ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

The Monitoring of Nizhniy Kaban Lake by Rbcl Gene of Fresh Water Ater Organisms Using Next-Generation Sequencing.

Ludmila L. Frolova*, and Anthony E. Sverdrup.

Kazan Federal University.

ABSTRACT

The ecological state of the environment, including water bodies, is evaluated by different methods, one of which is bioindication. The method of bioindication is based on the study of the state of biota in natural conditions, the monitoring of the composition and the number of indicator species inhabiting the medium under study, the identification of which has been carried out by researchers visually, using microscopy methods. As an alternative, the paper considers a method for identifying hydrobionts in a water sample using the *rbcL* marker gene of freshwater organisms on the basis of modern sequencing methods. The results of the analysis are presented and the water quality of Nizhniy Kaban lake (Kazan, Russia) is estimated on the basis of DNA-barcoding by a next-generation sequencing method. The sequenced sequences of the fragment of the *rbcL* hydrobiont gene of fresh-water Lake Nizhniy Kaban in the autumn (2016) and summer (2017) sampling periods in the fastq format are included in the international database on the NCBI's website with unique numbers SRR7470969, SRR7459788 and SRR7463326. A comparative analysis of metagenomic data shows that the majority of the organisms of Lower Kaban Lake are grouped by the *rbcL* gene near the b-a mesosaprobic zone. Lower Kaban lake can be characterized as contaminated, transitional from b-mesosaprobic to a-mesosaprobic by water quality.

Keywords: bioindication, saprobity, *rbcL* gene, next-generation sequencing.

Corresponding author

ISSN: 0975-8585

INTRODUCTION

It is known that one of the highly informative methods of environmental assessment of the aquatic environment is the use of the bioindication method, that is, the detection of the nonspecific reaction of living organisms to changes in their habitat, for example pollution or purification [1]. In comparison with chemical and physical methods of water reservoirs indication, bioindication methods have a significant advantage, since physical and chemical indices make an assessment of water quality only at a given time, whereas the presence of indicator species of plants, animals, fungi or bacteria makes it possible to more accurately determine the quality of water in a pond [2,3]. The method consists in determining the indicator organisms according to the morphological features, and the researcher's opinion may be subjective, since the identification of certain species requires the knowledge of narrow specialists because many species have very small dimensions, some species have strong sexual dimorphism, or vice versa, there are sibling species and fewer than all specialists can identify the organism at its larval stage [4].

At present, methods of molecular genetic analysis allow for instrumental identification of organisms accurate to species. Thus, for example, the DNA-barcoding method [5] is used for this purpose. Numerous DNA barcodes by kinds of organisms are accumulated in the international database of nucleotide sequences – GenBank [6]. The DNA-barcoding method is based on a sequence of nucleotides of the DNA-barcode, which is the same for individuals of only one species; for example, for animals it is a variable fragment of the CB1 gene with a length of 600-700 base pairs, for plants and algae – a fragment of the *rbcL* gene [7,8] . We previously used this technique to identify zooplankton organisms from the variable fragment of the *CO1* gene for assessing the environmental state of freshwater reservoirs by bioindication [9]. This paper provides an assessment of the water quality of Nizhniy Kaban lake (Kazan, Russia) using the *rbcL* marker gene of hydrobionts based on the next-generation sequencing method.

METHODS

Sampling from Nizhniy Kaban lake (Kazan) was conducted in September 2016 and in July 2017 in accordance with standard hydrobiological methods [10] and using the Apstein network by straining 100 liters of water.

Isolation of DNA from the precipitate obtained by centrifugation of 50 ml of the sample at a rate of 10,000 g for 15 min was carried out using the FAST DNA Kit (MP biomedicals) according to the manufacturer's protocol. Amplification of the isolated DNA was performed by Phusion High-Fidelity DNA polemerase (Thermo Fisher) using the primers (Table 1).

