ECOLOGICAL GENETICS
Original article

Вавиловский журнал генетики и селекции. 2023;27(6):694-702
DOI 10.18699/VJGB-23-80

Original Russian text https://vavilovj-icg.ru/

A study of macroinvertebrate communities in Bolshiye Koty Bay of Lake Baikal using DNA metabarcoding

L.S. Kravtsova[1] ✉, T.E. Peretolchina[1], T.I. Triboy[1], I.A. Nebesnykh[1], A.E. Tupikin[2], M.R. Kabilov[2]

[1] Limnological Institute of the Siberian Branch of the Russian Academy of Sciences, Irkutsk, Russia
[2] Institute of Chemical Biology and Fundamental Medicine of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
✉ lk@lin.irk.ru

Abstract. The diversity of macroinvertebrates, the structure of their communities in Bolshiye Koty Bay (Lake Baikal) was studied by a DNA metabarcoding approach using an Illumina MiSeq system. Internal primer *mlCOIintF* in combination with *jgHCO2198* of the Folmer fragment of the *COI* gene were used for macroinvertebrate metabarcoding. A total of 118009 reads of the *COI* gene fragment (at least 313 bp in length) were obtained. The correlation of the Spearman coefficient (S = 0.6, $p < 0.05$) with the abundance of macroinvertebrates in the samples before DNA extraction showed that the number of reads can serve as an indirect characteristic of the abundance of a species (operational taxonomic unit, OTU). 115 OTUs belonging to the higher taxa of macroinvertebrates were identified: Porifera, 1; Platyhelminthes, 3; Annelida, 38; Arthropoda, 55; Mollusca, 18. At a high level of resolution (with homology with GenBank reference sequences ≥ 95 %, coverage ≥ 90 %), 46 taxa of macroinvertebrates comprising three communities were registered: one dominated by molluscs (*Choanomphalus* conf. *maacki*) and two dominated by chironomids (*Orthocladius gregarius* Linev., *Sergentia baicalensis* Tshern.). Communities are characterized by low species diversity according to Shannon (from 0.7 to 1.2 bits), high concentration of dominance according to Simpson (from 0.5 to 0.7) and low evenness according to Pielou (from 0.3 to 0.4). Dominants and subdominants in the communities account for 91 to 96 % of *COI* gene fragment reads. The spatial distribution of the dominant species identified in the communities is influenced by the geomorphological features of the bottom and the composition of sediments in the area studied. The approach proposed for studying the structure of macroinvertebrate communities based on DNA metabarcoding and next generation sequencing can be recommended for express assessment of the state of aquatic ecosystems in the monitoring.

Key words: communities of macroinvertebrates; diversity; DNA metabarcoding; *COI*; high-throughput sequencing technologies; Lake Baikal.

For citation: Kravtsova L.S., Peretolchina T.E., Triboy T.I., Nebesnykh I.A., Tupikin A.E., Kabilov M.R. A study of macroinvertebrate communities in Bolshiye Koty Bay of Lake Baikal using DNA metabarcoding. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(6):694-702. DOI 10.18699/VJGB-23-80

Исследование сообществ макробеспозвоночных животных в бухте Большие Коты озера Байкал с использованием ДНК метабаркодинга

Л.С. Кравцова[1] ✉, Т.Е. Перетолчина[1], Т.И. Трибой[1], И.А. Небесных[1], А.Е. Тупикин[2], М.Р. Кабилов[2]

[1] Лимнологический институт Сибирского отделения Российской академии наук, Иркутск, Россия
[2] Институт химической биологии и фундаментальной медицины Сибирского отделения Российской академии наук, Новосибирск, Россия
✉ lk@lin.irk.ru

Аннотация. Приводятся сведения о разнообразии макробеспозвоночных животных, структуре их сообществ в бухте Большие Коты оз. Байкал, полученные методом ДНК метабаркодинга на основе NGS-технологии (Illumina, MiSeq). Для ДНК метабаркодинга макробеспозвоночных был использован внутренний праймер *mlCOIintF* в комбинации с *jgHCO2198* для амплификации фолмеровского фрагмента гена *COI*. Всего получено 118009 прочтений фрагмента гена *COI* (длиной не менее 313 п.н.). Показано, что количество прочтений может служить опосредованной характеристикой обилия вида (операционной таксономической единицы – ОТЕ). Корреляция количества прочтений с численностью макробеспозвоночных в пробах до экстракции ДНК по коэффициенту Спирмена составляет 0.6 ($p < 0.05$). Выявлено 115 ОТЕ, принадлежащих высшим таксонам макробеспозвоночных животных: Porifera – 1, Platyhelminthes – 3, Annelida – 38, Arthropoda – 55, Mollusca – 18. На видовом уровне (при гомологии с референсными последовательностями GenBank ≥ 95 % и покрытии не менее 90 %) зарегистрировано 46 таксонов макробеспозвоночных, формирующих три сообщества: одно – с доминированием моллюсков *Choanomphalus* conf. *maacki* и два – с доминированием хирономид *Orthocladius gregarius* Linev., *Sergentia*

baicalensis Tshern. Сообщества характеризуются невысоким видовым разнообразием по Шеннону (от 0.7 до 1.2 бит), высокой концентрацией доминирования по Симпсону (от 0.5 до 0.7) и низкой выравненностью по Пиелу (от 0.3 до 0.4). На долю доминантов и субдоминантов в сообществах приходится от 91 до 96 % прочтений фрагмента гена *COI*. На пространственное распределение доминирующих видов сообществ влияют геоморфологические особенности дна в исследуемом районе и состав донных отложений. Предложенный подход для изучения структуры сообществ макробеспозвоночных на основе ДНК метабаркодинга может быть рекомендован для экспресс-оценки состояния водных экосистем при мониторинге.
Ключевые слова: сообщества макробеспозвоночных; разнообразие; ДНК метабаркодинг; *COI*; высокопроизводительное секвенирование; Байкал.

Introduction

The structure of aquatic organisms communities, in particular macroinvertebrates, is one of the indicators characterizing the state of water bodies. Research in this direction is relevant in connection with global climate change and increasing anthropogenic impact on aquatic ecosystems (O'Reilly et al., 2003; Bonada et al., 2007; Burgmer et al., 2007; Moss et al., 2011; Hampton et al., 2018).

