Indian Journal of Biochemistry & Biophysics
Vol. 60, May 2023, pp. 423-429
DOI: 10.56042/ijbb.v60i5.1319

Molecular characterization of phycobiliproteins producing Cyanobacteria by using nanopore sequencer isolated from dairy effluent

Jayesh H Kabariya[1] & Vimal M Ramani[2]*

[1 & 2]Department of Dairy Microbiology, College of Dairy Science, Kamdhenu University, Amreli-365 601, Gujarat, India

Received 09 May 2023; *revised 03 July 2023*

Cyanobacterial taxonomy has been established based on morphological features, such as the shape and dimensions of the cells, and the presence of structurally differentiated cells. Since, the last decade's molecular techniques were used for the identification of cyanobacteria based on 16S rRNA genes. Nanopore sequencing technology is the most recent technique so, in the present study nanopore sequencing was explored for the identification of cyanobacterial species isolated from dairy effluent and screened based on effluent nutrients removal efficiency and it also phycobiliproteins producer. The steps were followed by DNA extraction, library preparation, barcoding, DNA amplification, 16S rRNA sequencing by using MinION Nanopore sequencer and data analysis was evaluated by Epi2Me software for cyanobacterial species identification and preparation of the phylogeny. The two cyanobacterial species (*Oscillatoria species*-1 & *Oscillatoria species*-2) were molecularly identified as *Oscillatoria acuminata* and *Oscillatoria sancta* by 16S rRNA sequencing using Nanopore sequencing Technology and the DNA sequence was submitted to NCBI with stain name VMRJHK2020 and VMRJHK012020 and the Gene bank accession are MW364270 and MW364271, respectively.

Keywords: 16S rRNA sequencing, Dairy effluent treatment, *Oscillatoria* species

Cyanobacteria, also known as microalgae, comprise a unique group of organisms with world-wide distribution. These were considered as algae because of their microscopic morphology, pigmentation and oxygen-evolving photosynthesis. Ecologists started giving them importance as primary producers and realized that without them no animal population would exist.

The role of cyanobacteria in the removal of various kinds of organic, inorganic and other related substances has been studied by several workers during the last several years. Cyanobacteria can utilized the organic as well as inorganic matters as a food and increase their biomass. Cyanobacteria can used in so many different fields, such as wastewater treatment, biofuel production, nutraceurical production, human nutrition (single cell protein (*spirulina*) bio-fertilizer and other area. Cyanobacterial species such as *Spirulina platensis* has potential antioxidative protein-pigment complex for food grade phycocyanin C. Scientist has extracted, purify and also checked antioxidative activity and biocompatibility[1]. Nutrient removal, particularly nitrogen and phosphorus from dairy effluent is a growing regulatory need and the use of cyanobacteria cultivation could create a unique combination between dairy effluent treatment and extraction of high-value compounds from produced cyanobacterial biomass. Scientist were also used the nanoparticles for treatment of effluent, PVA-Ag nanoparticles were reported the significant killing of microorganisms in a microbial assay using *E. coli* agar medium[2].

Cyanobacterial taxonomy has been established based on morphological features, such as the shape and dimensions of the cells, presence of structurally differentiated cells, and whether the cells grow as solitary cells or in colonies[3].

Molecular methods have become an indispensable tool for the characterization of cyano-prokaryotes and the assessment of evolutionary relations among them in recent decades. The direct sequencing of various genes is the most common method used in the taxonomy of cyanobacteria. However, RFLP (Restriction Fragment Length Polymorphism) is additionally widely applied, especially for a more detailed examination of the genetic variability of closely related taxa or to infer the extent of cyanobacterial diversity in nature[4]. Also, random amplified polymorphic DNA (RAPD) analysis is usually utilized to discriminate between genotypes of close

*Correspondence:
E-mail: vimalmramani@gmail.com

relatives. Novel primer designing and PCR-AFLP approach were also reported for the expeditious detection of coliforms in potable waters[5]. Gene specific PCR strategy has also used for cloning and structural elucidation as well as genetic transformation[6]. Researches has developed M-PCR mix with a unique combination of specificity and sensitivity coupled with great flexibility has led to a true revolution in molecular diagnostics[7].

