Bulgarian Chemical Communications, Special Issue F, (pp. 60 – 68) 2016

Prokaryotic pico- and nanoplankton community structure in the hypersaline lakes of the Qaidam Basin

L. Chu[1, 2], J. Wang[1, 3], F. Wang[1*], M. Chen[1], Z. Liu[4]

[1]*State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research, Beijing ,100038, China*
[2]*College of Hydrology and Water Resource, Hohai University, Nanjing ,210098, China*
[3]*Institute of Shandong River Wetlands, Laiwu ,271100, China*
[4]*Institute of Microbiology, Chinese Academy of Sciences, Beijing, 100101, China*

Received July 13, 2016; Revised December 13, 2016

Pico- and nanoplankton were the main drivers of the biogeochemical cycles over large areas of the world's waters, play key ecological roles in aquatic ecosystems. We performed a comprehensive analysis of the genetic diversity (16S rDNA gene) of the prokaryotic pico- and nanoplankton communities(PPNC, a size of member in PPNC range 0.2-20 μm) and the whole prokaryotic plankton community(WPPC, a size of member in WPPC >0.2 μm) in three hypersaline lakes located in the Qaidam Basin. Most of the 16S rDNA gene sequences obtained of the PPNC in Lake Gahai were closely related to Proteobacteria phylum. The most abundant sequences of the PPNC, however, primarily represented Euryarchaeota(78.60%) in Lake Gasikule. An obvious divergence between the structure of the PPNC and the WPPC was observed. The most common group of WPPC was assigned as unclassified prokaryotes in Lake Gahai, accounting for 23.27% of the total number of sequences. Altitude, temperature and TDS(total dissolved solids) were significantly correlated with the distribution of the PPNC. This study also shows important gaps in the current knowledge about PPNC inhabiting hypersaline lakes, highlighting the need for future, more detailed investigations to develop active conservation strategies to preserve the microbial biodiversity in these areas.

Keywords: Prokaryotic pico- and nanoplankton community, Hypersaline lake, Qaidam Basin.

INTRODUCTION

Inland saline lakes represent approximately 5% of modern drylands[1]. These water bodies are numerous and distributed worldwide in semi-arid or arid areas[2]. Studies have shown that salinity plays a dominant role in regulating the composition of prokaryotic plankton in inland waters[3-4]. The bacterial communities of freshwater and hypersaline lakes show only small taxonomic overlap[5]. Soda lake environments are good examples of extreme environments[6]. Studies of soda lakes have improved our understanding of the biology of extreme environments and have resulted in the identification of novel microorganisms and enzymes(extremozymes) with potential for biotechnological applications[7]. These enzymes are stable at high pH values, salt concentrations, or temperatures[8-9].

Prokaryotic pico- and nanoplankton is the smallest size fraction of prokaryotic plankton. Picoplanktonic cell sizes range from 0.2 to 2.0 μm[10](i.e., cells that can pass through a 3-μm pore-size filter[11]. Nanoplanktons are ubiquitous protozoan zooplankton in a size range of 2 to 20 μm, play key ecological roles in aquatic ecosystems[12]. In contrast to the macroscopic organisms studied for centuries, the microscopic prokaryotes have only received adequate attention in the last forty years. A number of studies have provided evidence that picoplankton are the most abundant organisms in the ocean, often dominating the planktonic biomass and primary production. Most studies[13-14] have almost entirely focused on marine members. However, only a few genomes are available from non-marine prokaryotic picoplankton and nanoplankton.

More recently, prokaryotic picoplankton members in both the summer and the winter communities in central European hypersaline lakes have been identified using molecular biological techniques, including denaturing gradient gel electrophoresis (DGGE) and sequence analysis[15]. These molecular analyses(16S rDNA gene and cpcBA-IGS region) have identified a dominant group of picocyanobacteria belonging to the Cyanobium gracile cluster(group A) of the picophytoplankton clade in shallow alkaline lake (Lake Fertő, Neusiedlersee) in April. The bacterial community is largely dominated by halophilic and halotolerant microorganisms in Isabel lake[6]. Molecular tools in prokaryotic picoplankton have substantially increased our knowledge of microbial community structures[16]; Our current knowledge

* To whom all correspondence should be sent:
E-mail: wangf@iwhr.com

on microorganisms isolated in culture, however, does not completely represent the microbial diversity in saline systems[17-20].

The Qaidam Basin, a large intermountain depression with an arid to semi-arid continental climate, is located in the northeastern margin of the Tibetan Plateau, China [21] and is surrounded by the Qilian, Kunlun, and Aerjin mountains. This basin contains dozens of freshwater to hypersaline lakes at high elevations from 2700 to 3200 m above sea level. The hypersaline lakes located in the Qaidam Basin represent a peculiar environment, increasing the potential of identifying previously undescribed halophilic species or isolates with interesting biochemical features[22-23]. In the present study, a massive 454 tag-based sequencing approach targeting the V3, V4 and V5 region of the 16S rDNA gene was used to obtain an overview of the prokaryotic pico- and nanoplankton community(PPNC, a size of member in PPNC range 0.2-20 μm) structure in the hypersaline lakes of the Qaidam Basin. The whole prokaryotic plankton community(WPPC, a size of member in WPPC >0.2 μm) were also analyszed. We also evaluated the distribution patterns of the PPNC along gradients of salinity and other physicochemical parameters.

