FEMS Microbiology Ecology, 97, 2021, fiab114

https://doi.org/10.1093/femsec/fiab114
Advance Access Publication Date: 2 August 2021
Research Article

RESEARCH ARTICLE

# Microbial eukaryote assemblages and potential novel diversity in four tropical East African Great Lakes

Paulina Fermani[1,*,†], Sebastián Metz[2], Vanessa Balagué[3], Jean Pierre Descy[4], Cédric Morana[4], Ramiro Logares[3], Ramon Massana[3] and Hugo Sarmento[5,‡]

[1]Instituto Tecnológico de Chascomús (INTECH), Universidad Nacional de San Martín-Consejo Nacional de Investigaciones Científicas y Técnicas, Av. Intendente Marino Km 8.200, (7130) Chascomús, Buenos Aires, Argentina, [2]University of Brest, CNRS, IRD, Ifremer, LEMAR, F-29280, Plouzané, France, [3]Institut de Ciències del Mar (CSIC), Passeig Marítim de la Barceloneta, 37-49, ES-08003, Barcelona, Catalonia, Spain, [4]Unit of Chemical Oceanography, University of Liège, Quartier Agora Allée du six Août 19 4000 Liège 1, Belgium and [5]Departamento de Hidrobiologia (DHb), Universidade Federal de São Carlos, Rodovia Washington Luís, Km 235 - Cx. Postal 676, CEP 13565-905 - São Carlos/SP, Brazil

*Corresponding author: Instituto Tecnológico de Chascomús (INTECH-UNSAM-CONICET), Av. Intendente Marino Km 8.2, Chascomús (7130), Buenos Aires, Argentina. Tel: (+54)-2241-430323, Ext. 110; E-mail: pfermani@intech.gov.ar

**One sentence summary:** Microbial eukaryotes diversity in four East African Great Lakes.

**Editor:** Martin W. Hahn

†Paulina Fermani, https://orcid.org/0000-0002-8036-9347
‡Hugo Sarmento, https://orcid.org/0000-0001-5220-7992

## ABSTRACT

East African Great Lakes are old and unique natural resources heavily utilized by their bordering countries. In those lakes, ecosystem functioning is dominated by pelagic processes, where microorganisms are key components; however, protistan diversity is barely known. We investigated the community composition of small eukaryotes (<10 μm) in surface waters of four African Lakes (Kivu, Edward, Albert and Victoria) by sequencing the 18S rRNA gene. Moreover, in the meromictic Lake Kivu, two stations were vertically studied. We found high protistan diversity distributed in 779 operational taxonomic units (OTUs), spanning in 11 high-rank lineages, being Alveolata (31%), Opisthokonta (20%) and Stramenopiles (17%) the most represented supergroups. Surface protistan assemblages were associated with conductivity and productivity gradients, whereas depth had a strong effect on protistan community in Kivu, with higher contribution of heterotrophic organisms. Approximately 40% of OTUs had low similarity (<90%) with reported sequences in public databases; these were mostly coming from deep anoxic waters of Kivu, suggesting a high extent of novel diversity. We also detected several taxa so far considered exclusive of marine ecosystems. Our results unveiled a complex and largely undescribed protistan community, in which several lineages have adapted to different niches after crossing the salinity boundary.

**Keywords:** protists; metabarcoding; diversity; novelty; Great Lakes; Africa

## INTRODUCTION

The East African Rift valley harbors some of the world's largest lakes, such as Lakes Tanganyika, Malawi, Edward, Albert, Kivu and Victoria, the second largest freshwater in the world (Ogutu-Ohwayo *et al.* 1997). These African Great Lakes provide essential ecosystem services to their bordering populations, like water supply, fisheries, recreation and tourism (Cohen, Kaufman

**Received:** 22 April 2021; **Accepted:** 30 July 2021

and Ogutu-Ohwayo 1996; Odada *et al.* 2003). Some of them are among the oldest freshwater systems in the world, displaying an array of limnological and biological features, ranging from productive waters to extremely oligotrophic systems (Bootsma and Hecky 1993; Loiselle *et al.* 2014; İnceoğlu *et al.* 2015). They have attracted biologists from the beginning of the 19th century for their very large vertebrate and invertebrate biodiversity, with for instance hundreds of endemic fish species (e.g. Brooks 1950; Fryer and Iles 1972; Seehausen *et al.* 1997). Large tropical lakes differ from the temperate ones mainly by the so-called endless summer (Kilham and Kilham 1990), with high temperatures throughout the year. Due to their large size, ecosystem function is dominated by pelagic processes, where microorganisms are key components (Descy and Sarmento 2008). However, the high human population density inhabiting near the coast generates high resource exploitation, pollution, species introduction and eutrophication, which together with climate change impacts, alter lake biological diversity and function (Hecky and Bugenyi 1992; Bootsma and Hecky 1993; Otu *et al.* 2011).

Although they share common characteristics, the African Great lakes display a variety of limnological conditions, which have to be taken into account for understanding their functioning and predicting their future, including their response to climate and other anthropogenic influences (Bootsma and Hecky 1993; Odada *et al.* 2003). In general, environmental heterogeneity produces community differences in water bodies, important structuring abiotic factors on the local scale being temperature, light, depth, pH, conductivity and nutrients, among others (Wu *et al.* 2009; Weisse *et al.* 2016). In particular, the deepest lakes (Tanganyika, Malawi and Kivu) are meromictic. As opposed to the shallower lakes (Victoria, Edward and Albert), which are holomictic, they never experience complete mixing of the water column, due to their great depth and the existence of a salinity gradient, separating the oxic surface waters (the mixolimnion) from the permanently anoxic deep waters (the monimolimnion). Because they have a chemocline with a redox gradient at the transition between the surface and deep waters, meromictic lakes have many ecological niches for microbial communities (Oikonomou, Pachiadaki and Stoeck 2014; Lepère *et al.* 2016). The anoxic and saline waters below the surface layers create another extreme environment (Charvet *et al.* 2012).

Lake Kivu is of particular interest because the vertical mixing and transport processes are different from most other large lakes in the world (Schmid and Wüest 2012). The oxic mixolimnion responds to the same atmospheric forcing as the other Rift lakes, with a relatively weak thermal stratification. On the other hand, the deep waters display a unique structure because they receive heat, salts and $CO_2$ from deep geothermal inflows, which increase the salinity, nutrients and methane concentrations (Schmid *et al.* 2002; Borges *et al.* 2011). The lake also shows significant spatial heterogeneity, as differences in the vertical structure of the water column between the Northern and Southern basins have been observed (Morana *et al.* 2015; Roland *et al.* 2017).

Single-celled eukaryotes (protists) are key for ecosystem functioning (Sherr and Sherr 1988). Autotrophs are the main carbon fixers in aquatic environments, whereas heterotrophs promote nutrient cycling by consuming bacteria and picophytoplankton, being also crucial in determining the transport of organic matter to higher trophic levels (Pomeroy 1974; Boenigk and Arndt 2002; Stenuite *et al.* 2009). Moreover, parasitic protists influence community dynamics of larger eukaryotic hosts (Worden *et al.* 2015). Several authors have observed that carbon transfer through this microbial food web to higher trophic levels is significantly more important in oligotrophic than eutrophic systems (Porter *et al.* 1988; Weisse *et al.* 1990; Domingues *et al.* 2017), as suspected by Hecky *et al.* (1981) and demonstrated by Pirlot *et al.* (2005) and Tarbe *et al.* (2011a) in Lake Tanganyika and Kivu (Sarmento *et al.* 2008). Nevertheless, due to the large number of interactions between planktonic components in the most productive systems, this trend is still under debate (Fermani *et al.* 2015; Xiong *et al.* 2021).

Over the last decades, studies of eukaryotic microbial communities in freshwater systems through molecular analysis have increased (Simon *et al.* 2014; Schloss *et al.* 2016; Debroas *et al.* 2017; Annenkova, Rodriguez-Giner and Logares 2020). Recently, two analyses across diverse habitats including freshwater, marine and soils, revealed clear differences in the taxonomic composition of the major protistan lineages as well as a higher $\beta$-diversity in freshwater than in the other systems (Singer *et al.* 2021; Xiong *et al.* 2021). Despite their crucial ecological roles, protist assemblages are barely known in African lakes (e.g. Tarbe *et al.* 2011b; Llirós *et al.* 2012), although they harbor an important part of the world biodiversity heritage (see e.g. Vadeboncoeur, Mcintyre and Vander Zander 2011), and habitats with oxygen gradients have received very little attention (e.g. Tarbe *et al.* 2011b; Oikonomou, Pachiadaki and Stoeck 2014; Lèpere *et al.* 2016).

Taking into account their importance in food webs and in order to fill these gaps, our first aim was to describe the small eukaryote (<10 μm) assemblage in four East African Great Lakes (Edward, Albert, Victoria and Kivu) using high-throughput amplicon tag pyrosequencing of the V4 region of the 18S rRNA gene. Because of their ancestry, these lakes might be able to support a high rate of endemism (e.g. Cichlids); therefore, we expect to find a high degree of potential novel diversity. Further, our second goal was to compare the surface layers of Lakes Edward, Albert, Victoria and two basins of Kivu (the main basin, which is the most exposed to the winds, and the southern basin, which is more wind protected) and to determine the main drivers shaping the structure of communities. Although these freshwater systems share common characteristics (same geological formation, tropical, large and old lakes), we hypothesize that their protist community may differ, due to morphological, physical and chemical differences among lakes. Moreover, considering the unique characteristics of Lake Kivu, our third objective was to compare the eukaryote microbial composition over a vertical profile in the redox gradient at two sites: Gisenyi (northern basin) and Ishungu (southern basin). We hypothesize that both sites exhibit a dissimilar protistan assemblage, due to the physical–chemical differences of the redox gradient. In turn, we expect to find a high and barely known diversity, especially in the deepest areas where the conditions are extreme.

## MATERIALS AND METHODS

### Study sites

East African Great Lakes are located between 4°35′N and 14°30′S of latitude, covering 2100 km north–south distance (Spigel and Coulter 1996). The region has a tropical climate with a bimodal distribution in rainfall, which defines two seasons: a dry season with south-eastern dominant winds (from June to September) and a rainy season (from October to May), calmer and warmer (Descy and Sarmento 2008). Some of the main lakes in the region are Kivu, Edward, Albert and Victoria (Fig. 1A). Each one is situated in a different hydrological basin differing

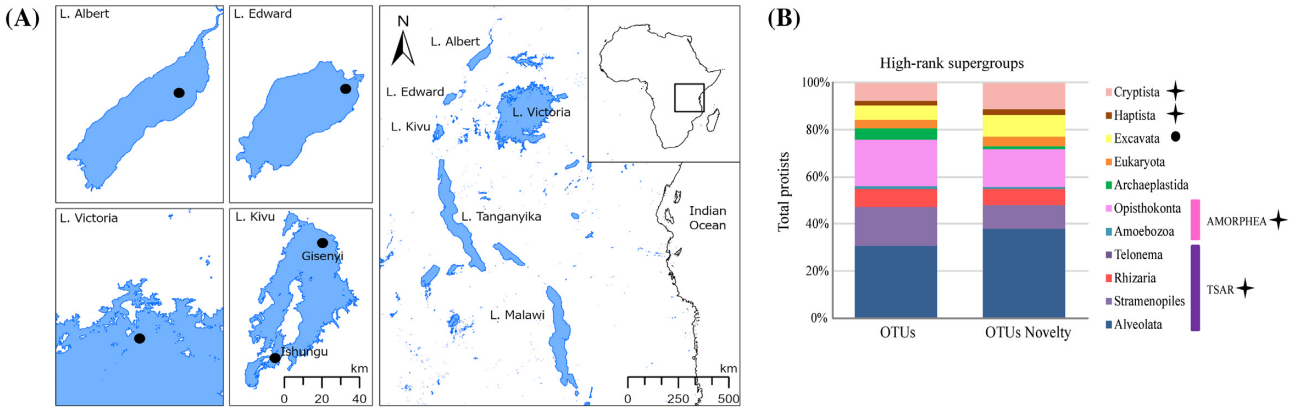


**Figure 1. (A)** Map of East African Great Lakes sampled during rainy season. The left panel shows the four lakes amplified: Albert, Edward, Victoria and the two basins of Kivu: Ishungu and Gisenyi. Black circles indicate sampling sites. **(B)** Percentage total of protist high-rank supergroups found in East Great African Lakes. Total operational taxonomic units (OTUs; $n = 779$); Total OTUs Novelty ($n = 316$). Star symbols denote taxa that were considered as supergroups in the recent classification based on new phylogenetic analyses, and circles show taxa that are still without a clear resolution (Burki *et al.* 2020).

in catchment dynamics, human impacts and limnological characteristics (Table 1) (Odada *et al.* 2003; Descy and Sarmento 2008; Morana *et al.* 2014).

## Sampling

Samples for limnological and biological parameters were collected during the rainy season in the pelagic zone of four lakes. Meromictic Lake Kivu (K) was sampled in February 2012 at two different sites: Ishungu (I) and the main basin of Gisenyi (G), located in the southern and northern part of the lake, respectively (Fig. 1; Table 1). Kivu samples were taken between 0 and 80 m depth, at intervals of 5 m (to cover the whole gradient of oxygen concentrations). Lakes Edward (E), Albert (A) and Victoria (V) were sampled in May 2012. They are shallower than Lake Kivu, holomictic and chlorophyll-*a* (Chl-*a*) concentrations are usually higher in their mixed layer (7.5, 20 and 15 m, respectively) (Morana *et al.* 2014). Based on *in situ* observed vertical profiles of temperature and oxygen, the mixed-layer depth was determined in each water body. Water samples were collected every 5 m from the surface to the bottom of the mixed layer using a 7-L Niskin bottle (Hydro-Bios, Apparatebau, Altenholz, Germany) and were then pooled to obtain a representative sample of this layer.

Limnological profiles were obtained in each lake using a Yellow Springs Instruments (Ohio, USA) 6600 v2 multiparameter probe. Water samples for chemical and microbiological analyses were collected using a 7.5-L Niskin bottle and stored in 4-L plastic containers for chemical analyses (except for $CH_4$) and 2-L Nalgene plastic bottles for biological analyses. Filtrations were carried out in the field; filters and water samples were kept frozen until analyses in the laboratory.

