UDC 579.25; 575.17; 574.5; 572.1/.4
IRSTI 34.27.23
DOI 10.37238/2960-1371.2960-138X.2025.99(3).147

Sergaliev N.Kh., Kakishev M.G.

West Kazakhstan University named after. M. Utemisova, Uralsk, The Republic of Kazakhstan

E-mail: sergaliev69@gmail.com , kakishev_murat@mail.ru

MICROBIOLOGICAL DETERMINANTS OF THE FUNCTIONING OF LAKES IN WESTERN KAZAKHSTAN FROM THE PERSPECTIVE OF THE FORESIGHT CONCEPT

Annotation. The article presents the results of microbiome studies conducted on Lakes Glubinnoye, Brusyanoe, and Karasu, located in the West Kazakhstan region, using high-throughput sequencing methods. The analysis revealed that unclassified bacteria, along with representatives of the phyla Proteobacteria, Bacteroidota, and Actinobacteriota, dominate in these aquatic ecosystems. The main differences among the microbiomes of the three lakes are associated with the proportion of unassigned bacterial taxa. It was also found that the taxonomic structures of the microbial communities across the lakes share common features. All samples contained representatives of Verrucomicrobiae, Gamma- and Alphaproteobacteria, Cyanobacteria, Bacteroidia, Actinobacteria, and Acidimicrobiia. Beta diversity analysis indicated that the microbiome composition varies depending on the sampling location. When assessing the taxonomic composition, it was observed that the microbiomes from Lake Glubinnoye are more similar to each other, while those from Lake Karasu, despite being grouped in the same cluster, exhibit greater diversity. Samples from Lakes Glubinnoye and Brusyanoe showed the presence of Candidatus Aquilinia (ranging from 2.6% to 7.5%) and relatively high levels of Algoriphagus (from 2% to 8%). A notable common feature between samples from Lakes Glubinnoye and Karasu is the relatively high abundance of the Methylophilaceae family (ranging from 1.3% to 3.8%).

Key words: microbiology; metagenomics; microbiome; lakes; taxonomic composition.

Introduction

The lakes located within the Ural (Zhayyk) River basin in Western Kazakhstan represent dynamic freshwater ecosystems characterized by pronounced physicochemical and biological gradients. These variations are driven by fluctuations in salinity, temperature, oxygen levels, nutrient concentrations, pH, and organic matter content. Due to the high environmental variability, the systematic collection of comprehensive datasets covering all ecosystem components is essential for effective monitoring of ongoing ecological transformations. In some cases, such transformations lead to severe

consequences such as eutrophication, shallowing, and even complete desiccation of water bodies.

In recent decades, the region has witnessed a decline in fish stocks and a sharp reduction in populations of commercially valuable fish species, primarily due to the combined impacts of climate change and anthropogenic pressure. Microorganisms are a vital component of aquatic biota, forming the base of the food web and serving as a primary source of biomass for higher trophic levels.

The study of aquatic microbiomes using metagenomic technologies has a long history. One of the pioneering studies in this field involved the microbiome of the Sargasso Sea, conducted even before the advent of high-throughput sequencing. Today, metagenomic research has significantly expanded, with modern technologies such as Illumina sequencing—capable of reading short DNA fragments up to 600 nucleotides in paired-end mode—and Oxford Nanopore sequencing, which allows for ultra-long reads up to 1 million nucleotides [1, 2].

Current metagenomic workflows include DNA extraction, library preparation, sequencing, and downstream bioinformatic analysis. These protocols have evolved into highly standardized and efficient methodologies. Key research directions include amplicon sequencing of taxonomically informative genes (e.g., 16S rRNA, 18S rRNA, ITS) for detailed profiling of prokaryotic and eukaryotic diversity, as well as whole-metagenome shotgun sequencing for functional characterization of microbial communities [3].

Particularly valuable are studies that aim to correlate microbial community dynamics with environmental variables, highlighting the potential of aquatic microbiomes as bioindicators. These indicators can be used in environmental assessment, detection of key functional taxa involved in processes such as pollutant degradation, and ecosystem health monitoring [4, 5, 6, 7].