EXPERIMENTAL

Materials

Analytical methods

Sample collection and DNA extraction

Three lakes located in the Qaidam Basin at elevations ranging from 2,853 to 3,170 m above sea level were investigated(Figure 1; Table S1). The selected lakes covered a TDS gradient from 93.60 to 466.00 g/L. Water samples were collected from surface waters(top 30 cm) using a 5 L Schindler sampler in August 2013 and were immediately filtered through a 20-μm mesh sieve to remove most of the mesozooplankton and large particles. Plankton samples(2000–2500 ml water) for the next-generation sequencing(NGS) analyses were collected on 0.2-μm pore-size Isopore filters using a hand pump at a pressure of less than 15 mmHg. Water temperature, pH and dissolved oxygen levels were measured using a Hydrolab sensor(Austin, TX, USA). Overall, five samples were analyzed(Table S1). The concentrations of the six major ions sodium(Na^+), calcium (Ca^{2+}), magnesium(Mg^{2+}) and sulfate(SO_4^{2-}), and the concentration of total nitrogen(TN) were measured according to standard methods[24] after transporting the samples to the laboratory. The total dissolved solids(TDS) of the investigated habitats were determined using a conventional conductivity meter(Table S1). Filters for DNA extraction were stored in liquid nitrogen during the field campaign and transported to the laboratory. The DNA was extracted after the cetyltrimethylammonium bromide extraction procedure[18].

16S rDNA Gene Sequencing

PCR was performed using 454 sequencing adaptor-linked primers flanking V3, V4and V5 region of the 16S rDNA gene by GeneWiz, Inc. (Beijing, China). The quality and the quantity of DNA were examined by agarose gel electrophoresis and spectrophotometrically quantified by Nano Drop ND 2000 (Thermo Scientific, DE, USA). Then the DNA was used as the template for amplifying 16S rDNA genes. PCR mixtures(50 μl) were prepared in duplicate and each contained 2 μl of DNA template, 5 μl of 10 × PCR buffer(50 mm KCl, 10 mm Tris-HCl and 1.5 mm $MgCl_2$), 200 μm of dNTP, 0.2 μm of each primer and 2.5 U Taq polymerase(Promega, Madison, WI, USA). The PCR thermal regime consisted of an initial denaturation of 3 min at 94 °C, followed by 30 cycles of 30 s at 94 °C, 30 s at 60 °C, 1 min at 72 °C and a final cycle of 5 min at 72 °C. PCR products were pooled and purified with the Qiaquick gel purification kit according to the manufacturer's instructions(Qiagen, Hilden, North Rhine-Westphalia, Germany). DNA concentration and quality were determined with a NanoDrop 1000 spectrophotometer (Wilmington, DE, USA).

Data Preprocessing

Paired-end reads was assigned to samples based on their unique barcode and truncated by cutting off the barcode and primer sequence. Paired-end reads were merged using FLASH (V1.2.7, http://ccb.jhu.edu/software/FLASH/)[25], a very fast and accurate analysis tool, which was designed to merge paired-end reads when at least some of the reads overlap the read generated from the opposite end of the same DNA fragment, and the splicing sequences were called raw tags. Quality filtering on the raw tags were performed under specific filtering conditions to obtain the high-quality clean tags[26] according to the QIIME(V1.7.0, http://qiime.org/index.html)[27] quality controlled process. The tags were compared with the reference database(Gold database, http://drive5.com/uchime/uchime_download.html)using UCHIME algorithm(UCHIME Algorithm, http://www.drive5.com/usearch/manual/uchime_algo.html) [28] to detect chimeric sequences, and then

the chimeric sequences were removed[29]. Then the Effective Tags finally obtained.

OTU Clustering and Taxonomy Assignment

Sequences analysis were performed by Uparse software (Uparse v7.0.1001, http://drive5.com/uparse/)[30]. Sequences with s were performed by Uparse software (Uparse v7.0.1001, http://drive5.com/uparse/) high-quality clean tagsrlap the read genFor each representative sequence, the Green Gene Database (http://greengenes.lbl.gov/cgi-bin/nph-index.cgi) [31] was used based on RDP.classifier(Version 2.2, http://sourceforge.net/projects/rdp-classifier/)algorithmto annotate taxonomic information.

Microbial Diversity and Statistical Analysis

OTUs abundance information were normalized using a standard of sequence number corresponding to the sample with the least sequences. Subsequent analysis of alpha diversity and beta diversity were all performed basing on this output normalized data. Alpha diversity is applied in analyzing complexity of species diversity for a sample through 6 indices, including Observed-species, Chao1, Shannon, PD-whole tree rarefaction. The Chao1 estimator (http://www.mothur.org/wiki/Chao); Two indices were used to identify Community diversity: The Shannon index Shannon (http://www.mothur.org/wiki/Shannon); The Simpson index (http://www.mothur.org/wiki/Simpson); One indice to characterized Sequencing depth: The Good's coverage (http://www.mothur.org/wiki/Coverage).

Other statistical analyses

For statistical analysis, the environmental parameters were transformed to avoid skewed data distributions: ion concentrations were arcsine transformed, other chemical parameters were log10 transformed; and the pH, elevation, and latitude were not transformed. Significant marginal effects were analyzed after running a separate canonical correspondence analysis (CCA) on the OTU using square root transformation for each of the environmental factors (i.e., ion percentages, pH, elevation, total phosphorus(TP), total nitrogen (TN), and (TDS). We ran a CCA on the OTU and selected the three most important factors based on automatic forward selection(i.e., TDS, Cl, and Na) of the CANOCO program. The data set generated in this study has been deposited at GenBank's Short Read Archive(SRA) under Accession No. SRA178606.

