



**HAL**  
open science

## Contrasting the genetic patterns of microbial communities in soda lakes with and without cyanobacterial bloom

Ana Paula Dini Andreote, Janaina Rigonato, Francisco Dini-Andreote, Gabriela Machineski Silva, Bruno Costa Evangelista Souza, Laurent Barbiero, Ary Tavares Rezende-Filho, Marli Fátima Fiore

### ► To cite this version:

Ana Paula Dini Andreote, Janaina Rigonato, Francisco Dini-Andreote, Gabriela Machineski Silva, Bruno Costa Evangelista Souza, et al.. Contrasting the genetic patterns of microbial communities in soda lakes with and without cyanobacterial bloom. *Frontiers in Microbiology*, 2018, 9, pp.244. 10.3389/fmicb.2018.00244 . hal-02084157

**HAL Id: hal-02084157**

**<https://hal.science/hal-02084157>**

Submitted on 29 Mar 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Contrasting the Genetic Patterns of Microbial Communities in Soda lakes with and without Cyanobacterial Bloom

Ana Paula Dini Andreote<sup>1</sup>, Janaina Rigonato<sup>1</sup>, Francisco Dini-Andreote<sup>2</sup>, Gabriela Machineski Silva<sup>1</sup>, Bruno Costa Evangelista Souza<sup>1</sup>, Laurent Barbiero<sup>3</sup>, Ary Tavares Rezende-Filho<sup>4</sup>, Marli Fatima Fiore<sup>1\*</sup>.

<sup>1</sup>University of São Paulo, Center for Nuclear Energy in Agriculture, Piracicaba, São Paulo, Brazil

<sup>2</sup>Microbial Ecology cluster, Genomics Research in Ecology and Evolution (GREEN), Groningen Institute for Evolutionary Life Sciences (GELIFES), University of Groningen, The Netherlands

<sup>3</sup>IRD, CNRS, UPS, OMP, Geoscience Environnement Toulouse, France

<sup>4</sup>Federal University of Mato Grosso do Sul, Faculty of Engineering, Architecture, Urbanism and Geography, Campo Grande, Mato Grosso do Sul, Brazil

## \*Correspondence:

Marli F. Fiore

fiore@cena.usp.br

**Article type:** Original Research

**Running title:** Soda lakes with cyanobacterial bloom

**Keywords:** arsenic, alkaline lakes, metagenomic, nitrogen, saline lakes, sulfur.

Length: 8,136 words. Include 4 figures and 3 tables.

## Abstract

Word count: 290

Soda lakes present high levels of sodium carbonates and are characterized by elevated salinity and pH. These ecosystems are found across Africa, Europe, Asia, Australia, North, Central and South-America. Particularly in Brazil, the Pantanal region presents a series of hundreds of shallow soda lakes (around 600) potentially colonized by a diverse haloalkaliphilic microbial community. Biological information of these systems is still elusive, in particular data on the description of the main taxa involved in the biogeochemical cycling of life-important elements. Here, we use metagenomic sequencing to compare the composition and functional profiles of two distinct soda lakes from the sub-region Nhecolândia, state of Mato Grosso do Sul, Brazil. These two lakes differ by the permanent cyanobacteria bloom (Salina Verde, green water lake) and by no record of cyanobacteria bloom (Salina Preta, black water lake). The dominant bacterial species in the Salina Verde water bloom was *Anabaenopsis elenkinii*. This cyanobacterium altered local abiotic parameters such as pH, turbidity and dissolved oxygen and consequently the overall structure of the microbial community. In the Salina Preta, the microbial community presented higher diversity and a structured profile. Therefore, the distribution of metabolic functions in the Salina Preta community encompassed a large number of taxa, whereas, in Salina Verde, the functionality was restrained across a specific set of existing taxa. Distinct signatures in the abundance of genes associated with the cycling carbon, nitrogen and sulfur were found. Interestingly, arsenic metabolism was present at higher abundance in the Salina Verde and it was associated with cyanobacteria bloom. Our data revealed a metabolic

connection between differences in the taxonomic community composition and the functional potential of these communities. Collectively, this study advances fundamental knowledge on the composition and functional potential of microbial communities in tropical soda lakes.

## Introduction

Soda lakes are characterized by high pH and a Na-HCO<sub>3</sub> chemical profile that result in saline and hypersaline alkaline waters. These ecosystems have been described worldwide, including the East African Rift Valley (Eugster and Jones, 1979; Jones et al., 1977; Ventosa and Arahal 1989, Duarte et al. 2012), North and Central America (Domagalski et al., 1989; Kazmierczak et al., 2011), Asia (Ma and Edmunds, 2006; Namsaraev et al., 2015), Australia (Hammer, 1986), and Europe (Felföldi et al., 2009; Stenger-kovács et al., 2014). The extreme environmental conditions of soda lakes (pH above 9 and elevated salinity) are known to be stressful for phytoplankton communities, but several microbial and microalgae taxa have adapted to live under these harsh abiotic conditions. It is worth noticing that, despite these extreme conditions, moderately saline alkaline lakes are recognized as among the most productive lakes worldwide (Hammer, 1981; Melack, 1981; Melack and Kilham, 1974). Often, seasonal and permanent microbial bloom events are observed in soda lakes (Jones, 1999; Grant, 2006). In African environments, blooms have been mostly dominated by the cyanobacteria genera *Spirulina-Cyanospira-Arthrospira* (Krienitz et al., 2013). Moreover, saline soda lakes harbor fully functional and diverse microbial communities responsible for the cycling of chemical elements, such as nitrogen, sulfur and carbon (Sorokin et al., 2014). However, the functional links between these microorganisms and their ecophysiological roles in the biogeochemical cycling are poorly understood (Paul Antony et al., 2013). In the last decade, studies in soda lakes have emphasized the importance of investigating the linkage between the biogeochemical processes and microbial community diversity in order to address the dynamics of these systems (Paul Antony et al., 2013; Sorokin et al., 2015).

In Brazil, soda lakes are located in the Pantanal region and present high alkalinity and salt concentration, similarly to Soda Lakes in East Africa (Ventosa and Arahal 1989, Duarte et al. 2012). The alkalinity is a result of the shift in the CO<sub>2</sub>/HCO<sub>3</sub><sup>-</sup>/CO<sub>3</sub><sup>2-</sup> equilibrium, and, due to the absence of Ca<sup>2+</sup> or Mg<sup>2+</sup>, CO<sub>3</sub><sup>2-</sup> remains soluble at elevated concentrations. In addition, there is a high concentration of Na-HCO<sub>3</sub> in these systems (Furian et al., 2013). Soda lakes are normally shallow and surrounded by a ground soil elevation that forms a barrier for exchanging water/flooding, which promotes the lake physical isolation. The occurrence of natural cyanobacteria blooms in Brazilian soda lakes are also common (De-Lamonica-Freire and Heckman, 1996; Oliveira and Calheiros, 2000; Medina-Júnior and Rietzler, 2005), being mostly associated with the cyanobacteria species *Anabaenopsis elenkinii* and *Arthrospira platensis* (Santos and Sant'anna, 2010; Costa et al., 2016). However, few culture-based studies on these soda lakes also recovered several other cyanobacteria, including new genera and species (Andreote et al., 2014; Genuário et al., 2017; Vaz et al., 2015).

The Brazilian Pantanal region has a sub-region called Nhecolândia, which presents a series of hundreds of soda lakes. These lakes are fed by water at several concentration stages of the Taquari river, arising from Mesozoic sandstone formations at the uplands (Rezende-Filho et al., 2012). In these lakes, the water column level varies by the process of evaporation and chemical concentration of solutes, being intrinsically influenced by the pronounced seasonal rainfall cycle, with the largest proportion of rainfall occurring in the period between October and April (Tarifa, 1986). Particularly of this sub-region, soda lakes have been studied regarding the hydrological functioning and inorganic geochemistry (Barbiero et al., 2007, 2008, 2016; Furian et al., 2013, Furquim et al., 2010a, 2010b). Although several biogeochemical functions have been described in Brazilian soda lakes (Martins, 2012), information on the microbial community diversity and functioning in these systems are scarce. As such, we set a particular focus on using comparative metagenomics analysis to foster our understanding of the microbial community composition and biogeochemical cycling of nutrients in

these systems by comparing two distinct soda lakes from sub-region Nhecolândia, one with permanent cyanobacteria bloom (Green water lake) and other with no record of cyanobacteria bloom (Black water lake).

## **Material and Methods**

### **Sample collection and characteristics of soda lakes**

Green water lake (Salina Verde, 19°28'13"S; 56°3'22"W) and Black water lake (Salina Preta, 19°26'56"S; 56°7'55"W) are located on the Centenário farm in the Pantanal region (Nhecolândia, Mato Grosso do Sul, Brazil). The Salina Verde was named based on the green color of its water ("verde" in Portuguese) due to the presence of continuous cyanobacteria bloom. A set of important characteristics and abiotic variables of both of these lakes were previously reported (Andreote et al., 2014). The sampling was performed on September 2012, during the dry season. For that, water samples were collected in triplicate using sterile bottles of 50 mL at 0.40 meters from the lake borders. Samples were collected at two depths (surface and bottom), at two-day times (10:00 am and 3:00 pm). The sampling time was determined according to the natural occurrence of oxygen saturation observed in the Salina Verde.