Table 1. Primers for PCR of rbcL gene of phytoplankton

Primers	Sequences
rbcL_AB_FI (forward)	5'-tcgtcggcagcgtcagatgtgtataagagacagtcigciaaraactayggtcg-3'
rbcL_AB_RI (reverse)	5'-gtctcgtgggctcggagatgtgtataagagacagggcatrtgccaiacrtgrat-3'
rbcL_D_FI (forward)	5'-tcgtcggcagcgtcagatgtgtataagagacaggatgatgaraayattaactc-3'
rbcL_D_RI (reverse)	5'-gtctcgtgggctcggagatgtgtataagagacagatttgdccacagtgdatacca-3'

A pair of primers rbcL_AB_FI (forward) and rbcL_AB_RI (reverse) were used to identify

Cyanobacteria, Chlorophyta and *Proteobacteria*; a pair of primers rbcL_D_FI (forward) and rbcL_D_RI (reverse) – to identify *Bacillariophyta, Pyrrophyta, Cryptophyta* and *Haptophyta.*

Purification of the PCR products was carried out with Agencourt AMPure XP beads (Beckman Coulter), followed by a second PCR for indexing the samples (Nextera XT indices).

The resulting DNA libraries were sequenced on an Illumina MiSeq (MiSeq Reagent kit v3). Metagenomic data were included in the international SRA database on the NCBI's website [6].

The obtained metagenomic data was aligned with the BLAST + program to establish the species diversity and for subsequent analysis.

RESULTS AND DISCUSSION

The sequenced sequences of the fragment of the *rbcL* hydrobiont gene of freshwater Nizhniy Kaban lake in the autumn (2016) and summer (2017) sampling periods in the fastq format are included in the SRA international database on the NCBI's website with unique numbers: SRR7470969 (2016), SRR7459788 (2017) *и* SRR7463326 (2017). After filtering the reads by quality, trimming the service sequences, and removing the chimeric sequences, the resulting nucleotide sequences were aligned with the BLAST + program to establish the taxonomic composition.

The sequencing of the sequences (rbcL_AB_FI (forward) and rbcL_AB_RI (reverse) primers) of the fragment of the *rbcL* hydrobiont gene of freshwater Lower Kaban lake in autumn (2016) and summer (2017) helped us to identify by years respectively: 62/40 species of *Bacteria* and 42/64 species of *Viridiplantae*. The percentage of taxonomic groups of Lower Kaban lake by species (autumn 2016 / summer 2017) is: 59.62% / 38.46% of *Bacteria*, 40.38% / 61.54% of *Viridiplantae* (Figure 1).

