Microbial Biogeography of Six Salt lakes in Inner Mongolia China and a Salt Lake in Argentina

Running Title: Microbial Biogeography of Salt Lakes

EULYN PAGALING,[1] MADELEINE VENABLES,[1] ANDREW WALLACE,[1] WILLIAM D. GRANT,[1] DON A. COWAN,[2] BRIAN E. JONES,[3] YANHE MA,[4] ANTONIO VENTOSA,[5] SHAUN HEAPHY[1]*

1. Department of Infection Immunity and Inflammation, University of Leicester, University Road, Leicester, LE1 9HN, UK.

2. Department of Biotechnology, University of the Western Cape, Bellville 7535, Cape Town, South Africa.

3. Genencor International B, V., Archimedesweg 30, 2333 CN Leiden, The Netherlands.

4. State Key Laboratory of Microbial Resource, Institute of Microbiology, Chinese Academy of Sciences, 100080, Beijing China.

5. Department of Microbiology and Parasitology, Faculty of Pharmacy, University of Sevilla, 41012, Sevilla, Spain.

Email addresses: Eulyn.Pagaling@ed.ac.uk, MMV543@bham.ac.uk, amw@le.ac.uk, wdg1@le.ac.uk, dcowan@uwc.ac.za, brian.e.jones@danisco.com, mayanhe@im.ac.cn, ventosa@us.es, sh1@le.ac.uk.

Shaun Heaphy* Corresponding author. Department of Infection Immunity and Inflammation, University of Leicester, University Road, Leicester, LE1 9HN, UK.

Tel + (0)44 116 252 2973; Fax 44 116 252 5030; Email sh1@le.ac.uk.

Journal section: Microbial Ecology

Key words: Archaea, bacteria, biogeography,halophilic, alkaliphilic

Abstract

We used cultivation-independent methods to investigate the prokaryotic biogeography of the water column in six salt lakes in Inner Mongolia China, and a salt lake in Argentina. These lakes had different salt compositions and pH values, and were at variable geographic distances, both at local and inter-continental scales, which allowed us to explore the microbial community composition within the context of both contemporary environmental conditions and geographic distance. Fourteen 16S rRNA gene clone libraries were constructed, and over 200 16S rRNA gene sequences were obtained. These sequences were used to construct biotic similarity matrices, which were used in combination with environmental similarity matrices and a distance matrix in the Mantel test to discover which factors significantly influenced biotic similarity. We showed that archaeal biogeography was influenced by contemporary environmental factors alone (Na^+, CO_3^{2-}, and HCO_3^- ion concentrations, pH and temperature). Bacterial biogeography was influenced both by contemporary environmental factors (Na^+, Mg^{2+} and HCO_3^- ion concentrations and pH) and by geographic distance.

Introduction

Biogeography aims to explain spatial patterns of diversity in the context of evolutionary events such as speciation, dispersal, extinction and species interactions [42]. Macro-ecologists have long studied the biogeography of higher plants and animals in various habitats [9, 13]. In contrast, there is very little information available on the biogeography of prokaryotes. This stemmed from the difficulty of assessing microbial communities by cultivation methods, which only sampled 0.1 – 10% of the microbial community [30]. However, with the advent of cultivation independent sequencing techniques, microbial communities of many environments have been characterised, including soil [43], the Arctic and Antarctic Oceans [5] and the Sargasso sea [61]. This in turn facilitated prokaryotic biogeography studies in a number of environments on scales ranging from 20 000 km to 0.002 km [42].

A study on the biogeography of soil bacteria across the Americas showed that differences were largely attributed to soil pH, with higher diversity observed in neutral soils [20]. Bacterial communities in an estuary in Massachusetts were found to vary with the salinity gradient [14]. Such studies demonstrated that environmental parameters influenced biogeographical patterns in microbial diversity. Further studies demonstrated that biogeography of hot spring cyanobacteria, hyperthermophilic archaea and *Pseudomonas* stains was influenced by geographic distance, which led to isolation of disparate populations and subsequent genetic divergence [12, 51, 63]. The apparent allopatric speciation demonstrated in these studies therefore contested the idea that prokaryotes were not affected by limits to dispersal due to their small size, abundance and metabolic plasticity (‘everything is everywhere’, see below) [21].

A simple framework was suggested to distinguish between the effects of evolutionary events and contemporary environmental conditions on the spatial variation of microbial diversity [42]. At the centre of this framework were four hypotheses. The null hypothesis stated that microorganisms were distributed randomly over space. Upon rejection of the null hypothesis, the second hypothesis stated that spatial variation reflected the influence of contemporary environmental variation. It assumed that geographic distance did not affect diversity due to the wide dispersal of microorganisms. This hypothesis is the famously quoted 'everything is everywhere; the milieu selects' by Baas-Becking [4, 6]. The third hypothesis stated that variation was shaped by evolutionary events [geographic distance] that limited dispersal, and past environmental conditions led to genetic divergence between different microbial assemblages. The fourth hypothesis stated that the biogeography of microorganisms was determined by both contemporary environmental conditions and past evolutionary events [geographic distance]. It is important to note here the possibility that evolutionary events can be represented by geographic distances. (For more details on this framework, see [42]).

Many studies have been carried out on salt lakes and salterns around the world [28], but few have tried to explain variations in microbial community composition. Those that did identified salinity, altitude, redox and ionic concentration, pO2 and seasonal events as relevant factors [7, 11, 16, 17, 34, 35, 38, 65]. To our knowledge, only two studies have looked at the effect of inter-continental geographic distances on microbial community composition in salt lakes. Foti and colleagues looked specifically at the biogeography of *Thioalkalivibrio* in soda lakes across Mongolia, Kenya, California, Egypt and Siberia, and found that these bacteria showed a

tendency for endemism, hence geographic distance was a significant factor in influencing community composition [22]. A further study looked at the biogeography of *Salinibacter ruber* strains from salterns in the Mediterranean, Atlantic and Peruvian regions using a metabolomic approach. Geographically distinct strains were distinguished by characteristic metabolites [58].

We examined the prokaryotic community composition in several salt lakes using ribosomal DNA methods. Six of the salt lakes in this study were situated on the Inner Mongolian steppe, north-west of Beijing, which had an average elevation of 1000-2000 m above sea level. The lakes were mostly several hundred km apart (0.147-395.2) and are in different climate and vegetation zones – typical grassland steppe in the north and east, to desert steppe bordering the Gobi desert in the south and west [70]. The lakes were Bagaejinnor, Chagannor, Ejinnor, Erliannor, Shangmatala and an unnamed lake near Xilinhot. Lakes Ejinnor and Erliannor were extensively developed into salterns. The salar Guayatayoc Lake was situated in the same basin as the Salinas Grandes in the Argentine Altiplano at an elevation of 3432 m, north-west of the city Salta, approximately 18000 km from the other lakes. All salt lakes were athalassohaline, located in arid climates, and subjected to high solar radiation and wide ranges of temperature. The lakes had different salt compositions and allowed us to explore the microbial community composition within the context of both contemporary environmental conditions and geographic distance.

Here we describe the microbial diversity of six salt lakes in Inner Mongolia and one salt lake in Argentina. Using the framework previously described, we present evidence that biogeography of *Archaea* in these salt lakes was significantly influenced

(P <0.05) by environmental factors (Na^{+}, CO_3^{2-}, and HCO_3^{-} ion concentrations, pH and temperature), but not geographic distance, consistent with the previously stated hypothesis 2. We also show that the biogeography of *Bacteria* was significantly influenced (P <0.05) by both environmental factors (Na^{+}, Mg^{2+} and HCO_3- ion concentrations and pH) and geographic distance, consistent with the previously stated hypothesis 4.

Materials and Methods

Descriptions of Sampling Sites

All sites were remote from centres of population and usually involved long drives over unmade roads. Our ability to transport equipment was limited. Lakes themselves were often fringed by deep mud making sampling hazardous. The first four lakes were in areas of grassland steppe, the last two lakes in desert steppe and the Argentine lake in an arid high altitude plateau. Temperature, pH and chemical analyses of the brines are shown in Table 1.

Lake Bagaejinnor (BJ) was a hypersaline lake, whose coordinates were N45^{o} 08.527' E116^{o} 36.167', north of the town Qog Ul. It has a surface area of 5 km^{2} during the wet season [66]. It had evaporated over the summer, exposing salt encrusted mud flats and had been reduced to a number of small pools and lagoons. The brine was colourless, but the salt crystals had a pink colouration, indicating the presence of haloarchaea.

Lake Ejinnor (EJ) was a hypersaline lake, with coordinates of N45 ^{o} 14.452' E116 ^{o} 32. 477' north of the town of Qog Ul, 40 km from lake BJ. It is a large shallow lake, 0.05-0.3M deep, with evaporating lagoons on the eastern side of the main body of water. The lake water sample was taken from a large saltern containing red brine and orange-pigmented salt crystals about 0.3 m deep.

An unnamed lake, located north-west of Xilin Hot with coordinates of N47 ^{o} 55.355' E115^{o} 36.757' was also sampled. It was a hypersaline lake situated near an abandoned soda works. The lake was divided by several causeways. A shallow lagoon was found

cut off from the rest of the lake, which was where one of the sampling sites was located (XH). The lake had a thick white salt crust, while the brine was clear and colourless and contained brine shrimps (*Artemia* sp.). Leading from the lake was a drainage channel that connected to a 15 cm deep pool of green brine (147 m from the lake), where a second sample was taken (X).

Lake Shangmatala a hypersaline lake was located in a shallow basin surrounded by hills at an elevation of 987 m with coordinates N43° 22.751' E114° 01. 361'. The lake had a surface area of 2.5 km^2 and a depth of 0.1-0.15 m. The lake was surrounded by lush grassland and vegetation, which grew almost up to the water's edge. The soil nearest the lake appeared to be soda soil, which had a layer of lichen growing on the surface. It was noted that an unpleasant smelling gas was emitted. A causeway led directly into the lake.

