

Article

Tracing the Origin of Planktonic Protists in an Ancient Lake

Natalia V. Annenkova^{1*}, Caterina R. Giner^{2,3}, Ramiro Logares^{2*}

¹Limnological Institute Siberian Branch of the Russian Academy of Sciences 3, Ulan-Batorskaya St., 664033, Irkutsk, Russia

² Institute of Marine Sciences (ICM), CSIC, Passeig Mar fim de la Barceloneta, 37-49, ES08003, Barcelona, Spain.

³ Institute for the Oceans and Fisheries, University of British Columbia, 2202 Main Mall, Vancouver, BC, V6T 1Z4, Canada

* Correspondence: tasha.annenkova@gmail.com (NVA); ramiro.logares@icm.csic.es (RL)

Abstract

Ancient lakes are among the most interesting models for evolution studies, because their biodiversity is the result of a complex combination of migration and speciation. Here, we investigate the origin of single celled planktonic eukaryotes from the oldest lake in the world – Lake Baikal. By using 18S rDNA metabarcoding we recovered 1,427 Operational Taxonomic Units (OTUs) belonging to protists populating surface waters (1-50 meters) and belonging to pico- and nano-plankton size fractions. The recovered communities resembled other lacustrine freshwater assemblages found elsewhere, especially the unclassified protists. However, our results suggest that a fraction of Baikal protists originated evolutionary recently from marine/brackish ancestors. Moreover, our results suggest that rapid radiation may have occurred among some protist taxa, partially mirroring what was already shown for multicellular organisms in Lake Baikal. We found potential species flocks in Stramenopiles, Alveolata, Opisthokonta, Archaeplastida, Rhizaria and Hacrobia. Putative flocks predominated in Chrysophytes, which are highly diverse in Lake Baikal. Some of the species, especially those from these flocks, may be endemic from Lake Baikal, because their 18S rDNA differed > 10% from known DNA. Overall, our study points to novel diversity of planktonic protists in Lake Baikal, some of which may have emerged *in situ* after evolutionary diversification.

Keywords: microbial eukaryotes; phylogeography; marine-freshwater transitions; evolutionary radiation; species flocks; Lake Baikal

1. Introduction

Even though single cell eukaryotes are fundamental for ecosystem function, a substantial fraction of their diversity remains poorly known. During the last decade, multiple studies have discovered new protist diversity, which was in part possible due to culture-independent molecular approaches using high-throughput sequencing (HTS) and metabarcoding [1,2]. Metabarcoding is highly relevant for pico/nano-sized (~0.8-5 μm , as defined in [3]) unicellular eukaryotes, as they are too small to show unambiguous morphological differences and are difficult to culture [3,4]. The key ecological role of these minute eukaryotes was shown in both oceans [3,5] and lakes [6,7]. Yet, we are still beginning to understand their diversity, evolution and dispersal.

In contrast to marine protists, we know less about the diversity of freshwater counterparts. Pico/nano-sized freshwater protists have been studied using Sanger sequencing in European ponds, lakes [8,9] and Lake George in USA [10]. More recently, small protists were investigated using metabarcoding in various lakes of Western Europe [11,12], as well as a few lakes in North- [13] and South-America [14]. Most of these studies pointed to diverse patterns of community structure and novel diversity. Yet, the previous reflects another bias: most studies of freshwater protists correspond mainly to Northern Europe and North America [15]. In contrast, other continents, like Asia, remain underexplored in terms of freshwater protist diversity, apart from a few lakes in China [16] and in the high-mountain Himalayan [17].

Lake Baikal (Siberia, Northern Asia) is the oldest (more than 26 million years), largest (23,615 km^3) and deepest (down to 1,642 m depth, with mean depth 744 m) freshwater lake in the world (Figure 1). As a comparison, the oldest basin of Lake Tanganyika is 9-12 million years old and the Great Lakes in North America began to form around 14,000 years ago. Baikal present-day communities inhabit the whole water-column and include more than 3,500 known species of eukaryotic organisms, of which more than half are endemic [18]. Such biodiversity is the result of a complex process including *in situ* evolutionary diversification as well as immigrations or extinctions in part related to geologic activity and climate change. Few relict groups are 70-30 M years old, while other organisms appeared 30 - 3.5 M years ago during the isolation of the Baikal depression. Furthermore, various taxa explosively developed 3.5 – 0.15 M years ago in response to the most dramatic tectonic changes in the region coupled to climate cooling [19]. One of the common mechanisms that lead to speciation within multicellular organisms in Lake Baikal was adaptive radiation, which resulted in endemic species flocks [18,20].

Compared to its eukaryotic macrobiota, much less is known on the diversity, evolution and ecology of Lake Baikal eukaryotic microbiota. Routine monitoring of the plankton in Lake Baikal has been carried out using light microscopy for more than 60 years [21,22,23]. These studies showed seasonally recurrent changes in phytoplankton as well as long-term changes that may be explained by global warming. Furthermore, the previous studies pointed to phytoplankton spatial heterogeneity in spring, indicating also a high productivity of picophytoplankton (up to 60% of phytoplankton biomass in some years). Microscopy studies of Baikal protists

indicated both cosmopolitan freshwater species as well as examples of endemism [e.g. 24,25]. More detailed studies of diatoms [26] and dinoflagellates [27] revealed the existence of species flocks. Recently, two studies have investigated the diversity of protist communities in Lake Baikal using metabarcoding. The first study [28] provided insights on the genetic diversity of protists from the southern basin of Lake Baikal (see Figure 1), pointing to novel diversity. Another study investigated co-occurrence networks of bacteria and microbial eukaryotes during the spring season [29] indicating positive correlations between specific OTUs as well as the influence of geography on microbial community structure. Despite the insights provided by these studies, more efforts are needed in order to understand protist diversity and evolution in Lake Baikal. The main reasons are related to the lake huge size, which provides many different niches in upper or deeper waters, in open waters or bays, as well as the lake's age, which, different to most widely-studied younger lakes, has allowed evolutionary diversification to take place *in situ*. Such evolutionary analyses, so far barely considered in previous studies, should provide important information about endemism in Lake Baikal protists as well as protist phylogeography in fresh waters.

The aim of this study is to investigate the origin of the planktonic protists in lake Baikal populating the pico-nano size fraction (0.4-8 μm in size). To do so, we analyzed the phylogeographic patterns, in particular, the evolutionary closeness of Baikal protists to known freshwater or marine taxa from all over the world as well as the potential existence of evolutionary radiations that may have led to species flocks and endemism.

2. Methods

2.1 Sample collection, DNA extraction and sequencing of 18S rRNA genes

Plankton samples were collected in July 2013 using a niskin bottle from 14 sites in Lake Baikal, Russia (Figure 1). In three sites located near the coast (depth was less than 40m), samples were taken at 5 m depth and in one coastal site (ICC) samples from 1 and 20 m were mixed. The rest of the sites were located in open Baikal waters, and samples were taken from 1, 15 and 50 m depth. These are typically used sampling depths for surface samples in Lake Baikal. Two samples were taken from 25 meters instead of 15 meters due to sampling logistics. Water samples were kept in dark and cold conditions before filtration. To collect microbial DNA, up to 5 L of water were pre-filtered through 70 μm mesh-size net to remove large non-protistan organisms and then sequentially filtered through 8 μm and 0.45 μm polycarbonate membrane filters (Sartorius, Göttingen, Germany) using a vacuum pump. Filters were immediately frozen at -20 $^{\circ}\text{C}$ and were stored at -70 $^{\circ}\text{C}$ until DNA extraction.

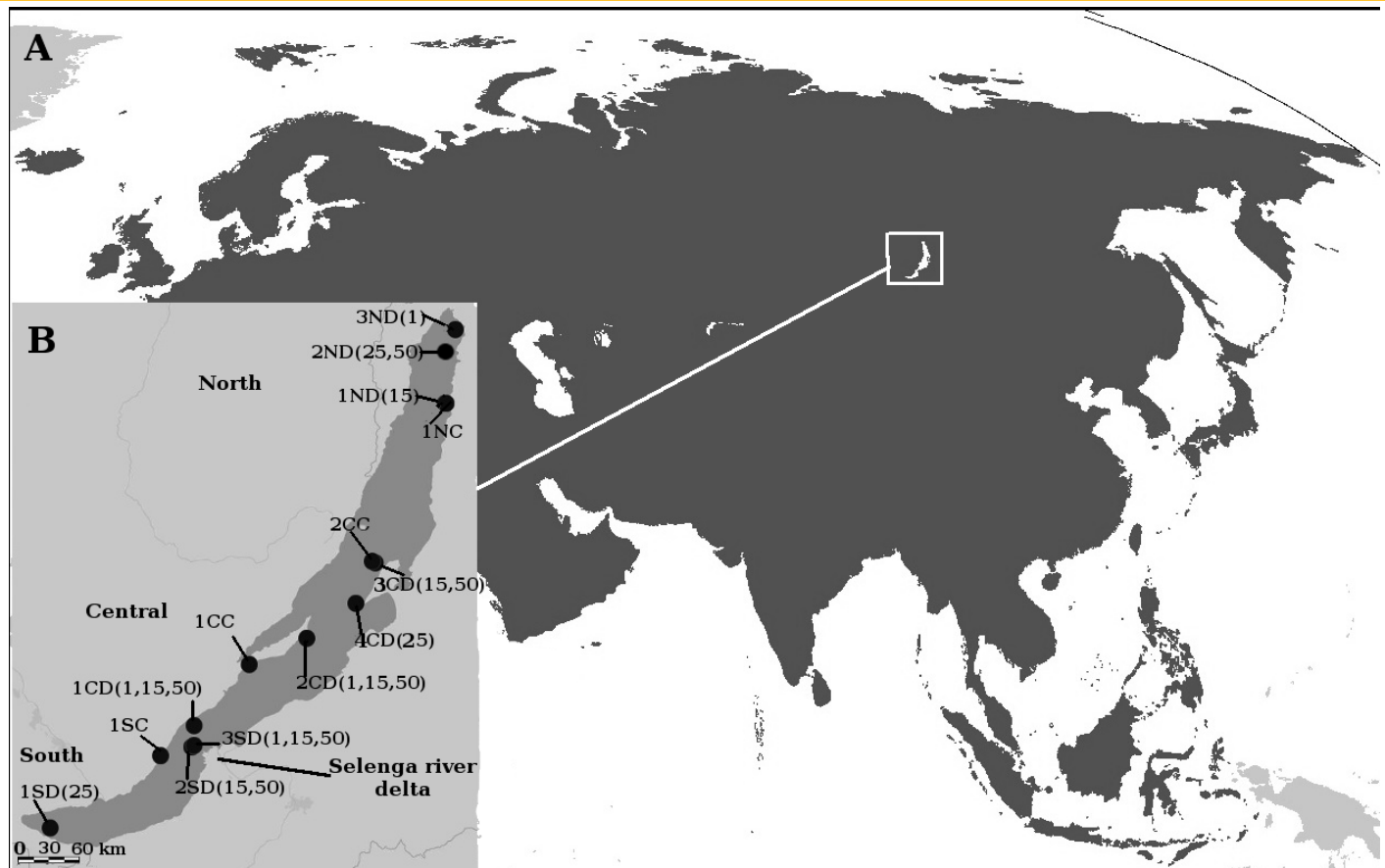


Figure 1. A. Lake Baikal position in Eurasia, B. Sampling sites in Lake Baikal. For each sample, the first number and letter indicates the number within one of the basins (S- South, C -Central, N -North), the second letter indicates position [coastal (C) , deep (D; more than 100m depth)] and numbers indicate sample depth (e.g. 1CD(1, 15, 50) – means, that samples from 1, 15 and 50 meters depth were taken in that site). No depth shown for the coastal samples.

