Global freshwater distribution of *Telonemia* protists

Roudaina Boukheloua[1,2], Indranil Mukherjee[1], Hongjae Park[1], Karel Šimek[1,2], Vojtěch Kasalický[1], Maxon Ngochera[3], Hans-Peter Grossart[4,5], Antonio Picazo-Mozo[6], Antonio Camacho[6], Pedro J. Cabello-Yeves[6,7], Francisco Rodriguez-Valera[8], Cristiana Callieri[9], Adrian-Stefan Andrei[10], Jakob Pernthaler[10], Thomas Posch[10], Albin Alfreider[11], Ruben Sommaruga[11], Martin W. Hahn[12], Bettina Sonntag[12], Purificación López-García[13], David Moreira[13], Ludwig Jardillier[13], Cécile Lepère[14], Corinne Biderre-Petit[14], Anna Bednarska[15], Mirosław Ślusarczyk[15,16], Viktor R. Tóth[17], Horia L. Banciu[18], Konstantinos Kormas[19], Sandi Orlić[20,21], Danijela Šantić[22], Gerard Muyzer[23], Daniel P.R. Herlemann[24,25], Helen Tammert[25], Stefan Bertilsson[26], Silke Langenheder[27], Thomas Zechmeister[28], Nico Salmaso[29,30], Nicola Storelli[31,32], Camilla Capelli[33], Fabio Lepori[33,34], Vojtěch Lanta[35], Helena Henriques Vieira[1], Fran Kostanjšek[1], Kateřina Kabeláčová[1], Maria-Cecilia Chiriac[1], Markus Haber[1], Tanja Shabarova[1], Clafy Fernandes[1,2], Pavel Rychtecký[1], Petr Znachor[1], Tiberiu Szőke-Nagy[1], Paul Layoun[1,2], Hon Lun Wong[1], Vinicius Silva Kavagutti[1,2], Paul-Adrian Bulzu[1], Michaela M. Salcher[1], Kasia Piwosz[36], Rohit Ghai[1,*]

[1]Department of Aquatic Microbial Ecology, Institute of Hydrobiology, Biology Centre of the Czech Academy of Sciences, 37005, České Budějovice, Czech Republic
[2]Faculty of Science, University of South Bohemia, 37005, České Budějovice, Czech Republic
[3]Department of Fisheries, Ministry of Natural Resources and Climate Change, 593 Lilongwe, Malawi
[4]Department of Plankton and Microbial Ecology, Leibniz Institute for Freshwater Ecology and Inland Fisheries, (IGB), Alte Fischerhuette 2, D-16775 Neuglobsow, Germany
[5]Institute of Biochemistry and Biology, Potsdam University, Maulbeerallee 2, D-14469 Potsdam, Germany
[6]Cavanilles Institute of Biodiversity and Evolutionary Biology, University of Valencia, E-46980 Paterna, Valencia, Spain
[7]School of Life Sciences, University of Warwick, CV4 7AL Coventry, United Kingdom
[8]Evolutionary Genomics Group, Departamento de Producción Vegetal y Microbiología, Universidad Miguel, Hernández, 03550, San Juan de Alicante, Alicante, Spain
[9]Water Research Institute, National Research Council (IRSA-CNR), Molecular Ecology Group (MEG), Largo Tonolli 50, Verbania 28922, Italy
[10]Limnological Station, Department of Plant and Microbial Biology, University of Zurich, 8802, Kilchberg, Switzerland
[11]Lake and Glacier Ecology Research Group, Department of Ecology, University of Innsbruck, A-6020, Innsbruck, Austria
[12]Research Department for Limnology, Mondsee, University of Innsbruck, A-5310, Mondsee, Austria
[13]Unité d'Ecologie Systématique et Evolution, CNRS, Université Paris-Saclay, AgroParisTech, 91190 Gif-sur-Yvette, France
[14]Laboratoire Microorganismes: Génome et Environnement, CNRS, Université Clermont Auvergne, 63000 Clermont-Ferrand, France
[15]Department of Hydrobiology, Faculty of Biology, Institute of Ecology, Biological and Chemical Research Centre, University of Warsaw, Żwirki i Wigury 101, 02-089 Warsaw, Poland
[16]Hydrobiological Station, Faculty of Biology, University of Warsaw, Pilchy 5, 12-200 Pisz, Poland
[17]Aquatic Botany and Microbial Ecology Research Group, HUN-REN Balaton Limnological Research Institute, 8237 Tihany, Hungary
[18]Department of Molecular Biology and Biotechnology, Faculty of Biology and Geology, Babeș-Bolyai University, 5-7 Clinicilor Street, 400006 Cluj-Napoca, Romania
[19]Department of Ichthyology and Aquatic Environment, School of Agricultural Sciences, University of Thessaly, 38446 Volos, Greece
[20]Division of Materials Chemistry, Ruđer Bošković Institute, Bijenička Cesta 54, 10000, Zagreb, Croatia
[21]Center of Excellence for Science and Technology-Integration of Mediterranean Region, Zagreb, Croatia
[22]Laboratory of Marine Microbiology, Institute of Oceanography and Fisheries, Šetalište Ivana Meštrovića 63, 21000 Split, Croatia
[23]Department of Freshwater and Marine Ecology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam 1098 XH, The Netherlands
[24]Leibniz Institute for Baltic Sea Research Warnemünde (IOW), Seestrasse 15, D-18119 Rostock, Germany
[25]Centre for Limnology, Estonian University of Life Sciences, 6117 Vehendi, Tartu County, Estonia
[26]Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences, 750 07 Uppsala, Sweden
[27]Department of Ecology and Genetics/Limnology, Uppsala University, SE-75236 Uppsala, Sweden
[28]Biological Station Lake Neusiedl, Seevorgelände 1, 7142 Illmitz, Austria
[29]Research and Innovation Centre, Fondazione Edmund Mach, Via E. Mach, 1, 38098 S. Michele all'Adige, Italy
[30]NBFC, National Biodiversity Future Center, 90133 Palermo, Italy
[31]Institute of Microbiology, University of Applied Sciences and Arts of Southern Switzerland, Campus Mendrisio, Via Flora Ruchat-Roncati 15, CH-6850 Mendrisio, Switzerland
[32]Department of Botany and Plant Biology, Microbiology Unit, University of Geneva, Sciences III, CH-1211 Geneva, Switzerland
[33]Institute of Earth Sciences, University of Applied Sciences and Arts of Southern Switzerland, Campus Mendrisio, Via Flora Ruchat-Roncati 15, CH-6850 Mendrisio, Switzerland
[34]État de Vaud, Direction de l'environnement industriel, urbain et rural (DGE-DIREV), 1066 Epalinges, Switzerland
[35]Department of Functional Ecology, Institute of Botany of the Czech Academy of Sciences, 252 43 Průhonice, Czech Republic
[36]Department of Fisheries Oceanography and Marine Ecology, National Marine Fisheries Research Institute, 81-332 Gdynia, Poland

*Corresponding author: Rohit Ghai, Department of Aquatic Microbial Ecology, Institute of Hydrobiology, Biology Centre of the Czech Academy of Sciences, Na Sádkách 702/7, 37005 České Budějovice, Czech Republic. Email: ghai.rohit@gmail.com
Members of The PELAGICS Consortium are listed after the Data availability information.

Abstract

Telonemia are one of the oldest identified marine protists that for most part of their history have been recognized as a distinct *incertae sedis* lineage. Today, their evolutionary proximity to the SAR supergroup (Stramenopiles, Alveolates, and Rhizaria) is firmly established. However, their ecological distribution and importance as a natural predatory flagellate, especially in freshwater food webs, still remain unclear. To unravel the distribution and diversity of the phylum *Telonemia* in freshwater habitats, we examined over a thousand freshwater metagenomes from all over the world. In addition, to directly quantify absolute abundances, we analyzed 407 samples from 97 lakes and reservoirs using Catalyzed Reporter Deposition-Fluorescence *in situ* Hybridization (CARD-FISH). We recovered *Telonemia* 18S rRNA gene sequences from hundreds of metagenomic samples from a wide variety of habitats, indicating a global distribution of this phylum. However, even after this extensive sampling, our phylogenetic analysis did not reveal any new major clades, suggesting current molecular surveys are near to capturing the full diversity within this group. We observed excellent concordance between CARD-FISH analyses and estimates of abundances from metagenomes. Both approaches suggest that *Telonemia* are largely absent from shallow lakes and prefer to inhabit the colder hypolimnion of lakes and reservoirs in the Northern Hemisphere, where they frequently bloom, reaching 10%–20% of the total heterotrophic flagellate population, making them important predatory flagellates in the freshwater food web.