RESULTS

Characteristics of the studied lakes

Gahai (TDS, 92-96 g/L), Xiaochaidan (TDS, 94-99 g/L) and Gasikule (TDS, 317-344 g/L) are typical hypersaline lakes situated in the Qaidam Basin of the Qinghai-Tibet Plateau, China. The salt lake Gasikule is located in the northwestern part of the Qaidam Basin at an elevation of approximately 2858 m above sea level, where less precipitation and high evaporation have resulted in the highest salinity (Figure 1). Gahai Lake is situated in the eastern part of the Qaidam Basin at an elevation of approximately 2853 m above sea level and is a hypersaline lake with abundant biological resources of Artemia (brine shrimp). The TDS of Gahai Lake has been stable in past years. Xiaochaidan Lake is a newly developed lake located in the northeastern part of the Qaidam Basin at an elevation of approximately 3170 m above sea level. Over the last century, the water area of Xiaochaidan Lake has increased two times. As a result, the TDS of Xiaochaidan Lake was 339.10 g/L in the 1970s, but was 94-99 g/L in 2013. Although these three lakes are all located in the Qaidam Basin of the Qinghai-Tibet Plateau, Xiaochaidan Lake and Gahai Lake are largely separated from Gasikule (408 and 603 km, respectively) (Figure 1). Five sample sites were investigated in the present study. TDS and certain physiochemical parameters varied widely along the spectrum from the Gahai sample site to the Gasikule sample site. Table S1 lists the sampling sites, the samples collected and the physicochemical parameters measured.

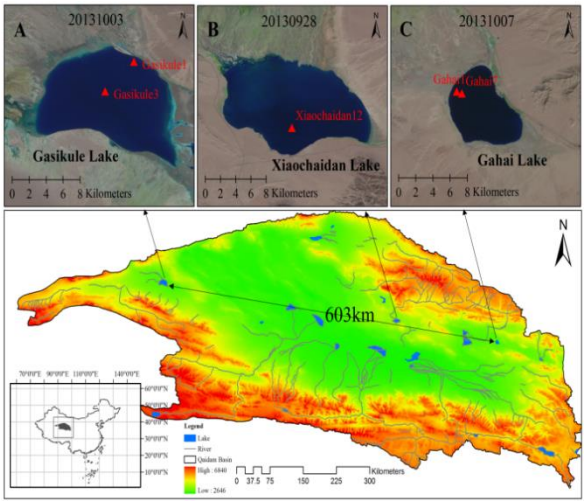

Fig. 1. Locations of three hypersaline lakes in Qaidam Basin. A, Gasikule Lake; B, Xiaochaidan Lake; C, Gahai Lake.

Composition and diversity of the PPNC

After quality filtering and preprocessing, 609,154 reads were obtained, with an average read length of 480 bp from the five sequenced samples examined in the present study; approximately forty-three percent, or a total of 260,201 reads, were assigned to prokaryotic pico- and nanoplankton assemblages. Sequencing yielded highly variable results among the samples, ranging from 85,336 to 153,144 total reads per sample site(Table 1). 97% sequence identity was considered a consensus threshold for reads belonging to the same OTU. The 16S rDNA gene sequences were distributed among nine high-rank taxonomic groups and matched 18 known prokaryote classes.

Table 1. Statistics of Taxonomic Composition in three hypersaline lakes of Qaidam Basin

Samples	Valid	TNR*	Phylum	Class	Order	Family	Genus	Other（ratio）
Gahai1	12827	648901	13	24	46	81	150	4.20%
Gahai7	15177	270250	12	25	49	90	152	55.50%
Xiaochadan12	15314	474351	11	21	49	96	154	15.50%
Gasikule1	85336	28873	11	21	38	62	103	0.70%
Gasikule3	90626	37826	13	23	41	66	114	5.00%

*TNR Total number of reads

Composition and diversity of the PPNC

After quality filtering and preprocessing, 609,154 reads were obtained, with an average read length of 480 bp from the five sequenced samples examined in the present study; approximately forty-three percent, or a total of 260,201 reads, were assigned to prokaryotic pico- and nanoplankton assemblages. Sequencing yielded highly variable results among the samples, ranging from 85,336 to 153,144 total reads per sample site(Table 1). 97% sequence identity was considered a consensus threshold for reads belonging to the same OTU. The 16S rDNA gene sequences were distributed among nine high-rank taxonomic groups and matched 18 known prokaryote classes.