Community of organisms is a set of populations of different species coexisting in space and time (Begon et al., 1986). Their structure is formed under the influence of both abiotic environmental factors (Brauns et al., 2007; McGoff et al., 2013; Rezende et al., 2014; Worrall et al., 2014) and biotic interactions (van den Berg et al., 1997; Arbačiauskas et al., 2008; Nalepa et al., 2009). As a rule, the species diversity and abundance of organisms are the basic characteristics of the structure of communities. The study of the diversity of organisms at a high-resolution species level requires the involvement of a large number of morphologists and is associated with a laborious process of taxa identification. Currently, molecular genetics methods such as DNA metabarcoding using high-throughput sequencing technologies are used as an alternative to the classical methods of the taxonomic diversity studying of aquatic ecosystems. This method is widely used to study the diversity of both marine and freshwater fauna (Porazinska et al., 2009; Hajibabaei et al., 2011; Aylagas et al., 2014; Elbrecht et al., 2017; Haenel et al., 2017; Kuntke et al., 2020).

In the early 2000s, the mitochondrial *COI* gene was adopted as a standard for DNA barcoding of animal taxa (Hebert et al., 2003). Despite the universality of primers (*LCO1490* and *HCO2198*) for the "Folmer fragment" of the mitochondrial *COI* gene amplification (Folmer et al., 1994), many researchers began to use different, more variable regions of it to obtain a clear phylogenetic signal. For example, for marine nematodes, region *I3–M11* was used (Derycke et al., 2010). The mini-barcode of the Folmer *COI* fragment using a combination of primers *mlCOIintF* and *jgHCO2198* has become popular to assess the diversity of Metazoa (Meusnier et al., 2008; Leray et al., 2013). The mini-barcode also proved to be effective in assessing the diversity of benthic invertebrates in the Listvennichny Bay of Lake Baikal (Kravtsova et al., 2021).

In the last decade, the application of environmental DNA (eDNA) approaches for rapid assessment of the biodiversity from water and sediment samples has been increased (Yu et al., 2012; Lacoursière-Roussel et al., 2018). The advantage of this method is a quick result, since no preliminary isolation of any organisms from the samples is required. However, it was less effective for studying the diversity of Metazoa in water bodies. It was shown that DNA metabarcoding of invertebrate tissues gives a more accurate estimation of the diversity of multicellular organisms (99 % of reads) than DNA from the environment (only 12 % of reads) (Gleason et al., 2021). In this study, we tried to find out the acceptability of the method of DNA metabarcoding from the tissues of organisms for assessing not only the diversity of macroinvertebrates, but also for the quantitative ratio of species that form communities.

The aim of this work is to study the features of the structural organization of macroinvertebrate communities distributed in the coastal zone of open Baikal using DNA metabarcoding.

Materials and methods

Quantitative samples of zoobenthos were collected in July 2019 in Bolshiye Koty Bay of Lake Baikal along the coastline over a 1 km long (Fig. 1).

Macroinvertebrates were collected from different types of bottom sediments at three stations (No. 1–3) (see Fig. 1). The first type of bottom sediment included large and small pebbles with individual boulders located in the subaqueous part of the beach (depths 0.3–0.4 m). The second type of bottom sediment was represented by boulders, unrounded rock fragments with crushed stone, and the third type was represented by silted sand. The last two types of sediments were found mainly on a shallow terrace at depths of 2–5 m (see Fig. 1). On the underwater part of the beach and on a shallow terrace samples were taken manually and with the help of divers, respectively. On each type of bottom sediment, five quantitative samples of zoobenthos were collected using a 0.1 m^2 counting frame. Invertebrates from the surface of stones were brushed off into a cuvette with water, and those from silted sand were removed by flotation in a saturated sugar solution with a specific gravity of 1.12 g/L. Samples were washed through a mill sieve No. 23 and fixed with 96 % ethanol. In total, 15 quantitative samples of zoobenthos were collected and sorted under laboratory conditions using an MBS-10 microscope (at 20× magnification).

According to Elbrecht et al. (2017), preliminary sorting of organisms by size significantly improves the result of sequencing of all taxa, regardless of the biomass of the organism. Since the zoobenthos in Baikal is represented by different size groups: mega-, macro-, meioorganisms, in this work we limited ourselves to only one size group – macroinvertebrates. Their sizes in samples varied from 2 to 50 mm. All macroinvertebrates found in the sample were used for DNA extraction. First, invertebrates were soaked in distilled water (1 hour), small organisms (2–3 mm) were taken as a whole, and tissue pieces of 2–3 mm were taken from large individuals (more than 5 mm) in order to level the scatter of biomass

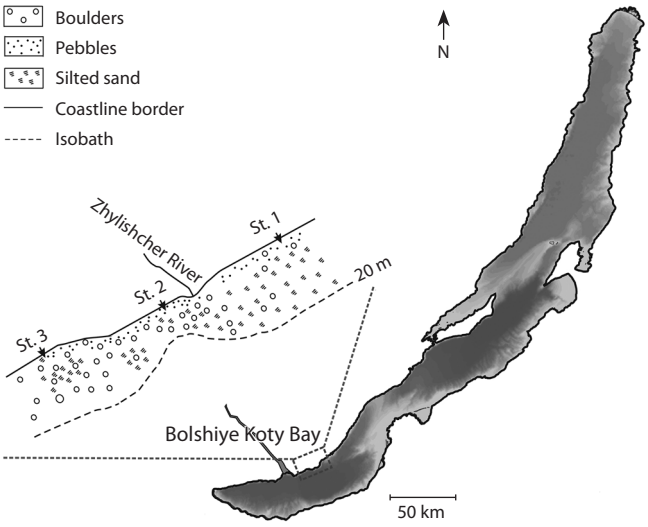

Fig. 1. Map-scheme of sampling localities of macrozoobenthos in Bolshiye Koty Bay, Lake Baikal (July, 2019).

in size. Then, pieces of tissue and small organisms collected from the same type of bottom sediment were combined into one sample, placed in a porcelain cup with 2 % CTAB solution, and ground with a pestle. DNA was extracted according to a modified (chloroform was used instead of a phenol-chloroform-isoamyl mixture) method of Doyle and Dickson (Doyle, Dickson, 1987). In total, three samples of genomic DNA (at least 20 ng each) of invertebrate animals were prepared for metabarcoding.