## Environmental variables

Water temperature (Temp), specific conductivity (Cond), depth (Z), pH and oxygen concentration (DO) were obtained from the limnological profiles. Total alkalinity (Alk) was carried out by open-cell titration with HCl 0.1 M on 50 mL water samples (Morana *et al.* 2014). Ammonia ($NH_4^+$) concentrations were determined using the dichloroisocyanurate-salicylate-nitroprussiate colorimetric method. Nitrite ($NO_2^-$) was calculated by the sulfanilamide coloration method, while nitrate ($NO_3^-$) was determined after cadmium reduction to $NO_2^-$ and quantified following the sulfanilamide coloration method (Llirós *et al.* 2010). $NO_x$ was calculated as the sum of $NO_2^-$ and $NO_3^-$, whereas DIN (dissolved inorganic nitrogen) was equivalent to $NH_4^+ + NO_2^- + NO_3^-$. Soluble reactive phosphorus (SRP) was determined using the molybdenum blue method (APHA 1992). Dissolved methane ($CH_4$) concentrations were measured by a headspace technique (Weiss 1981) using gas chromatography with flame ionization detection (GC-FID, SRI 8610C), as detailed by Borges *et al.* (2011). Chlorophyll-*a* (Chl-*a*) concentrations were determined by high-performance liquid chromatography (HPLC). Three liters of water were filtered on a Macherey-Nägel GF-5 filter (nominal porosity of 0.4 mm). Pigment extraction from the filters was carried out in 10 mL of 90% HPLC grade acetone. After two sonication steps of 15 min separated by an overnight period at 4°C, the extracts were stored in 2 mL amber vials at 25°C. HPLC analysis was performed following the method described in Sarmento, Isumbisho and Descy (2006) with a waters system comprising a photodiode array and fluorescence detectors. Calibration was made using commercial external standards (DHI Denmark Lab Products). Precision for Chl-*a* measurement was better than ±7%. Additional information on the analytical techniques and the data collected can be found in Morana *et al.* (2014).

## DNA extraction, amplification of the 18S rRNA gene and pyrosequencing

Approximately 500 mL of surface mixed-layer water for DNA extraction from Edward, Albert and Victoria was prefiltered through 10-µm pore-size membranes (ISOPORE, Millipore, MA) and cells were collected onto 0.8-µm pore-size filters (ISOPORE, Millipore, MA). Moreover, based on the physical and chemical data, six depths of each site of Kivu (I and G) were chosen to analyze the diversity of microbial eukaryotes: I-5, I-15, I-30, I-45, I-50 and I-70 m and G-5, G-15, G-25, G-30, G-35 and G-70 m. In these cases, 1000 mL of water was passed through 5-µm pore-size filters (ISOPORE, Millipore, MA) and planktonic cells were collected onto 0.2-µm pore-size membranes (ISOPORE, Millipore, MA). Each filter was placed in a cryovial contained 1.7 mL of lysis buffer (50 mM Tris, 40 mM EDTA and 0.75 M sucrose) and stored at –80°C until DNA extraction.

DNA was extracted from filters following the phenol-chloroform-isoamyl alcohol protocol. In brief, filters were

**Table 1.** Main geographical and limnological characteristics of East African Great Lakes.

| Lake | Kivu Ishungu (I) | Kivu Gisenyi (G) | Edward (E) | Albert (A) | Victoria (V) |
|---|---|---|---|---|---|
| Coordinates | 02°16′S, 28°59′E | 01°47′S, 29°12′E | 00°12′N, 29°49′E | 01°98′N, 31°16′E | 00°39′N, 33°16′E |
| Altitude (msnm) | 1463 | 1481 | 928 | 620 | 1151 |
| Lake area ($km^2$) | 2322 (main lake) | 2322 (main lake) | 2325[a] | 5300[a] | 68 800[a] |
| Maximum depth (m) | 160[b] | 440[b] | 117[a] | 58[a] | 79[a] |
| Mean euphotic depth 1% (m) | 18 (main lake)[a] | 18 (main lake)[a] | 13[a] | 12[a] | 9[a] |
| Mixing regime | Meromictic | Meromictic | Holomictic | Holomictic | Holomictic |

[a]Morana *et al.* (2014).
[b]Pasche *et al.* (2009).

incubated with Lysozyme (1 mg $mL^{-1}$ final concentration) at 37°C for 45 min while slightly shaken. In addition, Proteinase K (0.2 mg $mL^{-1}$) and sodium dodecyl sulfate (10% final concentration) were added and tubes were incubated at 55°C for 1 h while slightly shaken. The resulting lysate underwent two steps of phenol–chloroform-isoamyl alcohol (25:24:1) extraction: the phenol mixture was spin 10 min at 12 000 rpm. Then, one extract of equal volume of chloroform-isoamyl alcohol (24:1) was spin during 10 min at 12 000 rpm. After the last centrifugation step, the aqueous phase was collected, concentrated in an Amicon Ultra unit (Millipore) and washed with sterile deionized water. The total DNA extract was quantified using a Nanodrop-1000 spectrophotometer (Thermo Scientific) and stored at –80°C until further analysis (Casamayor *et al.* 2002; Massana *et al.* 2004).

The microbial eukaryote diversity was assessed by high-throughput sequencing the 18S rRNA gene. The eukaryotic universal primers TAReuk454FWD1 (5′-CCAGCASCYGCGGTAATTCC-3′) and TAReukREV3 (5′-ACTTTCGTTCTTGATYRA-3′) (Stoeck *et al.* 2010) were used to amplify the hypervariable V4 region (~380 bp). The primers included 454-specific adaptors, and the forward primer also carried the barcode (Table S1, Supporting Information). The procedures followed for PCR amplifications were previously described in detail by Pernice *et al.* (2016). In brief, we added 3.5 $\mu$L (15 ng $\mu L^{-1}$) of genomic DNA to PCR tubes containing dNTPs (0.2 mM), 1 μL of the primers R and F indicated above at 0.5 mM and the Phusion High-Fidelity DNA Polymerase kit (Thermofisher) following the manufacturer's recommendations in a final volume of 20 $\mu$L. Reactions were carried out in an automated thermocycler with the following cycle: initial denaturation step at 98°C for 3 min, followed by 30 cycles of denaturation at 98°C for 35 s., annealing at 51°C for 40 s., extension at 72°C for 45 s; and a final extension at 72°C for 10 min. Amplicons were checked in a 1.5% agarose gel for successful amplification. Due to low amount of DNA found, some samples (G-70, I-50, I-70 and E) were precipitated with ethanol-sodium acetate and resuspended in 13 $\mu$L of sterile water. PCRs were performed with 15 $\mu$L of PCR mix, 4 $\mu$L of the environmental DNA and 1 $\mu$L of Primer specific for 454-pyrosequencing. Triplicate amplicons were pooled, purified and quantified using the Qubit dsDNA broad-range (BR) assays (Invitrogen). About 10 ng $\mu L^{-1}$ (30 μL final volume) of PCR product were sent to the Research and Testing Laboratory (Lubbock, TX, USA; http://www.researchandtesting.com) for sequencing using 454 GS FLX (Roche-454 Life Sciences, 454 Life Sciences, Branford, CT, USA) with Titanium chemistry.

## Processing 454 sequences

Pyrosequencing data generated from the 454-sequencing runs were processed using QIIME (Caporaso *et al.* 2010). We obtained 145 236 sequences (pyrotags), which were demultiplexed using the barcode identifier in the forward primer. Pyrotags were between 150 and 600 bp long and those that were selected had no more than four mismatches in the primer and no homopolymers longer than 8 bp. For quality check, errors were computed in sliding windows of 50 bp and pyrotags containing a window with an error >1% and appearing only once in the dataset were removed. After eliminating possible PCR and pyrosequencing errors, a total of 117 432 reads from the 15 samples were left, which were denoised with DeNoiser and clustered into OTUs (operational taxonomic units) with UCLUST (99% similarity threshold) (Edgar *et al.* 2011). OTUs were taxonomically assigned using PR2 (version 4.13.0) database (Guillou *et al.* 2013). OTUs were assigned to a given group when its representative sequence had a BLAST hit with an e-value below $10^{-100}$ against a reference sequence. Chimera check and removal was performed using the SILVA 108 reference database (Quast *et al.* 2013) within the MOTHUR program (Schloss 2008). The commands 'align.seqs' and '**chimera.slayer**' were used (Haas *et al.* 2011) and Archaea, Bacteria and Metazoa sequences were eliminated. After removing singletons, doubletons, 1003 OTUs with No Blast Hit (related to sequencing errors) and the sample G-30 (because it had <1000 reads), the final out table included 86 463 reads from 14 samples distributed among 779 OTUs.

Raw data was deposited at the European Nucleotide Archive public database under the following accession number: PRJEB42983.

## Novelty analysis

To determine the degree of 'novelty' of the OTUs, the sequences were aligned to the NCBI database using BLASTN, with default parameters. Then we obtained the similarity of the first best hit corresponding to a cultured organism or to an environmental sequence, to assign the closest cultured match (CCM) and the closest environmental match (CEM), respectively (Massana *et al.* 2011; Triadó-Margarit and Casamayor 2012). After examining the averaged identity values for all sequences into separate taxonomic groups, both similarity values (CCM and CEM) were plotted in a 2D dispersion graph, giving the degree of 'novelty' of the dataset. OTUs with high percentages of CEM indicate that they are similar to sequences obtained in other environmental surveys, and those with low CEM similarity highlight true novel diversity detected here. Likewise, points with high CCM similarity indicate that sequences are close to cultured organism's sequences, whereas dots with low CCM similarity highlight environmental sequences with no cultured counterpart. Finally, for each plot we defined 'the highest novelty' as the area of the plot that contained OTUs with <90 or 95% (depending the supergroup) similarity to both CEM and CCM (del Campo and Massana 2011).

**Table 2.** Main physical, chemical and biological variables of East African Great Lakes in surface mixed-layer waters.

| Lake | Kivu Ishungu (I) | Kivu Gisenyi (G) | Edward (E) | Albert (A) | Victoria (V) |
|---|---|---|---|---|---|
| Temperature (°C) | 24.10 | 24.30 | 26.60 | 27.90 | 25.10 |
| pH | 9.50 | 8.40 | 9.10 | 8.90 | 8.30 |
| Conductivity (μS $cm^{-1}$) | 1155.00 | 1085.50 | 849.20 | 640.70 | 95.00 |
| Alkalinity (nmol $Kg^{-1}$) | 12.70 | 12.70 | 8.20 | 5.80 | 0.80 |
| DO (mg $L^{-1}$) | 5.70 | 7.50 | 7.60 | 8.50 | 7.20 |
| DIN (μM) | 0.36 | 0.13 | 2.42 | 2.74 | 3.88 |
| SRP (μM) | 0.20 | 0.20 | 0.43 | 0.19 | 0.82 |
| Chl-*a* (μg$L^{-1}$) | 2.10 | 2.10 | 9.90 | 5.90 | 6.00 |
| Trophic State (TSI) | Oligotrophic | Oligotrophic | Meso-eutrophic | Mesotrophic | Mesotrophic |

DO, dissolved oxygen; DIN, dissolved inorganic nitrogen; SRP, soluble reactive phosphorus; Chl-*a*, chlorophyll-*a*.

## Data analysis

The trophic status of the lakes was quantified by applying the Trophic State Index (TSI) developed by Carlson (1977), using Chl-*a* as a variable. Richness was evaluated by rarefaction analyses as the estimated number of OTUs in each lake (Hurlbert 1971). Rarefaction curves were constructed using 'rarecurve' function of Vegan version 2.5–6 (Oksanen *et al.* 2019) and R 3.5 (http://cran.r-project.org). Therefore, the relative abundance of each sample was resampled at the minimum sample size ($n = 1064$ and 1628 at superficial level and Kivu depths, respectively) with the 'rrarefy' function, R package Vegan version 2.5-6 (Oksanen *et al.* 2019). The $\alpha$-diversity (i.e. Shannon_H and Simpson_1-D indices, as well as Pielou's evenness) was determined using the Past 4.03 statistical package (Hammer 2016). Moreover, based on similar studies, we defined abundant OTUs as those with a mean relative abundance in all samples ≥1% (Logares *et al.* 2014; Du *et al.* 2019; Annenkova, Rodriguez-Giner and Logares 2020).

To explore the environmental drivers shaping protistan communities in surface mixed-layer waters and Kivu depths, we performed multivariate analyses. Ordinations by detrended correspondence analysis (DCA) indicated a linear distribution of the data (ter Braak and Šmilauer 1998). Then, a principal component analysis (PCA with supplementary variables) was conducted on microbial communities interpreting this variation with the help of environmental variables, in order to add supplementary information on the scatter plot for a better understanding of the data (Lê, Josse and Husson 2008). Supplementary variables were transformed with $y' = \ln(y + 1)$ in order to use linear ordination methods (ter Braak 1987). Variables characterized by high inflation factors were then dismissed. Since the microbiome data sets are compositional, they were not transformed and the PCA was the most accurate analysis representing the relationships between the eukaryotic microorganisms and the distances between samples on a common plot (Gloor *et al.* 2017). Multivariate analysis was performed using the CANOCO 5 software (ter Braak and Šmilauer 1998).

## Phylogenetic exploration of 'novel' diversity in Lake Kivu

According to '*Novelty analysis*', Alveolata and Stramenopiles were two of the supergroups that showed the highest numbers of novel OTUs. Thereby, OTUs belonging to those lineages in Lake Kivu were classified phylogenetically. First, we performed a reference alignment using the sequences from Mahé *et al.* (2017) reference tree and the best hit from each OTU obtained by the taxonomic assignment explained before. The alignment was performed using MAFFT version 7.453 (Katoh *et al.* 2002; Katoh and Standley 2013) with the E-INS-i algorithm and revised with Geneious version 9.0.5 (www.geneious.com). After obtaining a reference alignment, the OTUs sequences were added with the function –addfragments from MAFFT. The final alignment was trimmed automatically using TrimAl version 1.4.rev22 (Capella-Gutierrez, Silla-Martinez and Gabaldon 2009) with -gt 0.3 and -st 0.001 parameters. We inferred the maximum-likelihood tree using IQ-Tree version 1.6.12 (Nguyen *et al.* 2015) with 1000 bootstrap replicates for SH-aLRT. The best model selected automatically by IQ-Tree was GTR+F+R10 (Nguyen *et al.* 2015; Kalyaanamoorthy *et al.* 2017). The final phylogenetic tree was visualized and annotated using Itol (Letunic and Bork 2019) and inkscape (https://inkscape.org/).

# RESULTS

## Environmental parameters

The East African Great Lakes displayed some remarkable differences in lake area, water depth, physical, chemical and biological parameters (Tables 1 and 2).

### *Surface mixed-layer waters*

During the sampled period, the trophic status of these water bodies varied from meso-eutrophic systems (Lake Edward) to less productive waters (Lake Kivu). Water temperatures were typical of tropical lakes (24.1°C at Ishungu-27.9°C in Lake Albert), whereas DO values ranged from 5.7 mg $L^{-1}$ in Ishungu to 8.5 mg $L^{-1}$ in L. Albert. Moreover, a wide range of conductivity values was observed, L. Kivu being up to two orders of magnitude higher than the other lakes (95 μS $cm^{-1}$ at Victoria, 1155 μS $cm^{-1}$ at Ishungu) (Table 2).

