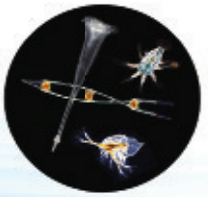
Journal of
Plankton Research

academic.oup.com/plankt



J. Plankton Res. (2019) 41(3): 241–256. First published online June 7, 2019 doi:10.1093/plankt/fbz019

ORIGINAL ARTICLE

Diversity patterns of planktonic microeukaryote communities in tropical floodplain lakes based on 18S rDNA gene sequences

KARINE BORGES MACHADO[1], CÍNTIA PELEGRINETI TARGUETA[1], ADRIANA MARIA ANTUNES[1], THANNYA NASCIMENTO SOARES[1], MARIANA PIRES DE CAMPOS TELLES[1,6], RAMIRO LOGARES[2], LUDGERO CARDOSO GALLI VIEIRA[3], VERA LÚCIA DE MORAES HUSZAR[4] AND JOÃO CARLOS NABOUT[5,*]

[1]UNIVERSIDADE FEDERAL DE GOIÁS. INSTITUTO DE CIÊNCIAS BIOLÓGICAS, CAMPUS II/UFG, CEP 74690-900, GOIÂNIA, GOIÁS, BRAZIL, [2]DEPARTAMENTO DE BIOLOGIA MARINHA E OCEANOGRAFIA CSIC, INSTITUTO DE CIÊNCIAS DEL MAR, 08003 BARCELONA, SPAIN, [3]UNIVERSIDADE DE BRASÍLIA (UNB), CAMPUS PLANALTINA, ÁREA UNIVERSITÁRIA I, VILA NOSSA SENHORA DE FÁTIMA, CEP 73345-010, PLANALTINA, DISTRITO FEDERAL, BRAZIL, [4]UNIVERSIDADE FEDERAL DO RIO DE JANEIRO, MUSEU NACIONAL, SÃO CRISTÓVÃO, CEP 20940-040, RIO DE JANEIRO, BRAZIL, [5]UNIVERSIDADE ESTADUAL DE GOIÁS, CAMPUS ANÁPOLIS DE CIÊNCIAS EXATAS E TECNOLÓGICAS, CEP 75132-903, ANÁPOLIS, GOIÁS, BRAZIL AND [6]PONTIFÍCIA UNIVERSIDADE CATÓLICA DE GOIÁS, ESCOLA DE CIÊNCIAS AGRÁRIAS E BIOLÓGICAS, CEP 74605-010, GOIÂNIA, GOIÁS, BRAZIL

*CORRESPONDING AUTHOR: joao.nabout@ueg.br

Received November 7, 2018; editorial decision April 3, 2019; accepted April 12, 2019

Corresponding editor: John Dolan

The aquatic microbiota plays key roles in ecosystem processes; however, the mechanisms that influence their biogeographic patterns are not yet fully understood. Using high-throughput 18S rDNA gene sequencing, we investigated the composition of planktonic microeukaryotes (organisms sampled using a 68-μm plankton net) in 27 floodplain lakes of the Araguaia River, central Brazil and explored the influence of environmental and spatial factors for communities considering taxonomic and trophic groups. Of the 807 operational taxonomic units (OTUs) observed, Chlorophyta and Charophyta were the groups with greater abundance. Beta diversity was high, and the similarity of communities decreased as the geographic distance increased. We found a shared explanation between environmental and spatial predictors for total and autotrophic microbiota. Environmental variables influence only mixotrophic microbiota. These results suggest an OTU turnover along the floodplain and a spatially structured composition. This spatial pattern can be derived from the association with extrinsic factors, such as spatially structured environmental

available online at academic.oup.com/plankt

variables, that generate spatial dependence. However, the relationship between the composition of microbiota and environmental conditions is still unclear.

KEYWORDS: Araguaia River; Illumina Miseq; metabarcoding; metacommunities

INTRODUCTION

The aquatic microbiota represents a considerable fraction of the total planktonic biomass (Cotner and Biddanda, 2002) and are key elements in aquatic food webs (Chen *et al.*, 2008). Aquatic microbes are responsible for nutrient cycling and degradation of natural and anthropogenic residues, acting in numerous biogeochemical processes (Rodríguez-Valera, 2004). Thus, understanding which factors shape microbial community structure is one of the greatest challenges in ecology (Logares *et al.*, 2013).

Local, regional and historical factors act at different spatial and temporal scales to determine species richness (e.g. alpha and gamma diversity) and turnover (e.g. beta diversity) (Martiny *et al.*, 2006; Lindström and Langenhender, 2012). Two components are typically used to explain the variation of organisms in space: environmental selection and dispersal capacity. According to the niche theory, environmental variables (e.g. habitat conditions, resource availability, abiotic factors) and ecological interactions determine species composition in a community (Hutchison, 1957). On the other hand, spatial variables, such as geographical distance, can influence the structuring of communities by limiting dispersal (e.g. Heino *et al.*, 2014; Ren *et al.*, 2015; Gong *et al.*, 2015). Finally, historical processes may also play a role in community structuring (e.g. Chase, 2003; Chase, 2010), such as dispersion limitation and stochastic events.

Dispersal processes and environmental factors are combined in the metacommunity theory, which considers four paradigms to explain community structuring: patch dynamics, species sorting, mass effects and neutral dynamics (Leibold *et al.*, 2004). The patch dynamic paradigm assumes that patches are identical and that the patterns of diversity are determined by the dispersal capacity of the species. In the species sorting paradigm, there is environmental filtering, and species occur only in environmentally appropriate sites. In this case, dispersion allows the species to migrate according to alterations in the environment, promoting changes in composition. In the mass effect paradigm, high dispersion rates promote a homogenization of the communities since species are able to occur even in sites that are not environmentally favorable. In the neutral perspective paradigm, species are ecologically equivalent, and the patterns of diversity are determined by the stochastic mechanisms of loss or gain of species.

Recently, mass effects and patch dynamics have been considered as subcategories of the species-sorting paradigm, with distinction by the dispersal rate that exists between the communities: low for most species in the patch dynamic paradigm, intermediate in the species sorting paradigm and high in the mass effect paradigm (Winegardner *et al.*, 2012). Thus, patterns of composition in metacommunities are determined by environmental conditions and species dispersion (Winegardner *et al.*, 2012). For aquatic ecosystems such as streams and lakes, species sorting predominates, although these patterns can be altered according to the type of system and spatial scale (Heino *et al.*, 2015).

Despite advances in studies on metacommunities, the processes that explain microbial biogeography are not fully understood (Chen *et al.*, 2008; Hanson *et al.*, 2012; Lindström and Langenhender, 2012). For a long time, it was considered that microbial distribution was determined only by environmental characteristics given their small size, high abundances and high dispersal rates (Finlay, 2002; Fenchel and Finlay, 2004). This is in line with the principle of Baas-Becking (1934), which states "everything is everywhere, but the environment selects". However, it has also been proposed that dispersal limitation may occur in microbes (Martiny *et al.*, 2006). Environmental characteristics of aquatic environments, such as pH (Heino *et al.*, 2014; Gong *et al.*, 2015), conductivity (Simon *et al.*, 2015a), phosphorus concentration (Wang *et al.*, 2015; Triadó-Margarit and Casamayor, 2012), luminosity (Charvet *et al.*, 2014), primary productivity (Bradford *et al.*, 2013; Simon *et al.*, 2015a; Wang *et al.*, 2015), temperature and depth (Gong *et al.*, 2015; Wang *et al.*, 2015) can affect microbial community composition. On the other hand, spatial variables have also been tested as predictors for community structuring and play an important role together with environmental conditions (e.g. Soininen *et al.*, 2011; Gong *et al.*, 2015).

For microorganisms, advances in DNA sequencing can help us to understand numerous ecological processes at the community level (Handelsman, 2009). Thus, by means of metabarcoding, environmental DNA can be amplified and high-throughput sequenced using universal molecular markers for both eukaryotic and prokaryotic genes (Pawlowski, 2014). Thus, DNA sequences corresponding to multiple organisms present in the communities are classified into operational taxonomic

units (OTUs) according to their similarity (e.g. Heino *et al.*, 2014; Fonseca *et al.*, 2014, Logares *et al.*, 2014). The latter strategy complements the information obtained via traditional morphological identification (Bradford *et al.*, 2013; Santoferrara *et al.*, 2015), allowing the study of rare (e.g. Logares *et al.*, 2014; Grattepanche *et al.*, 2016) or little-known taxa (Bik *et al.*, 2012) and enabling the evaluation of the effects of different functional groups (i.e. different trophic roles in ecosystems) simultaneously (e.g. Simon *et al.*, 2015b; Genitsaris *et al.*, 2016; Khomich *et al.*, 2017).

Comparisons between results obtained using high-throughput sequencing and classic microscopy-based morphological identification have shown similar spatial distribution patterns (e.g. see Santoferrara *et al.*, 2015 for ciliates; Hirai *et al.*, 2015 for copepods). The characterization of communities using high-throughput sequencing may facilitate the identification of very small taxa whose morphological characteristics are difficult to differentiate visually (Simon *et al.*, 2015b). Considering that DNA-based classification produces a large amount of information in a short period of time, its use can assist in the taxonomic characterization with higher resolution, reduce the time and costs employed in the process and enable a larger sampling effort (Keck *et al.*, 2017).

Although the number of studies using high-throughput sequencing has grown in recent years, the majority have focused on microorganisms in marine environments (e.g. Countway *et al.*, 2010; Bik *et al.*, 2012; Charvet *et al.*, 2014; Fonseca *et al.*, 2014; Logares *et al.*, 2014; Gong *et al.*, 2015; Genitsaris *et al.*, 2016; Grattepanche *et al.*, 2016; Zhang *et al.*, 2017), while investigations in freshwater environments are scarce (Simon *et al.*, 2015b) and have been mostly conducted in temperate zones (e.g. Amaral-Zettler, 2013; Leperé *et al.*, 2013; Heino *et al.*, 2014; Filker *et al.*, 2015; Kammerlander *et al.*, 2015). An exception was found in a tropical Australian river where it was observed that the composition of the eukaryote community was influenced by environmental variables (Bradford *et al.*, 2013).