Due to the cosmopolitan distribution of many microbial taxa, findings from such studies can be extrapolated across diverse geographic regions. The high metabolic activity of microbial networks supports self-purification processes in aquatic environments and offers potential for targeted biotechnological applications. Recognition of the microbiome's productive capacity, which often surpasses that of macro-organisms, also prompts a reconsideration of how trophic status is assessed in aquatic systems [8, 9, 10].

These considerations underscore the importance of developing new analytical frameworks for ecological data and their practical implementation.

Objective of the Study:

To conduct a metagenomic analysis of microbiomes in inland lakes of Western Kazakhstan, evaluating their potential use as bioindicators of environmental conditions.

Research Tasks:

- Collection and preservation of lake water samples;
- DNA extraction from lake water samples;
- Construction of amplicon libraries targeting the 16S rRNA gene;
- High-throughput sequencing using the Illumina platform;
- Analysis of alpha and beta diversity, as well as the taxonomic structure of lake microbiomes.

Material and research methods

For metagenomic analysis, microbial communities were sampled from Lakes Karasu, Glubinnoye, and Brusyanoe. Samples consisted of membrane filters containing organic sediment obtained through filtration of lake water.

DNA extraction was performed by mechanically disrupting filters (previously cut into four pieces) using glass beads (0.1 mm and 0.5 mm diameter) on a Precellys homogenizer (6000 rpm, 30 seconds) in 1 mL of CTAB extraction buffer (1×: 0.05 M Tris-HCl, pH 8.0; 0.7 M NaCl; 0.01 M EDTA; 1% CTAB), supplemented with 500 μL of chloroform (chloroform:isoamyl alcohol, 24:1). Chloroform extraction was carried out twice. DNA was precipitated using isopropanol (0.7 volume), washed with 70% ethanol, air-dried, and resuspended in 20 μL of water.

Purified DNA was used for PCR amplification of the V4 region of the 16S rRNA gene using universal primers F515 (GTGCCAGCMGCCGCGGTAA) and R806 (GGACTACVSGGGTATCTAAT) (Bates et al., 2010), along with Illumina adapters and unique barcode sequences. PCR was carried out in a 15 μL reaction mixture containing 0.5–1 U of Q5® High-Fidelity DNA Polymerase (NEB, USA), 5 pmol of each primer, 1–10 ng of DNA template, and 2 nM of each dNTP (LifeTechnologies). Thermal cycling conditions were: initial denaturation at 94 °C for 1 min, followed by 35 cycles of 94 °C for 30 s, 50 °C for 30 s, and 72 °C for 30 s, with a final elongation at 72 °C for 3 min. PCR products were purified using the AMPure XP system (Beckman Coulter, USA). Library preparation and sequencing were performed according to the manufacturer's protocols using the Illumina MiSeq platform with the MiSeq® Reagent Kit v3 (600-cycle paired-end reads, 2×300 bp).

Initial data processing, including demultiplexing and adapter trimming, was performed using Illumina software. Subsequent steps, including denoising, read merging, chimera removal, and generation of amplicon sequence variants (ASVs), as well as taxonomic classification based on the SILVA 16S rRNA reference database (release 132; Quast et al., 2013), were conducted using the **dada2**, **phyloseq**, and **DECIPHER** packages in the R programming environment.

Taxonomic visualization and summaries were generated using the QIIME software package (Caporaso et al., 2010). Alpha and beta diversity analyses were performed using QIIME2 (Bolyen et al., 2019). For normalization, rarefaction was applied to a depth of 8031 reads per sample. Alpha diversity was assessed using several metrics: **observed_ASVs**, **Chao1**, **Simpson's evenness index**, and **Shannon index**, the latter of which accounts for both richness and evenness. Beta diversity was estimated using both weighted and unweighted **UniFrac** distance metrics. A phylogenetic tree was constructed using the SEPP algorithm (fragment insertion of representative sequences into a reference tree) with the SILVA database, implemented via the *fragment-insertion sepp* plugin in QIIME2.

