CrossMark

ORIGINAL PAPER

Within versus between-lake variability of sedimentary diatoms: the role of sampling effort in capturing assemblage composition in environmentally heterogeneous shallow lakes

Gabriela S. Hassan

Abstract The effect of within-lake diatom assemblages variability on sample representativity and its subsequent impact on between-lake comparisons were addressed in three environmentally heterogeneous shallow lakes from the Argentinean Pampas. Surface sediment samples were collected from the open waters and the highly vegetated littoral areas on a seasonal basis and analyzed for diatom assemblages composition. Within-lake variability was assessed by comparing the Bray Curtis distances between original data and the Monte Carlo-simulated average assemblages composition through non-metric multidimensional scaling (NMDS). Diatom assemblages showed a high variability in composition, evidencing large dispersions of samples around the centroid in NMDS plots. Permutational multivariate analysis of variance tests signaled significant differences in average composition between the three lakes, related mainly to their differences in conductivity and depth. Representativity of original samples was assessed through principal coordinates analyses ordinations of the three lakes, being samples lying in the overlapping areas of the plot classified as poor representatives of between-lake differences. Several samples, both from littoral and open waters, were classified as poor representatives through this method. Simulation allowed us to evaluate the effect of sample replication on improving between-lake comparisons, and showed that collecting two littoral and two open-water samples allowed us to faithfully capture differences in average composition among the three lakes. Hence, the results suggest that using a single sample to estimate diatom assemblages composition in these lakes should be avoided, as it fails to capture between-lake differences, leading to biases in compositional comparisons among lakes and regions. Consequently, including multiple samples from each lake when constructing calibration sets would be the best option to obtain reliable paleoenvironmental reconstructions from single sediment cores in these environmentally heterogeneous shallow lakes.

G. S. Hassan (✉)
Instituto de Investigaciones Marinas y Costeras (IIMyC), Anexo Facultad de Ciencias Exactas y Naturales, CONICET-Universidad Nacional de Mar del Plata, Juan B Justo 2550, Mar del Plata, Argentina
e-mail: ghassan@mdp.edu.ar

🎺 Springer

Introduction

The study of modern species-environment relationships as a key to understand past environmental changes has become a central tenet in paleolimnology. The use of biological indicators to infer paleoenvironmental conditions has a long history, and evolved from qualitative studies based on a few indicator taxa to the calibration of thousands of taxa to a multitude of environmental variables (Smol 1992; Birks 1998, 2012; Juggins and Birks 2012). Such calibration data sets revolutionized paleolimnology by providing the means of deriving quantitative reconstructions of past environmental variables from the fossil assemblages preserved in lake sediments (Birks 1998).

Statistical calibration models (i.e. transfer functions) are based on surface-sediment samples collected from a large number of lakes covering the environmental gradient of interest for the reconstruction. In each lake, a single surface-sediment sample is usually collected from the center or deepest zone of the lake basin and death assemblages from the topmost sediments analyzed (Juggins and Birks 2012). In doing so, single death assemblages are considered to represent a good estimation of the species composition of one lake (Adler and Hübener 2007). This is a rough approximation, however, as practical considerations and resource limitations usually do not allow an intensive sampling of each lake along spatial and temporal scales. Nevertheless, this simplified methodology only works when the between-lake variability on diatom assemblages composition exceeds the internal variability within any of the lakes of the dataset (Link et al. 1994; Bunbury and Gajewski 2008). If the average composition of diatom assemblages significantly differs between two lakes of contrasting environmental characteristics, and no overlap among their multivariate dispersions is evident, then single samples can be faithfully used for environmental bioindication, provided the environmental variable of interest explains a significant portion of the diatom variance. If, on the other hand, diatom composition variability within these lakes is larger than its between-lake variance, then the value of single samples to represent environmental differences among lakes is highly dependent of the sampling design. In such cases, the role of sampling effort in capturing environmental differences becomes a key issue for diatom-based paleoenvironmental studies.

The distribution of death assemblages in lake sediments is a complex process, regulated not only by the ecology and distribution of the original living communities but also by taphonomic factors (Birks and Birks 1980). The point of deposition of sub-fossils and their possible resuspension and sedimentation can be influenced by lake morphometry, aquatic vegetation, waves, wind intensity and direction (Blais and Kalff 1995; Whitmore et al. 1996; Gilbert and Lamoureux 2004; Heggen et al. 2012). The source community and organism life-form are also critical issues, as floating remains (such as plankton) may tend to be easily reworked and redistributed through the lakes sub-environments, whereas benthic and periphytic taxa may be more easily sunk and incorporated into the sediments near their source areas (De Nicola 1986; Frey 1988; Heggen et al. 2012; Hassan 2015). Overall, all these factors influence the degree of variability exhibited by death assemblages, leading to unpredictable differences among lakes and proxies. This problem has been recognized both in biomonitoring and paleolimnology, and a series of studies addressing the impact of spatial variability of death assemblages on sample representativity have been published in the last decades (Anderson 1990; Link et al. 1994; Weilhoefer and Pan 2006; Eggermont et al. 2007; Bunbury and Gajewski 2008; Dennis et al. 2010; Heggen et al. 2012; Bennett et al. 2014).

Although the importance of diatom sample representativity has been addressed for deep, dimictic lakes (Charles et al. 1991; Adler and Hübener 2007) and rivers (Weilhoefer and Pan 2006), the extent of this problem in shallow lakes has received little attention (Adler and Hübener 2007). This is particularly true for the shallow, hypereutrophic and environmentally heterogeneous lakes characterizing the South American Pampas (Diovisalvi et al. 2015; García-Rodríguez et al. 2009). These lakes are characterized by an extensive development of the littoral zone relative to the pelagic zone, providing a wide suite of available microhabitats that promote the growth of diverse and productive periphytic and benthic communities (Wetzel 2001), and exerting differential taphonomic biases during diatom deposition (Hassan 2015). This environmental complexity originates an unpredictable variability on death diatom assemblages distribution in these lakes, leading to uncertainties on the sampling effort needed to faithfully represent within-lake composition. In this paper, the degree of

variability in sedimentary diatom assemblages composition is addressed in three Pampean shallow lakes exhibiting differences in their environmental characteristics, particularly in conductivity and hardness (Cristini et al. 2017). Previous studies evidenced significant relationships between diatom assemblages composition, conductivity, pH, and depth in these lakes, encouraging their use as indicators of past environmental fluctuations (Hassan 2015). Therefore, they bring an opportunity to assess the effect of within-lake variability on sampling representativity and its subsequent impact on between-lake comparisons. The results are expected to provide insights on the role of diatom variability on the formation of sedimentary assemblages in shallow lakes and to discuss the implications for diatom-based calibration and paleoenvironmental studies.

This study is part of a larger project focused on addressing the nature and magnitude of the taphonomic biases suffered by diatom remains in shallow lakes. In a previous work, the compositional and environmental fidelity exhibited by diatom death assemblages was evaluated through the multivariate analysis of live-dead data sets collected in three lakes from the Southern Pampas (Hassan 2015). Death assemblages captured environmental gradients better than life assemblages, being conductivity and depth the strongest environmental variables in explaining significant portions of diatom variance (12.26 and 10.66%, respectively; $p < 0.001$, Hassan 2015). In the present contribution, the extent to which single sedimentary samples do represent the within-lake average diatom composition is evaluated in the same three shallow lakes.

Study site

The lakes are located in the southeastern Pampa plain (Buenos Aires province, Argentina), close to the Atlantic coast: (1) Las Mostazas (37°9′57″S; 57°14′50″W), (2) Los Carpinchos (37°3′34″S; 57°19′56″W) and (3) Nahuel Rucá (37°37′21″S; 57°25′42″W). They are situated along a north-south gradient covering a distance of approximately 60 km (Fig. 1). As with most Pampean lakes, the studied sites are very shallow (max. depth = 90–150 cm; Table 1), showing considerable inter-annual differences in area and depth between wet and dry periods (Bohn et al. 2011). Their shallowness favors the interaction

between sediments and the water column through wind-driven turbulence. The dominant wind direction in the region is South-Southeast, making the fetch of Las Mostazas (1.4 km) higher than Nahuel Rucá (1.15) and Los Carpinchos (0.65; González-Sagrario et al. 2018). As most lakes occurring in the continental plain in the southern part of the Pampean region, they show typical pfanne or wanne profiles and tend to have rounded or slightly elongated shapes (Diovisalvi et al. 2015). Their littoral zones are densely populated by emergent and submersed macrophytes, forming a ring that usually surrounds entirely the lakes and clearly differentiates littoral from open waters areas (Fig. 1). When compared against other lakes worldwide, Pampean shallow lakes show limnological characteristics that depart from most of those located in temperate regions as having higher phosphorous, nitrogen and *chlorophyll a* concentrations and much lower transparency, and therefore they stand as extremes of the trophic-state continuum (Diovisalvi et al. 2015). This is the case of the studied lakes, which exhibited very eutrophic, turbid and alkaline conditions during the sampling period (Table 1). According to a recent study (Cristini et al. 2017), the three lakes differ significantly in their ionic composition, with Las Mostazas characterized by higher concentrations of chloride, carbonates, bicarbonates, calcium carbonate, calcium and conductivity, as it is a closed basin settled in a saline soil. Nahuel Rucá, on the other hand, is the most stable lake, as its water level is regulated by a floodgate. It exhibits the lowest conductivities and pH among the three studied lakes (Table 1). The sediments from the three lakes are muddy and organic, with maximum organic matter contents ranging between 30 and 60% (Cristini et al. 2017).