The 16S rDNA gene sequences were distributed among eight high-rank taxonomic groups in hypersaline Gahai Lake. *Proteobacteria* was the most common phylum in Gahai Lake, accounting for 33.60% of the total number of sequences (Figure 2A). *Bacteroidetes* was the second most prevalent phylum, accounting for 24.40% of the total number of sequences in Gahai Lake. *Cyanobacteria* was the third most prevalent phylum, accounting for 22.20% of the total number of sequences in Gahai Lake. Unexpectedly, *Cyanobacteria* accounted for only 0.05% and 6.10% of the total sequences in Gasikule Lake and Xiaochaidan Lake, respectively(Figure 2A). The 16S rDNA gene sequences were distributed among eight high-rank taxonomic groups in hypersaline Xiaochaidan Lake. Most sequences from the Xiaochaidan Lake sample were affiliated with *Proteobacteria*(56.40%)(Figure 2A). *Actinobacteria* was the next most prevalent phylum in the Xiaochaidan Lake samples, accounting for 20.20% of all sample sequences. The most abundant sequences in Gasikule Lake were primarily represented by *Euryarchaeota*(76.75%). The 16S rDNA gene sequences were distributed among nine high-rank taxonomic groups in hypersaline Gasikule Lake(Figure 2A). *Bacteroidetes* was the second most prevalent phylum, accounting for 11.60% of the total number of sequences in Lake Gasikule(Figure 2A). *Proteobacteria* was the third most prevalent phylum, accounting for 10.55% of the total number of sequences in Gasikule Lake. There were approximately 2.10%, 1.30%, and 0.10% sequences defined as unclassified prokaryotic pico- and nanoplankton clusters in Lakes Gahai, Xiaochaidan, and Gasikule, respectively.

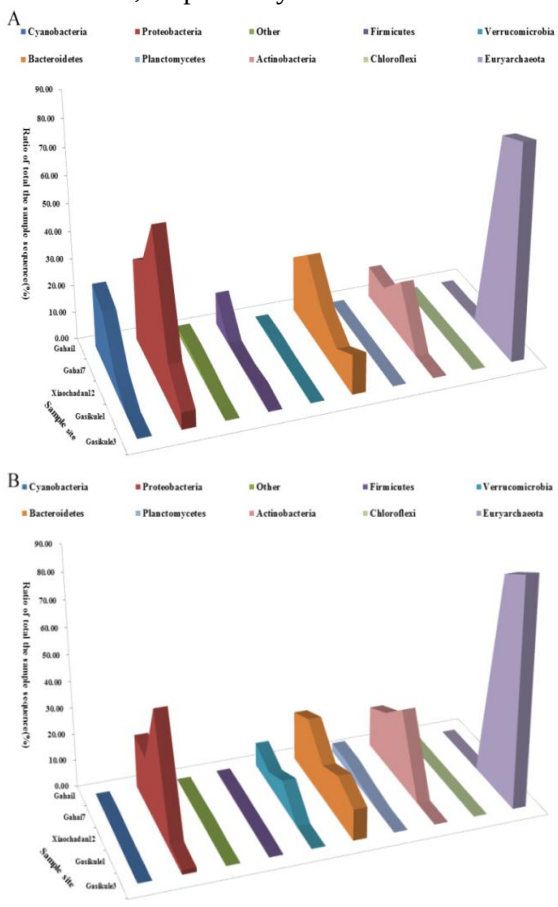

Fig. 2. The taxonomic composition of prokaryotic pico- and nanoplankton communities in three hypersaline lakes in Qaidam Basin. A, Prokaryotic pico- and nanoplankton community(a size range of 0.2-20 μm) of investigated sample site. B, prokaryotic plankton community (a size range >0.2 μm) of investigated sample site.

Compare the composition and diversity of the PPNC with WPPC

The diversity of the WPPC in the same sample site was significantly different from the diversity of the PPNC. As shown in Figure 2B, the most common group of WPPC was assigned as unclassified prokaryotes in Lake Gahai, accounting for 23.27% of the total number of sequences. *Bacteroidetes* was the second most prevalent Phylum, accounting for 20.67% of the total number of sequences in Gahai Lake. *Proteobacteria* was the third most prevalent phylum, accounting for 19.39% of the total number of sequences in Gahai Lake. *Euryarchaeota* was the most common phylum in Lake Gasikule, accounting for 82.85% of the total number of sequences(Figure 2B). *Bacteroidetes* was the second most prevalent phylum, accounting for 14.01% of the total number of sequences in Gasikule Lake. *Proteobacteria* was the third most prevalent phylum, accounting for 2.07% of the total number of sequences in Gasikule Lake.

Table 2. The most abundant taxa in in three hypersaline lakes of Qaidam Basin.

Sample site	Taxa	Ratio (%)
Gahai1	Bacteria;Cyanobacteria;Cyanobacteria;Chloroplast;Chlorophyta;Other	23.30
	Bacteria;Proteobacteria;Alphaproteobacteria;Rhodobacterales;Rhodobacteraceae;Other	18.50
	Bacteria;Firmicutes;Bacilli;Lactobacillales;Lactobacillaceae;Lactobacillus	11.80
Gahai7	Bacteria;Cyanobacteria;Cyanobacteria;Chloroplast;Chlorophyta;Other	18.50
	Bacteria;Proteobacteria;Gammaproteobacteria;Other;Other;Other	11.90
	Bacteria;Proteobacteria;Alphaproteobacteria;Rhodobacterales;Rhodobacteraceae;Other	11.30
*XCD12	Bacteria;Actinobacteria;Actinobacteria;Actinomycetales;Microbacteriaceae;Other	18.60
	Bacteria;Proteobacteria;Alphaproteobacteria;Rhodobacterales;Rhodobacteraceae;Other	16.50
	Bacteria;Proteobacteria;Betaproteobacteria;Burkholderiales;Alcaligenaceae;Other	13.10
**GSK1	Archaea;Euryarchaeota;Halobacteria;Halobacteriales;Halobacteriaceae;Haloquadratum	43.00
	Archaea;Euryarchaeota;Halobacteria;Halobacteriales;Halobacteriaceae;Other	31.00
	Bacteria;Bacteroidetes;Sphingobacteria;Sphingobacteriales;Rhodothermaceae;Salinibacter	7.60
GSK3	Archaea;Euryarchaeota;Halobacteria;Halobacteriales;Halobacteriaceae;Haloquadratum	47.90
	Archaea;Euryarchaeota;Halobacteria;Halobacteriales;Halobacteriaceae;Other	29.90
	Bacteria;Bacteroidetes;Sphingobacteria;Sphingobacteriales;Rhodothermaceae;Salinibacter	13.40