### *Vertical profile in Lake Kivu*

Lake Kivu profiles were characterized by a vertical gradient of several physical and chemical parameters, some of which differ between basins (Fig. 2). In general, surface waters displayed high DO but low nutrients concentrations; while $CH_4$ and conductivity grew gradually with depth, whereas pH decreased over the first 80-m depth, more drastically in Ishungu. The low mixing regime during the rainy season allowed the establishment of the epilimnion, with a mixed-layer depth that varied between 15 m in Ishungu to 25 m depth in Gisenyi (the main basin). This basin showed a steep oxycline from 25 to 40 m depth (transition zone), whereas the DO concentration profiles in Ishungu decreased more smoothly, with an oxycline extending from 15 to 50 m depth. Anoxic waters were detected from 45 and 55 m, in Gisenyi and Ishungu, respectively (Fig. 2A and B). $NO_x$ ($NO_2^-$ + $NO_3^-$) concentrations were low throughout most of the water

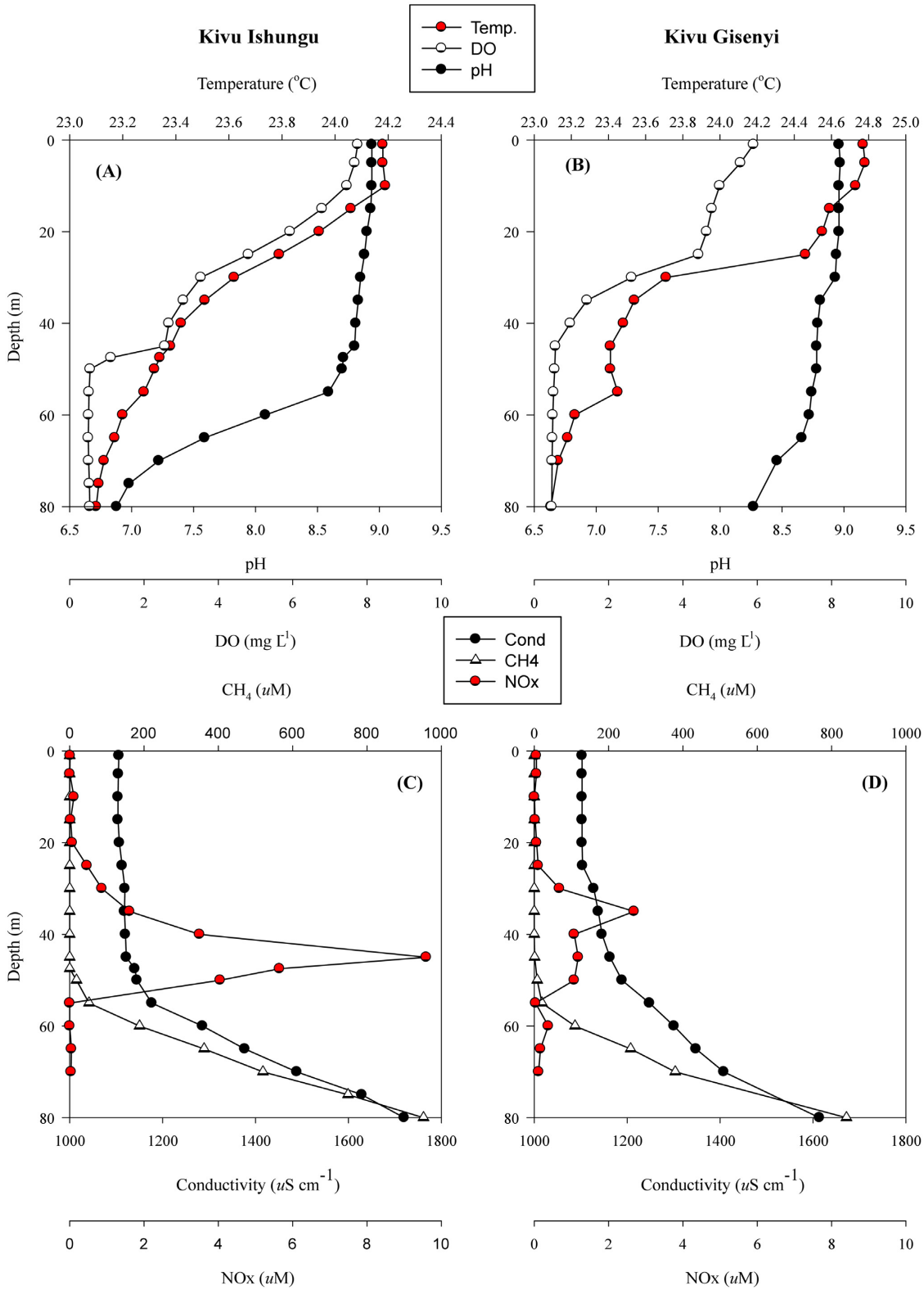


**Figure 2.** Vertical profiles of the main physical–chemical parameters of Lake Kivu at the Ishungu basin **(A, C)** and the Gisenyi basin **(B, D)**. **(A, B)** Temp.: temperature (°C), DO: dissolved oxygen (mg $L^{-1}$) and pH. **(C, D)** Cond: conductivity (µS $cm^{-1}$), $CH_4$: methane (µM) and $NO_x$: oxidized nitrogen (µM).

column for both basins, but an accumulation occurred in the oxic–anoxic transition zone with a peak at 35 m in Gisenyi (2.3 μM) and 45 m in Ishungu (9.6 μM) (Fig. 2C and D).

## Overall composition of protist communities

Across the 14 samples, we obtained a total of 86 463 quality-filtered sequences, derived from a variable sequencing effort among samples that ranged from 1064 sequences in Lake Victoria to 15 105 Ishungu at 15 m. These reads were grouped in 779 OTUs at 99% sequence identity. Rarefaction curves of OTU richness for each lake and depth indicated that some samples did not reach saturation, suggesting insufficient sequencing depth to capture all protist diversity (Fig. S1A and B, Supporting Information).

Total reads were spread among eleven high-rank supergroups or protistan lineages and grouped in 57 eukaryotic groups or classes. Most of the sequences were affiliated to Alveolata (31% OTUs, 44% reads), Opisthokonta (20% OTUs, 8% reads) and Stramenopiles (17% OTUs, 10% reads), while unclassified eukaryotes (Eukaryota) (3% OTUs, 0.6% reads), Amoebozoa (1% OTUs, 0.5% reads) and Telonema (0.3% OTUs, 0.09%) were the least represented (Fig. 1B). In terms of numbers of groups, the most diverse was Stramenopiles (15 groups), with Chrysophyceae displaying the highest percentage of OTUs (45%), followed by Alveolata (9 groups), dominated by Ciliophora (54% of OTUs) and Opisthokonta (8 groups), dominated by Basal_Fungi (47%), which comprises a diverse group of heterotrophic, saprophytic and parasitic organisms such as Microsporidiomycota and Chytridiomycota (Fig. S2, Supporting Information).

## Potential 'novel' diversity

Approximately 40.6% of OTUs (15.5% reads) displayed low CCM and CEM similarity (<90 or <95%, depending on the supergroup) with reported sequences in public databases (Fig. 3). Alveolata (38% OTUs, 35% reads), Opisthokonta (16% OTUs, 15% reads) and Stramenopiles (10% OTUs, 15% reads) showed the highest 'novel' diversity (Fig. 1B). Among alveolates, 29% of OTUs affiliated more closely with Ciliophora, within Opisthokonta, 18.7% of OTUs were associated with Choanomonada, whereas among Stramenopiles, 13% of OTUs were more closely affiliated to Chrysophyceae.

## Protist communities in surface mixed-layer waters

The $\alpha$-diversity indices differed in surface mixed-layer of East African Great Lakes (Table 3). L. Victoria displayed the most dissimilar values. Protists OTUs richness ranging from 96 in L. Kivu off Gisenyi to 125 in L. Victoria. Similarly, the main basin of L. Kivu had the lowest and highest diversity values, respectively (inverse Simpson index: 0.85–0.93, Shannon index: 2.97–3.59; respectively). While the evenness ranged from 0.17 in L. Edward to 0.29 in L. Victoria.

The number of reads varied widely, with the lowest number in L. Victoria (1064) and the highest in L. Edward (13 129) (Fig. S1A, Supporting Information). In order to compare the microbial eukaryote composition among all lakes, we standardized the samples at the smallest number. The standardized dataset resulted in an overall number of 5320 reads. The relative numbers of reads of taxonomic groups varied between lakes; however, reads affiliated to Alveolata (mean 40.4%), Ciliophora (mean 23%) and Dinoflagellata (mean 17.7%) being the most abundant, followed by Cryptista (mean 30.8%), mainly Cryptomonadales (mean 26.2%), accounted for the majority of sequences in all lakes (Fig. 4A). Amoebozoa being found in I-5 (0.28%) and V (0.09%), while only 2 OTUs more closely affiliated to Telonema were detected only in E (0.17%). Those were the least abundant supergroup (not displayed in the figure). On average, 12.7% of reads were considered 'novel' in the surface mixed-layer waters.

Considering the most abundant OTUs (OTUs ≥ 1%), only 14 (66.7% of total reads) were dominant (Fig. S3A, Supporting Information). Lakes Albert and Victoria, following by L. Edward had a similar composition in the proportion of eukaryotes, whereas both basins of L. Kivu differed in their protistan assemblage. Ishungu was characterized by the dominance of OTU_620, affiliated closely to Cryptomonadales, while Gisenyi for the occurrence of OTU_1404 (Dinoflagellata). The potential 'novel' OTU_1814 affiliated to Cryptomonadales with 86.7% of similarity (representing 6.6% of the total reads) was observed in different proportions in all lakes.

Multivariate analysis conducted on microbial eukaryote communities showed a segregation of lakes at the surface level (Fig. 4B). The PCA plot over the first two axes (70.5% of total variation) revealed L. Kivu as the most dissimilar, with the two sites separated from the rest of the water bodies. The first axis explained 40.55% of total cumulative variation and it was mainly associated positively with altitude ($r = 0.87$), maximum depth ($r = 0.71$) and conductivity ($r = 0.59$), and negatively with Chl-*a* ($r = -0.94$), DIN ($r = -0.84$), temperature ($r = -0.83$), DO ($r = -0.63$) and SRP ($r = -0.44$). The second axis (29.93% of total variance) was related mainly by conductivity ($r = 0.48$). The PCA showed a positive association between the variables SRP, DIN and Chl-*a* and the lakes Albert and Victoria, suggesting productive systems. These lakes displayed the most similar community structure distinguished by the presence of Diplonema and MALVs (marine alveolates), as well as a large number of Chrysophyceae and Kinetoplastida, among others. Diversity in L. Edward was mainly explained by Chl-*a* and showed a diverse microbial eukaryote community. On the other hand, L. Kivu was associated with higher values of conductivity, depth and altitude. In particular, at Ishungu a more heterotrophic community was found, including fungi and parasites, while Gisenyi displayed a higher abundance of MASTs (marine Stramenopiles), Charophytes and Dinoflagellates. Further, only 18 OTUs were shared by all the freshwaters systems, most of them being abundant (61.1% of reads) (Fig. 4C), with ciliates and Cryptomonadales as the most represented (27.7% of the total). In fact, ~64% of the OTUs observed in the surface mixed-layer waters was specific to a single lake.

## Vertical profiles of Lake Kivu's protist communities

In L. Kivu, the normalized dataset resulted in an overall number of 17 908 reads in 11 separate samples, corresponding to 544 OTUs. The southern part of the lake (Ishungu) was more diverse (Table 4), with higher values of $\alpha$-diversity indices at 70 m (Richness: 216, inverse Simpson index: 0.98, Shannon index: 4.45, Evenness: 0.40).

Most OTUs affiliated to Alveolata (mean 47.6%) and Cryptista (mean 22.6%). Along the first 30 m of Ishungu, Cryptista (40.3% of total reads) dominated, being Cryptomonadales (38.7% of total reads) the most abundant group (Fig. 5A). After 45 m depth, there was a change to a more heterotrophic protist community. At 45 m (I-45), 38.9% of reads corresponded to Dinoflagellata (Alveolata) and 36.9% affiliated to diverse heterotrophic groups; whereas in I-50 and I-70 Chrysophyceae (Stramenopiles) appeared in greater proportion (18.3%). Conversely, in Gisenyi

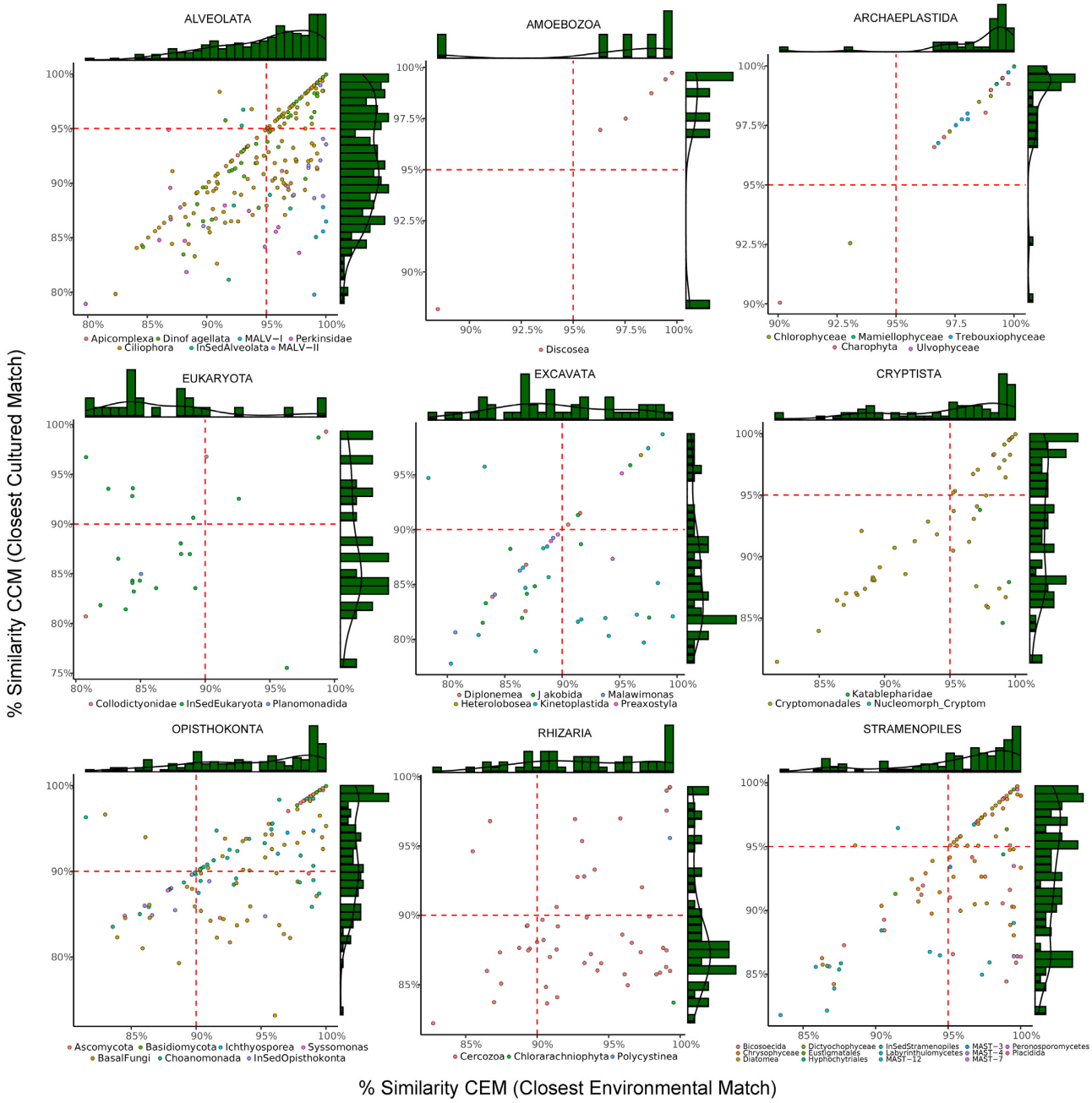


**Figure 3.** Novelty pattern derived from OTUs among supergroups. Dots represent the percentage of similarity with the CEM and the CCM for each OTU within each supergroup. Bars represents abundance frequencies.