Primers *mlCOIintF*: GGWACWGGWTGAACWGTW TAYCCYCC (Leray et al., 2013) and *jgHCO2198*: TAIA CYTCIGGRTGICCRAARAAYCA (Geller et al., 2013), where "I" is inosine, were used to obtain *COI* gene amplicons. Amplification was performed in a volume of 20 µl containing 0.2 mM of each dNTP, 0.5 µM of each primer, 2 mM $MgCl_2$, 5 µM SYTO9, 10 ng of DNA, 25 U/ml of Maxima Hot Start Taq DNA Polymerase (Thermo Scientific, Lithuania). Real-time PCR was performed on a CFX96 Touch Real-Time PCR system (Bio-Rad, USA) according to the program: initial denaturation at 95 °C, 4 min; 32 cycles at 95 °C – 30 s, 48 °C – 30 s and 72 °C – 30 s. The primer annealing temperature was selected using gradient PCR. PCR products were analyzed on an MCE-202 MultiNA Microchip Electrophoresis System using a DNA 12000 Reagent Kit (Shimadzu, Japan). The resulting amplicons were quantified on a Qubit 2.0 fluorimeter with a Qubit DNA High Sensitivity Assay Kits (Invitrogen, USA). NEBNext Ultra II DNA Library Prep Kit for Illumina and NEBNext Multiplex Oligos for Illumina (Dual Index Primers Set 1) (NEB, UK) were used to obtain DNA libraries from amplicons. The resulting libraries were quantified using the Kapa SYBR Fast Universal qPCR Kit (KapaBiosystems, USA). Library sequencing was performed using the MiSeq Reagent Standard Kit v3 PE300 on Miseq (Illumina, USA) at the Central Collective Use Center "Genomics" (ICBFM SB RAS).

The resulting paired reads were analyzed with UPARSE scripts (Edgar, 2013) using Usearch v11.0.667 (Edgar, 2010). Bioinformatics analysis included overlapping paired reads, filtering by quality and length, accounting for identical sequences, discarding singletons, removing chimeras, and obtaining an OTU (operational taxonomic unit). The taxonomic identification of the OTU sequences was determined using the SINTAX algorithms (Edgar, 2016) and the reference database MIDORI_UNIQUE_20180221_COI_SINTAX (Machida et al., 2017), as well as BLAST. The representativeness of OTU samples was analyzed using iNEXT 2.0.15 (Hsieh et al., 2016). The nucleotide sequences of the OTUs identified were tested for the presence of stop codons using the SeqKit v0.16.1 (https://doi.org/10.1371/journal.pone.0163962).

The Spearman correlation coefficient (S) was calculated to assess the relationship between the initial abundance of organisms of higher taxa in samples (before DNA extraction) and the number of reads of the *COI* gene fragment.

Communities of macroinvertebrates were identified by clustering the OTUs of the species rank and their *COI* reads by the Ward method; the Euclidean metric was used as a distance measure. Dominants and subdominants in the communities were determined by descending ranking of the modified density index (occurrence of OTUs in samples and their relative abundance = reads, in %) (Brotskaya, Zenkevich, 1939; Konstantinov, 1986). The communities of macroinvertebrates were named according to the dominant species with the highest density index. Species with a density index of ≥10 % were classified as subdominants. Species with a density index ≤10 % were considered minor. To characterize the degree of complexity of the community structure, the Shannon diversity (H, bit), Simpson dominance (D), and Pielou evenness (e) indices were used (Odum, 1983).

The spatial distribution of dominants and subdominants of communities in Bolshiye Koty Bay was analyzed using the principal component method, the relative abundance of OTUs and the composition of bottom sediments were used as variables. The data for calculations using multivariate statistics were previously transformed by the log(X+1) function. Calculations were carried out using the statistical programming environment R 3.0.0, package vegan 2.0-7.

Results

A total of 118,009 reads of the *COI* gene fragment (at least 313 bp in length) were obtained using high-throughput sequencing. Bioinformatics analysis showed that the proportion of unclassified *COI* sequences was less than 1 % of the total (1,157). 116,852 reads account for 115 OTUs belonging to the higher taxa of macroinvertebrates: Porifera – 1, Platyhelminthes – 3, Annelida – 38, Arthropoda – 55, Mollusca – 18.

The Spearman correlation coefficient (S) between the abundance of macroinvertebrates of higher taxa (Platyhelminthes, Hirudinea, Polychaeta, Oligochaeta, Isopoda, Amphipoda, Trichoptera, Chironomidae, Bivalvia, Gastropoda) found in samples before DNA extraction with the number of reads of the *COI* gene fragment is 0.6 ($p < 0.05$), which allows us to take the indices of the latter as equivalent to the relative abundance of organisms.

Forty-six taxa of species and genus ranks with ≥98 % and ≥95 % homology with GenBank sequences, respectively, were identified with coverage of at least 90 % (Table 1). They account for 88 % of the *COI* gene fragment reads.

Table 1. Composition of OTUs with ≥ 95% homology to GenBank reference sequences

Taxon	Species	GenBank numbers of reference sequences	Homology to the GenBank reference sequences, %
Porifera	*Baikalospongia fungiformis* Mak.	MH985288	99.7
Polychaeta/Fabriciidae	*Manayunkia baicalensis* (Nusb.)	MK393734	97.6
	Manayunkia godlewskii (Nusb.)	MK393737	99.3
	Manayunkia zenkewitschii Sit., Shcherb. et Kharch.	KF289863	99.2
Oligochaeta	*Rhynchelmis alyonae* Mart., Ferrag. et Kayg.	GU328670	98.3
	Rhynchelmis sp.	AJ577632	95.0
Amphipoda	*Baikalogammarus pullus* (Dyb.)	FJ756303	99.7
	Eulimnogammarus cyaneus (Dyb.)	MK887720	99.7
	Eulimnogammarus verrucosus (Gerstf.)	MK887569	100.0
	Eulimnogammarus vittatus (Dyb.)	MK887750	99.7
	Linevichella vortex (Dyb.)	MN148355	99.7
	Micruropus whalii (Dyb.)	MN148354	99.7
Diptera/Chironomidae	*Microtendipes* sp.	LC329125	95.0
	Paratanytarsus baicalensis (Tshern.)	MT020734	99.3
	Sergentia baicalensis Tshern.	AF116586	99.0
	Sergentia conf. *affinis*	AF116588	96.5
	Cricotopus fuscus (Kieff.)	MN673037	98.6
	Cricotopus conf. *intersectus*	MN683031	96.9
	Cricotopus sylvestris (Fabr.)	KC250789	100.0
	Cricotopus triannulatus (Macq.)	KJ439943	99.6
	Diplocladius cultriger Kieff.	HQ941599	98.6
	Orthocladius gregarius Linev.	KC879234	100.0
	Orthocladius nitidoscutellatus Lundstr.	MT048130	98.6
	Orthocladius sp.	KT248920	100.0
	Orthocladius sp. 2	KT248920	100.0
	Paratrichocladius rufiventris (Meig.)	HQ941597	99.3
Trichoptera	*Apatania majuscula* MacLach.	KX103052	99.3
	Baicalina bellicosa Mart.	KR153132	99.7
	Baicalina thamastoides Mart.	KR153144	99.6
	Baicalinella foliata (Mart.)	KR153101	99.7
	Protobaicalina spinosa (Mart.)	KR153124	98.9
Mollusca/Bivalvia	*Pisidium casertanum* (Poli)	KF483386	99.6
	Pisidium henslowanum (Shepp.)	KF483398	98.9
	Pisidium sp.	KF000182	99.3
	Pisidium sp. 1	KF483372	95.0
Mollusca/Gastropoda	*Benedictia fragilis* W. Dyb.	KX241839	99.3
	Choanomphalus amauronius Bourg.	Y14721	98.0
	Choanomphalus conf. *anomphalus*	Y14714	95.4
	Choanomphalus conf. *maacki*	LC429414	95.3
	Choanomphalus conf. *schrencki*	Y14713	96.1
	Gerstfeldtiancylus conf. *roepstorfi*	KR822550	96.2
	Maackia herderiana (Lindh.)	KY697388	99.7
	Megalovalvata baicalensis (Gerstf.)	LC377798	99.7
	Pseudancylastrum sibiricum (Gerstf.)	KR822557	100.0
	Pseudobaikalia contabulata (W. Dyb.)	Z92987	99.7
	Radix auricularia (Linn.)	MH190039	100.0

