# Genomic Fingerprinting Using Highly Repetitive Sequences to Differentiate Close Cyanobacterial Strains

*Rezvan Shokraei[a], Hossein Fahimi[a], Saúl Blanco[b], Bahareh Nowruzi[c] **

## Abstract

***Background***: Cyanobacterial taxonomy has experimented considerable changes due to the exploration of previously uninvestigated regions as well as the introduction of molecular tools. Challenges arose when strains collected from agricultural areas, salt waters and dry limestone did not reveal remarkable morphological differences and had a high level of similarity in the phylogeny of 16S rDNA gene sequences. The aim of the present investigation was to fingerprint members of the genera *Calothrix* and *Nostoc* based on the repetitive DNA sequences, as molecular markers for the detection of phylogenetic affinities and molecular diversity.

***Methods***: In this research, through a polyphasic approach, the differences in morphological and genotypic features of different strains were investigated. Bacteria free cyanobacterial clones were prepared followed by morphological characterization, genomic DNA extraction and PCR with 16S rRNA, ERIC, STRR1a and HIP primers. Then the phylogenetic analyses of partial 16S rRNA genes and fingerprints were performed.

***Results***: The results showed each marker producing unique and strain-specific banding pattern, thus highlighting the efficiency of this technique in the assessment of proximity between closely related cyanobacterial strains isolated from different climatic/geographic regions and habitats.

***Conclusions***: This case is the first documented genomic fingerprinting from seven close cyanobacterial strains in Iran.

***Significance*** | Genomic fingerprinting using highly repetitive sequences to differentiate close cyanobacterial strains

*Correspondence: **Bahareh Nowruzi**, PhD, Assistant Professor, Department of Biology, Science and Research Branch, Islamic Azad University, Tehran, Iran. E-mail: bahareh.nowruzi@srbiau.ac.ir; Contact no.: 00989113710956

Edited by **Md. Asaduzzaman Shishir, PhD, University of Dhaka, Dhaka, Bangladesh**, and accepted by the Editorial Board January 24, 2019 (Received for review December 26, 2018)

**Keywords:** Fingerprinting, Repetitive DNA fragments, Enterobacterial repetitive intergenic consensus (ERIC), Highly iterated palindrome, Close cyanobacteria.

**Abbreviations:** RFLP, Restriction Fragment Length Polymorphism; RAPD, Random Amplification of Polymorphic DNA; STRR, Short Tandemly Repeated Repetitive; HIP, Highly Iterated Palindromes; ERIC, Enterobacterial Repetitive Interspersed Consensus; EtBr, Ethidium Bromide; UV, ultraviolet; BLAST, Basic Local Alignment Search Tool; CCC, Cyanobacteria Culture Collection; *rpo*C, DNA-directed RNA polymerase subunit.

## Introduction

Cyanobacteria are an aged group of photosynthetic prokaryotes, their first molecular carbon skeletons can be identified in strata from around 2.75 billion years ago (Gould et al., 2008). The progenitors of current cyanobacteria originated oxygenic photosynthesis some 3. 6 billion years ago (Gould et al., 2008). Evaluative studies on firmly related cyanobacteria indicate rapid and highly variable gene fluxes in evolving microbial genomes (Walter et al., 2017). Cyanobacteria comprise both unicellular and colonial (including filamentous) forms. Taxonomically, cyanobacteria are classified into unicellular forms featuring binary fission (Order Chroococcales, or Bergey's Subsection I) or multiple fission (Order Pleurocapsales, or Bergey's Subsection II); and filamentous forms that are non-heterocystous (Order Oscillatoriales, or Bergey's Subsection III) or show heterocysts in non-branching (Order Nostocales, or Bergey's Subsection IV) or branching filaments (Order Stigonematales, or Bergey's Subsection V). A sixth cyanobacterial order, Gloeobacterales, was proposed by Cavalier-Smith (2002)- to accommodate the genus *Gloeobacter*, formerly included in the Chroococcales (Komárek, 2013).

**Author Affiliation:**
[a] Department of Genetics, Faculty of Advanced Science and Technology, Tehran Medical Sciences, Islamic Azad University, Tehran, Iran.
[b] Department of Biodiversity and Environmental Management, Faculty of Biological and Environmental Sciences, University of Leon, León, Spain.
[c] Department of Biology, Science and Research Branch, Islamic Azad University, Tehran, Iran.

**Please cite this article:**
Shokraei R, Fahimi H, Blanco S, Nowruzi B (2019). Genomic Fingerprinting Using Highly Repetitive Sequences to Differentiate Close Cyanobacterial Strains. *Microbial Bioactives*, 2(1), 068-075

Cyanobacteria represent a difficult group for the microbiologists. Their conventional taxonomy, in view of morphogenesis traits, doesn't reflect the results of phylogenetic analyses (Howard-Azzeh et al., 2014, Walter et al., 2017). The predominance of morphologic criteria assembled unrelated cyanobacteria under polyphyletic taxa which will require revisions later on (Komárek and Johansen., 2014). Morphologically similar strains might contrast extraordinarily at the molecular level and vice-versa. In a few occasions it will be not challenging distinguish cyanobacterial isolates to the genus level, especially where morphologic aspects are distinctive, e. g. *Calothrix* or *Nostoc*. However, to a number genera, including *Oscillatoria*, *Lyngbya*, and *Phormidium*, it may be frequently challenging to the non-expert to approach to convinced diagnoses. Identification problems expand further at the species level and little may be known about sub particular variability at the strain level. Despite these paramount traits and the expanding enthusiasm toward developing cyanobacterial strains for biotechnology, there is a shortage and disturbed dispensation of publicly accessible genomic data on Cyanobacteria. Improvements in the coverage of sequenced genomes will empower more accurate understanding of cyanobacterial niche-adaptation and evolution (Howard-Azzeh et al., 2014, Sánchez-Baracaldo et al., 2014, Schirrmeister et al., 2015, Shih et al., 2013, Uyeda et al., 2016).