In this study, we investigated the composition and diversity (alpha, beta and gamma) of planktonic microeukaryotes in 27 floodplain lakes in Araguaia River, central Brazil, using 18S rDNA gene high-throughput sequencing of targeted markers. We focused on the larger size fraction, retained in a 68-μm mesh plankton net. This strategy was adopted to capture a larger volume of water and consequently allows the sampling of most planktonic microeukaryote groups present in the lakes. To the best of our knowledge, this is one of the first studies to analyze the planktonic microeukaryotes in the Araguaia River, an important river in the Cerrado biome, through a metabarcoding approach and one of the few conducted in the tropical region. Thus, our aims are as follows: (i) to describe the diversity of planktonic microeukaryotes for this region, (ii) to determine the spatial diversity

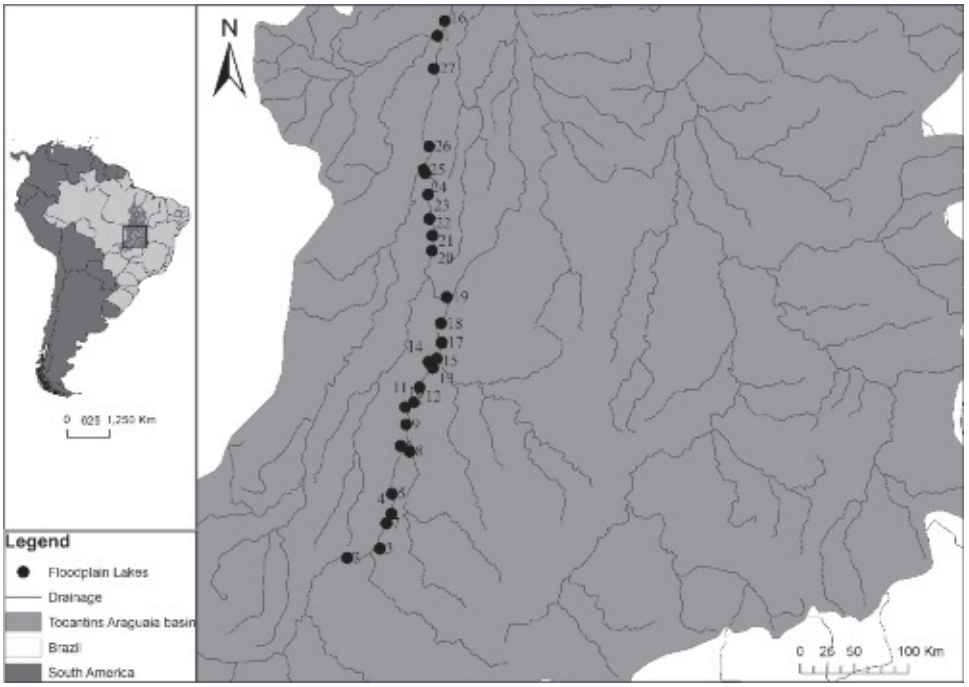


Fig. 1. Map and location of Araguaia River basin, showing the sampled lakes. The numbers indicate lake codes.

patterns of these communities along the floodplain and (iii) to evaluate the relative influence of environmental and spatial variables on the microbiota composition considering taxonomic and trophic groups. We hypothesized that environmental variables may be the major predictors of community composition, since floodplain lakes are connected to each other along the plain by the main channel of the river.

METHOD

Study area

The Araguaia River is located in the Tocantins–Araguaia basin and represents one of the most important rivers in the central region of Brazil (Valente *et al.*, 2013). Along its course, it is divided into three regions: high, medium and low Araguaia (Latrubesse and Steuvax, 2006). The medium course is characterized by the confluence of important tributaries such as the Crixás, Mortes and Cristalino rivers (Latrubesse and Steuvax, 2006) and by the presence of the floodplain with numerous lakes (Morais *et al.*, 2005), which contribute to maintaining the diversity and functioning of ecosystems in this region.

In this study, we sampled 27 lakes in the central region of Brazil along 500 km in the Araguaia River floodplain and its tributaries, including the Crixás, Vermelho and Mortes Rivers (Fig. 1). Sampling was performed in January 2012, which represents the flood period on the Araguaia River (Aquino *et al.*, 2008), a period that allows better access to the studied area. These lakes are considered oligotrophic, with low nutrient concentrations (Marcionilio *et al.*, 2016). Soil cover is predominantly composed of native Cerrado vegetation (a type of savanna), with some areas used for agriculture and livestock (Machado *et al.*, 2016).

DNA extraction, amplification, sequencing and bioinformatics

Aquatic microbial samples were collected from the water subsurface (0.5 m). Approximately 500 L of water was filtered through a plankton net (68 μm mesh size), concentrated to 250 mL. The concentrated plankton was stored in polyethylene bottles in a refrigerator (~2°C). The filtration of the concentrated sample was performed within 12 hours of sampling using Millipore cellulose filters (3 μm pore size) and a vacuum pump for DNA collection. Thus, the size fraction investigated was predominantly composed by microeukaryotes larger than 68 μm, although microeukaryotes smaller than 68 μm and nanoeukaryotes may also have been captured. The filters were stored in liquid nitrogen at −80°C.

Total genomic DNA was extracted and purified following the PowerWater DNA Isolation kit (MoBio, USA) protocol for each collection sample. The extracted DNA was

Table I: OTUs richness (OTUs number) and OTUs abundance (% of reads number) according to the taxonomic and trophic groups

Taxonomic group	OTUs Richness	Number of reads (%)	Trophic role	References
Amoebozoa	1	0.006	Heterotrophic	Lesen *et al.*, 2010
Apusozoa	1	0.01	Heterotrophic	Boenigk and Arndt, 2002
Bacillariophyta	11	0.41	Autotrophic	Khomich *et al.*, 2017
Bicosoecida	5	0.06	Heterotrophic	Khomich *et al.*, 2017
Centrohelida	3	0.06	Heterotrophic	Khomich *et al.*, 2017
Cercozoa	11	0.14	Heterotrophic	Khomich *et al.*, 2017
Charophyta	23	48	Autotrophic	Khomich *et al.*, 2017
Chlorophyta	44	42	Autotrophic	Simon *et al.*, 2015a; Khomich *et al.*, 2017
Choanoflagellata	2	0.01	Heterotrophic	Simon *et al.*, 2015a; Khomich *et al.*, 2017
Chrysophyta	28	0.86	Mixotrophic	Jones, 2000
Ciliophora	55	3	Heterotrophic	Beaver and Crisman, 1989; Simon *et al.*, 2015a
Colpodellida	1	0.01	Heterotrophic	Myl'nikova and Myl'nikov, 2009
Colponemidia	1	0.006	Heterotrophic	Tikhonenkov *et al.*, 2014
Cryptophyta	14	3	Mixotrophic	Simon *et al.*, 2015a
Dictyochopytes	1	0.01	Autotrophic	Khomich *et al.*, 2017
Dinoflagellata	6	0.75	Mixotrophic	Stoecker, 1999
Eustigmatophyceae	1	0.01	Autotrophic	Fietz *et al.*, 2005
Fungi	19	0.57	Heterotrophic	Simon *et al.*, 2015a
Hyphochytriomycota	1	0.03	Heterotrophic	Beakes and Thines, 2016
Ichthyosporea	2	0.01	Heterotrophic	Glockling *et al.*, 2013
Perkinsidae	4	0.47	Heterotrophic	Mangot *et al.*, 2011
Peronosporomycetes	9	0.55	Heterotrophic	Dick, 2001
Raphidophytes	1	0.03	Autotrophic	Khomich *et al.*, 2017

visualized on a 1% agarose gel. A hypervariable fragment, ~380 bp of the V4 region of the 18S rDNA gene was amplified using the universal primers TAReuk454FWD1 and TAReuKREV3 (Stoeck *et al.*, 2010), modified with addition of the sequences complementary to the Illumina indices (i7 and i5, Nextera XT). We used 5 ng/μL of DNA and 1 μM of each primer. Polymerase chain reaction (PCR) was performed using the Taq PCR Master Mix Kit (Qiagen) in triplicate for each sample to minimize the PCR bias. PCR cycles were as follows: 98°C for 1 minute, followed by 98°C for 30 seconds, 53°C for 30 seconds and 72°C for 30 seconds and a final extension at 72°C for 10 minutes.

Then, the Illumina indices (i7 and i5) were inserted into the fragments using a limited-cycle PCR program. We used the Nextera XT index kit v2 set B. Finally, the amplicon libraries were purified, and the short library fragments were removed using Agencourt AMPure XP Beads (Beckman Coulter). The libraries were quantified by real-time PCR with the KAPA Library Quantification Kit, and the amplicon size was estimated using an Agilent High-Sensitivity DNA Kit on a Bioanalyzer. The libraries were normalized to 4 nM and pooled for sequencing using the MiSeq Reagent Kit v3 (300 cycles) on the Illumina MiSeq platform.

The sequence quality was evaluated using FastQC software (available from https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) (Andrews, 2010). Sequences with reads <100 bp or bases with a Phred score of <20 were excluded using the Trimmomatic software (Bolger *et al.*, 2014). The OTU prediction was carried out following the UPARSE pipeline, available at http://drive5.com/usearch/manual/uparse_pipeline.html and in Edgar (2013), which consists of (i) merging sequences, (ii) grouping the sequences of all samples, (iii) identification of unique sequences (dereplication), (iv) identification of OTUs with representative sequences (clustering by 97% similarity) and chimera filtering and (v) construction of an OTU table by sample. The taxonomic prediction was performed by a BLAST search (Altschul *et al.*, 1990) of the representative OTU sequences against the Silva 119.1 database (available in https://www.arb-silva.de/) using a percent identity of 97%. The Metazoa sequences were removed from the total microbiota data set.