Lake Chagannor (CG) was a large hypersaline soda lake, situated near a soda works, 120 km south of Mandulatu, with coordinates of N43^o 16.131' E112^o 55.636'. Sampling took place on the south side of the lake. The brine appeared green and the mud was grey and viscous, with a layer of fine salt.

Lake Erliannor (EN) was a hypersaline lake, located north of Erenhot on the Mongolian border and the trans-Siberian railway, with coordinates of N43^o 44.426' E112 ^{o} 02. 081'. It is reported to have a surface area of 8.75 km^2 and a depth of 0.1-0.3 m. The natural lake was unrecognisable due to extensive development of salterns.

The lake water sample was taken from a saltern (0.1 m depth) that contained colourless brine and a white salt crust.

Salar Guayatayoc Lake (AG) was a hypersaline lake on the north edge of the Salinas Grandes, Argentina. Its coordinates were S23 ^{o} 36.604' W ^{o} 65 51.998'. It was locally reported to have a depth of 30 m and was covered by a ~ 1 m thick salt crust. Samples were taken through a hole in this crust.

Sample Collection

Biomass from the water column from the Argentinean salt lake was sampled in July 2003 and from the Inner Mongolian salt lakes in September 2003. In Inner Mongolia, brine was sampled at a distance in 250 ml stainless steel beakers suspended on the end of a 1 m pole. In Argentina, brine was collected through a hole in the ~ 1 m thick salt crust. Water was filtered through sterile 0.45 µm membrane filters (Millipore) in a 250 ml capacity polycarbonate filter unit (Sartorius) using a Nalgene hand pump, which produced a vacuum of 40-50 cm Hg under field conditions. Water was processed in this way until flow stopped, which suggested that sufficient biomass was captured on the filter. Membrane filters were removed from the apparatus using sterile tweezers and placed immediately in cold sterile stabilisation buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, 2 M NaCl) and agitated to resuspend the cells. This was immediately placed on ice until further processing.

Measuring Geographic Distances

GPS coordinates recorded at each sampling point were imported into MapSource™ according to the manufacturer's instructions to measure the geographic distances between the sites.

Measuring pH, Temperature and Salinity of the Salt Lakes

The pH of the salt lake water was measured using pH strips (Merck).

The temperature was measured using a Solomat 520C temperature monitor or a Hanna KType thermocouple with SP weighted tanker probe (Jencons, Leighton Buzzard, UK) according to manufacturer's instructions. The temperature was measured at a distance by attaching the probe to the end of a 1 m pole. Other physical and chemical analysis were performed on samples filtered through a 0.22 micron membrane, stored in sterile screw capped vials.

The salinity was measured using a Hanna HI 8633 or HI9033 multirange conductivity meter (Jencons, Leighton Buzzard, UK), which was calibrated to 20°C with a temperature coefficient of 2% according to the manufacturer's instructions. All salt lake water readings were off the scale, hence they were serially diluted with distilled water and readings made at the 199.9 mS/cm range. Water conductivity gave an indication of total salt concentration in g/L [64].

Determining Chemical Composition of Salt Lake Water

Chemical analysis of the salt lake water was carried out by Inductively Coupled Plasma Optical Emission Spectrometry (ICP-ES). Samples were sent for analysis to the Geology Department at the University of Leicester.

222

223 **Titrations of Carbonate and Bicarbonate**

224 Concentrations of carbonate ions (CO_3^{2-}) and bicarbonate ions (HCO_3-) were found by
225 titration of lake water with H_2SO_4 using a Digital Titrator Model 16900 (Hach
226 Systems for Analysis) according to manufacturer's instructions.

227

228 **Community DNA extraction and PCR Amplification of 16S rRNA genes**

229 Community DNA from the Inner Mongolian samples was extracted from
230 environmental samples using the GenomicPrep Cells and Tissue DNA Isolation Kit
231 (Amersham Biosciences). The initial stages of protein precipitation were carried out
232 on site and the sample stored at -20°C until DNA purification could be carried out in
233 the laboratory in Leicester. Community DNA from the Argentinean sample was
234 extracted by freezing the filter in a small amount of liquid nitrogen, which was then
235 homogenised in a pre-cooled pestle and mortar (at -80°C). This material was
236 transferred to a clean tube, and 960 µl of NET buffer (150 mM NaCl, 100 mM EDTA
237 pH 8.0, 50 mM Tris-HCl pH 8.0) containing 15 mg/ml of lysozyme was added, which
238 was incubated at 37°C for 10 min. 192 µl of 11mg/ml proteinase K and 128 µl of 10%
239 (w/v) SDS were added and incubated for a further 30 min at 65°C. DNA was
240 extracted by phenol chloroform and ethanol precipitation. The DNA pellet was
241 resuspended in 150 µl of Tris-EDTA. Archaeal and bacterial 16S rRNA genes were
242 PCR amplified as previously described [25].

243

244 **Construction of 16S rRNA Gene Libraries and Screening Inserts**

245 PCR products were ligated into pGEM T-easy cloning vector and transfected into
246 JM109 *E. coli* cells according to manufacturer's instructions to make 16S rRNA gene

libraries. Forty-eight white colonies containing recombinant plasmids were picked from each library and grown in Luria Bertani broth (1% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl and adjusted to pH 7.0 with 1 M NaOH) containing 100 µg/ml of ampicillin at 37°C overnight. Inserts were amplified by colony PCR by the following method: 2-3 µl cell culture was mixed with 20 µl 0.25% (v/v) Tween 20 and boiled for 20 min to denature cellular proteins. Cell debris was pelleted by centrifugation at 10000 x g for 10 min. 15 µl of this cell lysate was used in a second round of PCR using flanking vector primers M13F 5'-GTT TTC CCA GTC ACG AC-3' and M13R 5'-CAG GAA ACA GCT ATG AC-3' in a reaction previously described (see above), to a final volume of 25 µl. This gave an expected amplicon of 1.7 kb.

To look for the restriction fragment length polymorphism (RFLP) pattern of each insert, the 1.7 kb PCR products were digested with *Hae*III (NEB) in the manufacturer's reaction buffer at 37°C for 2 h. Fragments were visualised by electrophoresis in TAE using 2% agarose. Identical RFLPs were then identified by eye, and unique sequences were sent for sequencing to Lark Technologies Ltd (Essex).

Computer Analysis

The CHIMERA_CHECK program available from the Ribosomal Database Project (RDP) (http://rdp8.cme.msu.edu/cgis/news.cgi.) and Pintail [3] were used to check for chimeric sequences. Rarefaction curves were calculated from RFLP data using Analytical Rarefaction version 1.3 available at UGA Stratigraphy Lab (http://www.uga.edu/~strata/software/anRareReadme.html). 16S rRNA sequences

were searched using Blastn [1]. Sequences were aligned using MEGA version 3.1 [39]. The values for the Jaccard Index were determined using EstimateS (version 7.5; Department of Ecology and Evolutionary Biology, University of Connecticut [http://viceroy.eeb.uconn.edu/estimates]). The simple and partial Mantel Tests were carried out using the zt programme [8].

Definition of Operational Taxonomic Unit (OTU)

The 16S rRNA gene sequences were aligned using MEGA version 3.1 [39], and the output file was used to define Operational Taxonomic Units (OTUs) using DOTUR [59]. This was done using the furthest neighbour clustering algorithm (default setting). In this study, three of the commonly used OTU definitions were used (95%, 97% and 99%), which is equivalent to comparing taxonomic resolutions: at the genus, species and sub-species level [31]. From 217 non-chimeric sequences, 184 unique sequences were detected at an OTU definition of 99%, 135 unique sequences at 97% and 110 unique sequences at 95%.

Construction of Phylogenetic Trees

Phylogenetic analysis was done using MEGA version 3.1 [39] using the Jukes and Cantor nucleotide substitution model for sequence alignment and the Neighbour-Joining method of tree inference. The support for each node was determined by assembling a consensus tree of 1000 bootstrap replicates using the same phylogenetic settings. 16S rRNA gene sequences retrieved from the clone libraries were deposited into EMBL Nucleotide Sequence Database, with accession numbers FM210811 - FM211027.

Standard Normal Deviate Equivalents (SNDE)

Raw environmental data were standardised to make the different environmental factors comparable. This was done by the following equation: SNDE = (x – mean of the raw data) / standard deviation of the raw data, where x is the raw data for one sampling site.

Coverage

Library coverage was calculated using the following equation: $C = (1- (n_1 / N)) 100$ where n_1 is the number of RFLPs represented by a clone and N is the total number of clones in the library [24].

Results

Screening for Chimeras

The sequencing dataset (219 16S rRNA gene sequences) were screened for suspected chimeras using the CHIMERA_CHECK program available from the Ribosomal Database Project (RDP). Subsequently, 56 suspected chimeric sequences were analysed by Pintail [3]. No chimeric sequences were detected in the archaeal 16S rRNA gene libraries. However, 3 chimeric sequences were detected in the bacterial 16S rRNA gene libraries and were removed from further analysis. This observed frequency is less than previously reported [62]. However, since only partial sequences were used in the analysis, fewer chimeras were likely to be found. Chimeras are more likely to be detected in a dataset containing longer 16S rRNA gene sequences, i.e., if both ends of the 16S rRNA gene were sequenced.