Developments in molecular biology and bioinformatics permit mining the genome of an organism for the presence of unique sequences that can be used for recognizing a specific group of microorganism from its close relatives. Similar techniques have been developed based on primers that hybridize with repeated sequence structures present in bacterial DNA. These primers permit amplification of the DNA sequences between those adjacent repeated sequences happening in a suitable orientation and distance apart. Additionally, PCR-based techniques largely dependent on DNA polymorphism and fingerprinting of repetitive DNA fragments have been developed (Elhai, 2015). RFLP (Restriction Fragment Length Polymorphism)(Iteman et al., 2002), RAPD (Random Amplification of Polymorphic DNA) (Prabina et al., 2005; Shishir et al., 2015), STRR (Short Tandemly Repeated Repetitive)(Akoijam and Singh, 2011, Valerio et al., 2009, Wilson et al., 2000), HIP1 (Highly Iterated Palindromes) (Neilan et al., 2003, Orcutt et al., 2002, Wilson et al., 2005, Zheng et al., 2002) and ERIC (Enterobacterial Repetitive Interspersed Consensus) (Valério et al., 2005) have been attempted with an overall aim to provide better resolution among closely related species. These repetitive sequences were diagnosed in several cyanobacterial taxa, up to now broadly in heterocystous cyanobacteria (Lyra et al., 2005, Nilsson et al., 2000, Prasanna et al., 2006, Rasmussen and Svenning, 1998, Teaumroong et al., 2002, Wilson et al., 2000, Zheng et al., 1999). but also in some non-heterocystous ones (Rasmussen and Svenning, 1998). The conserved status of these repetitive sequences have made them ideal tools for diversity studies, and have brought forth a brand new PCR-based technique known as the rep-PCR technique that utilizes oligonucleotide-derived repetitive sequences present in bacterial strains to separate firmly related members of the same genus. This technique has been enormously productive in discriminating members of several eubacterial genera (Laguerre et al., 1996, Rodriguez-Barradas et al., 1995). The genera *Calothrix* and *Nostoc* are filamentous cyanobacteria belonging to the Order Nostocales. Members of this order exhibit a high level of morphological complexity (Nowruzi et al., 2012) and the incredible morphotype diversity observed in nature is typically underrepresented in cultures (Gugger and Hoffmann, 2004).

The aim of the present investigation was to fingerprint these strains using PCR based on the repetitive DNA sequences and the 16S rRNA gene as molecular markers to resolve close cogeneric strains, also we used these markers together with the phylogenetic affinities, to differentiate seven heterocystous cyanobacteria sharing similar stress tolerance profiles.

## Materials and methods

### Isolation and maintenance of clonal and axenic cultures of heterocystous cyanobacteria

Dry limestones harboring different *Nostoc* species have been observed in the North-west Mountains of Iran. Such dry and sometimes cold conditions are suitable for *Nostoc* growths (Helm et al., 2000, Hill et al., 1994, Potts, 2000, Shirkey et al., 2003).

Samples were collected from Cretaceous nodular chalk limestone rocks on the cliff face in the North-west Mountains of Khuzestan province, Iran (34°25′04″ N, 47°00′59″ W). Rocks were collected from the upper greensand layer, where limestone is predominant, together with glauconitic inclusions. *Nostoc* inhabits the surface and interior of the rocks, forming a homogenous epilithic covering. For the exposure experiments, rocks were cut into blocks with an upper surface area of 5 $cm^2$. Additionally, soil samples with different textures (according to the pedological map of the Kermanshah province) were selected and collected from agricultural areas (34°24′32″ N, 47°00′17″ W). Samples were collected from the surface up to five cm deep with a sterilized spatula after removing surface debris. Finally, saltwater samples (36°54′41″ N, 54°47′25″ W) were collected at *ca.* 30 cm depth and 1 m away from the shore using cone-shaped bottles. Samples were transferred to sterile Petri dishes with a suitable amount of BG11 nitrate-free liquid and solid media (Rippka et al., 1979). pH was adjusted to 7.1 after sterilization, and Petri dishes were incubated in a culture chamber at 28 °C, supplied with continuous artificial illumination (~1500–2000 lux) for two weeks (Kaushik et al., 2009). After 14 days, one or two colonies were isolated for purification, washed thrice with deionized water and transferred to fresh solid media. In order to keep bacteria-free cultures, the colonies were isolated and tested for bacterial contamination in dextrose-peptone broth and caseinate-glucose agar media. Thereafter, bacteria free clones were selected and maintained on different agar slants. No further analysis of the cultures was done until pure clonal cultures were established and examined microscopically.

### Morphological characterization

Cyanobacteria were characterized using the standard keys by Desikachary (1959) and Komárek (2013), considering only traditional morphological groups. Samples were analyzed in two stages: a first one in natural conditions and a second one when the growth phase begins under laboratory conditions. This ensures no differences between naturally occurring samples and laboratory-grown cultures. Phenotypic characters particularly observed were the shape and dimensions of vegetative and specialized cells (heterocytes, akinetes, baeocytes, arthrospores, hormocytes etc.), and some other features were observed depending on the type of cyanobacteria: division plane, branching and branching pattern, a polarity of trichomes, etc. Some

samples of Nostocaceae were grown for a longer period so as to document life cycle changes if any.

**Genomic DNA isolation and PCR conditions**

DNA was isolated from 8-10 days old cultures using the EZNA® SP Plant DNA kit (Omega Bio-Tek). Microtubes containing 100 mg wet cells were filled with 300 mg of two differently-sized acid-washed glass beads (180 and 425–600 µm in diameter, Sigma-Aldrich), adding lysis buffer and RNase solution as provided by the kit. In order to ensure proper disruption of the cells, tubes were homogenized three times for 20 s at 6.5 $ms^{-1}$ with a FastPrep homogenizer (Savant Instruments). The extraction procedure continued following the manufacturer protocol. DNA was quantified with a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Inc). 16S rRNA gene amplifications were done using the primers pA (5'-AGAGTTTGATCCTGGCTCAG-3') and B23S (5'-CTTCGCCTCTGTGTGCCTAGGT-3') (Taton et al., 2003). The PCR reaction starts with 1× Buffer solution (DyNAzyme™ PCR buffer, Finnzymes), 0.5 µM of forward primer, 0.5 µM of reverse primer, 0.5 U of *Taq* polymerase (DyNAzyme™ II DNA polymerase, Finnzymes), and 1 µL of template DNA, with sterile water until a total volume of 20 µL. Amplification reactions were conducted in a thermocycler (iCycler, Bio-Rad) with the following program: initial denaturation at 94 °C for 3 min, 30 cycles of denaturation at 94 °C for 30 s, annealing at 55 °C for 30 s and at 72 °C for 30 s, and a final annealing phase at 72 °C for 5 min. In order to amplify repetitive DNA fragments, reactions were performed in 25 µL aliquots containing 10-20 ng of DNA template, 0.5 µM of each ERIC, STRR1a and HIP primers, 1.5 mM of $MgCl_2$, 200 µM of dNTPs and 1U/µL of Taq DNA polymerase. ERIC1A (5'-ATGTAAGCTCCTGGGGATTCAC-3') and ERIC1B (5'-AAGTAAGTGACTGGGGTGAGCG-3') were used as ERIC primers (Valério et al., 2005). The first cycle at 95 °C for 7 min was followed by 30 cycles at 94 °C for 1 min, at 52 °C for 1 min, at 65 °C for 8 min, one cycle at 65 °C for 16 min, and a final incubation at 4 °C for 30 min (Valério et al., 2005). For the STRR1a primer (5'-CCARTCCCCARTCCCC-3'), cycles were as follows: initial denaturation at 95 °C for 6 min, 30 cycles at 94 °C for 1 min, at 56 °C for 1 min, and at 65 °C for 5 min, with a subsequent extension at 65 °C for 16 min and a final incubation at 4 °C for 30 min (Rasmussen and Svenning, 1998). For all the HIP variants (HIP-TG: 5'-GCGATCGCTG-3', HIP-GC: 5'-GCGATCGCGC-3' and HIP-CA: 5'-GCGATCGCCA-3'), thermal cycling conditions began with an initial denaturation at 95 °C for 5 min, 30 cycles at 95 °C for 30 s, at 30 °C for 30 s, at 72 °C for 60 s, and a final cycle at 72 °C for 5 min (Smith et al., 1998).