Phylogenetic analysis of cyanobacteria based on 16S rRNA genes showed that they are a diverse, monophyletic phylum of organisms within the bacterial radiation. Potential uses of cyanobacteria in biotechnology researchers are focused on cyanobacterial ecology, morphology, physiology, and 16S rRNA-based phylogeny in the last decades[8].

Molecular approaches have been divided into two classes: PCR independent and PCR based approaches. Molecular assessment of cyanobacterial biodiversity has been studied by using markers like 16S rRNA, phycocyanin locus, nif gene, rpo gene, ITS region, etc. An overall outline of biodiversity evaluation, molecular methods, and markers utilized for biodiversity analysis combinatorial methodology with various sub-molecular markers. It is likely to enhance the degree of resolution and supply as possible the broadest picture and thorough information about biodiversity documentation[9].

Various methods are available for DNA sequencings such as Maxam and gilbert, Sanger dideoxy sequencing, pyrosequencing, next-generation sequencing, and nanopore sequencing method[10].

Nanopore sequencing is extraordinary, versatile innovation that immediate, constant investigation of long fragments of DNA or RNA. It works by checking changes to an electrical flow as nucleic acids by protein nanopore. The resulting signal is decoded to supply the precise DNA or RNA sequence. The application of the nanopore sequencing method was useful for almost all forms of cyanobacteria species and strain discrimination[11]. The MinION is a pocket-sized device that – when paired with a laptop or mobile phone can sequence genetic information. It has already been harnessed during epidemics to track the genetic information of humans, animals and environmental sample[12].

Current approach for producing a bacterial genome assembly with the goal of perfection using a combination of Oxford Nanopore Technologies (ONT) long reads and Illumina short reads. Researches have chosed oxford nanopore technology for their availability and widespread adoption in microbial genomics[13-14]. Recently, scientist have also used the nanopore technology as a next-generation sequencing technology for modern agriculture[15].

In the present study nanopore sequencing technology was used for the identification of cyanobacterial species isolated from dairy effluent and screened based on nutrient removal efficiency from dairy effluent. The steps were followed by DNA extraction, library preparation, barcoding, DNA amplification, 16S rRNA sequencing by using MinION Nanopore sequence and in last step data analysis was evaluated by Epi2Me software for cyanobacterial species identification and preparation of the phylogeny.

Materials and Methods

DNA Extraction

Isolation of cyanobacteria from the dairy effluent sample was carried out using the modified method by using BG-11 (Blue Green-11, Hi-Media, India) medium[16]. After isolation, extraction of DNA from two cyanobacterial species screened on the basis of its dairy effluent treatment and its nutrient removal efficiency from dairy effluent *i.e. Oscillatoria species*-1 & *Oscillatoria species*-2 were carried out as described[17]. The cyanobacterial cultures were separated by using a centrifuge at 10000 rpm for 10 min. Culture was disrupted by using probe sonicator at 25 amplitute for 2-3 min with 4 second hold. The disrupted culture was transferred into a new tube and 400 μL lysis buffer was added followed by 4 μL RNase A. It was mixed properly and incubated for 10 min at 65°C. Then 130 μL buffer P3 was added and incubated for 5 min on ice. After that lysate was centrifuge for 5 min at 14000 rpm and transfer the lysate into a QIA shredder spin column then centrifuge again for 2 min at 14000 rpm.

The supernatant was transferred into a new tube without disturbing the pellet and wash buffer was added, mixed by pipetting. Transfer the 650 μL from the mixture into a DNeasy mini spin column and centrifuge it for 1 min at 8000 rpm then supernatant was discarded, this step was repeated for the remained sample**.** The DNA was eluted by adding 100 μL of elution buffer and incubated for 5 min at room temperature (25°C) followed by centrifuge at 8000 rpm for 1 min.

Quality check by gel electrophoresis & nanodrop spectrophotometer

5 µL of the extracted DNA sample was mixed with 2 µL of the gel loading dye. DNA sample was loaded into well then electrophoresis with 1% agarose gel was run at 50 volts for 30 min and the quality of DNA bands was checked.

The absorption of the extracted DNA sample was measured at A260 & A280nm wavelength by using NanoDrop spectrophotometer and the concentration of extracted DNA was also calculated.