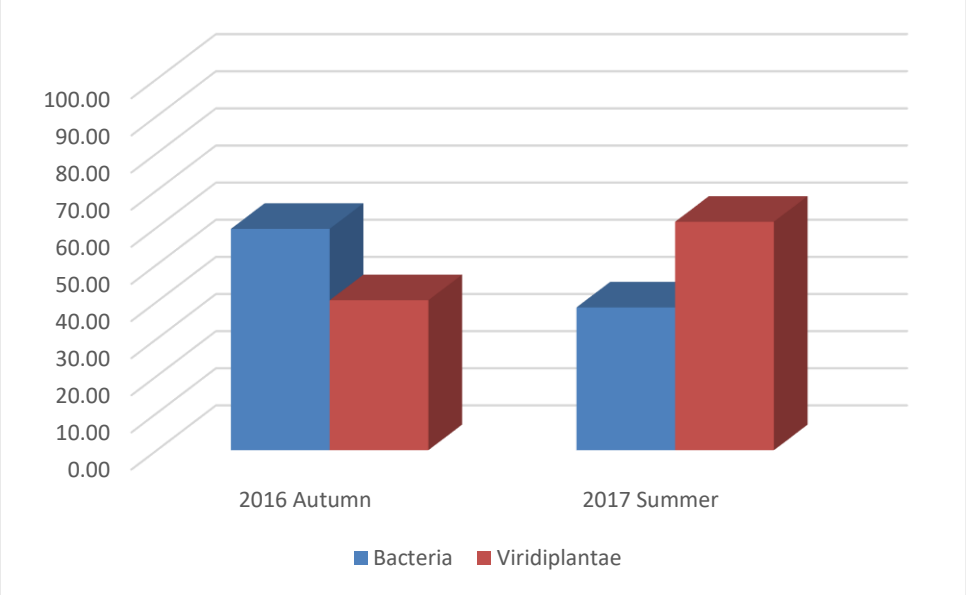

Figure 1. The percentage of taxon by species diversity of Nizhniy Kaban lake (Autumn 2016; Summer 2017)

The percentage of taxonomic groups of Lower Kaban lake by reads (autumn 2016 / summer 2017) is: 98.94% / 97.93% of *Bacteria*, and 1.06% / 2.07% of *Viridiplantae* (Figure 2).

ISSN: 0975-8585

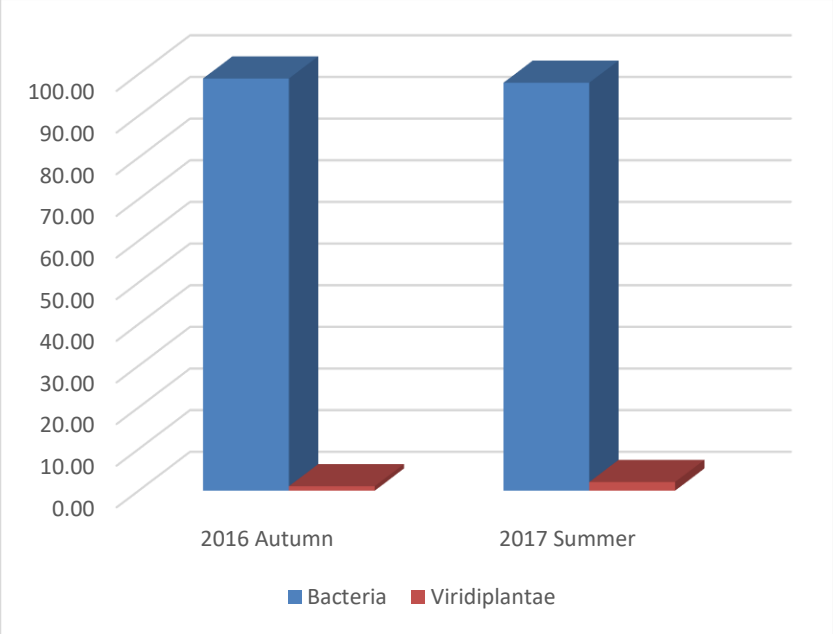

Figure 2. The percentage of taxon by reads of Nizhniy Kaban lake (Autumn 2016; Summer 2017)

In terms of the diversity of species we identified by years 2016/2017: 1/0 species of *Actinobacteria*, 2/0 species of *Bacteroidetes*, 38/27 species of *Cyanobacteria*, 9/9 species of *Proteobacteria*, 15/4 undetermined species of *Bacteria*, 37/61 species of *Chlorophyta* and 5/3 species of *Streptophyta*.

The results of the sequenced sequences of the fragment of the *rbcL* hydrobiont gene of freshwater Nizhniy Kaban lake in autumn (2016) and summer (2017) by diversity of species of *Bacteria* and *Viridiplantae* are given in percentage in Table 2 and are shown in Fig. 3.