PCR products were checked by electrophoresis on 1% agarose gels (SeaPlaque® GTG®, Cambrex Corporation) at 100 V, followed by 0.10 µg $mL^{-1}$ EtBr (Bio-Rad) staining. PCR products were visualized in the gel by UV light using a Molecular Imager® Gel Doc™ XR system (Bio-Rad). A digital image was obtained utilizing the QUANTITY ONE® 1-D V 4.6.7 analysis software. The size of the products was estimated by comparison with marker DNA (λ/HinfIII + φx/HaeIII, Finnzymes). The products were purified using the Geneclean® Turbo kit (Qbiogene, MP Biomedicals) and quantified with a Nanodrop™ ND-1000 spectrophotometer (Thermo Scientific). Sequencing of the partial 16S rRNA genes was subsequently performed using a BigDye® Terminator v3.1 cycle sequencing kit (Applied Biosystems, Life Technologies) (Elhai, 2015, Valerio et al., 2009).

**Phylogenetic analyses of partial 16S rRNA gene**

BLAST searches (http://www.ncbi.nlm.nih.gov/BLAST) for the partial 16S rRNA gene was performed to identify similar sequences deposited in the NCBI GenBank™ database. The 16S rRNA gene sequences obtained in this study, as well as reference sequences retrieved from GenBank, were first aligned with MUSCLE (Edgar, 2004) and maximum likelihood phylogenetic trees were inferred in IQ-Tree (Nguyen et al., 2014). The robustness of the tree was estimated by bootstrap percentages using 1000 replications. The root of the tree was determined using the 16S rRNA of *Aquifex aeolicus* and *Chloroflexus aurantiacus* as outgroups (Fig. 4). To prevent in group monophyly, 16S rRNA sequences of *Escherichia coli*, *Chloroflexus aurantiacus*, and *Agrobacterium tumefaciens* were included in the alignment.

**Phylogenetic analysis of fingerprints**

The generated HIP profiles were run on agarose gels with the same concentration in order to differentiate strong and doubtful signals/bands. Presence/absence of distinct and reproducible bands in each of the individual DNA fingerprinting pattern generated by HIP-AT, HIP-CA, HIP-GC, HIP-TG, ERIC, and STRR1a PCR profiles were converted into binary data (Selvakumar and Gopalaswamy, 2008), and the pooled binary data was used to construct a composite dendrogram (Abony et al., 2018). The BioDiversity Pro software (vers. 2) was used to perform the hierarchical analyses using the Jaccard cluster analysis option. All reactions were repeated three times.

**Nucleotide accession numbers**

Studied strains named *Calothrix* spp. R11 andR42, and *Nostoc* spp. FA1, FA3, FA5, F4, and F3 were registered in the DNA Data Bank of Japan (DDBJ) based on their partial 16S rDNA gene and under accession numbers MG356332, MG356333, MG385055, MG385056, MG385057, MG549315 and MG549314, respectively, and deposited at herbarium ALBORZ in Cyanobacteria Culture Collection (CCC) of the Science and Research Branch (Islamic Azad University, Iran) with herbarium numbers R11, R42, FA1, FA3, FA5, F4 and F3, respectively.

## Results

**Morphological characterization**

***Calothrix* spp. R11 and R42**

Trichomes in both strains always with basal, more or less spherical

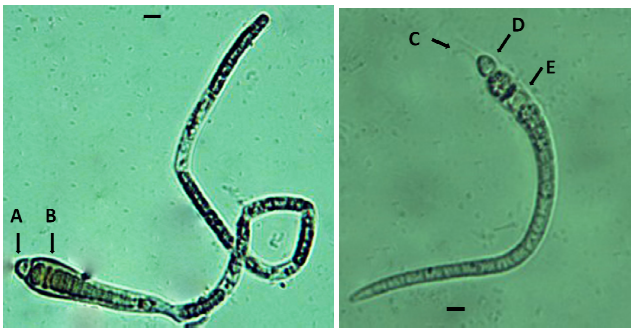


Figure 1: **Photomicrograph of *Calothrix* sp. R11 (right) and *Calothrix* sp. R42 (left). Bars, 10 µm. Spherical or hemispherical heterocysts (A and D). Thick sheaths (B). Trichomes ending in hair-like apical structures, composed of narrow, hyaline cells (C).**

or hemispherical heterocysts (A and D) (Fig. 1), yellow-brownish colored. Cells cylindrical or barrel-shaped. In *Calothrix* sp. R11, there is an immediate intercalary heterocysts near the basal heterocysts (E) (Fig. 1), while it is absent in *Calothrix* sp. R42. A distinctive feature of R42 is the presence of more swollen cells at the base of mature trichomes, with very thick sheaths (B) (Fig. 1), whereas trichomes in sp. R11 have a distinct basal-apical polarity and display a high degree of tapering. Trichomes ending in hair-like apical structures, composed of narrow, hyaline cells (C) (Fig. 1), are characteristic in *Calothrix* sp. R11. No obvious macroscopic colonies were visible on the agar plates for both strains. Microscopic colonies were light to dark green inmature *Calothrix* sp. R42, but light to dark brown in *Calothrix* sp. R11 (Fig. 1).

spherical or slightly oblong (3-5 µm broad, 4.5-7 µm long), olive and heterocysts were oval (4.5-6.5 µm broad, 4-7.5 µm long). Akinetes were rarely found (Fig. 3).

**Phylogenetic analysis of the 16S rRNA gene**

A section of the 16S rRNA gene was successfully amplified by the PCR technique. The resulting phylogenetic tree showed that, among available 16S rRNA sequences, studied strains formed three closely related clusters and were strictly separated from other members of the Nostocales clade (Fig. 4).

**Strain differentiation by Rep-PCR generated fingerprint profile**

In the present investigation, the maximal amplification was observed for the nontolerant cyanobacterial *Calothrix* sp. R11 culture. 28 amplified products with the size 400–1500 bp were obtained using

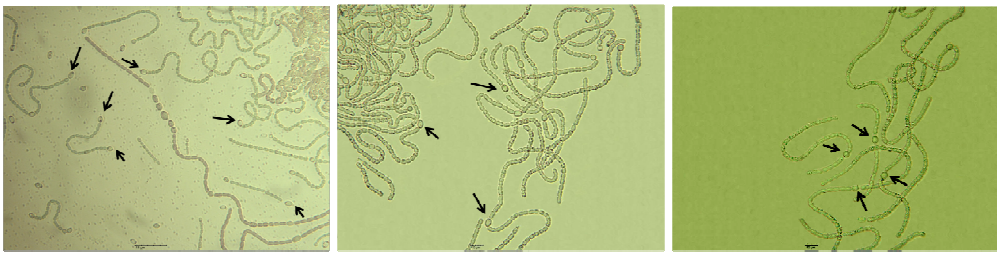

Figure 2: **Photomicrograph of *Nostoc* sp. A3 (left), *Nostoc* sp. A5 (right) and *Nostoc* sp. FA1 (middle). Bars, 10 µm. (Heterocytes are shown by arrows).**