DNA amplification and library preparation

16S Barcodes were thawed at room temperature and mixed by pipetting then barcodes were kept on ice until use. 10ng of extracted DNA was added into 0.2 mL PCR tubes containing 14 µL of nuclease-free water, 1 µL of 16S barcode, 25 µL of longAmp Taq 2X master mix, forward primer, reverse primer (frw1 primer (5'GACGGGTGAGTAATGCCTA-3') and rev1 (5'CACTGGTGTTCCTTCCTATA -3'), and dNTP’s then it was mixed gently by flicking. PCR protocol was set for 25 cycles with initial denaturation 1 min at 95°C, denaturation 20 secs at 95°C, annealing 30 secs at 55°C, extension 2 mins at 65°C, final extension 5 mins at 65°C and hold at 4°C.

Purification of Amplified DNA by Ampure XP Beads

The amplified DNA sample was transferred into a 1.5 mL DNA Eppendorf tube. AMPure XP beads (Beckman Coulter Genomics, CA, USA) were re-suspend by vortex then 30 µL of re-suspended AMPure XP beads were added into the tube, mixed it properly. The mixture was incubated on a hula mixer for 5 min at room temperature.

500 µL of 70% ethanol was prepared in nuclease-free water then beads were washed by 200 µL of freshly prepared 70% ethanol without disturbing the pellet. Removed all the 70% ethanol and the previous step was repeated then allowed it to dry for ~30 sec.

Pellet of DNA binded beads were re-suspend in 10 µL of 10 mM Tris-HCl pH 8.0 then incubate for 2 min at room temperature. Pelleted beads were placed on a magnetic stand until elute was clear and colourless. 10 µL of elute containing DNA was separated into a new Eppendorf tube.

Measurement of amplified DNA by Qubit Fluorometer

1000 µL of the working solution was prepared from qubit dsDNA BR reagent in the ratio of 1:200 with double distilled water. 190 µL of the above prepared working solution was taken and add 10 µL of DNA standard-1 & standard-2 solution. For unknown DNA samples, 198 µL of working solution and 2 µL of unknown DNA sample were added into new tubes. It was incubated for 5 min at room temperature. Reading of standard and unknown DNA samples were recorded by using qubit 4 fluorometer (ThermoFisher Scientific, Invitrogen) and measured the concentration of amplified DNA.

16S rRNA sequencing using Nanopore (Minion) Sequencer

The nanopore sequencing library was run on a MinION (MN29950) instrument with flowcell (FAK40529). The barcoded library was pooled in the ratio of total 50-100 fmoles in 10 µL of 10 mM Tris-HCl pH 8.0 with 50 mM NaCl. 1 µL of Rapid Sequencing Adapter (RAP) was added into the pooled library and it was mixed gently and incubated for 5 min at room temperature. The prepared library was kept on ice until the use.

Priming and Loading of DNA Library into Spoton Flowcell

Sequencing Buffer (SQB), Loading Beads (LB), Flush Tether (FLT), and one tube of Flush Buffer (FB) were thawed at room temperature and it was mixed by the vortex. For the preparation of flowcell priming mix 30 µL of FLT was added into FB tube, it was mixed by pipetting and kept on ice until use.

800 µL of priming mix was load into flowcell *via* priming port to remove earlier buffer of flowcell and incubated it for 5 min at room temperature, simultaneously, a total 75 µL of loading mix was prepared with 34 µL of SQB, followed by 25.5 µL of LB, 4.5 µL of nuclease-free water and 11 µL of barcoded DNA library. SpotON port of flowcell was opened before loading and then again 200 µL of the priming mix was added into the flowcell via the priming port. Then loading mix was mixed before use by pipetting and then load the 75 µL of loading mix dropwise into the flowcell *via* SpotON port.

Raw Data Collection of 16S rRNA Sequences by using Minknow Software

Loaded flowcell was first connected with MinKNOW software (versions 0.49.2.9 through 0.51.3.40 b201605171140) then flowcell was scan by software (mug scan) for available active pore into flowcell. After the mug scan, 623 active pores were available for sequencing. Then loading mix was sequenced with MinKNOW software settings appropriate for 2D and standard 48 h processing time for every run. Albacore 1.2.4 was used to convert FAST5 files into FASTQ files of each sequence. (Fig. 1) shows the live DNA sequencing process.