Library Coverage

Rarefaction curves are used to identify when sampling is sufficient to determine species diversity with some level of confidence. The numbers of unique clones were plotted against the number of unique species. Fig 1 shows that library coverage was generally approaching plateau stages, with coverage calculated between 44 – 80%. The bacterial library from Lake Bagaejinnor was the clear exception, with coverage estimated at just 8%. Coverage in the archaeal libraries was generally higher than in the bacterial libraries, reflecting lower diversity in the former. Bacterial diversity in Lake Ejinnor was strikingly lower than in the other lakes, although the reason for this is unclear.

Archaeal Diversity

Phylogenetic analysis showed the assignment of clone sequences into seven monophyletic assemblages within the order *Halobacteriales* (Fig 2). Clone sequences branched within the *Halorubrum*, *Natronomonas*, *Halogeometricum*, *Halobaculum*, *Haloarcula*, *Halorhabdus* and *Halosimplex* lineages. The range of haloarchaea identified in this study was entirely consistent with the saline conditions of the environments sampled. There were additional, well supported lineages that formed between these nodes, designated Clusters 1-4 (bootstrap values between 74 and 99), which showed that these sequences were significantly different to any known species. Other known genera within the order *Halobacteriales* were represented in the tree; however none of the sequences in the Clusters affiliated closely with any of them. Cluster 1 contained sequences that were 99% identical to a clone found in crystalliser ponds in Australia [10]. Clusters 2, 3 and 4 all showed low sequence similarity (<98%) to uncultured organisms (data not shown), so were therefore unique to the sites sampled. Moreover, sequence 'EJ22' found in Lake Ejinnor did not affiliate with any lineage, and was therefore unique to this saltern.

Twenty-eight clone sequences branched with the *Halorubrum* lineage. Sequences from both Inner Mongolia and Argentina were found in this group, which demonstrated its ubiquitous nature. Sequences from Lake Chagannor were most similar to haloalkaliphilic *Halorubrum vacuolatum*, which was consistent with the highly alkaline pH of this lake (pH 10.5). Similarly, sequences from Lake Chagannor and the small pool at the unnamed lake (pH 9.5) were affiliated with the haloalkaliphilic group *Natronomonas*. Twenty-five clone sequences branched with Cluster 1, a well supported lineage that is phylogenetically distinct from the

Halorubrum branch (bootstrap value of 98). It was the second largest haloarchaeal group in this study, and again was ubiquitous in the habitats studied. Cluster 4 contains 16 clone sequences from both Inner Mongolia and Argentina. It formed a lineage on the periphery of the *Halobacteriales* that does not show any resemblance to known *Euryarchaeota*, which suggested that adaptation to hypersaline environments may extend to *Archaea* outside the *Halobacteriales*.

Bacterial Diversity

Phylogenetic analysis showed the distribution of clone sequences into seven monophyletic assemblages – *Gammaproteobacteria*, *Alphaproteobacteria*, *Deltaproteobacteria*, *Bacteroidetes*, *Chlorophyceae*, *Cyanobacteria*, and *Aquificales* - and two paraphyletic groups of the *Firmicutes* (Fig 3). In addition, there were well supported lineages that formed between these nodes, designated Clusters 1-4 (bootstrap values between 50 and 99). Cluster 3 contained a sequence that was 99% identical to a clone found in Lake Chaka, an athalassohaline lake in China [35]. Sequences in Clusters 1, 2 and 4 were unique to these sites as they all had low sequence similarities to uncultured organisms (<98% identity).

DNA from chloroplasts in *Eukarya* was clearly extracted in community DNA preparations, which resulted in several chloroplast 16S rRNA genes in the clone libraries (*Chlorophyceae*). These were related to chloroplasts found in *Dunaliella salina*, a typical salt lake inhabitant [49].

The *Proteobacteria* were the largest group, containing 38 clone sequences. This group was divided into the *Gammaproteobacteria* (28 sequences), *Deltaproteobacteria* (6

sequences) and *Alphaproteobacteria* (4 sequences). Sequences from both Inner Mongolia and Argentina were only affiliated with the former two divisions, while only Inner Mongolian sequences were affiliated with the *Alphaproteobacteria*. Many clone sequences were related to *Halomonas*, also typical inhabitants of salt lake environments [27]. The second largest group was the *Bacteroidetes*, which consisted of 31 clone sequences from both Inner Mongolia and Argentina. Clone sequences from the saltern at Lake Ejinnor were affiliated with *Salinibacter ruber*, an extremely halophilic bacterium [2]. The third largest group was the *Firmicutes*, which consisted of 27 clone sequences, all from Inner Mongolia. The majority of these clones were related to anaerobic species related to the *Halanaerobiaceae*, with three clones related to *Bacillus* sp.

Other clone sequences from Inner Mongolia, BJ67, EN105 and SH116 were also affiliated with or branched near to the deeply branching *Aquificae*, see Fig 3, which usually inhabit hot spring environments [55]. This lineage appeared to be unique to the Inner Mongolian sites as no other *Aquificales* sequences have been found at other salt lakes. Owing to the low sequence similarity to existing 16S rRNA gene sequences, it may not be appropriate to construe the growth temperature ranges of these bacteria. However recently, mesophilic members of a deeply branching group have been discovered [46], so perhaps these clone sequences represent novel lineages distantly related to *Aquificae* that are adapted to lower temperatures.

Community Composition and Biotic Similarity Matrices

Comparison of the overall archaeal community composition (Table 2) and bacterial community composition (Table 3) of the salt lakes demonstrates that they were

unique. This was assessed using the Jaccard Index, calculated using the program EstimateS for each definition of OTU. No pair of lakes scored a value of 1 (1 = identical biotic composition) as can be seen from the biotic similarity matrices.

Geographic-Distance Matrix

The matrix for the distances between each of the environments in this study is found in the supplementary data, Table 1. This was calculated using the GPS coordinates measured during the expedition (see Methods).

Environmental Similarity Matrices

To construct environmental similarity matrices, all raw values had to be standardised to make them comparable since this accommodates different units of different variables. Therefore, Standard Normal Deviate Equivalents (SNDE) values were calculated for temperature, pH, and concentrations of Na^{+}, Mg^{2+}, K^{+}, Cl^{-}, S, CO_3^{2-} and HCO_3^{-} ions (supplementary data, Tables 2-10). In order to construct a similarity matrix with this new standardised data, values of one minus the Euclidean distance between the SNDE values of two lakes was calculated. A value of 1 therefore indicated that a particular environmental factor was identical for the two lakes. The environmental similarity matrices are found in the supplementary data.

Biogeography of *Archaea*

The simple Mantel test was carried out using the zt program [8]. This program calculated whether biotic similarity correlated with environmental factors, and whether this correlation was statistically significant. The value of r^2 was the correlation value; positive or negative values reflected the type of relationship

between the two matrices, while *P* was the probability associated with r^2. Values of *P* were significant if they were less than 0.05; values greater than 0.05 indicated that the null hypothesis applied. (The null hypothesis stated that distances in matrix A were independent of the distances in matrix B).

Table 4 shows the results of the simple Mantel test. The results showed that geographic distance was not a significant factor in influencing the archaeal community composition, since *P* values across all OTU levels were >0.05. Therefore historical events and geographic barriers to dispersal have not affected archaeal community composition.

Only contemporary environmental factors appear to significantly influence the archaeal community composition. At the genus (95% OTU) and species level (97% OTU) CO_3^{2-}, HCO_3^- ion concentrations and pH were significant factors in influencing archaeal community composition since *P* values were <0.05. Na+ ion concentration was significant at the genus level only. All r^2 values were positive, hence biotic similarity of the salt lakes increased as the similarity of the environmental factors increased. At the sub-species level (99% OTU), pH and temperature were the only significant factors. Again, all r^2 values were positive integers. Removal of the Argentine data set from the analysis did not significantly change these results [50].

Biogeography of *Bacteria*

Table 5 shows the results of the simple Mantel test. The results showed that geographic distance was a significant factor in influencing the bacterial community composition at the genus level (95% OTU) and species level (97% OTU), since *P*

values across these OTU levels were <0.05. Both r^2 values were negative, hence biotic similarity of the salt lakes increased as the geographic distance similarity between the lakes increased; i.e., the closer the lakes, the more similar the bacterial populations. The larger the distance between sites, the higher the likelihood of barriers to dispersal. We therefore concluded that historical events and geographic barriers affected bacterial community composition. At the sub-species level (99% OTU), geographic distance was no longer a significant factor affecting bacterial community composition. Only environmental factors were significant: i.e., HCO_3- ion concentration and pH. Omitting data from lake BJ due to the low coverage values does not affect the results at the 99% or 97% OTU definitions, but at the 95% level, temperature and pH are no longer significant (data not shown).

In order to disentangle the effects of the environment versus geographic distance, the partial Mantel test was used [42]. The effects of the environment on biotic similarity were tested at the genus and species level, while controlling the effects of geographic distance (Table 6). Na+, HCO_3- ion concentrations and pH had a significant influence at the genus level (95% OTU) and species level (97% OTU), with *P* values <0.05. Mg^{2+} ion concentration was significant at the species level only. Curiously, both HCO_3- and Na+ ion concentrations had negative r^2 values.

When the Argentinean data were removed from the analysis, environmental factors were again found to be significant, but not geographic distance [50].

Discussion

Archaeal Biogeography

We report the phylogeny and distribution of *Archaea* in seven salt lakes across two continents at almost antipodean positions to each other. Statistical analyses demonstrated that this distribution was significantly influenced by environmental factors (Na^+, CO_3^{2-}, HCO_3- ion concentrations, pH and temperature), see Table 4. All r^2 values were positive integers, which indicated that as environmental similarity increased, biotic similarity increased. Geographic distance was not a significant factor. This was supported by phylogenetic analysis, which showed that the Argentinean sequences were interspersed throughout the phylogenetic tree; therefore no lineages were specific to either the Inner Mongolian or Argentinean salt lakes, see Fig 2. Our results are in contrast to a previous finding that the distribution of hyperthermophilic archaea showed a tendency for endemism [63], despite the fact that some archaeal species can be air-borne [52]. This implies that unlike hyperthermophilic archaea, haloarchaea are more robust over long distance travel - one apparent explanation for this is that hyperthermophilic archaea are less likely to survive at ambient temperatures.