DNA was extracted using the DNeasy Plant Kit (QIAGEN GmbH, Hilden, Germany) following the manufacturer's instructions. The eukaryotic universal primers TAREukFWD1 (5'-CCAGCASCYGC GGTAAT) and TAREukREV3 (5'-ACTTTCGTTCTTGATYRA-3') were used for PCR amplification of the 18S [V4 region; ~380 bp] [30]. PCR products were sequenced at Fasteris (<https://www.fasteris.com>) on an *Illumina* MiSeq (2x250bp). Overall, 23 samples were successfully sequenced.

2.2 V4-amplicon data processing and statistical analyses

Reads were processed following an in-house pipeline [31]. Briefly, raw reads were corrected using BayesHammer [32] following the algorithm from [33]. Corrected paired-end reads were subsequently merged

with PEAR [34] and sequences > 200 bp were quality-checked (maximum expected errors 0.5) and de-replicated using USEARCH [35]. OTUs were delineated at 97-99% similarity using UPARSE V8.1.1756 [36]. To obtain OTU abundances, reads were mapped back to OTUs at 99% similarity using an exhaustive search (-maxaccepts 20 -maxrejects 100,000). Chimera check and removal was performed both *de novo* and using the SILVA reference database [37]. Taxonomic assignment of 18S OTUs was generated by BLASTing [38] OTU-representative sequences against four reference databases, SILVA123 [37], PR² [39] and two in-house marine protist databases (available at <https://github.com/ramalok>), one based on a collection of Sanger sequences from molecular surveys [MAS database] [40] and the other based on 454 reads from the BioMarKs project [41]. BLAST hits were filtered prior to taxonomy assignment using an in-house python script, considering a coverage >70%, a minimum alignment length of 200 bp and an e-value > 0.00001. OTUs whose representative sequence had a percentage of identity <90% with its best BLAST hit, were taxonomically assigned based on phylogenetic analyses. Metazoan and nucleomorph OTUs were excluded from phylogenetic analyses. Computing analyses were performed at the MARBITS bioinformatics platform of the Institut de Ciències del Mar (ICM; <http://marbits.icm.csic.es>). Sequences are publicly available at the European Nucleotide Archive (<http://www.ebi.ac.uk/ena>; accession numbers PRJEB24415).

Statistical analyses of the communities were run in the R environment [42] using the package *Vegan*. Bray-Curtis dissimilarities were used as an estimator of beta-diversity. Differences in community composition between samples was analysed using NMDS analyses in *Vegan*. Venn diagrams were constructed with the *VennDiagram* package.

2.3 Phylogenetic analysis

All Baikal OTUs were divided into six superphyla (Stramenopiles, Alveolates, Rhizaria, Hacrobia, Opisthokonta, Archaeplastida) according to BLAST analyses using SILVA and MAS databases. When downstream phylogenetic analyses showed uncertain placement of an OTU, we manually rechecked its superphylum affiliation. Sequences were aligned using MAFFT with E-INS-i strategy [43] and manually checked. TrimAl [44] was run to remove ambiguous regions from the alignment. Maximum Likelihood (ML) phylogenetic trees based on the generated alignments were inferred using RAxML with GTRCAT model and default parameters [45]. Bootstrap values were estimated with 1,000 pseudoreplicates using the rapid bootstrapping algorithm.

To construct phylogenetic trees we used Lake Baikal OTUs, closely related sequences from NCBI-nr, other well-known members of each phylum, and environmental OTUs from European shallow freshwater environments [from 12]. In the first step of the phylogenetic analysis, Baikal OTUs and OTUs from [12] (marked as OTU_EU_Pond) were used. Those OTUs from [12], which did not cluster with Baikal sequences, were excluded. In the second step, the closest BLAST hits as well as representatives of all classes from each

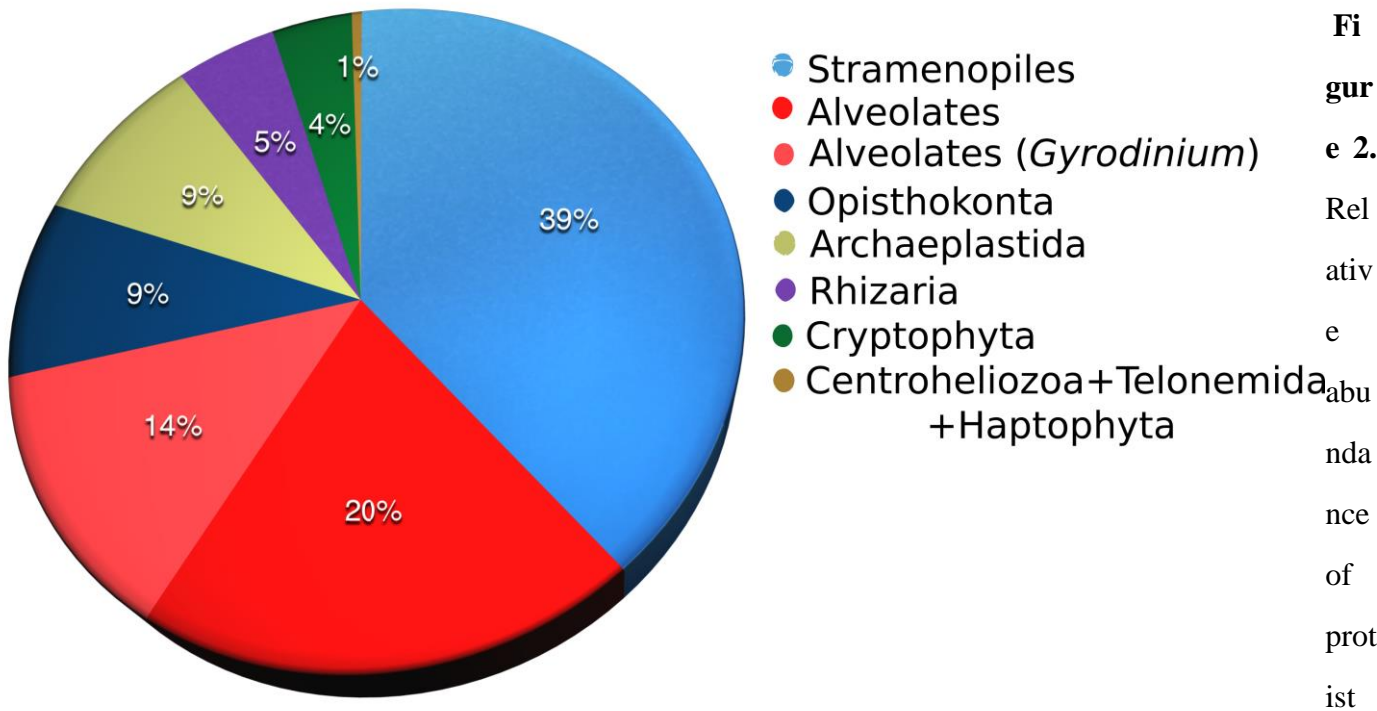
phylum were added to the analysis. Those representatives, which did not have any relationships with Baikal OTUs (e.g. raphidophytes) were excluded. In the last step additional members of classes, which showed similarity with Baikal OTUs, were added.

3. RESULTS

3.1 Protist composition

A total of 2 284 061 quality-filtered reads were obtained from 23 water samples (Figure 1). They grouped into 1,427 protistan and 37 metazoan OTUs delineated using a cut-off value of 99% of sequence identity. The metazoans belonged to known Baikal multicellular organisms (mainly copepoda). The protistan OTUs were analysed according to their relative abundance [46]. In total, 24 OTUs were abundant ($>1\%$ of all reads), 1,129 OTUs were rare ($<0.01\%$) and 317 OTUs had intermediate abundances ($0.01\% < \text{OTU} < 1\%$). Among the rare OTUs, only 344 OTUs were represented by less than 10 reads, while others were represented by up to 230 reads. The rarefaction curves (Figure S1) indicated that we approached to diversity saturation.

The phylogenetic diversity of protist OTUs covered Opisthokonta, Archaeplastida, SAR (Stramenopiles–Alveolata–Rhizaria) and taxa of uncertain position, that were combined in Hacrobia (Cryptophyta, Centroheliozoa, Telonemida, Haptophyta) (Figure 2A). Stramenopiles were the most diverse group (specially Chrysophyceae), encompassing 562 OTUs (39% of all protist OTUs, Figure 2A), being the second in abundance. Alveolates were the second most diverse, but the most abundant supergroup (up to 72% of reads per sample). However, its diversity may be overestimated due to the contribution of *Gyrodinium*-like sequences, which represented 40% of all alveolate sequences (Figure 2A): among them one OTU was represented by 214,236 reads in total and had 100% similarity with the known Baikal species *Gyrodinium helveticum*. The remaining 191 OTUs had less than 100 reads and had 95-99% similarity with *G. helveticum*. They likely represent variants of the rRNA gene, as dinoflagellates normally have many 18S rDNA copies per genome [47].



high-rank groups in Lake Baikal.

The prevalence of different taxa changed when considering the abundances of the reads or that of the OTUs (Figure 3). In particular, Eustigmatophyceae was represented only by two OTUs, but they corresponded to 3.3% of all reads. Prasinophytes had a similar abundance (3.1% of all reads), but they were more diverse (40 OTUs). Perkinsiids, Bicosoecida, Bolidophyceae, MAST and some fungi were rare in the samples, but included many different OTUs (Figure 3).