Table 2. Species diversity of *Bacteria* and *Viridiplantae* in the percentage of Nizhniy Kaban lake

Taxon	Autumn 2016	Summer 2017
Actinobacteria	0.93	0.00
Bacteroidetes	1.87	0.00
Cyanobacteria	35.51	25.96
Proteobacteria	8.41	8.65
Unclassified bacterium	14.02	3.85
Chlorophyta	34.58	58.65
Streptophyta	4.67	2.88

ISSN: 0975-8585

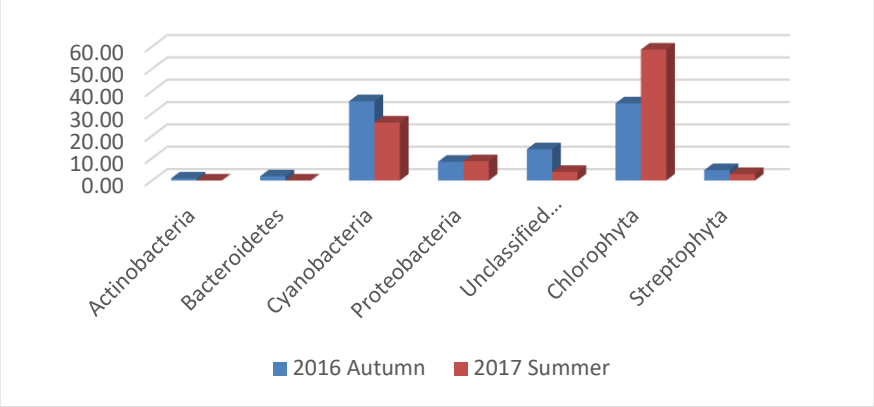

Figure 3. The percentage of species diversity of Bacteria and Viridiplantae of Nizhniy Kaban lake (Autumn 2016; Summer 2017)

In terms of the number of organisms we identified by years 2016/2017: 1/0 reads of *Actinobacteria*, 2/0 reads of *Bacteroidetes*, 39501/20119 reads of *Cyanobacteria*, 160/25 reads of *Proteobacteria*, 78/42 reads of the undetermined species *Bacteria*, 380/400 reads of *Chlorophyta* and 44/26 reads of *Streptophyta*.

The results of the sequenced sequences of the fragment of the *rbcL* hydrobiont gene of freshwater Nizhniy Kaban lake in autumn (2016) and summer (2017) by number of species of *Bacteria* and *Viridiplantae* are given in percentage in Table 3 and are shown in Fig. 4.

Table 3. The quantity of *Bacteria and Viridiplantae* (the quantity of reads) in the percentage of Nizhniy Kaban lake

Taxon	Autumn 2016	Summer 2017
Actinobacteria	0.00	0.00
Bacteroidetes	0.00	0.00
Cyanobacteria	98.34	97.61
Proteobacteria	0.40	0.12
Unclassified bacterium	0.19	0.20
Chlorophyta	0.95	1.94
Streptophyta	0.11	0.13

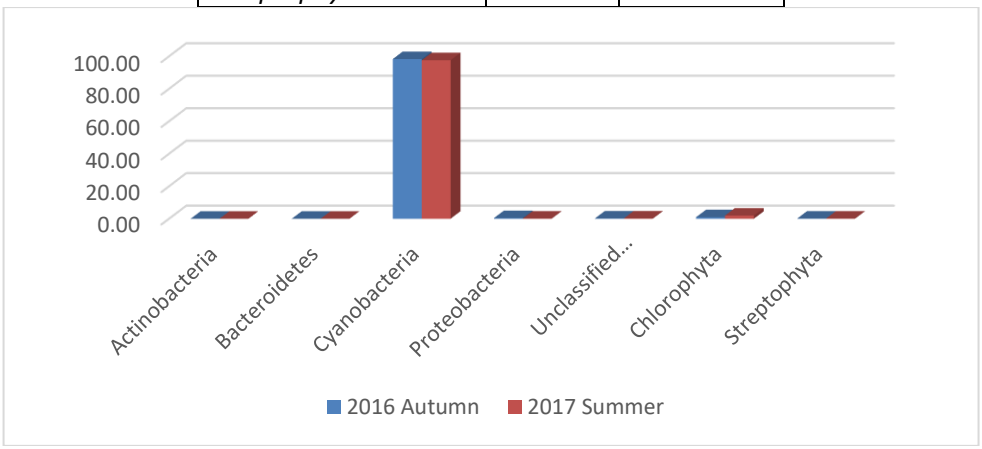

Figure 4. The quantity of *Bacteria and Viridiplantae* (the quantity of reads) in the percentage of Nizhniy Kaban lake (Autumn 2016; Summer 2017)

