

# The effect of sample volume and number of reads on the sequencing result of a microeukaryotic community from an oligotrophic lake

Bashenkhaeva M.V.\*

Limnological Institute Siberian Branch of the Russian Academy of Sciences, Ulan-Batorskaya Str., 3, Irkutsk, 664033, Russia

**ABSTRACT.** Microeukaryotes are an important component in aquatic ecosystems and can be used as an indicator of environmental conditions. Changes in their abundance or diversity can indicate changes in the habitat. In the last decade to addition to light microscopy method, DNA metabarcoding is used a monitoring of microeukaryote community. In order to optimise the metabarcoding method, it is important to determine the necessary sample volume and sequencing depth. This study compared samples with different filtered water volumes (50, 10 and 1.2 L) and sequencing depths to determine their impact on the diversity and taxonomic composition of the unicellular microeukaryotic community in Lake Baikal. ANOSIM analysis revealed similarity among samples with different filtered volumes. This suggests that a small filtered sample volume (1.2 L) is sufficient to reveal a high diversity of taxa in an oligotrophic reservoir, comparable to that obtained with a larger filter volume (50 L). Ciliophora species were detected the highest number of reads by metabarcoding at the large volume of filtered sample (50 L), while diatoms and green algae were better read in smaller volumes (10 and 1.2 L). It seems that samples volumes in 10 and 1.2 L are sufficient for microalgae monitoring. Samples with different sequencing depths produced similar results in the characterisation of dominant species. Differences were observed in minor taxa, some of which were only detected with a higher number of reads. These results are important for improving the monitoring of oligotrophic bodies using metabarcoding.

**Keywords:** diversity, sample volume, sequencing depth, metabarcoding, 18S rRNA, Lake Baikal

**For citation:** Bashenkhaeva M.V. The effect of sample volume and number of reads on the sequencing result of a microeukaryotic community from an oligotrophic lake // Limnology and Freshwater Biology. 2025. - № 5. - P. 1212-1226. DOI: [10.31951/2658-3518-2025-A-5-1212](https://doi.org/10.31951/2658-3518-2025-A-5-1212)

## 1. Introduction

Microeukaryotes, a group of unicellular organisms including microalgae, protozoa and microscopic fungi, are an important component of aquatic ecosystems. They participate in the global cycles of carbon (C), silicon (Si), and other nutrients, forming the basis of aquatic food webs (Falkowski et al., 2003; Calbet and Landry, 2004; Williams et al., 2008; Fuhrman, 2009). The structure of microeukaryotic communities is influenced by environmental changes such as temperature, nutrient content, salinity and light. Therefore, changes in their abundance and species composition can serve as early warning signs of disturbances to the ecosystem associated with pollution or climate change (Winder and Sommer, 2012). The classic method of monitoring unicellular microeukaryotes is microscopy, which is based on morphological features and requires specialist knowledge. It can also be complicated by the small size of the cells and the morphological variability of the

species. Over the past decade, metabarcoding has been employed for biomonitoring. This approach involves analysing the diversity of marker gene sequences in the total DNA of a community (Zimmermann et al., 2015; Abad et al., 2016; Andersson et al., 2023; Gelis et al., 2024; Mikhailov et al., 2025). Given the multi-stage nature of the metabarcoding method, it is important to select optimal conditions for each stage. The expected result can be influenced by various factors: 1) the volume of the filtered sample; 2) the selected DNA marker; 3) the methods used at different stages of molecular analysis (e.g. DNA extraction methods and sequencing technologies), and 4) the approaches to bioinformatics analysis (e.g. primary data processing methods, clustering algorithms, and taxonomic classification methods). For the purposes of biomonitoring, these stages must be standardised. Many studies have been conducted on the selection of marker genes for metabarcoding (Gran-Stadniczenko et al., 2017; Casey et al., 2021; Ficetola et al., 2021; Bukin et al., 2023; Kezlya et al., 2023), the

\*Corresponding author.

E-mail address: [maria.bashenkhaeva@gmail.com](mailto:maria.bashenkhaeva@gmail.com) (M.V. Bashenkhaeva)

**Received:** August 29, 2025; **Accepted:** October 27, 2025;

**Available online:** October 31, 2025

© Author(s) 2025. This work is distributed under the Creative Commons Attribution-NonCommercial 4.0 International License.



comparison of different DNA extraction methods (Mäki et al., 2017; Vasselon et al., 2017), and the optimisation of high-throughput sequencing data processing (Prodan et al., 2020; Abellan-Schneyder et al., 2021; Brandt et al., 2021). Sample volume also affects the final results, including the amount of extracted DNA and the quality of the samples obtained for sequencing. The choice of sample volume for filtration depends on the type of water body. Marine ecosystems are rich in phytoplankton and zooplankton; to capture the full diversity of organisms, large sample volumes of 50–100 litres are recommended (Govindarajan et al., 2022). In contrast, shallow mesotrophic estuaries require only a small volume of water (25–500 ml) for biodiversity assessment using metabarcoding (Andersson et al., 2023). In freshwater bodies, the required sample volume varies depending on the trophic status of the body of water and the cell concentration of various microorganisms.

Lake Baikal is an oligotrophic freshwater body with low nutrient and primary production concentrations (Votintsev et al., 1975). This study aimed to estimate the required sample volume and sequencing depth for determining microeukaryotic diversity using metabarcoding of 18S rRNA gene fragments.

## 2. Materials and methods

### 2.1. Field Sampling

Sampling was carried out at Station, 3 km from Maritui village (51°45.546'N; 104°13.222'E) in the southern basin of Lake Baikal, on 18 July 2023. Samples were collected from the research vessels “G.Yu. Vereshchagin” using an SBE 32 Carousel water sampler (Sea-Bird Electronics, USA), from depths of 0, 5, 10, 15, 20 and 25 m.

Phytoplankton were analysed in parallel using microscopy and DNA sequencing. Analysis was performed on integrated water samples (an equal volume of water from the different layers was combined into one sample). To quantify and identify phytoplankton using light microscopy, 1.2 L of each integrated sample was filtered through a 3 µm REATREK filter (Obninsk-3, Russia) and then fixed in 50 mL of formaldehyde solution to achieve a final concentration of 3.7%. Microalgae cells were counted using an Axiostar Plus microscope (Zeiss, Oberkochen, Germany) in two replicates, as previously described (Firsova et al., 2023).

### 2.2. DNA extraction, amplification and high-throughput sequencing (HTS)

For the DNA analysis, 1.2, 10 and 50 L of the integrated samples were pre-filtered using 100 µm nylon mesh to remove zooplankton, and then filtered through 3 µm pore-size filters (REATREK Filter, Obninsk-3, Russia). These volumes were chosen to represent those obtained by microscopy, which we typically use to assess biodiversity using DNA metabarcoding, and those used for shotgun sequencing. Filtration was carried out immediately after sampling. The biomass on the filters was washed into sterile bottles with 10 ml of sterile TE buffer (1 mM EDTA and 10 mM Tris-HCl; pH

7.5), after which it was stored at -20 °C, then -80 °C, until further analysis. Total DNA was extracted from the samples using lysozyme (1 mg/mL), proteinase K, 10% SDS, and a phenol:chloroform:isoamyl alcohol mixture (25:24:1) (Bukin et al., 2023). Amplification and sequencing of amplicons was performed on the Illumina MiSeq platform at the Genomic, Proteomic and Cell Biology Department of ARRIAM (Saint Petersburg, Russia). To compare samples of different volumes, universal primers targeting the V8–V9 region of the 18S rRNA gene were used. Sequencing was performed with different numbers of reads per sample: the most commonly used number was approximately 20,000 reads, and an increased number was approximately 50,000 reads.

### 2.3. Dataset Preprocessing

Sequencing data were analyzed using DADA2 v1.16 (the dada2 R package). Paired-end reads were filtered and merged, chimeric and short sequences were removed, and amplicon sequence variants (ASV) were generated (Callahan et al., 2016). Rarefaction curves, richness, and diversity indices (Chao1, Shannon, and Simpson) were calculated using R version 4.4.3. The ANOSIM test was used to compare samples with different volume and different number of reads. Analysis was performed in PAST 4.06b. Taxonomic identification was performed using the Silva 132 taxonomy (Bremen, Germany). The heatmap is based on the 80 most numerous ASVs in R. The pairwise distance matrix computed with the Bray-Curtis dissimilarity index was used for clustering community profiles by UPGMA in heatmap using vegan (Oksanen et al., 2019) and pheatmap (Kolde, 2019) packages.

Sequence data were uploaded to the Sequence Read Archive database (<https://www.ncbi.nlm.nih.gov/sra/>) of the National Center for Biotechnology Information (NCBI) under the accession number PRJNA1348943.

## 3. Results and Discussion

To assess the diversity and taxonomic composition of microeukaryotes in samples with different filtered volumes and sequencing depths, 972,147 reads were analysed. Samples with lower sequencing depths yielded 34,931–52,392 reads per sample, while samples with higher sequencing depths yielded 90,721–137,764 reads (Table). After quality filtering of the obtained reads, which included removing primer regions and sequences with fragment lengths of less than 250 bp, merging paired reads and removing short and chimeric sequences, 25,328–103,680 reads remained per sample. After removing sequences related to Metazoa, more than half of the reads in almost all samples were removed. Despite the mechanical removal of large eukaryotic DNA during sample filtering, it still makes its way into the samples and can account for the majority of reads. During ASV generation, a total of 67–121 ASVs were obtained per sample, with an average length of 331 bp.