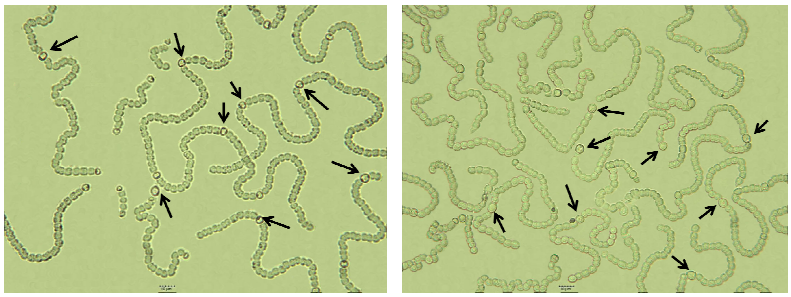

Figure 3: **Photomicrograph of *Nostoc* sp. F4 (left) and *Nostoc* sp. F3 (right). Bars, 10 µm. (Heterocysts are shown by arrows).**

***Nostoc* spp. FA1, FA3, and FA5**

The three *Nostoc* species collected from limestone did not exhibit remarkable morphological differences. Trichomes were isodiametrical throughout, composed of cylindrical and uniforms cells, 2.5-5 µm wide, 6-7 µm long, light blue-green or olive, heterocysts spherical or oblong, 4-6.5 µm wide, 7-9.5 µm long, spores ellipsoidal to oblong, 4-6.5 µm wide, 9-11 µm long (Fig. 2).

***Nostoc* spp. F4 and F3**

Cell dimensions were very similar in both strains, although in *Nostoc* sp. F4 filaments flexuously twisted. Vegetative cells were

the HIP TG primer (Fig. 5C). 29 amplified products with the size 300–1500 bp were obtained using HIP GC primer (Fig. 5C). Salt-tolerant *Nostoc* sp. F4 and F3 cultures produced the maximal number of amplified fragments with these primers. Finally, 33 amplified products with the size 300–1500 bp were obtained using STRR primer (Fig. 5D). Salt-tolerant and limestone cultures produced the maximal number of amplified fragments with these primers (Table 1).

Stress tolerant *Nostoc* sp. F4 and *Nostoc* sp. F3 cultures represented unique banding patterns (Fig. 3). A total of 13 amplified products with the size 400–3000 bp were obtained by using ERIC1A primer (Figs. 6A

Table 1: **Number of amplified products of studied strains by Rep-PCR generated fingerprint profile**

| Techniques | *Calothrix* sp. R42 | *Calothrix* sp. R11 | *Nostoc* sp. FA3 | *Nostoc* sp. FA5 | *Nostoc* sp. FA1 | *Nostoc* sp. F3 | *Nostoc* sp. F4 | Total Amplicons |
|---|---|---|---|---|---|---|---|---|
| ERIC1A | 1 | 1 | 2 | 1 | 3 | 1 | 3 | 12 |
| ERIC1B | 5 | 5 | 6 | 5 | 3 | 4 | 10 | 38 |
| HIP CA | 4 | 7 | 4 | 4 | 5 | 6 | 6 | 36 |
| HIP AT | 7 | 8 | 4 | 4 | 5 | 3 | 4 | 35 |
| HIP TG | 4 | 4 | 3 | 3 | 3 | 6 | 6 | 29 |
| HIP GC | 4 | 3 | 3 | 4 | 3 | 6 | 6 | 27 |
| STRR | 2 | 1 | 3 | 3 | 3 | 6 | 6 | 24 |

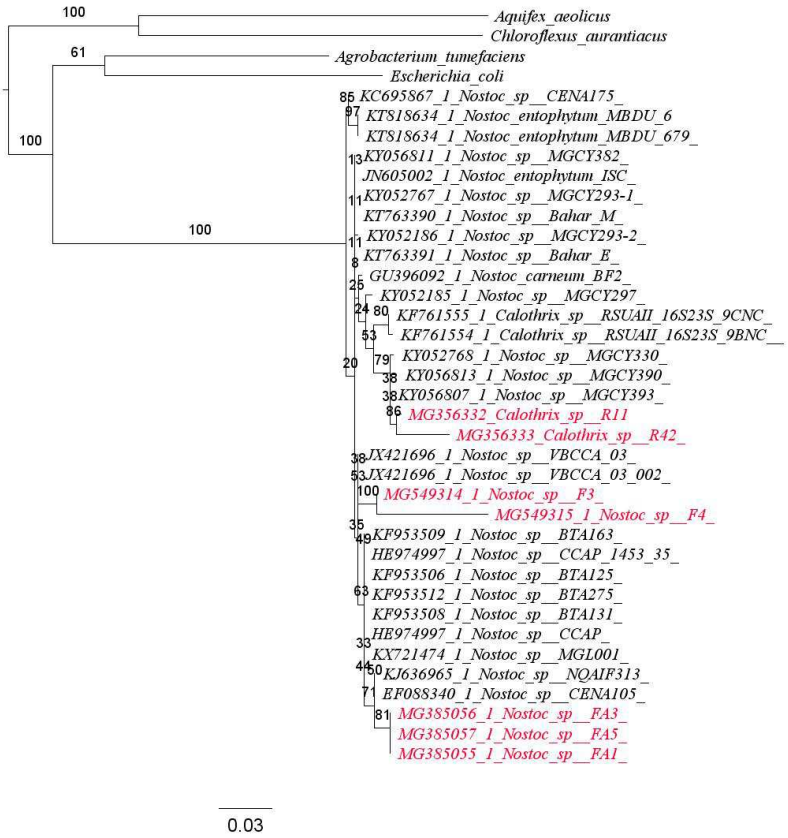


Figure 4: **Maximum-likelihood tree (IQ-Tree) based on the partial 16S rRNA gene sequenced in this study or taken from the GenBank. The studied strains are shown with red color. The scale bar represents 0.03 base substitutions per 1000 nucleotide position. Bootstrap percentages calculated from 1000 resembling are indicated at nodes.**

and 6B), whereas 43 amplified products (200–3000 bp in size) were obtained with the ERIC1B primer (Figs. 6A and 6B). ERIC1A led to the lowest number and the broadest range of bands. The largest number of amplified fragments was obtained for *Nostoc* sp. F4 with this primer. HIP primers used in the study were decamers with a common consensus sequence (5'GCGATCGC3') followed by a two-base tail of either AT, TG, GC or TC. A total of 30 amplified products with the size 200–1500 bp were obtained by using HIP CA primer (Figs. 6A and 6B). A total of 33 amplified products with the size 200–1500 bp were obtained using HIP AT primer (Figs. 6A and 6B).