The OTUs for which it was possible to assign a taxonomic classification were classified into trophic groups according to their nutrition type (see Simon *et al.*, 2015a; Genitsaris *et al.*, 2016; Khomich *et al.*, 2017). We considered autotrophic organisms those with chlorophyll-a, which act as primary producers (e.g. Chlorophyta, Charophyta, Glaucophyta, Rhodophyta). Organisms classified as mixotrophic (see Flynn *et al.*, 2013 for a detailed description of this nutrition mode) include those that acquire carbon through photosynthesis or heterotrophy (e.g. some putative members of the groups Cryptophyta, Chrysophyta, Ochrophyta, Dinophyta, etc.). Heterotrophic organisms were represented by predators, parasites and decomposers (e.g. fungi, Amoebozoa, Ciliophora, Choanoflagellata, etc.). See Table 1 for the complete OTU classification of trophic groups. Thus, four matrices for OTU abundance were used, hereinafter denominated as follows: (i) total microbiota (represented by all OTUs obtained in sequencing, excluding metazoans, plants and those organisms for which it was not possible to obtain a taxonomic classification); (ii) autotrophic microbiota (includes only primary producers); (iii) mixotrophic microbiota (includes only putative mixotrophic organisms); and (iv) heterotrophic microbiota (includes only heterotrophic organisms).

Environmental variables

In each lake, we measured conductivity, dissolved oxygen, pH, depth, water temperature, transparency, turbidity, oxygen saturation and total dissolved solids. The total nitrogen, total phosphorus and chlorophyll-a estimates were performed in the laboratory according to the methods described in Zagatto *et al.*, (1981) and Golterman *et al.*, (1978). We also determined the areas and widths of lakes and percentage of native Cerrado vegetation, pasture and agriculture. Details of the limnological, morphometric and land use variable estimations as well as the description of their values are presented in Machado *et al.*, (2016) and Marcionilio *et al.*, (2016).

We used a principal component analysis (PCA) to evaluate the environmental heterogeneity among lakes. The PCA was constructed using the correlation matrix and the data standardized by the *z*-score method (limnological variables) or transformed to arcsines of their square roots × $180/\pi$ (land use data). The variance inflation factor (VIF) was used to estimate the collinearity between the environmental variables, and those with VIF values greater than 10 were considered collinear (Alin, 2010, see more details in supplementary data). The variables conductivity, pH, water temperature, total dissolved solids, transparency, depth, total phosphorus, total nitrogen, chlorophyll-a and lake width did not present collinearity and thus constituted the set of environmental variables used in the redundancy analysis.

Spatial variables

In this study, we used two types of spatial filters, "*Principal Coordinates of Neighbour Matrices*" (PCNM, Borcard and

Legendre, 2002) and "*Asymmetric Eigenvector Maps*" (AEM, Blanchet *et al.*, 2008a). The PCNM considers the linear distances (Borcard and Legendre, 2002) and may indicate a non-directional dispersion process. The AEM is calculated through the directional connectivity between the sampling points (Blanchet *et al.*, 2008a) and can represent the dispersion along the river course (see more details in supplementary data). Thus, the PCNM and AEM filters represent different hypotheses about connectivity between sites (Heino *et al.*, 2015). The spatial component was represented by those PCNM and AEM filters that showed the highest correlation with OTU abundance. The filters were selected using the forward.sel function (Blanchet *et al.*, 2008b) of the Packfor package (Dray *et al.*, 2011). The PCNM filters were constructed using the package Vegan (Oksanen *et al.*, 2016) and the AEM filters using the package AEM (Blanchet *et al.*, 2008a), both in R software (R Core Team, 2016).

Data analysis

All statistical analyses were performed using the Vegan package (Oksanen *et al.*, 2016) of R software (R Core Team, 2016). We constructed a species accumulation curve to verify how well the OTUs were sampled. If the curve reached the asymptote before the inclusion of all the sampled sites, this indicated that the sampling effort may have been adequate. We used subsampling rarefaction to correct the bias that can be generated by comparing samples with different sizes since a larger number of sequences in a sample leads to a greater number of OTUs (e.g. Bradford *et al.*, 2013; Gong *et al.*, 2015). The rarefaction was conducted through a random subsampling in which the sample size was represented by the lowest number of sequences, i.e. 2 673 reads, recovered from our floodplain lakes (Hurlbert, 1971). This sample size was adequate to represent the community

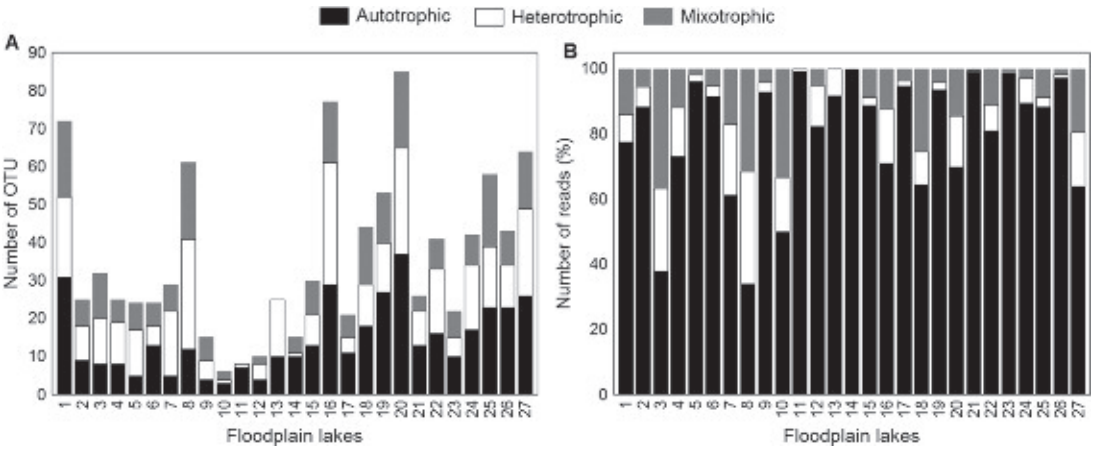


Fig. 2. Number of OTUs (A) and number of reads (B) per lake along the Araguaia River floodplain, divided into trophic groups. The sum of all groups indicates the OTUs richness in each floodplain lake (total microbiota).

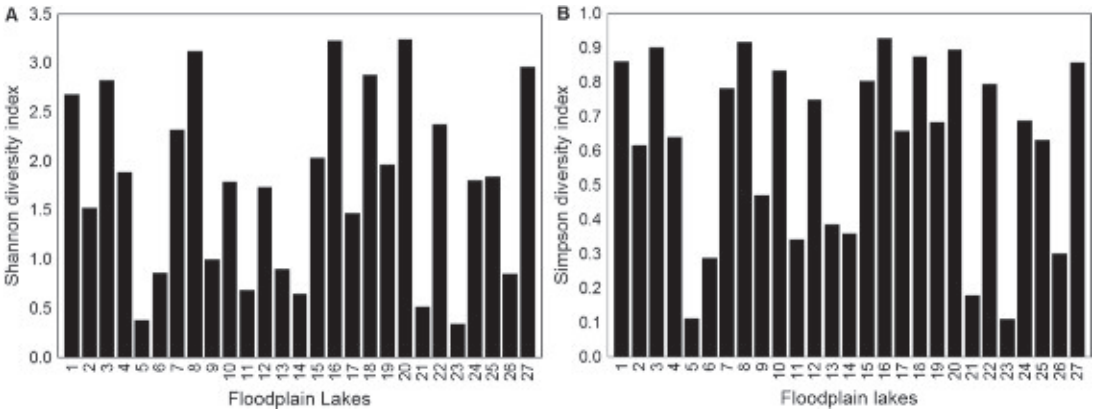


Fig. 3. Shannon and Simpson diversity indices for the total microbiota along the Araguaia River floodplain.

patterns using the Illumina platform (Caporaso *et al*., 2011) and similar to many studies for microeukaryotes (e.g. Bradford *et al*., 2013; Gong *et al*., 2015). The curve was constructed using the "exact" method of the "specaccum" function, and rarefaction analysis was performed using the "rarefy", "rrarefy" and "rarecurve" functions. The rarefied richness and abundance data were used for all the statistical analyses described below.

The alpha diversity was represented by the number of OTUs (richness) and the Shannon (Shannon, 1948) and Simpson indices (Simpson, 1949). The alpha diversity indices were calculated using the diversity function in Vegan package. The gamma diversity was expressed as the total number of OTUs obtained considering the whole set of samples. The total beta diversity was estimated using the Sørensen index and partitioned into turnover and nestedness (Baselga, 2010). The Sørensen index varies from 0 to 1, with a value closer to 1 representing greater total beta diversity (Baselga, 2010). High values for turnover indicate that changes in community composition occur mainly due to OTU replacement along the river, while a high value for nestedness indicates that lakes with low species richness are populated by a subset of OTUs that occur in lakes with higher species

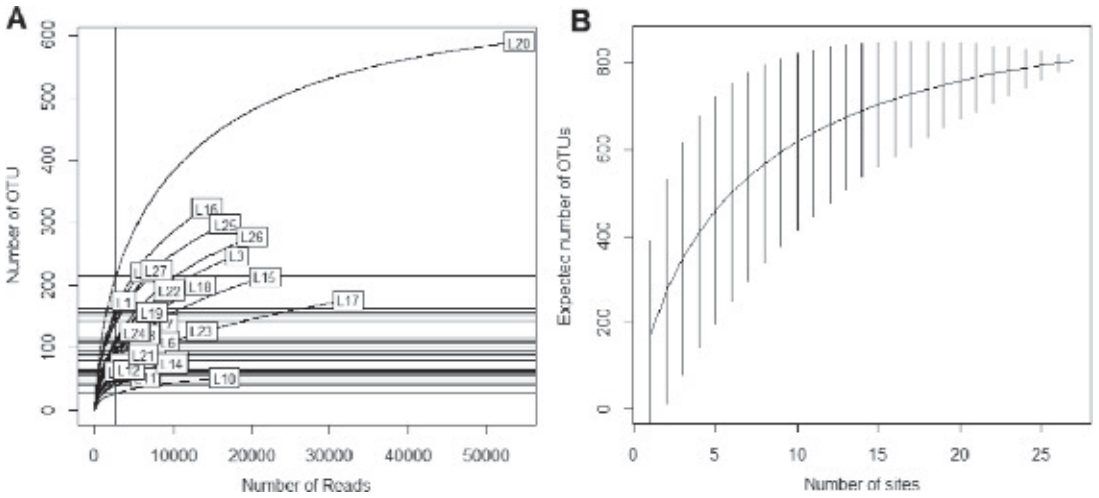


Fig. 4. Rarefaction curves for the 27 samples (**A**) and species accumulation curve for the total microbiota sampled in the floodplain of the Araguaia River (**B**). The vertical line in the rarefaction curve indicates the level of subsampling (2673 reads per sample). The vertical bars in species accumulation curve indicate the mean and standard error for sampled sites. Only L20 and L10 showed a plateau, suggesting that planktonic protist were not completely recovered in the most of lakes. However, the species accumulation curve stabilized when all samples were considered. L indicates lake.

