the branches. Sequences determined in the present study is highlighted.]

The cluster analysis resolved the selected *Spirulina/Arthrospira* strains into two main genotypic clusters, designated Clade I and II. Clade I contained

all Nostocales members and Clade II contained all Oscillatoriales, namely, CCC477 to CCC540. The bootstrap value between the Clade I and Clade II clusters was 60% in the phylogenetic tree, and the 16S rRNA gene similarity was 99.5%. *E. coli* K-12 was shown as outgroup in the tree. Thus, the clusters were poorly supported by the bootstrap analysis. In the 16S rRNA gene sequences for the strains, of 1,420 nt, the number of different nucleotides was less than 7. In a previous report, a complete analysis of the dendrogram structure grouped the strains into two well-separated genotypic groups. The genotypic diversity of several strains attributed to these two species was also previously investigated on the basis of morphological criteria using a very sensitive total DNA restriction profile analysis. In this case, the strains were also divided into two well-separated genotypic groups[19].

The similarity of the 16S rRNA genes between the *Spirulina* strains and *Lyngbya aestuarii* PCC 7419 was about 98%. However, the similarity of the 16S rRNA genes among the *Spirulina* strains, and with *Microcystis aeruginosa* NIES98 (U40337), *Synechococcus* sp. PCC7943 (AF216949) and *O. sancta* PCC7515 (AB039015) was 99%. Thus, as shown by the 16S rRNA gene sequences, it would seem that *L. aestuarii* PCC 7419 is also closely related to *Arthrospira* and a sister to the clade *Planktothrix/ Arthrospira.* All the strains of *Spirulina/Arthrospira*, used in the present study showed nearly 99% similarity amongst themselves and were also nearly 98% similar with members of clade I (Nostocales)[20] (Fig. 5). All the sequences were subjected to multiple sequence alignment with the help of T-Coffee (version 7.38) and BioEdit (version 7.1.9) softwares.

The analyses of the studied *Spirulina/Arthrospira* were conducted using both the coding sequences and the spacer, and the outgroup was the *cpcBA*-IGS from the chloroplast of *Cyanidium caldarium*. The NJ tree derived from the translated cpcB-IGS-cpcA sequences clustered the *Spirulina/Arthrospira* into two clades, clade I (Oscillatoriales) and clade II (Nostocales) where clade I clustered with the clade II clusters in 100% of the bootstrap trees and the cpcB-IGS-cpcA similarity was more than 98%. In clade I, two groups are shown; Group I contained CCC477, CCC479, CCC483, CCC538, CCC540, *A. platensis* FACHB439 (AY244669) and *Arthrospira* sp. (AJ310554), while Group II exhibited CCC478, CCC480, CCC481, CCC482, CCC539, *Arthrospira* sp. PK (AJ401179), *S. subsalsa* FACHB351 (AY244667 and *Arthrospira* sp. (FJ001915), and they represented 99% of the bootstrap replications[21] (Fig. 6).

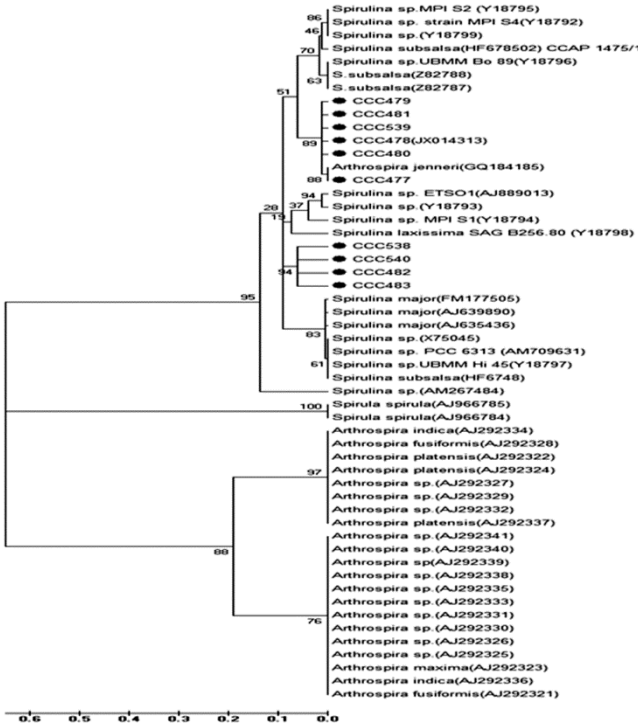

Fig. 5 — Phylogenetic tree showing relationship of 16S rRNA gene sequence of *Spirulina* strains with other *Spirulina/ Arthrospira* strains from NCBI database by using Neighbor Joining Program in MEGA 5.0. [Bootstrap values (≥50%) are indicated on the branches. Sequences determined in the present study is highlighted.]