Clustering of the PCR profiles revealed the presence of three major groups (Supp. Figs. 1-4). The two salt-tolerant cultures (*Nostoc* spp. F4 and F3) formed a first cluster for the ERIC1A and ERIC1B primers, sharing a 100% similarity (Supp. Fig. 1), while they exhibited the lowest similarity (50%) with the HIP-AT primer. The two nontolerant cultures (*Calothrix* spp. R11 and R42) formed the third cluster for the all primers (except HIP-AT) and were found to be distinct from the stress-tolerant isolates, sharing a 100% similarity with ERIC1A and HIP-CA primers but the lowest similarity (50%) with the STRR primer.

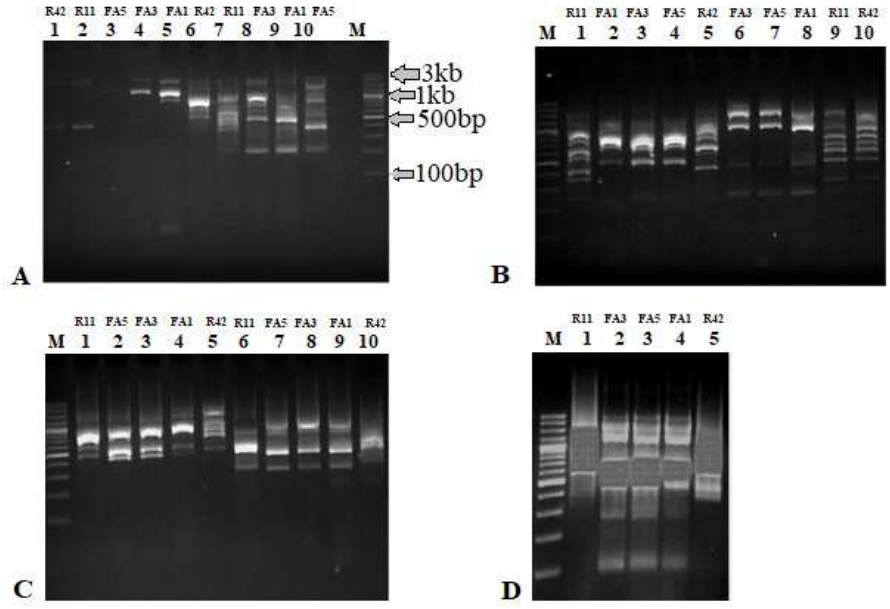


Figure 5. **PCR amplification pattern of the cyanobacterial cultures with the primer ERIC1A Lane (1-5) and ERIC1B Lane (6-10) (A). HIP-CA (1-5) and HIP-AT (6-10) (B). HIP-TG (1-5) and HIP-GC (6-10) (C). STRR1a (D). Lane M Molecular weight marker (100 bp Plus); Complete names are *Nostoc* sp. FA5; *Nostoc* sp. FA1; *Nostoc* sp. FA3; *Calothrix* sp. R11; *Calothrix* sp. R42; *Nostoc* sp. FA1; *Nostoc* sp. FA3; *Nostoc* sp. FA5; *Calothrix* sp. R11; *Calothrix* sp. R42.**

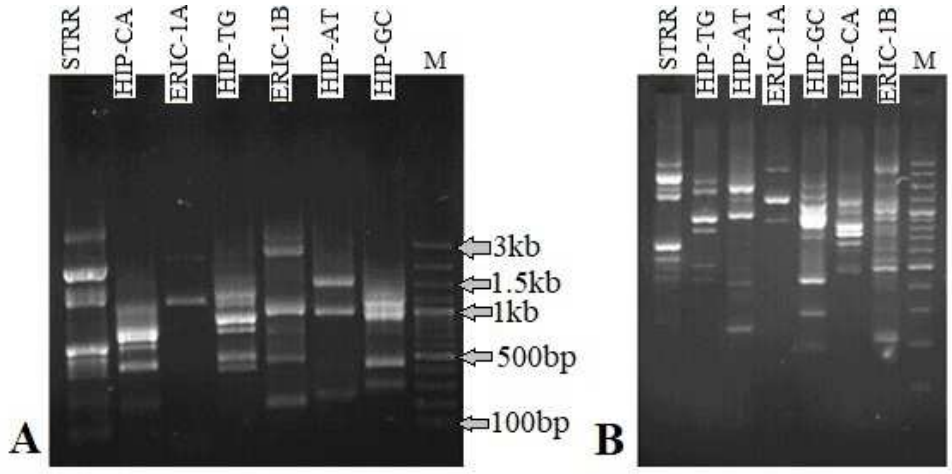


Figure 6. **PCR amplification pattern of the *Nostoc* sp.F3 (A) and *Nostoc* sp.F4 (B) with the primer HIP-CA and HIP-AT, HIP-GC, HIP-TG, STRR, ERIC1A, ERIC1B and STRR1. Lane M: Molecular weight marker (100 bp Plus).**

Limestone cultures were found to share a 100% similarity with HIP-GC and STRR primers; however, they had little similarity between themselves using other primers. For instance, *Nostoc* sp. FA3 and FA5 were found to share a 100% similarity, while *Nostoc* sp. FA1 shared a 60% similarity with both strains. Interestingly, the stress-tolerant *Nostoc* strains (saltwater and limestone cultures) did not share any clustering (Supp. Figs. 2-4).

**Discussion**

Generally, Cyanobacteria taxonomy has depended basically upon morphological and developmental characteristics ascertained under light microscopy (Komárek, 2010). Significant skill is required to distinguish species since both morphological and developmental characteristics shift alongside growth conditions. Such limitations forced by conventional techniques might be defeated by the use of molecular tools. Hence, utilization of phenotypic characters in a blend with molecular markers prompts a superior comprehension of molecular affinities and systematics within Cyanobacteria, and the composition of natural cyanobacterial communities (Nowruzi and Blanco, 2019).