### **Optical microscopy**

Fresh water samples were visualized under optical microscopy for the identification of the main dominant cyanobacterium species present in the Salina Verde samples. The morphological characterization was performed according to Komárek and Anagnostidis (1989).

### **Total DNA isolation and shotgun metagenomics**

Water samples were concentrated by centrifugation at  $7,000 \times g$  for 5 min and kept at  $-20^{\circ} C$  for further analysis. Total DNA was isolated using the MoBio Power Soil DNA isolation kit (MoBio Laboratories, Carlsbad, CA, USA), following the manufacturer's protocol. The quantification of the extracted DNA samples was performed by comparison to Low DNA mass ladder and Qubit dsDNA HS assay kit (Thermo Fisher Scientific, Carlsbad, CA, USA), and DNA integrity was checked by electrophoresis on agarose gel (1% w/v).

A total of 24 metagenomic libraries (2 lakes x 2 depths x 2 daytime x 3 replicates) were prepared using the Nextera XT DNA Sample Preparation kit (Illumina, Inc., San Diego, CA, USA), according to the manufacturer's protocol. The libraries were sequenced using the Illumina MiSeq platform (Illumina) and the MiSeq Reagent Kit V3 (600 cycles) (Illumina). Overlapping reads were paired using PEAR 0.9.5 (Zhang et al., 2014) and filtered for quality with Phred  $>30$  and length longer than 80 nt using Seqclean 1.9.8 (<https://bitbucket.org/izhbannikov/seqclean>). The sequences were compared by BLASTX against the NCBI-nr database (November, 2016) using DIAMOND (Buchfink et al., 2014), and analyzed using MEGAN6 (Huson et al., 2016). Sequences were matched to SEED database for functional assignment (Overbeek, 2005) and KEGG for metabolic pathways (Kanehisa and Goto, 2000). For the assignation of taxa associated with the distinct biogeochemical processes, we used the FOCUS tool (Silva et al., 2014) with default parameters.

### **Measurement of abiotic parameters**

The measurements of pH, electrical conductivity, temperature, water turbidity and dissolved oxygen (DO) were performed *in situ* at a distance of 20 m from the lake edges, in two different expeditions

during the dry season (2012 and 2013). Water samples were centrifuged at  $3,200 \times g$  for 30 minutes in the field, immediately after sampling. Samples were then filtered (0.45  $\mu\text{m}$  cellulose acetate filters) and stored in a cool and dark place for further analyses. The nitrate ( $\text{NO}_3^-$ ), ammonium ( $\text{NH}_4^+$ ), phosphate ( $\text{PO}_4^{3-}$ ), sulfate ( $\text{SO}_4^{2-}$ ), the ions  $\text{Cl}^-$ ,  $\text{Na}^+$ ,  $\text{Mg}^{2+}$ ,  $\text{Ca}^{2+}$  and the carbonate alkalinity were measured. Anions and cations were analysed by ion chromatography, and carbonate alkalinity by acid titration (HCl 0.1 N).

## Statistical analysis

The relationship between abiotic parameters and microbial community compositions were tested by Pearson correlation coefficient using ISwR package in R software version 3.3.2 (Dalgaard, 2008).

To test for statistical differences of SEED annotations across samples, we used the STAMP software package (Statistical Analysis of Metagenomic Profiles) (Parks et al., 2014). The  $P$ -values were calculated using the two-sided Welch's  $t$ -test followed by Benjamini-Hochberg FDR correction (Haynes, 2013).

The functional pathways of carbon, nitrogen and sulfur were annotated and mapped using the read counts from SEED affiliations and marker genes (Supplementary Table 1). The heatmap was constructed based on  $z$ -score transformed functional annotations to improve normality and homogeneity of the variances. Predicted functional annotations that segregated significantly between sample types were identified using random forest analysis (Breiman, 2001) with 1,000 trees followed by the Boruta algorithm for feature selection (average  $z$ -scores of 1,000 runs  $> 4$ ) (Kursa and Rudnicki, 2010).

## Results

### Differences in environmental parameters between lakes

The measurement of abiotic parameters in both lakes at the time of sampling revealed significant differences (see Table 1 for details). Taking into account that the lakes were sampled during the dry season when the water level was low (sampling site Salina Verde 0.25 m depth and Salina Preta 0.35 m), it was observed that water temperature and conductivity increased along the day in both lakes. An increase in water turbidity was observed in Salina Verde in the afternoon (Table 1), due to the natural oxygen saturation observed in this lake around 01h pm. Alkalinity, pH and conductivity parameters were relatively higher in Salina Verde than Salina Preta. The values of salinity, calculated based on ions' content of total dissolved solids (TDS), were  $2.21 \text{ g}\cdot\text{L}^{-1}$  and  $1.24 \text{ g}\cdot\text{L}^{-1}$  in the Salina Verde and Salina Preta, respectively. Worth noting, these values categorizes Salina Verde as presenting 'very high salinity', and Salina Preta presenting 'high salinity', as from the classification for lakes waters of the American guideline (APHA, 1995).

### Taxonomic composition of bacterial communities in soda lakes

The taxonomic composition of the communities based on annotated metagenomic reads showed that sequences affiliated to Bacteria were numerically dominant in the samples (99.56 – 99.39 %), with a minor representation of archaeal sequences (0.08 – 2.37 %), eukaryotes (0.39 – 2.49 %), and viruses (0.11 – 0.58 %) (Table 2). Salina Verde presented a higher relative abundance of bacteria compared to the Salina Preta ( $p = 1.59\text{e}^{-3}$ ). The relative abundance of eukaryotes ( $p=9.49\text{e}^{-3}$ ) and archaea ( $p=0.015$ ) were higher in the Salina Preta. The relative abundance of the archaeal phylum Thaumarchaeota, encompassing ammonia oxidizer organisms, was higher in the Salina Preta ( $p=7.02\text{e}^{-3}$ ), and those affiliated to Euryarchaeota encompassing methanogenic organisms had a

higher relative abundance in the Salina Verde ( $p=8.38e^{-3}$ ). No difference in community composition was observed between the two periods sampled within the same lake.

**TABLE 1** | Physicochemical data of Salina Verde and Salina Preta lakes measured at morning and afternoon (September, 2012).

	Salina Verde		Salina Preta	
	Morning	Afternoon	Morning	Afternoon
Depth (m)		0.25		0.35
Temperature (°C)	26.8		28.7	32.7
Conductivity ( $\mu\text{S}\cdot\text{cm}^{-1}$ )	2,528	2,741	1,465	1,587
Alkalinity (meq <sub>L</sub> )		22.5		10.6
pH	9.6	9.5	9.0	8.8
DO (%) <sup>a</sup>	172	327	86	104
Turbidity (NTU)	852	882	90	90
NO <sub>3</sub> <sup>-</sup> (mmol·L <sup>-1</sup> )		0.012		0.006
SO <sub>4</sub> <sup>2-</sup> (mmol·L <sup>-1</sup> )		2.02		2.33
PO <sub>4</sub> <sup>3-</sup> (mmol·L <sup>-1</sup> )		0.004		0.005
NH <sub>4</sub> <sup>+</sup> (mmol·L <sup>-1</sup> )		0.6		0.001
Na <sup>+</sup> (mmol·L <sup>-1</sup> )		18.3		11.4
K <sup>+</sup> (mmol·L <sup>-1</sup> )		8.24		4.39
Ca <sup>2+</sup> (mmol·L <sup>-1</sup> )		0.35		0.36
Mg <sup>2+</sup> (mmol·L <sup>-1</sup> )		0.15		0.21
Cl <sup>-</sup> (mmol·L <sup>-1</sup> )		1.64		1.49

<sup>a</sup>DO% in Salina Verde varied from 5% at night to 600% during the day (Barbiero et al., 2017).

The dominant phylum in the Salina Verde was Cyanobacteria (Figure 1), with the main genus accounting for the bloom event being affiliated to the nitrogen-fixing *Anabaenopsis* (Nostocales). The optical microscopy investigation of the water bloom samples identified as the dominant cyanobacterial species the planktonic heterocytous filamentous *Anabaenopsis elenkini*. In addition, members of the orders Oscillatoriales ( $p=8.01e^{-11}$ ), Spirulinales ( $p<0.028$ ), Synechococcales ( $p=2.51e^{-4}$ ), Pleurocapsales ( $p=1.32e^{-5}$ ) and Chroococcales ( $p=2.61e^{-9}$ ) were higher in terms of relative abundances in the Salina Verde comparatively to the Salina Preta. Alphaproteobacteria and Betaproteobacteria were also present at relatively high abundances, followed by Betaproteobacteria and Deltaproteobacteria.