ISSN: 0975-8585

The sequencing of the sequences (rbcL_D_FI (forward) and rbcL_D_RI (reverse) primers) of the fragment of the rbcL hydrobiont gene of freshwater Lower Kaban lake in autumn (2016) and summer (2017) helped us to identify 17 species of Bacteria, 138 species of Stramenopiles, 55 species of Viridiplantae, and 94 species of unclassified Eukaryota.

The percentage of taxonomic groups of Lower Kaban lake by species and reads (summer 2017) is, respectively: 5.59%/42.35% Bacteria, 18.09%/18.09% Viridiplantae, 45.39%/20.83% Stramenopiles and 30.92%/18.72% Unclassified Eukaryota (Figure 5).

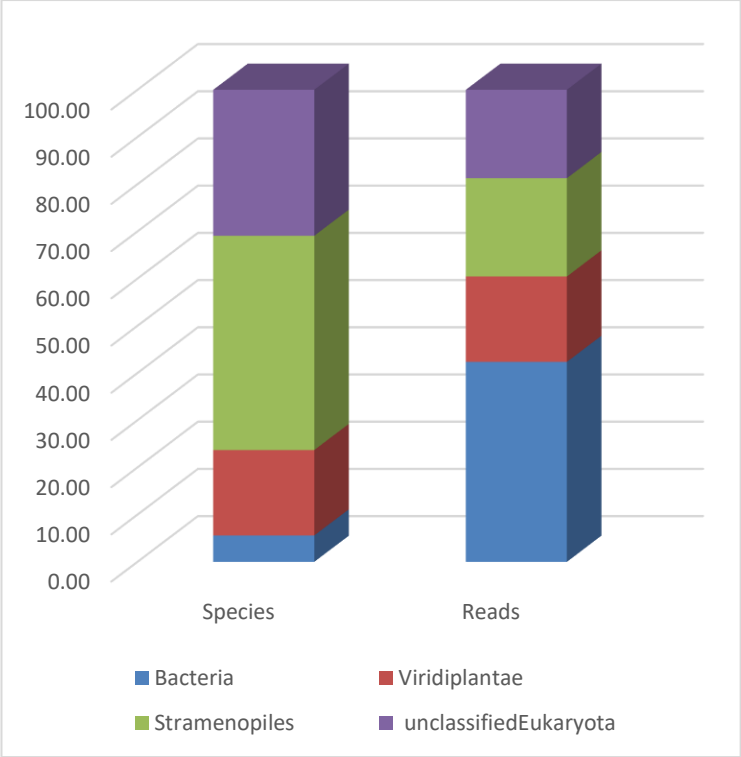

Figure 5. The percentage of taxon diversity and reads of hydrobionts of Nizhniy Kaban lake (Summer 2017)

The results of the sequenced sequences of the fragment of the *rbcL* hydrobiont gene of freshwater Nizhniy Kaban lake in summer (2017) by diversity of species and number of organisms, as well as in percentage are given in Table 3 and are shown in Fig. 6.

Table 3. The species diversity and quantity of freshwater organisms (the quantity of reads), also they are in the percentage of Nizhniy Kaban lake (Summer 2017)

Summer 2017	Species	Reads	Species, %	Reads, %
Uncultured Bacteria	2	6	0.66	0.02
Cyanobacteria	12	10512	3.95	42.24
Fusobacteria	1	1	0.33	0.00
Proteobacteria	2	22	0.66	0.09
Uncultured Viridiplantae	47	4386	15.46	17.62
Chlorophyta	8	117	2.63	0.47
Cryptophyta	9	1129	2.96	4.54
Haptophyta	2	3	0.66	0.01
Rhodophyta	1	1	0.33	0.00