Repetitive sequences comprise a crucial part of the prokaryotic genome (Elhai, 2015). Regardless of their obscure function and our lack of knowledge on how they are maintained and scattered, the widespread dispersion and high conservation ratios of these sequences make them methodologically imperative for DNA fingerprinting, permitting their utilization as an option for taxonomic identification and in diversity studies (Elhai, 2015; Valerio et al., 2009). Today, genomic fingerprints with ERIC, STRR1a, HIP-GC, HIP-CA, HIP-AT, and HIP-TG primers have been generally used for intraspecific comparisons in Cyanobacteria (Akoijam and Singh., 2011; Howard-Azzeh et al., 2014). Analysis of STRRs has been depicted as an amazing asset for taxonomic studies. Additionally, the particularity of these sequences has made STRRs helpful even for non-axenic cyanobacterial cultures (Liaimer et al., 2016). Also, STRR1A primer had the capacity to evaluate similarity within Cyanobacteria or in some *Gunnera* strains collected from various geographical areas. STRR sequences have been also used to segregate the *Nostoc* symbiont of *Blasia pusilla* L. (Liaimer et al., 2016). A modified STRR primer (STRR mod), based on the consensus sequence, was synthesized by Thajuddin et al., (2010) who reported the utility of this primer in differentiating symbiotic cyanobacteria from cycad roots. Valério *et al.* (2009) applied STRR along with other molecular targets such as 16S rRNA, *rpo*C and ERIC sequences for identification, typification, and traceability of freshwater cyanobacteria. Similarly, Thajuddin et al., 2010 used STRR along with other molecular and morphological attributes in order to assess the genetic diversity among symbiotic cyanobacterial genera isolated from cycads. Another repetitive eight-base sequence (5'-GCGATCGC-3'); alluded to as the HIP1, which is particularly however to a great extent overrepresented in cyanobacterial genomes, was first recognized at the borders of a gene deletion in a Cd-tolerant strain of *Synecococcus* PCC 6301 (Muralitharan, and Thajuddin., 2011). While there are no similar sequences to HIP1 in different organisms, the explanations behind this overrepresentation are obscure, although it is guessed that it might be a recombination hotspot. The HIP1 sequence can also be used to fingerprint organisms (Delaye and Moya, 2011). The HIP primers used in this investigation were constructed on the basis of this frequently occurring octa palindromic sequence, which permits DNA amplification between adjacent repeated HIP sequences. The utility of HIP sequences in differentiating members of *Anabaena* has been reported earlier (Delaye and Moya, 2011). In a further modification of this technique, Shalini et al. (2007) utilized a blend of HIP and random primer sets to create fingerprint patterns for *Calothrix*. They suggested that combining HIP CA and HIP TG primers generates a higher level of polymorphism than a single random primer.

Moreover, the techniques based on ERIC have been also used for identification and discrimination purposes in some Cyanobacteria (Ezhilarasi and An., 2010). The biochemical plasticity and diversity of Cyanobacteria have enabled them to occupy almost every conceivable habitat on the earth (Nowruzi et al., 2018; Nowruzi and Blanco, 2019). The present study embodies identification tools for filamentous cyanobacterial strains collected from different climatic/geographic regions and habitats, by comparing such isolates on the basis of their multiple DNA fingerprint profile. The strains did not actually reveal remarkable morphological differences so that their identification based on classical criteria was problematic. However, PCR fingerprinting based on the presence of STRR, ERIC and HIP sequences allowed the segregation of such closely related strains. These primers generated unique fingerprint patterns for individual *Nostoc* and *Calothrix* cultures with contrasting stress tolerance profiles and geographical origins. The dendrograms constructed showed overall significant similarities using distribution patterns (Supp. Figs. 1-4). This was evidenced in our results showing the clustering of isolates into three groups according to the ERIC, HIP and STRR

primers. However, *Nostoc* sp. F3 and F4 revealed different clustering patterns with all these primers. The limited number of PCR products obtained with ERIC A primer in contrast to ERIC B might be reflecting the position and orientation of the individual ERIC sequences in the genomes. The difference in banding patterns of limestone isolates compared to the other cultures was due to an absence of some PCR product. It has been postulated that these repetitive sequences might be the target of specific DNA-binding proteins responsible for chromosome condensation, but they might be also involved in the control of chromosome distribution and replication during heterocyst differentiation. Ever since their discovery, repetitive sequences have been widely used to resolve cyanobacterial isolates. Thereafter, repetitive sequences have been widely considered as one of the most accepted tools for assessing microbial diversity, particularly at a very high-resolution level in case of closely related cogeneric microbes (Prasanna et al., 2013). Clustering of cyanobacterial cultures with similar stress-tolerance properties using the rep-PCR fingerprint pattern shows that their genetic similarity could be mirroring their geographical origins. While the salt-tolerant cyanobacterial cultures used in this study were isolated from a saltwater reservoir near paddy fields in Golestan province, limestone-dwelling cultures were isolated from several locations at the North-west Mountains which continuously faced intense light, and nontolerant strains were isolated from agricultural areas of Kermanshah province. Earlier attempts of using rep-PCR profiles to cluster geographically distinct cyanobacterial isolates have relied on a limited number of primers, but we have successfully used a combination of STRR, ERIC and HIP rep-PCR profiles that considered many allelic positions for dendrogram construction. This has greatly expedited the grouping of the cultures into meaningful clusters. The correlation between stress tolerance and the utility of rep PCR amplification in Cyanobacteria is a future research target. From our results, it is evident that the repetitive sequences in the cyanobacterial genome facilitate the differentiation of strains inhabiting different environments. The dendrograms obtained (Supp. Figs. 1-4) reflect the phylogenetic relationships based on the 16S rRNA gene, thus supporting the use of tandem repeats as species markers for diversity assessment. The studied primer fingerprints discriminated successfully the seven analyzed strains collected from different habitats. Moreover, congruency was observed in the phylogeny reconstructed using repetitive primers and the 16S rRNA gene. A noticeable finding highlighted in this study is the use the rep-PCR profiles for clustering of cogeneric strains that differ in stress tolerance. To our knowledge, in Iran, this is the first report on the presence of repetitive sequences in cyanobacterial genomes, and on rep-PCR profiles generating closely genetic clusters based on identical stress-tolerant properties.

## Conclusion

From the present investigation, it is concluded that the repetitive sequences found in the genomes of Cyanobacteria are very useful in exploring genomic relationships among different strains, which supports their use for Cyanobacterial discrimination and identification in different climatic/geographic Regions and habitats.

## Author Contributions

BN designed the study, analyzed the data and drafted the manuscript; RS, HF and SB did meticulous revision of the manuscript.

## Acknowledgment

No acknowledgment.

## Competing financial interests

Authors have declared that no competing interest exists.

## References

Abony, M., Banik A., Shishir M. A., Akter N. J., Uddin M. E., Datta S. (2018). Physico-chemical Characterization of Indigenous *Streptomyces* and Influence of pH on Antimicrobial Activity. *Microbial Bioactives, 1*(2), 059-067. https://doi.org/10.25163/microbbioacts.12009A3010021118

Akoijam, C., & Singh, A. K. (2011). Molecular typing and distribution of filamentous heterocystous cyanobacteria isolated from two distinctly located regions in North-Eastern India. *World Journal of Microbiology and Biotechnology*, *27*(9), 2187-2194. https://doi.org/10.1007/s11274-011-0684-8

Cavalier-Smith, T. (2002). The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. *International Journal of Systematic and Evolutionary Microbiology, 52*(2), 297-354. https://doi.org/10.1099/00207713-52-2-297 PMid:11931142

Delaye, L., & Moya, A. (2011). Abundance and distribution of the highly iterated palindrome 1 (HIP1) among prokaryotes. *Mobile Genetic Elements*, *1*(3), 159-168. https://doi.org/10.4161/mge.1.3.18300
PMid:22312590 PMCid:PMC3271550

Desikachary, T.V. *Cyanophyta*, (Indian Council of Agricultural Research New Delhi, 1959).

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792-1797. https://doi.org/10.1093/nar/gkh340
PMid:15034147 PMCid:PMC390337

Elhai, J. (2015). Highly Iterated Palindromic sequences (HIPs) and their relationship to DNA methyltransferases. *Life*, *5*(1), 921-948. https://doi.org/10.3390/life5010921 PMid:25789551 PMCid:PMC4390886

Ezhilarasi, A., & An, N. (2010). Fingerprinting of repetitive DNA sequences in the genus Anabaena using PCR-based techniques. *African Journal of Microbiology Research*, *4*(8), 590-597.