*XCD, Xiaochadan12 Lake; **GSK, Gasikule Lake

Most abundant taxa of PPNC in different sampling sites

Table 2 shows the habitat distribution of abundant taxa at or below the phylum rank in different sampling sites, based on amplicon sequencing. Unclassified prokaryotic taxa(belonging to the class Cyanobacteria) exhibited the highest proportion in the Gahai1 Lake sample point, accounting for 18.60% of the total number of sequences(Table 2). The most common taxa in the Xiaochaidan12 sampling point belong to the class *Actinobacteria*, accounting for 39.79% of the total number of sequences. These taxa have not been well studied and cannot be clearly classified into any genus. At the Gasikule1 and Gasikule3 sampling sites, the prokaryote communities primarily consisted of microorganisms of the genus *Haloquadratum*. The next most abundant sequences were affiliated with unclassified species, accounting for 31% and 29.90% of the total number of sequences, respectively. Unclassified pico- and nanoplankton accounted for significant components of the investigated samples.

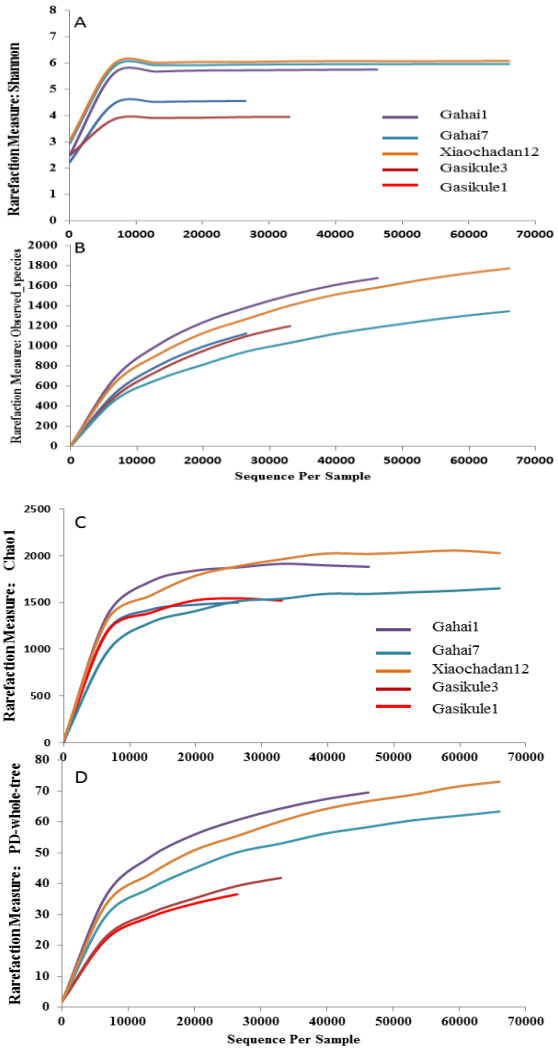

Fig. 3. Estimating species richness of three hypersaline lakes in Qaidam Basin. A, Shannon rarefaction curves; B Observed, species rarefaction curves; C, chao1 rarefaction curves; D PD-whole tree rarefaction curves.

Alpha diversity of the PPNC

Abundance OTU tables were used to calculate the Shannon's diversity index and Simpson's evenness. The Shannon rarefaction analysis(Figure 3) revealed exhaustive sequencing, even for the smallest dataset. As shown in Figure 3, the estimated number of observed species varied between 1124.80 and 1347.40. The highest diversity index was observed in Gahai7(Observed Species = 1347.40), and the lowest diversity index was observed in Lake Gasikule1(Observed Species = 1124.80). The Chao1 estimator was calculated to predict the total number of OTUs(richness) in the water samples from the studied lakes at a 97% similarity cutoff. The taxonomic richness levels estimated for the five samples were 1882.108, 1651.577, 2028.465, 1500.107, and 1520.658 for Lakes Gahai1, Gahai7, Xiaochaidan12, Gasikule1 and Gasikule3, respectively(Figure 3). Considering the mean estimated richness in DNA datasets, the Xiaochaidan 12 sample site was the most diverse lake, closely followed by the Gahai1 sample site.

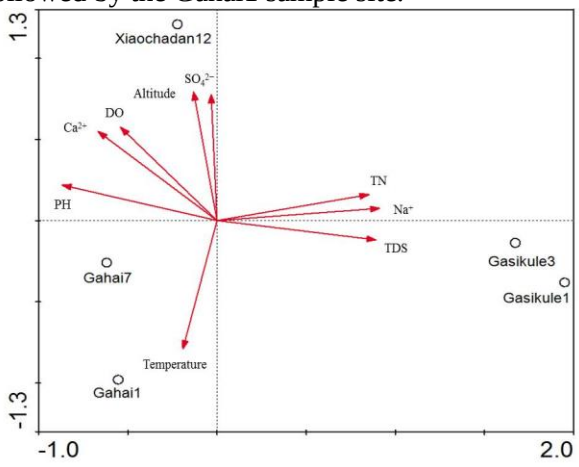

Fig. 4. CCA biplots based on OTU and geographical or selected chemical parameters. DO, dissolved oxygen, TN total nitrogen; TDS total dissolved solid.