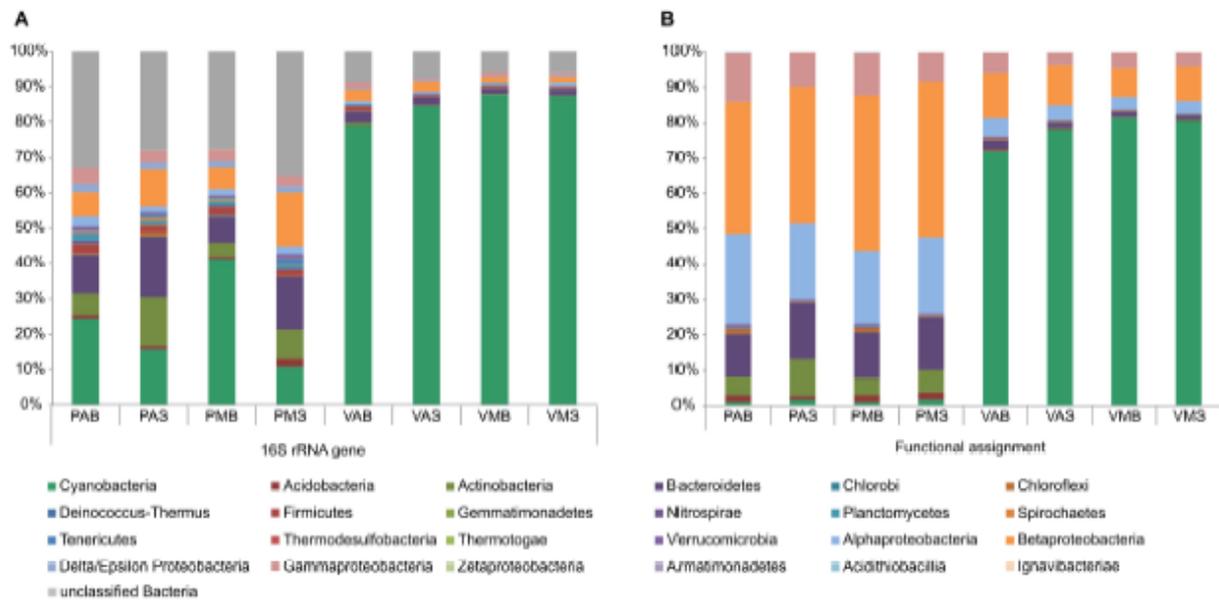
**TABLE 2** | Overall number of metagenomic reads obtained for Salina Verde and Salina Preta samples.

	PAB	PAS	PMB	PMS	VAB	VAS	VMB	VMS
Total nucleotide (Mb)	373,382	534,239	352,370	173,429	178,547	323,728	362,425	286,395
Average read length (bp)	202	214	212	219	196	189	163	167
Total number of reads	1,921,878	2,635,106	1,828,545	808,867	952,827	1,801,275	2,248,701	1,753,792
Taxonomic assigned reads (%)	28.4	33.2	30.1	38.0	47.6	50.4	46.9	39.7
Bacteria (%)	96.09	97.14	94.56	96.56	99.24	99.39	99.26	99.31
Archaea (%)	2.27	0.22	2.37	2.27	0.25	0.06	0.24	0.08
Eukaryota (%)	1.06	1.13	2.49	1.06	0.37	0.39	0.40	0.49
Viruses (%)	0.58	1.52	0.58	0.58	0.15	0.15	0.11	0.12
Functionally assigned reads (%)	37.05	43.54	38.12	52.35	44.52	44.16	39.44	39.24
16S rRNA genes in the metagenome	1477	2746	1947	754	1695	3068	3775	3139
Bacteria 16S rRNA (%)	98.9	100	99.28	100	100	100	99.9	100
Archaea 16S rRNA (%)	1.1	0	0.72	0	0	0	0.10	0

V: Salina Verde, P: Salina Preta, M: morning, A: afternoon, S: surface, and B: bottom.

The phylum Proteobacteria was dominant in the Salina Preta, being at a higher relative abundance than in the Salina Verde ( $p=6.37e^{-14}$ ). The most present classes were Betaproteobacteria, Alphaproteobacteria, Gammaproteobacteria and Deltaproteobacteria. Other phyla were also found at higher relative abundances in the Salina Preta, as follows: Actinobacteria ( $p=1.38e^{-5}$ ), Acidobacteria ( $p=1.71e^{-6}$ ), Bacteroidetes ( $p=6.25e^{-10}$ ), Planctomycetes ( $p=1.72e^{-6}$ ), Verrucomicrobia ( $p=2.05e^{-10}$ ), Gemmatimonadetes ( $p=1.23e^{-7}$ ), Chloroflexi ( $p=0.046$ ), Chlorobi ( $p=7.37e^{-3}$ ), Armatimonadetes ( $p=5.74e^{-3}$ ) and Nitrospirae ( $p=2.47e^{-4}$ ). It is worth mentioning that a higher relative abundance of the

green filamentous phototrophs Chloroflexi ( $p=2.42e^{-3}$ ) was found in the bottom samples of the Salina Preta (Fig. 2A). In addition, the archaeal phylum Thaumarchaeota ( $p=2.42e^{-3}$ ), and the bacterial groups Nitrospirae ( $p=0.041$ ), Ignavibacteriae ( $p=3.45e^{-3}$ ), Armatimonadetes ( $p=0.011$ ) and Spirochaetes ( $p=0.018$ ) were found to be higher in relative abundances in the bottom samples comparatively to the surface samples in the Salina Preta.



**FIGURE 1** | Relative abundances of the main bacterial phyla/classes in Salina Preta and Salina Verde lakes. Data are shown based on bacterial 16S rRNA gene sequences retrieved from the metagenomes (A) and functional metagenome annotations (B) (see section "Materials and Methods" for details). V: Salina Verde, P: Salina Preta, M: morning, A: afternoon, S: surface, B: bottom.

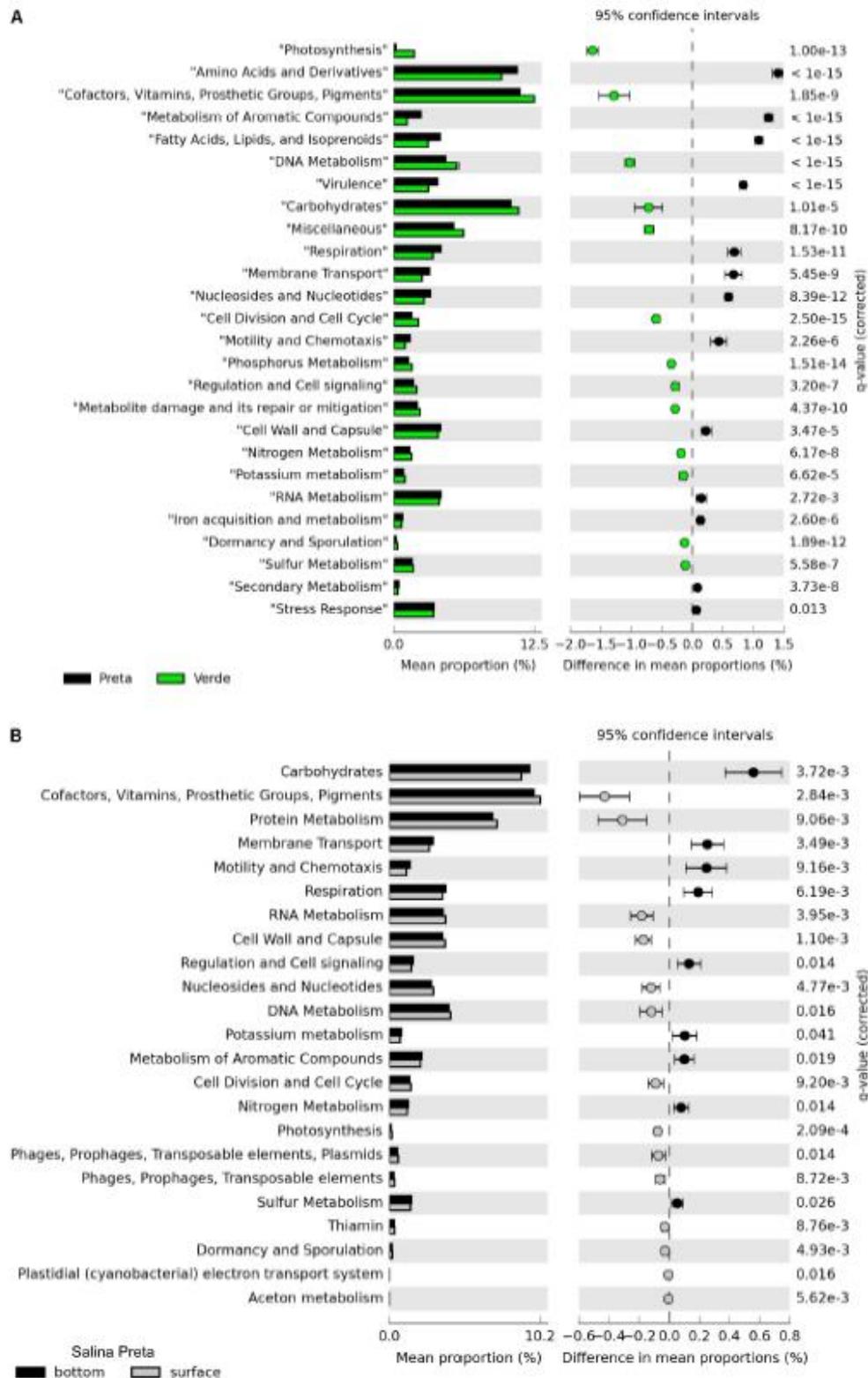
### Correlation between abiotic features and communities diversity