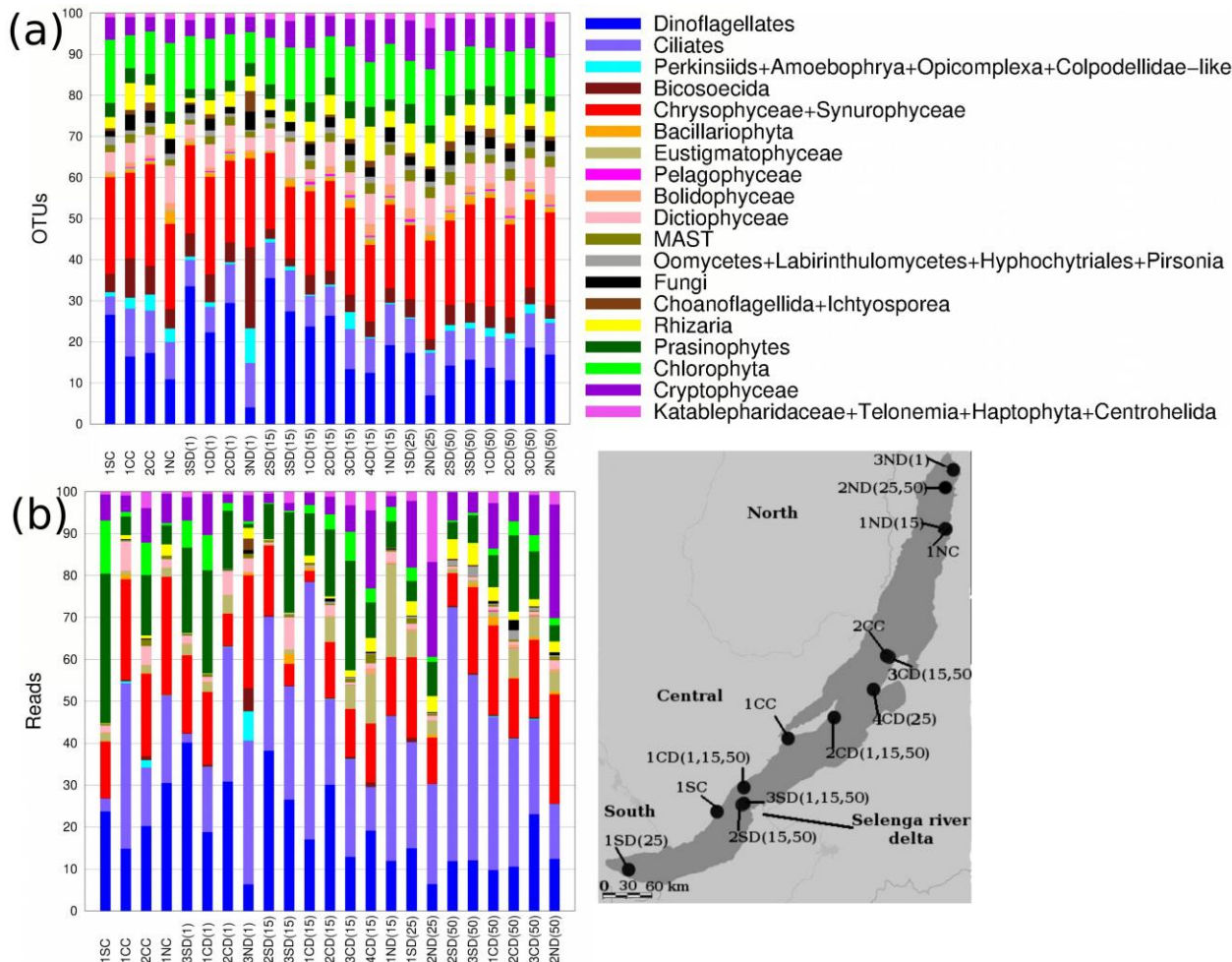


Figure 3. Histograms showing the relative proportion of 18S rDNA amplicon reads (A) and different OTUs (B), assigned to high-rank taxa across different sample sites.

3.2 Phylogenetic relationships

About 60% of Baikal OTUs clustered with environmental DNA sequences from diverse freshwater bodies (mainly lakes) with good statistical support (> 60% Bootstrap Support). These analyses indicated that Lake Baikal is populated by many protists that are evolutionary distantly related to cultured species, being instead related to unclassified microorganisms from diverse locations.

Interestingly, members of the so far exclusively-marine family Pelagophyceae were found in the Lake Baikal: two closely related OTUs (corresponding to 547 and 25 reads) clustered with marine *Ankylochrysis lutea* and two Arctic *Pelagophyceae* sp. (ML bootstrap support = 100) (Figure 4 group II, Figure S2). Specific taxa that are normally absent or infrequent in freshwater lakes were also found in Lake Baikal (Figure 4). In particular, a well-defined clade was composed by members of syndiniales and one Baikal OTU (Figure 4 group

V, Figure S3). The Baikal representative was very similar (99%) to an uncultured Amoebohyran from the brackish Arctic Ward Hunt Lake. This OTU was sequenced in four different stations in the Central Basin of Lake Baikal (totaling 308 reads).

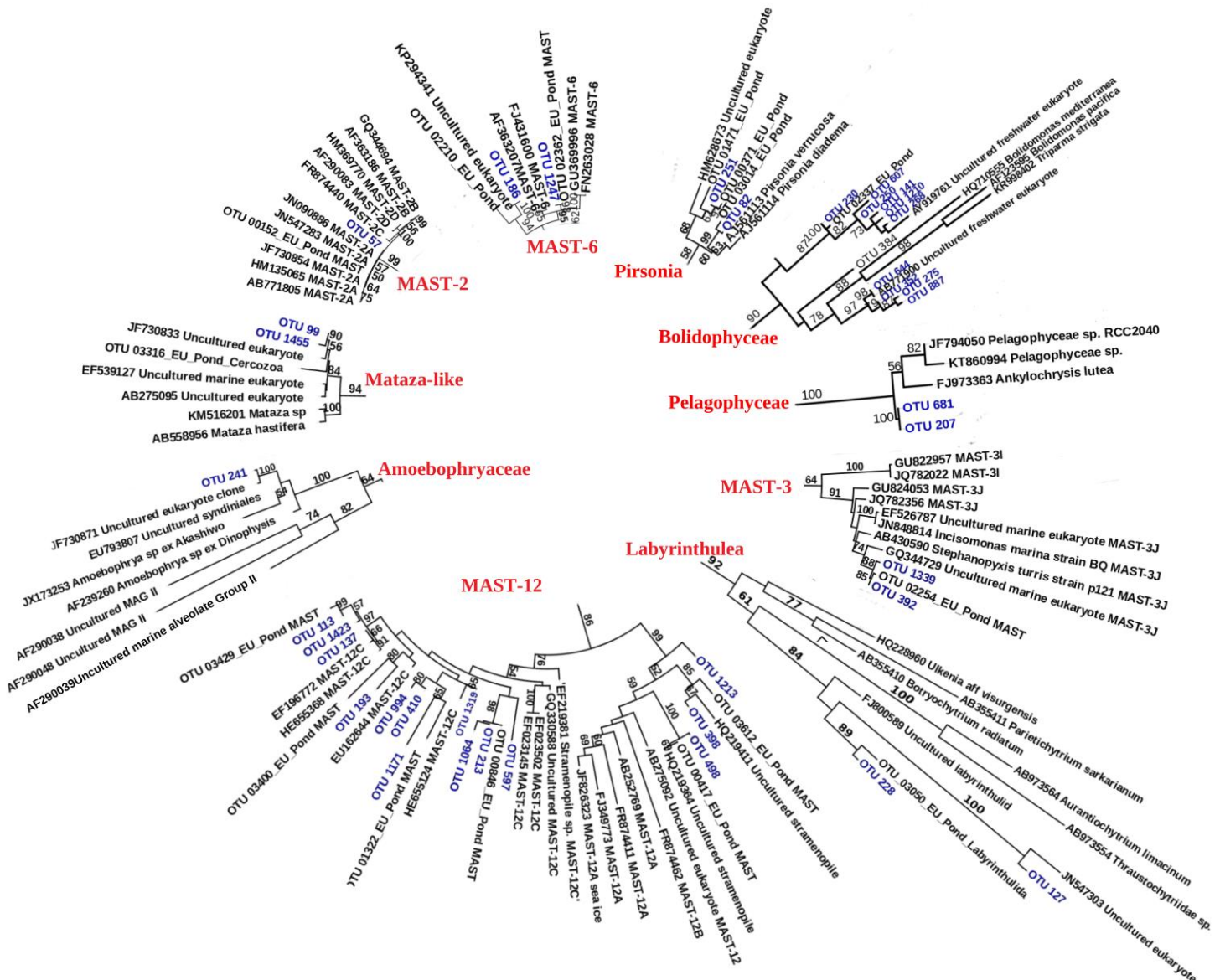


Figure 4. Portions of ML trees representing clades that are known mostly from marine waters and that include Baikal OTUs (in blue) [full versions of the trees are included in Supplementary Figures S2, S3, S4, S5]. Numbers next to branches indicate bootstrap support.

Among Stramenopiles, members of four MAST groups were found (Figure 4 group III, Figure S4). Baikal members of MAST-3J and MAST-6 were similar to sequences from a pond in the Natural Regional Park (France). Baikal MAST-2A was identical to a previously sequenced clone from Baikal (JN547283) and identical also to a sequence from the Lake Kusaki (Japan). MAST-12 was more diverse and included 11 OTUs

from MAST-12C and three OTUs from MAST-12A and -12B (Figure 4 group III-B, Figure S4). There were several OTUs which could also belong to MAST: one OTU clustered with MAST-6 and three OTUs clustered with MAST-24 with low bootstrap support (Figure S4). Labyrinthulomycetes, which are mostly marine protists, were represented in Baikal by two OTUs, one of which was similar to an unclassified freshwater sequence from French pond and another one that clustered with a sequence from a lake in Villerest (France) (Figure 4, group IV, Figure S4). A few OTUs had similarity with the Pirsonia clade, but bootstrap support was only 58 (Figure 4 group VII, Figure S4). Bolidophyceae, which are common in the oceans but mostly absent in freshwater environments [48], were represented by ten OTUs in Lake Baikal (Figure 4 group I, Figure S2). Lastly, two OTUs clustered with strong statistical support with marine cercozoa from the Mataza group (Figure group 4 VI, Figure S5).