Influence of physicochemical and chemical parameters on the diversity of PPNC

To analyze the influence of the prokaryotic pico- and nanoplankton community structure and other measured physicochemical parameters, a distinct CCA was generated from pooled habitat datasets. Three parameters showed significant correlations with this CCA: Altitude, temperature and TDS. These parameters were also significantly correlated with the CCA generated from unmerged datasets, and the fitted vectors are shown in Figure 4. The CCA analysis of chemical variables yielded three clusters (Figure 4), primarily separated according to elevation, temperature and TDS. The first cluster contained Gahai Lake sample points(Gahai1, Gahai7); the second contained large Xiaochaidan Lake sample points(Xiaochaidan 12); and the third cluster contained Gasikule Lake sample points(Gasikule1 and Gasikule3). Elevation affected the PPNC structure in Xiaochaidan Lake, whereas temperature was a major factor in Gahai Lake. calcium(Ca^{2+}) and sulfate(SO_4^{2-}) were also the factors that influence of the PPNC structure in Xiaochaidan Lake. TDS was the most important factor influencing the distribution of PPNC assemblages in Gasikule Lake (Figure 4).

DISCUSSION

The development of pyrosequencing as a technique for the deep sequencing of microbial communities has contributed a tremendous amount of new information to the current understanding of the diversity of these systems[32]. Extreme environments contain less diverse communities[33]. However, all hypersaline lakes examined in the present study harbored remarkably diverse microbial communities considering the high salinity of these water bodies(Figure 2). Microorganisms that inhabit these lakes are potentially valuable "gene reservoirs" for future biotechnological applications, particularly those involving saline conditions (e.g., microbial treatment of saline or high-salt waste water). Investigations of the community taxon diversity(microdiversity) of the bacterioplankton at higher salinities are necessary for understanding the patterns of the global distribution of microbial diversity.

Picocyanobacteria were prevalent in Gahai Lake

PPNC in the hypersaline lakes were dominated with bacteria, except for Gasikule lake, which was dominated with archaea at a proportion of 74.9-78.6%(Figure 2A). The competitive advantage of archaea over bacteria in extreme environments is associated with the degree of energetic[34] or salinity stress[35-36] experienced by microbes in lakes and the physiological adaptations of the two groups for dealing with such stresses.

Notably, the most prevalent class in the investigated sample site of Gahai Lake was picocyanobacteria, while only a few organisms were detected in Gasikule Lake. Although there were no obvious discrepancies in the TDS between Gahai Lake and Xiaochaidan Lake, the ratios of picocyanobacteria in PPNC showed obvious differences(22.20% and 6.10%, respectively). Cyanobacteria of picoplanktonic cell size (0.2 to 2.0 μm) are globally important primary producers in freshwater, brackish, and marine ecosystems[14]. Picocyanobacteria contribute as much as 70% of the

total primary production in lakes, particularly in oligotrophic high mountain lakes[37-38]. Furthermore, Picocyanobacteria incorporate dissolved organic matter (DOM) into the food web[39]. Because of the small size of picocyanobacteria, these organisms comprise the main food source of nanoplanktonic protozoans: Ciliata, Flagellata and larger zooplankton[40]. The differences among the compositions of the PPNC in the investigated lakes revealed that there were fewer food sources in Gasikule Lake and Xiaochaidan Lake. Moreover, as the primary producer in the ecosystem, picocyanobacteria play important roles in the material cycle and energy flow in hypersaline lakes. These results contributed to the explanation of why Artemia prevail in Gahai Lake but not in other hypersaline lakes situated in the Qaidam Basin of the Qinghai-Tibet Plateau.

Influences of TDS and Altitude on PPNC

Salinity might be the strongest stress factor limiting microbial diversity[17,41]. Our result confirmed that TDS is an environmental factor that strongly influences the taxonomic composition of prokaryotic picoplankton assemblages in inland waters[18]. Recent reports have suggested that salinity and oxygen are important factors that shape the microbial composition in aquatic habitats[18,42-4346-48]. As we know, the mechanisms controlling primary production might involve many factors, such as nutrient limitations, toxicity, or trophic interactions[44]. In the present study, CCA analysis revealed that temperature was significantly correlated with the distribution of prokaryotic picoplankton and assemblages in Gahai Lake, whereas altitude was a major factor influencing the taxonomic composition of prokaryotic pico- and nanoplankton assemblages in Xiaochaidan Lake (Figure 4). When compared with TDS, Gasikule Lake had the the highest levels. As expected, the prokaryotic pico- and nanoplankton assemblages in Gasikule Lake were significantly different than the two lakes. The results of the present study revealed that TDS defines distinct prokaryotic pico- and nanoplankton assemblages among lakes, whereas other factors affect the distribution of prokaryotic pico- and nanoplankton assemblages within one lake.