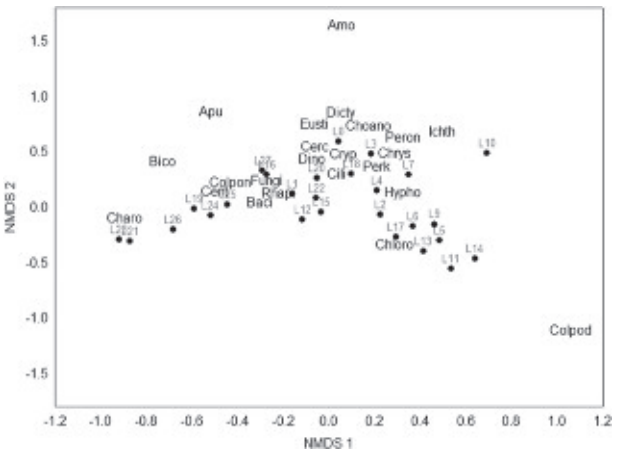


Fig. 5. NMDS based on Bray–Curtis dissimilarities. The codes used to describe the major taxonomic groups are as follows: Amo, Amoebozoa; Apu, Apusozoa; Baci, Bacillariophyta; Bico, Bicosoecida; Cent, Centrohelida; cerc, Cercozoa; Charo, Charophyta; Chloro, Chlorophyta; Choano, Choanoflagellata; Chrys, Chrysophyta; Cili, Ciliophora; Colpod, Colpodellida; Colpon, Colponemidia; Cryp, Cryptophyta; Dicty, Dictyochopytes; Dino, Dinoflagellata; Eusti, Eustignematophyceae; Fungi, Fungi; Hypho, Hyphochytriomycota; Ichth, Ichthyosporea; Perk, Perkinsidae; Peron, Peronosporomycetes; Raphi, Raphidophytes.

richness. Beta diversity was estimated using the Betapart package (Baselga and Orme, 2012). We used non-metric multidimensional scaling (NMDS) based on Bray–Curtis dissimilarities to evaluate the community composition along the floodplain. The abundance data were Hellinger transformed. Alpha and beta diversity were estimated for the total microbiota.

The Mantel correlogram was used to evaluate whether the planktonic microbiota is spatially structured, i.e. the lakes have similar or different OTU compositions according to the distance between them (see Legendre and Fortin, 1989; Fortin and Dale, 2005). This analysis was also conducted considering the total microbiota. The OTUs were correlated to a geographic matrix with five distance classes through a Mantel test, and the results were plotted in a correlogram. The number of classes was defined according to Sturge's rule (see Legendre and Legendre, 1998). The OTU abundance was Hellinger transformed and then converted into a Bray–Curtis distance matrix. The geographic coordinates was converted into a Euclidean distance matrix.

To determine the influence of environmental variables and geographic distance on the community composition, we used a partial redundancy analysis (pRDA, see Legendre and Legendre, 1998). Thus, we determined what proportion of the variation in community composition was explained by the environment (component a), spatial variables (component c), environment and spatial distance together (component b), and unexplained as the residual fraction (component d). The pRDAs were performed separately for the total, autotrophic, mixotrophic and heterotrophic microbiota. pRDA analysis was also performed between total microbiota, trophic groups (autotrophic, mixotrophic and heterotrophic microbiota) and the major taxonomic groups against each environmental variable separately. The purpose of this analysis was to verify whether specific environmental variables were more strongly related than others to the community structure. In this analysis, the environmental variables were standardized using the z score, and the OTU abundance data were Hellinger transformed.

RESULTS

Composition and diversity of microeukaryotes communities

The sequencing of all samples generated 3 287 448 reads. After quality filtering, these reads resulted in 357 933 merged sequences that were predicted in 807 OTUs. The mean number of OTUs was 163 and ranged from 52 to 589 OTUs between the samples. For reads, the mean number was 12 439, ranging from 2 673 to 54 176 between the floodplain lakes sampled. The rarefaction led to a reduction in the number of OTUs (630), with a mean richness of 95, ranging from 28 to 223 OTUs per lake. The OTU abundance was rarefied to 2 673 reads per sample. Of the 630 OTUs, it was possible to perform taxonomic annotation only for 342 OTUs (54%) since not all sequences had representatives in the reference database. Of the 342 taxonomically classified OTUs, 90 OTUs referred to non-protistan sequences (e.g. metazoan) and eight to non-planktonic organisms. Thus, they were excluded, resulting in 244 OTUs used for all analyses (Supplementary data, Table SI). Of these 244 OTUs, 81 (33%) referred to autotrophic organisms, 115 (47%)

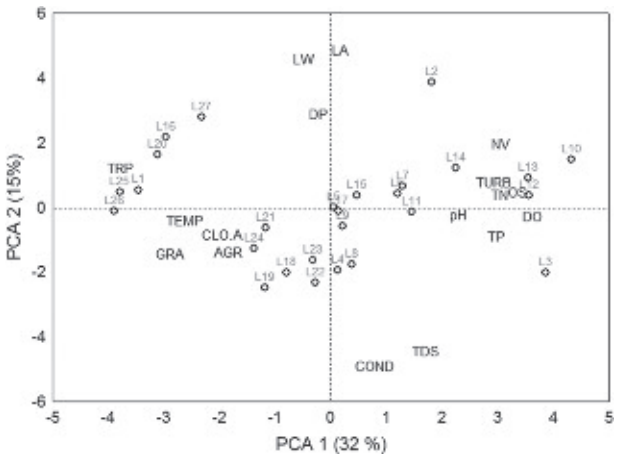


Fig. 6. First and second axes for the PCA applied to environmental variables in the Araguaia River floodplain lakes. The codes of the environmental variables are as follows: COND, conductivity; DO, dissolved oxygen; pH, potential of hydrogen; TEMP, temperature; TURB, turbidity; TDS, total dissolved solids; TRP, transparency; OS, oxygen saturation; DP, depth; TP, total phosphorus; TN, total nitrogen; CHL-A, chlorophyll-a; LA, lake area; LW, lake width; NV, native vegetation; GRA, grassland; AGR, agriculture.

to heterotrophic organisms and 48 (20%) to mixotrophic organisms.

Considering the total microbiota, the Charophyta and Chlorophyta groups presented the highest abundance, with 48% and 42% of reads, respectively (Table 1). Among the autotrophic organisms, the most relevant were also Charophyta (48%) and Chlorophyta (42%); among the mixotrophic, Cryptophyta (3%) and Chrysophyta (0.86%); and among the heterotrophic, Ciliophora (3%) featuring the highest number of reads (Table 1).

For the total microbiota, 23 taxonomic groups were identified, and Ciliophora (55 OTUs), Chlorophyta (44 OTUs) and Chrysophyta (28 OTUs) showed the largest number of OTUs considering all floodplain lakes evaluated (Table 1). Among the autotrophic organisms, the most representative were Chlorophyta (44 OTUs) and Charophyta (23 OTUs); among the mixotrophic, Chrysophyta (28 OTUs) and Cryptophyta (14 OTUs); while for the heterotrophic organisms, Ciliophora (55 OTUs) and fungi (19 OTUs) showed the greatest OTU richness (Table 1).

The alpha diversity was different between lakes. The OTU richness (Fig. 2A) as well as the Shannon and Simpson indices (Fig. 3) varied greatly between them. We also observed a predominance of autotrophic groups in relation to the OTU richness (Fig. 2A) and read number in most lakes (Fig. 2B). No mixotrophic organisms were found in lakes 11 and 13. The gamma diversity was 807 OTUs. After the rarefaction samples, this value decreased to 630. The rarefaction curves indicated that the diversity of the planktonic microbiota was not fully recovered in most of the lakes (Fig. 4A). However, the species accumulation curve plateaued when all samples were considered, indicating that the gamma diversity was recovered (Fig. 4B). This indicates that the OTU diversity in this region was well sampled. A high beta diversity value (βSorensen = 0.92) was observed, mostly associated with OTU replacement along the floodplain (β *turnover* = 0.85; β *nestedness* = 0.07). The NMDS analysis indicates that communities display biogeographical patterns in their composition since different taxonomic groups were associated with different lakes (Fig. 5).

Environmental variables

The first and second axes of the PCA represented 42% of the variability in the environmental data. Floodplain lakes displayed heterogeneous environmental characteristics (Fig. 6), and some lakes were more associated with certain environmental variables than others. In fact, lakes 6, 7, 10, 11, 12, 13, 14, and 15 were associated mainly with native vegetation, turbidity, total nitrogen, total phosphorus, dissolved oxygen, pH and oxygen saturation. Lakes 3, 4, 8, 9, 18, 19, 22, 23 were associated with conductivity and total dissolved solids. The variables transparency, temperature, chlorophyll-a, grassland and agriculture were important for lakes 1, 21, 24, 25 and 26. Lakes 16, 20 and 27 were associated with lake width.