Materials and methods

Sampling strategy

In each lake, sampling was carried out on a seasonal basis during a year. Given the very low water levels, Las Mostazas lake became very muddy during the summer, making it impossible to access the sampling area. This was mostly consequence of the settlement of suspended sediments in its extensive and very shallow littoral area during the dry season. Hence, the Las Mostazas data set included data from only 3 field
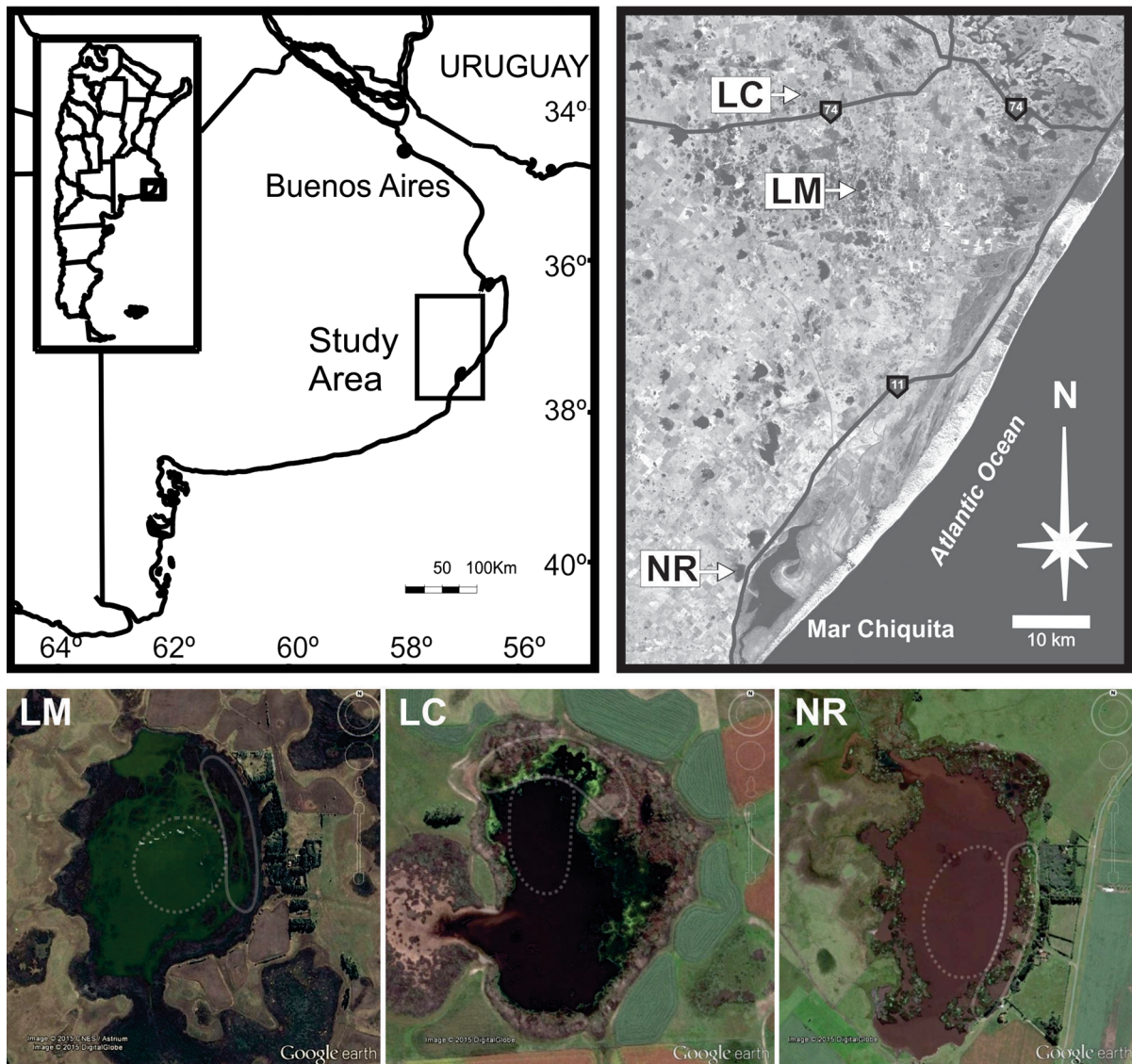
Fig. 1 Location of the studied lakes: Nahuel Rucá (NR), Los Carpinchos (LC) and Las Mostazas (LM). Lines indicate the zones of the open waters (dotted line) and littoral areas (continuous line) covered in the samplings

trips. The sampling was designed in order to allow the comparison between the two main habitats represented in each lake: the highly vegetated littoral zone and the open-water area (Fig. 1). The differentiation between littoral and open-water samples was based on the presence or absence of emerging macrophytes on the sampling point. Littoral samples were collected in the surroundings of *Schoenoplectus californicus* (Meyer) Sojak mats, which were absent of the zone classified as open waters. Zones with well-established mats were selected for sampling, in order to avoid the

regions of the lake that could be subjected to occasional desiccation. The percentage of each lake area covered by macrophytes was assessed from Google Earth® satellite images using klm polygons. The proportion of littoral area ranged between 50.2 and 55.3% (Table 1), supporting previous works on Nahuel Rucá and other Pampean lakes, in which macrophyte coverages were always around 50% (Federman 2003). Consequently, the sampling was designed in order to cover both littoral and open-water areas in similar proportions (50/50). In each visit, six

Table 1 Summary of environmental information taken during diatom samplings by Hassan (2015)

Environmental variables	Sampling Site		
	LM	LC	NR
Water chemistry			
Conductivity (mS/cm)	4.1 ± 1.5 (2.5–6.5)	2.2 ± 1.2 (1.3–4.3)	0.8 ± 0.5 (0.3–1.6)
pH	9.5 ± 0.5 (8.9–10.3)	8.7 ± 0.8 (7.7–10.0)	8.1 ± 0.3 (7.4–8.5)
Dissolved oxygen (ppm)	8.8 ± 4.4 (2.7–14.6)	5.6 ± 4.0 (0.1–9.3)	7.5 ± 2.7 (2.1–10.8)
Temperature (°C)	18.2 ± 4.1 (14.3–24.5)	18.5 ± 7.8 (10.2–32.5)	16.3 ± 7.3 (7.2–25.5)
CO_3^{2-} (mg L^{-1})	174 ± 145 (51–333)	46 ± 92 (0–183)	0
HCO_3^- (mg L^{-1})	664 ± 87 (569–740)	743 ± 135 (583–894)	529 ± 230 (303–849)
Cl^- (mg L^{-1})	822 ± 120 (750–961)	877 ± 1053 (270–2450)	144 ± 82 (91–266)
Hardness (mg L^{-1} of $CaCO_3$).	286 ± 83 (190–334)	128 ± 92 (13–229)	250 ± 127 (173–438)
Ca^{2+} (mg L^{-1})	60.4 ± 10.9 (54.1–73.0)	16.3 ± 11.5 (1.0–27.5)	30.7 ± 16.5 (18.1–53.2)
Mg^{2+} (mg L^{-1})	32.3 ± 17.5 (13.1–47.5)	20.9 ± 16.9 (2.4–38.5)	41.7 ± 29.7 (18.8–85.4)
SO_4^{2-} (mg L^{-1})	188 ± 34 (153–220)	90 ± 84 (1–188)	108.9 ± 61.5 (27.4–169)
SiO_2 (mg L^{-1})	18.7 ± 11.9 (9.7–32.3)	52.1 ± 48.3 (9.4–111.0)	17.7 ± 1.3 (16.8–18.6)
PO_4^{3-} (mg/L)	40.7 ± 8.9 (35–51)	18.1 ± 3.3 (15.8–20.4)	161.3 ± 88.7 (98.6–224.0)
NO_3^- (mg/L)	146.5 ± 30.4 (125–168)	45.4 ± 8.4 (40.0–55.1)	67*
Physical characteristics			
Surface (square kilometers)	1.61	0.53	2.45
Depth (Littoral samples, cm)	73.3 ± 13.8 (50–100)	57.4 ± 23.0 (20–90)	40 ± 14 (15–65)
Depth (Open waters samples, cm)	108.9 ± 13.9 (90–130)	92.7 ± 35.6 (40–150)	78.7 ± 8.8 (60–90)
Macrophyte coverage (%)	55.3	54.7	50.2
Secchi depth (cm)	28.6 ± 6.4 (20–40)	35.6 ± 21.2 (10–70)	16.7 ± 9.6 (5–40)

Values are mean ± SD of seasonal measurements; minimum and maximum values are given in parentheses. Summer conductivity and pH data for Las Mostazas lake was taken from Cristini et al. (2017). The remaining water chemistry data of Las Mostazas does not include summer measurements. *LM* Las Mostazas; *LC* Los Carpinchos; *NR* Nahuel Rucá. *Only one measurement available

surface sediment samples were randomly collected both from the littoral and open-water areas using a piston core. Hence, the final data set comprised a total of 132 surface sediment samples (Los Carpinchos: 48; Nahuel Rucá: 48; Las Mostazas: 36). The top 1 cm of each core was sliced in the field and preserved with 96% alcohol.