**Table 3.** $\alpha$-Diversity indices of East African Great Lakes in surface mixed-layer protistan assemblages.

| Sample | Kivu Ishungu (I-5) | Kivu Gisenyi (G-5) | Edward (E) | Albert (A) | Victoria (V) |
|---|---|---|---|---|---|
| Richness (OTUs numbers_S) | 122 | 96 | 113 | 98 | 125 |
| Diversity (Simpson 1-D) | 0.91 | 0.85 | 0.88 | 0.92 | 0.93 |
| Diversity (Shannon_H) | 3.35 | 2.97 | 2.98 | 3.30 | 3.59 |
| Evenness_e^H/S | 0.23 | 0.20 | 0.17 | 0.27 | 0.29 |

Kivu Ishungu: 5 m (I-5) (0.2–5 μm); Kivu Gisenyi: 5 m (G-5) (0.2–5 μm); Edward (E) (0.8–10 μm); Albert (A) (0.8–10 μm); Victoria (V) (0.8–10 μm).

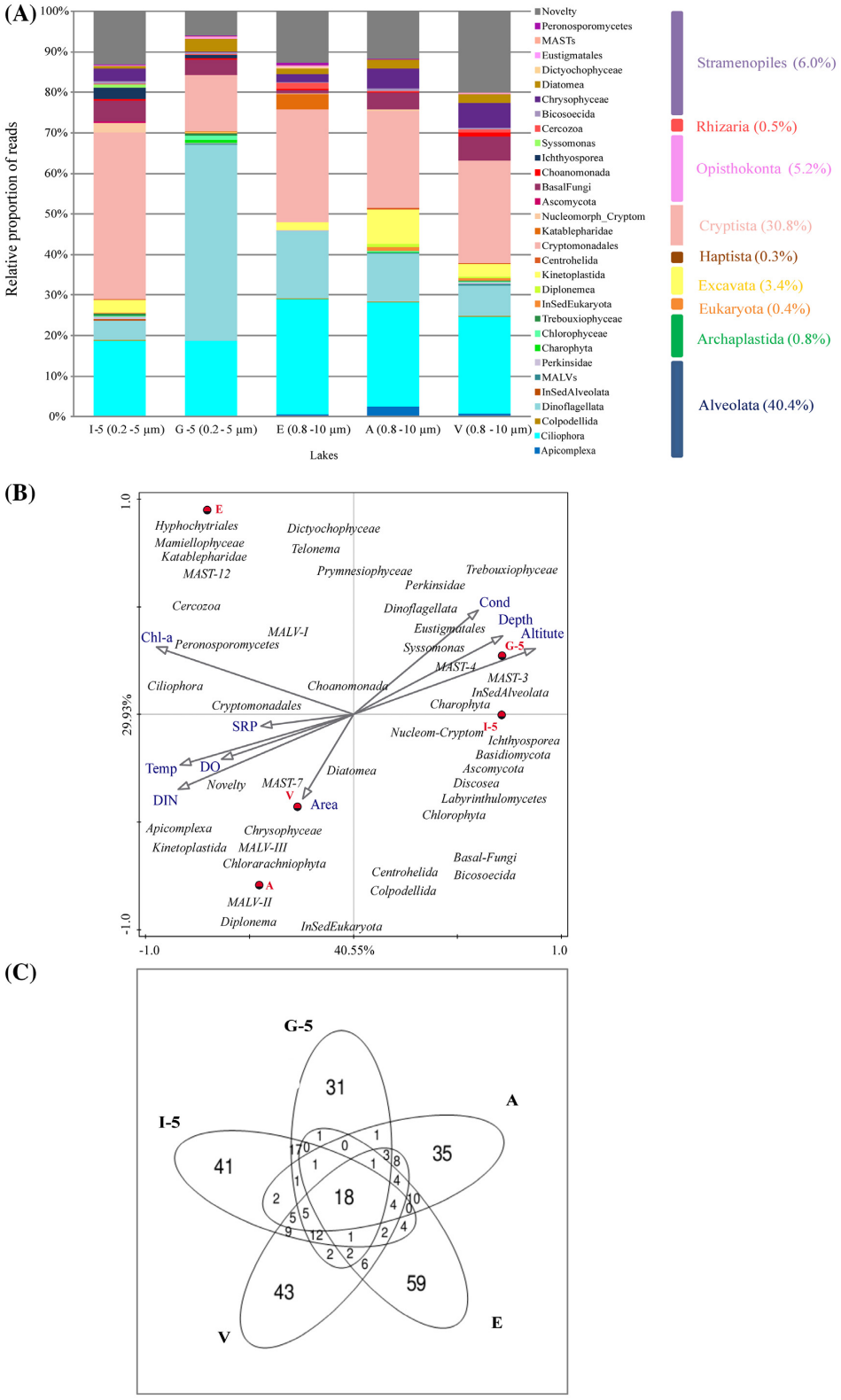


**Figure 4.** Community structure and taxonomic composition of surface protistan assemblages in the four East Great African Lakes (two sites in Kivu). **(A)** Relative proportion of reads within taxonomic groups. For a better interpretation, the graph was made with groups that presented >5 reads in all the samples. **(B)** PCA based on microbial communities interpreting the variation with the supplementary physico-chemical parameters. Samples appear in solid red circle and each arrow points in the direction of the steepest increase of the corresponding values. Taxonomic groups enriched in a given part of the plot are shown. Albert -A-, Victoria -V- and Edward -E-: 0.8–10 µm; I-5 and G-5 -Ishungu and Gisenyi at 5 m-: 0.2–5 µm. (Temp, water temperature; Cond, conductivity; Chl-*a*, chlorophyll *a*; DIN, dissolved inorganic nitrogen; SRP, soluble reactive phosphorous; DO, dissolved oxygen; Depth, maximum depth). **(C)** Venn diagram showing OTUs shared by the five surface samples.

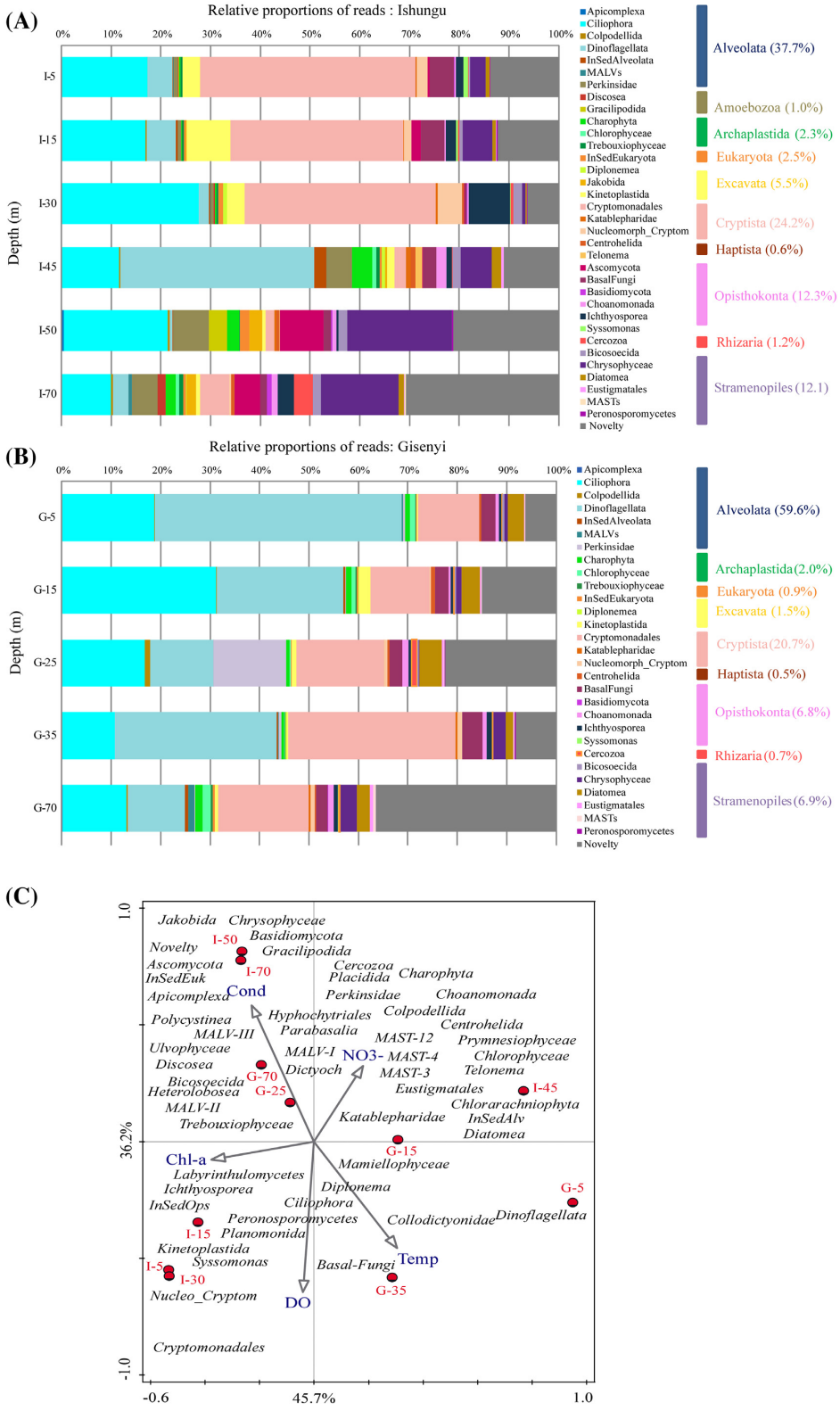


**Figure 5.** Community structure and taxonomic composition of protistan assemblages over vertical profiles of Lake Kivu. **(A)** Ishungu and **(B)** Gisenyi. For a better interpretation, histograms were made with taxonomic groups that presented >5 reads in all the samples. **(C)** PCA based on microbial communities interpreting the variation with the supplementary physico-chemical parameters. Samples appear in solid red circles and each arrow points to the direction of the steepest increase of the corresponding values. Taxonomic groups enriched in a given part of the plot are shown. Ishungu: I-5, I-15, I-30, I-45, I-50 and I-70. Gisenyi: G-5, G-15, G-25, G-35 and G-70. (Temp, water temperature; Cond, conductivity; Chl-*a*, chlorophyll *a*; DO, dissolved oxygen; $NO_3^-$, nitrates).

**Table 4.** $\alpha$-Diversity indices of Lake Kivu vertical protistan assemblages.

| Sample | I-5 | I-15 | I-30 | I-45 | I-50 | I-70 | G-5 | G-15 | G-25 | G-35 | G-70 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Richness (OTUs numbers_S) | 133 | 146 | 117 | 146 | 149 | 216 | 104 | 150 | 120 | 110 | 207 |
| Diversity (Simpson 1-D) | 0.90 | 0.93 | 0.89 | 0.86 | 0.95 | 0.98 | 0.85 | 0.92 | 0.92 | 0.83 | 0.97 |
| Diversity (Shannon_H) | 3.31 | 3.54 | 2.92 | 3.30 | 3.67 | 4.45 | 2.94 | 3.84 | 3.23 | 2.73 | 4.13 |
| Evenness_e^H/S | 0.21 | 0.24 | 0.16 | 0.18 | 0.26 | 0.40 | 0.18 | 0.22 | 0.21 | 0.14 | 0.30 |

Ishungu: I-5, I-15, I-30, I-45, I-50 and I-70. Gisenyi: G-5, G-15, G-25, G-35 and G-70.

the abundance of Alveolata gradually declined over the vertical profile, being Dinoflagellata (mean 26.4%) and Ciliophora (mean 18.3%) the most important groups (Fig. 5B). The potential microbial eukaryotic novelties were highest in deeper samples at both sites (33.4% on average at 70 m).

Multivariate analysis revealed a great dissimilarity between sites and depths (Fig. 5C). The PCA plot with supplementary variables accounted for 76.2% of the total variance. Axis 1 explained 45.7% and was positively related to Temp ($r = 0.30$) and negatively to Chl-*a* ($r = –0.38$). Axis 2 explained 36.3% and was positively related to conductivity ($r = 0.58$) and negatively to DO ($r = –0.64$) and Temp ($r = –0.46$). Microbial eukaryote communities were distributed according to physical–chemical parameters. The shallower samples (up to 35 m) corresponding to the two sites were distributed in the lower part of the diagram mainly associated with higher temperatures, DO and Chl-*a* concentrations. The deepest samples from Ishungu (I-50 and I-70) forming a separate group in the upper left side of the plot mainly related to high conductivity values; while I-45 were the most isolated, more associated with high nitrate concentrations, presenting a particular assemblage of microorganisms, many of which were heterotrophic.