The finding that pH was a significant environmental factor in influencing haloarchaeal biogeography at all three definitions of OTU was not unexpected. pH would allow different species to be selected in either slightly alkaline (pH 7.5) or highly alkaline (pH 10.5) environments, which was the pH range of the salt lakes in this study. Since pH is dependent on $CO_3^{2-}$and HCO_3^- ions, it was also not unexpected that these were also significant factors (although surprisingly, they were not significant at the 99%

OTU level). Phylogenetic analysis supported this finding as haloalkaliphilic species were only found in the alkaline lakes, while haloarchaeal species in the other lakes were consistent with environments of lower pH values. For example, clones relating to *Halorubrum vacuolatum* and *Natronomonas* sp. were only detected in Lake Chagannor (pH 10.5) and in the small pool at the unnamed lake (pH 9.5). However, it was unusual that other alkaliphilic groups within the *Halobacteriales* were not detected, such as those found, for example, in soda lakes in Kenya and Egypt [26, 36, 44, 53].

It appeared that temperature was a significant factor in driving haloarchaeal biogeography at the sub-species level only (99% OTU). This implied that seasonal changes in temperature were important in influencing haloarchaeal biogeography. Experiments with samples from a saltern showed that at 35°C, dense growths of haloarchaea were observed at 35% and 40% (w/v) salt, but at 25°C, very little haloarchaeal growth was observed [15]. Temperature may also play a role in competition - it was found that temperature was the deciding factor in competition between moderately halophilic bacteria and haloarchaea, with bacterial growth being favoured at lower temperatures [57].

Na^+ ion concentration was a significant factor affecting haloarchaeal community composition at the genus level (95% OTU). Haloarchaea adapt to high NaCl concentrations in the environment by accumulation of KCl to exclude NaCl from the cells thereby achieving osmotic equilibrium [40]. This is an adaptation that does not extend to *Archaea* outside the *Halobacteriales* [23]. In addition, some haloarchaeal enzymes have evolved a requirement for high Na+ concentrations [19, 45].

Bacterial Biogeography

Statistical analysis showed that the distribution of *Bacteria* in the six salt lakes across two continents was significantly influenced by geographic distance (Table 5). When this analysis was repeated for the Inner Mongolian samples alone, geographic distance was *not* a significant factor [data not shown; [50]], suggesting that geographic distance does not have a biogeographical effect at a local spatial scale. The strong winds observed on the steppe at the time of sampling could allow dispersal of microorganisms over long distances [18, 33, 37], and so the fact that geographic distance became a significant factor once the Argentinean data was added to the analysis suggested that there may be a tendency toward endemism in halophilic bacteria. The phylogenetic analysis implies certain lineages may be implicated in this endemism (Fig 3). The finding that geographic distance affects bacterial biogeography is consistent with the other studies e.g. [12, 22, 51], and has been explained by the fact that at large geographic distance, barriers to dispersal are more likely, and so evolutionary events such as speciation and extinction can give rise to differences in two populations separated by such barriers [32]. Our finding is consistent with previous studies on *Thioalkalivibrio* and *Salinibacter ruber*, which showed that strains of both bacterial species were endemic to certain regions, despite having cosmopolitan distributions [22, 58].

Geographic distance was not a significant factor at the sub-species level (99% OTU). Only pH and HCO_3- ion concentration were significant at this level of OTU. Again, the correlation to pH was not unexpected. The biogeography of bacteria in freshwater lakes has often been correlated with pH [41, 67]. It is not hard to imagine that the

same effect would occur with halophilic bacteria, but to our knowledge, no other studies have shown this biogeographical effect. Phylogenetic analysis supported this finding, which showed that clones from Lake Chagannor (pH 10.5) and the pool at the unnamed lake (pH 9.5) were related to haloalkaliphilic species such as *Halomonas campisalis* and "*Natronoanaerobium halophilum*."

Once the effects of geographic distance were controlled using the partial Mantel test, the contemporary environmental factors found to be significant in influencing bacterial biogeography at the genus level (95% OTU) and species level (97% OTU) were Na^+, Mg^{2+} and HCO_3- ion concentrations and pH. The r^2 values for pH and Mg^{2+} ion concentrations were positive (the higher the environmental similarity, the higher the biotic similarity), see Table 6. Surprisingly, the r^2 value for Na+ and HCO_3- ion concentrations were negative (the higher environmental similarity, the lower the biotic similarity), which does not fit what is currently known about microbial biogeography [42], where only a positive correlations with environmental parameters have been observed.

Na+ ions were significant in influencing bacterial community composition at both the genus (95% OTU) and species levels (97% OTU). However, unlike the haloarchaea, halophilic bacteria cope with high NaCl concentrations in the environment by accumulation of organic compatible solutes [56], with only a few exceptions (for example, the *Halanaerobiales* [19, 54] and *Salinibacter ruber* [48]). Other sodium salts may influence bacterial community composition since a previous study showed that the high salt requirement for a moderate halophilic bacterium was met by sodium salts other than NaCl [47]. In addition, Na^+ ions are important to some alkaliphilic

bacteria as they replace protons as the coupling ion to cope with the high external pH, rather than increasing the electric potential difference across the cytoplasmic membrane [60].

Mg^{2+} ion concentration was a significant factor in influencing the bacterial community composition at the species level only (97% OTU). Mg^{2+} favours the growth of haloarchaea [26], and so a possible explanation for this trend may be that only the halophilic bacteria that are tolerant to Mg^{2+} are able to proliferate and co-exist with the haloarchaea ($MgCl_2$ is a chaotropic agent, and is a limiting factor in the diversity of microbes in the environment [29]).

Conclusions

Unlike other saline and soda lakes, such as those in the East African Rift Valley, these particular lakes are formed in depressions entirely by runoff from the surrounding topography. As such they are influenced by seasons and vary in salinity depending on rainfall or spring melts. The lakes are shown as permanent sites on local maps of the areas. [68] [66, 69] We sampled these lakes in summer months at or close to maximum salinity levels which last for several months – the particular prokaryote population we detected reflects the particular set of conditions we measured at that time. Clearly, in an ideal world, sequential sampling over an extended time period would be appropriate, but logistical and financial considerations preclude repeated visits to these remote sites. The expedition undertaken was designed to access the lakes during a relatively stable period of water chemistry, but still must be a compromise in view of the seasonality of the sites.

606 Martiny *et al* [42] suggested that the relative effect of the environment on microbial
607 community composition relate to the geographical scale of sampling: at large
608 geographic scales, distance seems to influence the community composition more,
609 while environment seems not to have any effect; in contrast, at a small scale it is the
610 environment which has an effect but not distance [42]. However, due to the small
611 number of studies available thus far, this conclusion should be treated with caution.
612 For example, studies at large-scale distances have been carried out on prokaryotes
613 inhabiting extreme environments where the environmental parameters have a
614 relatively small range, and focus on a particular prokaryotic species [22, 51, 63] while
615 studies at small-scale distances target mixed bacterial populations [14]. It is possible
616 that this trend has arisen from the types of habitats sampled or the organism studied.
617
618 Similarly, we have presented a case where microbial biogeography is not so clear cut.
619 It can be argued that extremophiles have larger limits to dispersal due to the lack of a
620 suitable habitat. Certainly, this research has shown that bacterial biogeography is
621 significantly affected by geographic distance and therefore barriers to dispersal. This
622 implies that halophilic bacteria are less mobile or less robust across large distances.
623 On the other hand, geographic distance was not a significant factor in affecting
624 archaeal biogeography, but contemporary environmental factors were more
625 significant. This implies that *Archaea* are dispersed more easily or are more robust
626 over long distance travel. The study of microbial biogeography is still in its infancy.
627 Only further case studies will show any solid tends in this field.
628
629 **Acknowledgments**

630 This research was supported by the European Commission research programme
631 ‘Quality of life and management of living resources’, project Multigenome Access
632 Technology for Industrial Catalysts (QLRT-2001-01972). We thank Ms Huanzhi
633 Wang, a former MSc student at the University of Leicester for her help in sequencing
634 samples from X.
635

References

1. **Altschul, S.F., W. Gish, W. Miller, E.W. Myers and D.J. Lipman.** 1990. Basic local alignment tool. J. Mol. Biol. **215:**403-10.

2. **Antón, J., R. Rosselló-Mora, F. Rodríguez-Valera and R. Amann.** 2000. Extremely halophilic *Bacteria* in crystallizer ponds from solar salterns. Appl. Environ. Microbiol. **66:**3052-3057.

3. **Ashelford, K.E., N.A. Chuzhanova, J.C. Fry, A.J. Jones and A.J. Weightman.** 2005. At least 1 in 20 16S rRNA sequence records currently held in public repositories is estimated to contain substantial anomalies. Appl. Environ. Microbiol. **71:**7724-7736.

4. **Baas-Becking, L. G. M.** 1934. Geobiologie of inleidng tot milieukunde, p. 263. Serie 18/19, van Stockum's Gravenhange.

5. **Bano, N., S. Ruffin, B. Ransom and J.T. Hollibaugh.** 2004. Phylogenetic comparison of Arctic Ocean archaeal assemblages and comparison with Antarctic assemblages. Appl. Environ. Microbiol. **70:**781-789.

6. **Beijerinck, M. W.** 1913. De infusies en de ontdekking der backterien, p. 119-140. *In* Jaarboek van de Koninklijke Akademie voor Wetenschappen. F. Bruckmann, A. G. and J. B. Obernetter, Munchen, Germany.