In the laboratory, the fresh samples stored in alcohol were oxidized with 30% hydrogen peroxide and washed several times with distilled water. After homogenization, a subsample was transferred to a coverslip and air dried, and permanent slides were made using Naphrax®. On each slide, at least 500 diatom valves were counted across random transects using a Leica DM500 light microscope at X1000 magnification. Diatom taxa were identified according to Lange-Bertalot et al. (1996), Krammer and Lange-Bertalot (1997, 1999, 2004a, b), Metzeltin and Lange-

Bertalot (1998, 2007), Rumrich et al. (2000), Metzeltin et al. (2005), Levkov (2009) and Sar et al. (2009).

Data analyses

Within-lake variability in diatom composition

In order to assess the variability on within-lake diatom assemblages composition, surface-sediment samples from each lake were compared with Monte Carlo-simulated assemblages from the same species pool (Heggen et al. 2012). All diatom valves counted in each lake were pooled and randomly resampled 1000 times using Monte Carlo, with a minimum sample size of 500 valves. These randomized data sets were then averaged in order to calculate a simulated mean assemblage composition for each lake. The models

were performed using the sample function in R version 3.3.2 (R Development Core Team 2016). The ecological distances between the original samples and the simulated mean assemblage were assessed through Bray Curtis (Bray and Curtis 1957) as dissimilarity measurement. NMDS (Non-metric Multidimensional Scaling) based on Bray Curtis distances were used to summarize the multivariate data of each lake on a bidimensional ordination showing the variability of littoral and open waters samples from each season around the simulated mean assemblage. Permutational multivariate analysis of variance (PERMANOVA) tests were applied to search for differences between littoral and open water samples inside each lake, and the diatom taxa determining differences among both lake compartments were identified through Similarity Percentage (SIMPER) routines. This method computes the average dissimilarity between inter-group samples, and then assesses the separate contributions to this dissimilarity for each species (Clarke 1993).

Spatial differences in within-lake variability were also assessed through box-plots comparing the Bray Curtis distances to simulated mean assemblage for littoral and open waters samples separately. Additionally, temporal differences in within-lake variability were assessed through box-plots comparing the Bray Curtis distances among seasons for the three lakes. Significant differences in median values among littoral and open waters and among the four seasons were tested using Kruskal–Wallis tests followed by Mann–Whitney post-hoc comparisons. Differences in the variance between littoral and open waters were tested using F tests (Zar 2010). All the indices, analyses and graphs were performed using the software Past v. 3.11 (Hammer et al. 2001).

Between-lake differences in diatom composition

Relative abundances, richness, Shannon-Wiener diversity, and Simpson (1-D) evenness were used for comparison of different aspects of diatom assemblages between the three studied lakes (Magurran 2004). Significant differences in the median values among these indices were tested using Kruskal–Wallis analysis of variance followed by Mann–Whitney post-hoc tests (Zar 2010). In order to analyze the interaction among within and between-lake variability in the studied shallow lakes, the datasets from the three lakes were ordered in a bidimensional space using principal

coordinates analysis (PCoA) with Bray Curtis as dissimilarity measure. Significant differences between the centroids of the multivariate areas covered by the three lakes were tested with PERMANOVA, and interpreted as indicative of differences in the mean composition of diatom assemblages among lakes. The diatom taxa determining differences among the three lakes were identified through Similarity Percentage (SIMPER) routines. Kruskal–Wallis, PERMANOVA and SIMPER tests were performed using the software Past v. 3.11 (Hammer et al. 2001).

A test of homogeneity of multivariate dispersions (HMD, Anderson et al. 2006) was also applied in order to analyze for differences in the variances between the three lakes. Dispersions were represented by the Bray Curtis distances of samples to their centroid in a Euclidean space, calculated with PCoA. If two lakes differed in the location of their multivariate centroids, the degree of overlapping among the areas covered by these lakes were then considered as indicative of the probability of collecting a sample poorly representative of that between-lake difference. Here, the poor representative samples were defined as those single diatom assemblages lying in a portion of the multivariate diagram in which two or more lake polygons overlap. HMD and PCoA were conducted in R version 3.3.2 (R Development Core Team 2016). Dissimilarity matrices were constructed based on the Bray Curtis index and tested under 999 permutations using the betadisper function in the package "vegan" version 2.0–7 (Oksanen et al. 2013).

Effect of replicated sampling on representativity

Because of significant differences among between-lake average assemblages composition significant according to PERMANOVA results, a second modelling approach was applied to the original datasets from the three lakes in order to assess the minimum number of replicated samples needed to faithfully capture that difference. Monte Carlo resamplings were applied to the original data sets by randomly selecting n samples 100 times, being n pair and ranging between 2 and 20 samples. For each run, resamplings were constrained to choose equal numbers of samples from the littoral and open waters samples using the ddply function in the package "plyr" version 1.8.4 (Wickham 2016).

Fig. 2 Analyses of variability of diatom composition around ▶ the simulated average assemblage composition. **a** Boxplots showing Bray–Curtis distances to the simulated mean assemblage of samples from Las Mostazas (LM), Los Carpinchos (LC) and Nahuel Rucá (NR); B-D NMDS plots based on Bray–Curtis distances of the original datasets from littoral (circles) and open waters (squares) and the average composition of the modeled dataset (white cross) in Las Mostazas (**b**), Los Carpinchos (**c**) and Nahuel Rucá (**d**)

The capability of the simulated groups of replicates to capture the average composition of diatom assemblages within each lake was then evaluated by comparing the Bray Curtis distances of their centroid to the simulated mean assemblage. Those replicated mean assemblages lying within a Bray Curtis distance threshold of 0.3 from the simulated mean assemblage were considered as good representatives of the within-lake average composition (Kelly 2001). In order to evaluate the impact of increasing the number of replicates on between-lake comparisons, PCoA ordinations of the three lakes were repeated based on the models constructed with 2 and 4 replicated samples previously developed. The minimum number of replicates representative of between-lake differences was defined as the one which lead to no multivariate overlap between the modeled mean abundances of the three lakes. The models, analyses and graphs were prepared using R version 3.3.2 (R Development Core Team 2016).

Results

Within-lake variability in diatom composition

A total of 146 diatom species were identified in the three lakes (ESM1). The distribution of the dominant taxa in littoral and open waters sediments is presented in ESM 2–4. Assemblages in Las Mostazas were dominated by the tychoplanktonic *Cyclotella meneghiniana* Kützing, accompanied by epiphytic (*Cocconeis placentula* Ehrenberg, *Epithemia adnata* (Kützing) Brébisson and *E. sorex* Kützing) and benthic taxa (*Nitzschia amphibia* Grunow, *Hippodonta hungarica* (Grunow) Lange-Bertalot, Metzeltin and Witkowski and *Navicula peregrina* (Ehrenberg) Kützing). Differences in assemblages composition between littoral and open waters were significant (F: 11.89;
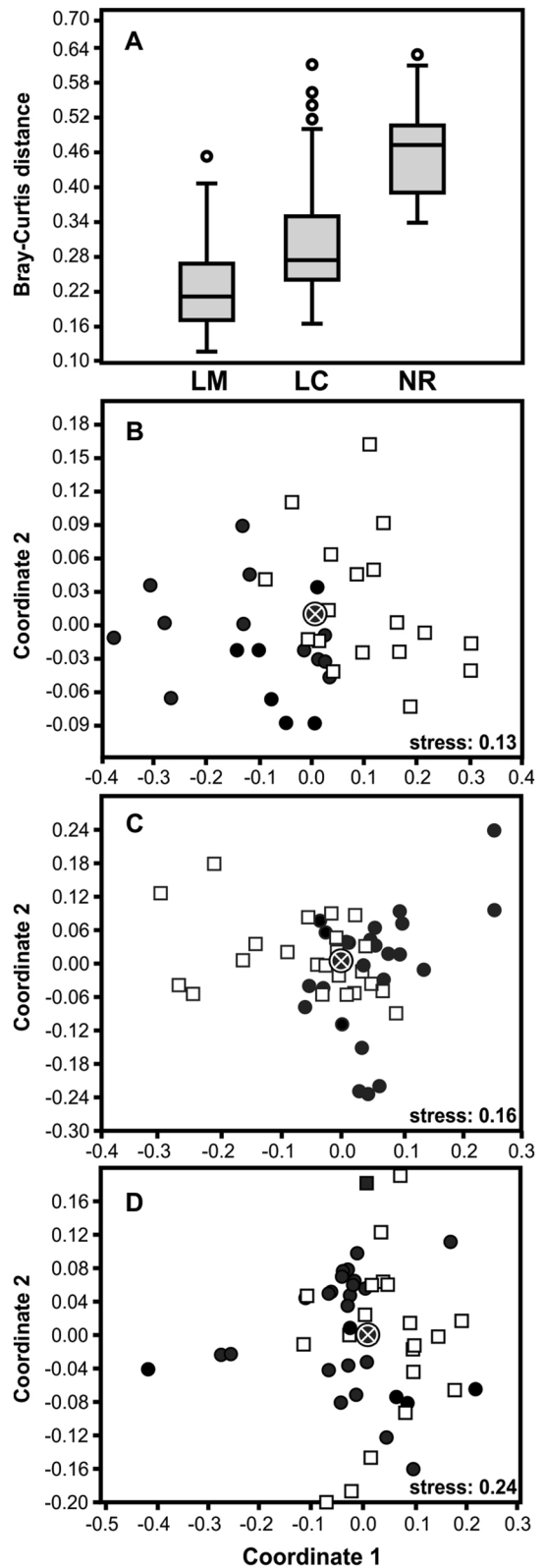
$p < 0.0001$), and mainly related to a higher percentage of *N. amphibia* (SIMPER contribution: 18.4%) in littoral sediments and a dominance of *C. meneghiniana* (19.37%) and *N. peregrina* (6.5%) in open waters (ESM2). This was also evidenced by the NMDS plot, in which a clear separation between littoral and open-water samples was observed (Fig. 2b). Bray Curtis similarity to within-lake simulated mean assemblage composition ranged between 0.54 and 0.87 (Fig. 2a). Littoral and open-water samples did not differ in the mean Bray Curtis similarities of samples to the simulated mean assemblage (H: 0.09, $p = 0.9$) nor in its variance (F = 1.74, $p = 0.3$; ESM5). There were no significant differences in the Bray Curtis similarities of samples among seasons in this lake (H: 0.57, $p = 0.75$; ESM5).