Spatial pattern and variance partitioning

We found a spatial pattern for aquatic microbiota along the Araguaia River (Fig. 7). The OTU composition was similar between floodplain lakes located <70 km from each other. Beyond this distance, the similarity in composition decreased drastically. The environmental and spatial predictors (both directional and non-directional) were not able to explain the variation in the total microbiota composition, as well as the autotrophic and heterotrophic variation (Table 2). However, we found that the envi-

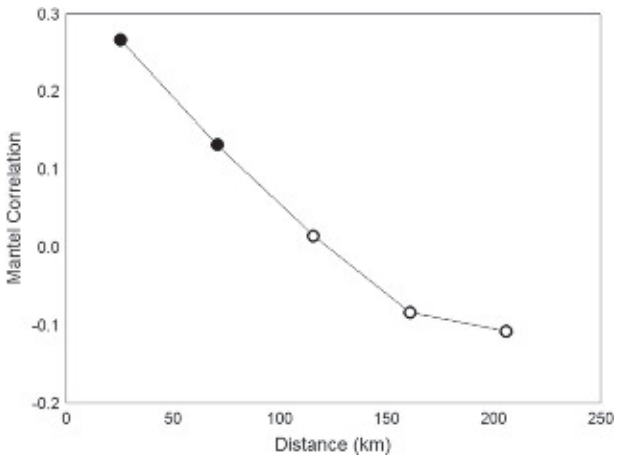


Fig. 7. Mantel correlogram for total microbiota in the Araguaia River floodplain. Bold symbols indicate significant results ($P \leq 0.05$).

Table II: Partial redundancy analysis performed between OTUs abundance (total, autotrophic, heterotrophic and mixotrophic microbiotas), environmental variables and spatial predictors

Groups	Environment variables (a)		Environment and spatial variables (b)	Spatial variables (c)		Residual (d)
	R^2 adj	*P*	R^2 adj	R^2 adj	*P*	R^2 adj
Total microbiota	0	0.74	0.22	0.07	0.13	0.71
Autotrophic	0	0.79	0.32	0.05	0.31	0.63
Heterotrophic	0.01	0.30	0.003	0.01	0.27	0.98
Mixotrophic	0.12	0.03	0.02	0.04	0.17	0.82

We considered a significance level of 5% ($P \leq 0.05$). In each analysis, the following spatial filters were selected by the 'for selection' function. Components (b) and (d) are not testable significance.
R^2 adj indicates R^2 adjusted; total microbiota = PCNM 1 and 2, AEM 1 to AEM 3; autotrophic = PCNM 1 and 4, AEM 1 to AEM 4; heterotrophic = PCNM 1, AEM 1; mixotrophic = PCNM 1, AEM 1 to AEM 3.

Table III: Coefficient of determination ($R^2adj = R^2$ adjusted) and P values obtained for the component 'a' (environment) of partial redundancy analysis performed between OTUs abundance (total, autotrophic, heterotrophic, mixotrophic) and each of the environmental variables

	Total microbiota		Autotrophic		Heterotrophic		Mixotrophic	
	R^2 adj	*P*	R^2 adj	*P*	R^2 adj	*P*	R^2 adj	*P*
Conductivity	0	0.90	0	0.93	0	0.78	0.02	0.16
pH	0	0.44	0	0.39	0.0004	0.37	0.05	0.02
Temperature	0	0.88	0	0.75	0	0.87	0.005	0.32
TDS	0	0.98	0	0.98	0	0.59	0.003	0.34
Transparency	0	0.63	0	0.41	0	0.87	0	0.69
Depth	0	0.75	0	0.71	0	0.40	0.02	0.10
Total phosphorus	0.02	0.16	0.04	0.07	0	0.71	0	0.76
Total nitrogen	0.003	0.37	0.003	0.35	0.02	0.08	0.01	0.22
Chlorophyll-a	0	0.65	0	0.69	0.02	0.10	0.01	0.23
Width	0.03	0.10	0.02	0.13	0.002	0.40	0.02	0.07

In this analysis, we considered one environmental variable at a time controlled by other environmental variables and spatial filters. Significant values are highlighted in bold ($P \leq 0.05$).
TDS: total dissolved solids

ronmental variables predicted the composition of the mixotrophic microbiota (Table 2). We also found considerable importance for the shared explanation component between environmental and spatial variables for the total and autotrophic microbiota (Table 2), although its significance cannot be statistically tested. The mixotrophic microbiota was weakly associated with pH (Table 3). The influence of the measured environmental variables on the structure of most taxonomic groups was weak to nonexistent (Supplementary data, Table SII).

DISCUSSION

High-throughput sequencing can help to unravel ecological factors that influence microbial biogeography (Leray and Knowlton, 2016). In this study, we evaluated the composition and diversity of planktonic microeukaryotes in floodplain lakes of the Araguaia River, exploring the influence of environmental and spatial factors on shaping these communities at different trophic levels. The floodplain lakes are environmentally heterogeneous and dominated by autotrophic organisms. We detected a spatial pattern in the microorganism distribution and a high beta diversity value. However, contrary to our initial expectation, the environmental variables and spatial filters did not influence the variation in OTU composition for total, autotrophic and heterotrophic microbiotas. We found effects of environmental variables only for mixotrophic microbiota; pH seems to be the most important variable associated with composition, although with a low percentage of explanation.

A high diversity of organisms can be revealed by the use of molecular techniques (see Medinger *et al.*, 2010; Bradford *et al.*, 2013 for comparisons between OTU-based taxonomic classification and traditional morphological studies). However, taxonomic annotation of most OTUs, beyond high-level groups, is not possible as sequences are not yet available in the databases (Richards *et al.*, 2005; Huson *et al.*, 2009). In this study, it was not possible to perform taxonomic annotation for 54% of the data. So far, most of the studies have been conducted in marine ecosystems, and although tropical ecosystems have great biodiversity, they are still rarely studied (Simon

et al., 2015b). As a consequence, most database sequences are from marine organisms. Therefore, due to similarity analyses with rigid criteria, there is a high number of freshwater eukaryotes that do not have corresponding OTUs in the SILVA 18S database. This underlines the importance of conducting further studies on freshwater ecosystems.

Considering the organisms for which it was possible to attribute a taxonomic group, representatives of Chlorophyta, Charophyta and Ciliophora were dominant in read numbers and OTU richness. This result is in agreement with studies carried out in rivers (Bradford *et al.*, 2013) and lakes (Medinger *et al.*, 2010; Simon *et al.*, 2015b; Schiaffino *et al.*, 2016) as well as with a taxonomic evaluation previously conducted in many of these floodplain lakes (e.g. Chlorophyceae; Nabout *et al.*, 2006; Machado *et al.*, 2015). During the flood pulse, organic matter enrichment is expected, increasing the particles in suspension, turbidity and consequent reduction in transparency, which could make the conditions of the environment unfavorable for primary productivity (Junk *et al.*, 1989) and lead to the replacement of photosynthetic organisms by mixotrophs during this period (Gallardo *et al.*, 2012). However, the number of primary producers in our lakes remained high. This probably occurred due to local environmental characteristics, such as the entrance of important tributaries of the Araguaia River, which bring more transparent waters (e.g. Mortes River). In addition, nutrients from flood pulses may intensify the primary productivity of phytoplankton in lakes with an intermediate connection level (Schiemer *et al.*, 2006).

The total OTU richness was lower than those found in a many studies conducted in marine and freshwater environments (Branford *et al.*, 2013, Fonseca *et al.*, 2014, Filker *et al.*, 2015, Genitsaris *et al.*, 2016, Kammerlander *et al.*, 2015, Schiaffino *et al.*, 2016, Das *et al.*, 2019). Most of these studies were performed using a smaller volume of water and focused mainly on microeukaryotes with a size smaller than 68 μm. Thus, the patterns of richness observed in our study may differ from the others only due to the fraction of the plankton size sampled and the greater volume of water used. However, the sampling strategy used here is also useful, as the filtration of a larger volume of water can capture organisms that would not be sampled with a smaller sample volume. The same pattern was observed when we evaluated the Simpson and Shannon indices that showed lower values compared to other studies (Bradford *et al.*, 2013, Gong *et al.*, 2015). This suggests that communities were composed of few abundant OTUs and many less-abundant OTUs. We found a high total beta diversity, which indicates that the OTUs composition is different between the floodplain lakes. This beta diversity was determined mainly by the substitution or replacement of OTUs along the floodplain. This suggests that these environments contain microbiotas that are lake specific (e.g. Schiaffino *et al.*, 2016). This was supported by the ordination pattern of lakes along the plain, in which some OTUs were more closely associated with certain floodplain lakes. In fact, we found a decrease in composition similarity, and geographically close lakes had more similar OTU compositions than distant lakes. A decrease in composition similarity is usually attributed to differences in environmental conditions (which generally increase as the geographic distance increases), landscape characteristics or biological limitations that regulate the dispersal rate of organisms (Soininen *et al.*, 2007). The effects of geographic distance (spatial factors) on communities have constantly been associated with processes that act on a regional scale, such as dispersal capacity (Lindström and Langenheder, 2012). However, in the Araguaia River floodplain, none of these variables could explain the OTU variation in the total microbiota.

In this study, we selected environmental variables that have already been considered important in shaping the planktonic communities in previous studies (Li *et al.*, 2012; Triadó-Margarit and Casamayor, 2012; Bradford *et al.*, 2013; Heino *et al.*, 2014; Simon *et al.*, 2015a; Wang *et al.*, 2015). However, we did not find significant effects of these predictors for the total microbiota, contrary to our initial expectations. In plankton, microbial eukaryotes are represented by taxonomically distinct groups (e.g. amoebas, fungi, ciliates, and primary producers, among others; Pawlowski, 2014), which differ in their morphological, genetic and functional characteristics (Countway *et al.*, 2010). Thus, environmental variables that are important to one group may not always be related to others, for example, depth is associated to Dinophyta, but not to Chlorophyta, fungi or Cercozoa; temperature is associated with Cryptophyta and Choanoflagellida, but not with Chlorophyta, Apicomplexa and Ciliophora; and chlorophyll-a is associated with Apicomplexa but not with Bacillariophyta or Perkinsea (see Gong *et al.*, 2015 and their supplementary material). Here, we found that most taxonomic groups were not associated with measured environmental variables.