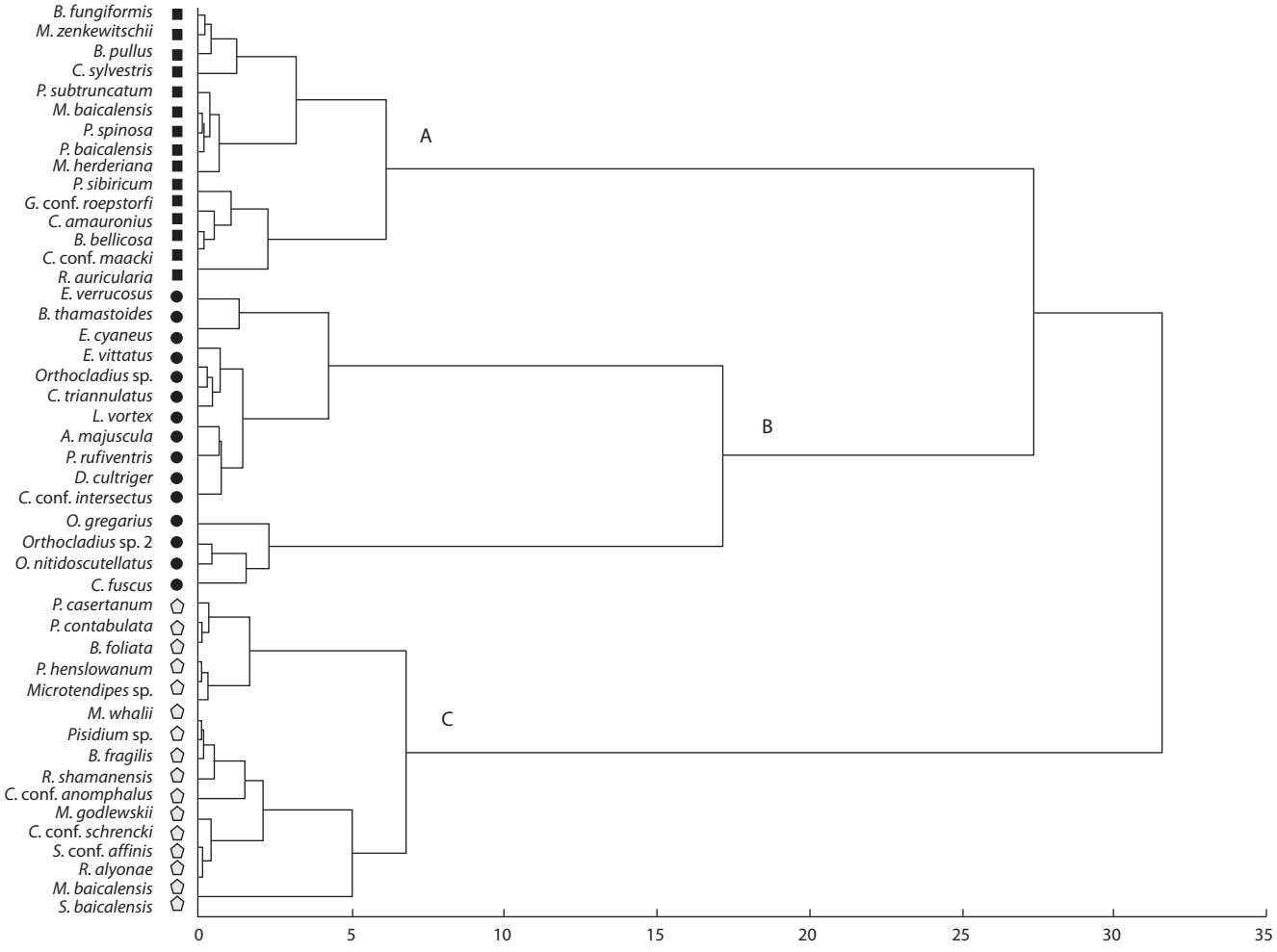

Fig. 2. Dendrogram of OTUs of macroinvertebrates constructed by the Ward clustering method using the Euclidean metric as a measure of distance. Clusters A, B, C on the dendrogram are macroinvertebrate communities identified based on reads of the *COI* gene fragment.

Table 2. Structural parameters of macroinvertebrate communities found in Bolshiye Koty Bay, Lake Baikal (July, 2019)

Community	Species number	Percentage of reads of dominants	H, bit	D	e	Number of reads per community
Orthocladius gregarius	15	96	1.20	0.45	0.44	39,941
Choanomphalus conf. *maacki*	15	93	0.74	0.70	0.28	32,573
Sergentia baicalensis	16	91	0.70	0.74	0.25	30,118

Note. H – species diversity according to Shannon, D – Simpson dominance index, e – Pielou's evenness index.

Three communities of macroinvertebrates were identified in the region studied, one of which was dominated by the molluscs *Choanomphalus* conf. *maacki* (A) and two were chironomid-dominated by *Orthocladius gregarius* Linev. (B), *Sergentia baicalensis* Tshern. (C) (Fig. 2).