Research results

In this study, 36 amplicon libraries were sequenced, each yielding at least 20,000 reads. Initial taxonomic analysis identified over 4,000 taxa belonging to approximately 400 genera. The dominant groups included unclassified bacteria as well as representatives of the phyla Proteobacteria, Bacteroidota, and Actinobacteriota. Differences among the microbiomes of the three lakes were mainly reflected in the

proportion of unassigned bacteria. However, at the lower taxonomic level of prokaryotic genera, it became clear that each lake has a highly specific taxonomic composition.

An alpha diversity analysis was also performed, assessing species richness (observed ASVs, chao1), phylogenetic diversity (Faith's PD), community evenness (Simpson), and the Shannon index, which accounts for both richness and evenness. The results showed that the highest species richness was found in microbiome samples from Lakes Glubinnoye and Brusyanoye, while the lowest richness was observed in samples from Lake Karasu (see Table 1).

Table 1. Alpha diversity indices characterizing species richness (observed ASVs, chao1, Shannon), phylogenetic diversity (Faith's PD), and evenness (Simpson, Shannon).

Lake	observed_ASVs	chao1	Faith's PD	Shannon	Simpson
Karasu	607.4	641.06	190.36	6.87	0.979
Brusyanoye	699.0	713.2	230.8	7.1	0.976
Glubinnoye	876.7	898.3	285.2	6.8	0.974

A similar pattern was observed for phylogenetic diversity, with the highest values in Glubinnoye samples, followed by Brusyanoye and Karasu. According to the Shannon index, Brusyanoye samples had the richest and most even communities, followed by Glubinnoye and Karasu. The Simpson index indicated that the most even communities were in Glubinnoye and Brusyanoye samples.

Overall, the highest species richness (including phylogenetic diversity) and evenness (Simpson index) were observed in samples from Lakes Glubinnoye and Brusyanoye. Samples from Lake Karasu showed the lowest species and phylogenetic diversity as well as lower evenness, as indicated by the Shannon index.

The taxonomic composition of microbiomes from the different lakes shared common features. All samples contained bacterial groups such as Verrucomicrobiae, Gamma- and Alphaproteobacteria, Cyanobacteria, Bacteroidia, Actinobacteria, and Acidimicrobiia. However, comparative analysis revealed differences already at the class level. Glubinnoye samples were characterized by a higher abundance of Acidimicrobiia (2.6–3.8%), a low proportion of Cyanobacteria (0.8–1.3%), and a relatively small fraction of Bacilli (0.5–0.7%) (see Figure 1).

Karasu samples showed an increased proportion of Cyanobacteria (8.2–13.18%), Alphaproteobacteria (7.6–12.1%), and Verrucomicrobiae (3.2–5.9%), presence of Rhodothermia representatives (3.3–5.5%), and a decrease in Bacteroidia (5.5–6.7%).

Brusyanoye samples were marked by a high proportion of Cyanobacteria (11.3%), the presence of unclassified Actinobacteriota (about 2%, not seen in other samples), a large fraction of Actinobacteria (21.3–22.9%), increased Acidimicrobiia (2–2.4%), appearance of Clostridia (3.7–4.2%) and Chlorobia (2.1–2.5%, though not in all replicates), a decreased proportion of Alphaproteobacteria (1.7–2.7%), and broad representation of Bacteroidia (16–30.3%). Bacilli were also present, comprising 1–3.4% of the community.