Considering trophic groups, the environmental component was weakly important only for mixotrophic microbiota. In fact, environmental variables such as those used in our study have already been identified as important to determine the composition of mixotrophic organisms (e.g. Genitsaris *et al.*, 2016; Saad *et al.*, 2016). However, the absence of an explanation of environmental and spatial components has also been observed for microbiota trophic groups (Khomich *et al.*, 2017). When evaluated alone, only pH was weakly associated with mixotrophic composition. Thus, due to the absence of a purely envi-

ronmental explanation, the relationship of the microbiota to environmental conditions in our lakes remains unclear.

Indeed, some studies have demonstrated the absence of a clear relationship between environmental variables and planktonic microorganisms (Simon *et al.*, 2015b, Genitsaris *et al.*, 2016; Grantepanche *et al.*, 2016), while others show a very low explanation percentage (Heino *et al.*, 2014; Khomich *et al.*, 2017). An absence of environmental and spatial predictors has also been found for phytoplankton in this same region using a traditional taxonomic approach (Nabout *et al.*, 2009). Thus, we believe that other variables that are not strictly limnological or morphometric, such as biotic interactions, can also be considered in future studies involving microbiota (e.g. Charvet *et al.*, 2014; Sullam *et al.*, 2017) or human degradation gradients (e.g. Tolkkinen *et al.*, 2015; Volant *et al.*, 2016). On the other hand, the presence of a shared explanation component and a spatial pattern of decay in the similarity allows us to assume the presence of spatially structured environmental variables acting on the determination of those communities (Borcard *et al.*, 1992), although their significance could not be tested (Legendre and Legendre, 1998).

The spatial patterns of communities can be derived from intrinsic factors of the organism (e.g. migration rate, dispersion capacity, competition, predation), which promote an autocorrelation in the data, or extrinsic factors (interaction with other spatially structured factors, such as environmental characteristics) that generate spatial dependence (Sokal and Oden, 1978; Legendre and Legendre, 1998). We consider that the spatial patterns observed in communities through the Mantel correlogram can be derived from spatial dependence. Thus, closer sites may feature similar environmental conditions and, consequently, similar OTU compositions (Soininen *et al.*, 2007). This fact is corroborated by the existence of a shared component between the environment and space, indicating the existence of spatially structured environmental variables. However, despite adopting spatial variables that represent directional (Borcard and Legendre, 2002) and non-directional (Blanchet *et al.*, 2008a) dispersion processes, we did not find evidence of a purely spatial explanation in the pRDA. This indicates the absence of spatial autocorrelation caused by intrinsic microbiota processes or by spatially structured environmental variables that were not included in the model (Legendre, 1993).

The reduction in the similarity of the community composition according to the increase in geographic distance is not enough to evaluate the effects of dispersion in the context of metacommunities since this reduction can be derived from environmental similarity or geographic distance (Moritz *et al.*, 2013). Thus, it was not possible to attribute a specific paradigm (i.e. patch dynamics, species sorting, mass effect or neutral dynamic) in the metacommunity context for the microeukaryotes evaluated. This indicates that a combination of different factors may be responsible for the decay in similarity (Soininen *et al.*, 2007; Moritz *et al.*, 2013), which is often specific to each area of study (Heino *et al.*, 2015). For the Araguaia River, our results demonstrate a combination of these two factors, with a reduction in environmental similarity according to geographic distance.

Previous studies have shown that environmental variables are important to explain the composition based on the morphological identification of autotrophic organisms at the taxonomic and functional levels (Machado *et al.*, 2016). Our results concerning microplankton demonstrate that this need not apply for composition determined by molecular techniques. A possible explanation for this discrepancy is the absence of a strong environmental gradient. Most lakes are oligotrophic (Marcionilio *et al.*, 2016) and are surrounded by native Cerrado vegetation (Machado *et al.*, 2016). Thus, although the environment explains the taxonomic and functional organism composition, there is no strong environmental pattern restricting certain groups to certain environments.

CONCLUSION

The metabarcoding approach is an efficient tool for the study of planktonic organisms (Hirai *et al.*, 2015) and may produce patterns distinct from the traditional taxonomic approach. In the Araguaia River, using this approach, we characterized the diversity of the larger planktonic microeukaryotes whose composition is spatially structured along the plain. Thus, our study adds information on the planktonic diversity of microeukaryotes, which is still rarely explored in tropical regions. Although we did not find direct effects of the environment and spatial distance on the OTU composition, we observed a considerable value for the shared explanation component, indicating that spatially structured environmental variables may be acting to determine the composition of these communities. Considering that shallow lakes are important reservoirs of eukaryotic diversity (Simon *et al.*, 2015b), this metabarcoding strategy should be combined with traditional taxonomic studies, seeking to predict the factors influencing communities in a more complete and efficient way.

DATA ARCHIVING

The sequences used in this study were deposited in the GenBank's Sequence Reads Archive database

(https://submit.ncbi.nlm.nih.gov/) under the access number SUB3701164: MH022894-MH023194 (Bioproject Accession PRJNA422037 and BioSample Accession SAMN08554969-SAMN08554997).

ACKNOWLEDGEMENTS

We thank Pedro Paulino Borges for the assistance in the construction of the study area map.

FUNDING

Conselho Nacional de Desenvolvimento Científico e Tecnológico and research productivity grant (project nos. 473730/2013-8; 563834/2010-2; 563839/2010-4 to T.N.S., M.P.C.T., V.L.M.H. and J.C.N.); Fundação de Amparo a Pesquisa do Estado de Goiás (FAPEG; project no. 201212267001071); National Institutes for Science and Technology in Ecology, Evolution and Biodiversity Conservation, supported by MCTIC/CNPq (proc. 465610/2014–5) and FAPEG for the PhD scholarship to K.B.M. and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (R.L., as a visiting researcher in the Science without Borders Program).

REFERENCES

Alin, A. (2010) Multicollinearity. *Wiley Interdiscip. Rev. Comp. Stat.*, **2**, 370–374.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) Basic local alignment Seash tool. *J. Mol. Biol.*, **215**, 403–410.

Amaral-Zettler, L. (2013) Eukaryotic diversity at pH extremes. *Front. Microbiol.*, **441**, 1–17.

Andrews, S. (2010) *FastQC: A Quality Control Tool for High Throughput Sequence Data*. http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Aquino, S., Latrubesse, E. M. and Souza-Filho, E. E. (2008) Relações entre o regime hidrológico e os ecossistemas aquáticos da planície aluvial do rio Araguaia. *Acta Sci. Biol. Sci.*, **30**, 361–369.

Baas-Becking, L. G. M. (1934) *Geobiologie of Inleiding tot de Milieukunde*, W.P. Van Stockum & Zoon, Den Haag.

Baselga, A. (2010) Partitioning turnover and nestedness components of beta diversity. *Glob. Ecol. Biogeogr.*, **19**, 134–143.

Baselga, A. and Orme, C. D. L. (2012) Betapart: an R package for the study of beta diversity. *Methods Ecol. Evol.*, **3**, 808–812.

Beakes, G. W. and Thines, M. (2016) Hyphochytriomycota and oomycota. In Archibald, J. M., Simpson, A. G. B. and Slamovits, C. H. (eds.), *Handbook of the Protists*, Springer, Switzerland, pp. 435–505.

Beaver, J. R. and Crisman, T. L. (1989) The role of ciliated protozoa in pelagic freshwater ecosystems. *Microb. Ecol.*, **17**, 111–136.

Bik, H. M., Sung, W., Ley, P. D., Baldwin, J. G., Sharma, J., Rocha-Olivares, C. R. and Thomas, W. K. (2012) Metagenetic community analysis of microbial eukaryotes illuminates biogeographic patterns in deep-sea and shallow water sediments. *Mol. Ecol.*, **21**, 1048–1059.

Blanchet, F. G., Legendre, P. and Borcard, D. (2008a) Modelling directional spatial processes in ecological data. *Ecol. Modell.*, **215**, 325–336.

Blanchet, F. G., Legendre, P. and Borcard, D. (2008b) Forward selection of explanatory variables. *Ecology*, **89**, 2623–2632.

Boenigk, J. and Arndt, H. (2002) Bacterivory by heterotrophic flagellates: community structure and feeding strategies. *Antonie van Leeuwenhoek*, **81**, 465–480.

Bolger, A. M., Lohse, M. and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120.

Borcard, D., Legendre, P. and Drapeau, P. (1992) Partialling out the spatial component of ecological variation. *Ecology*, **73**, 1045–1055.

Borcard, D. and Legendre, P. (2002) All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecol. Modell.*, **153**, 51–68.

Bradford, T. M., Morgan, M. J., Lorenz, Z., Hartley, D. M., Hardy, C. M. and Oliver, R. L. (2013) Microeukaryote community composition assessed by pyrosequencing is associated whit light availability and phytoplankton primary production along a lowland river. *Freshwater Biol.*, **58**, 2401–2413.

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., Fierer, N. and Knight, R. (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *PNAS*, **108**, 4516–4522.

Charvet, S., Vincent, W. F. and Lovejoy, C. (2014) Effects of light and prey availability on the Arctic freshwater protist communities examined by high-throughput DNA and RNA sequencing. *Microbiol. Ecol.*, **88**, 550–564.

Chase, J. M. (2003) Community assembly: when should history matter? *Oecologia*, **136**, 489–498.

Chase, J. M. (2010) Stochastic community assembly causes higher biodiversity in more productive environments. *Science*, **328**, 1388–1391.