Communities are characterized by a simple structure; 15–16 species were noted in their composition (Table 2). The Shannon index is low, ranging from 0.7 to 1.2 bits. At the same time, in the communities, there is a high concentration of dominance of one species (D varies from 0.5 to 0.7) and low evenness (e varies from 0.3 to 0.4).

In the community *C.* conf. *maacki* among the subdominants there are molluscs *Pseudancylastrum sibiricum* (Gerstf.), *Gerstfeldtiancylus* conf. *roepstorfi*. The subdominants of the *O. gregarius* community include the caddisflies *Baicalina thamastoides* Mart., chironomids *Orthocladius* sp. 2, *O. nitidoscutellatus* Lundstr., and *Cricotopus fuscus* (Kieff.). The community dominated by *S. baicalensis* contains numerous polychaetes *Manayunkia godlewskii* (Nusb.), molluscs *Choanomphalus* conf. *anomphalus*. Dominants and subdominants in communities account for 91 to 96 % of *COI* gene fragment reads.

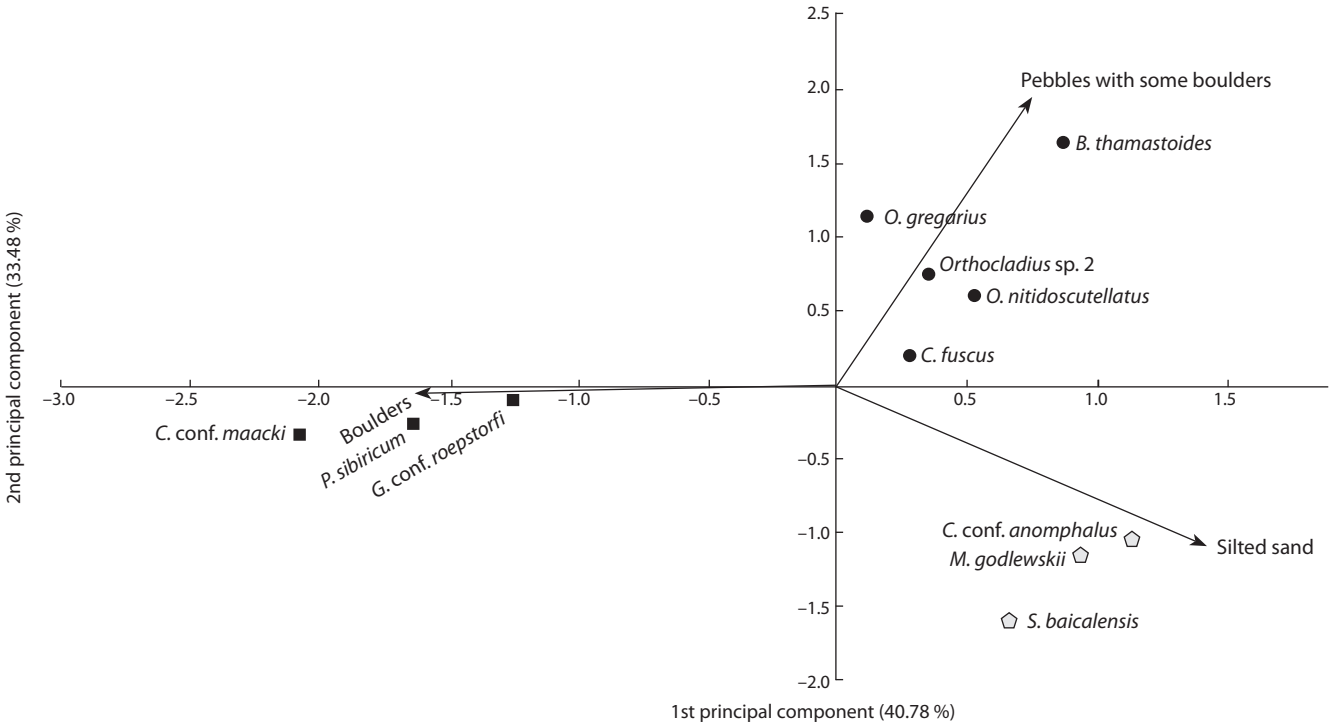

Fig. 3. Distribution of dominant species (OTUs) of macroinvertebrate communities (A–C) in the space of the first two principal components, taking into account 74 % of the variability of the relative abundance data set.

The spatial distribution of the dominant species of macroinvertebrate communities depending on environmental factors is shown in Figure 3. The first principal component characterizes the distribution of communities depending on the composition of bottom sediments. The dominant species of macroinvertebrate communities found on fine clastic material (pebbles, silted sand) have positive loads on the first principal component, and negative loads on coarse clastic material (boulders, rock fragments). The second principal component characterizes the distribution of dominants of the studied macroinvertebrate communities depending on the bottom geomorphology. The species found in the beach area have positive loads, while those found on the shallow-water terrace have negative loads.

Discussion

The studies of Bailey et al. (2001) indicate the effectiveness of monitoring observations of the state of water bodies using taxonomy at the genus or species level. Metabarcoding, as a modern genetic tool, is widely used for express assessment of biodiversity in ecosystems (Elbrecht, Leese, 2015). The authors of the work recommend using the presence/absence metric to characterize diversity, since the high resolution of primers makes it possible to take into account mass and minor species, but at the same time it does not make it possible to measure the absolute values of their abundance or biomass in samples. However, to study the structure of communities, indicators characterizing not only the composition of taxa, but also their quantitative ratio are needed. The number of reads, apparently, can be attributed to an indirect characteristic of species abundance (OTU). This confirms the presence of a positive correlation between the abundance of macroinvertebrates before DNA extraction and the abundance of OTUs from Lake Baikal not only in the samples collected in Bolshiye Koty Bay ($S = 0.6$, $p < 0.05$), but also in Listvennichny Bay ($S = 0.5$) (Kravtsova et al., 2021). In addition, a positive correlation was noted between the number of reads and the biomass of organisms collected in ponds in Germany (Elbrecht, Leese, 2015). Therefore, when studying the structure of macroinvertebrate communities, emphasis was placed on the number of reads per OTU, as well as the occurrence of OTUs in samples. To understand the extent to which the molecular genetics method using high-throughput sequencing technologies is applicable to the study of macroinvertebrate communities, the same approach was used for comparison as in ecological studies of water bodies based on classical hydrobiological methods (Brotskaya, Zenkevich, 1939; Begon et al., 1986; Konstantinov, 1986).