ISSN: 0975-8585

Bacillariophyta	106	3662	34.87	14.71
Uncultured Stramenopiles	5	5	1.64	0.02
Chrysophyceae	4	56	1.32	0.23
Ochrophyta	11	328	3.62	1.32
Uncultured Eukaryota	94	4660	30.92	18.72

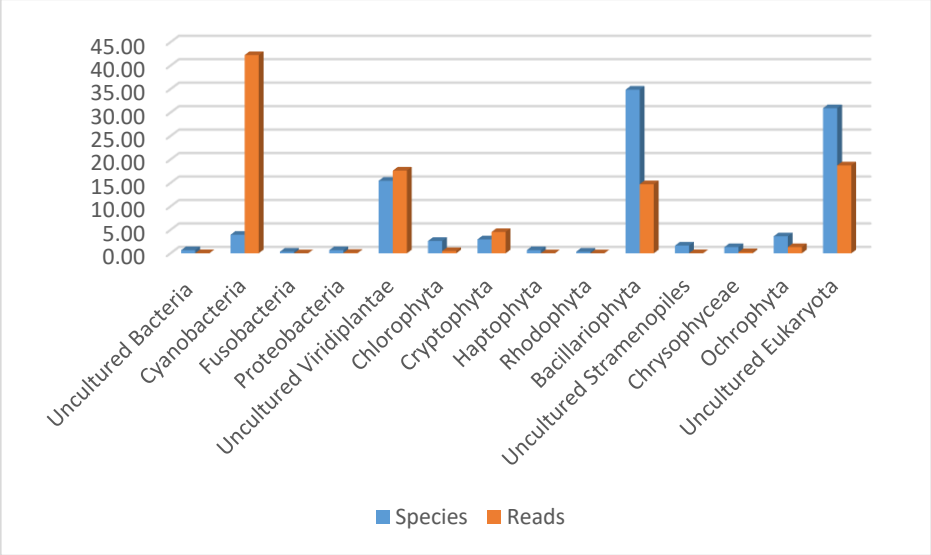

Figure 6. The percentage of species diversity and reads of hydrobionts of Nizhniy Kaban lake (Summer 2017)

The analysis of the metagenomic data on the *rbcL* gene fragment of the organisms of Nizhniy Kaban lake (autumn 2016) identified nine species classified as indicators as per V. Sladechek's list of the indicator organisms [11]:

2 species of *Bacteria*:

- *Microcystis aeruginosa* – indicator of the *b*-mesosaprobic zone with a weight of 1.75;
- *Oscillatoria tenuis* – indicator of the *a*-mesosaprobic zone with a weight of 2.85;

6 species of *Viridiplantae*:

- *Actinastrum hantzschii, Micractinium pusillum, Phacotus lenticularis* – indicators of the *b*-mesosaprobic zone with a weight of 2.00;
- *Ankistrodesmus falcatus* – indicator of the *b-a*-mesosaprobic zone with a weight of 2.35;
- *Chlorella vulgaris* – indicator of the *p-a*-saprobic zone with a weight of 3.6;
- *Chlorella pyrenoidosa* – indicator of the *p*-saprobic zone with a weight of 4.0;

The analysis of the metagenomic data on the *rbcL* hydrobiont gene of Nizhniy Kaban lake (summer 2017) identified 12 species classified as indicators as per V. Sladechek's list of the indicator organisms [11]:

1 species of *Bacteria*:

- *Microcystis aeruginosa* – indicator of the *b*-mesosaprobic zone with a weight of 1.75;

2 species of *Viridiplantae*:

- *Pandorina morum* – indicator of the *b*-mesosaprobic zone with a weight of 2.0;
- *Chlorogonium elongatum* – indicator of the *a*-mesosaprobic zone with a weight of 2.9;

9 species of *Stramenopiles*:

- *Gomphonema angustatum* – indicator of the *a*-saprobic zone with a weight of 1.15;
- *Amphora ovalis* – indicator of the *o*-*a*-mesosaprobic zone with a weight of 1.65;
- *Fragilaria capucina* – indicator of the *o*-*b*-mesosaprobic zone with a weight of 1.6;
- *Aulacoseira granulata* – indicator of the *b*-mesosaprobic zone with a weight of 1.8;
- *Gomphonema acuminatum* – indicator of the *b*-mesosaprobic zone with a weight of 1.7;
- *Cymbella cistula* – indicator of the *b*-mesosaprobic zone with a weight of 1.8;
- *Cyclotella meneghiniana* – indicator of the *a*-*b*-mesosaprobic zone with a weight of 2.6;
- *Nitzschia palea* – indicator of the *a*-mesosaprobic zone with a weight of 2.75;
- *Stephanodiscus hantzschii* – indicator of the *a*-mesosaprobic zone with a weight of 2.7.

The percentage of species of hydrobionts by zones of saprobity (autumn 2016, summer 2017) is shown in Fig. 7.

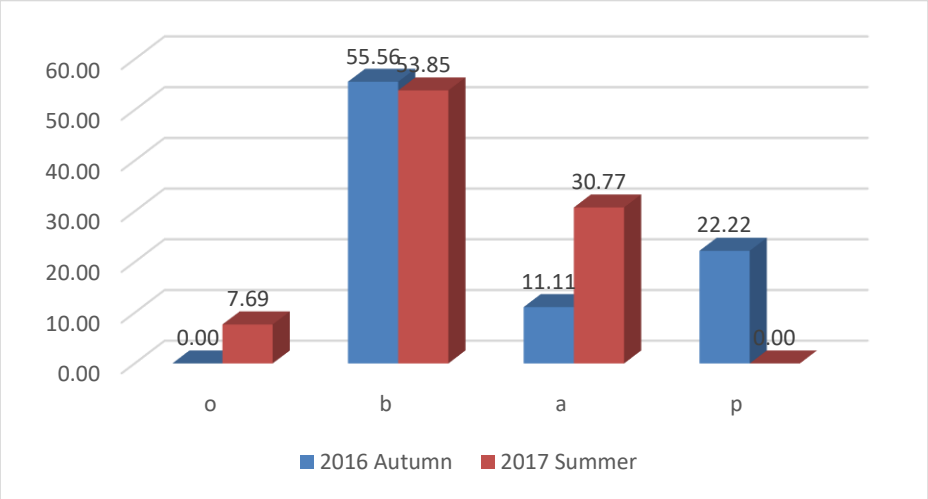

Figure 7. The percentage of hydrobionts saprobity by species of Nizhniy Kaban lake (Autumn 2016; Summer 2017)

The percentage of reads of hydrobionts by zones of saprobity (autumn 2016, summer 2017) is shown in Fig. 8.

A comparative analysis of metagenomic data on the *rbcL* hydrobiont gene shows that the majority of the organisms of Lower Kaban ake are grouped around the *b*-*a*-mesosaprobic zone (Fig. 7). In summer 2017, the diversity of species *Viridiplantae* (61.54%) is higher than the diversity of species *Bacteria* (38.46%), whereas in autumn 2016, bacterial organisms (59.62%) prevailed in Lower Kaban lake compared to *Viridiplantae* (40.38%) (Fig. 1), which was caused by intensive flowering of water during this period. The number of bacterial organisms significantly exceeded the number of *Viridiplantae*.