Keywords: freshwater lakes, microbial food webs, predatory flagellate, Telonemia, CARD-FISH, metagenomics

Introduction

Freshwaters are extremely diverse ecosystems, with a wide variety of trophic states along with substantial dynamics within their complex microbial food webs [1–9]. Bacterivorous protists are critical components of these food webs, estimated to predate upon one-fourth of free-living bacteria every day [10, 11]. It is assumed that the majority of such heterotrophic protists are generally <5 μm (HNF: heterotrophic nanoflagellates), such as the uncultured CRY1 lineage of cryptophytes that is one of the most widespread bacterivore and abundant lineages of HNF in freshwaters [12, 13]. Recent studies have suggested that middle-sized (5–20 μm) HNF are omnivores, feeding not only on bacteria, but also predating upon algae and other microbial eukaryotes [11, 14]. Predatory protists are capable of hunting or immobilizing their prey and ingesting cells of relatively large sizes [15]. Laboratory experiments on microbial food web manipulations with predatory flagellates revealed doubling times comparable to the bacterivorous HNF (i.e. hours to days), and appear ahead of the ciliates in the energy transfer [11]. Some of the known predatory lineages of flagellates are Diplonemea [16, 17], Cercozoa [13, 18], Katablepharida [19], MAST-6 lineage [20], and genus *Telonema* [21]. The predatory role of *Telonema* is mainly described in marine and brackish waters [21] and even though it has been frequently observed in freshwaters, its diversity, distribution, abundance, and ecological role in freshwater microbial communities remain less understood.

More than a century ago, *Telonema* was first described from the marine habitat as a relatively small (6–8 μm long), colorless, elliptical, rigid-bodied flagellate without a contractile vacuole and no close relationship with other known flagellates [22]. A few decades later, another report of *Telonema subtilis* appeared, where it was obtained in culture from brackish waters [23]. Based upon its morphological similarity to the then *Cyathomonas* (now *Goniomonas,* a colourless Cryptophyte), it was decided to place *Telonema* within the family Cyathomonadidae. Subsequently, *T. subtilis* was observed in geographically dispersed marine locations (Arctic, Mediterranean, Japanese coastal waters, etc.) [24, 25] at a wide temperature range (−1°C to 26°C) in both summer and autumn seasons. In general, *Telonema* was found ubiquitously but at low abundances, yet at times accounted for up to 10%–30% of total heterotrophic flagellates. A larger *Telonema* species (diameter 10–20 μm) was observed and provisionally described as *Telonema antarcticum* [26] and found to bloom in summer in an annual study at a bay in Greenland (ca. 200 cells ml^{-1}) [27]. *T. antarcticum* was cultured by providing *Rhodomonas* as a food source and its ultrastructure was described in detail [21]. This work [21], along with

others [28, 29], described the first molecular phylogenies using 18S rRNA genes and consistently concluded that the sequences appeared quite distinct from any other eukaryotic group, highlighting the still unresolved placement of *Telonema* spp. Subsequent ultrastructural analyses combined with molecular phylogenies of 18S rRNA, Hsp90, α-, and β-tubulin gene sequences suggested that *Telonemia* represents a deep branching group, placed within its own phylum, i.e. *Telonemia* [30]. The affinity of *Telonemia* derived 18S rRNA gene sequences to Cryptophytes (also speculated before based upon morphological evidence) and Haptophytes was noted, but not considered conclusive. The availability of additional *Telonemia* 18S rRNA gene sequences confirmed previous observations that *Telonemia* represents a widespread phylum and can be grouped into two clades: Group1 with *T. subtilis* and Group2 with *T. antarcticum* [31].

The first indication of freshwater *Telonemia* representatives stems from a microscopic examination of protist samples from Sombre Lake in Antarctica [25] and was later confirmed by sequencing of 18S rRNA gene clone libraries from Lake Pavin [32]. However, molecular phylogenies showed strong support for *Telonemia* being related to Stramenopiles, Alveolates, and Cercozoans, and not to Cryptophytes as was suggested before. Another multigene phylogeny (actin, α-tubulin, β-tubulin, cytosolic HSP70, BIP HSP70, and HSP90) also suggested that telonemids should be grouped together with the SAR supergroup (Stramenopiles, Alveolates, and Rhizaria) [33] rather than with Cryptophytes and Haptophytes. However, a larger phylogenetic analysis with more than a hundred genes recapitulated the *Telonemia* and Cryptophyte grouping [34], and this was retained in a later work combining ultrastructural analyses and multigene phylogenies [35]. These incongruencies were finally resolved with a robust multigene phylogeny, placing *Telonemia* as a sister group to the SAR supergroup, forming the TSAR assemblage (*Telonemia* + SAR) [36].

Multiple additional environmental surveys using amplicon sequencing or clone libraries have repeatedly detected *Telonemia* in a wide variety of habitats, ranging from marine [37–40] to brackish [41] and freshwaters [42, 43], though usually at low abundances. Recently, the dynamics during a spring phytoplankton bloom was reported with the use of a CARD-FISH probe specific for *Telonemia* [44]. A more focused study on both marine and freshwater *Telonemia* reiterated the presence of multiple clades of freshwater *Telonemia*, within the already defined groups Telo-1 and Telo-2, which suggests the possibility of various marine–freshwater transitions within this group [43]. One study also used network analysis of 18S rRNA gene amplicons to

show *Telonemia* associated with an unknown ciliate suggestive of predation of *Telonemia* itself [45]. Even more recently, multiple isolates of *Telonemia* have been obtained and new genera have been defined, e.g. *Lateronema*, *Arpakorses*, mostly from marine habitats, but also the first freshwater species: *T. rivulare* has been described [46]. Additionally, a limited number of studies have examined protist communities in rivers using 18S rRNA gene amplicon sequencing in the northern hemisphere including the Saint-Charles, Great Whale, Nelson, and Churchill rivers in Canada [47, 48], the Vistula river in Poland [41], and the Yangtze river in China [49]. These were not focused upon *Telonemia* per se, but have reported the presence of *Telonemia* OTUs, particularly in colder seasons or in brackish regimes.

A systematic examination of the prevalence of *Telonemia* in freshwaters (particularly in lentic habitats), has been missing until now. Moreover, it is unclear if novel, yet undescribed major clades of *Telonemia* thrive in freshwaters. In this work, we used catalyzed reporter deposition fluorescence *in situ* hybridization (CARD-FISH) and a specific probe to directly visualize and quantify *Telonemia* in 97 lakes across Europe, Africa, South America, Australia, and Japan. In addition, we examined its seasonal distribution at four different freshwater sites. Furthermore, we recovered almost 250 18S rRNA gene sequences from >1000 freshwater metagenomes, greatly expanding our knowledge on the ecology of *Telonemia* in freshwaters. Our analyses suggest that *Telonemia* diversity is restricted to the already described main clades (Telo-1 and Telo-2) with a worldwide distribution in freshwater lakes, typically in the cold hypolimnion, and they are largely absent from shallow or hypertrophic water bodies. Based on their ubiquitous presence and occasional peaks of very high abundances in this rather niche, we suggest that *Telonemia* might represent one of the major predatory flagellates of the deep microbial food web in some freshwater lakes.

Materials and Methods
Study sites and sampling
Samples for CARD-FISH were collected from 97 freshwater and brackish water habitats covering a broad diversity of trophic states (ultraoligotrophic, oligotrophic, mesotrophic, eutrophic, and dystrophic), continental biomes (e.g. arctic, alpine, continental, Mediterranean, and boreal) at elevations of up to 1921 m asl (Lake Cadagno), and depths of 2–300 m. The sampled habitats spanned across a wide geographical distribution covering five continents (Europe, Asia, Oceania, Africa, and South America). However, we must point out that tropical lakes are underrepresented in our sample collection. For each lake, water samples were collected from the epilimnetic and hypolimnetic layers (except for shallow lakes where samples were collected only from the surface). Several timelines were collected from specific lakes. Four hypertrophic ponds and three dimictic reservoirs of different trophic status in the Czech Republic were studied monthly for six and nine months, respectively. Monthly samples from a monomictic and oligo-mesotrophic lake (Lake Biwa, Japan) were collected for a whole year. Samples were also collected from a temporal high-resolution spring campaign (three times a week) in Řimov reservoir, Czech Republic. In total, we examined 407 CARD-FISH samples from these sites. A complete list of all samples used in this work, along with all physicochemical parameters measured for each sample is provided in Supplementary Table S1.