The fauna of macroinvertebrates on the shallow-water terrace and in the underwater part of the beach of the studied area of Lake Baikal in Bolshiye Koty Bay (excluding the population of the underwater slope and canyon) is quite diverse; in 1988, at least 177 species were recorded in its composition (Kravtsova et al., 2003). Most of the species found using DNA metabarcoding in 2019 were present here earlier, but in general, the diversity of benthic fauna (see Table 1) identified using molecular genetics methods is lower. This can be explained, on the one hand, by the smaller volume of collected quantitative samples of macrozoobenthos, and, on the other hand, by the absence in GenBank of *COI* sequences belonging to Baikal rich in species groups, for example, oligochaetes. In Baikal, 202 species of oligochaetes have been recorded, among them 165 are endemic (Semernoy, 2004). It is possible that the low representation of *COI* sequences of this group

of invertebrates in GenBank can be connected with the poor knowledge of the fauna due to its high endemism. In total, Annelida accounts for about 30 % of the total number of OTUs identified by DNA metabarcoding both from Bolshiye Koty Bay and Listvennichny Bay of Lake Baikal (Kravtsova et al., 2021).

Polychaeta, which are not rich in species, are also included in Annelida, but they were not previously included in the list of taxa from Bolshiye Koty Bay (Kravtsova et al., 2003), since their morphological identification had not been carried out. However, using DNA metabarcoding, in 2019, all three polychaete species found in Baikal were recorded in the fauna with high homology (98–99 %) with sequences from GenBank (see Table 1) (Pudovkina et al., 2016): *Manayunkia baicalensis* (Nusb.), *M. godlewskii*, *M. zenkewitschii* Sit., Shcherb. et Kharch.

Chironomids (Diptera) are among the objects, the identification of species of which by the morphological characters of the larvae is extremely difficult and they are often determined to a group of species (gr.) or species (sp.). It is possible that *O. gregarius*, the dominant of one of the three communities identified in 2019, was previously listed as *O.* gr. *thienemanni*, and the subdominant *O. nitidoscutellatus* was in *O.* gr. *olivaceus* (Kravtsova et al., 2003). Despite these species having been present in Bolshiye Koty Bay in 1988, they did not play a community-forming role among other macroinvertebrates (Kravtsova et al., 2004). The species *O. gregarius* and *O. nitidoscutellatus*, as shown by further molecular genetic studies (Kravtsova et al., 2014), have a rather long evolutionary history and their existence in Lake Baikal is beyond doubt (Makarchenko E.A., Makarchenko M.A., 2008).

The low diversity of species rank OTUs indicates that the macroinvertebrate communities dominated by *C*. conf. *maacki*, *O. gregarius*, *S. baicalensis* in Bolshiye Koty Bay are characterized by a simple structure (see Table 2) due to the above reasons.

It is known that abiotic environmental factors play an important role in the distribution and formation of the diversity of macroinvertebrate communities (Rezende et al., 2014). The spatial distribution of dominants and subdominants of macroinvertebrate communities from Bolshiye Koty Bay is consistent with the distribution of these species in the coastal zone of Lake Baikal. So, a community dominated by *C*. conf. *maacki* is confined, as before, to the rocky bottom sediments of the shallow terrace. The community of *O. gregarius* is distributed on pebbles with individual boulders in the beach zone, and with the dominance of *S. baicalensis*, on the silty sands of a shallow-water terrace, a typical biotope for this species. In Lake Baikal, representatives of the genus *Orthocladius* prefer to settle in a hydrodynamically active zone of wave mixing and water flow, while *Sergentia* prefer to settle in conditions where sedimentation processes prevail over erosion and transfer of terrigenous material, organic matter.

The study of the structure of macroinvertebrate communities in aquatic ecosystems using DNA metabarcoding has its own characteristics, in contrast to the rapid assessment of fauna diversity (based on the presence/absence metric). First of all, it is necessary to pay attention to the degree of study of the diversity of the bottom fauna of the reservoir, the size groups of its representatives. Of no small importance for the assessment of diversity is the presence of sequences of the studied gene fragment in the GenBank database. To obtain an adequate characterization of the relative abundance (reads of the *COI* gene fragment) of organisms in the community, quantitative sampling of macrozoobenthos should be carried out from a certain area, taking into account the biotopic heterogeneity of the bottom. In order to avoid the influence of body size of organisms on the number of reads per OTU, when preparing samples for DNA extraction, it is necessary to select tissue pieces of the same size from all individuals found in quantitative samples. This makes it possible to obtain an integral characteristic of the relative abundance of an organism (number of reads), taking into account its biomass, leveled by scatter (due to body size), as well as abundance. Since macroinvertebrates make up one third of the representatives of the unique Baikal fauna (2,565 species and subspecies (Timoshkin, 1997)), a more complete database of *COI* reference sequences is required for an objective assessment of α-diversity.

Conclusion

DNA metabarcoding using a primer (*mlCOIintF*) in combination with *jgHCO2198* to amplify the Folmer *COI* gene fragment showed its effectiveness in studies of the diversity and structure of Baikal macroinvertebrate communities. The fauna of Bolshiye Koty Bay contains typical representatives of most groups of macroinvertebrates inhabiting the coastal zone of Lake Baikal. It is shown that the number of reads, as a characteristic of the relative abundance of taxa, can be used to study the features of the structural organization of macroinvertebrate communities. In general, the proposed approach is quite acceptable for assessing the stability of macroinvertebrate communities in the temporal aspect in the monitoring of aquatic ecosystems.

References

Arbačiauskas K., Semenchenko V., Grabovski M., Leuven R., Paunović M., Son M., Csanyi B., Gumuliauskaitė S., Konopacka A., Nehring S., van der Velde G., Vezhnovetz V., Panov V. Assessment of biocontamination of benthic macroinvertebrate communities in European inland waterways. *Aquat. Invasions*. 2008;3(2):211-230. DOI 10.3391/ai.2008.3.2.12.

Aylagas E., Borja A., Rodríguez-Ezpeleta N. Environmental status assessment using DNA metabarcoding: towards a genetics based marine biotic index (gAMBI). *PLoS One*. 2014;9(3):e90529. DOI 10.1371/journal.pone.0090529.

Bailey R.C., Norris R.H., Reynoldson T.B. Taxonomic resolution of benthic macroinvertebrate communities in bioassessments. *J. North Am. Benthol. Soc.* 2001;20(2):280-286. DOI 10.2307/1468322.

Begon M., Harper J., Townsend K. Ecology: From Individuals to Ecosystems. Malden, MA: Blackwell Publishing, 1986. (Russ. ed.: Begon M., Harper J., Taunsend K. Ekologiya. Moscow: Mir Publ., 1989)

Bonada N., Dolédec S., Statzner B. Taxonomic and biological trait differences of stream macroinvertebrate communities between mediterranean and temperate regions: implications for future climatic scenarios. *Glob. Chang. Biol.* 2007;13(8):1658-1671. DOI 10.1111/j.1365-2486.2007.01375.x.