Taking into account the composition of most abundant OTUs (OTUs $\geq$ 1%), only 24 were dominant (62.5% of total reads), which also differed between sites and depths (Fig. S3B and C, Supporting Information). Usually, Ishungu displayed a more dissimilar eukaryotic microbial community through depths. The firsts 30 m of this basin were dominated by the OTU_620 (Cryptomonadales); whereas OTU_1404 (Dinoflagellata) abounded in Gisenyi. Besides, three 'novel' OTUs closely affiliated to Ciliates were also noteworthy (OTU_452, OTU_394 and OTU_2540).

The 'novel' OTUs affiliating to Alveolates and Stramenopiles in L. Kivu were explored by phylogeny. In total 327 sequences were retrieved from NCBI and aligned to construct two reference alignments. OTUs affiliated to Alveolata (201 OTUs) and Stramenopiles (98 OTUs) were added to the respective reference alignment for their phylogenetic classification. Most Stramenopiles OTUs were affiliated to Chrysophyceae (14 OTUs) and Bicosoecida (9 OTUs), mostly observed at great depths in both basins of L. Kivu (I-70 and G-70) (Fig. 6). The 'novel' OTUs 136, 917, 1941 and 2316, together with sequences from deep anoxic marine environments (HM749941 and AB505560), were grouped in a well-supported environmental cluster mostly observed in G-70. 'Novel' Chrysophyceae OTUs were related to Clades C, F2, F1 and A, formed mostly by heterotrophic organisms from freshwater and marine environments. Furthermore, 6 'novel' OTUs were affiliated to MASTs, 3 of which were grouped into MASTs 12 (OTU_1894 and 1644 to MAST-12C, OTU_1766 to MAST-12A), 2 OTUs were related to MAST 4/7 (OTU_1606 to MAST-7B and OTU_693 to MAST-4A) and OTU_1445 was affiliated to MAST-3J together with EU561853.2, a sequence from uncultured marine eukaryote.

Among Alveolata (Fig. 7), 87 OTUs were classified as 'novel', most of them related to Ciliophora (56 OTUs) and Dinoflagellata (19 OTUs). Dinoflagellate 'novel' OTUs were mostly found at surface level (at 5 and 15 m), while most Ciliophora 'novel' OTUs were observed in deeper samples (I-50, I-70 and G-70), highly supported by clades closer to Prostomatea (Fig. 8). Moreover, a great diversity of 'novel' OTUs was related to Oligohymenophorea and 15 OTUs affiliated to MALVs (Syndiniales). Among them, 4 to MALV-I, 10 to MALV-II and 1 to MALV-III, (Table S2, Supporting Information); most of them were observed with relative high abundance at I-70 and G-70 samples (Fig. 7).

## DISCUSSION

Despite the key role played by protists in aquatic ecosystems (Sherr and Sherr 1988), their diversity and function is still underestimated in freshwater (Schloss *et al.* 2016; Debroas *et al.* 2017), even more in meromictic lakes (e.g. Tarbe *et al.* 2011a; Lepère *et al.* 2016), although they provide a high diversity of niches along stratification gradients (Oikonomou, Pachiadaki and Stoeck 2014). East African Lakes are extremely old and their large biodiversity is supported by complex processes that include potential endemic organisms (Seehausen 2006; Descy and Sarmento 2008). Our study revealed a high and potential 'novel' diversity of small eukaryotes (<10 μm) in African Great Lakes, associated with different physicochemical variables, especially in the meromictic Lake Kivu.

Limnological features of these lakes are similar to those previously reported (Sarmento *et al.* 2008; Pasche *et al.* 2009; Morana *et al.* 2014). Although all systems are tropical, they differed concerning abiotic parameters, catchment dynamics, human impacts and food webs (Odada *et al.* 2003).

Protist communities are poorly known in African freshwaters (Llirós *et al.* 2012). A high diversity was found in an earlier study by Tarbe (2010) in surface waters of Tanganyika and Kivu. Accordingly, a large number of protistan lineages or supergroups were revealed in our work. Over the past decades, the eukaryote Tree of Life has been divided into several supergroups and multiple lineages of uncertain placement (Massana *et al.* 2011; Ortiz-Álvarez *et al.* 2018; Burki *et al.* 2020). The configuration of supergroups varies constantly as more studies are conducted. Recently, Burki *et al.* (2020) rearranged the 'tree' and the term Hacrobia, previously described by Okamoto *et al.* (2009), was split into Haptista and Cryptista, while Telonema became part of TSAR. Phylogenomic analyses infer a Metamonada plus Discoba clade corresponding to Excavata, and some recover instead a specific relationship with the excavate group malawimonads. As these relationships are still uncertain, we describe our results using the original term Excavata and we incorporate these new findings as the eukaryotic phylogeny progressively resolves (Fig. 1B). Thereby, in African lakes, assemblages


**Figure 6.** Phylogenetic tree and classification of stramenopile OTUs. From the phylogenetic tree, circles in red correspond to bootstrap support values >75%. Clusters highlighted in orange correspond to environmental clusters that were formed by novel OTUs. The collapsed clusters were annotated with the number of OTUs followed by the number of novel OTUs and the number of reference sequences (rs). The right panel shows the number of reads from each OTU in each sample from Ishungu (I-5, I-15, I-30, I-45, I-50 and I-70) and Gisenyi (G-5, G-15, G-25, G-35 and G-70). The circles represent the presence of the OTU in the corresponding sample and the values inside are the number of reads.

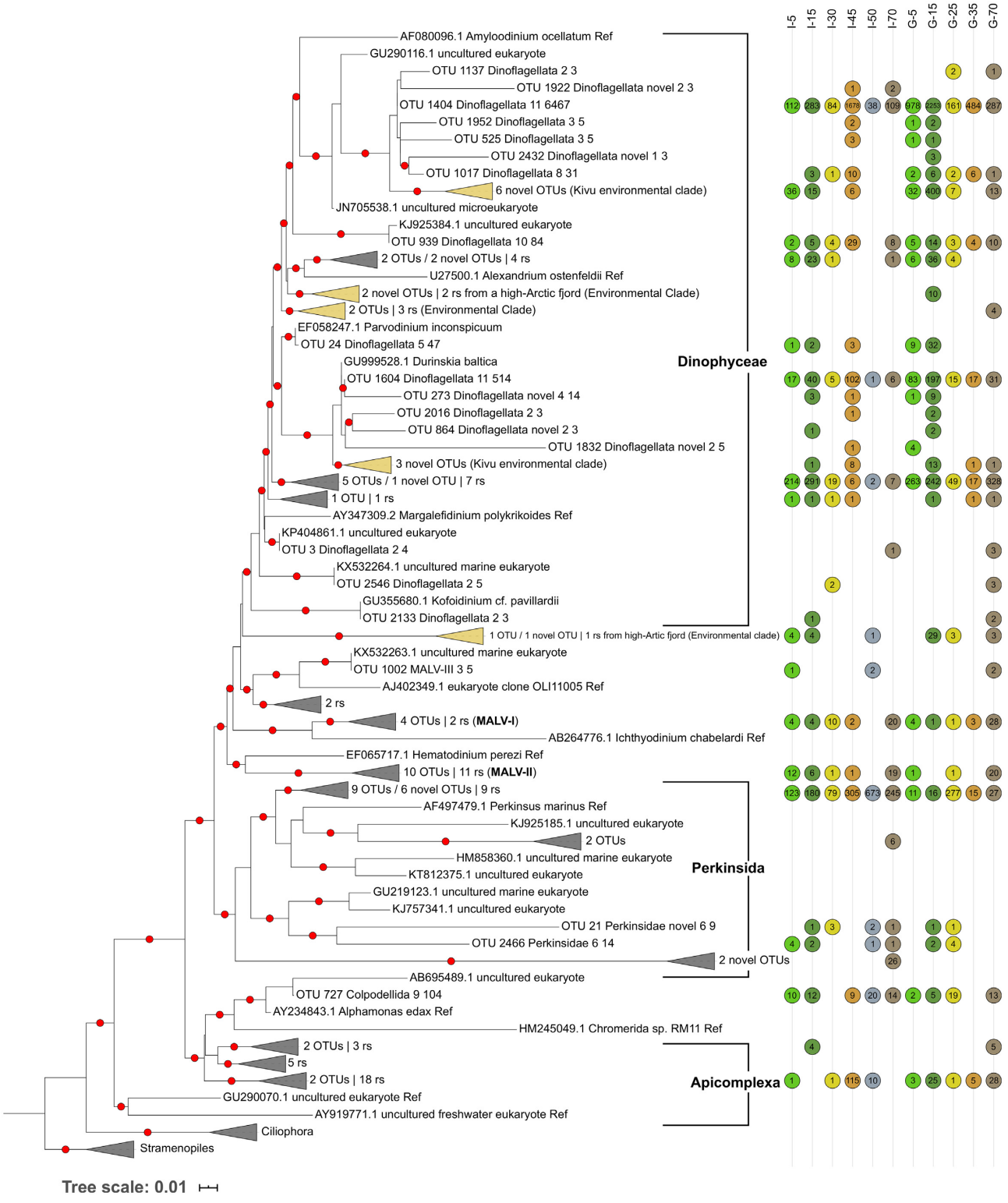


**Figure 7.** Phylogenetic tree and classification of Alveolata OTUs. From the phylogenetic tree, circles in red correspond to bootstrap support values >75%. Clusters highlighted in yellow correspond to environmental clusters that were formed by novel OTUs. The collapsed clusters were annotated with the number of OTUs followed by the number of novel OTUs and the number of reference sequences (rs). The right panel shows the number of reads from each OTU in each sample from Ishungu (I-5, I-15, I-30, I-45, I-50 and I-70) and Gisenyi (G-5, G-15, G-25, G-35 and G-70). The circles represent the presence of the OTU in the corresponding sample and the values inside are the number of reads.

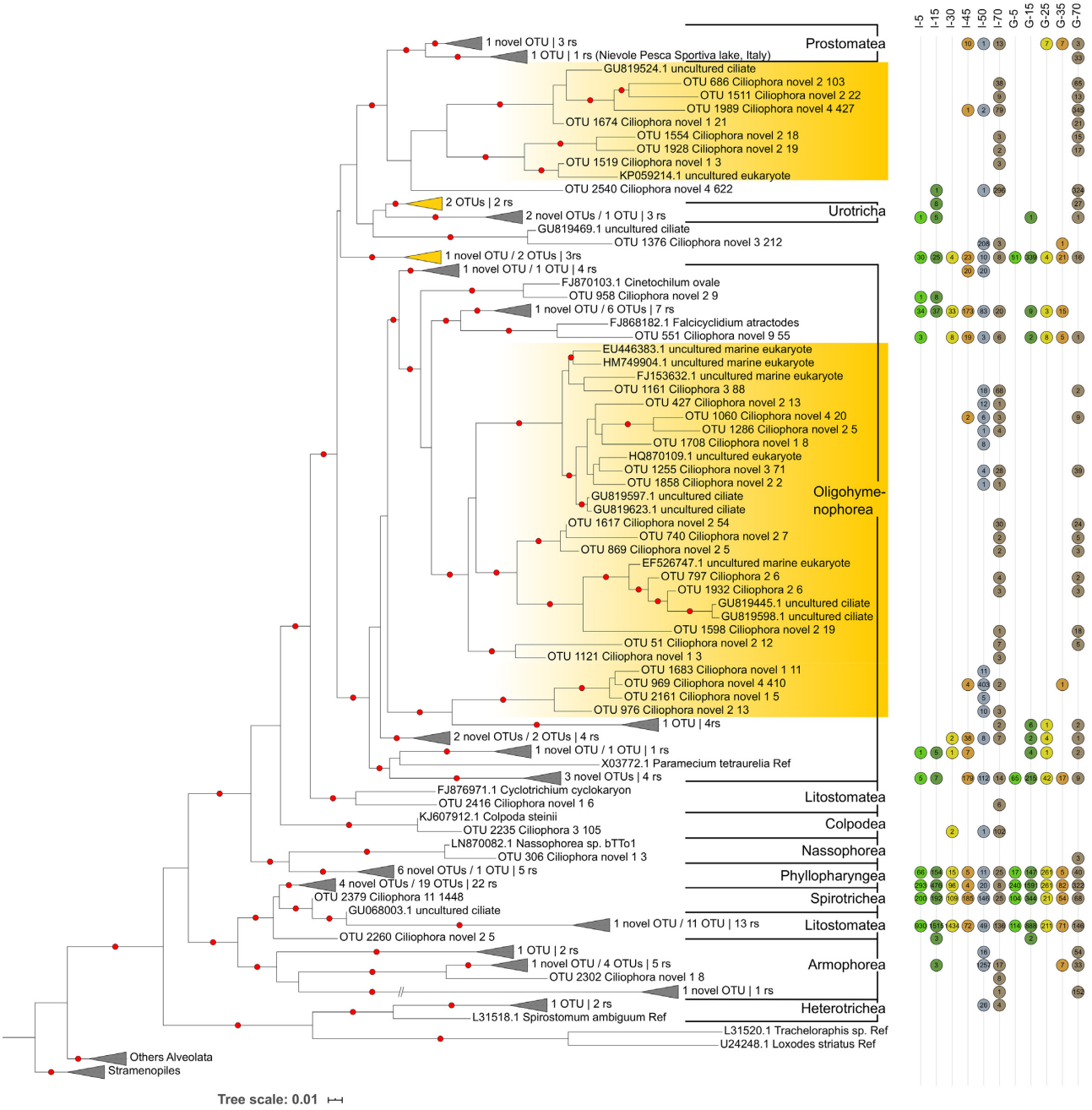


**Figure 8.** Phylogenetic tree and classification of ciliate OTUs. From the phylogenetic tree, circles in red correspond to bootstrap support values >75%. Clusters highlighted in yellow correspond to environmental clusters that were formed by novel OTUs. The collapsed clusters were annotated with the number of OTUs followed by the number of novel OTUs and the number of reference sequences (rs). The right panel shows the number of reads from each OTU in each sample from Ishungu (I-5, I-15, I-30, I-45, I-50 and I-70) and Gisenyi (G-5, G-15, G-25, G-35 and G-70). The circles represent the presence of the OTU in the corresponding sample and the values inside are the number of reads.

were dominated by Alveolata, Opisthokonta and Stramenopiles, according to Tarbe (2010) and those reported from other freshwater systems (Triadó-Margarit and Casamayor 2012; Simon *et al.* 2015; Ortiz-Álvarez *et al.* 2018). A recent study revealed that phototrophs dominated in freshwaters, while parasitic taxa represented roughly 15 to 20% of all sequences in marine and soil ecosystems (Singer *et al.* 2021). Here, a large number of heterotrophic phagotrophs and parasites organisms were also found and our results more closely resemble those found in Lake Baikal, the oldest lake in the world, as we will discuss in detail below (Annenkova, Rodriguez-Giner and Logares 2020; David *et al.* 2021).