Catalyzed reporter deposition-fluorescence *in situ* hybridization
Samples were fixed with formaldehyde (2% final concentration) for up to 24 h at 4°C. About, 30 ml of epilimnetic and 90 ml of hypolimnetic water were filtered on 0.8-μm polycarbonate filters (47 mm), and stored at −20°C for further processing. We used the oligonucleotide probe Telo-1250 (5′ CAGYCAAGGTGGACAAC-TYGTT 3′) targeting all *Telonemia* [44]. CARD-FISH was performed following the protocol described elsewhere [50] with fluorescein-labeled tyramides. CARD-FISH preparations were analyzed using an epifluorescence microscope (Olympus BX53, Japan) at 1000× magnification. Microscopic images were taken using Zeiss Imager Z2, Carl Zeiss, Oberkochen, DE equipped with a Colibri LED system.

Preprocessing and assembly of publicly available metagenomic datasets
Adaptor sequences and low-quality bases were removed from the (Illumina) sequences using the bbmap package (http://sourceforge.net/projects/bbmap/). Briefly, the reads were quality trimmed by bbduk.sh (using a Phred quality score of 18). Subsequently, bbduk.sh was used for trimming adapters, and also for the identification/removal of possible PhiX and p-Fosil2 contamination. *De novo* adapter identification with bbmerge.sh was also performed in order to ensure that the datasets meet the quality threshold necessary for assembly. Wherever necessary, metagenomic datasets were assembled independently with MEGAHIT (v1.1.5) (—-k-list 49 69 89 109 129 149) and default settings, otherwise previously available assemblies were used [4, 51–57]. A complete list of all metagenomes used in this work is provided in Supplementary Table S2.

Retrieval of Telonemia 18S rRNA gene sequences from assembled shotgun metagenomic datasets
Telonemia 18S rRNA gene sequences were gathered from previous publications [21, 30, 31, 43, 46]. The metagenomic assemblies were scanned for 18S rRNA gene sequences using ssu-align [58]. All recovered 18S rRNA gene sequences were submitted to the SILVA [59] online classification (https://www.arb-silva.de/aligner/) to identify bonafide *Telonemia* sequences. *Telonemia* 18S rRNA from metagenomic assemblies and published 18S rRNA gene sequences were clustered at 95% nucleotide identity and 100% coverage using cd-hit-est [60]. The representative sequences obtained (n = 21) were individually submitted to the IMG/ER service that allows retrieval of related sequences by megablast [61]. All retrieved sequences from IMG/ER were also submitted to the SILVA online classification to confirm that they belonged to *Telonemia*. Finally, only those *Telonemia* sequences with a minimum length of 400 bp were retained for further analysis. A complete table of all sequences (n = 771), their sources of origin, and the sequences is provided in Supplementary Table S3.

CARD-FISH probe specificity
The retrieved sequences were used to test the coverage of probe Telo-1250. After removing sequences that did not have the target region, the alignment of the remaining 476 sequences was imported into the software ARB [62]. As the probe has degenerated bases, all four non-degenerated sequences were tested using the probematch tool. The probe matched 90% of sequences when no mismatches were allowed, and 100% with 0 weighted mismatched option. The probe matched three nontarget sequences

retrieved from organisms with distinct morphology (*Symbiodinium*, Syndiniales Group, Cercozoa Novel Clade 2). It is unclear how prevalent Cercozoa Novel Clade 2 is in freshwaters, but it has been observed to be abundant in enrichments from brackish water [18] and the single sequence that matched the probe was retrieved from a coastal margin of the Columbia river [63], which is very different from other sequences in this lineage. Moreover, examination of 18S rRNA gene abundances of these groups in the metagenomes of the same samples, from which CARD-FISH was performed, revealed that most of the time, *Symbiodinium* was <1%. On the other hand, Cercozoa Clade 2 appeared to be quite widespread and abundant in these metagenomes; however, we detected no CARD-FISH signal from *Telonemia* at those sites with high cercozoan abundances (ca. 20%), strongly suggesting that overestimation of *Telonemia* using this probe is negligible (Supplementary Table S4). Moreover, *Telonemia has* larger, pear-shaped cells, a characteristic triangular nucleus, and can be distinguished from flagellates of Cercozoa Novel Clade 2.

18S rRNA phylogenetic tree construction

All *Telonemia* 18S rRNA gene sequences retrieved from literature or from metagenomes as described above (minimum sequence length 400 bp) were dereplicated at several nucleotide identity levels (95, 96, 97, 98, 99, and 100%). Alignments were created using mafft-linsi [64] and PASTA [65] at all these dereplication settings and maximum-likelihood phylogenetic trees were constructed using Iqtree2 v.2.2.2.6 31 (settings: -B 1000 –alrt 1000 -m MFP) [66, 67]. The best-fitting evolutionary models were chosen by ModelFinder [68] according to the Bayesian information criterion (BIC). Cryptophyte 18S rRNA gene sequences were used as outgroups for all phylogenetic trees. The delineation of clades Telo-1 and Telo-2 was based upon previous studies [43, 46]. All sequences, alignments, and phylogenetic trees are available at Zenodo (doi: 10.5281/zenodo.11237305). The resulting trees were visualized in iTOL (http://itol.embl.de).

Quantification of Telonemia 18S rRNA gene sequences in metagenomic datasets:

SILVA 138.1 eukaryotic 18S rRNA gene sequences (nr99) were downloaded locally. All *Telonemia* sequences gathered from literature and from locally assembled metagenomes or public servers (IMG/ER) were dereplicated at 99% identity using cd-hit-est [60] and added to the local SILVA 138.1 (nr99) database.

18S rRNA gene sequences were identified in the short-read metagenomes using ssu-align [58]. These short-read eukaryotic metagenomic 18S rRNA sequences were compared using MMseqs2 [69] to the *Telonemia* supplemented nr99 SILVA database (minimum %identity 80, minimum alignment length 100, e-value 1e-5) using a best-hit strategy. 18S rRNA gene sequences originating from organisms known for their extensive rRNA operon presence, such as Dinoflagellata and Ciliophora, were removed before further analysis. Additionally, sequences from multicellular organisms like Metazoa and Embryophyta, as well as those originating from nucleomorphs of Cryptophyceae, were excluded. The results were converted to percentages. The category "others" is a collection of all groups that were either unclassified or < 1% across all datasets.

Correlation analysis

Spearman correlations between *Telonemia* abundance (CARD-FISH) and environmental parameters were calculated using the R function "cor" [70]. The results of these correlations are provided in Supplementary Table S5.

Results and Discussion
Phylogenetic analysis of freshwater Telonemia 18S rRNA

In order to get an impression of the occurrence of *Telonemia* in highly diverse freshwaters, we used a collection of 1027 metagenomic assemblies, in addition to mining assemblies from public databases (see section Materials and Methods, Supplementary Table S1). We recovered 574 *Telonemia* 18S rRNA gene sequences, of which 249 sequences were derived from freshwaters (see section Materials and Methods, Supplementary Table S3). Currently, this represents the largest recovery of *Telonemia* sequences from freshwaters and suggests they are widespread in these habitats but clearly not uniformly distributed.