Diatom assemblages in Los Carpinchos Lake were also dominated by *Cyclotella meneghiniana*, accompanied by tychoplanctonic (*Aulacoseira granulata* (Ehrenberg) Simonsen and *A. granulata* var. *angustissima* (Müller) Simonsen), epiphytic (*Cocconeis placentula*, *Epithemia adnata*, *E. sorex*, *Pseudostaurosira americana* Morales and *Planothidium delicatulum* (Kützing) Round and Bukhtiyarova) and benthic taxa (*Nitzschia amphibia*, *Hippodonta hungarica*, *Navicula veneta* Kützing and *N. peregrina*). Diatom assemblage composition differed significantly between littoral and open-water compartments of the lake (F: 4.46; $p < 0.001$), as a consequence of the higher dominance of *C. meneghiniana* (SIMPER contribution: 14.3%), *C. placentula* (7.9%), *N. amphibia* (6.5%), *A. granulata* (5.5%), and *N. peregrina* (5%) in open waters (ESM3). Ordination of samples in the NMDS plot also evidenced a partial separation between littoral and open-water samples, although they show similar dispersions around the simulated mean assemblage (Fig. 2c). Accordingly, littoral samples did not differ from open-water ones in their Bray Curtis similarities to the simulated mean assemblage (H = 2.3, $p = 0.12$) nor in its variance (F = 1.05, $p = 0.92$; ESM5). Also, no significant differences in the Bray Curtis similarities among seasons were found in this lake (H = 2.3, $p = 0.5$; ESM5).

In Nahuel Rucá Lake, assemblages composition differed significantly between littoral and open waters (F: 3.56; $p < 0.001$), as the tychoplanktonic taxa *C. meneghiniana*, *A. granulata* and *A. granulata* var. *angustissima* were more abundant in open waters than

in littoral sediments (SIMPER cumulative contribution: 25.5%). The benthic taxa *N. amphibia*, *N. veneta* and *H. hungarica*, together with the epiphytics *E. adnata*, *E. sorex* and *P. americana*, exhibited their higher relative abundances in littoral sediments (SIMPER cumulative contribution: 24.8%). The compositional dissimilarity between littoral and open-water samples was not evident in the NMDS plot, where samples from both lake subenvironments occupied mostly overlapping multivariate spaces (Fig. 2d). No significant differences between the Bray Curtis distances (H = 1.74; $p = 0.18$) or the variances (F = 2.07; $p = 0.31$) of both sets of samples to the simulated mean assemblage were recorded (ESM5). However, Bray Curtis similarities to the simulated mean assemblage were significantly higher in autumn in this lake (H = 10.21; $p < 0.05$; ESM5).

Between-lake differences in diatom composition

Analyses of diversity indices indicated that Las Mostazas was the least diverse of the three studied lakes, as it exhibited significantly lower richness (R = 35 ± 7), Shannon diversity (H′ = 2.31 ± 0.25) and evenness values (E = 0.83 ± 0.04) than Los Carpinchos and Nahuel Rucá ($p < 0.0001$; ESM6). On the other hand, no significant differences among these two lakes were observed, as assemblages from Los Carpinchos and Nahuel Rucá exhibited similar values of richness (-Los Carpinchos = 54 ± 10, Nahuel Rucá = 57 ± 10), Shannon diversity (Los Carpinchos = 2.95 ± 0.43, Nahuel Rucá = 3.07 ± 0.42) and evenness (Los Carpinchos = 0.89 ± 0.08, Nahuel Rucá = 0.90 ± 0.08; ESM6).

According to the HMD tests, variance in Las Mostazas assemblages (D = 0.2339) was significantly lower than in Nahuel Rucá (D = 0.3348; $p < 0.0001$) and Los Carpinchos (D = 0.3121; $p < 0.0001$), while these two lakes did not show significant differences among them. PERMANOVA tests indicated that mean assemblages composition differed significantly among the three lakes (F = 24; $p < 0.0001$). According to SIMPER results, these differences were a consequence of both (1) differences in the relative abundances of the dominant diatom taxa (e.g. *C. meneghiniana*, *C. placentula* and *N. amphibia*; ESM7), and (2) the presence of some taxa exclusively represented in one lake (e.g. *Staurosira longirostris* in Nahuel Rucá, ESM7).

Examination of PCoA biplot, however, evidenced a strong degree of overlap between the multivariate spaces occupied by these lakes (Fig. 4a). The percentage of samples ordered in the multivariate space exclusive of each dataset varied among lakes: a 61% of Las Mostazas (n = 22; 11 littoral, 11 open waters), a 29% of Los Carpinchos (n = 14; 7 littoral, 7 open waters), and a 50% of Nahuel Rucá samples (n = 24; 5 littoral, 19 open waters), lied out of the overlapping area of the multivariate space and can be considered as fully representative of the assemblages from each lake (Fig. 4a). A nearly equivalent distribution of representative samples across seasons was found in Las Mostazas (Spring: 9, Autumn: 6, Winter: 7), and Nahuel Rucá (Spring: 6, Summer: 6, Autumn: 6, Winter: 6) Lakes, while in Los Carpinchos these were mostly collected in winter (Spring: 2, Summer: 1, Autumn: 4, Winter: 7). The degree of overlapping between the two lakes representing the two extreme conductivity conditions (the brackish Las Mostazas and the freshwater Nahuel Rucá, Table 1) was almost negligible, being only 1 open waters sample from Las Mostazas autumn ordered inside the multivariate space occupied by Nahuel Rucá samples (Fig. 4a).

Effect of replicated sampling on representativity

Monte Carlo resampling allowed identifying a strong influence of increasing the number of replicated samples on the representativity of diatom assemblages in the three lakes (Fig. 3). Average composition of the simulated groups of replicates decreased their distance to the simulated mean assemblage as the number of replicates increased. Collecting two replicates (one from littoral and one from open waters) lead to Bray Curtis distances to the centroid lower than 0.3 in Las Mostazas, while four replicated samples were needed to reach this threshold in Los Carpinchos and Nahuel Rucá (Fig. 3). Overall, collecting two samples from each subenvironment led to replicated groups of samples whose average composition closely resembled the average composition of the whole dataset from each lake, suggesting that this sampling intensity would be enough to capture a representative sample of within-lake mean assemblages composition.

The degree of overlapping among the three lakes decreased significantly when the models based on 2 replicated random samples were used in the PCoA ordination (Fig. 4b). The multivariate space occupied
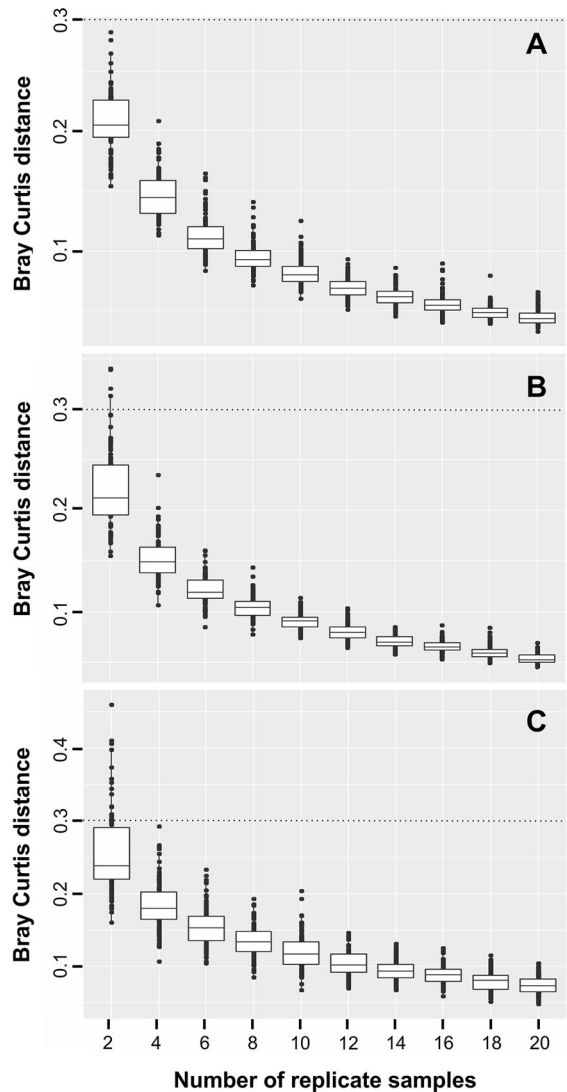
Fig. 3 Plots used to assess the effect of replicated sampling on representativity. Through Monte-Carlo simulations, simulated resamplings of 1–20 samples were generated, being half of the samples chosen from littoral and open waters, respectively. Their Bray–Curtis distances of the replicated groups of samples to the modeled average composition of diatom assemblages calculated. A critical distance to the centroid of 0.3 was used to separate poor (dotted line) and good representative samples