Besides high rank taxa, “marine elements” at the genus/species level were identified in Lake Baikal. In particular, two Baikal ciliates (OTU_27, 16,574 reads and OTU_579, 25 reads) and one green algae (OTU_532, 35 reads) were very similar to counterparts from the Baltic Sea (Figures S6 and S7). Also, OTU_80 (3,848 reads) from Lake Baikal was identical to the sequence of an unclassified ciliate from the Bering Sea, while another Baikal ciliate (OTU_1062, 6 reads) was similar to an environmental marine sequence collected near Norway (Figure S4). One of the *Telonema*-like OTUs (OTU_73) clustered exclusively with *Telonema antarcticum* and several uncultured marine eukaryotes (Figure S8). Among them, OTUs from the brackish Arctic Ward Hunt Lake were identical to the OTU_73 from Lake Baikal, which was found in 18 samples and corresponded to 0.2% of all reads in Lake Baikal (4,495 reads in total). The diatom *Thalassiosira pseudonana*, known from brackish waters [49] was also found in our data (Figure S2). OTU_299 (155 reads; Figure S10) belongs to an organism with potential polar origin: it is identical to *Naganishia vishniacii* (*Cryptococcus vishniacii*), which is an enigmatic protist from Antarctic soil [50] and also was found in the meltwater from Tianshan glacier [51].

3.3 Presence of species flocks

We found 97 clades with bootstrap support > 60% that included 3 or more Baikal OTUs. Among them, 42 clades (hereafter, potential species flocks) contained Baikal OTUs with the following characteristics: 1) < 3% of sequence divergence and 2) the abundance of at least three OTUs within the flock differed < 100 times from each other. Most of these putative flocks were associated with Stramenopiles (Chrysophyceae, Bicosoecids, Bolidophyceae and Dictyochophyceae), but other supergroups also featured them: Choanoflagellida (Opisthokonta), Pansomonadida (Rhizaria), Cryptophyceae, Prasinophytes (Archaeplastida), Perkinsiids, Ciliata and Dinoflagellata (Alveolata) (Table 1).

Table 1. Potential species flocks in the Lake Baikal.

phylum	Taxon	Number of potential species flocks
Stramenopiles	Chrysophyceae	11
	Dictyochophyceae	5
	Bicosoecida	3
	Bolidophyceae	2
Alveolata	Ciliata	4
	Perkinsiidae	5
Opisthokonta	Choanoflagellatea	2
	Chytridiomycota	2
	Rozellida	1
Archaeplastida	Chlorophyta	4
Rhizaria	Thecofilosea	1
	Sarcomonadea	1
	Novel Clade 10	1
Hacrobia	Cryptophyceae	1

Chrysophyceae was the richest group with 11 such potential flocks in different genera (Figure 5, Figure S11). Most of them included 3-6 Baikal OTUs. However, one flock (92 bootstrap support) contained 53 Baikal OTUs, which displayed 2.2% sequence divergence on average, including also 6 unclassified protists from other locations (one from Arctic Char Lake, two from Columbia River estuary and one from a shallow French pond). A flock within group Pansomonadida (Rhizaria) contained 7 Baikal OTUs, with sequences differing from each other 2.4% in average (Figure 6A). These Rhizarian OTUs may represent an endemic Lake Baikal species flock as their similarity to other known sequences is < 90%. Five potential species flocks within Dictyophyceae, which was previously known in Lake Baikal by one member *Pseudopedinella* sp., are shown in Figure 6B. Lastly, we found 24 clades where the differences between OTUs was > 3%, yet other features (bootstrap support, OTU abundance) were comparable to those observed in the species flocks indicated above. In particular, bicosoecids contained 11 such clades (Figure S9).

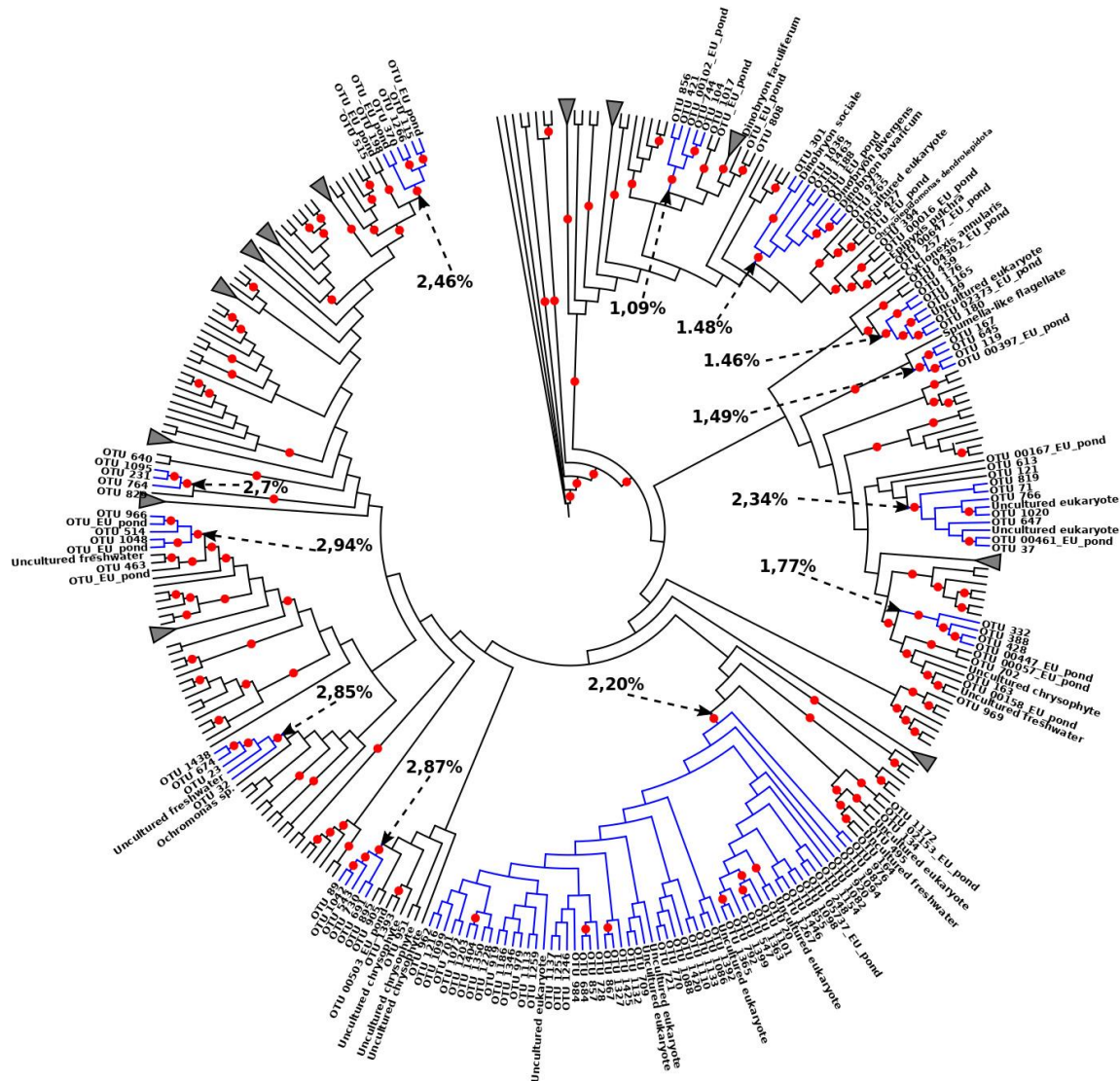


Figure 5. Reduced Chrisophyceae ML tree (See full version of the tree in Figure S11). Potential species flocks are indicated in blue. The percentages indicate the average divergence of the sequences within each flock. Red circles correspond to bootstrap supports > 60.

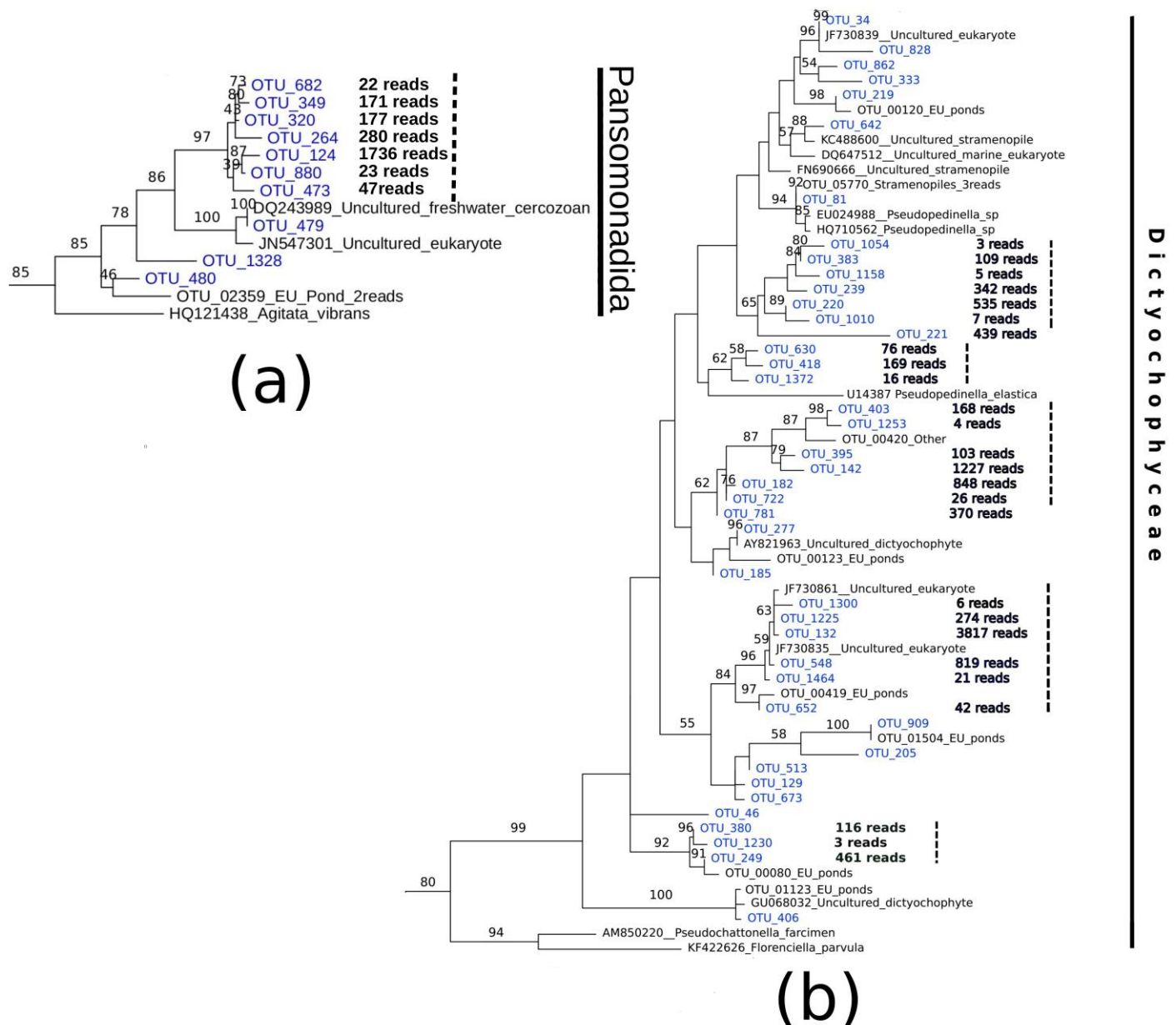


Figure 6. (a) Potential species flock in Pansomonadida (Rhizaria), which includes 7 OTUs differ > 10% from any known DNA sequence ; (b) Potential species flocks in Dictyochophyceae, (see full tree in Figure S3). Baikal OTUs are shown in blue. The amount of total reads, considering all samples, are shown for each Baikal OTU from the flocks. Flocks are shown by dashed line.