7. **Benlloch, S., A. López-López, E.O. Casamayor, L. Øvreås, V. Goddard, F.L. Daae, G. Smerdon, R. Massana, I. Joint, F. Thingstad, C. Pedrós-Alió and F. Rodríguez-Valera.** 2002. Prokaryotic genetic diversity throughout the salinity gradient of a coastal solar saltern. Environ. Microbiol. **4:**349-360.

8. **Bonnet, E. and Y. Van der Peer.** 2002. zt: a software tool for simple and partial Mantel tests. J. Stat. Software. **7:**1-12.

9. **Brown, J. H., B. R. Riddle, and M. V. Lomolino,** 2005. Biogeography. Sinauer Associates Inc.

10. **Burns, D.G., H.M. Camakaris, P.H. Janssen and M.L. Dyall-Smith.** 2004. Combined use of cultivation-dependent and cultivation-independent methods indicates that members of most haloarchaeal groups in an Australian crystallizer pond are cultivable. Appl. Env. Microbiol. **70**:5258-5265.

11. **Casamayor, E.O., R. Massana, S. Benlloch, L. Øvreås, B. Díez, V.J. Goddard, J.M. Gasol, I. Joint, F. Rodríguez-Valera and C. Pedrós-Alió.** 2002. Changes in archaeal, bacterial and eukaryal assemblages along a salinity gradient by comparison of genetic fingerprinting methods in a multipond solar saltern. Environ. Microbiol. **4:**338-348.

12. **Cho, J.-C. and J.M. Tiedje.** 2000. Biogeography and degree of endemicity of fluorescent *Pseudomonas* strains in soil. Appl. Environ. Microbiol. **66:**5448–5456.

13. **Cox, C.B. and Moore, P.D.** 2000. Biogeography: An Ecological and Evolutionary Approach. Blackwell Science Ltd.

14. **Crump, B.C., C.S. Hopkinson, M.L. Sogin and J.E. Hobbie.** 2004. Microbial biogeography along an estuarine salinity gradient: combined influences of bacterial growth and residence time. Appl. Environ. Microbiol. **70:**1494-1505.

15. **Del Moral, A., E. Quesada, V. Bejar and A. Ramos-Cormenzana.** 1987. Evolution of bacterial flora from a subterranean saline well by gradual salinity changes in enrichment media. J. Appl. Bacteriol. **62:**465-471.

16. **Demergasso, C.S., E.O. Casamayor, G. Chong, P.A. Galleguillos, L.V. Escudero and C. Pedrós-Alió.** 2004. Distribution of prokaryotic genetic

686 diversity in athalassohaline lakes of the Atacama Desert, Northern Chile.
687 FEMS Microbiol. Ecol. **48:**57–69.
688 17. **Dimitriu, P. A., H. C. Pinkart, B. M. Peyton and M. R. Mormile.** 2008.
689 Spatial and temporal patterns in the microbial diversity of a meromictic soda
690 lake in Washington state. Appl. Environ. Microbiol. **74:** 4877-4888.
691 18. **Echigo, A., M. Hino, T. Fukushima, T. Mizuki, M. Kamekura and R.**
692 **Usami.** 2005. Endospores of halophilic bacteria of the family *Bacillaceae*
693 isolated from non-saline Japanese soil may be transported by Kosa event
694 (Asian dust storm). Saline Syst. **1:**1-13.
695 19. **Empadinhas, N. and M.S. da Costa**. 2008. Osmoadaptation mechanisms in
696 prokaryotes: distribution of compatible solutes. Int. Microbiol. **11**:151-161.
697 20. **Fierer, N. and R.B. Jackson.** 2006. The diversity and biogeography of soil
698 bacterial communities. Proc. Natl. Acad. Sci. USA. **103:**626-631.
699 21. **Finlay, B.J.** 2002. Global dispersal of free-living microbial eukaryote species.
700 Science. **296:**1061-1063.
701 22. **Foti, M., S. Ma, D.Y. Sorokin, J.L.W. Rademaker, J.G. Kuenen and G.**
702 **Muyzer.** 2006. Genetic diversity and biogeography of haloalkaliphilic
703 sulphur-oxidizing bacteria belonging to the genus *Thioalkalivibrio*. FEMS
704 Microbiol. Ecol. **56:**95-101.
705 23. **Galinski, E.A. and H.G. Trüper.** 1994. Microbial behaviour in salt-stressed
706 ecosystems. FEMS Microbiol. Rev. **15:**95-108.
707 24. **Good, I.J.** 1953. The population frequencies of species and the estimation of
708 population parameters. Biometrika. **40:**237-264.
709 25. **Grant, S., W.D. Grant, B.E. Jones, C. Kato and L. Li.** 1999. Novel archaeal
710 phylotypes from an East African alkaline saltern. Extremophiles. **3:**139-145.

711 26. **Grant, W.D., M., Kamekura, T. J. McGenity and A. Ventosa.** 2001. Class
712 III Halobacteria class nov., p. 294-334. *In* D.R. Boone, R. Castenholz and
713 G.M. Garrity (ed.), Bergey's manual of systematic bacteriology, 2nd ed., vol. 1
714 Springer, New York, NY.
715 27. **Grant, W.D.** 2004. Life at low water activity. Phil. Trans. R. Soc. Lond. B.
716 **359:**1249-1267.
717 28. **Gunde-Cimerman, N., A. Oren, and A. Plemenitaš**. 2005. Adaptation to life
718 at high salt concentrations in Archaea, Bacteria, and Eukarya. Springer.
719 29. **Hallsworth, J.E., M.M. Yakimov, P.N. Golyshin, J.L.M. Gillion, G.**
720 **D'Auria, F. de Lima Alves, V. La Cono, M. Genovese, B.A. McKew, S.L.**
721 **Hayes, G. Harris, L. Giuliano, K.N. Timmis and T.J. McGenity.** 2007.
722 Limits of life in $MgCl_2$-containing environments: chaotropicity defines the
723 window. Environ. Microbiol. **9:**801–813.
724 30. **Head, I.M., J.R. Saunders and R.W. Pickup.** 1998. Microbial evolution,
725 diversity and ecology: a decade of ribosomal RNA analysis of uncultivated
726 microorganisms. Microbiol. Ecol. **35:**1-21.
727 31. **Horner-Devine, M.C., M. Lage, J.B. Hughes and B.J.M. Bohannan.** 2004.
728 A taxa-area relationship for bacteria. Nature. **432:**750-753.
729 32. **Horner-Devine, M.C., K.M. Carney and B.J.M. Bohannan.** 2004. An
730 ecological perspective on bacterial biodiversity. Proc. R. Soc. Lond. B.
731 **271:**113–122.
732 33. **Hua, N.-P., F. Kobayashi, Y. Iwasaka, G.-Y. Shi and T. Naganuma.** 2007.
733 Detailed identification of desert-originated bacteria carried by Asian dust
734 storms to Japan. Aerobiologia. **23:**291-298.

735 34. **Humayoun, S.B., N. Bano and J.T. Hollibaugh.** 2003. Depth distribution of
736 microbial diversity in Mono Lake, a meromictic soda lake in California. Appl.
737 Environ. Microbiol. **69:**1030-1042.
738 35. **Jiang, H., H. Dong, G. Zhang, B. Yu, L.R. Chapman and M.W. Fields.**
739 2006. Microbial diversity in water and sediment of Lake Chaka, an
740 athalassohaline lake in northwestern China. Appl. Environ. Microbiol.
741 **72:**3832-3845.
742 36. **Jones, B.E., W.D. Grant, A.W. Duckworth and G.G. Owenson.** 1998.
743 Microbial diversity of soda lakes. Extremophiles. **2:**191-200.
744 37. **Junfeng, L.** 1997. Renewable energy development in China: resource
745 assessment, technology status, and greenhouse gas mitigation potential. Appl.
746 Energy. **56:**381-394.
747 38. **Kulp, T. R., S. Han, C. W. Saltikov, B. D. Lanoil, K. Zargar and R. S.**
748 **Oremland.** 2007. Effects of imposed salinity gradients on dissimilatory
749 arsenate reduction, sulphate reduction, and other microbial processes in
750 sediments from two Californian soda lakes. Appl. Environ. Microbiol. **73:**
751 5130-5137.
752 39. **Kumar, S., K. Tamura and M. Nei.** 2004. MEGA 3: Integrated software for
753 molecular evolutionary genetics analysis and sequence alignment. Brief
754 Bioinform. **5:**150-163.
755 40. **Lanyi, J.K.** 1974. Salt-dependent properties of proteins from extremely
756 halophilic bacteria. Bacteriol. Rev. **38:**272-290.
757 41. **Lindstöm, E.S., M.P.K.-v. Agterveld and G. Zwart**. 2005. Distribution of
758 typical freshwater bacterial groups is associated with pH, temperature and lake
759 water retention time. Appl. Environ. Microbiol. **71**:8201-8206.

42. **Martiny, J.B.H., B.J.M. Bohannan, J.H. Brown, R.K. Colwell, J.A. Fuhrman, J.L. Green, M.C. Horner-Devine, M. Kane, J.A. Krumins, C.R. Kuske, P.J. Morin, S. Naeem, L. Øvreås, A. Reysenbach, V.H. Smith and J.T. Staley.** 2006. Microbial biogeography: putting microorganisms on the map. Nature Rev. Microbiol. **4:**102-112.

43. **McCraig, A.E., L.A. Glover and J.I. Prosser.** 1999. Molecular analysis of bacterial community structure and diversity in unimproved and improved upland grass pastures. Appl. Environ. Microbiol. **65:**1721-1730.