by Las Mostazas lake showed no overlapping with Nahuel Rucá, while only one of the samples from Los Carpinchos lied on the overlapping area. The polygons of Los Carpinchos and Nahuel Rucá lakes still showed a high overlapping, although this became strongly reduced when compared to the original dataset (Fig. 4a). Hence, although lower, there is still a
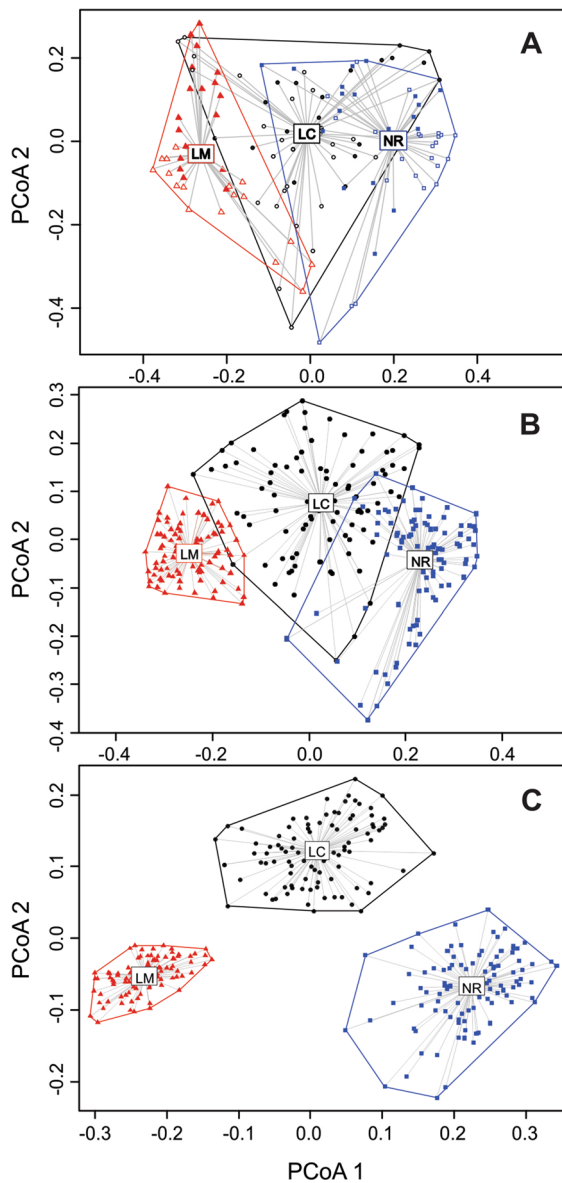
Fig. 4 PCoA plots used to assess the effect of representativity on between-lake comparisons. Ordinations of Las Mostazas (triangles), Los Carpinchos (circles) and Nahuel rucá (squares) lakes were performed based on the original datasets (**a**), modeled pools of 2 samples (**b**) and modeled pools of 4 samples (**c**). Littoral (filled symbols) and open waters (empty symbols) are indicated for the original dataset plot (**a**)

chance of collecting poor representative groups of samples when only one littoral and one open-water sample are collected in these lakes. No overlap among the multivariate spaces occupied by the three lakes was observed when 4 simulated replicates were used in the PCoA (Fig. 4c). Hence, the results indicate that

the probability of obtaining faithful representatives of average assemblages composition when collecting 4 samples (2 littoral and 2 open waters) is high, as average replicate composition of sampling groups is located in a narrow area around the centroid of the diatom assemblages distribution of each lake.

Discussion

Within-lake variability on diatom assemblages composition

Diatom assemblages preserved in surface sediments from Pampean shallow lakes were characterized by a high variability in species composition. This was evidenced by large dispersions of samples around the centroid in NMDS plots, related to large Bray Curtis distances among some samples and the modelled mean abundance composition. Examination of SIMPER results pointed the patchy distribution of some taxa as the main cause of this high variability, as evidenced by the differences on proportional abundances of dominant species. Heterogeneity of diatom communities on scales of centimeters or greater has been associated to changes in light and current regimes, grazing, successional stages and variation in substratum type and composition (Stevenson et al. 1985; Ledger and Hildrew 1998; Sommer 2000; Soininen 2003). Small- scale biotic interactions, such as competition, grazing, colonization processes, variations in recruitment, or low movement ability also affected composition of assemblages both in marine and freshwater habitats (Černá 2010). Physical and chemical parameters responsible of small-scale patchiness of microhabitats, substrate complexity and heterogeneity, and water currents also play an important role (Passy 2001; Soininen 2003 Moreno-Ostos et al. 2008; Černá 2010). In Pampean shallow lakes, these small-scale processes would be responsible for the patchy distribution of the living diatom communities associated to the observed within-lake variability. Moreover, additional variability can be added to the original community during the formation of sedimentary assemblages by taphonomic processes (Hassan 2015). Given their shallowness and high nutrient content, these lakes are prone to frequent resuspension cycles and intensive grazing by benthic invertebrates, which can cause breakage of diatom

frustules and diffusion of dissolved silica, making them more susceptible to differential fragmentation and dissolution (Bennion et al. 2010; Hassan 2015; Hassan and De Francesco 2018). The deposition of resuspended particles follows complex dynamics in these lakes, as their flat profiles and very shallow depths favors sediment trapping in peripheric embayments and sheltered areas, such as macrophytes patches, rather than focusing in the deepest point (Whitmore et al. 1996). Hence, the observed variability in diatom assemblages composition can be explained by a combination of the biological and ecological processes that structure the original communities and the physical and chemical taphonomic processes that act during the deposition of their dead remains.

Variability was high within both littoral and open-water assemblages from the three lakes, signaling that spatial heterogeneity of diatom assemblages composition was more related to micro-scale environmental diversity than to bathymetric position. Several studies pointed out the presence of highly variable surface sediment diatom assemblages unrelated to bathymetry in shallow waters (De Nicola 1986; Earle et al. 1988; Wolfe 1996; Weilhoefer and Pan 2006). In these studies, lateral environmental gradients, such as patchiness in substrate availability, were pointed out as the main forcings influencing the specific composition of communities in shallow, complex environments, as they exert strong influence not only by providing substrates for colonization, but also by changing the light environment and by mediating the available nutrients (Weilhoefer and Pan 2006). Besides their similar variability around mean composition, diatom assemblages from littoral and open water environments differed significantly in their taxonomic composition in the three lakes. This compositional difference was not surprising, however, as these shallow lakes are known to support distinctive littoral and open-water habitats that promote the growth of diverse and productive diatom assemblages adapted to the intrinsic characteristics of each lake subenvironment. In open waters, dominance of planktonic and tychoplanktonic taxa, such as *C. meneghiniana* and *A. granulata*, is promoted by intermittent mixing events that frequently resuspend the sediments, preventing their loss by sinking (Bennion et al. 2010; Hassan and De Francesco 2018). On the other hand, the development of extensive littoral zones covered by macrophytes provides a variety of habitats for epipelic and epiphytic taxa, leading to the growth of characteristic communities (Wetzel 2001, Rojas and Hassan 2017). These compositional differences have implications for paleoenvironmental studies, particularly for assessing past variations on macrophyte coverage and littoralization during the Holocene. Hence, compositional differences among littoral and open waters diatom assemblages point out the relevance of representing the main within-lake subenvironments when assessing the composition of diatom assemblages in these environmentally heterogeneous shallow lakes.

Within versus between-lake variability

Multivariate comparisons of diatom assemblages among the three lakes evidenced significant differences both in their variance and average compositions. Previous studies indicated conductivity and pH as the main environmental drivers of compositional differences in death diatom assemblages between these three lakes (Hassan 2015). These differences were mainly related to differences in the relative abundances of dominant taxa, rather than to species replacement among lakes. Hence, the brackish Las Mostazas lake was characterized by higher proportions of *C. meneghiniana* and *C. placentula,* which are known to display wider environmental tolerances and opportunistic strategies in the region (Rojas and Hassan 2017; Hassan and De Francesco 2018). This lake also exhibited the lower diversity and richness, which can be related to its higher values of conductivity, pH and hardness. Because these lakes are naturally exposed to episodes of drought and flooding, their environmental differences can become crucial in periods of drought, in which strong evaporation leads to the consequent substantial increase of ionic concentration (Cristini et al. 2017; Tietze and De Francesco 2017), as became evident during summer, when it was not possible to reach Las Mostazas lake as a consequence of low water levels. Under these circumstances, conductivity in Las Mostazas reaches values above 6 mS cm^{-1} (Cristini et al. 2017), surpassing the salinity tolerance of many freshwater diatoms, and posing physiological limitations to the development of their populations (Snoeijs and Weckström 2010). Hence, assemblages in this lake became mostly dominated by taxa adapted to higher

conductivities and capable to cope with environmental changes under periods of stress, resulting in a reduced variability in composition around the average when compared to the more stable Los Carpinchos and Nahuel Rucá lakes.