3.4 Potential endemic Baikal OTUs

About 8% of Lake Baikal OTUs had between 80 - 90% similarity to known sequences. Most of these OTUs were counted > 2 times within samples (from 3 to 1,848 reads). According to the phylogenetic analyses they belong to Stramenopiles (55 OTUs), Opisthokonta (26 OTUs), Alveolates (18 OTUs), Rhizaria (8 OTUs)

and Achaeplastida (1 OTU). Most of these potentially endemic OTUs clustered with sufficient statistical support with other taxa. Moreover, some of them were included within the mentioned putative species flocks and possibly originated via radiation, as may have occurred in Pansomonadida (Rhizaria, Figure 6A). Another example is represented by four novel alveolate OTUs (totaling 140 reads), which formed a separate group within the Gregarinasina clade (Apicomplexa, Figure S3, ML support values 100). Other potentially-endemic Alveolate OTUs belonged to perkinsiids, dinoflagellates and ciliates. Novel Stramenopile OTUs clustered mainly with bicosoecids (23 OTUs, Figure S9), but also with chrysophyceae (Figure S11) and MAST (Figure S4). OTU_153 (totaling 872 reads) had ~85% similarity to the closest reference sequence, which belongs to an unclassified marine organism from Ochrophyta, but it did not cluster with any other sequence in the phylogenetic analysis with sufficient bootstrap support (Figure S2). Potentially novel opisthokonts clustered with cryptomycota, choanomonada and chytridiomycota (Figure S10). One novel archaeplastid OTU was placed on the long branch within the family Mamiellophyceae (Prasinophytes) (Figure S7).

There were numerous Baikal OTUs that had > 90% similarity to known sequences, but did not cluster with them with good bootstrap support (>60). For example, OTU_154 (totaling 736 reads) was related to the Novel Clade 10 (Rhizaria), but did not cluster with other sequences within this clade (Figure S5). OTU_426 and OTU_327 clustered with Endomixa (Rhizaria), but with low bootstrap support, being located in the long branches (Figure S5). Several relatively common chlorophyte OTUs were related to Trebouxiophyceae with low support (Figure S7). Overall, such potentially novel Baikal OTUs could be found in most phyla.

3.5 Community composition

OTU richness per sample ranged between 200 – 400 OTUs with the exception of the community from the Northern basin (3ND(1)), which was more diverse, containing about 700 OTUs (Figure 7a). Community similarity was high (i.e. Bray Curtis dissimilarities below 0.37) in assemblages from similar depths influenced by the Selenga river (1SC and 3SD(1), 1SC and 1CD(1), 2SD(50) and 3SD(50); Figure 7b) as well as in assemblages from the same depths that were sampled from open waters in the South and Central basins (1SD(25) and 4CD(25); Figure 7b). Other assemblages differed substantially from each other (Bray–Curtis average ~ 0.7). Communities from deeper water layers (25-50 meters) were more similar to each other than to those from upper layers, even when they were taken in different geographic locations (Figure 7b).

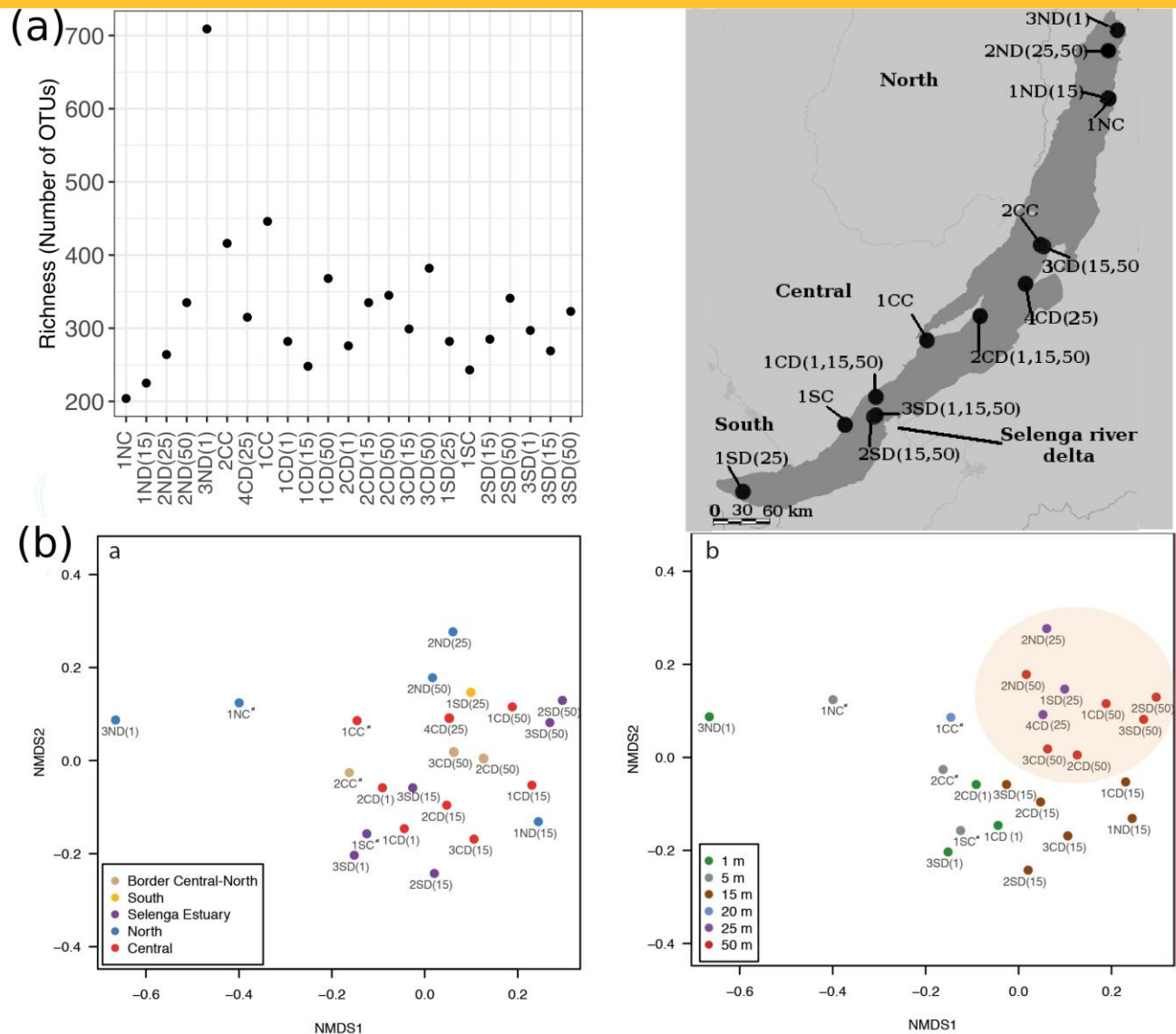


Figure 7. (a) OTUs richness in different samples with the map, showing locations of the samples. (b) NMDS analyses indicating the variable composition of protist assemblages within different basins (left) and different depth (right). The pink oval indicates samples from deeper waters (25 and 50m). Coastal samples marked by asterisk.

Many phylotypes were shared between the three Baikal basins (Figure S12). However, members of Bolidophyceae, Pelagophyceae and Choanoflagellida were not found in two samples that were taken near the Selenga river estuary (2SD(15) and 3SD(15)). Also members of Pelagophyceae were not found in samples near the Selenga river estuary (1SC and 3SD(1)), neither in samples from the North basin (1NC, 2ND(25) and 3ND(1)). Members of Choanoflagellida were absent as well in two Northern samples (1NC and 1ND(15)). The Northern basin contained many unique phylotypes (457 OTUs) due to the highly diverse 3ND(1) sample (Figure 7a, Figure S12).

4. Discussion

Most lakes in the world originated during the last glaciation, being <20,000 years old [53], therefore, our current knowledge on microbial populations inhabiting these waterbodies derive from research on geologically young habitats. Investigating microbial communities in ancient freshwater lakes such as Lake Baikal, which is ~26 million years old, can provide new insights on the evolution of microbes in freshwaters as well as on their immigration and local extinction history over millions years. Given that most lakes are younger than the majority of microbial lineages they contain, research on ancient lakes could provide crucial information that would help us understanding how the freshwater microbiota has evolved during the last million years and what evolutionary dynamics are prevalent.

A recent study using both plankton and benthic Baikal samples coupled to 454 DNA sequencing identified 644 protist OTUs (at 97% sequence similarity) and proposed the existence of novel lineages among parasites [27]. Our results indicate that Baikal protist communities are much richer. Only one fraction of the plankton (0.4 – 8 µm), collected during one week from a relatively small depth range (0-50 meters) contained between ~800 and ~1,400 OTUs (at 97% and 99% similarity respectively). For comparison, there were between 82 to 715 OTUs (at 97% similarity, 0.22 - 50 µm size-fraction, Illumina sequencing) per lake in one of the latest freshwater-protist metabarcoding studies including 227 Pyrenean mountain lakes [7]. Given its long history, and the fact that Lake Baikal could have served as a glacial refugia during the last glaciations [55], protists species may have accumulated in the lake more than in other younger waterbodies, thus partially explaining its higher richness. Furthermore, the complex relief of the lake (large bays, vast shoals, deep open waters, geothermal springs), provides multiple ecological niches for the immigrants. Rivers, also, have likely contributed to increase immigration into Lake Baikal (nowadays, ~300 rivers discharge into the lake) by bringing living cells or cysts. In particular, Popovskaya [56] has shown that the northern zone of Lake Baikal has a higher species richness of picophytoplankton due to rivers discharge. Concordantly, in our data, the sample from this northern part of the lake (3ND(1)) had the highest richness.