A set of abiotic parameters measured in the local sites were significantly correlated with the taxonomic community composition of these studied lakes (Table 3). The Cyanobacteria phylum, dominant in Salina Verde, was highly positively correlated with conductivity ( $r=0.814$ ,  $p<0.05$ ), turbidity ( $r=0.866$ ,  $p<0.05$ ) and pH ( $r=0.842$ ,  $p<0.05$ ); moderately positively correlated with dissolved oxygen ( $r=0.435$ ,  $p<0.01$ ) and alkalinity ( $r=0.474$ ,  $p<0.01$ ); and highly negatively correlated with nitrate ( $\text{NO}_3^-$ ) concentration ( $r=-0.870$ ,  $p<0.05$ ). Archaea and several bacterial taxa abundant in Salina Preta, such as Alphaproteobacteria, Betaproteobacteria, Gammaproteobacteria, Actinobacteria and Bacteroidetes, were positively-correlated with nitrate concentration in water and negatively-correlated with increases in conductivity, dissolved oxygen, turbidity, and pH of the water (Table 3). No significant correlation was found between the abiotic parameters temperature, phosphate and sulfate concentrations, and the microbial community composition.

### Functional genetic potential of microbial communities in soda lakes

A total of 13.9 million reads were generated in our metagenomics dataset. The percentage value of functional assigned reads varied between 39.24 to 52.35% based on the SEED database. The functional profiles of the lake metagenomes were compared using STAMP, and significant variations between the lakes with and without cyanobacterial bloom were found (Figure 3). Moreover, significant functional differences were found between the surface and bottom samples from the Salina Preta (Fig. 2B). These differences are illustrated in the metabolism of carbohydrates ( $p=0.034$ ) (mainly the fermentative and biodegradation processes), which showed high relative abundances in the bottom samples; and protein ( $p=0.010$ ), RNA ( $p=0.024$ ) and cell wall and capsule ( $p=0.046$ ) with

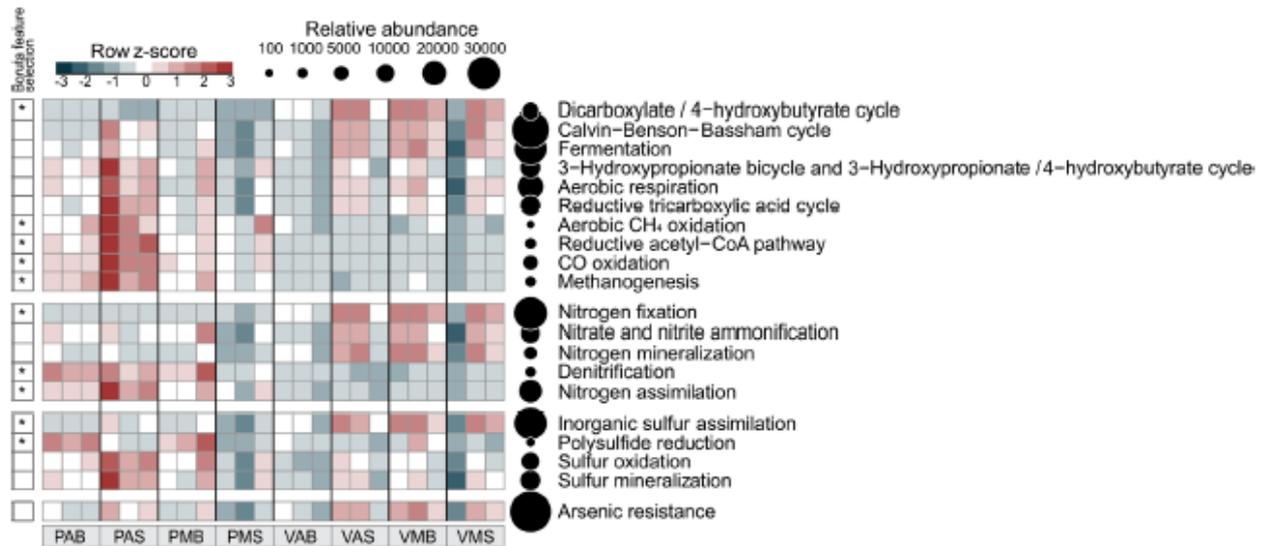
high relative abundances in the surface samples. No significant functional differences were observed for the two periods sampled within each lake.



**FIGURE 2 |** Functional distribution of the metagenomes annotated using the SEED subsystems. Statistically significant differences were found between Salina Preta and Salina Verde lakes (A), and between the bottom and surface water layers in Salina Preta (B) (see section "Results" for details).

The main metabolism of carbon pathway mapped in the Salina Verde was oxygenic photosynthesis (Calvin cycle), due to the dominant presence of the photoautotroph genus *Anabaenopsis*. In addition, the six well-characterized microbial carbon dioxide-fixation pathways were found in both lakes (Supplementary Figure 1), with some differences in terms of relative abundances between the two

lakes, as follows: the reductive citric acid cycle (major catabolic pathway in aerobic organisms); reductive acetyl CoA pathway (in anaerobic bacteria [acetogens] and archaea [methanogens]); 3-hydroxypropionate-bicycle (in green non-sulfur bacteria); and the archaean exclusives 3-hydroxypropionate/4-hydroxybutyrate and dicarboxylate/4-hydroxybutyrate cycles. Furthermore, the fermentative, methanogenic and methanotrophic processes were also mapped across these samples (see Figure 4 for details).



**FIGURE 3 |** Distribution of functional annotations involved in the C, N, S, and As cycles. The heatmap displays the relative abundance (row z-scores) of functional annotations across all samples (triplicate per site). \*Functional annotations that differentially segregated across sites were identified by random forest analysis with Boruta feature selection (average z-scores of 1,000 runs > 4). Circles are proportional to the relative abundance of each functional annotation in all samples. V: Salina Verde, P: Salina Preta, M: morning, A: afternoon, S: surface, and B: bottom.

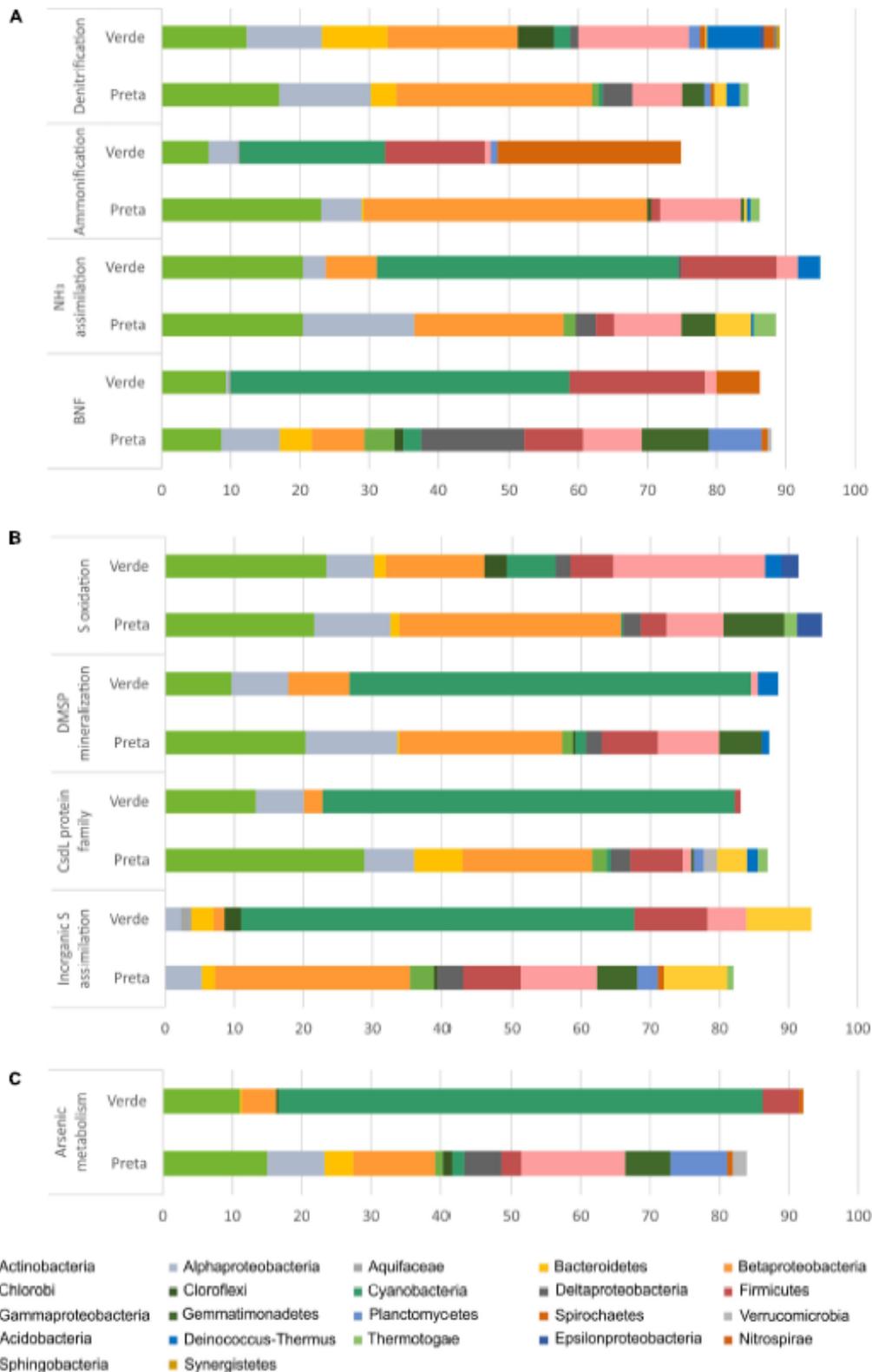
For the nitrogen (N) metabolism, most of the reads were related to the metabolic income of N via biological nitrogen fixation (BNF) (Supplementary Figure 3). The BNF process was higher in terms of relative abundance in the Salina Verde ( $p=1e^{-15}$ ), with the dominance of the diazotrophic *Anabaenopsis* (encompassing a total of 45% of the BNF affiliated sequences) (Supplementary Table 2). The other less abundant diazotrophic members were affiliated to Thermoanaerobacterales, Actinomycetales, and Spirochaetales (encompassing ca. 30% of the BNF affiliated sequences). In the Salina Preta, BNF was mostly affiliated with the members of the orders Gemmatimonadales (ca. 9%), Myxococcales, Syntrophobacterales and Bacillales (ca. 7% each).