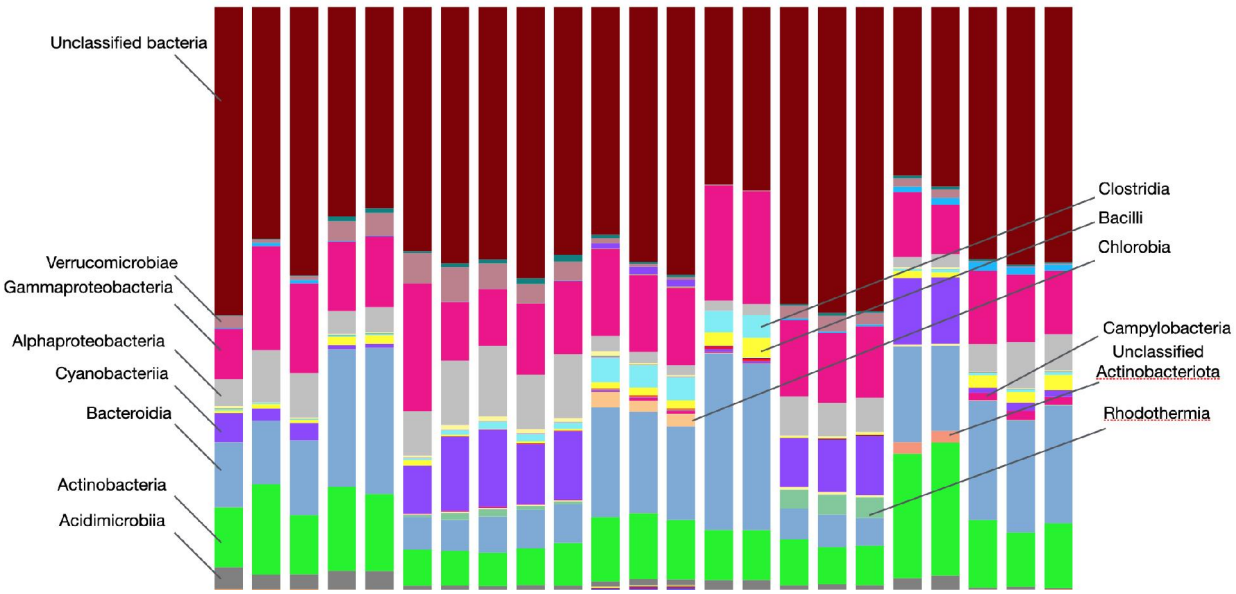

Figure 1 – Differences in the taxonomic composition of communities at the class level.

Figure 2 shows that differences in taxonomic composition are also evident at a lower taxonomic level (the genus level is illustrated for clarity). Based on these data, we can identify the dominant taxa for each community.

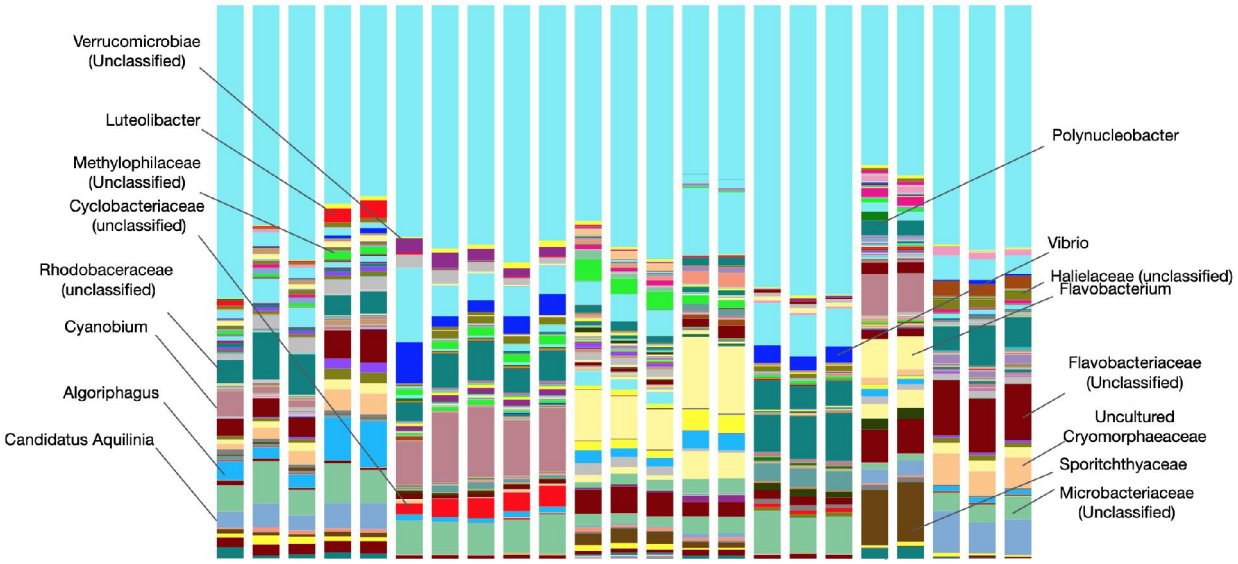

Figure 2 – Differences in the taxonomic composition of communities at the genus level.