Chen, M., Chen, F., Yu, Y., Ji, J. and Kong, F. (2008) Genetic diversity of eukaryotic microorganism in lake Taihu, a large shallow subtropical lake in China. *Microb. Ecol.*, **56**, 572–583.

Cotner, J. B. and Biddanda, B. A. (2002) Small players, large role: microbial influence on biogeochemical process in pelagic aquatic ecosystems. *Ecosystems*, **5**, 105–121.

Countway, P. D., Vigil, P. D., Schnetzer, A., Moorthi, S. D. and Caron, D. A. (2010) Seasonal analysis of protistan community structure and diversity at the USC microbial observatory (San Pedro Channel, North Pacific Ocean). *Limnol. Oceanogr.*, **55**, 2381–2396.

Das, P. B., Gauns, M. and Naqvi, S. W. A. (2019) Ecological diversity of planktonic protists in spatial regimes of the Arabian Sea reveled through next-generation sequence. *Reg. Stud. Mar. Sci.*, **25**, 1–10.

Dick, M. W. (2001) The peronosporomycetes. In McLaughlin, D. J., McLaughlin, E. G. and Lemke, P. A. (eds.), *Systematics and Evolution. The Mycota: A Comprehensive Treatise on Fungi as Experimental Systems for Basic and Applied Research*. **Vol. 7A**. Springer, Berlin, pp. 39 –72.

Dray, S., Legendre, P. and Blanchet, G. (2011) *Packfor: Forward Selection with Permutation (Canoco p. 46)*, (R package version 0.0–8/r100).

Edgar, R. C. (2013) UParse: highly accurate out sequences from microbial amplicon reads. *Nat. Methods*, **10**, 996–998.

Fenchel, T. and Finlay, B. J. (2004) The ubiquity of small species: patterns of local and global diversity. *BioScience*, **54**, 777–784.

Fietz, S., Bleiβ, W., Hepperle, D., Koppitz, L., Krienitz, L. and Nicklisch, A. (2005) First record of *Nannocloropsis limmetica*

(Eustigmatophyceae) autotrophic picoplankton from lake Baikal. *J. Phycol.*, **41**, 780–790.

Filker, S., Gimmler, A., Dunthorn, M., Mahé, F. and Stoeck, T. (2015) Deep sequencing uncovers protistan plankton diversity in the Portuguese Ria Formosa solar saltern ponds. *Extremophiles*, **19**, 283–295.

Finlay, B. J. (2002) Global dispersal of free-living microbial eukaryote species. *Science*, **296**, 1061–1063.

Flynn, K. J., Stoecker, D. K., Mitra, A., Raven, J. A., Glibert, P. M., Hansen, P. J., Granéli, E. and Burkholder, J. M. (2013) Misuse of the phytoplankton-zooplankton dichotomy: the need to assign organisms as mixotrophs within plankton functional types. *J. Plankton Res.*, **35**, 3–11.

Fonseca, V. G., Carvalho, G. R., Nichols, B., Quince, C., Johnson, H. F., Neill, S. P., Lambshead, J. D., Thomas, W. K. *et al.* (2014) Metagenetic analysis of patterns of distribution and diversity of marine meiobenthic eukaryotes. *Glob. Ecol. Biogeogr.*, **23**, 1293–1302.

Fortin, M. J. and Dale, M. (2005) *Spatial Analysis a Guide for Ecologists*, Cambridge University Press, New York.

Gallardo, B., Español, C. and Comin, F. A. (2012) Aquatic metabolism short-term response to the flood pulse in a Mediterranean floodplain. *Hydrobiologia*, **693**, 251–264.

Genitsaris, S., Monchy, S., Breton, E., Lecuyer, E. and Christaki, U. (2016) Small-scale variability of protistan planktonic pressures and biotic interactions at two adjacent coastal stations. *Mar. Ecol. Prog. Ser.*, **548**, 61–75.

Glockling, S. L., Marshall, W. L. and Gleason, F. H. (2013) Phylogenetic interpretations and ecological potentials of the Mesomycetozoa (Ichthyosporea). *Fungal Ecol.*, **6**, 237–247.

Golterman, H. L., Clymo, R. S. and Ohnstad, A. M. (1978) *Methods for Physical and Chemical Analysis of Freshwaters*, Blackwell Scientific Publication, Oxford.

Gong, J., Shi, F., Ma, B., Dong, J., Pachiadaki, M., Zhang, X. and Edgcomb, V. P. (2015) Depth shapes a- and β-diversities of microbial eukaryotes in surficial sediments of costal ecosystems. *Environ. Microbiol.*, **17**, 3722–3737.

Grattepanche, J. D., McManus, G. B. and Katz, L. A. (2016) Patchiness of ciliate communities sampled at varying spatial scales along the New England shelf. *PLoS One*, **11**, 1–16, e0167659. doi: 10.1371/journal.pone.0167659.

Handelsman, J. (2009) Metagenetics: spending our inheritance on the future. *Microb. Biotechnol.*, **2**, 128–156.

Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C. and Martiny, J. B. H. (2012) Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat. Rev. Microbiol.*, **10**, 497–506.

Heino, J., Tolkkinen, M., Pirttil, A. M., Aisala, H. and Mykrä, H. (2014) Microbial diversity and community-environment relationships in boreal streams. *J. Biogeogr.*, **41**, 2234–2244.

Heino, J., Melo, A. S., Siqueira, T., Soininen, J., Valanko, S. and Bini, L. M. (2015) Metacommunity organization, spatial extent and dispersal in aquatic systems: patterns, process and prospects. *Freshwater Biol.*, **60**, 845–869.

Hirai, J., Kuriyama, M., Ichikawa, T., Hidaka, K. and Tsuda, A. (2015) A metagenetic approach for revealing community structure of marine planktonic copepods. *Mol. Ecol. Resour.*, **15**, 68–80.

Hurlbert, S. H. (1971) The nonconcept of species diversity: a critique and alternative parameters. *Ecology*, **52**, 577–586.

Huson, D. H., Richter, D. C., Mitra, S., Auch, A. F. and Schuster, S. C. (2009) Methods for comparative metagenomics. *BMC Informatics*, **10**, S12, doi: 10.1186/1471-2105-10-S1-S12.

Hutchinson, G. E. (1957) Concluding remarks. *Cold Spring Harb. Symp. Quantitat. Biol.*, **22**, 415–427.

Jones, R. I. (2000) Mixotrophy in planktonic protists: an overview. *Freshwater Biol.*, **45**, 219–226.

Junk, W. J., Bayley, P. B. and Sparkes, R. E. (1989) The flood pulse concept in river floodplain systems. In Dodge, D. P. (ed.), *Proceedings of the International Large River Symposium*, Canadian Special Publication of Fisheries and Aquatic Sciences, Illinois, pp. 110–127.

Kammerlander, B., Breiner, H. W., Filker, S., Sommaruga, R., Sonntag, B. and Stoeck, T. (2015) High diversity of protistan plankton communities in remote high mountain lakes in the European Alps and the Himalayan mountains. *Microbiol. Ecol.*, **91**, 1–10.

Keck, F., Vasselon, V., Tapolczai, K., Rimet, F. and Bouchez, A. (2017) Freshwater biomonitoring in the information age. *Front. Ecol. Environ.*, **15**, 266–274.

Khomich, M., Kauserud, H., Logares, R., Rasconi, S. and Andersen, T. (2017) Planktonic protistan communities in lakes along a large-scale environmental gradient. *Microbiol. Ecol.*, **93**, fiw231.

Latrubesse, E. M. and Steuvax, J. C. (2006) Características físico-bióticas e problemas ambientais associados à planície aluvial do rio Araguaia, Brasil central. *Revista UnG-Geociências*, **5**, 65–73.

Legendre, P. (1993) Spatial autocorrelation: trouble or new paradigm? *Ecology*, **74**, 1659–1673.

Legendre, P. and Fortin, M. J. (1989) Spatial pattern and ecological analysis. *Vegetatio*, **80**, 107–138.

Legendre, P. and Legendre, L. (1998) *Numerical Ecology*, Elsevier Science, Amsterdam.

Leibold, M. A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J. M., Hoopes, M. F., Shurin, J. B., Law, R. *et al.* (2004) The metacommunity concept: a framework for multi-scale community ecology. *Ecol. Lett.*, **7**, 601–613.

Leperé, C., Domaizon, I., Taïb, N., Mangot, J. F., Bronner, G., Boucher, D. and Debroas, D. (2013) Geographic distance and ecosystem size determine the distribution of smallest protists in lacustrine ecosystems. *Microbiol. Ecol.*, **85**, 85–94.

Leray, M. and Knowlton, N. (2016) Censusing marine eukaryotic diversity in the twenty-first century. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, **371**, 20150331.

Lesen, A. E., Juhl, A. R. and Anderson, O. R. (2010) Heterotrophic microplankton in the lower Hudson river estuary: potential importance of naked, planktonic amebas for bacterivory and carbon flux. *Aquat. Microb. Ecol.*, **61**, 45–56.

Li, X., Yu, Y., Zhang, T., Feng, W., Ao, H. and Yan, Q. (2012) Seasonal variation of plankton communities influenced by environmental factors in an artificial lake. *Chi. J. Oceanol. Limnol.*, **30**, 397–403.

Lindström, E. S. and Langenheder, S. (2012) Local and regional factors influencing bacterial community assembly. *Environ. Microbiol. Rep.*, **4**, 1–9.

Logares, R., Lindström, E. S., Langenheder, S., Logue, J. B., Paterson, H., Laybourn-Parry, J., Rengefors, K., Tranvik, L. *et al.* (2013) Biogeography of bacterial communities exposed to progressive long-term environmental change. *ISME J.*, **7**, 937–948.

Logares, R., Audic, S., Bass, D., Bittner, L., Boutte, C., Christen, R., Claverie, J. M., Decelle, J. *et al.* (2014) Patterns of rare and abundant marine microbial eukaryotes. *Curr. Biol.*, **24**, 813–821.