**Table.** Sequencing metrics and biodiversity indices in different samples analysed by metabarcoding (50, 10 and 1 L of filtered water).

|                                                     | 50L_1 rep | 50L_2 rep | 10L_1 rep | 10L_2 rep | 1L_1 rep | 1L_2 rep | 50L_1 rep* | 50L_2 rep* | 10L_1 rep* | 10L_2 rep* | 1L_1 rep* | 1L_2 rep* |
|-----------------------------------------------------|-----------|-----------|-----------|-----------|----------|----------|------------|------------|------------|------------|-----------|-----------|
| Initial number of reads                             | 40,419    | 34,931    | 52,392    | 51,857    | 48,786   | 43,390   | 109,918    | 90,721     | 135,541    | 137,764    | 126,154   | 100,274   |
| Reads after filtration                              | 38,565    | 33,164    | 49,811    | 49,416    | 46,481   | 41,413   | 104,974    | 86,536     | 129,633    | 131,377    | 119,792   | 95,645    |
| Reads after removal of chimeric and short sequences | 26,993    | 25,328    | 39,534    | 37,841    | 38,886   | 36,945   | 72,053     | 64,074     | 103,680    | 100,215    | 98,822    | 85,129    |
| Reads after removed Metazoa sequences               | 11,709    | 8,084     | 18,553    | 19,953    | 12,361   | 13,381   | 33,241     | 23,400     | 51,697     | 55,129     | 35,839    | 30,087    |
| ASV                                                 | 73        | 67        | 84        | 87        | 78       | 76       | 107        | 95         | 112        | 121        | 105       | 105       |
| Simpson-index                                       | 73        | 67        | 84        | 87        | 78       | 76       | 107        | 95         | 112        | 121        | 105       | 105       |
| Shannon-index                                       | 2.7       | 2.72      | 2.85      | 2.88      | 2.93     | 3.1      | 2.82       | 2.76       | 2.95       | 2.96       | 3.01      | 3.16      |
| Chao1-index                                         | 0.84      | 0.86      | 0.87      | 0.88      | 0.90     | 0.91     | 0.85       | 0.86       | 0.88       | 0.88       | 0.9       | 0.92      |

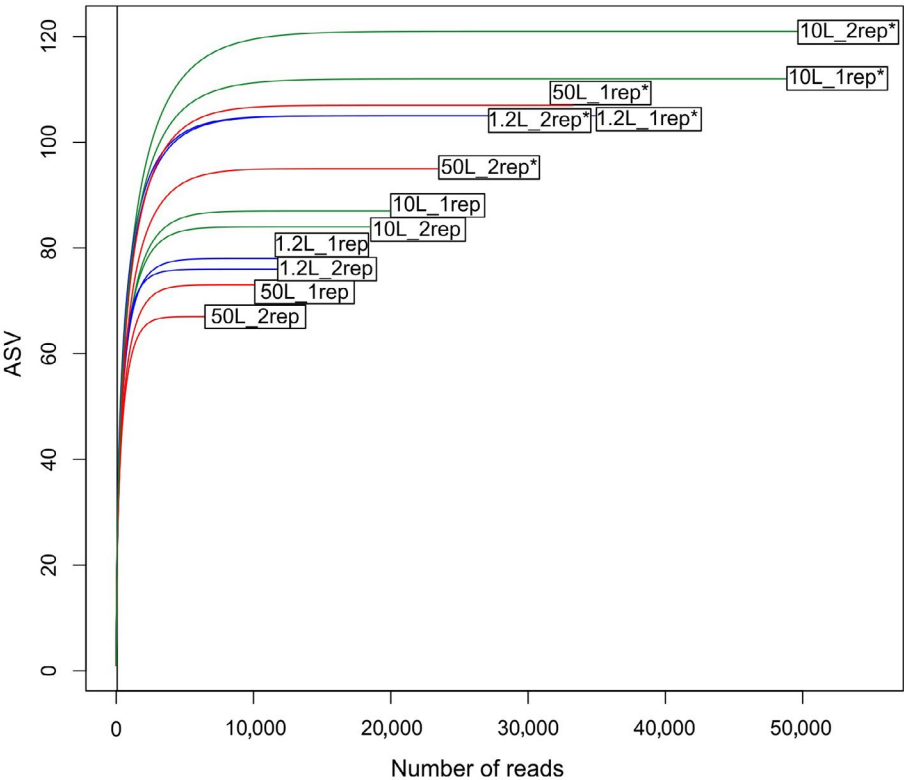
**Note:** samples with an increased number of reads are marked with an asterisk (\*).

3.1. Richness and Diversity Assessment

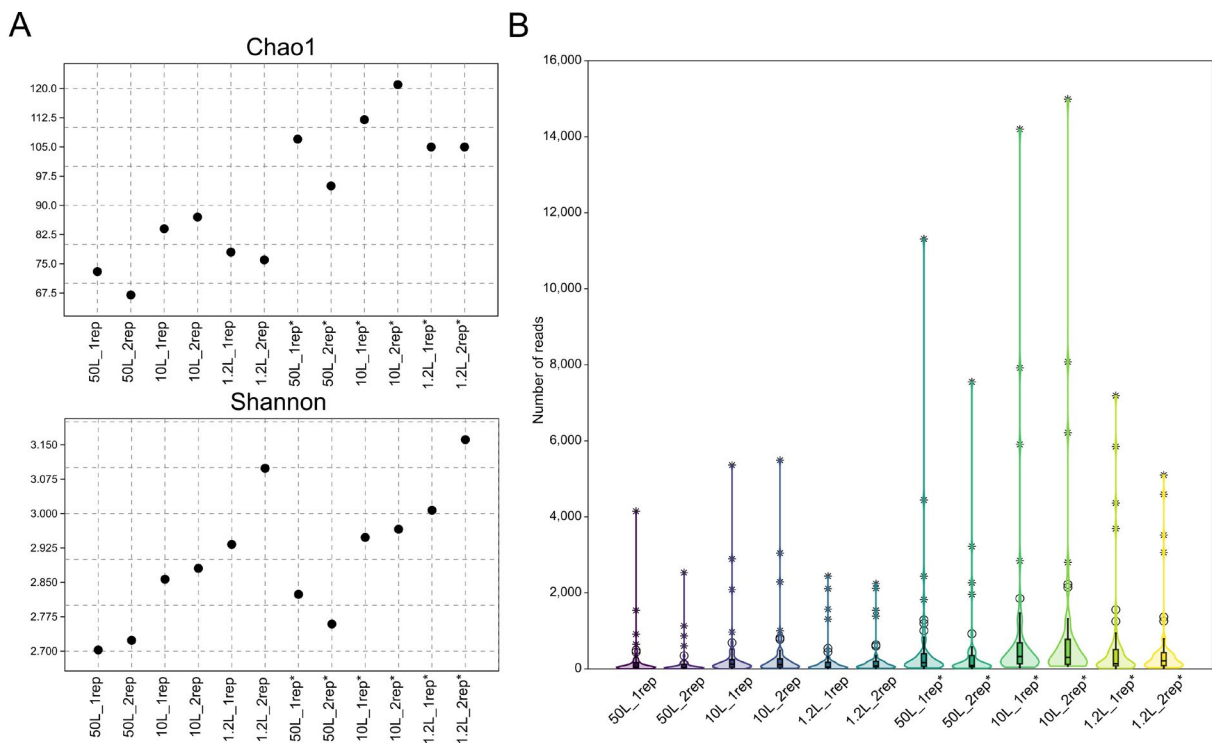
Alpha diversity analysis revealed similar ASV abundances across filtration volumes of 1.2, 10 and 50 L (Table). Rarefaction curves showing ASV numbers versus sample sequence depth indicated that the sequencing depth in our study was sufficient to estimate species richness (Fig. 1). The Chao1 index ranged from 67 to 121, with the highest values observed in samples with an increased number of reads. This suggests a correlation between species richness and sequencing depth, enabling the detection of rare species (Zaheer et al., 2018; Bardenhorst et al., 2022; Schmitz and Rahmann, 2025). For both sequencing depths, the highest richness was observed in the 10 L samples (Fig. 2A; Table). This result suggests that this volume may be the most

optimal of those selected. Filtration, which concentrates cells from the sample onto the filter, is faster and results in better sample preservation compared to 50 L. More species are also captured in the sample compared to 1.2 L.

The ANOSIM test showed slight differences in composition between samples with different filtered volumes ( $R=0.2384$ ,  $p=0.0482$ ). This demonstrates the sensitivity of the metabarcoding method: the small filter volume (1.2 L) revealed a taxon diversity comparable to that of the larger volume (50 L). This result was similar with studies examining samples from the mesotrophic Øre River estuary, which compared volumes of 0.25, 0.5, and 1.2 L (Andersson et al., 2023). In their study, diversity was detected even at the lowest sample volume examined – 0.25 L.



**Fig.1.** Rarefaction curves for samples with different volumes of filtered water (50, 10 and 1.2 L) and different sequencing depths. Samples with an increased number of reads are marked with an asterisk (\*).



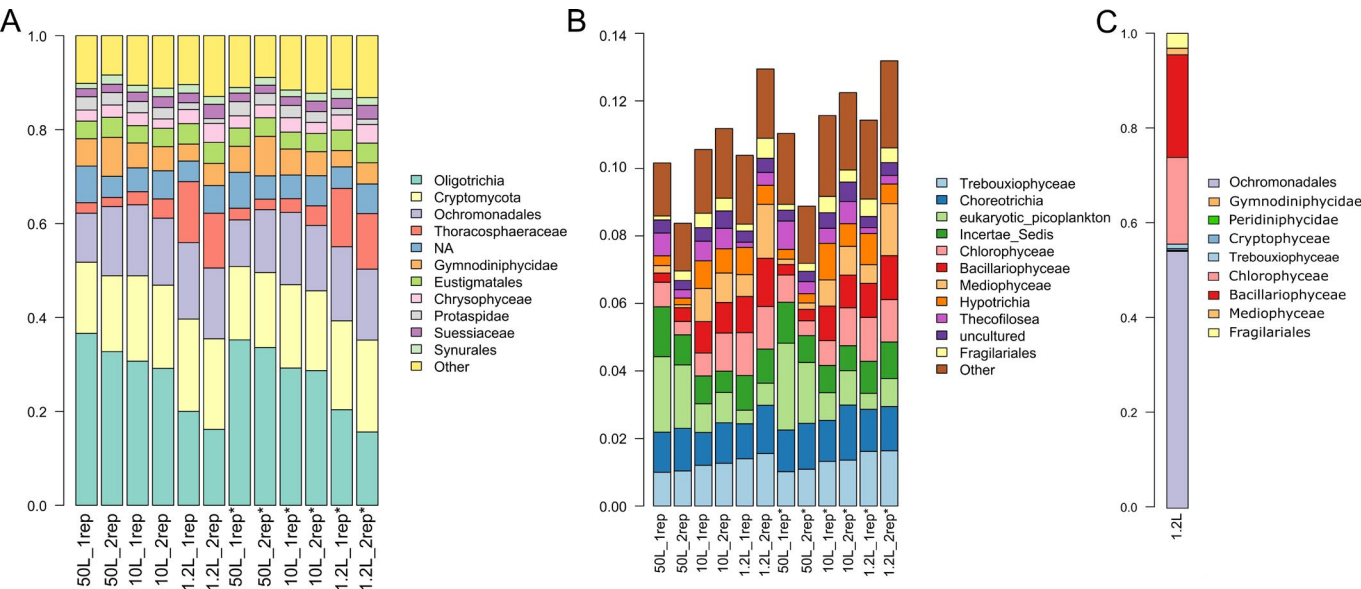
**Fig.2.** Diversity and richness indices for samples with different filtered volumes (50, 10 and 1.2 L) and sequencing depths (20,000 and 50,000 reads). (A) Chao1 and Shannon indices; (B) distribution of the 50 most abundant ASVs. Samples with an increased number of reads are marked with an asterisk (\*).