Phylogenetic analyses of these recovered sequences with reference sequences from cultured species recapitulated the groupings of Telo-1 (including *T. subtilis*) and Telo-2 (including *T. antarcticum*) as have been obtained before (Fig. 1, Supplementary Figs S1-S4). However, whereas Telo-1 appears to be a well-defined clade in these phylogenetic trees, this is not the case for Telo-2, which has been also earlier described as a polyphyletic clade [46]. Moreover, inferring phylogenetic relationships within *Telonemia* using 18S rRNA gene sequences alone is problematic owing to most clades having low bootstrap support (Fig. 1, Supplementary Figs S1-S4). Using 18S rRNA gene sequences for placing *Telonemia* within the tree of life has largely provided conflicting placements to multiple different groups, which were finally resolved through a phylogenomics approach. It also appears that there is little additional phylogenetic signal within the 18S rRNA sequences themselves for a robust within-phylum clade delineation. The relatively low diversity within the phylum can be illustrated by the recovery of only 17 representative sequences at 95% identity levels. At 96, 97, 98, 99, and 100% identity, we obtained 23, 36, 61, 134, and 430 representative sequences, respectively. No new consistently supported clades were observed. Given the wide diversity of habitats examined here, it is possible that 18S rRNA diversity within this phylum has been exhaustively sampled. We recovered more sequences from the Telo-2 clade (metagenomic: 177 marine, 194 freshwater; clone libraries: 105 marine, 35 freshwater) than for Telo-1 (metagenomic: 69 marine, 66 freshwater; clone libraries: 31 marine, 8 freshwaters), suggesting a generally wider distribution of Telo-2 in freshwaters than Telo-1.

Abundance of Telonemia in freshwater metagenomes

We examined 589 freshwater metagenomes to obtain a rough estimate of the relative abundances of *Telonemia* in freshwater habitats with respect to other protistan taxa (Supplementary Fig. S5, Supplementary Table S2). Reads of *Telonemia* were found at >1% in 117 samples, >5% in 21 samples, and >10% in six samples (Supplementary Table S2). The highest abundance (ca. 22%) was found in an under-ice sample of Lake Parsens (Sweden). Clustering of 18S rRNA gene taxonomy profiles of samples where *Telonemia* was present did not reveal any outstanding commonalities (Supplementary Fig. S5). These results suggest that *Telonemia* shows a preference for deeper waters as most samples with higher abundances were frequently derived from lake hypolimnia
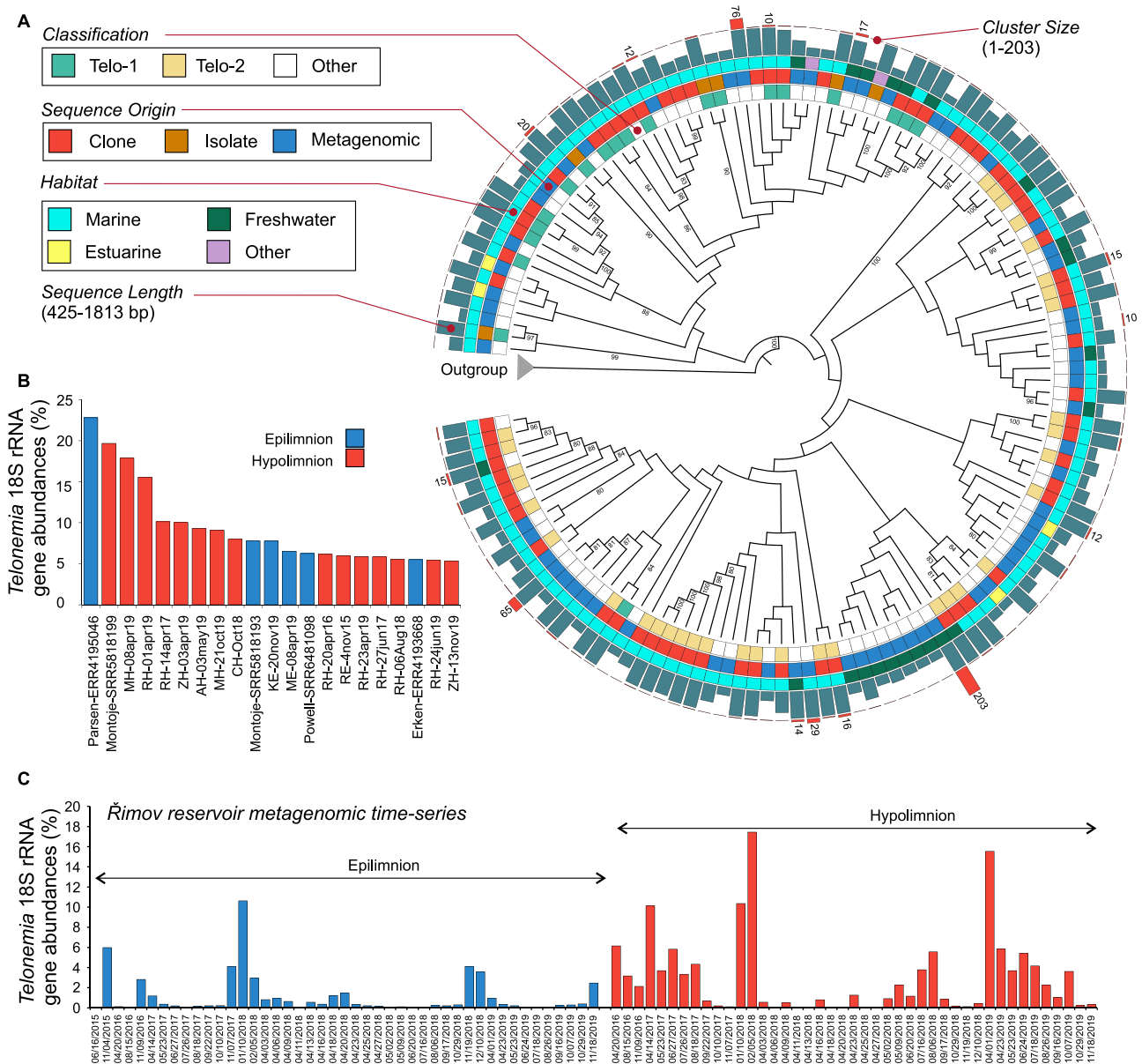
Figure 1. (A) Maximum likelihood phylogenetic tree of representative *Telonemia* 18S rRNA gene sequences clustered at 97% nucleotide identity (shown as a cladogram). The outgroup sequences are not shown but indicated by an arrow. Ultrafast bootstrap values are shown at each node. Two clades, Telo-1 and Telo-2 as defined in previous publications, are shown in different colors. Sequences not classified or <1% are shown as "other." Isolate sequences are shown in bold. The origin of each sequence (clone library, metagenomic, or isolate) is indicated by colored squares. The length of each sequence, and the number of sequences in each sequence cluster are shown at the right as barcharts. Number of sequences in clusters with >10 sequences is shown (B) *Telonemia* abundances estimated with 18S rRNA gene sequences from metagenomes and (C) in a metagenomic time-series of Řimov reservoir.

(Fig. 1). This was also supported by a weak, but significant correlation between *Telonemia* percentages by 18S rRNA gene and depth (*n* = 407, Spearman's *R* = 0.174, *P* value = 4.2e-04). Moreover, examination of longer metagenomic time-series of two sites (Řimov reservoir, Czechia, and Lake Mendota, Wisconsin, USA) showed quite different abundances even though both temperate water bodies are largely eutrophic. Read abundance levels of *Telonemia* 18S rRNA gene sequences in the Lake Mendota dataset were never >1% at any time (Supplementary Table S2). On the other hand, the Řimov reservoir had multiple time points with very high abundances (up to 18%), suggesting *Telonemia* are almost always present in the hypolimnion (Fig. 1). Furthermore, *Telonemia* declined during winter in the hypolimnion, whereas maxima

during winter were recorded for the epilimnion. This may also be due in part to the general higher abundance of prey in the epilimnion coupled with the more favorable lower temperatures in winter.

