Draft Genome Sequence of Filamentous Marine Cyanobacterium *Lyngbya confervoides* Strain BDU141951

Mathu Malar Chandrababunaidu, Diya Sen, Sucheta Tripathy

Structural Biology and Bioinformatics Division, CSIR, Indian Institute of Chemical Biology, Kolkata, West Bengal, India

Lyngbya confervoides strain BDU141951 is a fast-growing, unicellular, marine, nonheterocystous cyanobacterium forming long unbranched filaments inside sheaths. Here, we report the draft genome assembly of *Lyngbya confervoides* BDU141951 for the first time. The genome size is 8,799,693 bp and has 6,093 putative protein-coding genes assembled into 298 scaffolds.

Address correspondence to Sucheta Tripathy, tsucheta@gmail.com.

Cyanobacteria are well known for their ability to produce sustainable bioenergy [1]. Many efforts are on to reengineer these organisms for overproduction of biofuels from fast-growing strains. *Lyngbya* fits very well into this category with its fast growth and stress tolerance, as it can grow in saline as well as in freshwater conditions. Several strains of *L. aestuarii* are capable of producing hydrogen gas [2], and some strains of *L. aestuarii* are also known to produce various antifungal compounds [3]. Another species, *L. officinalis*, produces a toxin called *Lyngbya* toxin-A, which is characterized by its antiangiogenic and anti proliferative properties [4].

Lyngbya confervoides is a marine cyanobacterium that was cultured in ASN medium at room temperature (approximately 27°C) with 16-h light and 8-h dark periods. Cultures were gently shaken once a day. Morphologically, *L. confervoides* forms green clumps in liquid culture and is thick and matty in solid media. Genomic DNA was isolated using a Uniflex Bacterial DNA isolation kit (Genei, USA). The total amount of DNA isolated was approximately 3.6 μg, with a concentration of 330 ng/μl. Two libraries, one paired-end (300-bp insert size) and one mate-pair library (3-kb insert size), were constructed. Whole-genome sequencing was carried out using the Illumina HiSeq platform. Paired-end libraries were sequenced at 98× coverage (12.3 million reads), and the mate-pair library was sequenced at 40× coverage (5.6 million reads); the average read length was 101 bp for both libraries. Initially, the sequence artifacts and adapters were removed by using SGA [5] and Tagdust [6]. The clean reads from both libraries were assembled using Allpaths LG-49856 [7]. The final assembly was 8.7 Mb in length, having 298 scaffolds and an N_{50} of 5,207,129. The largest scaffold was 5,207,129 bp and the smallest scaffold was 3,620 bp in length. The calculated G+C content of the genome was 55%. The sequences were submitted to GenBank and were annotated using the PGAP pipeline (http://www.ncbi.nlm.nih.gov/genome/annotation_prok). There were 6,093 protein-coding genes, 1,096 pseudogenes, 4 CRISPR arrays, 70 tRNA genes, and 2 ncRNA genes predicted from this assembly. The protein sequences were submitted for KEGG analysis using the KAAS pipeline [8]. Pathway analysis revealed a nonribosomal peptide synthetic pathway leading to the production of viomycin (Vioj), which has antituberculosis activity. Since *Lyngbya* survives higher salinity conditions, several salt-tolerant genes, such as spermidine synthase spermidine transporters [9], are found in the genomes. Spermidine synthase protein from *Lyngbya confervoides* BDU141951 shares closest similarity with *Leptolyngbya* sp. PCC 6406 at 83% identity. Genome mining of this organism will help find many important genes for metabolite production as well as stress tolerance for biotechnology applications.

Nucleotide sequence accession number. The *Lyngbya confervoides* BDU141951 draft genome sequence data have been deposited in GenBank under the accession number JTHE00000000.

REFERENCES

1. **Nozzi NE, Oliver JW, Atsumi S.** 2013. Cyanobacteria as a platform for biofuel production. Front Bioeng Biotechnol **1**:7. http://dx.doi.org/10.3389/fbioe.2013.00007.
2. **Kothari A, Parameswaran P, Garcia-Pichel F.** 2014. Powerful fermentative hydrogen evolution of photosynthate in the cyanobacterium *Lyngbya aestuarii* BL J mediated by a bidirectional hydrogenase. Front Microbiol **5**:680. http://dx.doi.org/10.3389/fmicb.2014.00680.
3. **Kumar M, Singh P, Tripathi J, Srivastava A, Tripathi MK, Ravi AK, Asthana RK.** 2014. Identification and structure elucidation of antimicrobial compounds from *Lyngbya aestuarii* and *Aphanothece bullosa*. Cell Mol Biol (Noisy-le-grand) **60**:82–89.
4. **Kyadari M, Fatma T, Velpandian T, Malliga P, Bharat N, Fareha B.** 2014. Antiangiogenic and antiproliferative assessment of cyanobacteria. Indian J Exp Biol **52**:835–842.
5. **Simpson JT, Durbin R.** 2012. Efficient *de novo* assembly of large genomes using compressed data structures. Genome Res **22**:549–556. http://dx.doi.org/10.1101/gr.126953.111.
6. **Lassmann T, Hayashizaki Y, Daub CO.** 2009. TagDust—a program to

eliminate artifacts from next generation sequencing data. Bioinformatics **25:**2839 –2840. http://dx.doi.org/10.1093/bioinformatics/btp527.

7. **Ribeiro FJ, Przybylski D, Yin S, Sharpe T, Gnerre S, Abouelleil A, Berlin AM, Montmayeur A, Shea TP, Walker BJ, Young SK, Russ C, Nusbaum C, MacCallum I, Jaffe DB.** 2012. Finished bacterial genomes from shotgun sequence data. Genome Res **22:**2270 –2277. http://dx.doi.org/10.1101/gr.141515.112.

8. **Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M.** 2007. KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res **35**(Suppl 2):W182–W185. http://dx.doi.org/10.1093/nar/gkm321.

9. **Gupta B, Huang B.** 2014. Mechanism of salinity tolerance in plants: physiological, biochemical, and molecular characterization. Int J Genomics **2014:**701596. http://dx.doi.org/10.1155/2014/701596.