Brauns M., Garcia X.-F., Pusch M.T., Walz N. Eulittoral macroinvertebrate communities of lowland lakes: discrimination among trophic states. *Freshw. Biol.* 2007;52(6):1022-1032. DOI 10.1111/j.1365-2427.2007.01.

Brotskaya V.A., Zenkevich L.A. Quantitative accounting of the benthic fauna of the Barents Sea. *Proceedings of Russian Federal Research Institute of Fisheries and Oceanography*. 1939;4:5-98. (in Russian)

Burgmer T., Hillebrand H., Pfenninger M. Effects of climate-driven temperature changes on the diversity of freshwater macroinvertebrates. *Oecologia*. 2007;151(1):93-103. DOI 10.1007/s00442-006-0542-9.

Derycke S., Vanaverbeke J., Rigaux A., Backeljau T., Moens T. Exploring the use of cytochrome oxidase c subunit 1 (COI) for DNA barcoding of free-living marine nematodes. *PLoS One*. 2010;5(10): e13716. DOI 10.1371/journal.pone.0013716.

Doyle J.J., Dickson E.E. Preservation of plant samples for DNA restriction endonuclease analysis. *Taxon*. 1987;36(4):715-722. DOI 10.2307/1221122.

Edgar R.C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*. 2010;26(19):2460-2461. DOI 10.1093/bioinformatics/btq461.

Edgar R.C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods*. 2013;10(10):996-998. DOI 10.1038/nmeth.2604.

Edgar R.C. SINTAX, a Simple Non-Bayesian Taxonomy Classifier for 16S and ITS sequences. *bioRxiv.* 2016. DOI 10.1101/074161.

Elbrecht V., Leese F. Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass – sequence relationships with an innovative metabarcoding protocol. *PLoS One*. 2015;10(7):e0130324. DOI 10.1371/journal.pone.0130324.

Elbrecht V., Vamos E.E., Meissner K., Aroviita J., Leese F. Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. *Methods Ecol. Evol.* 2017;8(10):1265-1275. DOI 10.1111/2041-210X.12789.

Folmer O., Hoeh W.R., Black M.B., Vrijenhoek R.C. Conserved primers for PCR amplification of mitochondrial DNA from different invertebrate phyla. *Mol. Mar. Biol. Biotechnol.* 1994;3(5):294-299.

Geller J., Meyer C., Parker M., Hawk H. Redesign of PCR primers for mitochondrial cytochrome *c* oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Mol. Ecol. Resour.* 2013;13(5):851-861. DOI 10.1111/1755-0998.12138.

Gleason J.E., Elbrecht V., Braukmann T.W.A., Hanner R.H., Cottenie K. Assessment of stream macroinvertebrate communities with eDNA is not congruent with tissue-based metabarcoding. *Mol. Ecol.* 2021;30(13):3239-3251. DOI 10.1111/mec.15597.

Haenel Q., Holovachov O., Jondelius U., Sundberg P., Bourlat S.J. NGS-based biodiversity and community structure analysis of meiofaunal eukaryotes in shell sand from Hållö island, Smögen, and soft mud from Gullmarn Fjord, Sweden. *Biodivers. Data J.* 2017;5:e12731. DOI 10.3897/BDJ.5.e12731.

Hajibabaei M., Shokralla S., Zhou X., Singer G.A.C., Baird D.J. Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. *PLoS One*. 2011; 6(4):e17497. DOI 10.1371/journal.pone.0017497.

Hampton S.E., McGowan S., Ozersky T., Virdis S.G.P., Vu T.T., Spanbauer T.L., Kraemer B.M., Swann G., Mackay A.W., Powers S.M., Meyer M.F., Labou S.G., O'Reilly C.M., DiCarlo M., Galloway A.W.E., Fritz S.C. Recent ecological change in ancient lakes. *Limnol. Oceanogr.* 2018;63(5):2277-2304. DOI 10.1002/lno.10938.

Hebert P.D.N., Cywinska A., Ball S.L., DeWaard J.R. Biological identifications through DNA barcodes. *Proc. Biol. Sci.* 2003;270(1512): 313-321. DOI 10.1098/rspb.2002.2218.

Hsieh T.C., Ma K.H., Chao A. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods Ecol. Evol.* 2016;7(12):1451-1456. DOI 10.1111/2041-210X.12613.

Konstantinov A.S. General Hydrobiology. Moscow, 1986. (in Russian)

Kravtsova L.S., Karabanov E.B., Kamaltynov R.M., Mekhanikova I.V., Sitnikova T.Ya., Rozhkova N.A., Slugina Z.V., Izhboldina L.A., Weinberg I.V., Akinshina T.V., Krivonogov S.K., Shcherbakov D.Yu. Macrozoobenthos of subaqueous landscapes in the shoal of southern Lake Baikal: 1. Local diversity of bottom populations and features of their spatial distribution. *Zoologicheskiy Zhurnal = Zoological Journal*. 2003;82(3):307-317. (in Russian)

Kravtsova L.S., Kamaltynov R.M., Karabanov E.B., Mekhanikova I.V., Sitnikova T.Ya., Rozhkova N.F., Slugina Z.V., Izhboldina L.A., Weinberg I.V., Akinshina T.V., Sherbakov D.Yu. Macrozoobenthic communities of underwater landscapes in the shallow-water zone of southern Lake Baikal. *Hydrobiol*. 2004;522:193-205. DOI 10.1023/B:HYDR.0000029979.68265.3e.

Kravtsova L.S., Peretolchina T.E., Triboy T.I., Sherbakov D.Y. The evolutionary history of two species of Orthocladiinae (Diptera: Chironomidae) from Lake Baikal (Eastern Siberia). *Aquat. Insects*. 2014;36(3-4):171-185. DOI 10.1080/01650424.2015.1062111.

Kravtsova L.S., Peretolchina T.E., Triboy T.I., Nebesnykh I.A., Kupchinskiy A.B., Tupikin A.E., Kabilov M.R. The study of the diversity of hydrobionts from Listvennichny Bay of Lake Baikal by DNA metabarcoding. *Russ. J. Genet*. 2021;57(4):460-467. DOI 10.1134/S1022795421040050.