In the microbiome samples from Lake Glubinnoye, dominant groups include unclassified members of the Rhodobacteraceae family (3.6–8.5%), the genus Algoriphagus (2–8%), Microbacteriaceae family (4.7–7.6%), and unclassified members of the Flavoracteriaceae family (3.3–6%). There is also notable dominance of unclassified Candidatus Aquiluna (5.5–7.5%), uncultivable representatives of Cryomorphaeaceae (4.5–5.7%), and unclassified Gammaproteobacteria (3.3–4.4%). A substantial portion of the community is comprised of unclassified representatives of the kingdom Bacteria (35–53%).

In samples from Lake Karasu, the dominant taxa, besides unclassified Bacteria (44–46.5%), include bacteria such as Synechococcus (6.7–8.7%), unclassified Gammaproteobacteria (5.1–13.4%), Cyanobium (8–12.6%), Microbacteriaceae (6–7%), Vibrio (1.9–7.4%), and unclassified Rhodobacteraceae (3.3–7%).

For Lake Brusyanoye, dominance is observed in unclassified Bacteria (44–46%), Sporichthyaceae hgI clade (10–10.7%), Cyanobium PCC-6307 (about 7%), Flavobacterium (6–7%), and a group of unclassified Actinobacteria (around 6%).

Beta-diversity analysis showed that the taxonomic composition of microbial communities varies depending on the sampling site (see Figure 3). According to the weighted UniFrac metric, which reflects the qualitative composition of the community (presence or absence of taxa), similarity in taxonomic composition is found only in the microbiome of Lake Karasu. When comparing quantitative composition (weighted UniFrac), similarity in taxonomic structure is observed in samples from Lake Glubinnoye, whereas the microbiomes of samples from Lake Karasu, although clustered together, show considerable divergence.

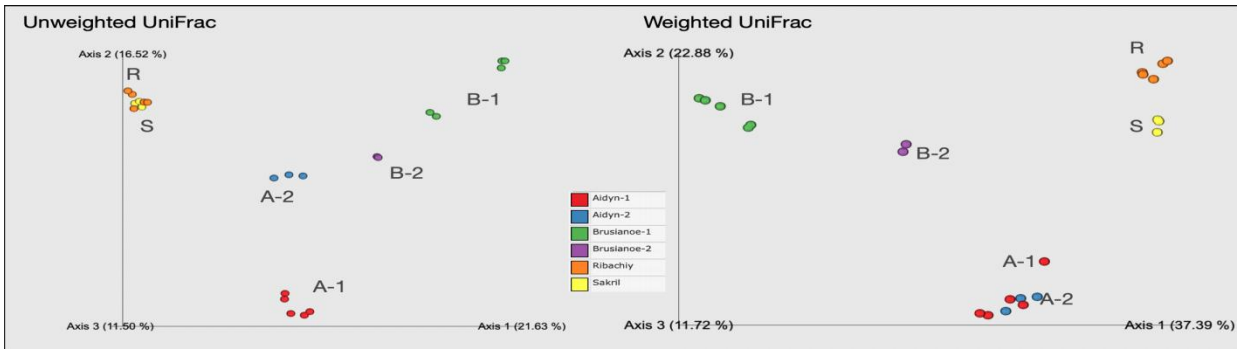

Figure 4 – Beta diversity, weighted and unweighted UniFrac.

The results of the comparative analysis of community structures indicate that water samples from different locations exhibit distinct uniqueness; however, groups of samples with similar compositions can be identified, including those from Lakes Glubinnoye and Karasu.

Conclusion

During the metagenomic analysis, dominant groups were identified, including unclassified bacteria and representatives of the phyla Proteobacteria, Bacteroidota, and

Actinobacteriota. Clear differences were observed among the three water bodies: Karasu, Glubinnoye, and Brusyanoe.

Distinctive features can also be highlighted for the studied communities. For example, samples from Lake Karasu are characterized by a high abundance of cyanobacteria, predominantly from the genera Cyanobium and Synechococcus. A similar pattern is seen in samples from Lakes Glubinnoye and Brusyanoe. Among the Bacteroidota representatives, Algoriphagus is more typical for Glubinnoye, while Flavobacterium dominates in Brusyanoe samples.