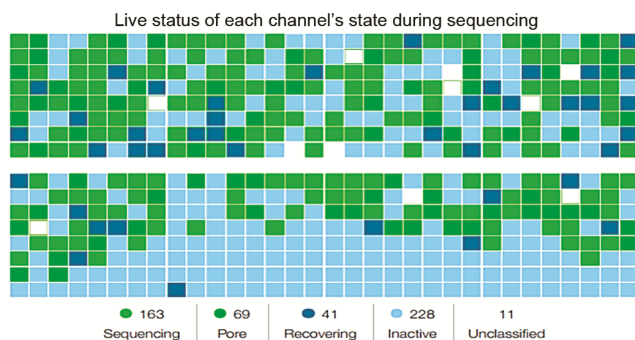

Fig. 1 — Live DNA sequencing process *via* MinKNOW software

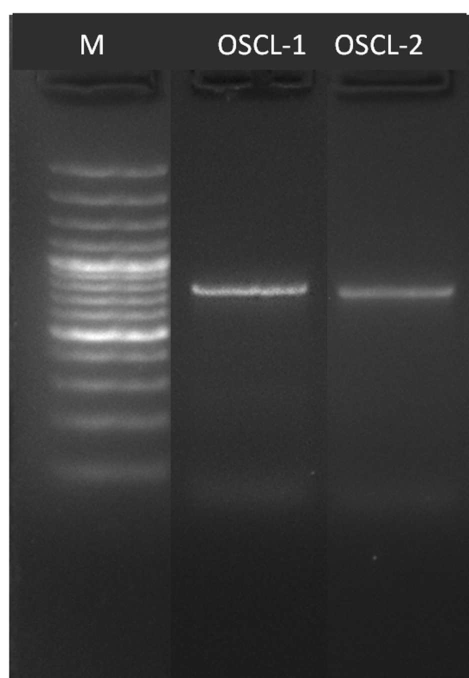

Fig. 2 — DNA bands of extracted DNA on 1% agarose gel

Data Processing by Metrichor Software

Raw files (FASTQ file) were uploaded into cloud-based bioinformatics software *i.e.* Metrichor agent developed by Nanopore Technology *via* 16S DNA sequencing portal. The bioinformatics analysis took time according to the amount of DNA present in the library. The bioinformatics analyzed data was downloaded from the cloud after the completion of analysis by using the same software.

Result and Discussion

DNA Extraction and Gel Electrophoresis

The concentration and purity of extracted DNA were evaluated based on optical density at A260/A280 ratio and agarose gel electrophoresis. 5.1 µg/µL and 8.1 µg/µL of DNA was extracted from *Oscillatoria species*-1 (OSCL-1) & *Oscillatoria species*-2 (OSCL-2), respectively. (Table 1) shows the concentration of extracted DNA were evaluated based on A260/A280.

Table 1 — Concentration of extracted DNA on NanoDrop spectrophotometer

Sl. No.	Sample id	NanoDrop readings (µg/µL)	NanoDrop OD $A_{260/280}$	NanoDrop OD $A_{260/230}$
1	OSCL-1	5.1	1.55	0.50
2	OSCL-2	8.1	1.85	0.72

Table 2 — Statistics of nanopore sequencer (MinION) data

Sample	#Reads	Total Bases	Data in Mb
OSCL1	77,251	10,00,12,524	~ 100
OSCL2	1,00,106	16,49,96,394	~ 164

RNA, proteins and other impurities were migrated at different rates compared to DNA, so, DNA was easily detected. Extracted DNA bands indicated that any impurities were not present in samples. (Fig. 2) shows the extracted DNA bands on an agarose gel.

Qubit Fluorometer Analysis after DNA Amplification

After purification of amplified DNA by AMPure XP beads, the concentration of PCR amplified DNA was evaluated on qubit 4 fluorometer with standard samples. The concentration of amplified DNA of OSCL-1 and OSCL-2 were found 1.07 ng/µL and 1.59 ng/µL, respectively, as indicated in qubit 4 fluorometer display.

16S rRNA Sequencing Analysis by Epi2Me Software

The nanopore sequencing generated 77,251 and 100,106 sequencing reads from 623 active pores for OSCL-1 and OSCL-2, respectively. Total 264 Mb data were generated and among them, 100 Mb and 164 Mb bases data of OSCL-1 and OSCL-2 were generated, respectively. (Table 2) shows the statistics of nanopore (MinION) generated data *i.e.* a number of reads, Total bases and data in Mb. The average quality score of DNA sequencing process was 8.70 with 100% workflow. The average sequence length of sequenced DNA was 1493 bases.