## Potential 'novel' diversity

A high percentage of potential 'novel' organisms was found, since 40.6% of OTUs had low similarity with reported sequences in public databases (Triadó-Margarit and Casamayor 2012; Massana *et al.* 2015). Previous studies in some groups have already shown that 'novelty' is larger in freshwater than in

marine systems (del Campo and Massana 2011), probably due to its heterogeneous abiotic parameters compared with the relatively buffered conditions in the ocean (Singer *et al.* 2021). In high mountain lakes, a high 'novel' protists, mainly for Rhizaria and Opisthokonta supergroups, as well as Alveolata and Stramenopiles was reported (Triadó-Margarit and Casamayor 2012; Ortiz-Alvárez *et al.* 2018). A high degree of 'novelty' has been also found in picoeukaryotes from pampean lakes in Argentina (Metz *et al.* 2019). The percentage found in our work is a slightly greater than that recorded, likely explained by the fact that freshwater tropical environments are undersampled (Llirós *et al.* 2012; Sarmento 2012). Likewise, some groups could be endemic, since in ancient lakes microorganisms could have had time to diversify, as already reported for other species in these lakes (Seehausen 2006) and for protists in Baikal (Annenkova, Rodriguez-Giner and Logares 2020). The environmental heterogeneity and persistently extreme conditions might promote an adapted specialist microbiota (Catalan *et al.* 2006). Regardless, all these studies highlight that freshwater planktonic diversity remains largely to be explored.

## Protist community in surface mixed-layer waters

A high microbial diversity and richness in surface mixed-layer waters was found, probably to a wide variation in the physical and chemical properties (Debroas *et al.* 2017; Boenigk *et al.* 2018). The eukaryote assemblage was dominated by Ciliophora and Dinoflagellata, followed by Cryptomonadales (Fig. 4A; Fig. S3A, Supporting Information), as reported in tropical and freshwaters systems (Tarbe *et al.* 2011b; Kammerlander *et al.* 2015; Simon *et al.* 2016; Debroas *et al.* 2017). Usually, alveolate reads are found in smaller fractions because they have a high copy number of the SSU rRNA gene, that presumably break down and pass through the filter (Dyal *et al.* 1995; Galluzi *et al.* 2004). Nevertheless, small organisms were also observed under the microscope and the relative abundance of groups varied between lakes.

The mesotrophic lakes Albert and Victoria presented a similar eukaryote community structure. Among Discoba (Excavates), Diplonemea and Kinetoplastida were distinguished. These two heterotrophic flagellates are generally abundant in the sea, but were recently reported in freshwaters (Boenigk, Wodniok and Glücksman 2015; Pernice *et al.* 2016; Mukherjee *et al.* 2019). In particular, some groups previously thought to be exclusively marine have been identified in the plankton of Lake Baikal (David *et al.* 2021). In general, Excavata are very poorly represented when the V4 region is amplified (Obiol *et al.* 2020), therefore our abundances are likely to be underestimated. Moreover, MALV-I, -II and -III, inhabiting frequently marine environments (Massana *et al.* 2011; de Vargas *et al.* 2015), were also found. Those groups are extremely rare in freshwaters. However, MALVs were recorded in a meromictic Arctic lake (Charvet *et al.* 2012) as well as in Baikal (Annenkova, Rodriguez-Giner and Logares 2020; David *et al.* 2021). The potentially 'novel' OTU_1814, closely affiliated to Cryptomonadales (86.7% of similarity), was rather abundant in L. Victoria (Fig. S3A, Supporting Information). On the other hand, L. Edward was associated with higher Chl-*a*, temperature, DO and nutrient values in the PCA (Fig. 4B). These characteristics support the growth of a large number of microorganisms including aerobic and heterotrophic protists, like Ciliophora and Cercozoa (Liu, Li and Chai 2021). In this lake, within minority groups also highlights the small generalist Katablepharidae, which is usually found in low abundances in lakes and oceans (Simon *et al.* 2015; Boenigk *et al.* 2018) and Dictyochophyceae, almost exclusively marine photosynthetic algae, also found in lakes Taihu (Chen *et al.* 2008) and Baikal (Annenkova, Rodriguez-Giner and Logares 2020). These organisms possibly belong to Pedinellalaes, common in freshwaters (Liu, Li and Chai 2021).

Both basin of Lake Kivu exhibited a dissimilar community structure. Associated with high conductivity, scarce nutrients, low temperature and Chl-*a* concentrations, the oligotrophic L. Kivu harbors unique limnological conditions and particular protistan assemblages. Surface waters are usually colder than the rest of the Rift lakes, as a consequence of its elevation (1463 m) (Schmid *et al.* 2005). Ishungu is located in the smaller southern basin, which is less exposed to winds and harbors a large amount of Cryptomonadales. These mixotrophic flagellates can survive under low light condition, allowing the exploitation of oligotrophic environments (Oikonomou, Pachiadaki and Stoeck 2014; Boenigk, Wodniok and Glücksman 2015). Here, the percentage of alveolates decreased to 24.5% and Perkinsidae (0.10%) were detected. Opisthokonta reached 9.6%, being Ascomycota, Basal_Fungi, Ichthyosporea and Syssomonas the most abundant groups. A large number of fungi were observed, in contrast to temperate lakes (Lefèvre *et al.* 2007; Lepère, Domaizon and Debroas 2008) and the findings by Tarbe *et al.* (2011b) in Tanganyika. Fungi is one of the most understudied microbial groups in aquatic systems, therefore its distribution and role in freshwaters environments are largely unknown. Recently, Lepère *et al.* (2019) found a high diversity in lakes and rivers, suggesting that freshwater fungi have been undersampled. In addition, unlike Tanganyika (Tarbe *et al.* 2011b), Bicosoecida were also registered. Ishungu is located in a bay close to the only outflow through the Ruzizi River that discharges into Lake Tanganyika (Sarmento, Darchambeau and Descy 2012). Its proximity and the same trophic state between both lakes could suggest a similar protist assembly. Nevertheless, as food webs are different, small eukaryotes communities seem to be specific, as found by Tarbe (2010) with less than 11% of shared sequences. The thermal stratification of the water column reduces nutrient supply in the euphotic zone, which is generally shallower at Ishungu, limiting phytoplankton growth (Descy, Darchambeau and Schmid 2012; Sarmento, Darchambeau and Descy 2012). All these characteristics could favor mixotrophy, phagotrophy and parasitism as a strategy of organisms to obtain energy in this site.

The deepest basin, Gisenyi, presented a dissimilar microbial eukaryotic community compared with Ishungu. Conductivity was higher in this site (Schmid *et al.* 2005; Olapade and Omitoyin 2012), favoring thecate organisms. Among most abundant taxa, the OTU_2553 was related to a dinoflagellate (99.5% of similarity), already found in Tanganyika (Tarbe *et al.* 2011b). Dinoflagellates distribution is known to be influenced by light, temperature and salinity, being commonly distributed in the sea as well as in large freshwater bodies (Tarbe *et al.* 2011a; Khomich *et al.* 2017; Órtiz-Alvarez *et al.* 2018). Among less frequent groups, two Archaeplastida (Chlorophyceae and Charophyceae) and two Stramenopiles (Diatomea and MASTs) were detected. Chlorophyceae are a very important group of algae in continental aquatic systems (Khomich *et al.* 2017; Metz *et al.* 2019), but in terms of abundance, they tend to be poorly represented in most African great lakes (Tarbe *et al.* 2011b). Charophyta are structurally complex green algae, frequently found in hard waters, like Baltic Sea (Schubert and Blindow 2004) and in tropical brackish lagoons impacted by humans (Palma-Silva, Albertoni and Esteves 2004). Diatoms are among the major groups of photosynthetic organisms, contributing to a large part of the Earth's primary production, inhabit a wide variety of environments (Verleyen *et al.* 2009) and already observed by Sarmento, Isumbisho and Descy (2006).

On the other hand, we found a few OTUs belonging to 'typically marine' lineages MAST-3 and MAST-4 (Massana *et al.* 2004; Logares *et al.* 2012), although some of them have already been registered in freshwaters (Charvet *et al.* 2012; Simon *et al.* 2014; Arroyo *et al.* 2018).

Overall, the lakes showed a small eukaryote community structure associated with conductivity values and productivity gradients. A similar assemblage associated with the most productive lakes Albert, Victoria and Edward was found in temperate lakes (Lefranc *et al.* 2005; Simon *et al.* 2015). On the other hand, differences between sites in the oligotrophic L. Kivu seemed to be due to OTUs ≥1%, which generally contribute to biomass production, as well as minor groups, that could regulate the functioning of aquatic habitats (Debroas, Hugoni and Domaizon 2015; Du *et al.* 2019). Indeed, Ishungu presented a more heterotrophic community, whereas the microbial food web would be supported by autochthonous organic carbon sources in Gisenyi.

## How is the community structure along the depth gradient in both basins of Lake Kivu?

The water column structure in L. Kivu is different from most other large lakes in the world (Schmid and Wüest 2012). Because of their morphometry, the influence of the nearby volcanoes, groundwater inflows and anthropogenic activities, both basins differed in vertical abiotic parameters (Fig. 2) (Llirós *et al.* 2012; İnceoğlu *et al.* 2015), providing different conditions for life by microbial eukaryote communities down the water column. These conditions were reflected in our study, with noticeable differences between sites and depth, as demonstrated in an Arctic meromictic lake (Charvet *et al.* 2012) and in Lake Baikal (Annenkova, Rodriguez-Giner and Logares 2020).

The first 30 m of Ishungu were dominated by Cryptomonadales (∼40% of the total reads), generalist taxa occurring in many environmental conditions (Boenigk *et al.* 2018). Moreover, we highlight the presence of Kinetoplastida, being OTU_1531 among the most abundant (Fig. S3B, Supporting Information). This group is ecologically important because of their diverse mode of feeding (Salani *et al.* 2012), containing many bacterivorous (e.g. *Bodo*) as well as endoparasitic lineages (e.g. *Trypanosoma*), who could infect fishes, with the consequent economic losses (Corrêa *et al.* 2016). Associated with different habitats within marine and freshwater environments (Yubuki and Leander 2018), Kinetoplastida were previously reported by Tarbe *et al.* (2011a) in the epilimnion of Tanganyika and Kivu, and in our study *Trypanosoma* spp. was detected with 85% of CCM similarity. From 45 m depth down, microorganism's diversity changes abruptly to a more heterotrophic community (Table 4; Fig. 5A). At 45 m depth, a $NO_x$ peak was recorded (Fig. 2C) and the oxic–anoxic transition zone was dominated by the nitrite-oxidizing *Nitrospira* (İnceoğlu *et al.* 2015). Here, dinoflagellates (38.8% of total reads) dominated, being OTU_1404 the most abundant (Fig. S3B, Supporting Information). Below this zone (I-50 and I-70), anoxic waters prevailed. Others studies have shown that a diverse methanotrophic bacterial community was highly active in these layers (İnceoğlu *et al.* 2015; Morana *et al.* 2016), which could support a complex assemblage of protists (Taylor *et al.* 2006; Llirós *et al.* 2012). Indeed, a greater diversity was found, dominated by heterotrophic and mixotrophic organisms (Ciliates and Chrysophyceae), saprotrophic (Fungi), parasitic (Perkinzoa) and many other minor heterotrophic groups.

In the epilimnion of Gisenyi, the community structure was dominated by dinoflagellates and ciliates. The decline of these groups at 25 m could have been due to an increase of Perkinsidae, which reached 14% of total reads (Fig. 5B). Perkinsidae was currently found in marine systems (Massana *et al.* 2015), but many of them also began to be reported in freshwaters (Lefranc *et al.* 2005; Ortiz-Alvarez *et al.* 2018), and also in the water column of L. Baikal (David *et al.* 2021). This parasitic group can infect a wide variety of organisms. Parasites infecting protists as ciliates and dinoflagellates could directly influence the food web (Lepère, Domaizon and Debroas 2008; Ortiz-Alvarez *et al.* 2018; Cruaud *et al.* 2019). Thus, the high relative abundance found specifically at this depth could imply the importance of parasitism in regulating these protists. As in Ishungu, a high percentage of 'novel' OTUS were found in the anoxic waters.

Overall, the protistan assemblages in both basins in Lake Kivu were dissimilar, reflecting differences in abiotic parameters along the depth profile. Studies in tropical and meromictic lakes are scarce, therefore this work provides new insights into these specific ecosystems. We revealed a high genetic diversity, particularly in anoxic waters, as was already reported in marine (Pernice *et al.* 2016; Zhao *et al.* 2017) and in some meromictic systems (Oikonomou, Pachiadaki and Stoeck 2014; Lepère *et al.* 2016). Photosynthetic algae can inhabit extreme environments, being able to survive as cysts (Lepère *et al.* 2016). However, the community of the lower mixolimnion layers was composed mainly by heterotrophic, mixotrophic and saprotrophic organisms, highlighting their important role in the food web (Lepère, Domaizon and Debroas 2008). Furthermore, the extremely high potential 'novelty' of taxa recorded in the deeper layers was remarkable, particularly in Alveolata and Stramenopiles supergroups. These taxa also showed a substantial proportion of clones (10–20%; c. 162 sequences) with low relatedness with any previously reported sequence in GenBank in Pyrenean lakes (Triadó-Margarit and Casamayor 2012). Also, in a meromictic lake, unclassified ciliate had highest representation at 29 m (Charvet *et al.* 2012). Our results seem to be more closely similar to those also found in L. Baikal, where also depth, and consequently nutrients, light, temperature and DO had a strong effect on the protistan assemblage, large size and longevity allowing *in situ* evolution (Annenkova, Rodriguez-Giner and Logares 2020; David *et al.* 2021).

## Phylogenetic exploration of 'novel' diversity in Lake Kivu

Lake Kivu hosted a great phylogenetic diversity with 280 OTUs classified as 'novel'. Chrysophyceae OTUs were related to clades formed mostly by heterotrophic organisms from freshwater and marine environments (del Campo and Massana 2011). Bicosoecida, a common group in freshwater, was also found in hypersaline systems (Debroas *et al.* 2017; Filker *et al.* 2017) (Fig. 6). Most of them were related to freshwater reference sequences and distributed at different depths. OTU_69 and OTU_1968 were distantly related to *Cafeteria*, a common marine genus; however, they present a low similarity to their reference sequence (<90%). Moreover, 'novel' Ciliophora OTUs were many and diverse at great depths (I-50, I-70 and G-70) (Fig. 8). Ciliophora is a ubiquitous, complex and highly diverse group, already observed to display high 'novelty' in other extreme environments (Reboul *et al.* 2018). 'Novel' OTUs closer to Prostomatea affiliated to sequences from lake sediment (KP059214) and anoxic waters (GU819524). 'Novel' OTUs related

to Oligohymenophorea were mostly obtained from anoxic deep marine samples and three 'novel' OTUs were classified as abundant (≥1%): Oligohymenophorea (OTU_452), Phyllopharyngea (OTU_394) and an environmental OTU (OTU_2540). OTU_452 presented a relatively high contribution in I-45 and OTU_2540 in I-70 and G-70 samples, highlighting their importance in these environments and how little we know about the diversity of microbial eukaryotes in anoxic freshwater systems.