Diatom assemblages from Pampean shallow lakes clearly exemplified the effect that within-lake variability exerts on between-lake comparisons. Even when the multivariate clouds of the three lakes differed significantly in their mean position, a high chance of collecting samples that are poorly representative of that difference does exist, as evidenced by the high overlap among their multivariate spaces in PCoA analysis. The extent and characteristics of this overlap can be related to environmental similarities between these lakes, as it became more evident as lakes displayed more environmental similarities. Accordingly, assemblages from Los Carpinchos, which showed intermediate conductivities, were highly overlapped with both Las Mostazas and Nahuel Rucá, while these two extreme lakes shared only a small part of their multivariate space. Hence, it becomes evident that the impact of within-lake variability in obscuring between-lake differences is related to environmental characteristics, suggesting that the problems associated to sampling representativity could tend to be reduced at longer environmental gradients. Consequently, the significant role of sampling effort in faithfully capturing assemblages composition may become emphasized when between-lake comparisons intend to capture subtle environmental differences among sites.

The results also indicated that the comparison of diatom assemblages can be significantly affected when within-lake variability differs among sites. Accordingly, lakes exhibiting higher internal variability (Nahuel Rucá and Los Carpinchos) showed also a higher proportion of samples lying in the overlapping areas of the plot. If overlooked, this variability might influence the outcome of multivariate analyses, and the consequent detection and quantification of differences between communities (Podani et al. 1993; Cao et al. 2002b; Schmera and Eros 2006). As differences among communities can be underestimated at low representativities (Cao et al. 2007), the impact of within-lake variability on between-lake comparisons can have noteworthy consequences for paleolimnological studies, not only for the construction of calibration sets but also for qualitative and semi-

quantitative interpretations. Ordination techniques are widely used in paleolimnology, both to assess similarities in assemblage composition between adjacent points (modern and fossil samples) and to detect the underlying latent structure in the data (ter Braak and Prentice 1988; Legendre and Birks 2012). At low representativities, such ordination patterns could become more related to sampling biases than to real spatial or environmental gradients, with the consequent impact on the quality of the inferences derived from the datasets. Hence, obtaining representative sampling sets for assessing within-lake diatom assemblages composition should become an essential first step in order to construct reliable and robust datasets when conducting paleolimnological reconstructions from biological data in these environmentally heterogeneous environments.

Effect of replicate sampling on diatom representativity

Because diatom assemblages were found to be heterogeneously structured and exhibited high within-lake variability in Pampean shallow lakes, the accuracy of estimations on ecological parameters (such as diversity, richness and relative abundances) will greatly depend on sample representativity, and consequently, on sampling effort. The importance of sample size on ecological estimations has been largely recognized, as well as the difficulties that this sampling effort or spatial scale effect have for community estimates and interpretations (Cao et al. 2002a). The observed small-scale patchiness in the distribution of diatom assemblages points out the relevance of sampling design to obtain representative sedimentary samples. The use of single samples as representatives of the average community overlooks within-site variability, which can constitute a significant part of the total variation (Link et al. 1994; Barker et al. 2010; Dennis et al. 2010). Hence, sampling should be focused on elucidating and representing within-lake average assemblage composition, which can only be achieved through replication (Link et al. 1994). The importance of replication as a way to cope with all sources of variability that lead to observation errors has been largely recognized in ecological and monitoring studies (Link et al. 1994; Cao et al. 2002a, b; Weiloefer and Pan 2006; Barker et al. 2010; Dennis et al. 2010); but scarcely

considered in paleolimnological and paleoenvironmental research (Earle et al. 1988; Charles et al. 1991; Bennington and Rutherford 1999; Heiri et al. 2003).

In Pampean shallow lakes, the strong spatial variability exhibited by sedimentary diatom assemblages constitutes a potential source of error to deal with in order to obtain full representation of within-lake average composition. Given the high prevalence of poor representative assemblages, single-sample based studies should be avoided. Here, multi-sample collection and replication arise as a promising approach to reduce uncertainties in sampling designs. Previous studies have pointed out the advantages collecting multiple samples from different parts of the lakes in order to improve paleoenvironmental reconstructions, particularly for quantitative inferences (Dixit and Evans 1986; Jones and Flower 1986; Earle et al. 1988; Charles et al. 1991). In the present study, simulation allowed us to evaluate the effect of replication on improving assemblages representativity, leading to the conclusion that collecting four samples would reduce noticeably the variability on diatom assemblages composition around the average, being those samples equally collected from both main lake subenvironments (i.e. littoral and open waters areas). In doing so, the average within-lake assemblage composition can be faithfully captured, even in highly variable lakes. These findings are in agreement with the previous results obtained by Weilhoefer and Pan (2006), who simulated the compositing of diatom samples and its impact on assemblages richness in morphologically complex wetlands, demonstrating that at least five composite samples were needed to characterize the diatom assemblage adequately.

The impact of replicated sampling in between-lake comparisons was also evident in the dataset, as results of PCoA demonstrated that collecting four samples was enough not only to capture within-lake average assemblages composition, but also to accurately reflect the significant compositional differences among lakes detected by the PERMANOVA. It is usually accepted that increasing sampling effort asymptotically increases the representativeness of the samples and also the separation of samples originating from different communities in multivariate spaces (Cao et al. 2002b; Schmera and Eros 2006). This is the case of the studied lakes, in which replication demonstrated to be useful not only to avoid the problems related to differences in within-

lake variability, but also to capture between-lake differences at short environmental gradients. The impact of environmental gradient lengths in these comparisons was evidenced by the results of Monte Carlo resamplings: whereas only 2 replicates were necessary to completely avoid overlapping between assemblages average composition of Nahuel Rucá and Las Mostazas, the sampling effort required to capture Los Carpinchos compositional differences was of 4 replicates. Moreover, stratifying the samplings in order to cover main lake subenvironments also allowed to capture within-lake differences among littoral and open waters assemblages, highlighting the importance of replication to capture the main compositional characteristics of sedimentary diatom assemblages related to both local and regional environmental gradients.

Temporal changes did not seem to have an impact in diatom assemblages representativity, as no clear relationship among variability and season was evident in the dataset. Except for littoral samples from Los Carpinchos, which were mostly collected in winter, representative samples were almost equally distributed among seasons in these lakes. Bray Curtis distances to average composition were also similar among seasons, being autumn samples from Nahuel Rucá the only exception. This is not surprising, however, as death assemblages constitute time-averaged representations of living communities and integrate dead-valve inputs over long periods of time, capturing demographic and environmental stochasticity and leading to time-averaged species richness (Kidwell 2002; Tomasových and Kidwell 2011). In a previous example, Bunbury and Gajewski (2008) demonstrated no significant effect of temporal variablility at the scale of years on the results of diatom-based transfer functions, associated to high time-averaging in shallow lakes. Although sedimentation rates for Pampean lakes are unknown, age models for Nahuel Rucá suggested a span of 2.5–3 years for the first centimeters of sediment (Stutz et al. 2014). Moreover, the low live-dead fidelity exhibited by diatom composition in surface sediments suggested that time-averaging in death assemblages exceeds the timeframe of 1 year (Hassan 2015). Hence, results suggest that representative samples in these lakes can be regarded as independent of sampling season, since replication over spatial rather than temporal scales was required to capture average composition. Therefore,

collecting a minimum of two littoral and two open waters samples from a single season would be enough to faithfully represent the average composition of diatom assemblages in these lakes, as demonstrated by Monte Carlo modeled replications. Certainly, more samples will always provide more information and reduce uncertainties, although the time and resources invested in obtaining such information will also depend on the balance between the costs of collecting and analyzing a higher number of samples and the improvements obtained (Heiri et al. 2003; Bennett et al. 2014). Moreover, even as the actual number of samples required will vary according to lake size and habitat complexity (Weilhoefer and Pan 2006), the minimum threshold of 4 samples identified in the present study seems to work well for lakes differing both in their environmental and compositional variability, encouraging the application of the results to other lakes and environments. Nevertheless, further studies covering different habitats and regions are needed in order to improve our knowledge of within and between-lake diatom assemblages variability and to evaluate the intensity of the sampling efforts required under contrasting environmental situations.

Conclusions

Diatom assemblages from Pampean shallow lakes surface sediments exhibited strong within-lake variability in composition. This spatial heterogeneity implied large variations in the representativity of single assemblages when compared to average composition, as a number of samples departed significantly from the lake's mean assemblage composition and were ordered in the areas of the PCoA plot shared by two lakes. This finding impacts directly on the use of surface sediment assemblages as modern analogues in paleoenvironmental reconstructions, as low representativities could mask true ecological patterns inside datasets, leading to errors in estimations of ecological parameters and biasing the outcomes of multivariate ordinations. This problem is particularly important for diatoms, as they exhibit a high variability and diversity when compared to other bioindicators, such as plant remains, cladocera or chironomids (Heggen et al. 2014). Clearly, choosing single samples as a proxy for diatom assemblage composition in these lakes would

lead to unpredictable errors in between-lake comparisons of assemblages and should not be recommended.