CARD-FISH analyses of Telonemia in freshwater lakes

CARD-FISH counts of 407 samples confirmed the results already seen in the 18S rRNA gene abundance results (Fig. 1), and *Telonemia* are more abundant in the deeper layers of lakes, where temperatures are generally below 10°C (Fig. 2). In a more extreme case, relative abundances of up to 50% of all eukaryotes were
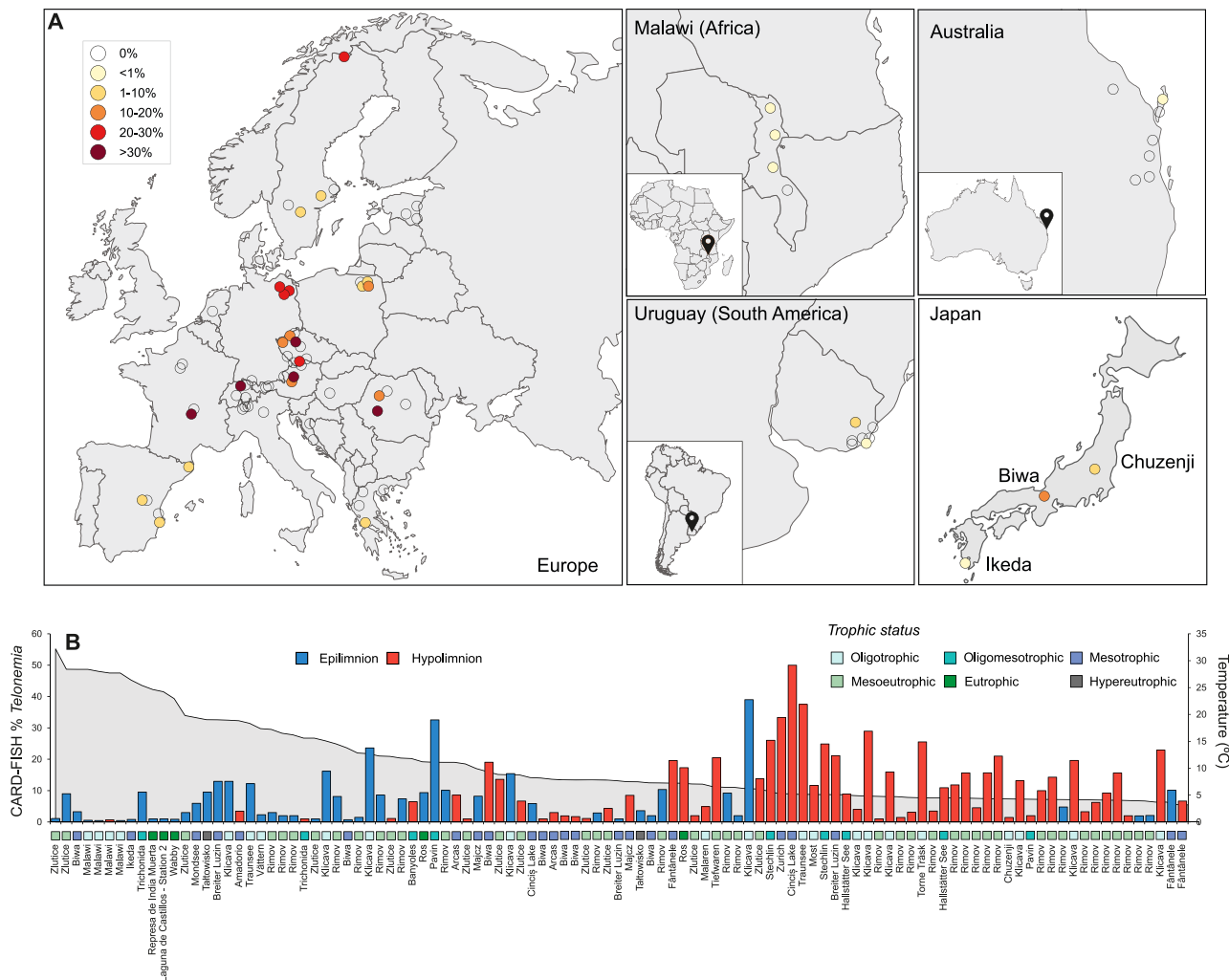
Figure 2. (A) Geographic locations of lake samples used for CARD-FISH in this work. Samples are color-coded according to the maximum %*Telonemia* found at that site using CARD-FISH (see key top left). Sites where *Telonemia* was not detected at all are shown as empty circles. A complete list of all samples is provided in Supplementary Table S1. (B) CARD-FISH counts (%*Telonemia*) in CARD-FISH filters (only those more than 1% are shown here), sorted by decreasing temperature. Epilimnion and hypolimnion samples, along with lake trophic status are shown in different colors.

recorded in the hypolimnion of mesotrophic Lake Cinciş (Romania) (Figs 2 and 3, Supplementary Fig. S6). We also did not observe *Telonemia* in lakes from Africa (Lake Malawi), or from Australia and South America (Fig. 2, Supplementary Table S1) likely because of their elevated temperature (>20°C). We found a strong correlation between the abundance assessments (as % *Telonemia*) between CARD-FISH and from metagenomes (*n* = 51, Spearman's *R* = 0.715, *P* value = 3.84e-09, Supplementary Fig. S7), often not observed for many protist groups [71]. Some possible reasons for concordance may be most likely related to the relatively limited diversity of *Telonemia* 18S rRNA gene sequences, the divergence from other protist groups, and the high specificity of the CARD-FISH probe for this group. Additionally, *Telonemia* were completely absent at sampling sites with a maximum depth of ca. 12 m (79 samples) except for two instances where it was still <1% (Supplementary Table S1).

These counts reveal that *Telonemia* are widely distributed in freshwater habitats (in particular deeper ones) at given conditions, but do not reveal any distinct seasonal patterns that are better examined using time-series analyses. To discover such patterns, we conducted monthly sampling for one year from four distinct water bodies. Seasonal profiles of *Telonemia*

abundance differed greatly. In Řimov reservoir, *Telonemia* appeared to be present in the hypolimnion throughout the entire year reaching low levels only in the coldest time of the year, concomitantly with maxima in the epilimnion (Fig. 3D). In the Kličava reservoir, *Telonemia* reached maxima of up to 40% of all eukaryotes in the epilimnion during the colder months but could also be simultaneously detected in the hypolimnion (unlike in Řimov). *Telonemia* were completely absent in Kličava in summer and autumn when the reservoir is strongly stratified even though the temperatures in the hypolimnion remained stable around 5°C. In Žlutice reservoir, *Telonemia* almost completely disappeared in the autumn–winter months, and appeared again in spring and further increased in the hypolimnion until autumn. Thereafter, they disappeared from the hypolimnion for almost the entire winter period. Increasing temperatures in summer also coincided with the disappearance of *Telonemia* in the Žlutice reservoir at around 10°C in the epilimnion (Fig. 3D). The hypolimnion in this reservoir is anoxic at this time of the year, which may also explain *Telonemia*'s disappearance there. The lowest abundances, however, were recorded in Lake Biwa, where *Telonemia* were undetected most of the time, or at very low abundances (<5%), but reached
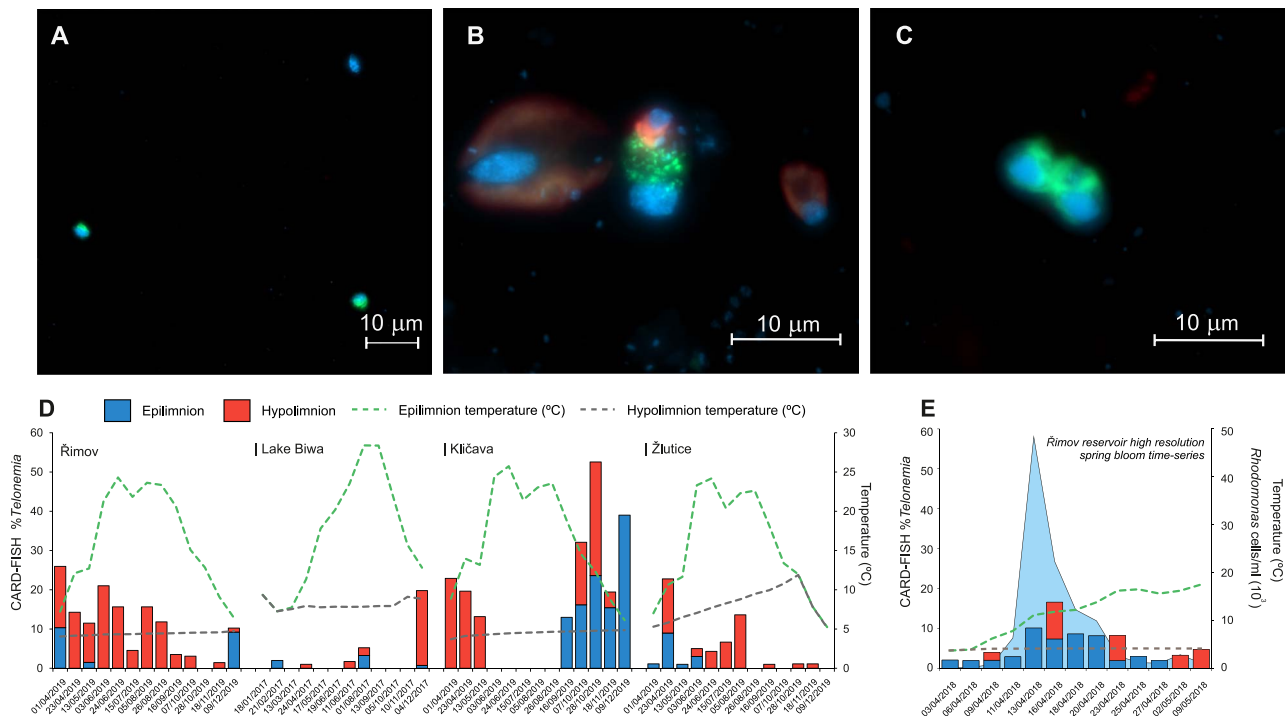
Figure 3. CARD-FISH images of *Telonemia* targeted by the probe Telo-1250. Green: CARD-FISH probe, blue: DAPI, red: autofluorescence (A) from Lake Cinciş (20 m) (B) from Řimov reservoir (0.5 m), *Telonemia* ingesting a *Rhodomonas*. A larger *Cryptomonas* and smaller *Rhodomonas* are also seen left and right, respectively, and (C) a dividing *Telonemia* cell from Breiter Luzin hypolimnion (50 m). All scale bars are 10 μm, panel (A) is in magnification 40X, and panels (B) and (C) are in 100X magnification. (D) Four annual time-series from Řimov reservoir, Lake Biwa, Kličava reservoir and Žlutice reservoir showing relative abundances of *Telonemia* (using CARD-FISH) in epilimnion and hypolimnion. Temperatures are shown as green (epilimnion) and gray-dotted (hypolimnion) lines. (E) Relative abundances of *Telonemia* in epilimnion and hypolimnion during a high-resolution sampling of a spring phytoplankton bloom in Řimov reservoir. Counts of the most abundant Cryptophyte (*Rhodomonas*) are shown as a blue background and temperatures as lines.