The Shannon and Simpson diversity indices were similar between samples with different sequencing depths (Fig. 2A and Table). The Shannon index ranged from 2.7 to 3.1 with 20,000 reads and from 2.76 to 3.16 with 50,000 reads. The Simpson index varied from 0.84 to 0.91 with 20,000 reads, and from 0.85 to 0.92 with 50,000 reads. Thus, the overall community structure (i.e. relative species abundance) remains stable even with increasing sequencing depth. The highest diversity values were observed in the 1.2 L samples, which are characterised by a more uniform distribution of species within the sample. The minimum Shannon

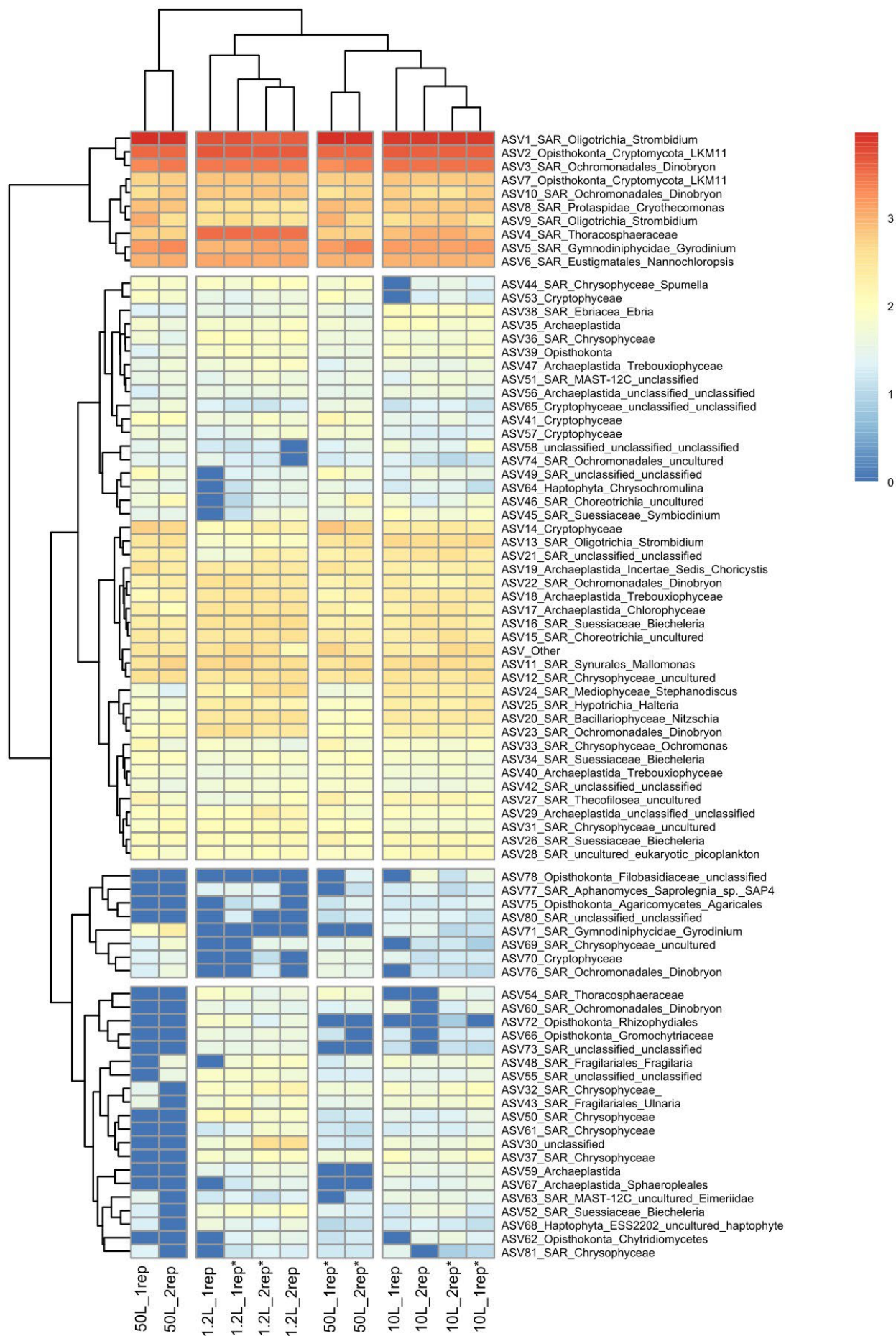
and Simpson index values were observed in the 50 L samples. Low values of these indices indicate the dominance of individual taxa in the community (Genomics, 2024). Comparing the ASV composition of the samples reveals that more than half of the reads in the 50 L samples belong to a single ASV (Fig. 2B).

**3.2. Taxonomic Structure**

To estimate the impact of volume filtration on the taxonomic composition of the microeukaryotic community, we compared the prevalence of dominant and



**Fig.3.** Taxonomic structure of samples with different filtered sample volumes (50, 10, and 1.2 L of filtered water). Family-level metabarcoding analysis of 11 dominant ASVs (A) and minor ASVs; results of microscopic analysis of the 1.2 L sample (C). Samples with an increased number of reads are marked with an asterisk (\*).



**Fig.4.** Heatmap of microeukaryotic community samples with different filtered sample volumes (50, 10, and 1.2 L of filtered water), created using a set of the 80 most common ASVs. UPGMA-assisted clustering of the Bray-Curtis pairwise distance matrices was used to produce the sample-wise and ASV-wise clustering trees. Samples with an increased number of reads are marked with an asterisk (\*).

minor ASVs at the family level (Fig. 3). Thus, ASVs were related to the Oligotrichia (Ciliophora) represented the largest share in the 50 and 10 L samples. The heatmap also shows that ASV1, which belongs to *Strombidium* genus, is dominant in the 50 and 10 L samples, and there are significant rate of the 1.2 L samples (Fig. 4). *Cryptomycota* and *Ochromonadales* are represented in roughly equal proportions in the samples (Fig. 3A), and are either subdominant or dominant. ASV3, classified as a representative of the genus *Dinobryon*, is one of the most abundant ASVs in all samples. This is comparable with the results obtained by phytoplankton microscopy (Fig. 3B), in which *Ochromonadales*, *Dinobryon cylindricum* ( $65.4 \times 10^3$  cells L<sup>-1</sup>) and *Dinobryon sociale* ( $48.2 \times 10^3$  cells L<sup>-1</sup>) also dominated in abundance. In contrast to the 50 and 10 L samples, there is a significant share of the Thoracosphaeraceae (Dinoflagellata) family in the 1.2 L samples. Dinoflagellates inhabit cold water bodies and are highly sensitive to increased temperature, which causes their rapid destruction (Kobanova, 2009). It is possible their cell were destructed and DNA degradation due to a longer filtration time of the 50 and 10 L samples. Thus, according to microscopy data, the share of dinoflagellates was small, while ones were among the 11 dominant families by metabarcoding data (Fig. 3A, 3B).

The dinoflagellate genome is multicopy and dinoflagellates often have a 'mesokaryotic' nucleus in which the chromosomes are not packaged in histones as they are in other eukaryotes. This results in a more extended state which may facilitate more frequent and easier copying of chromosomal material (Wisecaver and Hackett, 2011). As result, dinoflagellates can account for a significant share of the sequence data even at low abundance.

The share of reads of diatoms (Bacillariophyceae and Mediophyceae families, and Fragilariales order) and green algae (Trebouxiophyceae and Chlorophyceae families) were high in the 1.2 L samples. In contrast, the contribution of these microalgal groups were small in the 50 L samples (Fig. 3B), despite these groups being subdominant microscopically (Fig. 3C).

Comparing samples with different sequencing depths reveals that most taxa have a similar proportion (Fig. 3 and 4). Some differences were occurred in minor taxa, some of which are only detected with a larger number of sequences obtained (Fig. 4). ANOSIM testing confirmed differences in sample composition at different sequencing depths ( $R = 0.513$ ,  $p = 0.0029$ ). The heatmap shows that the microeukaryotic community is most fully represented in the 10 L samples with an increased sequencing depth, and it is likely optimal condition for detection of microeukaryotic.

## 4. Conclusion

It was found that the optimal approach for the monitoring of unicellular microeukaryotes in the oligotrophic Lake Baikal during the summer were: i) to collect 10 L of water; ii) perform sequencing at a depth of 50,000 reads per sample using the 18S rRNA gene metabarcoding. A relatively small water volume (1.2 L) is sufficient to characterise dominant taxa. In the largest

samples (50 L), over half of all sequences belonged to Ciliophora, while the number of reads belonging to diatoms and green algae was underestimated. This makes collecting such a large sample volume unjustifiable for phytoplankton research. Increasing the sequencing depth allows more minor taxa to be covered, while the proportion of dominant taxa remains similar at different sequencing depths. It is required to optimise the metabarcoding method for use in other seasons.

## Acknowledgements

The author is grateful to Dr., prof. E.V. Likhoshway for valuable advice and assistance in preparing the manuscript; Dr. D.P. Petrova for participation in planning the study, Dr.V.V. Blinov and the team of the research vessel "Akademik G.Yu. Vereshchagin" (The Center for Collective Use «Research vessels Center of LIN SB RAS on Lake Baikal») for assistance in sampling. This work was supported by the Ministry of Science and Higher Education of the Russian Federation, project no. 121032300186-9. The study was carried out using the equipment of the "Genomic Technologies, Proteomics and Cell Biology" of ARRIAM.

## Conflict of Interest

The author declares no conflicts of interest.