Gould, S. B., Waller, R. F., & McFadden, G. I. (2008). Plastid evolution. *Annual Review of Plant Biology*, *59*, 491-517. https://doi.org/10.1146/annurev.arplant.59.032607.092915 PMid:18315522

Gugger, M. F., & Hoffmann, L. (2004). Polyphyly of true branching cyanobacteria (Stigonematales). *International Journal of Systematic and Evolutionary Microbiology*, *54*(2), 349-357. https://doi.org/10.1099/ijs.0.02744-0 PMid:15023942

Helm, R. F., Huang, Z., Edwards, D., Leeson, H., Peery, W., & Potts, M. (2000). Structural characterization of the released polysaccharide of desiccation-tolerant *Nostoc commune* DRH-1. *Journal of Bacteriology*, *182*(4), 974-982. https://doi.org/10.1128/JB.182.4.974-982.2000 PMid:10648523 PMCid:PMC94373

Hill, D. R., Peat, A., & Potts, M. (1994). Biochemistry and structure of the glycan secreted by desiccation-tolerant *Nostoc commune* (Cyanobacteria). *Protoplasma*, *182*(3-4), 126-148. https://doi.org/10.1007/BF01403474

Howard-Azzeh, M., Shamseer, L., Schellhorn, H. E., & Gupta, R. S. (2014). Phylogenetic analysis and molecular signatures defining a monophyletic clade of heterocystous cyanobacteria and identifying its closest relatives. *Photosynthesis Research*, *122*(2), 171-185. https://doi.org/10.1007/s11120-014-0020-x PMid:24917519

Iteman, I., Rippka, R., de Marsac, N. T., & Herdman, M. (2002). rDNA analyses of planktonic heterocystous cyanobacteria, including members of the genera *Anabaenopsis* and *Cyanospira*. *Microbiology*, *148*(2), 481-496. https://doi.org/10.1099/00221287-148-2-481 PMid:11832512

Kaushik, P., Chauhan, A., Chauhan, G., & Goyal, P. (2009). Antibacterial potential and UV-HPLC analysis of laboratory-grown culture of *Anabaena variabilis*. *International Journal of Food Safety*, *11*, 11-18.

Komárek J and Johansen JR. (2014) Filamentous cyanobacteria. In *Freshwater Algae of North America (Second Edition)*. Elsevier, pp. 135-235.

Komárek, J. (2010). Recent changes (2008) in cyanobacteria taxonomy based on a combination of molecular background with phenotype and ecological consequences (genus and species concept). *Hydrobiologia*, *639*(1), 245-259. https://doi.org/10.1007/s10750-009-0031-3

Komárek J. (2013): Cyanoprokaryota. 3. Heterocystous genera. – In: Büdel B., Gärtner G., Krienitz L. & Schagerl M. (eds), Süswasserflora von Mitteleuropa/Freshwater flora of Central Europe, p. 1130, Springer Spektrum Berlin, Heidelberg.

Laguerre, G., Mavingui, P., Allard, M. R., Charnay, M. P., Louvrier, P., Mazurier, S. I., ... & Amarger, N. (1996). Typing of rhizobia by PCR DNA fingerprinting and PCR-restriction fragment length polymorphism analysis of chromosomal and symbiotic gene regions: application to *Rhizobium leguminosarum* and its different biovars. *Applied and Environmental Microbiology*, *62*(6), 2029-2036. PMid:8787401 PMCid:PMC167981

Liaimer, A., Jensen, J. B., & Dittmann, E. (2016). A genetic and chemical perspective on symbiotic recruitment of cyanobacteria of the genus *Nostoc* into the host plant *Blasia pusilla* L. *Frontiers in Microbiology*, *7*, 1693. https://doi.org/10.3389/fmicb.2016.01693

Lyra, C., Laamanen, M., Lehtimäki, J. M., Surakka, A., & Sivonen, K. (2005). Benthic cyanobacteria of the genus *Nodularia* are non-toxic, without gas vacuoles, able to glide and genetically more diverse than planktonic *Nodularia*. *International Journal of Systematic and Evolutionary Microbiology*, *55*(2), 555-568. https://doi.org/10.1099/ijs.0.63288-0 PMid:15774625

Muralitharan, G., & Thajuddin, N. (2011). Rapid differentiation of phenotypically and genotypically similar *Synechococcus elongatus* strains by PCR fingerprinting. *Biologia*, *66*(2), 238-243. https://doi.org/10.2478/s11756-011-0003-8

Neilan, B. A., Saker, M. L., Fastner, J., Törökné, A., & Burns, B. P. (2003). Phylogeography of the invasive cyanobacterium *Cylindrospermopsis raciborskii*. *Molecular Ecology*, *12*(1), 133-140. https://doi.org/10.1046/j.1365-294X.2003.01709.x PMid:12492883

Nguyen, L. T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2014). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, *32*(1), 268-274. https://doi.org/10.1093/molbev/msu300 PMid:25371430 PMCid:PMC4271533

Nilsson, M., Bergman, B., & Rasmussen, U. (2000). Cyanobacterial diversity in geographically related and distant host plants of the genus *Gunnera*. *Archives of Microbiology*, *173*(2), 97-102. https://doi.org/10.1007/s002039900113 PMid:10795680

Nowruzi, B., Khavari-Nejad, R. A., Sivonen, K., Kazemi, B., Najafi, F., & Nejadsattari, T. (2012). Identification and toxigenic potential of a *Nostoc* sp. *Algae*, *27*(4), 303-313. https://doi.org/10.4490/algae.2012.27.4.303

Nowruzi, B., & Blanco, S. (2019). *In silico* identification and evolutionary analysis of candidate genes involved in the biosynthesis methylproline genes in cyanobacteria strains of Iran. *Phytochemistry Letters*, *29*, 199-211. https://doi.org/10.1016/j.phytol.2018.12.011

Nowruzi, B., Blanco, S., & Nejadsattari, T. (2018). Chemical and Molecular Evidences for the Poisoning of a Duck by Anatoxin-a, Nodularin and Cryptophycin at the Coast of Lake Shoormast (Mazandaran Province, Iran). *International Journal on Algae*, *20*(4). https://doi.org/10.1615/InterJAlgae.v20.i4.30

Orcutt, K. M., Rasmussen, U., Webb, E. A., Waterbury, J. B., Gundersen, K., & Bergman, B. (2002). Characterization of *Trichodesmium* spp. by genetic techniques. *Applied and Environmental Microbiology*, *68*(5), 2236-2245. https://doi.org/10.1128/AEM.68.5.2236-2245.2002 PMid:11976093 PMCid:PMC127538

Prabina, B. J., Kumar, K., & Kannaiyan, S. (2005). DNA amplification fingerprinting as a tool for checking genetic purity of strains in the cyanobacterial inoculum. *World Journal of Microbiology and Biotechnology*, *21*(5), 629-634. https://doi.org/10.1007/s11274-004-3566-5