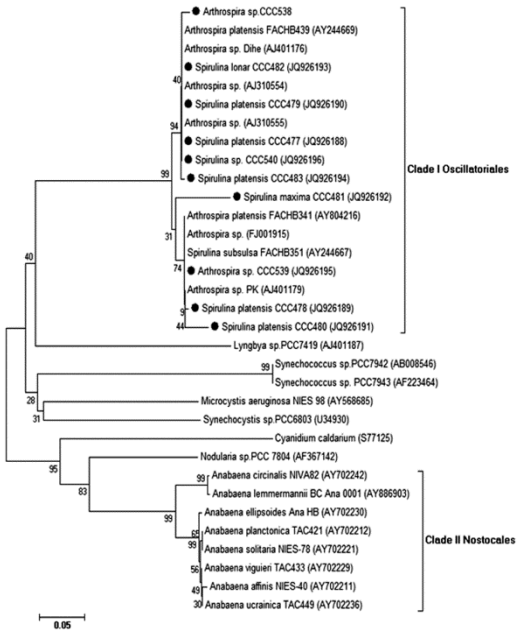

Fig. 6 — Phylogenetic tree showing relationship of cpcB-IGS-cpcA gene sequence of *Spirulina/Arthrospira* strains with other cyanobacteria using Neighbor Joining Program. [Bootstrap values (≥50%) are indicated on the branches. Sequences determined in the present study is highlighted.]

Fig. 7 — Phylogenetic tree showing relationship of cpcB-IGS-cpcA gene sequence of *Spirulina* strains with other *Spirulina/Arthrospira* strains from NCBI database by using Neighbor Joining Program in MEGA 5. [Bootstrap values (≥50%) are indicated on the branches. Sequences determined in present study is highlighted.]

All the strains of *Spirulina/Arthrospira*, which were used in the present study, showed nearly 99% similarity amongst themselves and were also nearly 98% similar with members of Clade II (Nostocales). *Spirulina/Arthrospira* strains used in the present study also showed 99% similarity with 50 sequences (cpcB-IGS-cpcA) of *Spirulina* strains from NCBI database on the basis of NJ tree developed with the help of MEGA 5.0 (Fig. 7). All the sequences were subjected to Multiple Sequence Alignment with the help of T-Coffee (version 7.38) and BioEdit (version 7.1.9) softwares. Both *Arthrospira* and *Spirulina* are the member of Cyanophyceae family but in some text, *Spirulina* is considered as a supplement material of *A. platensis.* According to some reports, rpoC1 gene has more advantage to distinguish the strains in the same genus than that of 16S rRNA gene.

Acknowledgement

The facilities provided by the CCUBGA, IARI, New Delhi for carrying out the present investigation is gratefully acknowledged. The authors are also thankful to the Director, IARI, New Delhi for giving permission to carry out the study at IARI and Indian Council of Agricultural Research, New Delhi for financial support.

References

1 Pulz O, Gross W, Valuable products from biotechnology of microalgae, *Appl Microbiol Biotechnol*, 65 (2004) 635-648.

2 Vonshak A, Outdoor mass production of *Spirulina*: The basic concept, in *Spirulina platensis* (*Arthrospira*): *Physiology, cell biology and biotechnology*, edited by A Vonshak (CRC Press, Taylor & Francis Ltd., London) 1997, 79-99.

3 Tomaselli L, Morphology, ultrastructure and taxonomy of *Arthrospira* (*Spirulina*) *maxima* and *Arthrospira* (*Spirulina*) *platensis*, in *Spirulina platensis (Arthrospira*): *Physiology, cell biology and biotechnology*, edited by A Vonshak (CRC Press, Taylor & Francis Ltd., London) 1997, 1-15.

