Draft genome sequence of the thermotolerant cyanobacterium *Desertifilum* sp. IPPAS B-1220

Kirill S. Mironova¹#, Maria A. Sinetovaa¹, Kenzhegul Bolatkhanb, Bolatkhan K. Zayadanb, Vera V. Ustinovac, Elena V. Kupriyanovaa¹, Alexandra N. Skrypkina, Natalya E. Gogolevida, Yuriy V. Gogolevd, Dmitry A. Losa

Institute of Plant Physiology, Russian Academy of Science, Moscow, Russiaa; Department of Biotechnology, Faculty of Biology and Biotechnology, Al-Farabi Kazakh National University, Almaty, Kazakhstanb; Microbiology Department, Central Tuberculosis Research Institute RAMS, Moscow, Russiac; Kazan Institute of Biochemistry and Biophysics, Kazan Science Centre, Russian Academy of Sciences, Kazan, Russiad

Running Head: Draft genome of the cyanobacterium *Desertifilum* sp.

#Correspondence to Kirill S. Mironov: kirill.mironoff@gmail.com
Abstract

We report the draft genome of the filamentous cyanobacterium *Desertifilum* sp. strain IPPAS B-1220 isolated from Lake Shar-Nuur, Mongolia. The genome of 6.1 Mb codes for 5113 genes. Genome mining revealed 10 clusters for the synthesis of bioactive compounds (non-ribosomal peptides, polyketides, bacteriocins, lantipeptides) with potential biotechnological or medical importance.

Cyanobacteria produce a broad range of secondary metabolites with diverse chemical structures. The majority of these metabolites are the products of non-ribosomal peptide synthetase (NRPS) or polyketide synthase (PKS) pathways (1). We sequenced the genome of the cyanobacterium *Desertifilum* sp. strain IPPAS B-1220 (Oscillatoriaceae), which was newly isolated from fresh-water Lake Shar-Nuur, Mongolia.

Genomic DNA was isolated from cells grown at stationary phase by incubation with saturated iodide solution followed by lysozyme treatment and 2% SDS lysis at 70°C (2). DNA was purified by phenol/chloroform extraction. Isolated DNA was fragmented by adaptive focused acoustics technology using Covaris S220 ultrasonicator (Covaris, Woburn, MA, USA). Parameters for DNA shearing were adapted to obtain 500 bp fragments. DNA library was prepared with NEBNext® Ultra™ DNA Library Prep Kit for Illumina (NEB, Ipswich, MA, USA). Sequencing was performed on MiSeq system with MiSeq reagent kit v3 in 600 cycle paired-end format. Sequence quality were analysed by FastQC (3). Adapters and low-quality nucleotides were trimmed with Trimmomatic (4). Genome was assembled using AbySS (5) and SPAdes (6) software. The draft genome assembly quality was analyzed by QUAST (7). Draft genome coverage
was no less than 30, N50 statistics $> 10^5$. Approximate genome size is 6.1 Mb, the average G+C content was estimated as 48.7%.

Genome was annotated using automated NCBI Prokaryotic Genome Annotation Pipeline. It contains 5113 genes: 4964 genes coding for proteins, 86 pseudo genes, 6 rRNA-coding sequences (4 for 5S, 1 for 16S, and 1 for 23S), 53 tRNAs, 4 non-coding RNAs. Seven CRISPR arrays have been found in the genome. The genome sequence was analyzed with antiSMASH tool (8) for the presence of gene clusters coding for secondary metabolite biochemical pathways. Eight contigs contained gene clusters for secondary metabolites biosynthesis (non-ribosomal peptides, polyketides, bacteriocins, lantipeptides, etc.). Thus, genome mining of Desertifilum sp. IPPAS B-1220 revealed a set of putative bioactive compounds that may have further biotechnological or pharmaceutical applications.

**Accession number(s).** Genome sequence of Desertifilum sp. strain IPPAS B-1220 was deposited at NCBI with the following attributes: SAMN05788062 for BioSample, PRJNA343432 for BioProject, and SRR4255595 for SRA database. Genome PGAP file accession number is MJGC00000000.

**Acknowledgements**

The authors are grateful to Common use center "Biotechnology" of All-Russian Scientific Research Institute of Agricultural Biotechnology (Moscow, Russia) for the help with bioinformatic analysis.

**Funding information**
This work was supported by a grant from Russian Science Foundation (no. 14-24-00020) to D.A.L. and by a grant from Ministry of Education and Science of Republic of Kazakhstan to B.K.Z.

References