44. **Mesbah, N.M., S.H. Abou-El-Ela and J. Wiegel**. 2007. Novel and unexpected prokaryotic diversity in water and sediments of the alkaline, hypersaline lakes of the Wadi An Natrun, Egypt. Microbiol. Ecol. **54**:598-617.

45. **Mevarech, M., F. Frolow and L.M. Gloss.** 2000. Halophilic enzymes: proteins with a grain of salt. Biophys. Chem. **86:**155-164.

46. **Nesbø, C.L., M. Dlutek, O. Zhaxybayeva and W.F. Doolittle.** 2006. Evidence for existence of "Mesotogas," members of the order *Thermotogales* adapted to low-temperature environments. Appl. Environ. Microbiol. **72:**5061-5068.

47. **O'Connor, K. and L.N. Csonka.** 2003. The high salt requirement of the moderate halophile *Chromohalobacter salexigens* DSM3043 can be met not only by NaCl but by other ions. Appl. Environ. Microbiol. **69:**6334–6336.

48. **Oren, A., M. Heldal, S. Norland and E.A. Galinski.** 2002. Intracellular ion and organic solute concentrations of the extremely halophilic bacterium *Salinibacter ruber*. Extremophiles. **6:**491-498.

49. **Oren, A.** 2005. A hundred years of *Dunaliella* research: 1905–2005. Saline Syst. **1:**1-14.

785 50. **Pagaling, E.** 2007. Ph.D. thesis. University of Leicester, UK.

786 51. **Papke, R.T., N.B. Ramsing, M.M. Bateson and D.M. Ward.** 2003.
787 Geographic isolation in hot spring cyanobacteria. Environ. Microbiol. **5:**650-
788 659.

789 52. **Radosevich, J.L., W.J. Wilson, J.H. Shinn, T.Z. DeSantis and G.L.**
790 **Andersen**. 2002. Development of a high-volume aerosol collection system for
791 the identification of air-bourne micro-organisms. Letters in Appl. Microbiol.
792 **34**:162-167.

793 53. **Rees, H.C., W.D. Grant, B.E. Jones and S. Heaphy**. 2004. Diversity of
794 Kenyan soda lake alkaliphiles assessed by molecular methods. Extremophiles.
795 **8**:63-71.

796 54. **Rengpipat, S., S.E. Lowe and J.G. Zeikus.** 1988. Effect of extreme salt
797 concentrations on the physiology and biochemistry of *Halobacteroides*
798 *acetoethylicus*. J. Bacteriol. **170:**3065-3071.

799 55. **Reysenbach, A.** 2001. Phylum BI. Aquificae phy. nov., p. 359-367. *In* D.R.
800 Boone, R. Castenholz and G. Garrity (ed.), Bergey's manual of systematic
801 bacteriology, 2nd ed., vol. 1. Springer, New York, NY.

802 56. **Roberts, M.F.** 2005. Organic compatible solutes of halotolerant and
803 halophilic microorganisms. Saline Syst. **1**(5).

804 57. **Rodríguez-Valera, F., F. Ruiz-Berraquero and A. Ramos-Cormenzana.**
805 1980. Behaviour of mixed populations of halophilic bacteria in continuous
806 cultures. Can. J. Microbiol. **26:**1259-1263.

807 58. **Rosselló-Mora, R., M. Lucio, A. Pena, J. Brito-Echeverría, A. López-**
808 **López, M. Valens-Vadell, M. Frommberger, J. Antón and P. Schmitt-**

Kopplin. 2008. Metabolic evidence for biogeographic isolation of the extremophilic bacterium *Salinibacter ruber*. The ISME Journal. **2:**242-253.

59. **Schloss, P.D. and J. Handelsman.** 2005. Introducing DOTUR, a computer program for defining Operational Taxonomic Units and estimating species richness. Appl. Environ. Microbiol. **71:**1501-1506.

60. **Skulachev, V.P., H. Kobayashi, T.A. Krulwich, G. Schafer, R.H. Fillingame, R.K. Poole, G.M. Cook, M.J. Dimroth, W.N. Konings and J.B. Stock.** 1999. Bacterial energetics at high pH: what happens to the H+ cycle when the extracellular H+ concentration decreases? Bacterial response to pH—Novartis Foundation Symposium. **221:**200–217.

61. **Venter, J.C., K. Remington, J.F. Heidelberg, A.L. Halpern, D. Rusch, J.A. Eisen, D. Wu, I. Paulsen, K.E. Nelson, W. Nelson, D.E. Fouts, S. Levy, A.H. Knap, M.W. Lomas, K. Nealson, O. White, J. Peterson, J. Hoffman, R. Parsons, H. Baden-Tillson, C. Pfannkoch, Y.-H. Rogers and H.O. Smith.** 2004. Environmental genome shotgun sequencing of the Sargasso Sea. Science. **304:**66-74.

62. **Wang, G.C. and Y. Wang.** 1996. The frequency of chimeric molecules as a consequence of PCR co-amplification of 16S rRNA genes from different bacterial species. Microbiology. **142:**1107-1114.

63. **Whitaker, R.J., D.W. Grogan and J.W. Taylor.** 2003. Geographic barriers isolate endemic populations of hyperthermophilic archaea. Science. **301:**976-978.

64. **Williams, W.D. and J.E. Sherwood.** 1994. Definition and measurement of salinity in salt lakes. Int. J. Salt Lake Res. **3:**53-63.

65. **Wu, Q.L., G. Zwart, M. Schauer, M.P.K.-v. Agterveld and M.W. Hahn.** 2006. Bacterioplankton community composition along a salinity gradient of sixteen high-mountain lakes located on the Tibetan plateau, China. Appl. Environ. Microbiol.**72:**5478-5485.

66. **Xiyu, Z., Z. Minggang and D. Jihe.** 1992. Salt Lakes in Inner Mongolia. Beijing: Science Press.

67. **Yannarel, A.C. and E.W. Triplett**. 2005. Geographic and environmental sources of variation in lake bacterial community composition. Appl. Environ. Microbiol. 71:227-239.

68. **Yu. G., S. P. Harrison and B. Xue.** 2001. Lake status records from China: data base documentation. MPI-BGC Tech Rep 4.

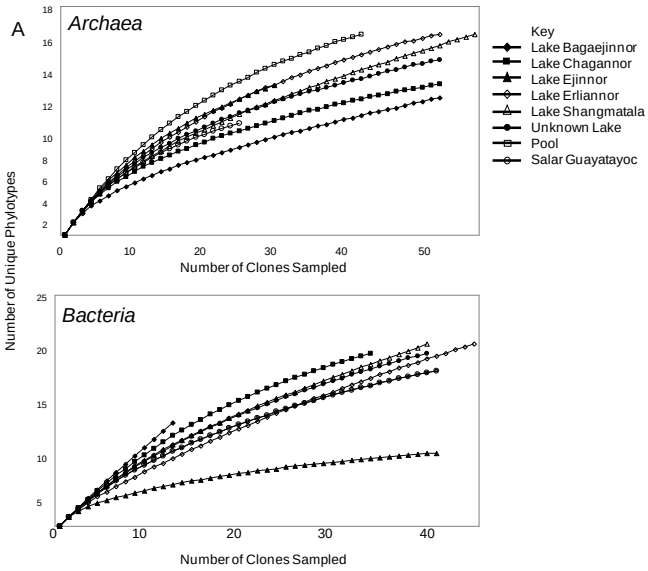
69. **Zheng, X., M. Zhang, J. Dong, Z. Gao, C. Xu, Z. Han, B. Zhang, D. Sun and K. Wang.** 1992. Salt lakes in Inner Mongolia of China. Science Press, Beijing

70. **Zhongking, L.** 1963. Stipa steppes in Inner Mongolia. Acta Phytoecologica et Geobotanica Sinica. No 1-2.

Table 1 Chemical composition of the salt lake brines.

Sample	Environment	Temp (°C)	pH	Conductivity (mS/cm) / Total salinity (g/L)	Cl (M)	K (mM)	Mg (M)	Na (M)	S (M)	HCO_3 (mM)	CO_3 (mM)
Lake Bagaejinnor	Salt lake	20.5	8.5	474 / 333	4.61	33.2	0.35	5.32	1.07	7.4	3.3
Lake Chagannor	Salt lake	17.1	10.5	202 / 118	1.08	14	0.001	2.89	0.43	360	410
Lake Ejinnor	Saltern	27.6	7.5	464 / 397	4.36	68.9	2.08	2.82	0.94	9.84	23.3
Lake Erliannor	Saltern	17.9	8	482 / 312	5.33	39	0.86	4.2	0.48	4.1	8.3
Lake Shangmatala	Salt lake	20.8	8.5	487 / 346	4.69	150	0.26	5.38	0.81	7.4	13
Unknown Lake	Salt lake	23.6	8.5	463 / 356	5.4	53.1	0.085	5.06	0.33	13.9	1.7
Pool	Small pool	21	9.5	287 / 196	2.21	17.1	0.035	2.01	0.06	14.3	0
Salar Guayatayoc	Salt lake	10	7.5	- / -	5.5	173	0.061	4.9	0.031	9.84	0

A



B

Environment	**Archaeal Library (%)**	**Bacterial Library (%)**
Lake Bagaejinnor	75	8
Lake Chagannor	73	44
Lake Ejinnor	52	80
Lake Erliannor	70	56
Lake Shangmatala	67	50
Unknown lake	68	53
Pool	54	56
Salar Guayatayoc	63	46

Fig 1 Rarefaction curves for sampling of the archaeal and bacterial 16S rRNA gene libraries (A), and a table of library coverage (B).