## References

- Abad D., Albaina A., Aguirre M. et al. 2016. Is metabarcoding suitable for estuarine plankton monitoring? A comparative study with microscopy. *Marine Biology* 163(7): 149. DOI: [10.1007/s00227-016-2920-0](https://doi.org/10.1007/s00227-016-2920-0)
- Abellan-Schneyder I., Machado M.S., Reitmeier S. et al. 2021. Primer, pipelines, parameters: issues in 16S rRNA gene sequencing. *Msphere* 6(1): 10–1128. DOI: [10.1128/msphere.01202-20](https://doi.org/10.1128/msphere.01202-20)
- Andersson A., Zhao L., Brugel S. et al. 2023. Metabarcoding vs microscopy: comparison of methods to monitor phytoplankton communities. *ACS ES&T Water* 3(8): 2671–2680. DOI: [10.1021/acsestwater.3c00176](https://doi.org/10.1021/acsestwater.3c00176)
- Bardenhorst S.K., Vital M., Karch A. et al. 2022. Richness estimation in microbiome data obtained from denoising pipelines. *Computational and Structural Biotechnology Journal* 20: 508–520. DOI: [10.1016/j.csbj.2021.12.036](https://doi.org/10.1016/j.csbj.2021.12.036)
- Brandt M.I., Trouche B., Quintric L. et al. 2021. Bioinformatic pipelines combining denoising and clustering tools allow for more comprehensive prokaryotic and eukaryotic metabarcoding. *Molecular Ecology Resources* 21(6): 1904–1921. DOI: [10.1111/1755-0998.13398](https://doi.org/10.1111/1755-0998.13398)
- Bukin Y.S., Mikhailov I.S., Petrova D.P. et al. 2023. The effect of metabarcoding 18S rRNA region choice on diversity of microeukaryotes including phytoplankton. *World Journal of Microbiology and Biotechnology* 39(9): 229. DOI: [10.1007/s11274-023-03678-1](https://doi.org/10.1007/s11274-023-03678-1)
- Calbet A., Landry M.R. 2004. Phytoplankton growth, microzooplankton grazing, and carbon cycling in marine systems. *Limnology and Oceanography* 49(1): 51–57. DOI: [10.4319/lo.2004.49.1.0051](https://doi.org/10.4319/lo.2004.49.1.0051)
- Callahan B.J., McMurdie P.J., Rosen M.J. et al. 2016. "DADA2: High-Resolution Sample Inference from Illumina Amplicon Data." *Nature Methods* 13: 581–583. DOI: [10.1038/nmeth.3869](https://doi.org/10.1038/nmeth.3869)

- Casey J.M., Ransome E., Collins A.G. et al. 2021. DNA metabarcoding marker choice skews perception of marine eukaryotic biodiversity. *Environmental DNA* 3(6): 1229–1246. DOI: [10.1002/edn3.245](https://doi.org/10.1002/edn3.245)
- Falkowski P.G., Laws E.A., Barber R.T. et al. 2003. Phytoplankton and their role in primary, new, and export production. In: *Ocean biogeochemistry: The role of the ocean carbon cycle in global change*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 99–121.
- Ficetola G.F., Boyer F., Valentini A. et al. 2021. Comparison of markers for the monitoring of freshwater benthic biodiversity through DNA metabarcoding. *Molecular Ecology* 30(13): 3189–3202. DOI: [10.1111/mec.15632](https://doi.org/10.1111/mec.15632)
- Firsova A., Galachyants Y., Bessudova A. et al. 2023. Environmental Factors Affecting Distribution and Diversity of Phytoplankton in the Irkutsk Reservoir Ecosystem in June 2023. *Diversity* 15(10): 1–20. DOI: [10.3390/d15101070](https://doi.org/10.3390/d15101070)
- Fuhrman J.A. 2009. Microbial community structure and its functional implications. *Nature* 459:193–199. DOI: [10.1038/nature08058](https://doi.org/10.1038/nature08058)
- Gelis M.M.N., Canino A., Bouchez A. 2024. Assessing the relevance of DNA metabarcoding compared to morphological identification for lake phytoplankton monitoring. *Science of the Total Environment* 914: 169774. DOI: [10.1016/j.scitotenv.2023.169774](https://doi.org/10.1016/j.scitotenv.2023.169774)
- Genomics C.D. 2024. The use and types of alpha-diversity metrics in microbial NGS.
- Govindarajan A.F., McCartin L., Adams A. 2022. Improved biodiversity detection using a large-volume environmental DNA sampler with *in situ* filtration and implications for marine eDNA sampling strategies. *Deep Sea Research Part I: Oceanographic Research Papers* 189: 103871. DOI: [10.1016/j.dsr.2022.103871](https://doi.org/10.1016/j.dsr.2022.103871)
- Gran-Stadniczenko S., Šupraha L., Egge E.D. 2017. Haptophyte diversity and vertical distribution explored by 18S and 28S ribosomal RNA gene metabarcoding and scanning electron microscopy. *Journal of Eukaryotic Microbiology* 64(4): 514–532. DOI: [10.1111/jeu.12388](https://doi.org/10.1111/jeu.12388)
- Kezlya E., Tseplik N., Kulikovskiy M. 2023. Genetic Markers for Metabarcoding of Freshwater Microalgae: Review. *Biology* 12: 1038. DOI: [10.3390/biology12071038](https://doi.org/10.3390/biology12071038)
- Kobanova G.I. 2009. Morphology and life cycle of *Gymnodinium baicalense* Ant. (Dinophyceae) from Lake Baikal. *Contemporary Problems of Ecology* 2(6): 581–585. DOI: [10.1134/S1995425509060150](https://doi.org/10.1134/S1995425509060150)
- Kolde R. 2019. Pheatmap: Pretty Heatmaps. R Package Version 1.0.12. URL: <https://CRAN.R-project.org/package=pheatmap>
- Mäki A., Salmi P., Mikkonen A. 2017. Sample preservation, DNA or RNA extraction and data analysis for high-throughput phytoplankton community sequencing. *Frontiers in Microbiology* 8: 1848. DOI: [10.3389/fmicb.2017.01848](https://doi.org/10.3389/fmicb.2017.01848)
- Mikhailov I.S., Bukin Yu.S., Firsova A.D. 2025. Comparison of Relative and Absolute Abundance and Biomass of Freshwater Phytoplankton Taxa Using Metabarcoding and Microscopy. *Ecology and Evolution* 15(3): 1–24. DOI: [10.1002/ece3.70856](https://doi.org/10.1002/ece3.70856)
- Oksanen J., Blanchet F.G., Friendly M. et al. 2019. Vegan: Community Ecology Package. R Package Version 2.5-6. URL: <https://CRAN.R-project.org/package=vegan>
- Prodan A., Tremaroli V., Brolin H. 2020. Comparing bioinformatic pipelines for microbial 16S rRNA amplicon sequencing. *PloS one* 15(1): e0227434. DOI: [10.1371/journal.pone.0227434](https://doi.org/10.1371/journal.pone.0227434)
- Schmitz J.E., Rahmann S. 2025. A comprehensive review and evaluation of species richness estimation. *Briefings in Bioinformatics* 26(2): bbaf158. DOI: [10.1093/bib/bbaf158](https://doi.org/10.1093/bib/bbaf158)
- Vasselon V., Domaizon I., Rimet F. 2017. Application of high-throughput sequencing (HTS) metabarcoding to diatom biomonitoring: Do DNA extraction methods matter?. *Freshwater Science* 36(1): 162–177. DOI: [10.1086/690649](https://doi.org/10.1086/690649)
- Votintsev K.K., Meshcheryakova A.I., Popovskaya G.I. 1975. Cycle of organic matter in Lake Baikal. Novosibirsk: Nauka.
- Williams P.J.L.B., Thomas D.N., Reynolds C.S. 2008. Phytoplankton productivity: carbon assimilation in marine and freshwater ecosystems. John Wiley & Sons.
- Winder M., Sommer U. 2012. Phytoplankton response to a changing climate. *Hydrobiologia* 698(1): 5–16. DOI: [10.1007/s10750-012-1149-2](https://doi.org/10.1007/s10750-012-1149-2)
- Wisecaver J.H., Hackett J.D. 2011. Dinoflagellate genome evolution. *Annual review of microbiology* 65(1): 369–387. DOI: [10.1146/annurev-micro-090110-102841](https://doi.org/10.1146/annurev-micro-090110-102841)
- Zaheer R., Noyes N., Ortega Polo R. 2018. Impact of sequencing depth on the characterization of the microbiome and resistome. *Scientific reports* 8(1): 5890. DOI: [10.1038/s41598-018-24280-8](https://doi.org/10.1038/s41598-018-24280-8)
- Zimmermann J., Glöckner G., Jahn R. 2015. Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies. *Molecular ecology resources* 15(3): 526–542. DOI: [10.1111/1755-0998.12336](https://doi.org/10.1111/1755-0998.12336)

# Влияние объема пробы и количества прочтений на результат секвенирования микроэукариотического сообщества из олиготрофного озера

Башенхаева М.В.\*

Лимнологический институт Сибирского отделения Российской академии наук, ул. Улан-Баторская, д. 3, Иркутск, 664033, Россия

**АННОТАЦИЯ.** Микроэукариоты играют ключевую роль в водных экосистемах и являются эффективными индикаторами их состояния. Изменения в их численности или разнообразии могут говорить об изменениях в среде обитания. В последние годы, помимо классического методом микроскопии, для мониторинга микроэукариот начали применять метод ДНК метабаркодирования. Для оптимизации метода метабаркодирования актуально установить необходимый объем отобранного образца и глубину секвенирования. В данной работе проведен сравнительный анализ образцов с разным объемом отфильтрованной пробы (50 л, 10 л и 1.2 л), а также разной глубиной секвенирования и влияния этих факторов на разнообразие и таксономический состав сообщества одноклеточных микроэукариот в олиготрофном озере Байкал. Тест ANOSIM показал сходство образцов с разными отфильтрованными объемами. То есть даже небольшого объема пробы для фильтрации (1.2 л) хватает, чтобы выявить высокое разнообразие таксонов, сопоставимое с разнообразием при фильтрации большого объема, 50 л. При большом объеме отфильтрованной пробы по результатам метабаркодирования наибольшее количество прочтений относилось к представителям Ciliophora. Диатомовые и зелёные водоросли были лучше детектированы при меньшем объеме – 10 л и 1.2 л, поэтому для мониторинга микроводорослей этого объема достаточно. Сравнение образцов с разной глубиной секвенирования показало сходные результаты в характеристике доминирующих видов. Разница проявляется в минорных таксонах, некоторые из которых выявлялись только при большем количестве прочтений. Полученные результаты важны для оптимизации мониторинга олиготрофных водоемов с применением метода метабаркодирования.