4 Viti C, Ventura S, Lotti F, Capolino E, Tomaselli L *et al*, Genotypic diversity and typing of cyanobacterial strains of the genus *Arthrospira* by very sensitive total DNA restriction profile analysis, *Res Microbiol*, 148 (1997) 605-611.

5 Neilan B A, Jacobs D & Goodman A E, Genetic diversity and phylogeny of toxic cyanobacteria determined by DNA polymorphisms within the phycocyanin locus, *Appl Environ Microbiol*, 61 (1995) 3875-3883.

6 Anagnostidis K & Komarek J, Modern approach to the classification system of cyanophytes 1. Introduction, *Archiv Hydrobiol Suppl*, 71 (1985) 291-302.

7 Manen J F & Falquet J, The cpcB-cpcA locus as a tool for the genetic characterization of the genus *Arthrospira* (Cyanobacteria): Evidence for horizontal transfer, *Int J Syst Evol Microbiol*, 52 (2002) 861-867.

8 Li X, *Phenotypic and genotypic characterization of Nostoc species from 6 different sites in the Mojave desert*. Master's Thesis, John Carroll University, Cleveland, Ohio, 2000.

9 Ishida T, Yokota A & Sugiyama J, Phylogenetic relationships of filamentous cyanobacterial taxa inferred from 16S rRNA sequence divergence, *J Gen Appl Microbiol*, 43 (1997) 237-241.

10 Kim S G, Rhee S K, Ahn C Y, Ko S R, Choi G G *et al*, Determination of cyanobacterial diversity during algal blooms in Daechung Reservoir, Korea, on the basis of cpcBA intergenic spacer region analysis, *Appl Environ Microbiol*, 72 (2006) 3252-3258.

11 Wilmotte A & Golubic S, Morphological and genetic criteria in the taxonomy of cyanophyta/cyanobacteria, *Archiv Hydrobiol Suppl*, 64 (1991) 1-24.

12 Zarrouk C (1966), Contribution a l'etud d'une facteure physiques et la. Photosynthesis de *Spirulina platensis* (Setch et Gardner) Geitler. Ph D Thesis, University of Paris, France.

13 Kumar D, Dhar D W, Pabbi S, Kumar N & Walia S, Protocol optimization for enhanced production of pigments in *Spirulina*, *Indian J Plant Physiol*, 18 (2013) 308-312.

14 Otsuka S, Suda S, Li R, Watanabe M, Oyaizu H *et al*, 16S rDNA sequences and phylogenetic analyses of *Microcystis* strains with and without phycoerythrin, *FEMS Microbiol Lett*, 164 (1998) 119-124.

15 Weisburg W G, Barns S M, Pelletier D A & Lane D J, 16S ribosomal DNA amplification for phylogenetic study, *J Bacteriol*, 173 (1991) 697-703.

16. Singh P, Singh S S, Elster J & Mishra A K, Molecular phylogeny, population genetics and evolution of heterocystous cyanobacterial using *nifH* gene sequences, *Protoplasma*, 250 (2013) 751-764.

17 Kumar D, Dhar D W, Pabbi S, Kumar N & Walia S, Extraction and purification of C-phycocyanin from *Spirulina platensis* (CCC_{540}), *Indian J Plant Physiol*, 19 (2014) 184-188.

18 Catling D C, Zahnle K J & McKay C, Biogenic methane, hydrogen escape, and the irreversible oxidation of early earth, *Science*, 293 (2001) 839-843.

19 Colla L M, Reinehr C O, Reichert C & Costa J A V, Production of biomass and nutraceutical compounds by *Spirulina platensis* under different temperature and nitrogen regimes, *Bioresour Technol*, 98 (2007) 1489-1493.

20 Altschul S F, Madden T L, Schäffer A A, Zhang J, Zhang Z *et al*, Gapped BLAST and PSI-BLAST: A new generation of protein database search programs, *Nucleic Acids Res*, 25 (1997) 3389-3402.

21 Dwivedi P P, Patel B K C, Rees G N & Ollivier B, A rapid method for sequencing of rRNA gene(s) amplified by polymerase chain reaction using an automated DNA sequencer, *Indian J Microbiol*, 36 (1996) 9-12.