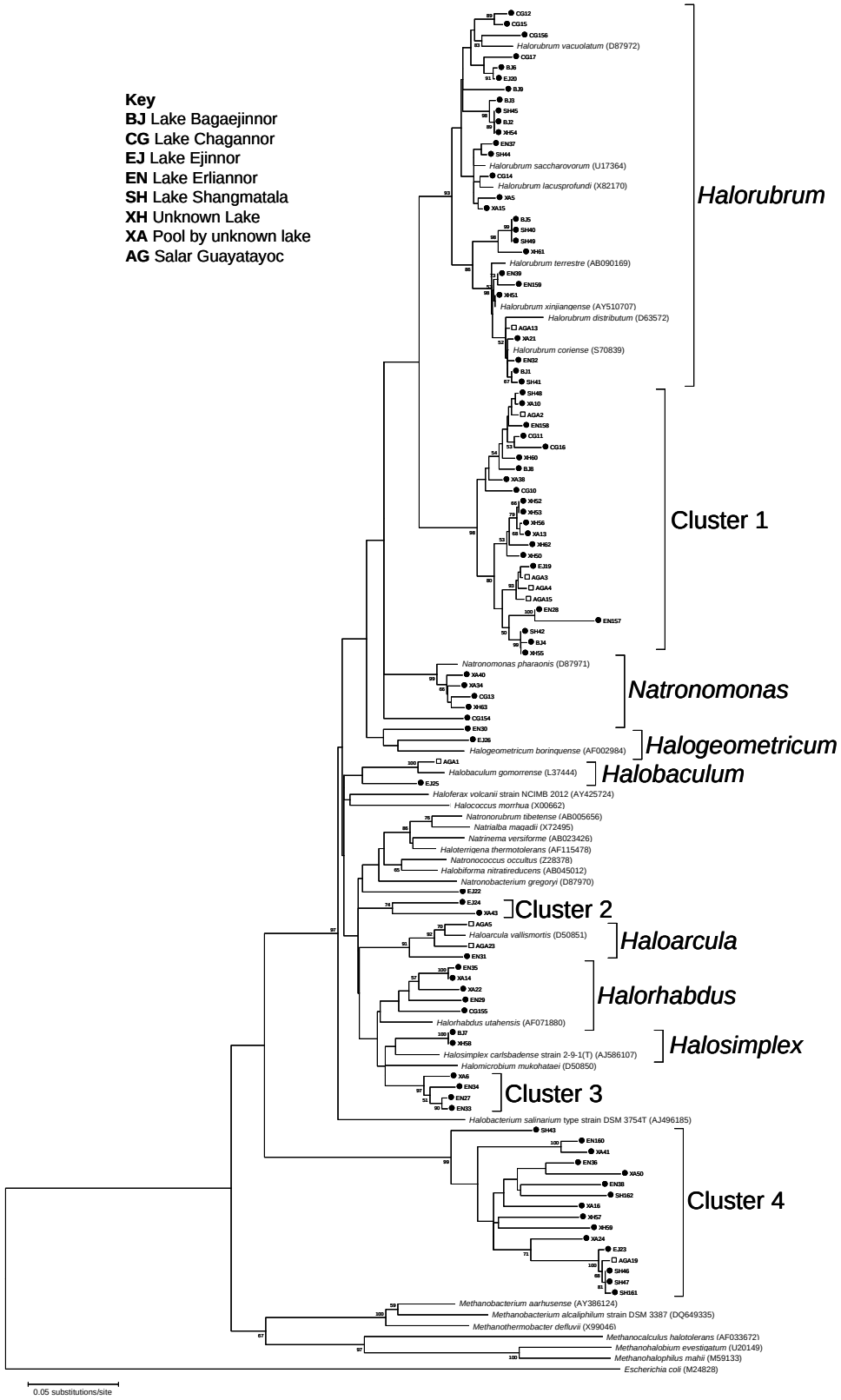


Fig 2 Phylogenetic tree of the archaeal population. Closed circles indicate sequences from Inner Mongolia, and open squares indicate sequences from Argentina.

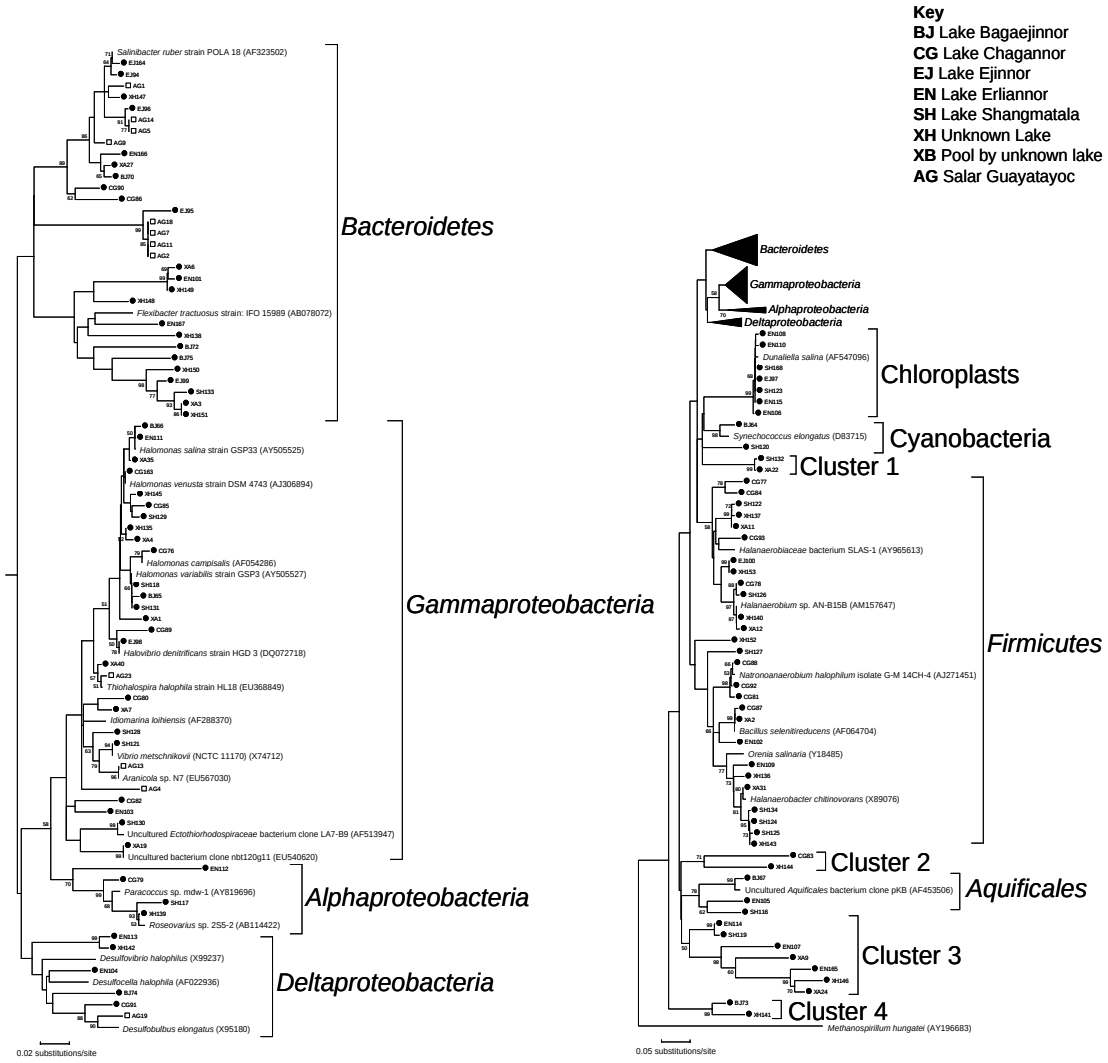


Fig 3 Phylogenetic tree of the bacterial population. Closed circles indicate sequences from Inner Mongolia, and open squares indicate sequences from Argentina. Black wedges on the right hand side are shown in more detail on the left hand side.

Table 2 Biotic similarity matrix for *Archaea*[a]

	BJ	**CG**	**EJ**	**EN**	**SH**	**XH**	**X**	**AG**
BJ	1							
CG	0 0 0	1						
EJ	0.066 0.071 0.166	0 0 0	1					
EN	0.041 0.045 0.117	0 0.043 0.052	0 0 0	1				
SH	0.187 0.307 0.454	0 0.058 0.062	0.062 0.066 0.230	0 0.142 0.166	1			
XH	0.157 0.285 0.545	0 0.055 0.058	0 0 0.133	0.035 0.086 0.1	0.095 0.266 0.357	1		
X	0.041 0.043 0.111	0 0.041 0.105	0 0 0	0.066 0.192 0.315	0 0.086 0.157	0.035 0.130 0.210	1	
AG	0 0.076 0.3	0 0 0	0 0.181 0.181	0 0.105 0.125	0 0.25 0.363	0 0.066 0.230	0 0.1 0.117	1

[a]These matrices show biotic similarity determined by the Jaccard Index for *Archaea* at three definitions of Operational Taxonomic Units (OTUs); listed in each box from top to bottom 99, 97 & 95% respectively. A value of 1 indicates identical microbial communities. BJ: Lake Bagaejinnor, CG: Lake Chagannor, EJ: Lake Ejinnor, EN: Lake Erliannor, SH: Lake Shangmatala, XH: unknown lake, X: pool by unknown lake, AG: Salar Guayatayoc.