Under this scenario, replication arises as an alternative to cope with representativity biases. Collection of a number of replicated samples from within-lake subenvironments improved significantly the representativity of the obtained diatom assemblages, as it allowed us to capture good estimations of their mean composition in the modeled data. As the source of variability was mainly the patchy distribution of certain diatom taxa, collecting replicate samples from different points of the lakes integrated mean diatom composition better than single assemblages. In the present study, a minimum of four samples was required in order to reduce uncertainties until negligible levels in shallow lakes, being half of these collected from each of both dominant lake subenvironments (littoral and open waters). However, as diatom-assemblage variability strongly depends on the intrinsic characteristics of each lake, the required number of samples for other lakes would be variable and should be addressed individually. In the present work, the lower variability exhibited by Las Mostazas diatom assemblages led to a lower number of replicated samples threshold. Although further work is needed in order to evaluate if these findings are applicable to other lakes or regions, it becomes evident that within-lake variability needs to be carefully considered if sedimentary assemblages are being used as modern analogues in paleoenvironmental research.

The obtained results have strong implications for the quantitative reconstruction of past environmental changes from fossil diatoms in these lakes. If compositional variability is assumed to have been similar in past and modern environments, hence fossil assemblages from a single core should be regarded as one of an equally wide range of possible past compositions. Moreover, compositional changes between some successive sedimentary levels could be in some cases consequence of past within-lake variability rather than responses to paleoenvironmental changes. Under these circumstances, including multiple samples in calibration sets would improve the chances of good analogy between modern and fossil assemblages, and using MAT (Modern Analog Technique, Overpeck et al. 1985) as transfer function method could lead to reliable quantitative estimations of past conductivity fluctuations. Alternatively, the widely used WA (Weighted Averaging, ter Braak and Looman 1986)

and WA-PLS (Weighted Averaging-Partial Least Square, ter Braak and Juggins 1993) methods can be applied to down-weight species according to their indicative power, allowing to suppress the species with wide and overlapping environmental optima (Juggins and Birks 2012). Nevertheless, replication would be still needed in order to faithfully capture the optima and tolerances of single species in these environmentally heterogeneous lakes, as calculation over single samples could lead to biased inferences of species distributions over environmental gradients. For instance, the presence of *S. longirostris* in half of the samples of Nahuel Rucá indicates that by collecting a single sample a 50% chance of overlooking the presence of this indicator taxa on the lake exists, while collecting 4 replicated samples allowed us to capture this taxa in all the simulated resamplings. Hence, it can be concluded that including multiple samples from each lake when constructing calibration sets would be the best option to obtain reliable paleoenvironmental reconstructions from single sediment cores in these environmentally heterogeneous shallow lakes.

References

Adler S, Hübener T (2007) Spatial variability of diatom assemblages in surface lake sediments and its implications for transfer functions. J Paleolimnol 37:573–590

Anderson NJ (1990) Variability of sediment diatom assemblages in an upland, wind-stressed lake (Loch Fleet, Galloway, S.W. Scotland). J Paleolimnol 4:43–59

Anderson MJ, Ellingsen KE, McArdle BH (2006) Multivariate dispersion as a measure of beta diversity. Ecol Lett 9:683–693

Barker TOM, Irfanullah HM, Moss B (2010) Micro-scale structure in the chemistry and biology of a shallow lake. Fresh Biol 55:1145–1163

Bennett JR, Sisson DR, Smol JP, Cumming BF, Possingham HP, Buckley YM (2014) Optimizing taxonomic resolution and sampling effort to design cost-effective ecological models for environmental assessment. J Appl Ecol 51:1722–1732

Bennington JB, Rutherford SD (1999) Precision and reliability in paleocommunity comparisons based on cluster-confidence intervals; how to get more statistical bang for your sampling buck. Palaios 14:506–515

Bennion H, Sayer CD, Tibby J, Carrick HJ (2010) Diatoms as indicators of environmental change in shallow lakes. In: Smol JP, Stoermer EF (eds) The diatoms: applications for the environmental and earth sciences. Cambridge University Press, New York, pp 152–173

Birks HJB (1998) Numerical tools in palaeolimnology—progress, potentialities, and problems. J Paleolimnol 20:307–332

Birks HJB (2012) Overview of numerical methods in palaeolimnology. In: Birks HJB, Lotter AF, Juggins S, Smol JP (eds) Tracking environmental change using lake sediments. Springer, Dordretcht, pp 19–92

Birks HJB, Birks HH (1980) Quaternary palaeoecology. University Park Press, Baltimore

Blais JM, Kalff J (1995) The influence of lake morphometry on sediment focusing. Limnol Oceanogr 40:582–588

Bohn VY, Perillo GME, Piccolo M (2011) Distribution and morphometry of shallow lakes in a temperate zone (Buenos Aires Province, Argentina). Limnetica 30:89–102

Bray JR, Curtis JT (1957) An ordination of the upland forest communities of Southern Wisconsin. Ecol Monogr 27:325–349

Bunbury J, Gajewski K (2008) Does a one point sample adequately characterize the lake environment for paleoenvironmental calibration studies? J Paleolimnol 39:511–531

Cao Y, Williams DD, Larsen DP (2002a) Comparison of ecological communities: the problem of sample representativeness. Ecol Monogr 72:41–56

Cao Y, Larsen DP, Hughes RM, Angermeier PL, Patton TM (2002b) Sampling effort affects multivariate comparisons of stream assemblages. J North Am Benthol Soc 21:701–714

Cao Y, Hawkins CP, Larsen DP, Van Sickle J (2007) Effects of sample standardization on mean species detectabilities and estimates of relative differences in species richness among assemblages. Am Nat 170:381–395

Černá K (2010) Small-scale spatial variation of benthic algal assemblages in a peat bog. Limnologica 40:315–321

Charles DF, Dixit SS, Cumming BF, Smol JP (1991) Variability in diatom and chrysophyte assemblages and inferred pH: paleolimnological studies of Big Moose Lake, New York, USA. J Paleolimnol 5:267–284

Clarke KR (1993) Non-parametric multivariate analyses of changes in community structure. Aust J Ecol 18:117–143

Cristini PA, Tietze E, De Francesco CG, Martínez DE (2017) Water geochemistry of shallow lakes from the southeastern Pampa plain, Argentina and their implications on mollusk shells preservation. Sci Tot Environ 603:155–166

De Nicola DM (1986) The representation of living diatom communities in deep-water sedimentary diatom assemblages in two Maine (U.S.A.) lakes. In: Smol JP, Battarbee RW, Davis RB, Meriläinen J (eds) chapter 7, Diatoms and lake acidity. Dr W Junk Publishers, Dordrecht, pp 73–85

Dennis B, Ponciano JM, Taper ML (2010) Replicated sampling increases efficiency in monitoring biological populations. Ecology 91:610–620

Diovisalvi N, Baigu C, Bohn VY, Piccolo MC, Perillo GME, Zagarese HE (2015) Shallow lakes from the Central Plains of Argentina: an overview and worldwide comparative analysis of their basic limnological features. Hydrobiologia 752:5–20

Dixit SS, Evans RD (1986) Spatial variability in sedimentary algal microfossils and its bearing on diatom-inferred pH reconstructions. Can J Fish Aquat Sci 43:1836–1845

Earle JC, Duthie HC, Glooschenko WA, Hamilton PB (1988) Factors affecting the spatial distribution of diatoms on the surface sediments of three Precambrian Shield lakes. Can J Fish Aquat Sci 45:469–478

Eggermont H, De Deyne P, Verschuren D (2007) Spatial variability of chironomid death assemblages in the surface sediments of a fluctuating tropical lake (Lake Naivasha, Kenya). J Paleolimnol 38:309–328

Federman M (2003) Mapeo y caracterización de la comunidad de macrófitas en tres lagos someros del sudeste bonaerense. Tesis de grado de la licenciatura en Ciencias Biológicas. Universidad Nacional de Mar del Plata. Argentina

Frey D (1988) Littoral and offshore communities of diatoms, cladocerans and dipterous larvae, and their interpretation in paleolimnology. J Paleolimnol 1:179–191

García-Rodríguez F, Piovano E, del Puerto L, Inda H, Stutz S, Bracco R, Panario D, Córdoba F, Sylvestre F, Ariztegui D (2009) South American lake paleo-records across the Pampean Region. PAGES News 17:115–118

Gilbert R, Lamoureux S (2004) Processes affecting deposition of sediment in a small, morphologically complex lake. J Paleolimnol 31:37–48

González-Sagrario MA, Rodríguez Golpe D, La Sala L, Sánchez Vuichard G, Minottic P, Panarello H (2018) Lake size, macrophytes, and omnivory contribute to food web linkage in temperate shallow eutrophic lakes. Hydrobiologia 818:87–103

Hammer Ø, Harper DAT, Ryan PD (2001) PAST: Paleontological statistics software package for education and data analysis: palaeontologia electronica, 4, no. 1, 184 Kb, http://palaeo-electronica.org/2001_1/past/issue1_01.htm

Hassan GS (2015) On the benefits of being redundant: low compositional fidelity of diatom death assemblages does not hamper the preservation of environmental gradients in shallow lakes. Paleobiology 41:154–173