a maximum of ca. 20% in a single winter sample. Lake Biwa hypolimnion temperatures appear higher than other lakes (7°C), which is still well below 10°C. Thus, although temperature does appear to be an important factor shaping the occurrence of *Telonemia,* it does not explain entirely the observed patterns in the time series. We also found no correlations with Chl-a, dissolved oxygen, or any other physicochemical parameters (e.g. ammonia, nitrites, nitrates, pH) with relative abundances of *Telonemia* (Supplementary Table S5).

There were several occasions on which *Telonemia* were also found in the epilimnion. On one such occasion, i.e. the spring phytoplankton bloom, they reached up to 15% of all HNF [44]. We analyzed CARD-FISH samples from a high-frequency sampling during the phytoplankton spring bloom (ca. every 3 days) [4] and observed *Telonemia* peaking in the epilimnion (up to 10%) and even ingesting the most abundant Cryptophyte *Rhodomonas* (Fig. 3B). This is likely due to the still relatively low temperatures in the epilimnion in this season and also to the high availability of prey organisms such as cryptophytes. Later in the season, *Telonemia* abundance seems to decrease with increasing temperatures and decreasing abundance of its prey.

Our analyses provide new insights into the distribution and seasonal patterns (at both long and short time intervals) offering clues on possible niche preferences of *Telonemia* in lentic water bodies, which appear to be predicated by a low-temperature regime frequently associated with higher depth (<12 m). It was almost completely absent from shallow or hypertrophic sites. Additionally, *Telonemia* appeared also absent from sites with temperatures >10°C. It also seems that potential prey availability does not appear to be the main driving agent as habitats with extremely high microbial and flagellate populations (e.g. fish ponds) appear totally devoid of *Telonemia* and lake hypolimnia where they are usually resident have lower bacterial and flagellate abundances than surface layers. This is in contrast to the omnipresent occurrence pattern of major bacterivorous flagellates affiliated to the CRY1 lineage, which can be found in a wide temperature range, stretching from the cold hypolimnion to the relatively warmer fish ponds [13]. Indeed, it is likely that *Telonemia* predate upon the CRY1 lineage, which is almost always found in deep hypolimnion at high abundances [13]. The preference of *Telonemia* for deeper water bodies also suggests they are primarily a resident in the hypolimnion, and depending upon favorable environmental conditions (e.g. lower surface temperatures in winter or spring bloom) can transition to the epilimnion. It may also be derived from this that algae like *Rhodomonas* are not their primary prey. However, *Telonemia* show much larger fluctuations in population size, suggesting that even within what appears as a relatively stable hypolimnion, there is sufficient instability in resources through different sedimentation rates, and invasions of additional prey during seasonal algal bloom, thus eventually promoting sudden *Telonemia* blooms (i.e. >20% of all HNF at several sites).

Freshwater food webs, particularly in the hypolimnion, remain little understood. Classical models in ecology have focused largely upon surface layers that show dramatic changes in response to environmental factors [1]. Recently, the distribution and dynamics of protist groups in deeper water layers are increasingly studied revealing a host of diverse flagellates (e.g. kinetoplastids, katablepharids, cercozoans) preying both upon bacteria and other

flagellates or smaller algal cells [11, 13, 72]. However, many of these lineages are known solely from sequence data and cultured representatives are scarce (unlike *Telonemia*). This work shows that *Telonemia* are mostly found in cold, deeper freshwaters making them likely significant predatory flagellates in the food web of deep lakes. However, interactions of *Telonemia* flagellates with the larger microbial community still remain obscure. The general approach taken in this work applied to other important and as yet not fully understood lineages in freshwaters will be key to unravel their identities, lifestyles, and dynamics in the largest, yet less studied habitat of deep lakes.

Acknowledgements

Author contributions

R.B., K.P., and R.G. conceived the work. The PELAGICS Consortium helped organize and sample the European lakes. M.M.S., M.H.,. T.S., V.K., P.L., M.N., and C.F. sampled lakes in Uruguay, Australia, and Malawi. R.B., I.M., K.S., V.K., and K.P. performed CARD-FISH enumerations. R.B., H.P., and R.G. performed the data analysis. R.G. wrote the initial draft with input from all authors. All authors approved the final version of the manuscript.

Supplementary material

Supplementary material is available at *The ISME Journal* online.

Conflicts of interest

The authors declare no conflict of interest.

Funding

Data availability

All *Telonemia* sequences used in this work, along with all alignments and phylogenetic trees are available in Zenodo (doi: 10.5281/zenodo.11237305). Phylogenetic trees are also available for visualization at https://itol.embl.de/shared/rohitghai.

The PELAGICS Consortium

Hans-Peter Grossart[4,5], Antonio Picazo-Mozo[6], Antonio Camacho[6], Pedro J. Cabello-Yeves[6,7], Francisco Rodriguez-Valera[8], Cristiana Callieri[9], Adrian-Stefan Andrei[10], Jakob Pernthaler[10], Thomas Posch[10], Albin Alfreider[11], Ruben Sommaruga[11], Martin W. Hahn[12], Bettina Sonntag[12], Purificación López-García[13], David Moreira[13], Ludwig Jardillier[13], Cécile Lepère[14], Corinne Biderre-Petit[14], Anna Bednarska[15], Mirosław Ślusarczyk[15,16], Viktor Tóth[17], Horia L. Banciu[18], Konstantinos Kormas[19], Sandi Orlić[20,21], Danijela Šantić[22], Gerard Muyzer[23], Daniel P. R. Herlemann[24,25], Helen Tammert[25], Stefan Bertilsson[26], Silke Langenheder[27], Thomas Zechmeister[28], Nico Salmaso[29,30], Nicola Storelli[31,32], Camilla Capelli[33], Fabio Lepori[33,34], Vojtěch Lanta[35], Helena Henriques Vieira[1], Fran Kostanjšek[1], Kateřina Kabeláčová[1], Maria-Cecilia Chiriac[1], Markus Haber[1], Tanja Shabarova[1], Clafy Fernandes[1,2], Pavel Rychtecký[1], Petr Znachor[1], Tiberiu Szőke-Nagy[1], Paul Layoun[1,2], Hon Lun Wong[1], Vinicius Silva Kavagutti[1,2], Paul-Adrian Bulzu[1], Roudaina Boukheloua[1,2], Indranil Mukherjee[1], Hongjae Park[1], Karel Šimek[1,2], Vojtěch Kasalický[1], Michaela M. Salcher[1], Rohit Ghai[1,*].