Table 3 Biotic similarity matrix for *Bacteria*[a]

	BJ	**CG**	**EJ**	**EN**	**SH**	**XH**	**X**	**AG**
BJ	1							
CG	0 0.04 0.043	1						
EJ	0 0 0	0 0 0	1					
EN	0.04 0.043 0.09	0 0.032 0.071	0.041 0.047 0.05	1				
SH	0.037 0.041 0.086	0 0.064 0.107	0.038 0.045 0.047	0.028 0.068 0.107	1			
XH	0 0.076 0.076	0 0.058 0.062	0.038 0.04 0.086	0.028 0.096 0.096	0 0.129 0.206	1		
X	0 0.130 0.130	0 0.096 0.142	0 0 0	0 0.066 0.103	0 0.137 0.269	0.090 0.2 0.241	1	
AG	0 0 0	0 0 0	0.066 0.076 0.2	0 0 0	0 0 0	0 0 0.041	0 0 0.045	1

[a]These matrices show biotic similarity determined by the Jaccard Index for *Bacteria* at three definitions of Operational Taxonomic Units (OTUs); listed in each box from top to bottom 99, 97 & 95% respectively. A value of 1 indicates identical microbial communities. BJ: Lake Bagaejinnor, CG: Lake Chagannor, EJ: Lake Ejinnor, EN: Lake Erliannor, SH: Lake Shangmatala, XH: unknown lake, X: pool by unknown lake, AG: Salar Guayatayoc.

Table 4 Simple Mantel test for the archaeal population[a]

Factor	99% OTU		97% OTU		95% OTU	
	r	**p**	**r**	**p**	**r**	**p**
Cl^-	0.360	0.056	0.371	0.076	0.493	0.052
CO_3^{2-}	0.342	0.067	0.443	0.032	0.532	0.027
Distance	-0.336	0.116	0.094	0.473	0.087	0.486
HCO_3^-	0.336	0.118	0.440	0.032	0.527	0.035
K^+	0.072	0.373	-0.164	0.289	-0.110	0.391
Mg^{2+}	0.139	0.393	0.306	0.166	0.279	0.176
pH	0.431	0.010	0.497	0.005	0.478	0.019
S^{2-}	0.089	0.334	-0.045	0.415	-0.137	0.264
Temp	0.401	0.043	0.090	0.378	0.102	0.352
Na^+	0.245	0.113	0.322	0.051	0.343	0.038

[a] r^2 is the correlation value; positive or negative values reflect the type of relationship between the two matrices, while *P* is the probability associated with r^2. Values of *P* are significant if it is less than 0.05 (shaded).

Table 5 Simple Mantel test for the bacterial population[a]

Factor	99% OTU		97% OTU		95% OTU	
	r	**p**	**r**	**p**	**r**	**p**
Cl^-	0.226	0.147	-0.069	0.464	-0.043	0.479
CO_3^{2-}	0.350	0.090	0.135	0.275	0.166	0.399
Distance	-0.125	0.166	-0.484	0.0016	-0.328	0.002
HCO_3^-	0.362	0.002	0.131	0.319	0.168	0.370
K^+	0.118	0.290	0.210	0.180	0.000	0.417
Mg^{2+}	-0.230	0.127	0.282	0.159	0.226	0.145
pH	0.358	0.033	0.306	0.119	0.325	0.045
S^{2-}	0.068	0.359	-0.056	0.404	-0.083	0.367
Temp	-0.068	0.396	0.503	0.022	0.372	0.048
Na^+	-0.184	0.192	-0.381	0.002	-0.374	0.010

[a] r^2 is the correlation value; positive or negative values reflect the type of relationship between the two matrices, while *P* is the probability associated with r^2. Values of *P* are significant if it is less than 0.05.

Table 6 The partial Mantel test: testing the effect of the following factors, while controlling the effect of geographic distance[a]

Factor	97% OTU		95% OTU	
	r	**p**	**r**	**p**
CO_3^{2-}	0.235	0.190	0.228	0.262
HCO_3^-	-0.346	0.022	-0.342	0.024
pH	0.475	0.014	0.425	0.026
Temp	0.256	0.137	0.208	0.185
Na^+	-0.346	0.024	-0.342	0.025
Cl^-	-0.028	0.472	-0.013	0.507
K^+	-0.105	0.295	-0.253	0.091
Mg^{2+}	0.426	0.032	0.305	0.059
S^{2-}	-0.164	0.214	-0.151	0.237

[a] r^2 is the correlation value; positive or negative values reflect the type of relationship between the two matrices, while *P* is the probability associated with r^2. Values of *P* are significant if it is less than 0.05 (shaded).

Supplementary data

Table 1 Geographic Distance (km) matrix

	BJ	**CG**	**EJ**	**EN**	**SH**	**XH**	**X**	**AG**
BJ	X							
CG	360.5	X						
EJ	12.5	362.2	X					
EN	395.7	88.4	395.2	X				
SH	284.9	89.8	288.8	165.1	X			
XH	158.6	227.2	165.4	286.6	140.2	X		
X	157.6	229.1	164.5	297.8	142	0.147	X	
AG	17632.3	17845.5	17593.6	17763.8	17805.7	17733.2	17742.7	X

Table 2 Temperature (°C) matrix

	BJ	**CG**	**EJ**	**EN**	**SH**	**XH**	**X**	**AG**
BJ	X							
CG	0.3411	X						
EJ	-0.376	-1.035	X					
EN	0.438	0.903	-0.938	X				
SH	0.9419	0.283	-0.318	0.38	X			
XH	0.3992	-0.26	0.2248	-0.16	0.457	X		
X	0.9031	0.244	-0.279	0.341	0.961	0.496	X	
AG	-1.035	-0.38	-2.411	-0.473	-1.09	-1.64	-1.13	X

Table 3 pH matrix

	BJ	**CG**	**EJ**	**EN**	**SH**	**XH**	**X**	**AG**
BJ	X							
CG	-1.077	X						
EJ	-0.039	-2.116	X					
EN	0.221	-1.856	0.74	X				
SH	1	-1.077	-0.039	0.221	X			
XH	1	-1.077	-0.039	0.221	1	X		
X	0.803	-1.27	0.159	0.418	0.803	0.803	X	
AG	-0.039	0.7403	1	0.74	-0.039	-0.039	0.159	X

Table 4 Carbonate ion concentration (mM) matrix

	BJ	**CG**	**EJ**	**EN**	**SH**	**XH**	**X**	**AG**
BJ	X							
CG	-1.851	X						
EJ	0.86	-1.71	X					
EN	0.965	-1.816	0.895	X				
SH	0.932	-1.783	0.928	0.9671	X			
XH	0.989	-1.862	0.849	0.9537	0.9208	X		
X	0.977	-1.87	0.837	0.942	0.909	0.988	X	
AG	0.977	-1.874	0.837	0.942	0.9089	0.9881	1	X

Table 5 K+ concentration (mM) matrix

	BJ	**CG**	**EJ**	**EN**	**SH**	**XH**	**X**	**AG**
BJ	X							
CG	0.6975	X						
EJ	0.5446	0.242	X					
EN	0.9186	0.616	0.626	X				
SH	-0.918	-1.22	-0.463	-0.84	X			
XH	0.4215	0.387	0.8554	0.771	-0.61	X		
X	0.7324	0.965	0.2769	0.651	-1.19	0.422	X	
AG	-1.324	-1.63	-0.868	-1.242	0.594	-1.01	-1.59	X

Table 6 Mg^{2+} concentration (mM) matrix

	BJ	**CG**	**EJ**	**EN**	**SH**	**XH**	**X**	**AG**
BJ	X							
CG	0.5084	X						
EJ	-1.437	-1.929	X					
EN	0.2815	-0.21	-0.719	X				
SH	0.8732	0.635	-1.564	0.155	X			
XH	0.6265	0.882	-1.811	-0.09	0.753	X		
X	0.5564	0.952	-1.881	-0.16	0.683	0.93	X	
AG	0.5929	0.916	-1.844	-0.126	0.72	0.966	0.964	X

Table 7 Cl- concentration (M) matrix

	BJ	**CG**	**EJ**	**EN**	**SH**	**XH**	**X**	**AG**
BJ	X							
CG	-1.171	X						
EJ	0.8463	-1.017	X					
EN	0.5573	-1.61	0.4035	X				
SH	0.9508	-1.22	0.7971	0.606	X			
XH	0.5142	-1.66	0.3605	0.957	0.563	X		
X	-0.476	0.305	-0.322	-0.92	-0.52	-0.96	X	
AG	0.4527	-1.72	0.299	0.8955	0.502	0.939	-1.02	X

Table 8 S^{2-} concentration (M) matrix

	BJ	**CG**	**EJ**	**EN**	**SH**	**XH**	**X**	**AG**
BJ	X							
CG	-0.645	X						
EJ	0.6659	-0.311	X					
EN	-0.517	0.871	-0.182	X				
SH	0.3317	0.023	0.6659	0.152	X			
XH	-0.902	0.743	-0.568	0.614	-0.23	X		
X	-1.596	0.049	-1.262	-0.08	-0.93	0.306	X	
AG	-1.671	-0.03	-1.336	-0.154	-1	0.231	0.925	X

Table 9 Bicarbonate ion concentration (mM) matrix

	BJ	**CG**	**EJ**	**EN**	**SH**	**XH**	**X**	**AG**
BJ	X							
CG	-1.845	X						
EJ	0.980	-1.825	X					
EN	0.973	-1.871	0.954	X				
SH	1	-1.845	0.980	0.973	X			
XH	0.948	-1.792	0.967	0.921	0.948	X		
X	0.944	-1.789	0.964	0.918	0.944	0.997	X	
AG	0.980	-1.825	1	0.954	0.980	0.967	0.964	X

Table 10 Na+ concentration (M) matrix

	BJ	**CG**	**EJ**	**EN**	**SH**	**XH**	**X**	**AG**
BJ	X							
CG	-0.843	X						
EJ	-0.897	0.947	X					
EN	0.150	0.0062	-0.047	X				
SH	0.954	-0.889	-0.942	0.105	X			
XH	0.803	-0.646	-0.699	0.348	0.757	X		
X	-1.511	0.332	0.386	-0.661	-1.557	-1.314	X	
AG	0.681	-0.525	-0.578	0.469	0.636	0.879	-1.192	X