**Ключевые слова:** разнообразие, объем пробы, глубина секвенирования, метабаркодирование, 18S рРНК, озеро Байкал

**Для цитирования:** Башенхаева М.В. Влияние объема пробы и количества прочтений на результат секвенирования микроэукариотического сообщества из олиготрофного озера // Limnology and Freshwater Biology. 2025. - № 5. - С. 1212-1226. DOI: [10.31951/2658-3518-2025-A-5-1212](https://doi.org/10.31951/2658-3518-2025-A-5-1212)

## 1. Введение

Микроэукариоты – группа одноклеточных организмов, которая включает микроводоросли, простейшие и микроскопические грибы – важный компонент водных экосистем. Они участвуют в глобальном цикле C, Si и других биогенных элементов и являются основой пищевых сетей в водной среде (Falkowski et al., 2003; Calbet and Landry, 2004; Williams et al., 2008; Fuhrman, 2009). Поскольку изменения в окружающей среде, такие как температура, содержание биогенных веществ, соленость и освещенность влияют на структуру

сообществ микроэукариот, данные об их обилии и видовом составе могут служить ранним предупреждением о нарушениях в экосистеме, связанных с загрязнением или изменением климата (Winder and Sommer, 2012). Классическим методом мониторинга одноклеточных микроэукариот является микроскопия, которая основана на морфологических признаках, что требует экспертных знаний и может усложняться из-за малых размеров клеток и морфологической изменчивости видов. В последнее десятилетие для биомониторинга стали применять метод метабаркодирования, в основе которого

\*Автор для переписки.

Адрес e-mail: [maria.bashenkhaeva@gmail.com](mailto:maria.bashenkhaeva@gmail.com) (М.В. Башенхаева)

**Поступила:** 29 августа 2025; **Принята:** 27 октября 2025;  
**Опубликована online:** 31 октября 2025

© Автор(ы) 2025. Эта работа распространяется под международной лицензией Creative Commons Attribution-NonCommercial 4.0.



лежит анализ разнообразия последовательностей маркерных генов в суммарной ДНК (Zimmermann et al., 2015; Abad et al., 2016; Andersson et al., 2023; Gelis et al., 2024; Mikhailov et al., 2025). Учитывая многоэтапность метода метабаркодирования, актуален выбор оптимальных условий для каждого из этапов. На полученный результат могут повлиять различные факторы, такие как: 1) объем фильтруемого образца; 2) выбранный ДНК-маркер; 2) методы, используемые на различных этапах молекулярного анализа (методы выделения ДНК, технологии секвенирования) и 3) различные подходы биоинформатического анализа (методы первичной обработки данных; алгоритмы кластеризации и методы таксономической классификации). Для целей биомониторинга нужно провести стандартизацию этих этапов. На данный момент существует немало работ по выбору маркерного гена для метабаркодирования (Gran-Stadniczeńko et al., 2017; Casey et al., 2021; Ficetola et al., 2021; Bukin et al., 2023; Kezlya et al., 2023); сравнению различных методов выделения ДНК (Mäki et al., 2017; Vasselon et al., 2017) и оптимизации обработки последовательностей данных высокопроизводительного секвенирования (Prodan et al., 2020; Abellan-Schneyder et al., 2021; Brandt et al., 2021). Объем пробы также влияет на итоговые результаты, включая количество выделенной ДНК и качество полученных образцов для секвенирования. Выбор объема пробы для фильтрования зависит от типа водоема. Морские экосистемы богаты фито- и зоопланктоном, и чтобы учесть всё разнообразие организмов рекомендуется отбирать большие объемы – от 50 до 100 литров (Govindarajan et al., 2022). Для неглубоких мезотрофных эстуарий достаточно небольшого объема воды (25 – 500 мл), чтобы оценить биоразнообразие с помощью метабаркодирования (Andersson et al., 2023). В пресных водоемах необходимый объем пробы может варьировать в зависимости от трофности водоема и концентрации клеток разнообразных микроскопических организмов.

Озеро Байкал является олиготрофным водоемом с низкой концентрацией биогенных веществ и небольшой первичной продукцией (Вотинцев и др., 1975). Целью данной работы была оценка необходимого объема проб и глубины секвенирования для определения разнообразия микроэукариот с помощью метабаркодирования фрагментов гена 18S рРНК.

## 2. Материалы и методы

### 2.1. Отбор проб

Отбор проб воды проводили на станции, расположенной в 3 км от пос. Маритуй (51°45.546' с.ш.; 104°13.222' в.д.), в южной котловине озера Байкал, 18 июля 2023 г. Пробы были отобраны с борта научно-исследовательского судна «Г.Ю. Верещагин» с помощью водосборника SBE 32 Carousel (Sea-Bird Electronics, США) из верхних слоев 0, 5, 10, 15, 20 и 25 м. Для дальнейшей работы равные объемы воды из разных слоев объединяли в интегральную

пробу. Фитопланктон анализировали параллельно с помощью микроскопии и ДНК метабаркодирования. Для количественного определения и идентификации фитопланктона методом световой микроскопии 1,2 л интегральной пробы фильтровали через фильтр с размером пор 3 мкм (РЕАТРЕК-Фильтр, Обнинск-3, Россия) и фиксировали формальдегидом в объеме 50 мл до конечной концентрации 3,7%. Подсчет клеток микроводорослей проводили с помощью микроскопа Axiostar Plus (Zeiss, Германия) в двух повторностях, как описано ранее (Firsova et al., 2023).

### 2.2. Выделение ДНК, амплификация и метабаркодирование

Для ДНК метабаркодирования 1,2, 10 и 50 л интегральных проб предварительно фильтровали через газовый фильтр с размером пор 100 мкм (для удаления крупной фракции зоопланктона), а затем фильтровали через фильтры с размером пор 3 мкм (РЕАТРЕК-Фильтр, Обнинск-3, Россия). Объемы были выбраны таким образом: объем 1,2 л равный объему пробы для микроскопии; объем 10 л, который обычно используют в работе для оценки биоразнообразия с помощью метабаркодирования ДНК, и объем 50 л, используемый для полногеномного секвенирования. Фильтрацию проводили сразу после отбора проб. Биомассу с фильтра смывали в стерильные флаконы с 10 мл стерильного ТЕ буфера (1 мМ ЭДТА и 10 мМ Трис-НCl; pH 7,5) и хранили при -20°C, а затем при -80°C до дальнейшего анализа. Суммарную ДНК из образцов экстрагировали лизоцимом (1 мг/мл), протеиназой K, 10% SDS и смесью фенола, хлороформа и изоамилового спирта (25:24:1) (Bukin et al., 2023). Амплификацию и секвенирование на платформе Illumina MiSeq проводили в Центре геномных технологий, протеомики и клеточной биологии ВНИИСХМ (Санкт-Петербург, Россия). Для сравнения образцов разного объема использовали универсальный праймер 18S рРНК участок V8–V9 (V8f(ATAACAGGTCTGTGATGCCCT)/1510R(CCCTTCYGCAGGTTACCTAC)). Секвенирование проводили с разным количеством прочтений на образец с чаще используемым (примерно 20 тыс. прочтений) и с увеличенным (примерно 50 тыс. прочтений).

### 2.3. Обработка данных секвенирования

Данные секвенирования анализировали с помощью DADA2 v1.16 (R-пакет dada2). Парные концы прочтений были отфильтрованы, объединены, химерные и короткие последовательности удалены, сгенерированы варианты последовательностей ампликонов (по-английски amplicon sequence variants (ASV)) (Callahan et al., 2016). Кривые разрежения, индексы богатства и разнообразия (Чео1, Шеннон и Симпсон) рассчитаны с использованием R версии 4.4.3. Для сравнения образцов с разным объемом и количеством прочтений использовался тест ANOSIM. Анализ выполнялся в PAST 4.06b. Для

таксономической идентификации использовали таксономию Silva 132 (Бремен, Германия). Тепловая карта построена на основе 80 наиболее многочисленных ASV в R. Матрица парных расстояний, рассчитанная с помощью индекса Брея-Кертиса, использовалась для кластеризации профилей сообществ в тепловой карте UPGMA с использованием пакетов vegan (Oksanen et al., 2019) и pheatmap (Kolde, 2019).

Данные о последовательностях были загружены в базу данных Sequence Read Archive (<https://www.ncbi.nlm.nih.gov/sra/>) Национального центра биотехнологической информации (NCBI) под регистрационным номером PRJNA1348943.

3. Результаты и обсуждения

Для оценки разнообразия и таксономического состава микрорукриот в образцах с разным объемом отфильтрованной пробы и с разной глубиной секвенирования всего было проанализировано 972 147 прочтений. Для образцов с меньшей глубиной секвенирования было получено от 34 931 до 52 392 прочтений на образец, для образцов с увеличенной глубиной секвенирования – от 90 721 до 137 764 (Таблица). После фильтрации полученных прочтений по качеству, которая заключалась в удалении праймерных областей, удалении последовательности при длине фрагмента меньше 250 п.н.; объединении парных прочтений; удалении коротких и химерных последовательностей осталось от 25 328 до 103 680 на образец. Больше половины прочтений практически во всех образцах было удалено после удаления последовательностей, относящихся к Metazoa. Несмотря на механическое удаление

крупных эукариот при фильтрации проб, их ДНК всё равно попадает в образцы и способна заполнять большую часть прочтений. При формировании ASV всего было получено от 67 до 121 ASV на образец, со средней длиной 331 п.н.