Distinctive groups for Glubinnoye samples include representatives of Luteolibacter; for Karasu samples, Cyclobacteriaceae (about 4% of the community) and an unclassified group of Verrucomicrobiae are notable; these are also characteristic for Brusyanoe samples.

Additionally, the microbiomes of Lakes Glubinnoye and Brusyanoe commonly contain Candidatus Aquilina (ranging from 2.6 to 7.5%) and relatively high levels of Algoriphagus (2 to 8%). Among the shared traits between Glubinnoye and Karasu samples is a relatively high presence of Methylophilaceae (1.3–3.8%).

Acknowledgments

The research was conducted within the framework of the budget program 217 "Development of Science," under subprogram 102 "Grant Funding for Scientific Research," with the priority area: Ecology, Environment, and Rational Nature Management, and the sub-priority: Fundamental and Applied Research in Ecology, Environment, and Rational Nature Management. This was part of the project AP23483830 "Ecohydrological Determinants of the Functioning of Lakes in Western Kazakhstan from the Perspective of the Foresight Concept."

REFERENCES

[1] Sergaliyev N.Kh. et al., The structure of the natural microbiome of sturgeon / N.Kh. Sergaliyev, M.G. Kakishev, N.S. Ginayatov, E.E. Andronov, A.G. Pinaev // International Journal of Engineering and Advanced Technology (IJEAT), 2019. ISSN: 2249-8958, – Vol.-9. Issue-2. – P. 166-174.

[2] Venter J.C., Remington K., Heidelberg J.F., et al. Environmental genome shotgun sequencing of the Sargasso Sea. Science. 2004;304(5667):66-74. doi:10.1126/science.1093857

[3] Jo J., Oh J., Park C. Microbial community analysis using high-throughput sequencing technology: a beginner's guide for microbiologists. J Microbiol. 2020;58(3):176-192. doi:10.1007/s12275-020-9525-5

[4] Gerhard W.A., Gunsch C.K. Microbiome composition and implications for ballast water classification using machine learning. Sci Total Environ. 2019;691:810-818. doi:10.1016/j.scitotenv.2019.07.053

[5] Langenheder S., Lindström E.S. Factors influencing aquatic and terrestrial bacterial community assembly. Environ Microbiol Rep. 2019;11(3):306-315. doi:10.1111/1758-2229.12731

[6] Garcia C.A., Hagstrom G.I., Larkin A.A., et al. Linking regional shifts in microbial genome adaptation with surface ocean biogeochemistry. Philos Trans R Soc Lond B Biol Sci. 2020;375(1798):20190254. doi:10.1098/rstb.2019.0254

[7] Fernandes G.L., Shenoy B.D., Damare S.R. Diversity of bacterial community in the oxygen minimum zones of Arabian Sea and Bay of Bengal as deduced by Illumina sequencing. Front Microbiol. 2020;10:3153. Published 2020 Jan 21. doi:10.3389/fmicb.2019.03153

[8] Vorobev A., Dupouy M., Carradec Q., et al. Transcriptome reconstruction and functional analysis of eukaryotic marine plankton communities via high-throughput metagenomics and metatranscriptomics. Genome Res. 2020;30(4):647-659. doi:10.1101/gr.253070.119

[9] Nurlan Khabibullovich Sergaliyev, Murat Galikhanovich Kakishev, Nurbek Satkanuly Ginayatov, Evgeny Evgenievich Andronov, Alexander Georgievich Pinaev. Examination of the microbiome of bastard sturgeon cultivated in the conditions of recirculated water. International Journal of Engineering and Advanced Technology (IJEAT), №9:1, 2019, pp. 4536–4542.

[10] Nurlan Khabibullovich Sergaliev, Murat Galikhanovich Kakishev, Nurbek Satkanuly Ginayatov, Farida Khamidullievna Nurzhanova, Evgeny Evgenievich Andronov. Microbiome structure in a recirculating aquaculture system and its connection to infections in sturgeon fish. Veterinary World, №14, 2021, pp. 166–174.

Серғалиев Н.Х., Қакишев М.Г.