4.1 Freshwater and marine lineages in Lake Baikal

A substantial fraction of Baikal OTUs clustered with DNA sequences from other freshwater protists. This could be explained by past migrations, with Lake Baikal acting as a source and sink of immigrants. Given its potential role as a glacial refugia for freshwater protists, along with its huge size, Lake Baikal may have been an important source of immigrants that colonized the newly formed lakes after the last glaciation in Northern Eurasia and North America. These events may partially explain the presence of identical or highly similar protistan OTUs in Lake Baikal and in other younger glacial lakes.

Freshwater origin was established for most studied multicellular organisms in Lake Baikal (e.g. Baikal sponges, Baikal oilfish), but not for all of them. In particular, the neoendemic Baikal seal *Pusa sibirica* diverged

1.60–0.72 million years ago from a marine Arctic ancestor [57]. Similarly, besides freshwater protists, we found other protists that are evolutionary close ($\leq 1\%$ divergence in 18S rDNA) to brackish/marine species. These marine-derived Baikal protists occurred in various lineages, such as ciliates, dinoflagellates, chlorophyta, telonemids, fungi, pelagophyceae, amoebophrya and most of them could be traced back to Arctic waters or the Baltic Sea. This agrees with other results indicating close evolutionary relationships between specific bacteria from Lake Baikal and the Baltic Sea [58] as well as close relationships between dinoflagellates in Lake Baikal, the Baltic Sea and the Arctic area [27,59].

How arctic species overcame the salinity barrier [60] may be explained by two reasons. First, the salinity of polar seas fluctuates highly (nowadays and in the past) and could be low due to riverine discharge and ice melting [61]. Thus, their microbial inhabitants often have a wide salinity tolerance [62] that may allow them to colonize freshwater environments. In particular, we have found wide salinity tolerance in two Arctic ancestors of Baikal dinoflagellates [27,59]. Second, certain ecological niches in the pelagic zone of Lake Baikal were empty during the last cooling period in the Pleistocene due to drastic tectonic and climate changes [55], which may have promoted new “marine” invaders to settle in the lake. In sum, the planktonic protist assemblages from Lake Baikal seem to contain freshwater as well as marine/brackish species that may have crossed the salinity boundary relatively recently. Interestingly, the existence of taxa with close marine relatives in Lake Baikal has already been suggested [63], but such scenario was not popular among scientists. In the light of our results, the previous could have more relevance for protists than for multicellular organisms.

Additionally, we found in Lake Baikal certain high rank taxa that are considered predominantly marine (Figure 4). In particular, we found freshwater representatives of the group Pelagophyceae, which were close to other Pelagophyceae from the Beaufort Sea (Arctic). Marine groups such as Labyrinthulomycetes, various MAST, pirsoniids and Bolidophyceae were found both in ephemeral freshwater bodies [12,64], in large shallow eutrophic [16] and oligotrophic [10] lakes and in the ancient and stable oligotrophic waters of Lake Baikal. This suggest dispersal but perhaps no long-term colonization (i.e. in ephemeral waters) as well as cases of long-term colonization via adaptive evolution.

4.2 Do Lake Baikal protists radiate?

Adaptive (and possibly non-adaptive) radiation is one of the main mechanisms responsible for the ecological and phenotypic diversity of multicellular life [65, 66]. Yet, much less is known about the operation of this process in protists. Recent radiation has been proposed for dinoflagellates in Lake Baikal [26] as well as for a diatom species flock in Lake Ohrid [67]. Furthermore, the existence of species flocks in diatoms has been extensively discussed in a number of ecosystems [68,69,70]. In addition, based on metabarcoding data, novel radiations within Lake Baikal fungal parasites was suggested [28].

According to the definition of Schluter [71], adaptive radiations are characterized by four features in the

resulting species flock: 1) common ancestry, 2) rapid speciation, 3) phenotype-environment correlation, and 4) trait utility. Commonly, these flocks are endemic to a specific region, but this is not a mandatory feature [66]. The previous should be especially true for protists, which have high dispersal abilities. Here, we investigated 42 bootstrap-supported clades, which contained similar ($> 97\%$ identity in average) but not identical Baikal OTUs. We suggest that at least some of them may be regarded as species flocks that originated via radiation, though it is impossible to check based on DNA data the two last requisites previously mentioned [71]. The reasons supporting these species flocks are: 1) these OTUs constituted monophyletic clades with $> 60\%$ bootstrap support and their similarity indicated that they are closely related, yet distinct, taxa; 2) OTUs were located on relatively short branches, suggesting that all species within the flock may have originated over a short period of time. Most of our potential species flocks contained 3 to 10 OTUs, however one Chrysophyceae flock contained 53 OTUs. Moreover, Chrysophyceae was the taxon featuring most potential flocks (Table 1, Figure 5). This agrees with results from [72], who proposed based on SEM microscopy, that Lake Baikal is a hotspot of silica-scaled chrysophyte diversity.

An alternative explanation for the previous could be intragenomic or intercellular variation [47]. However, concerted evolution should eliminate different copies of the rRNA gene within genomes in evolutionary times [73]. Still, certain protist lineages may keep different copies of the 18S rDNA in a single genome [74,75]. Indeed, we found 192 OTUs in Lake Baikal, which were similar to the dinoflagellate *Gyrodinium helveticum*. Among them one OTU was abundant (214,236 reads in total) and identical to *G. helveticum* from Lake Baikal [76], while others were represented by < 100 reads. Potential intragenomic variability was also evidenced in Katablepharidaceae: seven OTUs were included in a well-supported clade, but one of them was found 30,958 times, being present in all samples. OTUs that were similar to it were much rarer (2-86 reads). A similar example was found within Cryptophyceae. Excluding previous cases suggesting intragenomic variation, we are still left with multiple examples that can be regarded as species flocks (Table 1).

In addition to the cases described above, we found monophyletic clades where the difference between Baikal OTUs was $> 3\%$. It is hard to say if their origin was continuous or if they represent ancient radiations. Our analytical approach was based on the conserved 18S rRNA gene, thus, we were not able to consider the most recent radiations, which may have happened < 1 million years ago, when a fraction of Baikal neoendemics, including protists, originated [27,59,77,78]. Nevertheless, we found potential species flocks in diverse lineages (Table 1) suggesting that evolutionary radiations could be as important for single-cell eukaryotes as for multicellular counterparts [20, 78].

4.3 Potential endemism within Baikal protists

The amount of endemism in microbiotas is a matter of debate [79]. Large ancient lakes can generate novel insights on the previous discussion, given that, different to most lakes, microbes have had the time to

diversify in them [e.g. 70]. Endemism in Lake Baikal has been already suggested based on scanning electron microscopy for certain relatively large protists [e.g. 24,27,67]. We found several types of 18S rDNA sequences in Lake Baikal that could belong to endemic organisms. These are: 1) OTUs displaying <90% similarity to other known protists, 2) OTUs that did not fall into clades with sufficient statistical support, and 3) OTUs that seem to have originated in the lake via adaptive radiation. These potentially-endemic OTUs were obtained from different samples and passed our strict quality filters, thus they should not constitute errors, but represent true diversity. A good example of endemism in Lake Baikal is likely represented by parasitic Gregarinozina (Apicomplexa). Six Baikal Gregarinozina species are known based on microscopy [24]. All of them are endemic parasites of endemic amphipodes, and three parasite species seem to belong to a species flock. We found seven OTUs belonging to Gregarinozina. Among them, four OTUs differed from other known DNA sequences by 11.8 – 13.7% and clustered into a putative species flock. Also, two additional OTUs differed from other known sequences by 6.2 and 13.0%. Thus, a total of six potentially endemic Baikal Gregarinozina including one species flock were independently predicted by microscopy and DNA metabarcoding.

4.4 Community structure of small Baikal protists

Baikal piconanoplankton has been predominantly associated with cyanobacteria and “some tiny flagellates”. Our approach allowed us to identify some of these “tiny flagellates” as highly diverse bicosoecids, choanoflagellates, specific cercozoa (like Novel clade 10), bolidophyceae, pelagophyceae, telonemids, perkinsiids and fungi-like parasites. Previously, it was shown that *Nannochloropsis limnetica* from Eustigmatophyceae plays a prominent role in Baikal picoplankton [80]. Indeed, this family was represented by many reads in most samples and corresponded to two different OTUs (one was *N. limnetica* and another was close to *Monodopsis subterranea*).

It has been shown that the structure of Baikal microbial communities in spring is affected by lake topology [28]. In our study, summer communities differed from each other, but we did not observe a clear correlation with lake's topology. However, we found that Baikal communities may be influenced by riverine inputs, as surface communities near the delta of the Selenga river clustered together. Also, the community from the 3ND site (Figure 7a), which is under the influence of two big rivers, had a higher species richness than the rest of the samples pointing to community coalescence [81]. In addition, we found that depth influenced community composition, as surface communities differed from those at 25-50 m depth. Previously it was shown that picophytoplankton decreased in the 20–30 m samples compared to the 5 m samples in July [80].

5. Conclusion

Lake Baikal is one of the most important lakes in the world considering its long ecological history and economic relevance. Its study can help understanding the evolution and community structure of lacustrine

microbiotas around the world, as well as their susceptibility to global change. Our results indicate that the richness of Lake Baikal protists is higher than in other lakes in temperate zones, which may reflect Lake Baikal long history or its role as a glacial refugia. In addition, we found that the Lake Baikal protist microbiota contains marine-derived species, pointing to relatively recent colonizations of its fresh waters by marine protists. The fact that these marine-derived protists have not been widely reported in other freshwater lakes suggests that these colonizers may have become locally adapted to Lake Baikal ecosystem. Last but not least, we found potential examples of evolutionary radiations within the lake that agree with other radiations reported for multicellular organisms in Lake Baikal. If future studies confirm these putative radiations, then Lake Baikal will become a paramount example of evolutionary radiations across the macrobial and microbial worlds.