Machado, K. B., Borges, P. P., Carneiro, F. M., Santana, J. F., Vieira, L. C. G., Huszar, V. L. M. and Nabout, J. C. (2015) Using lower taxonomic resolution and ecological approaches as a surrogate for plankton species. *Hydrobiologia*, **743**, 255–267.

Machado, K. B., Teresa, F. B., Vieira, L. C. G., Huszar, V. L. M. and Nabout, J. C. (2016) Comparing the effects of landscape and local environmental variables on taxonomic and functional composition of phytoplankton communities. *J. Plankton Res.*, **38**, 1334–1346.

Mangot, J. F., Debroas, D. and Domaizon, I. (2011) Perkinsozoa, a well-know marine protozoan flagellate parasite group, newly identified in lacustrine systems: a review. *Hydrobiologia*, **659**, 37–48.

Marcionilio, S. M. L. O., Machado, K. B., Carneiro, F. M., Ferreira, M. E., Carvalho, P., Vieira, L. C. G., Huszar, V. L. M. and Nabout, J. C. (2016) Environmental factors affecting chlorophyll-a concentration in tropical floodplain lakes, Central Brazil. *Environ. Monit. Assess.*, **188**, 611–620.

Martiny, J. B. H., Bohannan, B. J. M., Brown, J. H., Colwell, R. K., Fuhrman, J. A., Green, J. L., Horner-Devine, M. C., Kane, M. *et al.* (2006) Microbial biogeography: putting microorganism on the map. *Nat. Rev. Microbiol.*, **4**, 102–112.

Medinger, R., Nolte, V., Pandey, R. V., Jost, S., Ottenwälder, B., Schlötterer, C. and Boenigk, J. (2010) Diversity in a hidden world: potential and limitation of next-generation sequencing for surveys of molecular diversity of eukaryotic microorganism. *Mol. Ecol.*, **19**, 32–40.

Morais, R. P., Oliveira, L. G., Latrubesse, E. M. and Pinheiro, D. C. R. (2005) Morfometria de sistemas lacustres da planície aluvial do médio Araguaia. *Acta Sci. Biol. Sci.*, **27**, 203–213.

Moritz, C., Meynard, C. N., Devictor, V., Guizien, K., Labrune, C., Guarini, J. M. and Mouquet, N. (2013) Disentangling the role of connectivity, environmental filtering, and spatial structure on metacommunity dynamics. *Oikos*, **122**, 1401–1410.

Myl'nikova, Z. M. and Myl'nikov, A. P. (2009) The morphology of predatory flagellate *Colpodella pseudoedax*, 2007 (Colpodellida, Alveolata). *Inland Water Biol.*, **2**, 199–204.

Nabout, J. C., Nogueira, I. S. and Oliveira, L. G. (2006) Phytoplankton community of floodplain lakes of the Araguaia river, Brazil, in the rainy and dry seasons. *J. Plankton Res.*, **28**, 181–193.

Nabout, J. C., Siqueira, T., Bini, L. M. and Nogueira, I. S. (2009) No evidence for environmental and spatial processes in structuring phytoplankton communities. *Acta Oecol.*, **35**, 720–726.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P. *et al.* (2016) *Vegan: Community Ecology Package. R Package Version 2.0–6*. http://CRAN.R-project.org/package=vegan (25 November 2017, data last accessed).

Pawlowski, J. (2014) *Protist Evolution and Phylogeny. eLS*, Chichester: John Wiley & Sons Ltd, doi:10.1002/9780470015902.a0001935.pub2.

R Core Team (2016) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. http://www.R-project.org/ (25 November 2017, data last accessed).

Ren, L., Jeppesen, E., He, D., Wang, J., Liboriussen, L., Xing, P. and Wu, Q. L. (2015) pH influences the importance of niche-related and neutral processes in lacustrine bacterioplankton assembly. *Appl. Environ. Microbiol.*, **81**, 3104–3114.

Richards, T. A., Vepritskiy, A. A., Gouliamova, D. E. and Nierzwicki-Bauer, S. A. (2005) The molecular diversity of freshwater picoeukaryotes from an oligotrophic lakes reveals diverse, distinctive and globally dispersed lineages. *Environ. Microbiol.*, **7**, 1413–1425.

Rodríguez-Valera, F. (2004) Environmental genomics, the big picture? *Microbiol. Lett.*, **231**, 153–158.

Saad, J. F., Unrein, F., Tribelli, P. M., López, N. and Iziguirre, I. (2016) Influence of lake trophic conditions on the dominant mixotrophic algal assemblages. *J. Plankton Res.*, **38**, 818–829.

Santoferrara, L. F., Grattepanche, J. D., Katz, L. A. and McManus, G. B. (2015) Patterns and processes in microbial biogeography: do molecules and morphologies give the same answers. *ISME J.*, **10**, 1779–1790.

Schiaffino, M. R., Lara, E., Fernández, L. D., Balagué, V., Singer, D., Seppey, C. C. W., Massana, R. and Izaguirre, I. (2016) Microbial eukaryote communities exhibit robust biogeographical patterns along a gradient of Patagonian and Antarctic lakes. *Environ. Microbiol.*, **18**, 5249–5264.

Schiemer, F., Hein, T. and Peduzzi, P. (2006) Hydrological control of system characteristics of floodplain lakes. *Ecohydrol. Hydrobiol.*, **6**, 7–18.

Shannon, C. E. (1948) A Mathematic Theory of Communication. *Bell Syst. Tech. J.*, **27**, 379–423.

Simon, M., López-García, P., Deschamps, P., Moreira, D., Restoux, G., Bertolino, P. and Jardillier, L. (2015a) Marked seasonality and high spatial variability of protest communities in shallow freshwater systems. *ISME J.*, **9**, 1941–1953.

Simon, M., Jardillier, L., Deschamps, P., Moreira, D., Restoux, G., Bertolino, P. and López-García, P. (2015b) Complex communities of small protest and unexpected occurrence of typical marine lineages in shallow freshwater systems. *Environ. Microbiol.*, **17**, 3610–3627.

Simpson, E. H. (1949) Measurement of diversity. *Nature*, **163**, 688.

Soininen, J., Mcdonald, R. and Hillebrand, H. (2007) The distance decay of similarity in ecological communities. *Ecography*, **30**, 3–12.

Soininen, J., Korhonen, J. J., Karhu, J. and Vetterli, A. (2011) Disentangling the spatial patterns in community composition of prokaryotic and eukaryotic lake plankton. *Limnol. Oceanogr.*, **56**, 508–520.

Sokal, R. R. and Oden, N. L. (1978) Spatial autocorrelation in biology 2. Some biological implications and four applications of evolutionary and ecological interest. *Biol. J. Linn. Soc.*, **10**, 229–249.

Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M. D. M., Breiner, H. W. and Richards, T. A. (2010) Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol. Ecol.*, **19**, 21–31.

Stoecker, D. K. (1999) Mixotrophy among dinoflagellates. *J. Eukaryotic Microbiol.*, **46**, 397–401.

Sullam, K. E., Matthews, B., Aebischer, T., Seehausen, O. and Bürgmann, H. (2017) The effect of top-predator presence and phenotype on aquatic microbial communities. *Ecol. Evol.*, **7**, 1572–1582.

Tikhonenkov, D. V., Kovec, J. J., Mylnikov, A. P., Mikhailov, K. V., Simdyanov, T. G., Aleoshin, V. V. and Keeling, P. J. (2014)

Description of *Coponema vietnamica* sp.n. and *Acavomonas peruviana* n. gen. n. Sp., tho new Alveolate phyla (Colponemidia nom. Nov. and Acavomonidia nom. Nov.) and their contributions to reconstructing the ancestral state of alveolates and eukaryotes. *PLoS One*, **9**, e95467.

Tolkkinen, M., Mykra, H., Annala, M., Markkola, A. M., Vuori, K. M. and Muotka, T. (2015) Multi-stressor impacts on fungal diversity and ecosystem functions in streams: natural vs. anthropogenic stress. *Ecology*, **96**, 672–683.

Triadó-Margarit, X. and Casamayor, E. O. (2012) Genetic diversity of planktonic eukaryotes in high mountain lakes (central Pyreness, Spain). *Environ. Microbiol.*, **14**, 2445–2456.

Valente, C. R., Latrubesse, E. M. and Ferreira, L. G. (2013) Relationships among vegetation, geomorphology and hydrology in the Bananal Island wetlands, Araguaia River basin, Central Brazil. *J. South Am. Earth Sci.*, **46**, 150–160.

Volant, A., Héry, M., Desoeuvre, A., Casiot, C., Morin, G., Bertin, P. N. and Bruneel, O. (2016) Spatial distribution of eukaryotic communities using high-throughput sequencing along pollution gradient in the arsenic-rich creek sediments of Carnoulès mine, France. *Environ. Microbiol.*, **72**, 608–620.

Wang, Y., Liu, L., Chen, H. and Yang, J. (2015) Spatiotemporal dynamics and determinants of planktonic bacterial and microeukaryotic communities in a Chinese subtropical river. *Appl. Microbiol. Biotechnol.*, **99**, 9255–9266.

Winegardner, A. K., Jones, B. K., Ng, I. S. Y., Siqueira, T. and Cottenie, K. (2012) The terminology of metacommunity ecology. *Trends Ecol. Evol.*, **27**, 253–254.

Zagatto, E. A. G., Jacintho, A. O., Reis, B. F., Krug, F. J., Bergamin, H., Pessanda, L. C., Mortatti, J. and Gine, F. (1981) *Manual de Análises de Plantas Empregando Sistemas de Injeção em Fluxo*, Universidade de São Paulo, Piracicaba.

Zhang, W., Pan, Y., Yu, L. and Liu, L. (2017) Genetic diversity patterns of microeukaryotic plankton communities in Shenhu Bay, southeast China. *Cont. Shelf Res.*, **141**, 68–75.