### Marine–freshwater transitions

A large number of protist lineages, mainly related to marine environment, like MASTs, Perkinsidae, marine Ciliates, Chlorarachniophyta and MALVs among others, were found. With the advance of molecular techniques, MASTs are already being detected in different freshwater systems (Charvet *et al.* 2012; Simon *et al.* 2014; Arroyo *et al.* 2018). On the other hand, successful transitions from marine waters to freshwaters in Perkinsidae have been reported in detail (Bråte *et al.* 2010). This group also began to be reported from high-mountain oligotrophic lakes (Ortiz-Álvarez *et al.* 2018) and recently a large distribution across freshwaters was observed, suggesting potential parasitic associations with phytoplankton (Jobard *et al.* 2020) and emphasizing the need for further studies of these parasitic organisms. Most 'novel' ciliates were abundant in L. Kivu deep samples and they mostly affiliated to marine sequences. Those anoxic deep waters were characterized by high nutrient concentrations and conductivity (Pasche *et al.* 2009; Roland *et al.* 2017), harboring diverse and complex bacterial and archaeal assemblages (Llirós *et al.* 2010; İnceoğlu *et al.* 2015). The high density of ciliates in deep waters could be explained by the fact that many of them could prey on prokaryotes (Fenchel and Finlay 2010; Charvet *et al.* 2012). Likewise, anaerobic ciliates have been able to adapt to anoxic conditions by harboring methanogenic symbionts (Oikonomou, Pachiadaki and Stoeck 2014). Other exclusive marine algae found are chlorarachniophytes, widely distributed in tropical and temperate waters (Ishida, Yabuki and Ota 2007; Gile *et al.* 2010) and, until now, not registered in other freshwater habitats. Members of MALVs were registered in L. Baikal, and in association with MASTs and other Cercozoa, they are likely to participate in complex interactions (Genitsaris *et al.* 2020). Hence, our results demonstrate a complex eukaryotic diversity, in which several lineages have adapted to completely different niches and have succeeded to overcome the salinity boundary (Simon *et al.* 2014; Annenkova *et al* 2015; Arroyo *et al.* 2018).

## ACKNOWLEDGEMENTS

We thank Pascal Isumbisho (Institut Supérieur Pédagogique, Bukavu, Democratic Republic of Congo) and William Okello (National Fisheries Resources Research Institute, Jinja, Uganda), and their respective teams for logistic support during the cruises for help during fieldwork. We are grateful to Marc Llirós for help with laboratory and fieldwork and the rest of the team of the University of Namur, Belgium. In addition, we are thankful to Caterina Rodiguez-Giner for their help in molecular analyses and the rest of the team at the Institut de Ciències del Mar, CSIC, Barcelona. Bioinformatics analyses were performed at the MARBITS platform (http://marbits.icm.csic.es), ICM-CSIC. We are also thankful to the two anonymous reviewers whose thorough work greatly improved the early versions of this manuscript.

## SUPPLEMENTARY DATA

Supplementary data are available at FEMSEC online.

## FUNDING

This study was supported by the Spanish Ministry of Science and Technology (AGLOM project CGL2010-11556-E), the Belgian Federal Science Policy Office—BELSPO (Project EAGLES), the Fonds de la Recherche Scientifique—FNRS (Project MICKI—Microbial diversity and processes in Lake Kivu), 'Damas' mission, by FNRS and Fundaçao de Amparo a` Pesquisa do Estado de Sao Paulo (FAPESP process 2014/14139-3). PF was supported by the fellow 'Bourse de séjour scientifique', by FNRS and Programa de Financiamiento Parcial de Estadías Breves en el Exterior para Becarios Postdoctorales by Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET—Argentina). HS was supported by Conselho Nacional de Desenvolvimento e Pesquisa Tecnolo´gica (CNPq productivity grant 309514/2017-7).

***Conflict of interest.*** None declared.

## REFERENCES

Annenkova NV, Hansen G, Moestrup Ø *et al.* Recent radiation in a marine and freshwater dinoflagellate species flock. *ISME J* 2015;**9**:1821–34.

Annenkova NV, Rodriguez-Giner C, Logares R. Tracing the origin of planktonic protists in an ancient lake. *Microorganisms* 2020;**8**:543.

APHA. Inorganic non metallic constituents. In: Greenberg AE Clesceri LS Eaton AD (eds.). *Standard Methods for the Examination of Water and Wastewater*. New York, USA: American Public Health Association, 1992, 129.

Arroyo AS, López-Escardó D, Kim E *et al.* Novel diversity of deeply branching, Holomycota and unicellular Holozoans revealed by metabarcoding in Middle Paraná River, Argentina. *Front Ecol Evol* 2018;**6**:99.

Boenigk J, Arndt H. Bacterivory by heterotrophic flagellates: community structure and feeding strategies. *Antonie Van Leeuwenhoek* 2002;**81**:465–80.

Boenigk J, Wodniok S, Bock C *et al.* Geographic distance and mountain ranges structure freshwater protist communities on a European scale. *Metab Metagen* 2018;**2**:e21519.

Boenigk J, Wodniok S, Glücksman E. *Biodiversity and Earth History*. Berlin: Springer, 2015.

Bootsma HA, Hecky RE. Conservation of the African Great Lakes: a limnological perspective. *Conserv Biol* 1993;**7**:644–56.

Borges AV, Abril G, Delille B *et al.* Diffusive methane emissions to the atmosphere from Lake Kivu (Eastern Africa). *J Geophys Res* 2011;**116**:G03032.

Bråte J, Logares R, Berney C *et al.* Freshwater Perkinsea and marine-freshwater colonizations revealed by pyrosequencing and phylogeny of environmental rDNA. *ISME J* 2010;**4**:1144–53.

Brooks JL. Speciation in ancient lakes. *Q Rev Biol* 1950;**25**:131–76.

Burki F, Roger AJ, Brown MW *et al.* The new tree of eukaryotes. *Trends Ecol Evol* 2020;**35**:43–55.

Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. TrimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 2009;**25**:1972–3.

Caporaso JG, Kuczynski J, Stombaugh J *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010;**7**:335.

Carlson RE. A trophic state index for lakes. *Limnol Oceanogr* 1977;**22**:361–9.

Casamayor EO, Massana R, Benlloch S *et al.* Changes in archaeal, bacterial and eukaryal assemblages along a salinity gradient by comparison of genetic fingerprinting methods in a multipond solar saltern. *Environ Microbiol* 2002;**4**:338–48.

Catalan J, Camarero L, Felip M *et al.* High mountain lakes: extreme habitats and witnesses of environmental change. *Limnetica* 2006;**25**:551–84.

Charvet S, Vincent WF, Comeau AM *et al.* Pyrosequencing analysis of the protist communities in a High Arctic meromictic lake: DNA preservation and change. *Front Microbiol* 2012;**3**:422.

Chen M, Chen F, Yu Y *et al.* Genetic diversity of eukaryotic microorganisms in Lake Taihu, a large shallow subtropical lake in China. *Microb Ecol* 2008;**56**:572–83.

Cohen AS, Kaufman L, Ogutu-Ohwayo R. Anthropogenic threats, impacts and conservation strategies in the African Great Lakes: a review. In: Johnson TC, Odada E (eds). *The Limnology, Climatology and Paleoclimatology of the East African Lakes*. Toronto: Gordon and Breach, 1996, 575–624.

Corrêa LL, Oliveira MSB, Tavares-Dias M *et al.* Infections of Hypostomus spp. by Trypanosoma spp. and leeches: a study of hematology and record of these hirudineans as potential vectors of these hemoflagellates. *Rev Bras Parasitol Vet* 2016;**25**:299–305.

Cruaud P, Vigneron A, Fradette M-S *et al.* Annual protist community dynamics in a freshwater ecosystem undergoing contrasted climatic conditions: the Saint-Charles River (Canada). *Front Microbiol* 2019;**10**:2359.

David GM, Moreira D, Reboul G *et al.* Environmental drivers of plankton protist communities along latitudinal and vertical gradients in the oldest and deepest freshwater lake. *Environ Microbiol* 2021;**23**:1436–51.

de Vargas C, Audic S, Henry N *et al.* Eukaryotic plankton diversity in the sunlit ocean. *Science* 2015;**348**:1261605.

Debroas D, Domaizon I, Humbert J-F *et al.* Overview of freshwater microbial eukaryotes diversity: a first analysis of publicly available metabarcoding data. *FEMS Microbiol Ecol* 2017;**93**:fix023.

Debroas D, Hugoni M, Domaizon I. Evidence for an active rare biosphere within freshwater protists community. *Mol Ecol* 2015;**24**:1236–47.

del Campo J, Massana R. Emerging diversity within chrysophytes, choanoflagellates and bicosoecids based on molecular surveys. *Protist* 2011;**162**:435–48.

Descy JP, Darchambeau F, Schmid M. *Lake Kivu: Limnology and Biogeochemistry of a Tropical Great Lake*. Berlin: Springer, 2012.

Descy JP, Sarmento H. Microorganisms of the East African Great Lakes and their response to environmental changes. *Freshw Rev* 2008;**1**:59–73.

Domingues CD, Silva LHS, Rangel LM *et al.* Microbial food-web drivers in tropical reservoirs. *Microb Ecol* 2017;**73**:505–20.

Du P, Jiang ZB, Wang YM *et al.* Spatial heterogeneity of the planktonic protistan community in a semi-closed eutrophic bay, China. *J Plankton Res* 2019;1–17.

Dyal PL, Hope S, Roberts DM *et al.* Use of the PCR and fluorescent-probes to recover SSU ribosomal-RNA gene-sequences from single cells of the ciliate protozoan *Spathidium*. *Mol Ecol* 1995;**4**:499–503.

Edgar RC, Haas BJ, Clemente JC *et al.* UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 2011;**27**:2194–200.

Fenchel T, Finlay BJ. Free-living protozoa with endosymbiotic methanogens. In Hackstein JHP (ed). *(Endo)symbiotic Methanogenic Archaea*, Berlin: Springer Heidelberg, 2010, 1–11.

Fermani P, Torremorell A, Lagomarsino L *et al.* Microbial abundance patterns along a transparency gradient suggest a weak coupling between heterotrophic bacteria and flagellates in eutrophic shallow Pampean lakes. *Hydrobiologia* 2015;**752**:103–23.

Filker S, Forster D, Weinisch L *et al.* Transition boundaries for protistan species turnover in hypersaline waters of different biogeographic regions. *Environ Microbiol* 2017;**19**:3186–200.

Fryer G, Iles TD. *The Cichlid Fishes of the Great Lakes of Africa*. Edinburgh: Oliver and Boyd, 1972.

Galluzzi L, Penna A, Bertozzini E *et al.* Development of a real-time PCR assay for rapid detection and quantification of *Alexandrium minutum* (a dinoflagellate). *Appl Environ Microbiol* 2004;**70**:1199–206.

Genitsaris S, Stefanidou N, Moustaka-Gouni M *et al.* Variability and community composition of marine unicellular eukaryote assemblages in a Eutrophic Mediterranean Urban Coastal Area with marked plankton blooms and red tides. *Diversity* 2020;**12**:114.

Gile GH, Stern RF, James ER *et al.* DNA barcoding of chlorarachniophytes using nucleomorph ITS sequences. *J Phycol* 2010;**46**:743–50.

Gloor GB, Macklaim JM, Pawlowsky-Glahn V *et al.* Microbiome datasets are compositional: and this is not optional. *Front Microbiol* 2017;**8**:2224.

Guillou L, Bachar D, Audic S *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.

Haas BJ, Gevers D, Earl AM *et al.* Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res* 2011;**21**:494–504.

Hammer Ø. *PAST: Paleontological Statistics. Reference Manual*. Natural History Museum, University of Oslo, 2016.

Hecky RE, Bugenyi FWB. Hydrology and chemistry of the Great Lakes and water quality issues: problems and solutions. *Mitt internat Verein Limnol* 1992;**23**:45–54.

Hecky RE, Fee EJ, Kling HJ *et al.* Relationship between primary production and fish production in Lake Tanganyika. *Trans Am Fish Soc* 1981;**110**:336–45.

Hurlbert SH. The nonconcept of species diversity: a critique and alternative parameters. *Ecology* 1971;**52**:577–86.

İnceoğlu O, Llirós M, Crowe SA *et al.* Vertical distribution of functional potential and active microbial communities in meromictic Lake Kivu. *Microb Ecol* 2015;**70**:596–611.

Ishida KI, Yabuki A, Ota S. The chlorarachniophytes: evolution and classification. *Syst Ass Vol Ser* 2007;**75**:171.

Jobard M, Wawrzyniak I, Bronner G *et al.* Freshwater Perkinsea: diversity, ecology and genomic information. *J Plankton Res* 2020;**42**:3–17.

Kalyaanamoorthy S, Minh BQ, Wong TK *et al.* ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* 2017;**14**:587–9.

Kammerlander B, Breiner HW, Filker S *et al.* High diversity of protistan plankton communities in remote high mountain lakes in the European Alps and the Himalaya mountains. *FEMS Microbiol Ecol* 2015;**91**:fiv010.

Katoh K, Misawa K, Kuma L *et al.* MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 2002;**30**:3059–66.

Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 2013;**30**:772–80.

Khomich M, Kauserud H, Logares R *et al.* Planktonic protistan communities in lakes along a large-scale environmental gradient. *FEMS Microbiol Ecol* 2017;**93**:fiw23.

Kilham P, Kilham SS. Endless summer: internal loading processes dominate nutrient cycling in tropical lakes. *Freshw Biol* 1990;**23**:379–89.

Lê S, Josse J, Husson F. FactoMineR: an R package for multivariate analysis. *J Stat Softw* 2008;**25**:1.

Lefèvre E, Bardot C, Noël C *et al.* Unveiling fungal zooflagellates as members of freshwater picoeukaryotes: evidence from a molecular diversity study in a deep meromictic lake. *Environ Microbiol* 2007;**9**:61–71.

Lefranc M, Thénot A, Lepère C *et al.* Genetic diversity of small eukaryotes in lakes differing by their trophic status. *Appl Environ Microbiol* 2005;**71**:5935–42.

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**:2940–9.