Supplementary Materials: The following materials are available online at www.mdpi.com/xxx/s1, Figure S1: Rarefaction curves of all samples; Figure S2 – S11: Maximum Likelihood phylogenetic trees based on Baikal OTUs (in blue) and 18S RNA genes of other protists obtained from databases; Figure S12: Venn diagram showing the number of OTUs shared by Lake Baikal's basins or specific to some of them.

Author Contributions: Conceptualization, N.V.A. and R.L.; laboratory work N.V.A. and C.R.G.; sequence analysis R. L.; phylogenetic analyses N.V.A.; ecological analyses C.R.G., N.V.A., R.L.; data curation, R. L.; writing—original draft preparation, N.V.A.; writing—review and editing, N.V.A., R.L., C.R.G.; visualization, N.V.A., C.R.G. All authors have read and agreed to the published version of the manuscript.

Funding: N.V.A was supported by a grant of the Russian Science Foundation (18-74-00054). R.L. was supported by a Ramón y Cajal fellowship (RYC-2013-12554, MINECO, Spain). Lake Baikal expedition was supported by Russian state assignment to N.V.A. (0345-2016-0009). DNA sequencing was supported with research start-up funding to R.L. (RYC-2013-12554, MINECO, Spain).

Acknowledgments: We thank Dr. Purificación López-García for providing the list of representative DNA sequences from shallow water ponds in France. Bioinformatics analyses were performed at the MARBITS platform (ICM; <http://marbits.icm.csic.es>). We thank the captain and crew of the research vessel “G. Titov” for their help during the expedition.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Logares, R.; Haverkamp, T.H.; Kumar, S.; Lanzón, A.; Nederbragt, A.J.; Quince, C.; Kauserud, H. Environmental microbiology through the lens of high-throughput DNA sequencing: synopsis of current platforms and bioinformatics approaches. *J Microbiol methods* **2012**, *91*, 106-113.
2. Pedrós-Alió C.; Acinas, S.G.; Logares, R.; Massana, R. Marine microbial diversity as seen by high-throughput sequencing. In *Microbial Ecology of the Oceans*, 3rd ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2018. pp. 47-97, ISBN: 978-1-119-10718-7.
3. de Vargas, C.; Audic, S.; Henry, N.; Decelle, J.; Mahé F, Logares, R.; Lara, E.; Berney, C.; Le Bescot, N.; Probert, I. et al. Eukaryotic plankton diversity in the sunlit global ocean. *Science* **2015**, *348*, 1261605-126160.
4. Vaultot, D.; Eikrem, W.; Viprey, M.; Moreau, H. The diversity of small eukaryotic phytoplankton ($\leq 3 \mu\text{m}$) in marine ecosystems. *FEMS Microbiol Rev* **2008**, *32*, 795–820.
5. Massana, R. Eukaryotic Picoplankton in Surface Oceans. *Annu Rev Microbiol* **2011**, *65*, 91-110.
6. Lepère, C.; Domaizon, I.; Debroas, D. Unexpected importance of potential parasites in the composition of the freshwater small-eukaryote community. *Appl Environ Microbiol* **2008**, *74*, 2940-2949.
7. Ortiz-Álvarez, R.; Catalan, J.; Triadó-Margarit, X.; Camarero, L.; Casamayor, E.O.; Catalan, J. High planktonic diversity in mountain lakes contains similar contributions of autotrophic, heterotrophic and parasitic eukaryotic life forms. *Sci rep* **2018**, *8*, 4457.
8. Lefranc, M.; Thénot, A. Genetic diversity of small eukaryotes in lakes differing by their trophic status. *Appl. Environ. Microbiol.* **2005**, *71*, 5935–5942.
9. Šlapeta, J.; Moreira, D.; López-García, P. The extent of protist diversity: insights from molecular ecology of freshwater eukaryotes. *Proc R Soc Lond [Biol]* **2005**, *272*, 2073-2081.
10. Richards, T.A.; Vepritskiy, A.A.; Gouliamova, D.E.; Nierzwicki-Bauer, S.A. The molecular diversity of fresh-water picoeukaryotes from an oligotrophic lake reveals diverse, distinctive and globally dispersed lineages. *Environ Microb* **2005**, *7*, 1413–1425.
11. Oikonomou, A.; Filker, S.; Breiner, H.; Stoeck, T. Eukaryotic communities in a meromictic lake. *Environ Microbiol* **2015**, *17*, 2144-2157.

12. Simon, M.; Jardillier, L.; Deschamps, Ph.; Moreira, D.; Restoux, G.; Bertolino, P.; López-García, P. Complex communities of small protists and unexpected occurrence of typical marine lineages in shallow freshwater systems. *Environ Microb* **2015**, *17*, 3610–3627.
13. Cahoon, A. B.; Huffman, A. G.; Krager, M. M.; Crowell, R. M. A meta-barcoding census of freshwater planktonic protists in Appalachia–Natural Tunnel State Park, Virginia, USA. *MBMG* **2018**, *2*, e26939.
14. Lentendu, G.; Buosi, P. R.; Cabral, A. F.; Trevizan Segóvia, B.; Ramos Meira, B.; Lansac-Tôha, F. M.; Velho, L. F.; Ritter, C. D.; Dunthorn, M. Protist Biodiversity and Biogeography in Lakes From Four Brazilian River–Floodplain Systems. *J Eukaryot Microbiol* **2019**, *66*, 592-599.
15. Debroas, D.; Domaizon, I.; Humbert, J-F; Jardillier, L.; Lepère, C.; Oudart, A.; Tañ. N. Overview of freshwater microbial eukaryotes diversity: a first analysis of publicly available metabarcoding data. *FEMS Microbiol Ecol* **2017**, *93*, fix023.
16. Li, S.; Bronner, G.; Lepère, C.; Kong, F.; Shi, X. Temporal and spatial variations in the composition of freshwater photosynthetic picoeukaryotes revealed by MiSeq sequencing from flow cytometry sorted samples. *Environ Microbiol*, **2017**, *19*, 2286-2300.
17. Kammerlander, B.; Breiner, H-W.; Filker, S.; Sommaruga, R.; Sonntag B.; Stoeck T. High diversity of protistan plankton communities in remote high mountain lakes in the European Alps and the Himalayan mountains. *FEMS Microbiol Ecol* **2015**, *91*, DOI 10.1093/femsec/fiv010.
18. Timoshkin, O.A. Main Tendencies in Research of Ancient Lake Biodiversity; most Interesting Recent Discoveries in Biodiversity of Lake Baikal. In *Index of Animal Species Inhabiting Lake Baikal and Its Catchment Area*; Timoshkin, O.A., et al.(Eds.); Nauka: Novosibirsk, Russia, 2011 Volume 2, pp. 1423–1428
19. Mats, V.D.; Shcherbakov, D.Y.; Efimova, I.M. Late Cretaceous-Cenozoic history of the Lake Baikal depression and formation of its unique biodiversity. *Stratig Geo Correl* **2011**, *19*, 404-23.
20. Shcherbakov, D.Y. Molecular phylogenetic studies on the origin of biodiversity in Lake Baikal. *Trends Ecol Evol* **1999**, *14*, 92–94.
21. Hampton, S.E.; Izmet'seva, L.R.; Moore, M.V.; Katz, S.L.; Dennis, B.; Silow, E.A. Sixty years of environmental change in the world's largest freshwater lake – Lake Baikal, Siberia. *Glob Change Biol* **2008**, *14*, 1947-1958.

22. Popovskaya, G.I. Ecological monitoring of phytoplankton in Lake Baikal. *Aquat Ecosyst Health Manag* **2000**, *3*, 215-225.
23. Bondarenko N.A.; Logacheva N.F. Structural Changes in Phytoplankton of the Littoral Zone of Lake Baikal. *Hydrobiol J* **2017**, *53*, 16-24.
24. Obolkina, L.A. New species of the family Colepidae (Prostomatida, Ciliophora) from Lake Baikal. *Zool Zh* **1995**, *74*, 3-19.
25. Pronin, N.M. List of parasitofauna species. In *Lake Baikal: Evolution and biodiversity*, 2nd ed.; O. M. Kozhova, L. R. Izmet'seva, Eds.; Backhuys Publishers: Leiden, NL, 1998, pp 417-447, ISBN 9789057820014
26. Kulikovskiy, M.S.; Lang-Bertalot, H.; Kuznetsova, I.V. *Lake Baikal: hotspot of endemic diatoms: II. Iconographia Diatomologic*; Koeltz Scientific Books, Oberreifenberg, Ger., 2015, ISBN 9783874294904.
27. Annenkova, N. V.; Hansen, G.; Moestrup, Ø.; Rengefors, K. Recent adaptive radiation in a marine and freshwater dinoflagellate species flock. *ISME J* **2015**, *9*, 1821-1834.
28. Yi, Z.; Berney, C.; Hartikainen, H.; Mahamdallie, Sh.; Gardner, M.; Boenigk, J.; Cavalier-Smith, T.; Bass, D. High throughput sequencing of microbial eukaryotes in Lake Baikal reveals ecologically differentiated communities and novel evolutionary radiations. *FEMS Microbiol Ecol* **2017**, *93*, doi: 10.1093/femsec/fix073.
29. Mikhailov, I.S.; Zakharova, Y.R.; Bukin, Yu.S.; Galachyants, Yu.P.; Petrova, D.P.; Sakirko, M.V.; Likhoshway, E.V. Co-occurrence Networks Among Bacteria and Microbial Eukaryotes of Lake Baikal During a Spring Phytoplankton Bloom. *Microbial Ecol* **2019**, *77*, 96-109.
30. A Stoeck, T.; Bass, D.; Nebel, M.; Christen, R.; Jones, M.D.; Breiner, H.W.; Richards, T.A. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol Ecol* **2010**, *19*, 21-31.
31. Logares R. Workflow for analysing MiSeq Amplicons based on Uparse v1. 5. *Zenodo10* **2017**, 5281.
32. Nikolenko, S.I.; Korobeynikov, A.I.; Alekseyev, M.A. BayesHammer: Bayesian clustering for error correction in single-cell sequencing. *BMC Genomics* **2013**, *14*, S7.
33. Schirmer, M.; Ijaz, U.Z.; D'Amore, R.; Hall, N.; Sloan, W.T.; Quince, C. Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. *Nucleic Acids Res* **2015**, *43*, e37.