The potential for allantoin utilization and cyanate hydrolysis were also found in the two lakes. Both of these processes were present at higher relative abundances in the Salina Preta comparatively to the Salina Verde ( $p=1.47e^{-12}$  and  $p=5.07e^{-9}$ , respectively).

The ammonia assimilation was the second most abundant N metabolism, being higher in the Salina Preta ( $p=3.96e^{-11}$ ). For this lake, the majority of the sequences were affiliated to members of the orders Acidimicrobiales (15%), Burkholderiales (11%) and Rhizobiales (10%). In the Salina Verde, the majority of the sequences assigned to this metabolism were affiliated to the Cyanobacteria (45%), mostly *Anabaenopsis* (19%) and *Arthrospira* (13%), followed by sequences affiliated to Actinomycetales (18%) and Thermoanaerobacterales (11%).

The nucleotide sequences coding for the nitrate and nitrite ammonification (mineralization process) were mostly affiliated to Spirochaetales (ca. 25%), Chroococcales and Nostocales (Cyanobacteria) (~20%) and to the Archaea domain (ca. 15%) in the Salina Verde. No sequences of the genera *Anabaenopsis* and *Arthrospira* were found associated with the ammonification process. In the Salina Preta, this process was mostly affiliated to Burkholderiales (ca. 25%) and Bifidobacteriales (ca. 15%). In addition, a higher potential for nitrate and nitrite ammonification was observed in the bottom of

this lake ( $p=0.029$ ). Some minor abundant taxa (such as Coriobacteriales, Desulfomonadales and Myxococcales) were found exclusively at the bottom samples from this lake.



**FIGURE 4 |** Relative abundances (%) of the main bacterial phyla/classes in Salina Preta and Salina Verde lakes linked to Nitrogen (A), Sulfur (B), and Arsenic (C) metabolisms. Data obtained using the SEED genomic approach and the FOCUS tool. This information at order level can be seen in Supplementary Table 2.

The potential for denitrification was low in both lakes, being slightly higher in the Salina Preta ( $p=2.55e^{-9}$ ). Most of the reads were affiliated to Burkholderiales (ca. 16%) in the Salina Preta and Actinomycetales (ca. 15%) in the Salina Verde. Additionally, sequences retrieved from the bottom

samples of both lakes and associated with denitrification were affiliated to the domain Archaea (ca. 5 – 11%).

For the sulfur (S) pathway (Supplementary Figure 4), the highest percentage of the reads matched with inorganic S assimilation and CsdL protein family (sulfur acceptors) (Supplementary Figure 5), and both were higher in relative abundances in the Salina Verde compared to the Salina Preta ( $p=1.07e^{-11}$  and  $p=6.07e^{-8}$ , respectively). For the Salina Verde, the majority of the reads related to these metabolisms (ca. 40 and 36%, respectively) were affiliated to the bloom-forming *Anabaenopsis*. Also, 8% of the reads affiliated to inorganic S assimilation and 19% of the reads affiliated to CsdL protein family were assigned to the genus *Arthrospira*. In the Salina Preta, these metabolisms were affiliated to a set of different bacterial taxa. In addition, archaeal members performing inorganic S assimilation were found solely in this lake (ca. 5%).

The cellular transporters of taurine, glutathione, and alkanesulfonate were found in both lakes, suggesting these molecules as additional sulfur sources in these soda lakes.

The dimethylsulfoniopropionate (DMSP) mineralization (demethylation pathway) was the third most abundant metabolism of S cycling in these systems. Cyanobacteria members were identified as the main responsible for this process in the Salina Verde (>50%), with reads affiliated to *Anabaenopsis* and *Arthrospira*. In the Salina Preta, this metabolism was affiliated to other bacterial taxa, being the majority affiliated to Burkholderiales (44%).

Sulfur oxidation was mapped in both lakes and affiliated to a particular set of taxa. In the Salina Verde, sequences were mostly affiliated to Actinomycetales (18%), whereas in the Salina Preta, sequences affiliated to Burkholderiales (29%). The Rhodospirillales members were abundant across all the samples (ca. 4.7 – 10.2%) and assigned as potentially important taxa involved in S oxidation in these systems.

The genetic potential for arsenic resistance was mapped in both lakes, being relatively higher in the Salina Verde ( $p=2.95e^{-14}$ ). Genes encoding arsenate reductase (ArsC), cellular extrusion (ArsA, ArsB and ACR3) and arsenic resistance (ArsD and ArsH) were mapped and presented different relative abundances between the two lakes (Supplementary Figure 6). Bacteria and Archaea members were found as potential arsenic metabolizers (Salina Verde: 98% Bacteria and 2% Archaea; Salina Preta: 93% Bacteria and 7% Archaea). The main taxa affiliated with this resistance mechanism in the Salina Verde was *Anabaenopsis* (ca. 50%), followed by Actinomycetales (11%) Burkholderiales, Clostridiales, Chroococcales (total of 5%) and Oscillatoriales (ca. 3%). In the Salina Preta, the taxonomic diversity of potential arsenic metabolizers was higher and reads were mostly affiliated to the orders Burkholderiales, Planctomycetales, Pseudomonadales and Actinomycetales (total of ca. 8%).

## Discussion

In this study, two soda lakes, Salina Preta (black water lake) and Salina Verde (green water lake - with continuous cyanobacteria bloom), were investigated in terms of microbial community structures and their potential functional contributions to biogeochemical cycles. The results here obtained have shown that community composition and relative abundance patterns of organisms are different in the two lakes. The higher bacterial relative abundance in the Salina Verde was due to the intensive cyanobacteria proliferation with dominance of the species *Anabaenopsis elenkinii*. This date corroborated previous studies showing the dominance of *A. elenkinii* in certain Pantanal lakes (Santos and Sant'anna, 2010; Costa et al., 2016). The genus *Anabaenopsis* comprises a planktonic species with a known distribution in soda lakes located in tropical regions (Ballot et al., 2005, 2008; Komárek,

2005). Besides the autotrophic lifestyle, *Anabaenopsis* also is capable of biological nitrogen fixation, which confers an adaptive advantage to this genus. Cyanobacteria seem to play a major role in the geochemistry of Pantanal lakes, but determining the cause of the persistent cyanobacterial blooms in this hypersaline water is a daunting task. The intensive photosynthesis of cyanobacterial blooms results in a sharp increase in pH due to the consumption of CO<sub>2</sub> and HCO<sub>3</sub><sup>-</sup> and the consequent increase in CO<sub>3</sub><sup>2-</sup> (Schneider and Le Campion-Alsumard, 1999; Almeida et al., 2011). The elevation of pH is known to enhance nutrient release from bottom sediments, thus impacting nutrient biogeochemical processes in the system (Gao et al., 2012). For instance, these authors stated that enhanced nutrient release from bottom sediments, high pH and alkalinity can boost the growth and persistence of cyanobacterial blooms in shallow water ecosystems. Both pH and alkalinity were high in the Salina Verde with persistent *Anabaenopsis* bloom and both parameters were statistically correlated to the cyanobacteria. Among the nutrient elements required for cyanobacteria, phosphorus and nitrogen are often the growth-limited ones (Paerl, 2008).

Despite the dominance of *Anabaenopsis* in the Salina Verde, the *in silico* reconstruction of the metabolic pathways of N and S cycles (i.e. ammonia assimilation and inorganic S assimilation, which supply N and S for biosynthesis) was linked to *Arthrospira* genus, suggesting that *Anabaenopsis* and *Arthrospira* genera coexist in this lake. Furthermore, 16S rRNA gene fragments recovered from metagenomic data showed sequences matching *Arthrospira* and spiral-shaped filaments typically of this genus were visualized in water samples examined under optical microscopy (data not shown). These findings corroborated a previous study conducted in other green water lakes (Salina do Meio) showing the presence of *Arthrospira* populations together with *Anabaenopsis* bloom (Santos and Sant'Anna, 2010); a finding also reported in African soda lakes (Ballot et al., 2008).