БОЛАШАҚ КОНЦЕПЦИЯСЫНЫҢ ҚАРАСТАУЫНАН БАТЫС ҚАЗАҚСТАН ӨҢІРІНДЕГІ КӨЛДЕРДІҢ ФУНКЦИЯЛЫҚ ЖҰМЫСЫНЫҢ МИКРОБИОЛОГИЯЛЫҚ ДЕТЕРМИНАНТТАРЫ

Аннотация. Мақалада Батыс Қазақстан облысындағы Глубинное, Брусяное және Карасу көлдерінің микробиомын жоғары өткізу қабілетті секвенирлеу әдісімен зерттеу нәтижелері берілген. Зерттелген су айдындарында классификацияланбаған бактериялармен қатар Proteobacteria, Bacteroidota және Actinobacteriota филумдарына жататын өкілдердің басым екендігі анықталды. Үш көлдің микробиомалары арасындағы негізгі айырмашылықтар классификацияланбаған бактериялардың үлесіне қатысты. Сондай-ақ, әртүрлі су айдындарының микробтық қауымдастықтарының таксономиялық құрылымында ұқсас белгілер бар екені белгілі болды. Барлық үлгілерде Verrucomicrobiae, Gamma- және Alphaproteobacteria, Cyanobacteria, Bacteroidia, Actinobacteria және Acidimicrobiia топтары анықталды. Бета-әртүрлілік талдауы микробиомалардың құрамы үлгі алу орнына байланысты екенін көрсетті. Таксономиялық құрамды бағалау барысында Глубинное көлінің микробиомалары бір-біріне көбірек ұқсас болса, Карасу көлінің микробтық қауымдастықтары, бір кластерге кірсе де, әртүрлілігі жоғары екені анықталды. Глубинное және Брусяное көлдерінің үлгілерінде Candidatus Aquilinia (2,6-дан 7,5%-ға дейін) және салыстырмалы түрде жоғары Algoriphagus мөлшері (2-ден 8%-ға дейін) анықталды. Сонымен қатар, Глубинное және Карасу көлдерінің үлгілерінде Methylophilaceae отбасына жататын микроорганизмдердің айтарлықтай болуы (1,3-тен 3,8%-ға дейін) ортақ белгі ретінде атап өтілді.

Кілт сөздер: микробиология; метагеномика; микробиом; көлдер; таксономиялық құрам.

Сергалиев Н.Х., Какишев М.Г.

МИКРОБИОЛОГИЧЕСКИЕ ДЕТЕРМИНАНТЫ ФУНКЦИОНИРОВАНИЯ ОЗЕР ЗАПАДНОГО КАЗАХСТАНА С ТОЧКИ ЗРЕНИЯ КОНЦЕПЦИИ ФОРСАЙТА

Аннотация. В статье представлены результаты исследований микробиома озёр Глубинное, Брусяное и Карасу, расположенных в Западно-Казахстанской области, выполненных с использованием метода высокопроизводительного секвенирования. Установлено, что в исследуемых водоёмах преобладают неклассифицированные бактерии, а также представители филумов Proteobacteria, Bacteroidota и Actinobacteriota. Основные различия между микробиомами трёх озёр касаются доли неатрибутированных бактерий. Также выявлено, что таксономическая структура микробных сообществ различных водоёмов имеет сходные черты. Во всех пробах обнаружены такие группы, как Verrucomicrobiae, Gamma- и Alphaproteobacteria, Cyanobacteria, Bacteroidia, Actinobacteria и Acidimicrobiia. Анализ бета-разнообразия показал, что состав микробиомов зависит от места отбора проб. При оценке таксономического состава было выявлено, что микробиомы из озера Глубинное обладают большей схожестью между собой, в то время как микробные сообщества из озера Карасу, несмотря на попадание в один кластер, демонстрируют большее разнообразие. В пробах из озёр Глубинное и Брусяное обнаружено присутствие Candidatus Aquilinia (от 2,6 до 7,5%) и относительно высокое содержание Algoriphagus (от 2 до 8%). Кроме того, как общая черта проб из озёр Глубинное и Карасу отмечается значительное присутствие семейства Methylophilaceae (от 1,3 до 3,8%).

Ключевые слова: микробиология; метагеномика; микробиом; озера; таксономический состав.