Kuntke F., de Jonge N., Hesselsøe M., Nielsen J.L. Stream water quality assessment by metabarcoding of invertebrates. *Ecol. Indic.* 2020; 111:105982. DOI 10.1016/j.ecolind.2019.105982.

Lacoursière-Roussel A., Howland K., Normandeau E., Grey E.K., Archambault P., Deiner K., Lodge D.M., Hernandez C., Leduc N., Bernatchez L. eDNA metabarcoding as a new surveillance approach for coastal Arctic biodiversity. *Ecol. Evol.* 2018;8(16):7763-7777. DOI 10.1002/ece3.4213.

Leray M., Yang J.Y., Meyer C.P., Mills S.C., Agudelo N., Ranwez V., Boehm J.T., Machida R.J. A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Front. Zool.* 2013;10(1):34. DOI 10.1186/1742-9994-10-34.

Machida R.J., Leray M., Ho S.-L., Knowlton N. Metazoan mitochondrial gene sequence reference datasets for taxonomic assignment of environmental samples. *Sci. Data*. 2017;4(1):170027. DOI 10.1038/sdata.2017.27.

Makarchenko E.A., Makarchenko M.A. New findings of chironomids (Diptera, Chironomidae, Orthocladiinae) in Far East and bordering territories. III. *Orthocladius* van der Wulp. *Evraziatskii Entomologicheskii Zhurnal = Euroasian Entomological Journal*. 2008; 7(2):243-262. (in Russian)

McGoff E., Aroviita J., Pilotto F., Miler O., Solimini A.G., Porst G., Jurca T., Donohue L., Sandin L. Assessing the relationship between the Lake Habitat Survey and littoral macroinvertebrate communities in European lakes. *Ecol. Indic.* 2013;25:205-214. DOI 10.1016/j.ecolind.2012.09.018.

Meusnier I., Singer G.A.C., Landry J.-F., Hickey D.A., Hebert P.D.N., Hajibabaei M. A universal DNA mini-barcode for biodiversity analysis. *BMC Genomics*. 2008;9(1):214. DOI 10.1186/1471-2164-9-214.

Moss B., Kosten S., Meerhoff M., Battarbee R.W., Jeppesen E., Mazzeo N., Havens K., Lacerot G., Liu Z., de Meester L., Paerl H., Scheffer M. Allied attack: climate change and eutrophication. *Inland Waters*. 2011;1(2):101-105. DOI 10.5268/IW-1.2.359.

Nalepa T.F., Fanslow D.L., Lang G.A. Transformation of the offshore benthic community in Lake Michigan: recent shift from the native amphipod *Diporeia* spp. to the invasive mussel *Dreissena rostriformis bugensis*. *Freshw. Biol.* 2009;54(3):466-479. DOI 10.1111/j.1365-2427.2008.02123.x.

Odum Yu. Basic Ecology. Philadelphia–New York–Chicago–San Francisco–Montreal–Toronto–London–Sydney–Tokyo–Mexico City–Rio de Janeiro–Madrid: Saunders College Publ., 1983. (Russ. ed.: Odum Yu. Ekologiya. Moscow, 1986)

O'Reilly C.M., Alin S.R., Plisnier P.-D., Cohen A.S., McKee B.A. Climate change decreases aquatic ecosystem productivity of Lake Tanganyika, Africa. *Nature*. 2003;424(6950):766-768. DOI 10.1038/nature01833.

Porazinska D.L., Giblin-Davis R.M., Faller L., Farmerie W., Kanzaki N., Morris K., Powers T.O., Tucker A.E., Sung W.A.Y., Thomas W.K. Evaluating high-throughput sequencing as a method for metagenomic analysis of nematode diversity. *Mol. Ecol. Resour.* 2009;9(6):1439-1450. DOI 10.1111/j.1755-0998.2009.02611.x.

Pudovkina T.A., Sitnikova T.Y., Matveyev A.N., Shcherbakov D.Y. Kindred relations of Baikal polychaete of the *Manayunkia* genus [Polychaeta: Sedentaria: Sabellidae] according to CO1 and settlement history analysis. *Russ. J. Genet. Appl. Res*. 2016;6(2):129-137. DOI 10.1134/S207905971602009X.

Rezende R.S., Santos A.M., Henke-Oliveira C., Gonçalves J.F., Jr. Effects of spatial and environmental factors on benthic a macroinvertebrate community. *Zoologia.* 2014;31(5):426-434. DOI 10.1590/S1984-46702014005000001.

Semernoy V.P. Oligochaeta of Lake Baikal. Novosibirsk: Nauka Publ., 2004. (in Russian)

Timoshkin O.A. Biodiversity of Baikal fauna: state-of-the-art (Preliminary analysis). In: New Scope on Boreal Ecosystems in East Siberia. Proc. of the Intern. Workshop, Kyoto, Japan, 23–25 Nov. 1994. Novosibirsk: Russ. Acad. Sci. Publ. Siberian Branch, 1997;35-76.

van den Berg M.S., Coops H., Noordhuis R., van Schie J., Simons J. Macroinvertebrate communities in relation to submerged vegetation in two Chara-dominated lakes. *Hydrobiologia*. 1997;342:143-150. DOI 10.1023/A:1017094013491.

Worrall T.P., Dunbar M.J., Extence C.A., Laize C.L.R., Monk W.A., Wood P.J. The identification of hydrological indices for the characterization of macroinvertebrate community response to flow regime variability. *Hydrol. Sci. J.* 2014;59(3-4):645-658. DOI 10.1080/02626667.2013.825722.

Yu D.W., Ji Y., Emerson B.C., Wang X., Ye C., Yang C., Ding Z. Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods Ecol. Evol.* 2012;3(4):613-623. DOI 10.1111/j.2041-210X.2012.00198.x.

ORCID ID

L.S. Kravtsova orcid.org/0000-0003-0862-4726
T.E. Peretolchina orcid.org/0000-0002-2950-9762
T.I. Triboy orcid.org/0000-0002-4830-0155
I.A. Nebesnykh orcid.org/0000-0001-5961-0514
A.E. Tupikin orcid.org/0000-0002-8194-0322
M.R. Kabilov orcid.org/0000-0003-2777-0833

Acknowledgements. The work was supported by the project of the Russian Foundation for Basic Research No. 19-05-00398_a. Collection and analysis of samples was partially carried out within the state-funded project conducted by LIN SB RAS No. 121032300196-8; bioinformatics analysis was partially supported by the state-funded project conducted by LIN SB RAS No. 121031300042-1.

Conflict of interest. The authors declare no conflict of interest.

Received February 20, 2023. Revised May 16, 2023. Accepted May 16, 2023.