Prasanna, R., Kumar, R., Sood, A., Prasanna, B. M., & Singh, P. K. (2006). Morphological, physiochemical and molecular characterization of *Anabaena* strains. *Microbiological Research*, *161*(3), 187-202. https://doi.org/10.1016/j.micres.2005.08.001 PMid:16765835

Prasanna, R., Babu, S., Rana, A., Kabi, S. R., Chaudhary, V., Gupta, V., ... & Pal, R. K. (2013). Evaluating the establishment and agronomic proficiency of cyanobacterial consortia as organic options in wheat-rice cropping sequence. *Experimental Agriculture*, *49*(3), 416-434. https://doi.org/10.1017/S001447971200107X

Rasmussen, U., & Svenning, M. M. (1998). Fingerprinting of cyanobacteria based on PCR with primers derived from short and long tandemly repeated repetitive sequences. *Applied and Environmental Microbiology*, *64*(1), 265-272. PMid:16349487 PMCid:PMC124704

Rippka, R., Deruelles, J., Waterbury, J. B., Herdman, M., & Stanier, R. Y. (1979). Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *Microbiology*, *111*(1), 1-61. https://doi.org/10.1099/00221287-111-1-1

Rodriguez-Barradas, M. C., Hamill, R. J., Houston, E. D., Georghiou, P. R., Clarridge, J. E., Regnery, R. L., & Koehler, J. E. (1995). Genomic fingerprinting of *Bartonella* species by repetitive element PCR for distinguishing species and isolates. *Journal of Clinical Microbiology*, *33*(5), 1089-1093. PMid:7615711 PMCid:PMC228110

Sánchez-Baracaldo, P., Ridgwell, A., & Raven, J. A. (2014). A neoproterozoic transition in the marine nitrogen cycle. *Current Biology*, *24*(6), 652-657. https://doi.org/10.1016/j.cub.2014.01.041 PMid:24583016

Schirrmeister, B. E., Gugger, M., & Donoghue, P. C. (2015). Cyanobacteria and the Great Oxidation Event: evidence from genes and fossils. *Palaeontology*, *58*(5), 769-785. https://doi.org/10.1111/pala.12178; https://doi.org/10.1111/pala.12193 PMid:26924853 PMCid:PMC4755140

Selvakumar, G., & Gopalaswamy, G. (2008). PCR based fingerprinting of *Westiellopsis* cultures with short tandemly repeated repetitive (STRR) and highly iterated palindrome (HIP) sequences. *Biologia*, *63*(3), 283-288. https://doi.org/10.2478/s11756-008-0065-4

Shih, P. M., Wu, D., Latifi, A., Axen, S. D., Fewer, D. P., Talla, E., Herdman, M. (2013). Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proceedings of the National Academy of Sciences*, *110*(3), 1053-1058. https://doi.org/10.1073/pnas.1217107110 PMid:23277585 PMCid:PMC3549136

Shirkey, B., McMaster, N. J., Smith, S. C., Wright, D. J., Rodriguez, H., Jaruga, P., ... & Potts, M. (2003). Genomic DNA of *Nostoc commune* (Cyanobacteria) becomes covalently modified during long-term (decades) desiccation but is protected from oxidative damage and degradation. *Nucleic Acids Research*, *31*(12), 2995-3005. https://doi.org/10.1093/nar/gkg404 PMid:12799425 PMCid:PMC162238

Shishir, M. A., Pervin, S., Sultana, M., Khan, S. N., & Hoq, M. M. (2015). Genetic Diversity of Indigenous *Bacillus thuringiensis* Strains by RAPD-PCR to Combat Pest Resistance. *Bt Research*, *6*(8), 1–16. https://doi.org/10.5376/bt.2015.06.0008

Smith, J. K., Parry, J. D., Day, J. G., & Smith, R. J. (1998). A PCR technique based on the HipI interspersed repetitive sequence distinguishes cyanobacterial species and strains. *Microbiology*, *144*(10), 2791-2801. https://doi.org/10.1099/00221287-144-10-2791 PMid:9802020

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. *Applied and Environmental Microbiology*, *69*(9), 5157-5169. https://doi.org/10.1128/AEM.69.9.5157-5169.2003 PMid:12957897 PMCid:PMC194958

Teaumroong, N., Innok, S., Chunleuchanon, S., & Boonkerd, N. (2002). Diversity of nitrogen-fixing cyanobacteria under various ecosystems of Thailand: I. Morphology, physiology and genetic diversity. *World Journal of Microbiology and Biotechnology*, *18*(7), 673-682. https://doi.org/10.1023/A:1016812116538

Thajuddin, N., Muralitharan, G., Sundaramoorthy, M., Ramamoorthy, R., Ramachandran, S., Akbarsha, M. A., & Gunasekaran, M. (2010). Morphological and genetic diversity of symbiotic cyanobacteria from cycads. *Journal of Basic Microbiology*, *50*(3), 254-265. https://doi.org/10.1002/jobm.200900343 PMid:20473963

Uyeda, J. C., Harmon, L. J., & Blank, C. E. (2016). A comprehensive study of cyanobacterial morphological and ecological evolutionary dynamics through deep geologic time. *PloS One*, *11*(9), e0162539. https://doi.org/10.1371/journal.pone.0162539 PMid:27649395 PMCid:PMC5029880

Valerio, E., Chambel, L., Paulino, S., Faria, N., Pereira, P., & Tenreiro, R. (2009). Molecular identification, typing and traceability of cyanobacteria from freshwater reservoirs. *Microbiology*, *155*(2), 642-656. https://doi.org/10.1099/mic.0.022848-0 PMid:19202113

Walter, J. M., Coutinho, F. H., Dutilh, B. E., Swings, J., Thompson, F. L., & Thompson, C. C. (2017). Ecogenomics and taxonomy of Cyanobacteria phylum. *Frontiers in Microbiology*, *8*, 2132. https://doi.org/10.3389/fmicb.2017.02132

Wilson, K. M., Schembri, M. A., Baker, P. D., & Saint, C. P. (2000). Molecular characterization of the toxic cyanobacterium Cylindrospermopsis raciborskii and design of a species-specific PCR. *Applied and Environmental Microbiology*, *66*(1), 332-338. https://doi.org/10.1128/AEM.66.1.332-338.2000 PMid:10618244 PMCid:PMC91826

Zheng, W. W., Nilsson, M., Bergman, B., & Rasmussen, U. (1999). Genetic diversity and classification of cyanobacteria in different Azolla species by the use of PCR fingerprinting. *Theoretical and Applied Genetics*, *99*(7-8), 1187-1193. https://doi.org/10.1007/s001220051323

Zheng, W., Song, T., Bao, X., Bergman, B., & Rasmussen, U. (2002). High cyanobacterial diversity in coralloid roots of cycads revealed by PCR fingerprinting. *FEMS Microbiology Ecology*, *40*(3), 215-222. https://doi.org/10.1111/j.1574-6941.2002.tb00954.x PMid:19709229