Data Article

Metagenomic data of the microbial community of the chemocline layer of the meromictic subarctic Lake Bolshie Hruslomeny, North European Russia

Vitaly V. Kadnikov a,*, Alexander S. Savvichev b, Andrey V. Mardanov a, Alexey V. Beletsky a, Nikolai V. Ravin a, c, Nikolai V. Pimenov b

a Institute of Bioengineering, Research Center of Biotechnology of the Russian Academy of Sciences, Moscow, Russia
b Winogradsky Institute of Microbiology, Research Center of Biotechnology of the Russian Academy of Sciences, Moscow, Russia
c Faculty of Biology, Lomonosov Moscow State University, Moscow, Russia

Article history:
Received 7 November 2018
Received in revised form 18 February 2019
Accepted 20 February 2019
Available online 2 March 2019

ABSTRACT

The Lake Bolshie Hruslomeny is located on the shores of the Kandalaksha Bay of the White Sea, North European Russia. This lake, formed from the sea bay and still retaining the subsurface connection with the sea, is meromictic, with a fresh oxygenated upper layer and an anoxic brackish hypolimnion with high concentrations of methane and hydrogen sulphide. To characterize the microbial communities involved in the carbon and sulfur cycles in the lake, we sequenced the metagenome of a water sample collected at the chemocline level. At the phylum level, Chlorobi, Proteobacteria, Bacteroidetes and Firmicutes were the most numerous groups. The obtained data will help investigate the diversity and ecological role of the microbial community in the Lake Bolshie Hruslomeny and provide insight into the biogeochemical processes in subarctic lakes. The raw sequencing data is available from the NCBI Sequence Read Archive (SRA) database under the BioProject PRJNA503531.

© 2019 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
1. Data

Meromictic lakes are the subject of research in the field of limnology, biogeochemistry and microbiology [1,2]. Bolshie Hruslomeny Lake is located on the shores of the Kandalaksha Bay of the White Sea, North European Russia. This lake was originally a sea gulf, artificially separated from the sea by a dam at the beginning of the 20th century. The maximum depth of the lake is about 18 m. The upper epilimnion layer (0–3.5 m) is fresh and oxygenated. The water in the hypolimnion (below 5 m) is anoxic, brackish (up to 22‰) due to infiltration of seawater, has high concentrations of methane (up to 1.8 mM) and hydrogen sulphide (up to 17.8 mM). The highest rates of methane oxidation and anoxygenic photosynthesis were observed in the chemocline zone. This lake is an interesting object for studying the microbial processes related to the methane and sulfur cycles. To characterize the microbial communities involved in these processes, we performed a metagenomic shotgun sequencing of a water sample collected at the chemocline level.

Of the 10,845,687 sequencing reads, 51.33% were assigned to Bacteria, 0.90% to Archaea, 0.11% to Eukaryota, 0.18% to viruses, while other reads were not classified. At the phylum level, Chlorobi (28.98%), Proteobacteria (9.47%, mostly members of the class delta), Bacteroidetes (3.68%), and Firmicutes (1.57%) were the most abundant, as shown in Fig. 1. At the species level, a single bacterium, Chlorobium phaeovibrioides, accounted for 23.69% of all reads. The obtained data will help investigate the diversity and ecological role of the microbial community in the Lake Bolshie Hruslomeny and provide insight into the various biogeochemical processes in subarctic lakes.

2. Experimental design, materials, and methods

2.1. Sample collection and preparation

The water sample was collected from the chemocline level (at a depth of 4.25 m) of the Lake Bolshie Hruslomeny in March 2017 using sterile plastic containers. The concentration of microorganisms, according to microscopic estimation, was about $1.9 \times 10^7$ cells/ml. The sample was delivered to the
laboratory on ice for 3 h. Cells from 75 ml of water were collected on 0.22 μm cellulose nitrate membranes (Sartorius, Germany) using a Sartorius filtration unit.

2.2. DNA extraction

The filters were frozen in liquid nitrogen and then ground and mixed with TE buffer (pH 8.0) in a 37°C water bath. Total DNA was extracted using Power Soil DNA Isolation Kit (MO BIO Laboratories Inc, Carlsbad, USA). The quality and concentration of the extracted DNA sample was measured using Qubit® dsDNA HS Assay Kit (Life Technologies), followed by agarose gel electrophoresis. About 1 μg of total DNA (20 ng/μL) was isolated.

2.3. Sequencing and taxonomic analysis

Sequencing library was prepared using 100 ng of DNA with the Nextera DNA Library preparation kit (Illumina Inc., USA) following the manufacturer’s instructions. The sequencing of this library on the Illumina HiSeq2500 platform using HiSeq Rapid Run v2 sequencing reagents generated 11,061,025 single-end reads with a length of 250 nt (2.8 Gbp in total). Primer and quality trimming were performed with Cutadapt v. 1.17 [3] and Sickle v. 1.33 (https://github.com/najoshi/sickle), respectively. Cutadapt was used with the default settings, and Q30 score was used for the Sickle. 10,845,687 reads with an average length of 193 nt were retained after processing.

Taxonomic classification of filtered reads was carried out using the Kaiju program [4], with default parameters. A microbial subset of the NCBI non-redundant protein database, also including fungi and microbial eukaryotes (nr + euk option), was used as a reference database. 5,695,944 of 10,845,687 reads were classified.

Acknowledgments

This study was performed using the scientific equipment of the Core Research Facility ‘Bioengineering’ (Research Center of Biotechnology RAS) and financially supported by the Russian Science Foundation [grant number 16-14-10201].
Transparency document

Transparency document associated with this article can be found in the online version at https://doi.org/10.1016/j.dib.2019.103800.

References

[1] B. Boehrer, M. Schultze, Stratification of lakes, Rev. Geophys. 46 (2008) RG2005.
[2] A.C. Lehours, C. Bardot, A. Thenot, D. Debroas, G. Fonty, Anaerobic microbial communities in Lake Pavin, a unique meromictic lake in France, Appl. Environ. Microbiol. 71 (2005) 7389–7400.
[3] M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads, EMBnet J. 17 (2011) 10–12.
[4] P. Menzel, K.L. Ng, A. Krogh, Fast and sensitive taxonomic classification for metagenomics with Kaiju, Nat. Commun. 7 (2016) 11257.