The obvious divergence between the structures of the PPNC and WPPC

The sequence length obtained through 454 pyrosequencing, originally of 100 bp, now exceeds 400 bp (Titanium chemistry, 454 Life Sciences, Basel, Switzerland), enabling a more precise taxonomic classification of the reads[45]. The most common group of prokaryotic communities in Lake Gahai was assigned as unclassified prokaryotes, accounting for 23.27% of the total number of sequences[23]. However, the results of our study suggest that Proteobacteria was the most common phylum in Lake Gahai, accounting for 33.60% of the total number of sequences. As shown in Figure 2B, Proteobacteria was the third most prevalent phylum (20.67% of the total number of sequences) among the WPPC in Gahai Lake. Unexpectedly, our results revealed that Picocyanobacteria was the third most prevalent phylum of PPNC in Gahai Lake (22.20%). These results confirm that caution must be taken before making conclusions regarding the geographic restriction of novel clades, reflecting the biases of the methods used, i.e., data search or taxon under-sampling. One should consider that diversity analyses based on single genes require a more careful interpretation of the results (e.g., phylogenetic resolution of the region analyzed or the possibility of horizontal gene transfer). These studies also reveal important gaps in the current knowledge concerning planktonic microbial inhabitants in hypersaline water bodies.

Most saline lakes are defined according to endorheic drainage basins in dry areas world wide. Considering the scarcity of water in arid lands, temporary water has greater ecological significance in arid regions compared with in wet regions. The results of the present study showed that hypersaline lakes in the Qaidam Basin most likely contain a significant number of novel species, which must be cultured for detailed ecophysiological studies. The present study is the first attempt to characterize picoprokaryotic diversity in the hypersaline lakes of the Qaidam Basin, setting the basis for future studies describing new bacterial species or isolates with biotechnological applications and stressing the need to preserve extreme ecosystems with undescribed diversity.

CONCLUSION

Proteobacteria were the common phylum of PPNC in Lake Gahai. The most abundant sequences, however, primarily represented Euryarchaeota (78.60%) of PPNC in Lake Gasikule. An obvious divergence between the structure of the PPNC and the WPPC was observed. The most common group of WPPC was assigned as unclassified prokaryotes in Lake Gahai. *Bacteroidetes* was the second most prevalent Phylum, accounting for 20.67% of the total number of sequences in Gahai Lake. Altitude, temperature and TDS(total dissolved solids) were

significantly correlated with the distribution of the PPNC.

Acknowledgements: *The authors would like to thank Lanping Zhou, Qiang Zhang and Peng Zhang for their assistance with the collection of samples from the waters of saline lakes. This work was financed through a grant from the Independent Study Project of the State Key Laboratory of the Simulation and Regulation of the Water Cycle in a River Basin (NO.2013ZY06), a Pre-project of The Qinghai Department of water resources, The Self-Regulation Project of State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research, Grant(NO. 2016TS01), the Open Research Fund of State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research, Grant(NO. IWHR-SKL-201620), The National Natural Science Funds Fund(NO.51309247).*

Supporting Information: *One files containing The geographical, physical, and chemical characteristics of three hypersaline lakes in Qaidam Basin are available as supplemental materials at Online. Table S1 The geographical, physical, and chemical characteristics of three hypersaline lakes in Qaidam Basin.*