Lepère C, Domaizon I, Hugoni M *et al.* Diversity and dynamics of active microbial eukaryotes in the anoxic zone of a freshwater meromictic lake (Pavin, France). *Front Microbiol* 2016;**7**:130.

Lepère C, Domaizon I, Humbert JF *et al.* Diversity, spatial distribution and activity of fungi in freshwater ecosystems. *PeerJ* 2019;**7**:e6247.

Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* 2019;**47**:W256–9.

Liu J, Li X, Chai B. Effects of seasonal freezing-thawing on the protistan communities in a mountain lake. 2021. DOI: https://doi.org/10.21203/rs.3.rs-360259/v1.

Llirós M, Descy JP, Libert X *et al.* Microbial ecology of Lake Kivu. In: Descy J-P, Darchambeau F, Schmid M (eds.). *Lake Kivu: Limnology and Biogeochemistry of a tropical great lake. Aquatic Ecology Series*. The Netherlands: Springer, 2012, 85–105.

Llirós M, Gich F, Plasencia A *et al.* Vertical distribution of ammonia-oxidizing Crenarchaeota and methanogens in the epipelagic waters of Lake Kivu. *Appl Environ Microbiol* 2010;**76**:6853–63.

Logares R, Audic S, Bass D *et al.* Patterns of rare and abundant marine microbial eukaryotes. *Curr Biol* 2014;**24**:813–21.

Logares R, Audic S, Santini S *et al.* Diversity patterns and activity of uncultured marine heterotrophic flagellates unveiled with pyrosequencing. *ISME J* 2012;**6**:1823–33.

Loiselle S, Cózar A, Adgo E *et al.* Decadal trends and common dynamics of the bio-optical and thermal characteristics of the African Great Lakes. *PLoS One* 2014;**9**:e93656.

Mahé F, de Vargas C, Bass D *et al.* Parasites dominate hyperdiverse soil protist communities in Neotropical rainforests. *Nat Ecol Evol* 2017;**1**:0091.

Massana R, Castresana J, Balagué V *et al.* Phylogenetic and ecological analysis of novel marine Stramenopiles. *Appl Environ Microbiol* 2004;**70**:3528–34.

Massana R, Gobet A, Audic S *et al.* Marine protist diversity in European coastal waters and sediments as revealed by high-throughput sequencing. *Environ Microbiol* 2015;**17**:4035–49.

Massana R, Pernice M, Bunge JA *et al.* Sequence diversity and novelty of natural assemblages of picoeukaryotes from the Indian Ocean. *ISME J* 2011;**5**:184–95.

Metz S, Lopes dos Santos A, Castro-Berman M *et al.* Diversity of photosynthetic picoeukaryotes in eutrophic shallow lakes as assessed by combining flow cytometry cell-sorting and high throughput sequencing. *FEMS Microbiol Ecol* 2019;**95**:fiz038.

Morana C, Borges AV, Roland FA *et al.* Methanotrophy within the water column of a large meromictic tropical lake (Lake Kivu, East Africa). *Biogeosciences* 2015;**12**:2077–88.

Morana C, Roland FA, Crowe SA *et al.* Chemoautotrophy and anoxygenic photosynthesis within the water column of a large meromictic tropical lake (Lake Kivu, East Africa). *Limnol Oceanogr* 2016;**61**:1424–37.

Morana C, Sarmento H, Descy JP *et al.* Production of dissolved organic matter by phytoplankton and its uptake by heterotrophic prokaryotes in large tropical lakes. *Limnol Oceanogr* 2014;**59**:1364–75.

Mukherjee I, Hodoki Y, Okazaki Y *et al.* Widespread dominance of kinetoplastids and unexpected presence of diplonemids in deep freshwater lakes. *Front Microbiol* 2019;**10**:2375.

Nguyen LT, Schmidt HA, Von Haeseler A *et al.* IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 2015;**32**:268–74.

Obiol A, Giner CR, Sánchez P *et al.* A metagenomic assessment of microbial eukaryotic diversity in the global ocean. *Mol Ecol Resour* 2020;**20**:718–31.

Odada EO, Olago DO, Bugenyi F *et al.* Environmental assessment of the East African Rift Valley lakes. *Aquat Sci* 2003;**65**:254–71.

Ogutu-Ohwayo R, Hecky RE, Cohen AS *et al.* Human impacts on the African Great Lakes. *Environ Biol Fishes* 1997;**50**:117–31.

Oikonomou A, Pachiadaki M, Stoeck T. Protistan grazing in a meromictic freshwater lake with anoxic bottom water. *FEMS Microbiol Ecol* 2014;**87**:691–703.

Okamoto N, Chantangsi C, Horák A *et al.* Molecular phylogeny and description of the novel katablepharid *Roombia truncata* gen. et sp. nov., and establishment of the Hacrobia taxon nov. *PLoS One* 2009;**4**:e7080.

Oksanen J, Blanchet FG, Friendly M *et al.* Vegan: community ecology package. *R Package Version 2.5-6*. 2019. https://github.com/vegandevs/vegan/.

Olapade OJ, Omitoyin BO. Anthropogenic pollution impact on physico-chemical characteristics of lake Kivu, Rwanda. *Afr J Food Agric Nutr Dev* 2012;**12**:6517–36.

Ortiz-Álvarez R, Triadó-Margarit X, Camarero L *et al.* High planktonic diversity in mountain lakes contains similar contributions of autotrophic, heterotrophic and parasitic eukaryotic life forms. *Sci Rep* 2018;**8**:4457.

Otu MK, Ramlal P, Wilkinson P *et al.* Paleolimnological evidence of the effects of recent cultural eutrophication during the last 200 years in Lake Malawi, East Africa. *J Great Lakes Res* 2011;**37**:61–74.

Palma-Silva C, Albertoni EF, Esteves FA. Charophytes as nutrient and energy reservoir in a tropical coastal lagoon impacted by humans (RJ, Brazil). *Braz J Biol* 2004;**64**:479–87.

Pasche N, Dinkel C, Müller B *et al.* Physical and bio-geochemical limits to internal nutrient loading of meromictic Lake Kivu. *Limnol Oceanogr* 2009;**54**:1863–73.

Pernice MC, Giner CR, Logares R *et al.* Large variability of bathypelagic microbial eukaryotic communities across the world's oceans. *ISME J* 2016;**10**:945–58.

Pirlot S, Vanderheyden J, Descy JP *et al.* Abundance and biomass of heterotrophic micro-organisms in Lake Tanganyika. *Freshw Biol* 2005;**50**:1219–32.

Pomeroy LR. The ocean's food web, a changing paradigm. *Bioscience* 1974;**24**:499–504.

Porter KG, Paerl H, Hodson R *et al.* Microbial interactions in lake food webs. In: Carpenter SR (ed). *Complex Interactions in Lake Communities*. New York: Springer-Verlag, 1988.

Quast C, Pruesse E, Yilmaz P *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 2013;**41**:D590–6.

Reboul G, Moreira D, Bertolino P *et al.* Microbial eukaryotes in the suboxic chemosynthetic ecosystem of Movile Cave, Romania. *Environ Microbiol Rep* 2018;**11**:464–73.

Roland FAE, Darchambeau F, Borges AV *et al*. Denitrification, anaerobic ammoniumoxidation, and dissimilatory nitrate reduction to ammonium in an east African Great Lake (Lake Kivu). *Limnol Oceanogr* 2017;**63**:687–701.

Salani FS, Arndt H, Hausmann K *et al*. Analysis of the community structure of abyssal kinetoplastids revealed similar communities at larger spatial scales. *ISME J* 2012;**6**:713–23.

Sarmento H, Darchambeau F, Descy JP. Phytoplankton of Lake Kivu. In: Descy JP, Darchambeau F, Schmid M (eds.). *Lake Kivu: Limnology and Biogeochemistry of a Tropical Great Lake*. Dordrecht: Springer, 2012, 67–83.

Sarmento H, Isumbisho M, Descy JP. Phytoplankton ecology of Lake Kivu (Eastern Africa). *J Plankton Res* 2006;**28**:815–29.

Sarmento H, Unrein F, Isumbisho M *et al*. Abundance and distribution of picoplankton in tropical, oligotrophic Lake Kivu, eastern Africa. *Freshw Biol* 2008;**53**:756–71.

Sarmento H. New paradigms in tropical limnology: the importance of the microbial food web. *Hydrobiologia* 2012;**686**: 1–14.

Schloss PD, Girard R, Martin T *et al*. The status of the microbial census: an update. *mBio* 2016;**7**:e00201–16.

Schloss PD. Evaluating different approaches that test whether microbial communities have the same structure. *ISME J* 2008;**2**:265–75.

Schmid M, Halbwachs M, Wehrli B *et al*. Weak mixing in Lake Kivu: new insights indicate increasing risk of uncontrolled gas eruption. *Geochem Geophys Geosyst* 2005;**6**:Q07009.

Schmid M, Tietze K, Halbwachs M *et al*. How hazardous is the gas accumulation in Lake Kivu? Arguments for a risk assessment in light of the Nyiragongo volcano eruption of 2002. *Acta Vulcanol* 2002;**14**:115–22.

Schmid M, Wüest A. Stratification, mixing and transport processes in Lake Kivu. In: Descy JP, Darchambeau F, Schmid M (eds). *Lake Kivu: Limnology and Biogeochemistry of a Tropical Great Lake*. Dordrecht: Springer, 2012, 67–83.

Schubert H, Blindow I. *Charophytes of the Baltic Sea*. Königstein: Koeltz Scientific, 2004.

Seehausen O, Witte F, Katunzi EE *et al*. Patterns of the remnant cichlid fauna in southern Lake Victoria. *Conserv Biol* 1997;**11**:890–904.

Seehausen O. African cichlid fish: a model system in adaptive radiation research. *Proc Biol Sci* 2006;**273**:1987–98.

Sherr EB, Sherr BF. Role of microbes in pelagic food webs: a revised concept. *Limnol Oceanogr* 1988;**33**:1225–7.

Simon M, Jardillier L, Deschamps P *et al*. Complex communities of small protists and unexpected occurrence of typical marine lineages in shallow freshwater systems. *Environ Microbiol* 2014;**17**:3610–27.

Simon M, López-García P, Deschamps P *et al*. Marked seasonality and high spatial variability of protist communities in shallow freshwater systems. *ISME J* 2015;**9**:1941–53.

Simon M, López-García P, Deschamps P *et al*. Resilience of freshwater communities of small microbial eukaryotes undergoing severe drought events. *Front Microbiol* 2016;**7**:812.

Singer D, Seppey CVW, Lentendu G *et al*. Protist taxonomic and functional diversity in soil, freshwater and marine ecosystems. *Environ Int* 2021;**146**:106262.

Spigel RH, Coulter GW. Comparison of hydrology and physical limnology of the East African Great Lakes: Tanganyika, Malawi, Victoria, Kivu and Turkana (with reference to some North American Great Lakes). In: Johnson TC, Odada E (eds.). *The Limnology, Climatology and Paleoclimatology of East African Lakes*. Toronto: Gordon and Breach, 1996, 103–39.

Stenuite S, Tarbe AL, Sarmento H *et al*. Photosynthetic picoplankton in Lake Tanganyika: biomass distribution patterns with depth, season and basin. *J Plankton Res* 2009;**31**:1531–44.

Stoeck T, Bass D, Nebel M *et al*. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol Ecol* 2010;**19**:21–31.

Tarbe AL, Stenuite S, Balagué V *et al*. Molecular characterization of the small-eukaryote community in a tropical Great Lake (Lake Tanganyika, East Africa). *Aquat Microb Ecol* 2011b;**62**:177–90.

Tarbe AL, Unrein F, Stenuite S *et al*. Protist herbivory: a key pathway in the pelagic food web of Lake Tanganyika. *Microb Ecol* 2011a;**62**:314–23.

Tarbe AL. Les protistes: acteurs-clés du réseau trophique pélagique au Lac Tanganyika. Ph.D. Thesis, University of Namur. 2010.

Taylor JW, Turner E, Townsend JP *et al*. Eukaryotic microbes, species recognition and the geographic limits of species: examples from the kingdom Fungi. *Philos Trans R Soc Lond B Biol Sci* 2006;**361**:1947–63.

ter Braak CJF, Smilauer P. *CANOCO Reference Manual and User's Guide to Canoco for Windows—Software for Canonical Community Ordination (version 4)*. Ithaca: Microcomputer Power, 1998.

ter Braak CJF. Ordination. In: Jongman RHG, ter Braak CJF, van Tongeren OFR (eds). *Data Analysis in Community and Landscape Ecology*. Wageningen: Cambridge University Press, 1987, 91–173.

Triadó-Margarit X, Casamayor EO. Genetic diversity of planktonic eukaryotes in high mountain lakes (Central Pyrenees, Spain). *Environ Microbiol* 2012;**14**:2445–56.

Vadeboncoeur Y, Mcintyre PB, Vander Zander MJ. Borders of biodiversity: life at the edge of the world's large lakes. *Bioscience* 2011;**61**:7.

Verleyen E, Vyverman W, Sterken M *et al*. The importance of dispersal related and local factors in shaping the taxonomic structure of diatom metacommunities. *Oikos* 2009;**118**: 1239–49.

Weiss RF. Determinations of carbon dioxide and methane by dual catalyst flame ionization chromatography and nitrous oxide by electron capture chromatography. *J Chromatogr Sci* 1981;**19**:611–6.

Weisse T, Anderson R, Arndt H *et al*. Functional ecology of aquatic phagotrophic protists - Concepts, limitations, and perspectives. *Eur J Protistol* 2016;**55**:50–74.

Weisse T, Müller H, Pinto-Coelho RM *et al*. Response of the microbial loop to the phytoplankton spring bloom in a large prealpine lake. *Limnol Oceanogr* 1990;**35**:781–94.

Worden AZ, Follows MJ, Giovannoni SJ *et al*. Rethinking the marine carbon cycle: factoring in the multifarious lifestyles of microbes. *Science* 2015;**347**:1257594.

Wu QL, Chatzinotas A, Wang J *et al*. Genetic diversity of eukaryotic plankton assemblages in Eastern Tibetan Lakes differing by their salinity and altitude. *Microb Ecol* 2009;**58**:569–81.

Xiong W, Jousset A, Li R *et al*. A global overview of the trophic structure within microbiomes across ecosystems. *Environ Int* 2021;**151**:106438.

Yubuki N, Leander BS. Diversity and evolutionary history of the Symbiontida (Euglenozoa). *Front Ecol Evol* 2018;**6**:100.

Zhao F, Filker S, Xu K *et al*. Patterns and drivers of vertical distribution of the ciliate community from the surface to the abyssopelagic zone in the Western Pacific Ocean. *Front Microbiol* 2017;**8**:2559.