3.1. Оценка богатства и разнообразия

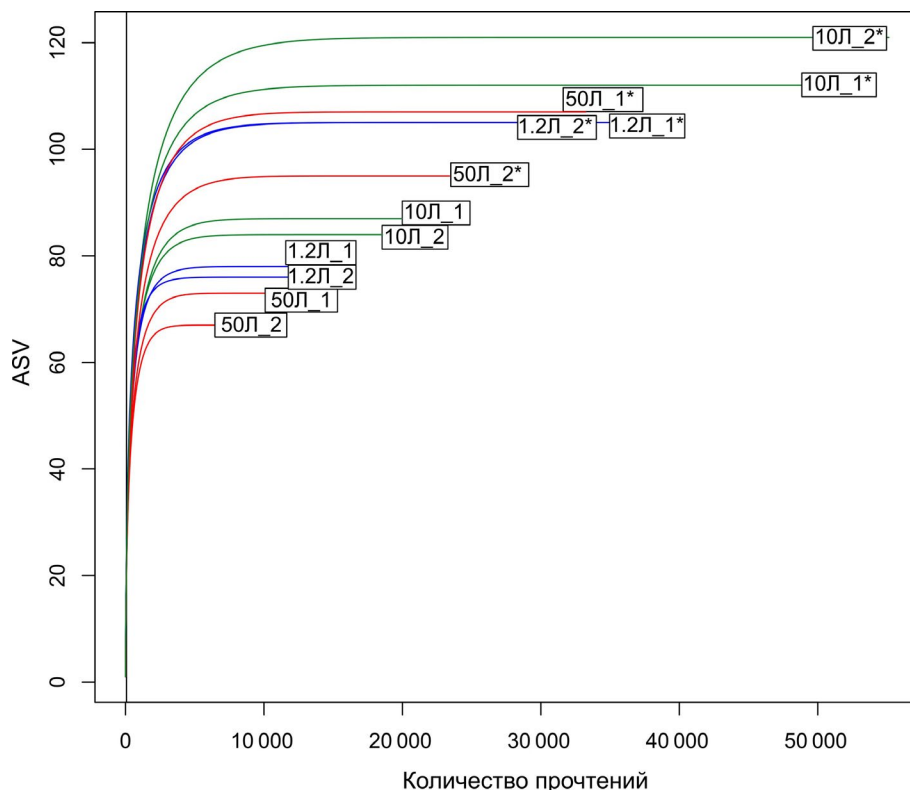
Анализ альфа-разнообразия показал, что количество ASV не различалось при объемах фильтрации воды 1,2 л, 10 л и 50 л (Таблица). Кривые разрежения, рассчитанные для всех образцов, достигали насыщения (Рис. 1), что указывает на то, что количество прочтений было достаточным для оценки видового богатства. Индекс Чао1 варьировал от 67 до 121, наибольшие значения были в образцах с увеличенным количеством прочтений, что говорит о корреляции видового богатства с глубиной секвенирования, которое позволяет выявлять больше редких видов (Zaheer et al., 2018; Bardenhorst et al., 2022; Schmitz and Rahmann, 2025). При небольшой и увеличенной глубине секвенирования наибольшее богатство было отмечено в образцах 10 л (Рис. 2А; Таблица). Этот результат говорит о том, что данный объем может быть наиболее оптимальным среди исследуемых. Фильтрация, в процессе которого клетки из пробы концентрируются на фильтре, идет быстрее, что делает более сохраняемыми образцы, по сравнению с 50 л, и в пробу попадает больше видов при сравнении с 1,2 л.

Тест ANOSIM показал, что образцы с разными отфильтрованными объемами проб имеют едва различимые различия по составу между собой ( $R = 0,2384$ ,  $p = 0,0482$ ). Это говорит о чувствительности метода метабаркодирования, что при неболь-

Таблица. Показатели секвенирования и индексы биоразнообразия в образцах с разным объемом отфильтрованной пробы (50, 10 и 1.2 л фильтрованной воды), проанализированных с помощью метабаркодирования.

|                                                                                | 50Л_1  | 50Л_2  | 10Л_1  | 10Л_2  | 1.2Л_1 | 1.2Л_2 | 50Л_1*  | 50Л_2* | 10Л_1*  | 10Л_2*  | 1.2Л_1* | 1.2Л_2* |
|--------------------------------------------------------------------------------|--------|--------|--------|--------|--------|--------|---------|--------|---------|---------|---------|---------|
| Исходное количество прочтений                                                  | 40 419 | 34 931 | 52 392 | 51 857 | 48 786 | 43 390 | 109 918 | 90 721 | 135 541 | 137 764 | 126 154 | 100 274 |
| Количество прочтений после фильтрации                                          | 38 565 | 33 164 | 49 811 | 49 416 | 46 481 | 41 413 | 104 974 | 86 536 | 129 633 | 131 377 | 119 792 | 95 645  |
| Количество прочтения после удаления химерных и коротких последовательностей    | 26 993 | 25 328 | 39 534 | 37 841 | 38 886 | 36 945 | 72 053  | 64 074 | 103 680 | 100 215 | 98 822  | 85 129  |
| Количество прочтений после удаления последовательностей, относящихся к Metazoa | 11 709 | 8 084  | 18 553 | 19 953 | 12 361 | 13 381 | 33 241  | 23 400 | 51 697  | 55 129  | 35 839  | 30 087  |
| ASV                                                                            | 73     | 67     | 84     | 87     | 78     | 76     | 107     | 95     | 112     | 121     | 105     | 105     |
| Индекс Чао1                                                                    | 73     | 67     | 84     | 87     | 78     | 76     | 107     | 95     | 112     | 121     | 105     | 105     |
| Индекс Шеннона                                                                 | 2.7    | 2.72   | 2.85   | 2.88   | 2.93   | 3.1    | 2.82    | 2.76   | 2.95    | 2.96    | 3.01    | 3.16    |
| Индекс Симпсона                                                                | 0.84   | 0.86   | 0.87   | 0.88   | 0.90   | 0.91   | 0.85    | 0.86   | 0.88    | 0.88    | 0.9     | 0.92    |

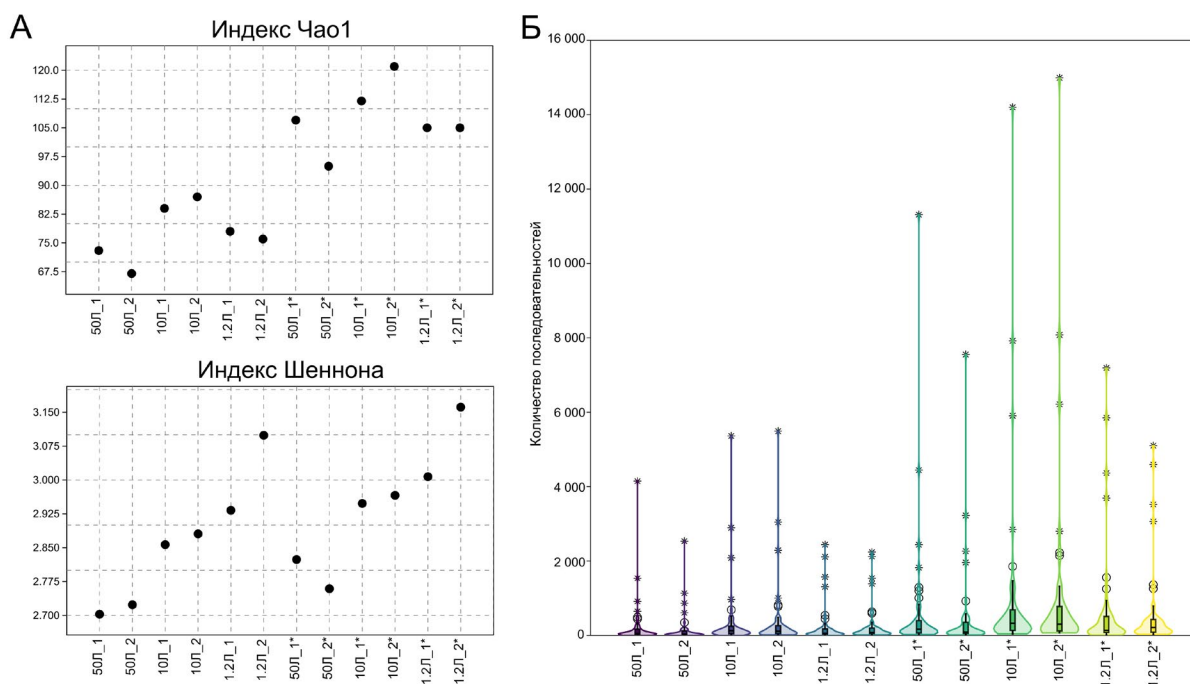
Примечание: образцы с увеличенным количеством прочтений отмечены звездочкой (\*).



**Рис.1.** Кривые разряжения для образцов с разным объемом отфильтрованной пробы воды (50 л, 10 л и 1.2 л) и с разной глубиной секвенирования. Образцы с увеличенным количеством прочтений отмечены звездочкой (\*).

шом объеме фильтрации (1,2 л) выявляется высокое разнообразие таксонов, равное большому объему (50 л). Этот результат соотносится с результатами, полученными на образцах мезотрофного эстуария реки Эре, где сравнивали объемы 0,25 л, 0,5 л и 1,2 л (Andersson et al., 2023). В их работе разнообразие было зафиксировано уже при самом низком исследованном объеме пробы – 0,25 л.

Индексы разнообразия Шеннон и Симпсон между образцами с разными количествами прочтений были сходными (Рис. 2А; Таблица). Индекс Шеннона при 20 тыс. прочтений варьировал от 2,7 до 3,1; при 50 тыс. прочтений от 2,76 до 3,16. Индекс Симпсона изменялся от 0,84 до 0,91 при 20 тыс. прочтений и от 0,85 до 0,92 при 50 тыс. прочтений. То есть общая структура сообщества (относительное



**Рис.2.** Показатели богатства и разнообразия для образцов с разным объемом отфильтрованной пробы воды (50 л, 10 л и 1.2 л) и с разной глубиной секвенирования (20 тыс. и 50 тыс. прочтений). А. Значения индексов Чао1 и индекса Шеннона; Б. Распределение 50 наиболее многочисленных ASV. Образцы с увеличенным количеством прочтений отмечены звездочкой (\*).