REFERENCES

1. R.G. Bryant, *International J of Remote Sensing*., **17**, 315 (1996).
2. W.D. Williams, *Verh Internat Verein Limnol*., 26, 61 (1996).
3. S. Benlloch, A. López-López, E.O. Casamayor, L. Øvreås, V. Goddard, F.L. Daae, G. Smerdon, R. Massana, I. Joint, F. Thingstad, C. Pedrós-Alió, F. Rodríguez-Valera, *Environ. Microbiol*., **4**(6), 349 (2002).
4. H. Jiang, H. Dong, B. Yu, G. Lv, S. Deng, Y. Wu, M. Dai, N. Jiao, *FEMS Microbiol Ecol*., **67**, 268 (2009).
5. Q.L. Wu, G. Zwart, M. Schauer, Kamst-van, M.P. Agterveld, M.W. Hahn, *Applied & Environm. Microbiol*., **72**(8), 5478 (2006).
6. J.F. Aguirre-Garrido, H.C. Ramírez-Saad, N.Toro, F. Martínez-Abarca, *Microb. Ecol*., **71**(1), 68 (2016).
7. C.D. Litchfield, *Meteorit Planet Science*., **33**, 813 (1998).
8. E.B. Hollister, A.S. Engledow, A.J. Hammett, T.L. Provin, H.H. Wilkinson, T.J. Gentry, *Isme Journal*., **4** (6), 829 (2010).
9. C.M. De, M. Staiano, M. Rossi, S. D'Auria, *Journal of the Royal Society Interface*., **4**(13), 183 (2007).
10. J.M. Sieburth, V. Smatacek, J. Lenz, *Limnol and Oceanography*., **23**,1256 (1978).
11. M.1. Viprey, L. Guillou, M. Ferréol, D. Vaulot, *Environm Microbiol*., **10**(7), 1804 (2008).
12. N. Das, A. Pandey, *International Letters of Natural Sciences*., **43**, 38 (2015).
13. O. Zhaxybayeva, W.F. Doolittle, R.T. Papke, J. Gogarten, *Genome Biol and Evolu*., **1**, 325 (2009).
14. D.J. Scanlan, M. Ostrowski, S. Mazard, A. Dufresne, L. Garczarek, W.R. Hess, A.F. Post, M. Hagemann, I. Paulsen, F. Partensky, *Microbiol & Molecular Biol R*., **73**(2), 249 (2009).
15. Z.G. Keresztes, T. Felföldi, B. Somogyi, G. Székely, N. Dragoş, K. Márialigeti, C. Bartha, L. Vörös, *Extremophiles*., **16**(5), 759 (2012).
16. J. Handelsman, *Microb and molecular biol reviews*., 68, 669 (2004).
17. A. Oren, B.K. Baxter, B.C. Weimer, Natural Resources & Environmental Issues, Iss 1 (2009).
18. Q.L. Wu, A. Chatzinotas, J. Wang, J. Boenigk, *Micro Ecol*., **58**, 569 (2009).
19. X.Triadó-Margarit, E.O. Casamayor, Fems *Microb ecol*., **85**, Early View. doi: 10.1111/1574-6941.12095 (2013).
20. E.O. Casamayor, *FEMS microbiol ecol*., **85**, 503 (2013).
21. M.P. Zheng, X.F. Liu, *Aquat Geochem*., **15**, 293 (2009).
22. J.L. Wang, F. Wang, L.M. Chu, H. Wang, Z. Zhong, Z. Liu, J. Gao, H. Duan, *Plos One*., **9**(11), e112812 (2014).
23. Z.P. Zhong, Y. Liu, L.L. Miao, F.Wang, L.M. Chu, J.L. Wang, Z.P. Liu, *Applied & Environ Microbiol*., **82**(6), 1846 (2016).
24. A.E. Greenberg, L.S. Clesceri, A. DEaton, American Public Health Association., Washington (1992).
25. T. Magoč, S.L. Salzberg, *Bioinformatics*., **27**(21), 2957 (2011).
26. N.A. Bokulich, S. Subramanian, J.J. Faith, D. Gevers, J.I. Gordon, R. Knight, D.A. Mills, J.G. Caporaso, *Nature Methods*., **10**(1), 57 (2013).
27. J.G. Caporaso, J. Kuczynski, K. Bittinger, F.D. Bushman, E.K. Costello, N. Fierer, A.G. Peña, J.K. Goodrich, J.I. Gordon, G.A. Huttley, S.T. Kelley, D. Knights, J.E. Koenig, R.E. Ley, C.A. Lozupone, D. McDonald, B.D. Muegge, M. Pirrung, J. Reeder, J.R. Sevinsky, P.J. Turnbaugh, W.A. Walters, J. Widmann, T. Yatsunenko, J. Zaneveld, R. Knight, *Nature Methods*., **7**(5), 335 (2010).
28. R.C. Edgar, B.J. Haas, J.C. Clemente, C. Quince, R. Knight, Bioinformatics, **27**(16), 2194 (2011).
29. B.J. Haas, D. Gevers, A.M. Earl, D.V. Ward, G. Giannoukos, D. Ciulla, D. Tabbaa, S K. Highlander, E. Sodergren, B. Methé, T.Z. DeSantis, J.F. Petrosino, R. Knight, B.W. Birren, *Genome Research*., **21**(3), 494 (2011).
30. R.C. Edgar, *Nature Methods*., **10**(10), 996 (2013).
31. T.Z. Desantis, P. Hugenholtz, N. Larsen, M. Rojas, E.L. Brodie, K. Keller, T. Huber, D. Dalevi, P. Hu, G.L. *Andersen, Greengenes, Applied & Environm Microbiol.,* **72**(7), 5069 (2006).
32. L.J. Yan, M.P. Zheng, Z.J. Yuan, *Mineral Deposits*., **33** (5), 921 (2014).
33. A. Sengupta, W.A. Dick, Microb. Ecol., **70**(3), 1 (2015).

34. M.J. Fawley, K.P. Fawley, M.A. Buchheim, *Microb Ecol*., 48,489 (2004).
35. A. Oren, *Microbiol and molecular boil reviews*., **63**, 334 (1999).
36. D.L. Valentine, *Nature Reviews Microbio*., **5**, 316 (2007).
37. Q.L. Wu, P. Xing, W.T. Liu, *Micro Ecol*., **59**, 614 (2010).
38. T. Weisse, and A. Schweizer, *Verh Internat Verein Limnol*., 24, 821 (1991).
39. C. Ruiz-González, R. Simó, R. Sommaruga, J.M. Gasol, *Frontiers in Microbiol*., **4**(4), 131 (2013).
40. N. Jakubowska, E. Szeląg-Wasielewska, *Marine Drugs*., **13**(3), 1497 (2015).
41. T.N. Zhilina, D.G. Zavarzina, A.N. Panteleeva, G.A. Osipov, N.A. Kostrikina, T.P. Tourova, G.A. Zavarzin, *International Journal of Systematic and Evolutionary Microbiol*., **62**, 1666 (2012).
42. D.P. Herlemann, M. Labrenz, K. Jürgens, S. Bertilsson, J.J. Waniek, A.F Andersson, *Isme Journal*., **5**, 1571 (2011).
43. A. Barberán, and E.O. Casamayor, *Microb Ecol*., **61**, 465 (2011).
44. G.A. Zavarzin, T.N. Zhilina, Kluwer Academic Publishers., 2000, p.191.
45. C. Bachy, J.R. Dolan, P. López-García, P. Deschamps, D. Moreira, *Isme J*., **7**(2), 244 (2012).