ISSN: 0975-8585

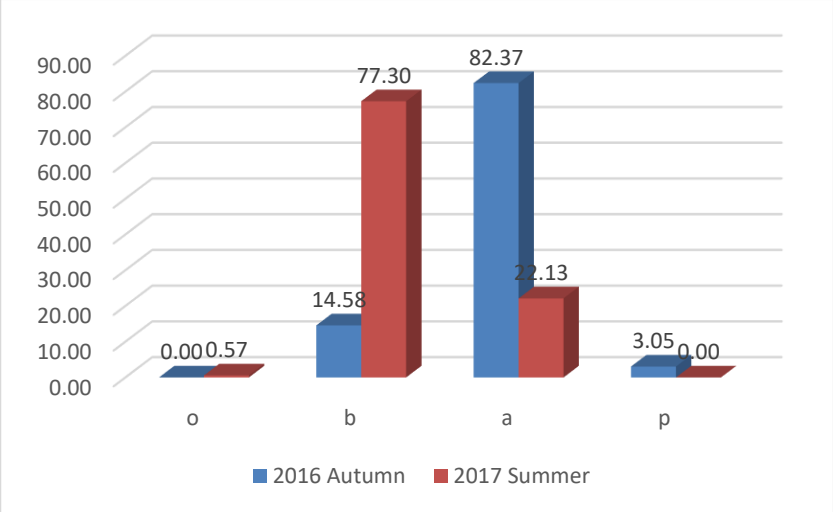

Figure 8. The percentage of hydrobionts saprobity by reads of Nizhniy Kaban lake (Autumn 2016; Summer 2017)

SUMMARY

Based on the results of the research using modern, next-generation sequencing methods, molecular and bioinformational analysis of water quality, Lower Kaban lake is transitional from *b*-mesosaprobic to *a*-mesosaprobic and can be characterized as contaminated.

CONCLUSIONS

One of the methods for assessing water quality is bioindication, which is based on the study of the state of biota in natural conditions, the monitoring of the composition and the number of indicator species inhabiting the medium under study, the identification of which has been carried out by researchers visually, using microscopy methods. The recent achievements in molecular biology have improved the identification of organisms by marker genes. The use of modern methods of molecular biology for bioindication gives positive results, increases the effectiveness and reliability of assessing the environmental state of water bodies. The obtained results are of great practical interest in monitoring of water bodies in particular, and the environment in general.

ACKNOWLEDGEMENTS

The work is performed according to the Russian Government Program of Competitive Growth of Kazan Federal University.

BIBLIOGRAPHY

[1] V.V. Kurilenko, O.V. Zaitseva, E.A. Novikova, N.G. Osmolovskaia, M.D. Ufimtseva, "Fundamentals of ecology, bioindication and biotesting of aquatic ecosystems", p.448, 2003.
[2] G.N. Miseiko, D.M. Bezmaternykh, G.I. Tushkova, "Biological test of fresh water quality", p.201, 2001.
[3] A.N. Sharov, «Indicator role of phytoplankton in the assessment of long-term changes concerning the quality of large lake waters», *Water resources*, vol. 35 (6), pp 679-702, 2008.
[4] A.Gomez, «Mating trials validate the use of DNA barcoding to reveal cryptic speciation of a marine bryozoans taxon», *Proceedings B of The Royal Society*, vol. 274, pp. 199-207, 2007.
[5] P.Hebert, S.Ratnasingham, J.R.deWaard, «Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species», Proc Roy Soc Lond B, vol. 270, pp. 96– 99, 2003.
[6] National Center for Biotechnology Information – http://www.ncbi.nlm.nih.gov/
[7] O.Folmer, M.Black, W.Hoeh, R.Lutz, R.Vrijenhoek, «DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates», Mol Mar BiolBiotechnol, vol. 3, pp. 294–299, 1994.

[8] CBOL Plant Working Group, «A DNA barcode for land plants», *Proc Natl AcadSci U S A*, vol.106 (31), pp 12794-12797, 2009.

[9] L.L.Frolova, A.M.Husainov, «Identification of indicator species of zooplankton organisms by COI gene fragment for estimation of ecological state of a water body», *International Journal of Pharmacy & Technology*, vol. 8, № 4, pp. 24477-24486, 2016.

[10] V.V. Bulion, "Methodological recommendations for collecting and processing materials in hydrobiological studies in freshwater. Phytoplankton and its products." - *L.: GosNIORKh,* p.32. 1981.

[11] V.Sladechek, «System of water quality from the biological point of view», *Arch. Hydrobiol. Ergeb. Limnol*, pp. 179-191, 1973.