The Salina Preta, with no record of microbial blooms, has black waters (turbidity around 90 NTU) due to the presence of clay and high amount of dissolved organic matter and almost constant DO around 7 mg·L<sup>-1</sup>. Disregarding the presence of cyanobacteria bloom, the Salina Verde presents crystalline waters, but the bloom causes an increase in turbidity due to the high cell density and a high variation of DO. In this lake, oxygen concentration gradually increases throughout the day as a result of oxygenic photosynthesis before decrease at night as a result of microbial respiration, thus altering water abiotic parameters and consequently affecting the coexisting microbial taxa.

A connection between differences in the community composition and the functional potential of the two lakes investigated was found. Although, overall, the same metabolic pathways were mapped on the SEED database, the abundance of almost all the metabolic process varied between Salina Verde and Salina Preta. This data suggest that the extreme conditions of these environments lead functional adaptations allowing the two lakes to work in a similar metabolic 'core' manner. Nevertheless, the dominance of one species in a high proliferative state in Salina Verde elevated, for example, the processes involved in cell proliferation, DNA replication and nitrogen and sulfur assimilation. The distribution of metabolic functions in the Salina Preta community encompassed a large number of taxa, whereas, in Salina Verde, the functionality was restrained across a specific set of existing taxa. The dominance of cyanobacteria bloom derived by the environment selective pressure reduces the community diversity, which can be ecologically linked to a decrease in community resilience (Shade et al., 2012). Moreover, the highly complex microbial community of the Salina Preta is more structurally distributed in the surface and bottom layers even in this shallow lake, displaying taxonomic and functional differences. However, the non-stratification of the microbial community in the Salina Verde could be a consequence of cyanobacteria migration through the water column and the oxygen saturation, which revolve the water column, preventing any stratification profile.

Nitrogen has been found as a limiting nutrient in Brazilian soda lakes due to the low nitrogen/phosphorous ratio that reduces the bioavailability of nitrogenous compounds in these systems (Mourão, 1989). In this study, the cyanobacteria bloom was related to a high BNF potential

in the Salina Verde, and, in spite of the mechanisms of nitrogen fixation in soda lakes under oxic conditions are unknown, several bacterial taxa of the Salina Preta were linked to this metabolic process. The total nitrogen content in these two lakes was previously quantified, revealing a ten-fold higher value in Salina Verde (Machado, 2016). *Anabaenopsis* bloom certainly is an important input of nitrogen to the Salina Verde through BNF. Additionally to the BNF, the N bio-availability by mineralization, together to the potential for allantoin utilization (nitrogen source in anaerobic conditions) and cyanate hydrolysis (nitrogen source in alkaline conditions) conferred an additional N supply in these lakes.

The conversion of ammonia to nitrate ( $\text{NO}_3^-$ ) and nitrite ( $\text{NO}_2^-$ ) (nitrification) was not detected in the metagenomic dataset. However, ammonia, nitrite, and nitrate assimilation processes were abundant in both lakes. It is known that nitrification can decrease or cease at high salt concentrations (Sorokin, 1998), and that the lakes with high salinity can have an external source of  $\text{NO}_2^-$  and  $\text{NO}_3^-$  (Oren, 2001). The nitrogen output by denitrification was poorly mapped since it is a metabolic process restrict to anoxic environments, which could be occurring in the sediments of these lakes. Nevertheless, low nitrous oxide ( $\text{N}_2\text{O}$ ) emission, a product of denitrification, and also an important greenhouse gas, was recently reported in the Salina Verde and Salina Preta lakes (Machado, 2016).

The functional potential of sulfur metabolism was detected in both lakes. This was mostly limited to the oxidative part of the sulfur cycle, driven by chemolithoautotrophic haloalkaliphilic sulfur-oxidizing bacteria, like the S mineralization by Bacteria using the demethylation pathway of DMSP. Some genes related to polysulfide reduction were also mapped, although the lake sediments were not evaluated in this study. The sulfur cycle is one of the most active in soda lakes (Sorokin et al., 2011) and the reduction part occurs in the anoxic layers of the sediments of soda lakes (Sorokin et al., 2014).

Soda lakes are among the most productive natural ecosystems in the world, active on atmospheric carbon sequestration, but also on the production and emission of greenhouse gases (GHG) (Paul Antony, et al. 2013). The role in the GHG emission of Salina Preta and Salina Verde is totally different. In the Pantanal, lakes with cyanobacteria bloom are considered as  $\text{CO}_2$  sinks, reaching a dissolved organic carbon (DOC) content around ten times higher than the black water lakes, which are considered  $\text{CO}_2$  sources (Machado, 2016; Barbiero et al., 2017). In the two studied Brazilian soda lakes, the six pathways of atmospheric  $\text{CO}_2$  fixation (Sato and Atomi, 2010) were mapped, illustrating a diverse microbial composition potentially involved in these processes. As expected, the Calvin-Benson-Bassham cycle (performed by photosynthetic bacteria) was highly abundant in the Salina Verde, suggesting the cyanobacteria bloom as responsible for the major  $\text{CO}_2$  fixation in this lake.

Cyanobacteria bloom may be an important event for the dynamic of carbon sequestration and bio-availability in the studied soda lakes. Considering the methane metabolism, both processes methanotrophy (aerobic  $\text{CH}_4$  oxidation) and methanogenesis (anaerobic process preferentially active on sediments and performed by Archaea members) were found in the metagenomes. Salina Verde and Salina Preta are both considered  $\text{CH}_4$  sources, with Salina Verde presenting a high  $\text{CH}_4$  emission than Salina Preta (Barbiero et al., 2017). In the Salina Verde, due to the photosynthetic and respiratory activity, the percentage of dissolved oxygen varied from 7 during the night to 600 during the day, this shift leads to an anaerobic condition allowing methanogenesis during the night, and a bubble in the water due to the oxygen saturation, causing a release of the  $\text{CH}_4$  during the day. These two processes may explain the high  $\text{CH}_4$  emission in this lake. Is important noticing that high abundances of methanogenic taxa can explain the high  $\text{CH}_4$  content in green-waters, however, methanogenesis process can be properly evaluated in sediments. Additionally, in saline and alkaline lakes methanotrophs being responsible by a significant  $\text{CH}_4$  consume (Joye et al., 1999; Khmelenina et al., 2000).

Several microbial groups have the potential for arsenic resistance, including those inhabiting soda lakes. The water of these lakes presents arsenic (As) content values ranging from 0.05 to 0.5 mg·L<sup>-1</sup>, usually higher in the Salina Verde, mainly in the form of As(III) in the sediments and As(V) in the waters (Barbiero et al., 2007). The dominance of *Anabaenopsis* in the lake with higher As and the high number of reads linked to arsenic metabolism affiliated to Cyanobacteria suggest this genus as a potentially important As metabolizer in this environment. Information on cyanobacterial strains performing As resistance metabolism is still elusive. However, in soda lakes, Cyanobacteria can have a direct role in the As cycle, for instance by carrying out resistance-based (ArsC) reduction of As(V) to As(III) or using As(III) as an electron donor to sustain anoxygenic photosynthesis (Kulp et al., 2006).

This study is the first report exploring the microbial community composition and functional pathways in Brazilian soda lakes. The metagenomics approach unveiled the relations among distinct microbial taxa and their potential role in the biogeochemical transformations of C, N and S, in a comparative framework that included two soda lakes (Salina Verde and Salina Preta). Our data provided valuable insights into the metabolisms responsible for the maintenance of nutrient and life metabolic processes taking place in these extreme environments. Future efforts should be made in investigating the taxonomic and functional profiles of additional soda lakes, with and without bloom, including their seasonal variations. In addition, sediments of these lakes should be taken into consideration to assess important anoxic processes. Thus, collectively contributing for a comprehensive overview of the biogeochemical cycling in these systems.

#### **Nucleotide sequence accession number**

All sequence data (24 libraries) have been deposited in the MG-RAST database under the project named Pantanal and accession numbers: mgm4575354.3 to mgm4575377.3.

#### **Author Contributions**

MFF conceived the study. LB and ATRF collected the samples. GMS and JR performed the experiments. APDA, JR, GMS, FDA and BCES analyzed data. All authors were involved in writing the paper and had final approval of the manuscript.

#### **Funding**

This work was supported by the São Paulo Research Foundation (FAPESP n°2013/50425-8) to MFF, FAPESP (2013/09192-0) to LB and by the Federal University of Mato Grosso do Sul (PROPP n°2011/0338) to ATRF. GMS was supported by FAPESP graduate scholarship (2013/20142-4). JR received a postdoctoral fellowship from the Brazilian Federal Agency for the Support and Evaluation of Graduate Education (CAPES- PNP20131744 USP/CENA program). BCES was supported by CAPES graduate scholarship (1727082). LB was supported by grants awarded by São Paulo University (USP), Campinas University (UNICAMP) and the French Consulate in São Paulo. MFF thanks Brazilian National Research Council (CNPq) for research fellowship (#310244/2015-3).