Sequencing of the full-length 16S rRNA gene was achieved in Nanopore long reads. The read length has narrow distribution and the average read length was estimated (N50) 1360bp, which was nearly the full length of the 16S rRNA gene (Fig. 3).

Identification of Cyanobacterial species through 16S rRNA Sequencing

The isolated and screened cyanobacterial species from dairy effluent *i.e.* OSCL-1 and OSCL-2 were

identified as *Oscillatoria acuminata* and *Oscillatoria sancta,* respectively by 16S rRNA sequencing using nanopore technology. Based on the obtained sequence and similarities matching by Metrichor software, the sequence was showing maximum similarity with the sequence of *Oscillatoria acuminate* and *Oscillatoria sancta.*

The 16S rRNA sequence of these identified cyanobacterial species were successfully submitted to NCBI. The Gene bank accession number of *Oscillatoria acuminata* species was MW364270 with strain name VMRJHK2020 and the accession number of *Oscillatoria sancta was* MW364271 with strain name VMRJHK012020.

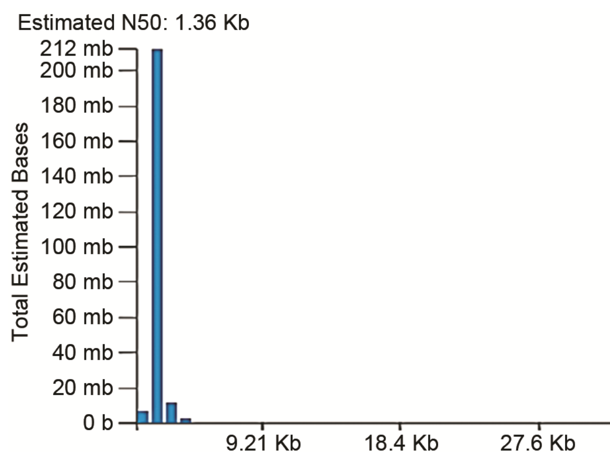

Fig. 3 — Read length histogram of sequencing data

The results of the 16S rRNA sequence were confirmed the species-level identification and the isolated cyanobacteria species was *Oscillatoria acuminata* and *Oscillatoria sancta.* (Fig. 4) shows the cladogram of both identified cyanobacterial species.

Cyanobacterial species was reported that the *Oscillatoria acuminata* and *Scenedesmus armatus* have potential for use in effluent treatment[19]. The latest technology *i.e.* nanopore sequencing technology for the identification of mouse gut microbiome using full-length 16S rRNA amplicon sequencing[20].

Scientist was reported the nanopore technology can be used at the extreme environmental condition[21]. They used these techniques for a field trial of metagenome profile to characterize the microbiota of the high Arctic glacier. The profile of the microbial communities produced by this technology was coherent with those from conventional amplicon and shotgun metagenome sequencing of glacier environments and prompt