References

1. Sommer U, Adrian R, De Senerpont DL *et al.* Beyond the plankton ecology group (PEG) model: mechanisms driving plankton succession. *Annu Rev Ecol Evol Syst* 2012;**43**:429–48. https://doi.org/10.1146/annurev-ecolsys-110411-160251

2. Šimek K, Nedoma J, Znachor P *et al.* A finely tuned symphony of factors modulates the microbial food web of a freshwater reservoir in spring. *Limnol Oceanogr* 2014;**59**:1477–92. https://doi.org/10.4319/lo.2014.59.5.1477

3. Shabarova T, Salcher MM, Porcal P *et al.* Recovery of freshwater microbial communities after extreme rain events is mediated by cyclic succession. *Nat Microbiol* 2021;**6**:479–88. https://doi.org/10.1038/s41564-020-00852-1

4. Kavagutti VS, Bulzu P-A, Chiriac CM *et al.* High-resolution metagenomic reconstruction of the freshwater spring bloom. *Microbiome* 2023;**11**:15. https://doi.org/10.1186/s40168-022-01451-4

5. Park H, Shabarova T, Salcher MM *et al.* In the right place, at the right time: the integration of bacteria into the plankton ecology group model. *Microbiome* 2023;**11**:112. https://doi.org/10.1186/s40168-023-01522-0

6. Wetzel RG. *Limnology: Lake and River Ecosystems*. 3rd edn. Academic Press, San Diego, California, 2001.

7. Lampert W, Sommer U, Ulrich LS. *Limnoecology: The Ecology of Lakes and Streams*. Oxford, Oxford University Press, 1997.

8. Kalff J. *Limnology: Inland Water Ecosystems*. Prentice Hall, New Jersey 2002.

9. Villena-Alemany C, Mujakić I, Fecskeová LK *et al.* Phenology and ecological role of aerobic anoxygenic phototrophs in freshwaters. *Microbiome* 2024;**12**:65. https://doi.org/10.1186/s40168-024-01786-0

10. Ram ASP, Palesse S, Colombet J *et al.* The relative importance of viral lysis and nanoflagellate grazing for prokaryote mortality in temperate lakes. *Freshw Biol* 2014;**59**:300–11. https://doi.org/10.1111/fwb.12265

11. Šimek K, Grujčić V, Mukherjee I *et al.* Cascading effects in freshwater microbial food webs by predatory Cercozoa, Katablepharidacea and ciliates feeding on aplastidic bacterivorous cryptophytes. *FEMS Microbiol Ecol* 2020;**96**:fiaa121. https://doi.org/10.1093/femsec/fiaa121

12. Grujcic V, Nuy JK, Salcher MM *et al.* Cryptophyta as major bacterivores in freshwater summer plankton. *ISME J* 2018;**12**:1668–81. https://doi.org/10.1038/s41396-018-0057-5

13. Šimek K, Mukherjee I, Szöke-Nagy T *et al.* Cryptic and ubiquitous aplastidic cryptophytes are key freshwater flagellated bacterivores. *ISME J* 2022;**17**:84–94. https://doi.org/10.1038/s41396-022-01326-4

14. Mukherjee I, Hodoki Y, Nakano S. Seasonal dynamics of heterotrophic and plastidic protists in the water column of Lake Biwa, Japan. *Aquat Microb Ecol* 2017;**80**:123–37. https://doi.org/10.3354/ame01843

15. Leander BS. Predatory protists. *Curr Biol* 2020;**30**:R510–6. https://doi.org/10.1016/j.cub.2020.03.052

16. Tashyreva D, Prokopchuk G, Votýpka J *et al.* Life cycle, ultrastructure, and phylogeny of new Diplonemids and their endosymbiotic bacteria. *MBio* 2018;**9**. https://doi.org/10.1128/mBio.02447-17

17. Mukherjee I, Salcher MM, Andrei A-Ş *et al.* A freshwater radiation of diplonemids. *Environ Microbiol* 2020;**22**:4658–68. https://doi.org/10.1111/1462-2920.15209

18. Piwosz K, Pernthaler J. Enrichment of omnivorous cercozoan nanoflagellates from coastal Baltic Sea waters. *PLoS One* 2011;**6**:e24415. https://doi.org/10.1371/journal.pone.0024415

19. Šimek K, Mukherjee I, Nedoma J *et al.* CARD-FISH and prey tracer techniques reveal the role of overlooked flagellate groups as major bacterivores in freshwater hypertrophic shallow lakes. *Environ Microbiol* 2021;**24**:4256–73. https://doi.org/10.1111/1462-2920.15846

20. Piwosz K, Pernthaler J. Seasonal population dynamics and trophic role of planktonic nanoflagellates in coastal surface waters of the southern Baltic Sea. *Environ Microbiol* 2010;**12**:364–77. https://doi.org/10.1111/j.1462-2920.2009.02074.x

21. Klaveness D, Shalchian-Tabrizi K, Thomsen HA *et al.* Telonema antarcticum sp. nov., a common marine phagotrophic flagellate. *Int J Syst Evol Microbiol* 2005;**55**:2595–604. https://doi.org/10.1099/ijs.0.63652-0

22. Griessmann K. Über marine Flagellaten. *Arch Protistenkd* 1913;**32**:1–78.

23. Hollande A, Cachon J. Structure et affinités d'un flagellé marin peu connu: Telonema subtilis Griesm. *Ann Sci Nat Zool Biol Anim* 1950;109–13.

24. Vørs N. Heterotrophic amoebae, flagellates and Heliozoa from the Tvärminne area, Gulf of Finland, in 1988–1990. *Ophelia* 1992;**36**:1–109. https://doi.org/10.1080/00785326.1992.10429930

25. Tong S, Vørs N, Patterson DJ. Heterotrophic flagellates, centrohelid heliozoa and filose amoebae from marine and freshwater sites in the Antarctic. *Polar Biol* 1997;**18**:91–106. https://doi.org/10.1007/s003000050163

26. Thomsen HA. *Plankton i Indre Danske Farvande. En Analyse Af Forekomsten Af Alger Og Heterotrofe Protister (Ekskl. Ciliater) i Kattegat.* København: Miljøstyrelsen, Miljøministeriet, 1992.

27. Trier H. *Marint Phytoplankton i Subarktis – Kvantitative Og Kvalitative Aspekter.* M. Sc. University of Copenhagen, 1998.

28. Massana R, Balagué V, Guillou L *et al.* Picoeukaryotic diversity in an oligotrophic coastal site studied by molecular and culturing approaches. *FEMS Microbiol Ecol* 2004;**50**:231–43. https://doi.org/10.1016/j.femsec.2004.07.001

29. Romari K, Vaulot D. Composition and temporal variability of picoeukaryote communities at a coastal site of the English Channel from 18S rDNA sequences. *Limnol Oceanogr* 2004;**49**:784–98. https://doi.org/10.4319/lo.2004.49.3.0784

30. Shalchian-Tabrizi K, Eikrem W, Klaveness D *et al.* Telonemia, a new protist phylum with affinity to chromist lineages. *Proc Biol Sci* 2006;**273**:1833–42. https://doi.org/10.1098/rspb.2006.3515

31. Shalchian-Tabrizi K, Kauserud H, Massana R *et al.* Analysis of environmental 18S ribosomal RNA sequences reveals unknown diversity of the cosmopolitan phylum Telonemia. *Protist* 2007;**158**:173–80. https://doi.org/10.1016/j.protis.2006.10.003

32. Lefèvre E, Roussel B, Amblard C *et al.* The molecular diversity of freshwater picoeukaryotes reveals high occurrence of putative parasitoids in the plankton. *PLoS One* 2008;**3**:e2324. https://doi.org/10.1371/journal.pone.0002324

33. Reeb VC, Peglar MT, Yoon HS *et al.* Interrelationships of chromalveolates within a broadly sampled tree of photosynthetic protists. *Mol Phylogenet Evol* 2009;**53**:202–11. https://doi.org/10.1016/j.ympev.2009.04.012

34. Burki F, Inagaki Y, Bråte J *et al.* Large-scale phylogenomic analyses reveal that two enigmatic protist lineages, telonemia and centroheliozoa, are related to photosynthetic chromalveolates. *Genome Biol Evol* 2009;**1**:231–8. https://doi.org/10.1093/gbe/evp022