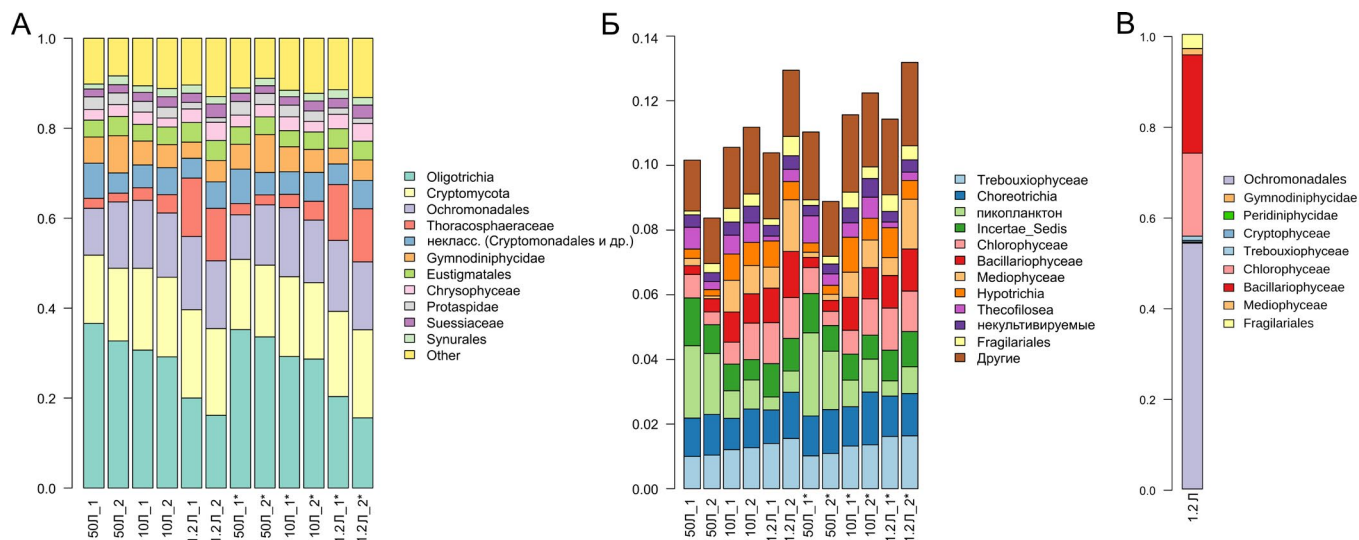
обилие видов) остается стабильной, даже при увеличении глубины секвенирования. Максимальные значения разнообразия отмечены в образцах 1,2 л, что характеризуется более равномерным распределением видов в образце. Минимальные значения индексов Шеннона и Симпсона были в образцах 50 л. Низкие показатели данных индексов свидетельствуют о доминировании в сообществе отдельных таксонов (Genomics, 2024). При сравнении состава ASV между образцами видно, что в образцах 50 л больше половины прочтений относится к одной ASV (Рис. 2Б).

### 3.2. Таксономическая структура

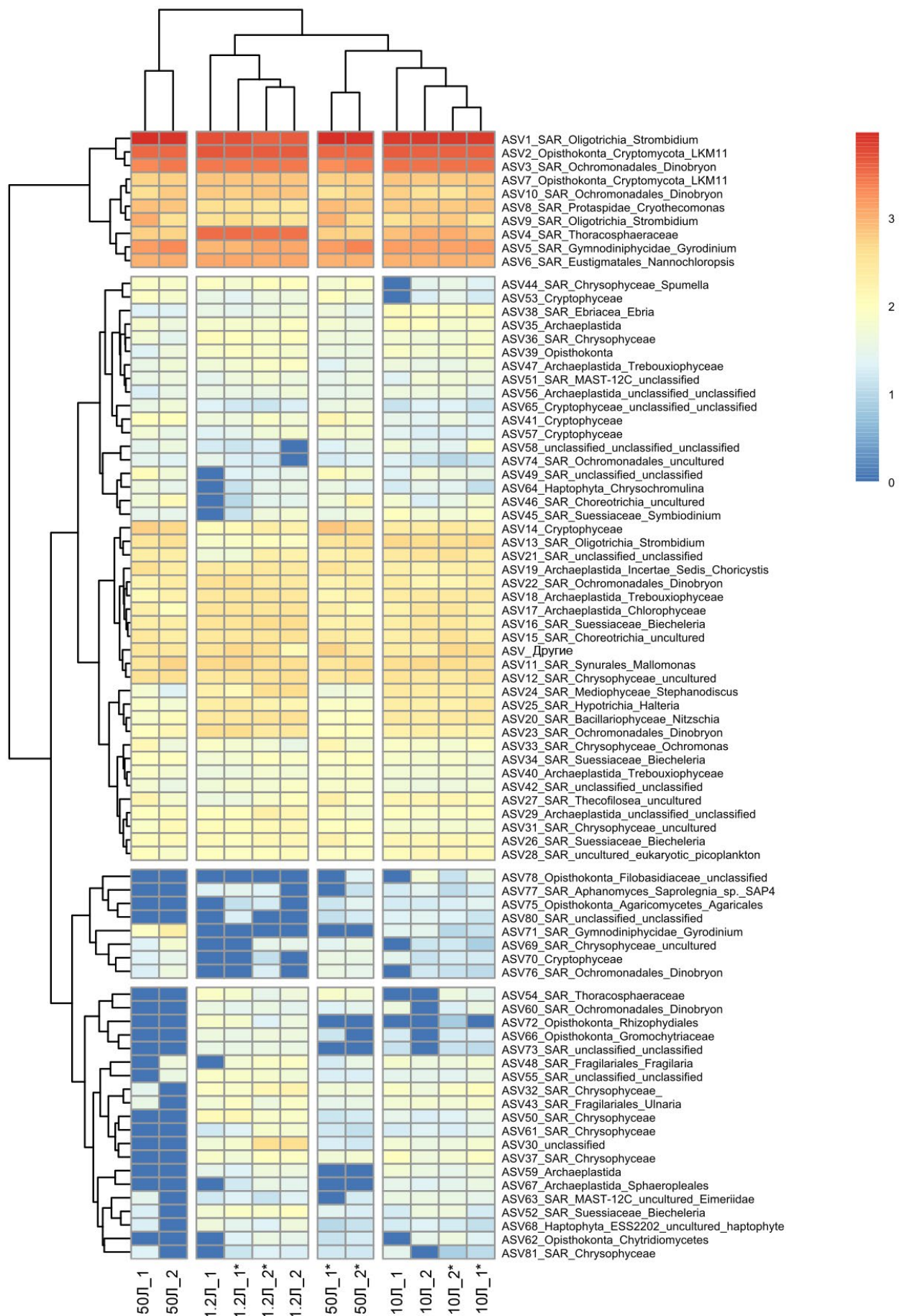
Для оценки влияния объема фильтрованных проб на таксономическую структуру сообщества микроэукариот, мы сравнили отдельно доминирующие и минорные ASV на уровне семейства (Рис. 3). В пробах 50 л и 10 л наибольшую долю составляли ASV, относящиеся к Oligotrichia (Ciliophora). Из тепловой карты также видно, что ASV1, относящееся к роду *Strombidium*, является доминирующей во всех образцах 50 л и 10 л и составляет значительную долю в образцах 1,2 л (Рис. 4). Приблизительно одинаково представлены в пробах Cryptomycota и Ochomonadales (Рис. 3А), данные семейства являются субдоминантами или доминантами в пробах. ASV3, классифицированное как представитель рода *Dinobryon*, является одним из наиболее многочисленных среди всех образцов. Данный результат сопоставим с результатом, полученным с помощью микроскопии фитопланктона (Рис. 3Б), где по численности тоже доминировали представители Ochomonadales, *Dinobryon cylindricum* ( $65,4 \times 10^3$  кл/л) и *Dinobryon sociale* ( $48,20 \times 10^3$  кл/л). В образцах 1,2 л значительная доля относится к семейству Thoracosphaeraceae (Dinoflagellata), в отличие от образцов 50 л и 10 л, где их было в два или три раза меньше.

Динофлагелляты, обитающие в холодных водоемах, очень чувствительны к повышению температуры, которое вызывает быстрое разрушение жгутиконосцев (Кобанова, 2009). Возможно, более длительное фильтрование 50 л и 10 л способствовало разрушению клеток и деградации ДНК. По сравнению с данными метабаркодирования, где представители динофлагеллят были в числе 11 доминирующих семейств (Рис. 3А), по данным микроскопии, доля динофлагеллят в общей численности была небольшая (Рис. 3Б). Геном динофлагеллят является многокопийным, а также динофлагелляты часто обладают «мезокариотическим» ядром, в котором хромосомы не упакованы в гистоны, как у других эукариот, и находятся в более растянутом состоянии, что может способствовать более частому и легкому копированию хромосомного материала (Wisecaver and Hackett, 2011). Соответственно даже при низкой численности динофлагелляты могут иметь значительную долю при секвенировании. Доля прочтений, относящихся к диатомовым водорослям (семейства Bacillariophyceae, Mediophyceae и порядка Fragilariiales), а также зеленым водорослям (семейства Trebouxiophyceae и Chlorophyceae), была в образцах 1,2 л. В образцах 50 л вклад обеих групп микроводорослей был небольшим (Рис. 3Б), несмотря на то, что по микроскопии эти группы являлись субдоминантами (Рис. 3Б).

При сравнении образцов с разной глубиной секвенирования видно, что большинство таксонов имеют приблизительно одинаковое соотношение (Рис. 3, 4). Разница проявляется в минорных таксонах, некоторые из которых детектируются только при большем количестве полученных последовательностей (Рис. 4). Тест ANOSIM подтвердил наличие различий в составе образцов с разной глубиной секвенирования ( $R = 0,513$ ,  $p = 0,0029$ ). Из тепловой карты видно, что наиболее полно сообщество микроэукариот представлено в образцах объемом 10 л с увеличенной глубиной секвенирования, что в данном случае является оптимальными условиями.



**Рис.3.** Таксономическая структура образцов с разным объемом отфильтрованной пробы (50, 10 и 1.2 л фильтрованной воды). Анализ с помощью метабаркодирования на уровне семейств 11 доминирующих ASV (А) и минорных ASV; результаты анализа пробы 1.2 л методом микроскопии (В). Образцы с увеличенным количеством прочтений отмечены звездочкой (\*).



**Рис.4.** Тепловая карта образцов микробиоты с разным объемом отфильтрованной пробы (50, 10 и 1.2 л фильтрованной воды), созданная на основе набора из 80 наиболее распространенных ASV. Для построения кластерных деревьев по выборкам и по ASV использовалась кластеризация матриц парных расстояний Брея–Кертиса с помощью UPGMA. Образцы с увеличенным количеством прочтений отмечены звездочкой (\*).