#### **Acknowledgments**

We wish to thank the Center of Functional Genomics Applied to Agriculture and Agroenergy (USP, Campus “Luiz de Queiroz”) for generating Illumina MiSeq data. We thank the owner of the Centenário farm for permission to collect the water samples and D.B. Genuário and M. G. M. V. Vaz for the samples collection logistic support.

## Supplementary Material

The Supplementary Material for this article can be found online at:

## References

- Almeida, T. I. R., Calijuri, M. C., Falco, P. B., Casali, S. P., Kupriyanova, E., Paranhos Filho et al. (2011) Biogeochemical process and the diversity of Nhecolândia lakes, Brazil. *An. Acad. Bras. Ciênc.* 83, 391-407. <http://dx.doi.org/10.1590/S0001-37652011000200004>.
- Andreote, A. P. D., Vaz, M. G. M. V., Genuário, D. B., Barbiero, L., Rezende-Filho, A. T., Fiore, M. F. (2014). Nonheterocytous cyanobacteria from Brazilian saline-alkaline lakes. *J. Phycol.* 50, 675-684. doi:10.1111/jpy.12192.
- APHA – AMERICAN PUBLIC HEALTH ASSOCIATION. (1995). Standard methods for the examination of water and wastewater. Springfield, USA: Byrd Prepress.
- Ballot, A., Dadheech, P. K., Haande, S., Krienitz, L. (2008). Morphological and Phylogenetic Analysis of *Anabaenopsis abijatae* and *Anabaenopsis elenkinii* (Nostocales, Cyanobacteria) from Tropical Inland Water Bodies. *Microb. Ecol.* 55, 608-618. doi:10.1007/s00248-007-9304-4.
- Ballot, A., Krienitz, L., Kotut, K., Wiegand, C., Pflugmacher, S. (2005). Cyanobacteria and cyanobacterial toxins in the alkaline crater lakes Sonachi and Simbi, Kenya. *Harmful Algae* 4, 139-150. doi:10.1016/j.hal.2004.01.001.
- Barbiero, L., Berger, G., Rezende Filho, A. T., Meunier, J.-F., Martins-Silva, E. R., Furian, S. (2016). Organic control of dioctahedral and trioctahedral clay formation in an alkaline soil system in the Pantanal wetland of Nhecolândia, Brazil. *PLoS ONE* 11. doi:10.1371/journal.pone.0159972.
- Barbiero, L., Furquim, S., Valles, V., Furian, S., Sakamoto, A., Filho, A. et al. (2007). Natural arsenic in groundwater and alkaline lakes at the upper Paraguay basin, Pantanal, Brazil. *Tr. Met. Env.* 101-126. doi 10.1016/S0927-5215(06)09004-7.
- Barbiero, L., Rezende Filho, A., Furquim, S. A. C., Furian, S., Sakamoto, A. Y., Valles, V. et al. (2008). Soil morphological control on saline and freshwater lake hydrogeochemistry in the Pantanal of Nhecolândia, Brazil. *Geoderma* 148, 91-106. doi: 10.1016/j.geoderma.2008.09.010.
- Barbiero, L., Siqueira Neto, M., Braz, R. R., Carmo, J. B., Rezende Filho, A. T., Mazzi et al. (2017). Biogeochemical diversity and hot moments of GHG emissions from shallow alkaline lakes in the Pantanal of Nhecolândia, Brazil. *Biogeosci. Discuss.* <https://doi.org/10.5194/bg-2017-108>.
- Breiman, L. (2001). Random Forests. *Mach. Learn.* 45, 5-32. doi:10.1023/A:1010933404324.
- Buchfink, B., Xie, C., Huson, D. H. (2014). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59-60. doi:10.1038/nmeth.3176.
- Costa, N. B., Kolman, M. A., Giani, A. (2016). Cyanobacteria diversity in alkaline saline lakes in the Brazilian Pantanal wetland: a polyphasic approach. *J. Plankton Res.* 38, 1389-1403. doi:10.1093/plankt/fbw066.
- Dalgaard, P. (2008). *Introductory Statistics with R*. Copenhagen, Denmark: Springer Verlag.
- De-Lamonica-Freire, E. M. and Heckman, C. W. (1996). The Seasonal Succession of Biotic Communities in Wetlands of the Tropical Wet-and-Dry Climatic Zone: III. The Algal Communities in the Pantanal of Mato Grosso, Brazil, with a Comprehensive List of the Known Species and Revision of two Desmid Taxa. *Int. Rev. Gesamten Hydrobiol. Hydrogr.* 81, 253–280. doi:10.1002/iroh.19960810209.
- Domagalski, J. L., Orem, W. H. and Eugster, H. P. (1989). Organic geochemistry and brine composition in Great Salt, Mono, and Walker Lakes. *Geochim. Cosmochim. Acta.* 53, 2857-2872. doi.org/10.1016/0016-7037(89)90163-4.
- Duarte, R.T.D., Nóbrega, F., Nakayama, C.R. Pellizari, V.H. (2012). Brazilian research on extremophiles in the context of astrobiology. *Int. J. Astrobiol.*, 1-9. doi:10.1017/S1473550412000249.
- Eugster, H. P. and Jones, B. F. (1979). Behavior of major solutes during closed-basin brine evolution. *Am. J. Sci.* 279, 609-31. doi: 10.2475/ajs.279.6.609.

- Felföldi, T., Somogyi, B., Márialigeti, K., Vörös, L. (2009). Characterization of photoautotrophic picoplankton assemblages in turbid, alkaline lakes of the Carpathian Basin (Central Europe). *J. Limnol.* 68, 385–395. doi: 10.3274/JL09-68-2-21.
- Furian, S., Martins, E. R. C., Parizotto, T. M., Rezende-Filho, A. T., Victoria, R. L., Barbiero, L. (2013). Chemical diversity and spatial variability in myriad lakes in Nhecolândia in the Pantanal wetlands of Brazil. *Limnol. Oceanogr.* 58, 2249–2261. doi:10.4319/lo.2013.58.6.2249.
- Furquim, S. A. C., Barbiero, L., Graham, R. C., de Queiroz Neto, J. P., Ferreira, R. P. D., Furian, S. (2010a). Neof ormation of micas in soils surrounding an alkaline-saline lake of Pantanal wetland, Brazil. *Geoderma* 158, 331–342. <https://doi.org/10.1016/j.geoderma.2010.05.015>.
- Furquim, S. A. C., Graham, R. C., Barbiero, L., Queiroz Neto, J. P., Vidal-Torrado, P. (2010b). Soil mineral genesis and distribution in a saline lake landscape of the Pantanal Wetland, Brazil. *Geoderma* 154, 518–528. <https://doi.org/10.1016/j.geoderma.2009.03.014>.
- Gao, Y., Cornwell, J. C., Stoecker, D. K., Owens, M. S. (2012). Effects of cyanobacterial-driven pH increases on sediment nutrient fluxes and coupled nitrification-denitrification in a shallow freshwater estuary. *Biogeosciences* 9, 2697–2710. doi:10.5194/bg-9-2697-2012.
- Genuário, D. B., Andreote, A. P. D., Vaz, M. G. M. V., Fiore, M. F. (2017). Heterocyte-forming cyanobacteria from Brazilian saline-alkaline lakes. *Mol. Phylogenet. Evol.* 109, 105–112. doi:10.1016/j.ympev.2016.12.032.
- Grant, W. D. (2006). “Alkaline environments and biodiversity”, in *Extremophilies*, Encyclopedia of Life Support Systems (EOLSS), eds. Gerday, C. and Glansdorff, N. (Oxford, UK: Eolss Publishers), [<http://www.eolss.net>].
- Hammer, U. T. (1986). *Saline lake ecosystems of the world*, Netherlands, Dordrecht: Springer.
- Hammer, U. T. (1981). “Primary production in saline lakes”, in *Salt Lakes*, ed. Williams, W. D. (Dordrecht, Netherlands: Springer), 47-57.
- Haynes, W. (2013). “Benjamini–Hochberg Method,” in *Encyclopedia of Systems Biology*, eds. Dubitzky, W., Wolkenhauer, O., Cho, K.-H., Yokota, H. (New York, NY: Springer), 78–78.
- Huson, D. H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., et al. (2016). MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. *PLOS Comput. Biol.* 12, e1004957. doi: 10.1371/journal.pcbi.1004957.
- Jones, B. F., Eugster, H. P., Rettig, S. L. (1977). Hydrochemistry of the lake Magadi basin, Kenya. *Geochim. Cosmochim. Ac.* 41, 53–72. [https://doi.org/10.1016/0016-7037\(77\)90186-7](https://doi.org/10.1016/0016-7037(77)90186-7).
- Jones, B. E., Grant, W. D., Duckworth, A. W., Owenson, G. G. (1998). Microbial diversity of soda lakes. *Extremophiles* 2, 191-200.
- Joye, S. B., Connell, T. L., Miller, L. G., Oremland, R. S., Jellison, R. S. (1999) Oxidation of ammonia and methane in an alkaline, saline lake. *Limnol. Oceanogr.* 44, 178-188.
- Kanehisa, M. and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30. <https://doi.org/10.1093/nar/28.1.27>.
- Kazmierczak, J., Kempe, S., Kremer, B., Lopez-Garcia, P., Moreira, D., Tavera, R. (2011). Hydrochemistry and microbialites of the alkaline crater lake. *Facies.* 57, 543–570. doi: 10.1007/s10347-010-0255-8.
- Khmelenina, V. N., Eshinimaev, B., Kalyuzhnaya, M. G., Trotsenko, Y.A. (2000). Potential activity of methane and ammonium oxidation by methanotrophic communities from the soda lakes of Southern Transbaikal. *Mikrobiologiya* (English translation) 69, 460-465.
- Komárek, J. and Anagnostidis, K. (1989). Modern approach to the classification system of cyanophytes. Nostocales 4. *Algol. Stud.* 56, 247-345.
- Komárek, J. (2005). Phenotype diversity of the heterocytous cyanoprokaryotic genus *Anabaenopsis*. *Fottea* 5, 1-35.
- Krienitz, L., Dadheech, P. K., Kotut, K. (2013). Mass developments of the cyanobacteria *Anabaenopsis* and *Cyanospira* (Nostocales) in the soda lakes of Kenya: ecological and systematic implications. *Hydrobiologia.* 703, 79-93. doi: 10.1007/s10750-012-1346-z.
- Kulp, T. R., Hoefl, S. E., Miller, L. G., Saltikov, C., Murphy, J. N., Han, S., et al. (2006). Dissimilatory arsenate and sulfate reduction in sediments of two hypersaline, arsenic-rich soda lakes:

Mono and Searles lakes, California. *Appl. Environ. Microbiol.* 72, 6514-6526. doi:10.1128/AEM.01066-06.

Kursa, M. B. and Rudnicki, W. R. (2010). Feature Selection with the Boruta Package. *J. Stat. Softw.* 36, 1-13. doi:10.18637/jss.v036.i11.

Ma, J. and Edmunds, W. M. (2006). Groundwater and lake evolution in the Badain Jaran Desert ecosystem, Inner Mongolia. *Hydrogeol. J.* 14, 1231-1243. doi: 10.1007/s10040-006-0045-0.

Machado, R. R. B. (2016) Estimativa de emissão de gases de efeito estufa por lagoas salinas no pantanal da Nhecolândia (MS). [doctoral thesis]. [Piracicaba (SP)]: University of São Paulo.

Martins, E. R. C. (2012). Tipologia de lagoas salinas no Pantanal da Nhecolândia (MS). [doctoral thesis]. [Piracicaba (SP)]: University of São Paulo.

Medina-Júnior, P. B. and Rietzler, A. C. (2005). Limnological study of a Pantanal saline lake. *Braz. J. Biol.* 65, 651-659. doi:10.1590/S1519-69842005000400013.

Melack, J. M. (1981). Photosynthetic activity of phytoplankton in tropical African soda lakes. *Hydrobiologia* 81, 71-85.

Melack, J. M. and Kilham, P. (1974). Photosynthetic rates of phytoplankton in East African alkaline, saline lakes. *Limnol. Oceanogr.* 19, 743-755. doi: 10.4319/lo.1974.19.5.0743.

Mourão, G. M. (1989). Limnologia comparativa de três lagoas (duas “baías” e uma “salina”) do Pantanal da Nhecolândia (MS). [masters dissertation]. [São Carlos (SP)]: Federal University of São Carlos.

Oliveira, M. D., and Calheiros, D. F. (2000). Flood pulse influence on phytoplankton communities of the south Pantanal floodplain, Brazil. *Hydrobiologia* 427, 101-112. doi:10.1023/A:1003951930525.

Oren, A. (2001). The bioenergetic basis for the decrease in metabolic diversity at increasing salt concentrations: implications for the functioning of salt lake ecosystems. *Hydrobiologia* 466, 61-72. doi: 10.1023/A:1014557116838.

Overbeek, R. (2005). The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res.* 33, 5691-5702. doi:10.1093/nar/gki866.

Paerl, H. (2008). “Nutrient and other environmental controls of harmful cyanobacterial blooms along the freshwater-marine continuum”, in: *Cyanobacterial Harmful Algal Blooms: State of the Science and Research Needs*, ed. Hudnell, K. (North Carolina, USA: Springer), 217-237.

Parks, D. H., Tyson, G. W., Hugenholtz, P., Beiko, R. G. (2014). STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics* 30, 3123-3124. doi:10.1093/bioinformatics/btu494.

Paul Antony, C., Kumaresan, D., Hunger, S., Drake, H. L., Murrell, J. C., Shouche, Y. S. (2013). Microbiology of Lonar Lake and other soda lakes. *ISME J.* 7, 468-476. doi:10.1038/ismej.2012.137.

Rezende Filho, A. T., Furian, S., Victoria, R. L., Mascré, C., Valles, V., Barbiero, L. (2012). Hydrochemical variability at the Upper Paraguay Basin and Pantanal wetland. *Hydrol. Earth Syst. Sci.* 16, 2723-2737. doi:10.5194/hess-16-2723-2012.

Santos, K. R. S. and Sant’anna, C. L. (2010). Cianobactérias de diferentes tipos de lagoas (“salina”, “salitrada” e “baía”) representativas do Pantanal da Nhecolândia, MS, Brasil. *Braz. J. Bot.* 33, 61-83. doi:10.1590/S0100-84042010000100007.

Sato, T. and Atomi, H. (2010). “Microbial Inorganic Carbon Fixation,” in *Encyclopedia of Life Sciences*. (Chichester, UK: John Wiley & Sons, Ltd). doi: 10.1002/9780470015902.a0021900.

Schneider, J. and Le Campion-Alsumard, T. (1999). Construction and destruction of carbonates by marine and freshwater cyanobacteria, *Eur. J. Phycol.* 34, 417-426, doi: 10.1080/09670269910001736472.

Shade, A., Peter, H., Allison, S. D., Baho, D. L., Berga, M., Bürgmann, H. et al. (2012). Fundamentals of microbial community resistance and resilience. *Front. Microbiol.*, 19, 1-19. <https://doi.org/10.3389/fmicb.2012.00417>.

Silva, G. G. Z., Cuevas, D. A., Dutilh, B. E., Edwards, R. A. (2014). FOCUS: an alignment-free model to identify organisms in metagenomes using non-negative least squares. *Peer J.* 2:e425. doi:10.7717/peerj.425.

- Sorokin, D. Y., Banciu, H. L., Muyzer, G. (2015). Functional microbiology of soda lakes. *Curr. Opin. Microbiol.*, 25, 88-96. doi:10.1016/j.mib.2015.05.004.
- Sorokin, D. Y., Berben, T., Melton, E. D., Overmars, L., Vavourakis, C. D., Muyzer, G. (2014). Microbial diversity and biogeochemical cycling in soda lakes. *Extremophiles*. 18, 791-809. doi 10.1007/s00792-014-0670-9.
- Sorokin, D. Y., Kuenen, J. G., Muyzer, G. (2011). The microbial sulfur cycle at extremely haloalkaline conditions of soda lakes. *Front. Microbiol.* 2. doi:10.3389/fmicb.2011.00044.
- Sorokin, D.Y. (1998) Occurrence of nitrification in extremely alkaline natural habitats. *Microbiology* 67, 404-408.
- Stenger-Kovács, C., Lengyel, E., Buczkó, K., Tóth, F. M., Crossetti, L. O., Doma, Z. Z. et al. (2014). Vanishing world : alkaline, saline lakes in Central Europe and their diatom assemblages. *Inl. Waters*. 4, 383-396. doi: 10.5268/IW-4.4.722.
- TARIFA, J. R. (1986). O sistema climático no Pantanal: Da compreensão do sistema a definição de prioridade de pesquisa climatológica. In: EMBRAPA/DDT, p. 9–27. Anais do Simpósio sobre Recursos Naturais e Socio-econômicos do Pantanal. Brazilian Corporation for Agricultural Research. [The climate system in the Pantanal: From understanding of the system to defining climatological research priority.]
- Vaz, M. G. M. V., Genuário, D. B., Andreote, A. P. D., Malone, C. F. S., Sant’Anna, C. L., Barbiero, L. et al. (2015). *Pantanalinema* gen. nov. and *Alkalinema* gen. nov.: novel pseudanabaenacean genera (Cyanobacteria) isolated from saline-alkaline lakes. *Int. J. Syst. Evol. Microbiol.* 65, 298-308. doi:10.1099/ijs.0.070110-0.
- Ventosa, A. and Arahál, D.R. (1989). “Physico-chemical characteristics of hypersaline environments and their biodiversity”, in *Extremophiles: Encyclopedia of Life Support Systems (EOLSS)*, eds. Gerday C. and Glansdorff, N. (Liege, Belgium: Eolss Publishers), [<http://www.eolss.net>].
- Zhang, J., Kobert, K., Flouri, T., Stamatakis, A. (2014). PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* 30, 614-620. doi:10.1093/bioinformatics/btt593.