34. Zhang, J.; Kobert, K.; Flouri, T.; Stamatakis, A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **2014**, *30*, 614-620.
35. Edgar, R.C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **2010**, *26*, 2460-2461.
36. Edgar, R.C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* **2013**, *10*, 996-998.
37. Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glöckner, F.O. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **2013**, *41*, D590-596.
38. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *JMB* **1990**, *215*, 403-410.
39. Guillou, L.; Bachar, D.; Audic, S.; Bass, D.; Berney, C.; Bittner, L.; Boutte, C.; Burgaud, G.; de Vargas, C.; Decelle, J. *et al.* The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res* **2013**, *41*, D597-604.
40. Pernice, M.C.; Logares, R.; Guillou, L.; Massana, R. General patterns of diversity in major marine microeukaryote lineages. *PLoS One* **2013**, *8*, e57170.
41. Massana, R.; Gobet, A.; Audic, S.; Bass, D.; Bittner, L.; Boutte, C.; Chambouvet, A.; Christen, R.; Claverie, J.-M.; Decelle, J. *et al.* Marine protist diversity in European coastal waters and sediments as revealed by high-throughput sequencing. *Environ Microbiol* **2015**, *17*, 4035-4049.
42. Team, R.C.R: A language and environment for statistical computing. **2013**.
43. Katoh, S. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **2013**, *30*, 772-780.
44. Capella-Gutierrez, S.; Silla-Martinez, J.M.; Gabaldon, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **2009**, *25*, 1972-1973.
45. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312-1313.

46. Logares, R.; Mangot, J.F.; Massana, R. Rarity in aquatic microbes: Placing protists on the map. *Res Microbiol* **2015**, *166*, 831–841.
47. Miranda, L.N.; Zhuang, Y.Y.; Zhang, H.; Lin, S. Phylogenetic analysis guided by intragenomic SSU rDNA polymorphism refines classification of "Alexandrium tamarense" species complex. *Harmful Algae* **2012**, *16*, 35-48.
48. Ichinomiya, M.; dos Santos, A.; Gourvil, P.; Yoshikawa, S.; Kamiya, M.; Ohki, K.; Audic, S.; de Vargas, C.; Nođ, M.-H.; Vaulot, D.; Kuwata, A. Diversity and oceanic distribution of the Parmales (Bolidophyceae), a picoplanktonic group closely related to diatoms. *ISME J* **2016**, *10*, 2419–2434.
49. Hevia-Orube, J.; Orive, E.; David, H.; D éz, A.; Laza-Mart ínez, A.; Miguel, I.; Seoane, S. Molecular and morphological analyses of solitary forms of brackish Thalassiosiriod diatoms (Coscinodiscophyceae), with emphasis on their phenotypic plasticity. *Eur J Phycol* **2016**, *51*, 11-30.
50. Vishniac, H.A.; Hempfling, W.P. *Cryptococcus vishniacii* sp. nov., an Antarctic yeast. *Int J Syst Evol Microbiol* **1979**, *29*, 153-158.
51. Luo, B.; Sun, H.; Zhang, Y.; Gu, Y.; Yan, W.; Zhang, R.; Ni, Y. Habitat-specificity and diversity of culturable cold-adapted yeasts of a cold-based glacier in the Tianshan Mountains, northwestern China. *Appl Microbiol Biotechnol* **2019**, *103*, 2311–2327.
52. Rengefors, K.; Logares, R.; Laybourn-Parry, J. Polar lakes may act as ecological islands to aquatic protists. *Molec Ecol* **2012**, *21*, 3200-3209.
53. Khublaryan, M.G. Surface waters: Rivers, Streams, Lakes and Wetlands. In *Types and Properties of Water - Vol. 1*; Khublaryan, M.G. Ed.; EOLS Publishers Co. Ltd.: Oxford, UK, 2009, ISBN: 978-1-84826-922-4
54. Martens, K. Speciation in ancient lakes. *Trends Ecol. Evol.* **1997**, *12*, 177-18.
55. Karabanov, E.; Douglas, W.; Kuzmin, M.; Sideleva, V.; Khursevich, G.; Prokopenko, A.; Solotchina, E.; Tkachenko, L.; Fedenya, S.; Kerber, E. et al. Ecological collapse of Lake Baikal and Lake Hovsgol ecosystems during the Last Glacial and consequences for aquatic species diversity. *Palaeogeography, Palaeoclimatology, Palaeoecology* **2004**, *209*, 227-243.
56. Popovskaya, G.I. 1983. Phytoplankton. In: Limnology of Northern Baikal. Novosibirsk; Galazii, G.I., Bogdanov, V. T.; Nauka: Novosibirsk, Russia, 85-93.

57. Fulton, T.L.; Strobeck, C. Multiple fossil calibrations, nuclear loci and mitochondrial genomes provide new insight into biogeography and divergence timing for true seals (Phocidae, Pinnipedia). *J Biogeogr* **2010**, *37*, 814-829.
58. Cabello-Yeves, P.J.; Zemskaya, T.I.; Rosselli, R.; Coutinho, F.H.; Zakharenko, A.S.; Blinov, V.V.; Rodriguez-Valera, F. Genomes of Novel Microbial Lineages Assembled from the Sub-Ice Waters of Lake Baikal. *AEM* **2017**, *84*, e02132-17.
59. Annenkova, N.V.; Hansen, G.; Rengefors, K. Closely related dinoflagellate species in vastly different habitats – an example of a marine-freshwater transition. *Eur J Phycol*, under review.
60. Logares, R.; Bråte, J.; Bertilsson, S.; Clasen, J. L.; Shalchian-Tabrizi, K.; Rengefors, K. Infrequent marine–freshwater transitions in the microbial world. *Trends Microbiol* **2009**, *17*, 414-422.
61. Spielhagen, R.F.; Erlenkeuser, H.; Siegert, Ch. History of freshwater runoff across the Laptev Sea (Arctic) during the last deglaciation. *Global and Planetary Change* **2005**, *48*: 187-207.
62. Kirst, G.O; Wiencke, C. Ecophysiology of polar algae. *J Phycol* **1995**, *31*, 181-99.
63. Vereshchagin, G. Yu. Origines et histoire de la faune et de la flora du lac Baikal. *Int Rev d gesamten Hydrobiol*, **1940**, *5*, 390-419.
64. Velasco-González, I.; Sanchez-Jimenez, A.; Singer, D.; Murciano, A.; D éz-Hermano, S.; Lara, E.; Mart ín-Cereceda, M. Rain-Fed Granite Rock Basins Accumulate a High Diversity of Dormant Microbial Eukaryotes. *Microb Ecol* **2019**, doi: 10.1007/s00248-019-01463-y. [Epub ahead of print]
65. Gavrillets, S.; Losos, J.B. Adaptive Radiation: contrasting theory with data. *Science* **2009**, *323*, 732–737.
66. Schön, I.; Martens, K. Adaptive, pre-adaptive and non-adaptive components of radiations in ancient lakes: a review. *Org Divers Evol* **2004**, *4*, 137–156.
67. Stelbrink, B.; Jovanovska, E.; Levkov, Z.; Ognjanova-Rumenova, N.; Wilke, T.; Albrecht, C. Diatoms do radiate: evidence for a freshwater species flock. *J Evol Biol*, **2018**, *31*, 1969-1975.
68. Kulikovskiy, M.S.; Lange-Bertalot, H.; Metzeltin, D.; Witkowski, A. *Lake Baikal: hotspot of endemic diatoms: I. Iconographia Diatomologica*; Koeltz Scientific Books, Koenigstein, Ger., 2012, ISBN 978-3-905-997-09-5.

69. Zhang, W.; Li, Y.L.; Kociolek, J.P.; Zhang, R.L.; Wang, L.Q. *Oricymba tianmuensis* sp. nov., a new cymbelloid species (Bacillariophyceae) from Tianmu Mountain Zhejiang Province, China. *Phytotaxa* **2015**, *236*, 257–265.
70. Kociolek, J.P.; Hamsher, S.E.; Kulikovskiy, M.; Bramburger, A.J. Are there species flocks in freshwater diatoms? A review of past reports and a look to the future. *Hydrobiologia* **2017**, *792*, 17–35.
71. Schluter, D. *The Ecology of Adaptive Radiation*. Oxford University Press: Oxford, UK, 2000.
72. Bessudova, A.Yu.; Sorokovikova, L.M.; Tomberg, I.V.; Likhoshway, Y.V. Silica-Scaled Chrysophytes in Large Tributaries of Lake Baikal. *Cryptogamie Algal* **2018**, *39*, 145–165.
73. Ganley, A.R.D.; Kobayashi, T. Highly efficient concerted evolution in the ribosomal DNA repeats: Total rDNA repeat variation revealed by whole-genome shotgun sequence data. *Genome Res* **2007**, *17*, 184–191.
74. Gong, J.; Dong, J.; Liu, X.; Massana R. Extremely high copy numbers and polymorphisms of the rDNA operon estimated from single cell analysis of oligotrich and peritrich ciliates. *Protist* **2013**, *164*, 369–79.
75. Kudryavtsev, A.; Gladkikh, A. Two new species of *Ripella* (Amoebozoa, Vannellida) and unusual intragenomic variability in the SSU rRNA gene of this genus. *Eur J Protistol* **2017**, *61*, 92–106.
76. Annenkova, N.V.; Belykh, O.I.; Denikina, N.N.; Belikov, S.I. Identification of dinoflagellates from the Lake Baikal on the basis of molecular genetic data. *Dokl Biol Sci* **2009**, *426*, pp. 253–256.
77. Khursevich, G.K.; Karabanov, E.B.; Prokopenko, A.A.; Williams, D.F.; Kuzmin, M. I.; Fedenya, S.A.; Gvozdkov, A.N.; Kerber, E.V. Detailed diatom biostratigraphy of Baikal sediment during the Brunhes Chron and climatic factors of species formation. *Russian Geology and Geophysics* **2011**, *42*, 108 – 129.
78. Stelbrink, B.; Shirokaya, A.A.; Clewing, C.; Sitnikova, T.Y.; Prozorova, L.A.; Albrecht C. Conquest of the deep, old and cold: an exceptional limpet radiation in Lake Baikal. *Biol lett* **2015**, *11*, 20150321.
79. Whittaker, K.A.; Ryneerson, T.A. Evidence for environmental and ecological selection in a microbe with no geographic limits to gene flow. *Proc Natl Acad Sci USA* **2017**, *114*, 2651–2656.
80. Fietz, S., Bleiß, W., Hepperle, D., Koppitz, H., Krienitz, L. and Nicklisch, A., 2005. First record of *Nannochloropsis limnetica* (Eustigmatophyceae) in the autotrophic picoplankton from lake baikal. *J Phycol* **2005**, *41*, 780–790.
81. Rillig, M.C.; Tsang, A.; Roy, J. Microbial community coalescence for microbiome engineering. *Front*

Microbiol **2016**, 7, 1967.