## 4. Заключение

Таким образом, для мониторинга одноклеточных микроэукариот в олиготрофном озере Байкал в летний период методом метабаркодирования по гену 18S рРНК оптимально отобрать пробу объемом 10 л и провести секвенирование с глубиной 50 тысяч прочтений на образец, чтобы более детально охарактеризовать сообщество. Для характеристики доминирующих таксонов достаточно относительно небольшого объема воды (1,2 л). В пробах с наибольшим объемом (50 л) больше половины всех секвенированных последовательностей, относилась к Ciliophora, а количество прочтений, относящихся к диатомовым и зеленым водорослям, снижалось, что не оправдывает отбор такого большого объема пробы для исследования фитопланктона. Увеличение глубины секвенирования позволяет охватить больше минорных таксонов, в то время как доля доминирующих таксонов при разной глубине секвенирования была сходна. Для оптимизации метода метабаркодирования для исследований и в другие сезоны необходимо проведение дополнительных работ.

## Благодарности

Автор выражает благодарность д.б.н., проф. Е.В. Лихошвай за ценные советы и помощь в подготовке рукописи; к.б.н. Д.П. Петровой за участие в планировании исследования; к.г.н. В.В. Блинову и команде НИС «Академик Г.Ю. Верецагин» (ЦКП «Центр комплексных судовых исследований Байкала») за помощь в отборе проб. Работа выполнена при поддержке Министерства науки и высшего образования Российской Федерации, проект № 121032300186-9. Исследование проведено с использованием оборудования ресурсного центра «Геномные технологии, протеомика и клеточная биология» ВНИИСХМ.

## Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

## Список литературы

Вотинцев К.К., Мещерякова А.И., Поповская Г.И. 1975. Круговорот органического вещества в озере Байкал. Новосибирск: Наука.

Кобанова Г.И. 2009. О морфологии и жизненном цикле *Gymnodinium baicalense* Ant. (Dinophyta) из озера Байкал. Сибирский экологический журнал 16(6): 899–905. DOI: [10.1134/S1995425509060150](https://doi.org/10.1134/S1995425509060150)

Abad D., Albaina A., Aguirre M. et al. 2016. Is metabarcoding suitable for estuarine plankton monitoring? A comparative study with microscopy. Marine Biology 163(7): 149. DOI: [10.1007/s00227-016-2920-0](https://doi.org/10.1007/s00227-016-2920-0)

Abellan-Schneyder I., Matchado M.S., Reitmeier S. et al. 2021. Primer, pipelines, parameters: issues in 16S rRNA gene sequencing. Msphere 6(1): 10–1128. DOI: [10.1128/msphere.01202-20](https://doi.org/10.1128/msphere.01202-20)

Andersson A., Zhao L., Brugel S. et al. 2023. Metabarcoding vs microscopy: comparison of methods to monitor phytoplankton communities. ACS ES&T Water 3(8): 2671–2680. DOI: [10.1021/acsestwater.3c00176](https://doi.org/10.1021/acsestwater.3c00176)

Bardenhorst S.K., Vital M., Karch A. et al. 2022. Richness estimation in microbiome data obtained from denoising pipelines. Computational and Structural Biotechnology Journal 20: 508–520. DOI: [10.1016/j.csbj.2021.12.036](https://doi.org/10.1016/j.csbj.2021.12.036)

Brandt M.I., Trouche B., Quintric L. et al. 2021. Bioinformatic pipelines combining denoising and clustering tools allow for more comprehensive prokaryotic and eukaryotic metabarcoding. Molecular Ecology Resources 21(6): 1904–1921. DOI: [10.1111/1755-0998.13398](https://doi.org/10.1111/1755-0998.13398)

Bukin Y.S., Mikhailov I.S., Petrova D.P. et al. 2023. The effect of metabarcoding 18S rRNA region choice on diversity of microeukaryotes including phytoplankton. World Journal of Microbiology and Biotechnology 39(9): 229. DOI: [10.1007/s11274-023-03678-1](https://doi.org/10.1007/s11274-023-03678-1)

Calbet A., Landry M.R. 2004. Phytoplankton growth, microzooplankton grazing, and carbon cycling in marine systems. Limnology and Oceanography 49(1): 51–57. DOI: [10.4319/lo.2004.49.1.0051](https://doi.org/10.4319/lo.2004.49.1.0051)

Callahan B.J., McMurdie P.J., Rosen M.J. et al. 2016. “DADA2: High-Resolution Sample Inference from Illumina Amplicon Data.” Nature Methods 13: 581–583. DOI: [10.1038/nmeth.3869](https://doi.org/10.1038/nmeth.3869)

Casey J.M., Ransome E., Collins A.G. et al. 2021. DNA metabarcoding marker choice skews perception of marine eukaryotic biodiversity. Environmental DNA 3(6): 1229–1246. DOI: [10.1002/edn3.245](https://doi.org/10.1002/edn3.245)

Falkowski P.G., Laws E.A., Barber R.T. et al. 2003. Phytoplankton and their role in primary, new, and export production. In: Ocean biogeochemistry: The role of the ocean carbon cycle in global change. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 99–121.

Ficetola G.F., Boyer F., Valentini A. et al. 2021. Comparison of markers for the monitoring of freshwater benthic biodiversity through DNA metabarcoding. Molecular Ecology 30(13): 3189–3202. DOI: [10.1111/mec.15632](https://doi.org/10.1111/mec.15632)

Firsova A., Galachyants Y., Bessudova A. et al. 2023. Environmental Factors Affecting Distribution and Diversity of Phytoplankton in the Irkutsk Reservoir Ecosystem in June 2023. Diversity 15(10): 1–20. DOI: [10.3390/d15101070](https://doi.org/10.3390/d15101070)

Fuhrman J.A. 2009. Microbial community structure and its functional implications. Nature 459:193–199. DOI: [10.1038/nature08058](https://doi.org/10.1038/nature08058)

Gelis M.M.N., Canino A., Bouchez A. 2024. Assessing the relevance of DNA metabarcoding compared to morphological identification for lake phytoplankton monitoring. Science of the Total Environment 914: 169774. DOI: [10.1016/j.scitotenv.2023.169774](https://doi.org/10.1016/j.scitotenv.2023.169774)

Genomics C.D. 2024. The use and types of alpha-diversity metrics in microbial NGS.

Govindarajan A.F., McCartin L., Adams A. 2022. Improved biodiversity detection using a large-volume environmental DNA sampler with *in situ* filtration and implications for marine eDNA sampling strategies. Deep Sea Research Part I: Oceanographic Research Papers 189: 103871. DOI: [10.1016/j.dsr.2022.103871](https://doi.org/10.1016/j.dsr.2022.103871)

Gran-Stadniczenko S., Šupraha L., Egge E.D. 2017. Haptophyte diversity and vertical distribution explored by 18S and 28S ribosomal RNA gene metabarcoding and scanning electron microscopy. Journal of Eukaryotic Microbiology 64(4): 514–532. DOI: [10.1111/jeu.12388](https://doi.org/10.1111/jeu.12388)

Kezlya E., Tseplik N., Kulikovskiy M. 2023. Genetic Markers for Metabarcoding of Freshwater Microalgae: Review. Biology 12: 1038. DOI: [10.3390/biology12071038](https://doi.org/10.3390/biology12071038)

Kolde R. 2019. Pheatmap: Pretty Heatmaps. R Package Version 1.0.12. URL: <https://CRAN.R-project.org/package=pheatmap>

Mäki A., Salmi P., Mikkonen A. 2017. Sample preservation, DNA or RNA extraction and data analysis for high-throughput phytoplankton community sequencing. *Frontiers in Microbiology* 8: 1848. DOI: [10.3389/fmicb.2017.01848](https://doi.org/10.3389/fmicb.2017.01848)

Mikhailov I.S., Bukin Yu.S., Firsova A.D. 2025. Comparison of Relative and Absolute Abundance and Biomass of Freshwater Phytoplankton Taxa Using Metabarcoding and Microscopy. *Ecology and Evolution* 15(3): 1–24. DOI: [10.1002/ece3.70856](https://doi.org/10.1002/ece3.70856)

Oksanen J., Blanchet F.G., Friendly M. et al. 2019. Vegan: Community Ecology Package. R Package Version 2.5-6. URL: <https://CRAN.R-project.org/package=vegan>

Prodan A., Tremaroli V., Brolin H. 2020. Comparing bioinformatic pipelines for microbial 16S rRNA amplicon sequencing. *PloS one* 15(1): e0227434. DOI: [10.1371/journal.pone.0227434](https://doi.org/10.1371/journal.pone.0227434)

Schmitz J.E., Rahmann S. 2025. A comprehensive review and evaluation of species richness estimation. *Briefings in Bioinformatics* 26(2): bbaf158. DOI: [10.1093/bib/bbaf158](https://doi.org/10.1093/bib/bbaf158)

Vasselon V., Domaizon I., Rimet F. 2017. Application of high-throughput sequencing (HTS) metabarcoding to diatom biomonitoring: Do DNA extraction methods matter?. *Freshwater Science* 36(1): 162–177. DOI: [10.1086/690649](https://doi.org/10.1086/690649)

Williams P.J.L.B., Thomas D.N., Reynolds C.S. 2008. Phytoplankton productivity: carbon assimilation in marine and freshwater ecosystems. John Wiley & Sons.

Winder M., Sommer U. 2012. Phytoplankton response to a changing climate. *Hydrobiologia* 698(1): 5–16. DOI: [10.1007/s10750-012-1149-2](https://doi.org/10.1007/s10750-012-1149-2)

Wisecaver J.H., Hackett J.D. 2011. Dinoflagellate genome evolution. *Annual review of microbiology* 65(1): 369–387. DOI: [10.1146/annurev-micro-090110-102841](https://doi.org/10.1146/annurev-micro-090110-102841)

Zaheer R., Noyes N., Ortega Polo R. 2018. Impact of sequencing depth on the characterization of the microbiome and resistome. *Scientific reports* 8(1): 5890. DOI: [10.1038/s41598-018-24280-8](https://doi.org/10.1038/s41598-018-24280-8)

Zimmermann J., Glöckner G., Jahn R. 2015. Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies. *Molecular ecology resources* 15(3): 526–542. DOI: [10.1111/1755-0998.12336](https://doi.org/10.1111/1755-0998.12336)