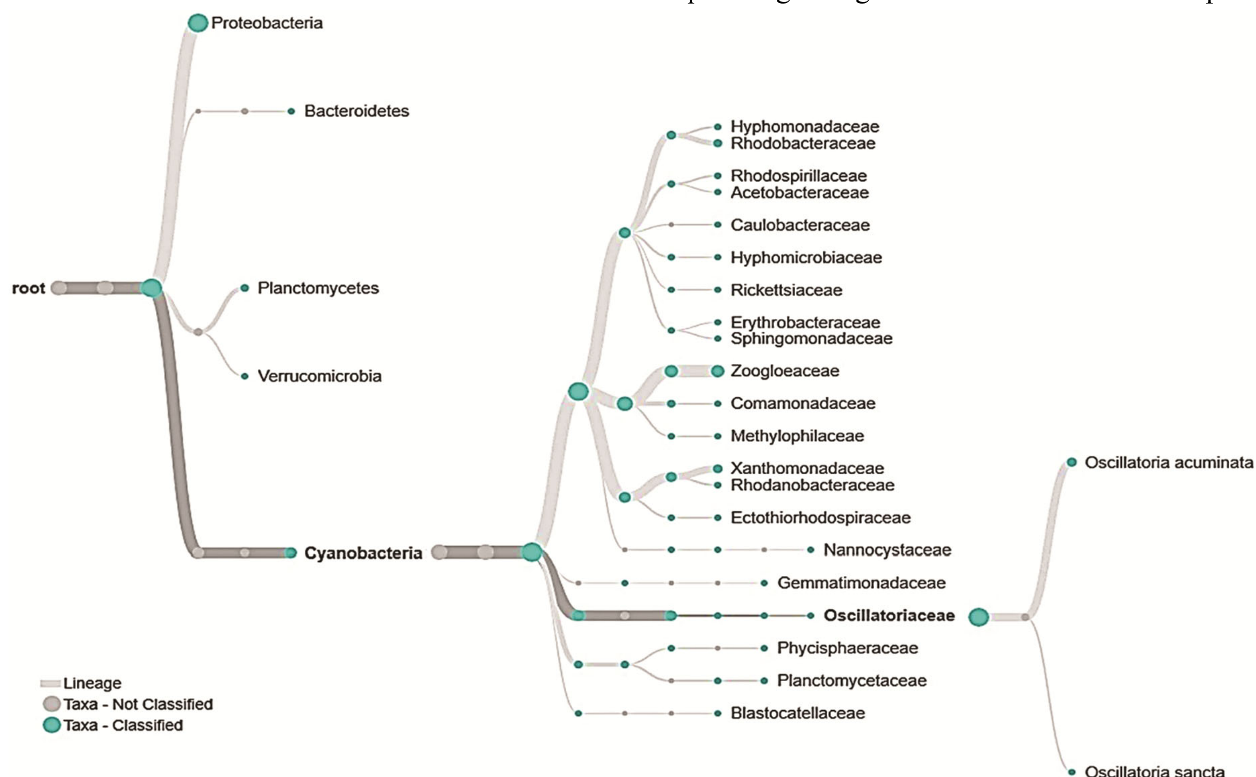

Fig. 4 — Cladogram of identified cyanobacterial species

testable hypotheses regarding microbial community structure and function based on data generated and analysed while in the field.

Characterized and estimated the diversity of cyanobacteria in biological soil crust in the sacred grove forest of Tamil Nadu, India[22]. They found 16 different genera commonly associated with biological soil crust and *Oscillatoria sancta* and *Phormidium autumnale* were found dominated species. Researchers was identified the cyanobacterial/ microalgae species *i.e*. *Senedesmus obliguus* and *Scenedesmus elongatus* by 16S rRNA & 18s rRNA sequencing[21].

Recently, scientist was reported the 3.5-Mb draft genome sequence of the cyanobacterium *Synechococcus sp*. strain Nb3U1 by using nanopore sequencer, which was isolated from a microbial mat sample collected from Nakabusa Hot Spring, Nagano, Japan[23].

In this study, the two efficient cyanobacterial species (*Oscillatoria species*-1 & *Oscillatoria species*-2) were molecular identified as *Oscillatoria acuminata* and *Oscillatoria sancta* by 16S rRNA sequencing using Nanopore sequencing Technology and DNA sequence was submitted to NCBI with stain name VMRJHK2020 and VMRJHK012020 and the Gene bank accession are MW364270 and MW364271, respectively.

Conflict of interest

The authors declare no conflict of interest.

References

1 Karthick SRN, Shivaramakrishnan K & Bharani ARS, Potential antioxidative protein-pigment complex Spirulina platensis mediated food grade phycocyanin C -Extraction, purification, antioxidative activity and biocompatibility. *Indian J Biochem Biophys*, 56 (2019) 230.

2 Dutta A, Dutta P, Kayal T, Ray S, Dutta M, Ghosh K, Biswas S, Sarkar S, Bera A, Saha M, Mandal K, Bhar A & Paul S, Fabrication of PVA-Silver nanoparticle composite film for elimination of microbial contaminant from effluent. *Indian J Biochem Biophys*, 59 (2022) 936.

3 Paerl HW, Nuisance phytoplankton blooms in coastal, estuarine, and inland waters. *Limnol Oceanogr*, 33 (1988) 823.

4 Neilan BA, Identification and phylogenetic analysis of toxigenic cyanobacteria by multiplex randomly amplified polymorphic DNA PCR. *Appl Environ Microbiol,* 61 (1995) 2286.