Hassan GS, De Francesco CG (2018) Preservation of *Cyclotella meneghiniana* Kützing (Bacillariophyceae) along a continental salinity gradient: implications for diatom-based paleoenvironmental reconstructions. Ameghiniana. https://doi.org/10.5710/AMGH.20.11.2017.3144

Heggen MP, Birks HH, Heiri O, Grytnes JA, Birks HJB (2012) Are fossil assemblages in a single sediment core from a small lake representative of total deposition of mite, chironomid, and plant macrofossil remains? J Paleolimnol 48:669–691

Heiri O, Birks HJB, Brooks SJ, Velle G, Willassen E (2003) Effects of within-lake variability of fossil assemblages on quantitative chironomid-inferred temperature reconstruction. Palaeogeogr Palaeoclimatol Palaeoecol 199:95–106

Jones VJ, Flower RJ (1986) Spatial and temporal variability in periphytic diatom communities: palaeoecological significance in an acified lake. In: Smol JP, Battarbee RW, Davis RB, Meriläinen J (eds) Diatoms and lake acidity. Dr W Junk Publishers, Dordrecht, pp 87–94

Juggins S, Birks HJB (2012) Quantitative environmental reconstructions from biological data. In: Birks HJB, Lotter AF, Juggins S, Smol JP (eds) Tracking environmental change. Springer, Netherlands, pp 431–494

Kelly MG (2001) Use of similarity measures for quality control of benthic diatom samples. Water Res 35:2784–2788

Kidwell SM (2002) Time-averaged molluscan death assemblages: palimpsests of richness, snapshots of abundance. Geology 30:803–806

Krammer K, Lange-Bertalot H (1997) Bacillariophyceae 2. Teil: Bacillariophyceae, Epithemiaceae, Surirellaceae. In: Ettl H, Gerloff J, Heynig H, Mollenhauer D (eds) Süsswasserflora von Mitteleuropa, 2/2 edn. Spektrum Akademischer Verlag, Stuttgart, p 610

Krammer K, Lange-Bertalot H (1999) Bacillariophyceae 1. Teil: Naviculaceae. In: Ettl H, Gerloff J, Heynig H, Mollenhauer D (eds) Süsswasserflora von Mitteleuropa, 2/1 edn. Spektrum Akademischer Verlag, Stuttgart, p 876

Krammer K, Lange-Bertalot H (2004a) Bacillariophyceae 3. Teil: Centrales; Fragilariaceae, Eunotiaceae. In: Ettl H, Gerloff J, Heynig H, Mollenhauer D (eds) Süsswasserflora von Mitteleuropa, 2/3 edn. Spektrum Akademischer Verlag, Stuttgart, p 596

Krammer K, Lange-Bertalot H (2004b) Bacillariophyceae 4. Teil: Bacillariophyceae, Epithemiaceae, Surirellaceae. In: Ettl H, Gerloff J, Heynig H, Mollenhauer D (eds) Süsswasserflora von Mitteleuropa, 2/4 edn. Spektrum Akademischer Verlag, Stuttgart, p 596

Lange-Bertalot H, Külbs K, Lauser T, Nörpel-Schempp M, Willmann M (1996) Diatom taxa introduced by Georg Krasske documentation and revision. Iconogr Diatomol 3:1–358

Ledger M, Hildrew AG (1998) Temporal and spatial variation in the epilithic biofilm of an acid stream. Freshw Biol 40:655–670

Legendre P, Birks HJB (2012) From classical to canonical ordination. In: Birks HJB, Lotter AF, Juggins S, Smol JP (eds) Tracking environmental change using lake sediments, vol 5. Springer, Dordrecht, pp 201–248

Levkov Z (2009) *Amphora* sensu lato. In: Lange-Bertalot H (ed) Diatoms of Europe: diatoms of European inland waters and comparable habitats, vol 5. ARG GantnerVerlag KG, Ruggell, p 916

Link WA, Barker RJ, Sauer JR, Droege S (1994) Within-site variability in surveys of wildlife populations. Ecology 75:1097–1108

Magurran AE (2004) Measuring biological diversity. Blackwell, London, p 132

Metzeltin D, Lange-Bertalot H (1998) Tropical diatoms of South America I. About 700 predominantly rarely known or new taxa representative of the neotropical flora. Iconogr Diatomol 5:1–695

Metzeltin D, Lange-Bertalot H (2007) Tropical diatoms of South America II. Special remarks on biogeographic disjunction. Iconogr Diatomol 18:1–876

Metzeltin D, Lange-Bertalot H, García-Rodríguez F (2005) Diatoms of Uruguay compared with other taxa from South America and elsewhere. Iconogr Diatomol 15:1–736

Moreno-Ostos E, Cruz-Pizarro L, Basanta A, Glen G (2008) The spatial distribution of different phytoplankton functional groups in a Mediterranean reservoir. Aquat Ecol 42:115–128

Oksanen JF, Blanchet G, Kindt R, Legendre P, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner S (2013) Vegan: community ecology package. R package, Version 2. 0–7. http://CRAN.R-project.org/package=vegan. Accessed 17 Jun 2015

Overpeck JT, Webb T, Prentice IC (1985) Quantitative interpretation of fossil pollen spectra—dissimilarity coefficients and the method of modern analogs. Quat Res 23:87–108

Passy SI (2001) Spatial paradigms of lotic diatom distribution: a landscape ecology perspective. J Phycol 37:370–378

Podani J, Czárán T, Bartha S (1993) Pattern, area and diversity: the importance of spatial scale in species assemblages. Abst Bot 17:37–51

R Development Core Team (2016) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. www.R-project.org

Rojas L, Hassan GS (2017) Distribution of epiphytic diatoms on five macrophytes from a Pampean shallow lake: host-specificity and implications for paleoenvironmental reconstructions. Diat Res 32:263–275

Rumrich U, Lange-Bertalot H, Rumrich M (2000) Diatoms of the Andes. From Venezuela to Patagonia/Tierra del Fuego. Iconogr Diatomol 9:1–673

Sar EA, Sala SE, Sunesen I, Henninger S, Montastrut M (2009) Catálogo de los géneros, especies y taxa intraespecíficos erigidos por J. Frenguelli. In: Witkowski A (ed) Diatom monographs, vol 10. Koeltz Scientific Books, Koenigstein, p 419

Schmera D, Eros T (2006) Estimating sample representativeness in a survey of stream caddisfly fauna. Ann Limnol 42:181–187

Smol J (1992) Paleolimnology: an important tool for effective ecosystem management. J Aquat Ecosyst Health 1:49–58

Snoeijs P, Weckström K (2010) Diatoms and environmental change in large brackish-water Ecosystems. In: Smol JP, Stoermer EF (eds) The diatoms: applications for the environmental and earth sciences. Cambridge University Press, New York, pp 287–308

Soininen J (2003) Heterogeneity of benthic diatom communities in different spatial scales and current velocities in a turbid river. Arch Hydrobiol 156:551–564

Sommer U (2000) Benthic microalgal diversity enhanced by spatial heterogeneity of grazing. Oecologia 122:284–287

Stevenson RJ, Singer R, Roberts DA, Boylen CW (1985) Patterns of epipelic algal abundance with depth, trophic status, and acidity in poorly buffered New Hampshire lakes. Can J Fish Aquat Sci 42:1501–1512

Stutz S, Tonello M, González-Sagrario MA, Navarro D, Fontana S (2014) Historia ambiental de los lagos someros de la llanura Pampeana (Argentina) desde el Holoceno medio: inferencias paleoclimáticas. Lat Am J Sed Bas Anal 21:119–138

ter Braak CJF, Juggins S (1993) Weighted averaging partial least squares regression (WA-PLS): an improved method for reconstructing environmental variables from species assemblages. Hydrobiolgia 269(270):485–502

ter Braak CJF, Looman CWN (1986) Weighted averaging, logistic regression and the Gaussian response model. Vegetatio 65:3–11

ter Braak CJF, Prentice IC (1988) A theory of gradient analysis. Adv Ecol Res 18:271–317

Tietze E, De Francesco CG (2017) Compositional fidelity and taphonomy of freshwater mollusks from three pampean shallow lakes of Argentina. Ameghiniana 54:208–226

Tomasových A, Kidwell SM (2011) Accounting for the effects of biological variability and temporal autocorrelation in assessing the preservation of species abundance. Paleobiology 37:332–354

Weilhoefer C, Pan Y (2006) Diatom-based bioassessment in wetlands: how many samples do we need to characterize the diatom assemblage in a wetland adequately? Wetlands 26:793–802

Wetzel R (2001) Limnology: lake and river ecosystems. Academic Press, San Diego, p 1006

Whitmore T, Brenner M, Schelske CL (1996) Highly variable sediment distribution in shallow, wind-stressed lakes: a case for sediment-mapping surveys in paleolimnological studies. J Paleolimnol 15:207–221

Wickham H (2016) Plyr: Tools for Splitting, Applying and Combining Data. R package, Version 1. 8.4. http://CRAN.R-project.org/package=plyr. Accessed 23 Oct 2017

Wolfe AP (1996) Spatial patterns of modern diatom distribution and multiple paleolimnological records from a small arctic lake on Baffin Island, Arctic Canada. Can J Bot 74:435–449

Zar JH (2010) Biostatistical analysis. Prentice Hall, New Jersey