35. Cavalier-Smith T, Chao EE, Lewis R. Multiple origins of Heliozoa from flagellate ancestors: new cryptist subphylum Corbihelia, superclass Corbistoma, and monophyly of Haptista, Cryptista Hacrobia and Chromista. *Mol Phylogenet Evol* 2015;**93**:331–62. https://doi.org/10.1016/j.ympev.2015.07.004

36. Strassert JFH, Jamy M, Mylnikov AP *et al.* New Phylogenomic analysis of the enigmatic phylum Telonemia further resolves the eukaryote tree of life. *Mol Biol Evol* 2019;**36**:757–65. https://doi.org/10.1093/molbev/msz012

37. Thaler M, Lovejoy C. Biogeography of heterotrophic flagellate populations indicates the presence of generalist and specialist taxa in the Arctic Ocean. *Appl Environ Microbiol* 2015;**81**:2137–48. https://doi.org/10.1128/AEM.02737-14

38. Bachy C, López-García P, Vereshchaka A *et al.* Diversity and vertical distribution of microbial eukaryotes in the snow, sea ice and seawater near the north pole at the end of the polar night. *Front Microbiol* 2011;**2**:106. https://doi.org/10.3389/fmicb.2011.00106

39. Ramond P, Siano R, Sourisseau M *et al.* Assembly processes and functional diversity of marine protists and their rare biosphere. *Environ Microbiome* 2023;**18**:59. https://doi.org/10.1186/s40793-023-00513-w

40. Orsi WD, Wilken S, Del Campo J *et al.* Identifying protist consumers of photosynthetic picoeukaryotes in the surface ocean using stable isotope probing. *Environ Microbiol* 2018;**20**:815–27. https://doi.org/10.1111/1462-2920.14018

41. Piwosz K, Całkiewicz J, Gołębiewski M *et al.* Diversity and community composition of pico- and nanoplanktonic protists in the Vistula River estuary (gulf of Gdańsk, Baltic Sea). *Estuar Coast Shelf Sci* 2018;**207**:242–9. https://doi.org/10.1016/j.ecss.2018.04.013

42. Kammerlander B, Breiner H-W, Filker S *et al.* High diversity of protistan plankton communities in remote high mountain lakes in the European alps and the Himalayan mountains. *FEMS Microbiol Ecol* 2015;**91**. https://doi.org/10.1093/femsec/fiv010

43. Bråte J, Klaveness D, Rygh T *et al*. Telonemia-specific environmental 18S rDNA PCR reveals unknown diversity and multiple marine-freshwater colonizations. *BMC Microbiol* 2010;**10**:168. https://doi.org/10.1186/1471-2180-10-168

44. Mukherjee I, Grujčić V, Salcher MM *et al*. Integrating depth-dependent protist dynamics and microbial interactions in spring succession of a freshwater reservoir. *Environmental Microbiome* 2024;**19**:1–16. https://doi.org/10.1186/s40793-024-00574-5

45. Qu Z, Forster D, Bruni EP *et al*. Aquatic food webs in deep temperate lakes: key species establish through their autecological versatility. *Mol Ecol* 2021;**30**:1053–71. https://doi.org/10.1111/mec.15776

46. Tikhonenkov DV, Jamy M, Borodina AS *et al*. On the origin of TSAR: morphology, diversity and phylogeny of Telonemia. *Open Biol* 2022;**12**:210325. https://doi.org/10.1098/rsob.210325

47. 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. https://doi.org/10.3389/fmicb.2019.02359

48. Blais M-A, Vincent WF, Vigneron A *et al*. Diverse winter communities and biogeochemical cycling potential in the under-ice microbial plankton of a subarctic river-to-sea continuum. *Microbiol Spectr* 2024;**12**:e0416023. https://doi.org/10.1128/spectrum.04160-23

49. Li F, Peng Y, Fang W *et al*. Application of environmental DNA metabarcoding for predicting anthropogenic pollution in rivers. *Environ Sci Technol* 2018;**52**:11708–19. https://doi.org/10.1021/acs.est.8b03869

50. Piwosz K, Mukherjee I, Salcher MM *et al*. CARD-FISH in the sequencing era: opening a new universe of Protistan ecology. *Front Microbiol* 2021;**12**:640066. https://doi.org/10.3389/fmicb.2021.640066

51. Layoun P, López-Pérez M, Haro-Moreno JM *et al*. Flexible genomic island conservation across freshwater and marine Methylophilaceae. *ISME J* 2024;**18**. https://doi.org/10.1093/ismejo/wrad036

52. Chiriac M-C, Bulzu P-A, Andrei A-S *et al*. Ecogenomics sheds light on diverse lifestyle strategies in freshwater CPR. *Microbiome* 2022;**10**:84. https://doi.org/10.1186/s40168-022-01274-3

53. Tran PQ, Bachand SC, McIntyre PB *et al*. Depth-discrete metagenomics reveals the roles of microbes in biogeochemical cycling in the tropical freshwater Lake Tanganyika. *ISME J* 2021;**15**:1971–86. https://doi.org/10.1038/s41396-021-00898-x

54. Andrei A-Ş, Salcher MM, Mehrshad M *et al*. Niche-directed evolution modulates genome architecture in freshwater Planctomycetes. *ISME J* 2019;**13**:1056–71. https://doi.org/10.1038/s41396-018-0332-5

55. Cabello-Yeves PJ, Zemskaya TI, Zakharenko AS *et al*. Microbiome of the deep Lake Baikal, a unique oxic bathypelagic habitat. *Limnol Oceanogr* 2019;**65**:1471–88. https://doi.org/10.1002/lno.11401

56. Mehrshad M, Salcher MM, Okazaki Y *et al*. Hidden in plain sight-highly abundant and diverse planktonic freshwater Chloroflexi. *Microbiome* 2018;**6**:176. https://doi.org/10.1186/s40168-018-0563-8

57. Garner RE, Kraemer SA, Onana VE *et al*. A genome catalogue of lake bacterial diversity and its drivers at continental scale. *Nat Microbiol* 2023;**8**:1920–34.

58. Nawrocki EP, Structural RNA. *Homology Search and Alignment Using Covariance Models*. Ph.D. Washington University in Saint Louis School of Medicine, 2009.

59. 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. https://doi.org/10.1093/nar/gks1219

60. Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 2006;**22**:1658–9. https://doi.org/10.1093/bioinformatics/btl158

61. Chen I-MA, Chu K, Palaniappan K *et al*. The IMG/M data management and analysis system v.7: content updates and new features. *Nucleic Acids Res* 2023;**51**:D723–32. https://doi.org/10.1093/nar/gkac976

62. Ludwig W, Strunk O, Westram R *et al*. ARB: a software environment for sequence data. *Nucleic Acids Res* 2004;**32**:1363–71. https://doi.org/10.1093/nar/gkh293

63. Kahn P, Herfort L, Peterson TD *et al*. Discovery of a Katablepharis sp. in the Columbia River estuary that is abundant during the spring and bears a unique large ribosomal subunit sequence element. *Microbiology* 2014;**3**:764–76. https://doi.org/10.1002/mbo3.206

64. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 2013;**30**:772–80. https://doi.org/10.1093/molbev/mst010

65. Mirarab S, Nguyen N, Guo S *et al*. PASTA: ultra-large multiple sequence alignment for nucleotide and amino-acid sequences. *J Comput Biol* 2015;**22**:377–86. https://doi.org/10.1089/cmb.2014.0156

66. Nguyen L-T, 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. https://doi.org/10.1093/molbev/msu300

67. Hoang DT, Chernomor O, von Haeseler A *et al*. UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol* 2018;**35**:518–22. https://doi.org/10.1093/molbev/msx281

68. Kalyaanamoorthy S, Minh BQ, Wong TKF *et al*. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* 2017;**14**:587–9. https://doi.org/10.1038/nmeth.4285

69. Steinegger M, Söding J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat Biotechnol* 2017;**35**:1026–8. https://doi.org/10.1038/nbt.3988

70. Core Team RR. *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2013.

71. Piwosz K, Shabarova T, Pernthaler J *et al*. Bacterial and eukaryotic small-subunit amplicon data do not provide a quantitative picture of microbial communities, but they are reliable in the context of ecological interpretations. *mSphere* 2020;**5**. https://doi.org/10.1128/mSphere.00052-20

72. 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. https://doi.org/10.3389/fmicb.2019.02375