5 Shiva S & Pavankumar P, Novel primer designing and PCR-AFLP approach for an expeditious detection of coliforms in potable waters. *Indian J Biochem Biophys*, 60 (2023) 65.

6 Hay VD, Sai KR, Payal G, Rohini S, Devendra KY & Prasanta KD, Cloning and structural elucidation of a brassinosteroids biosynthetic gene (Atdwarf4) and genetic transformation of Indian mustard (*Brassica juncea* L.). *Indian J Biochem Biophys*, 59 (2022) 320.

7 Pooja GT, Jaydeep BG, Fesal IG & Gaurang MS, Simultaneous detection of thirteen exons of dystrophin gene by optimized multiplex PCR assay to screen Duchenne/Becker muscular dystrophy. Indian Journal of *Indian J Biochem Biophys*, 60 (2023) 31.

8 Ernst A, Marschall P & Postius C, Genetic diversity among *Synechococcus* spp. (cyanobacteria) isolated from the pelagial of Lake Constance. *FEMS Microbiol Ecol*, 17 (1995) 197.

9 Thajuddin N & Subramanian G, Cyanobacterial biodiversity and potential applications in biotechnology. *Curr Sci*, 4 (2005) 47.

10 Kumari N, Srivastava A, Bhargava P & Rai LL, Molecular approaches towards assessment of cyanobacterial biodiversity. *Afr J Biotech*, 8 (2009) 228.

11 Johnson JS, Spakowicz DJ, Hong BY, Petersen LM, Demkowicz P, Chen L, Leopold S R, Hanson BM, Agresta HO, Gerstein M & Sodergren E Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat Commun*, 10 (2019) 1.

12 Yunhao W, Yue Z, Audrey B, Yuru W & Kin FA, Nanopore sequencing technology, bioinformatics and applications. *Nat Biotech*, 39 (2021) 1348.

13 Lara U, Andre Ho, Jotautas B, Michael BH, Philipp BW, Michael JS, Daniel JK, Surangi NP, Daniel EM, Edward TT, Susannah JS & Maximilian RS, Freshwater monitoring by nanopore sequencing. *eLife*, 10 (2021) 1.

14 Tvedte ES, Gasser M, Sparklin BC, Michalski J, Hjelmen CE & Johnston JS, Comparison of longread sequencing technologies in interrogating bacteria and fly genomes. *G3 (Bethesda)*, 11 (2021) 1.

15 Wick RR, Judd LM, Gorrie CL & Holt KE, Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol*, 13 (2017) 1.

16 Joko PT, Faisal A & Bens P, Identifying Cyanobacteria through Next-Generation Sequencing Technology for Modern Agriculture. *E3S Web of Conf*, 388 (2023) 1.

17 Ferris MJ, Muyzer G & Ward DM, Denaturing gradient gel electrophoresis profiles of 16S rRNA-defined populations inhabiting a hot spring microbial mat community. *Appl Environ Microbiol*, 62 (1996) 340.

18 Gopinath SM, Niruba J & Meghana R, Isolation, Molecular Identification and Comparative Lipid Profiling of Microalgae and Cyanobacteria. *Int J Sci Res*, 3 (2012) 2353.

19 Vanithasree V & Murugesan S, Growth and nutrient removal rates of *Oscillatoria acuminata* and *Scenedesmus armatus* in aquaculture wastewater: a laboratory scale study. *Biosci Biotechnol Res Asia,* 7 (2010) 701.

20 Shin J, Lee S, Go MJ, Lee SY, Kim SC, Lee CH & Cho BK, Analysis of the mouse gut microbiome using full-length 16S rRNA amplicon sequencing. *Sci Rep*, 6 (2016) 29681.

21 Arwyn E, Aliyah RD, Birgit S, Luis AJM & Andrew JH, Extreme metagenomics using nanopore DNA sequencing. *Field report from Svalbard 78 °N*, (2016) 073965.

22 Vinoth M, Jeevanantham, G, Muruganantham P & Mohammed HJ, Characterization and estimated diversity of cyanobacteria in biological soil crust in sacred grove forest of Tamil Nadu, India. *J Pharmacogn Phytochem*, 8 (2019) 849.

23 Christopher LP, Satoshi O, Hideaki M & Scott RM, Draft Genome Sequence of the Cyanobacterium *Synechococcus* sp. Strain Nb3U1. *Am J Microbiol*, 11 (2022) 1.