

ORIGINAL ARTICLE

From community approaches to single-cell genomics: the discovery of ubiquitous hyperhalophilic *Bacteroidetes* generalists

María Gomariz^{1,2}, Manuel Martínez-García², Fernando Santos², Francisco Rodríguez³, Salvador Capella-Gutiérrez⁴, Toni Gabaldón^{4,5}, Ramon Rosselló-Móra⁶, Inmaculada Meseguer¹ and Josefa Antón²

¹Department of Materials, Optics and Electronics, University Miguel Hernández of Elche, Alicante, Spain;

²Department of Physiology, Genetics, and Microbiology, University of Alicante, Alicante, Spain;

³Departament of Applied Mathematics, University of Alicante, Alicante, Spain; ⁴Bioinformatics and Genomics Programme, Centre for Genomic Regulation (CRG), Barcelona, Spain; ⁵Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra (UPF), Barcelona, Spain and ⁶Marine Microbiology Group, Department of Ecology and Marine Resources, Institut Mediterrani d'Estudis Avançats IMEDEA (CSIC-UIB), Esporles, Spain

The microbiota of multi-pond solar salterns around the world has been analyzed using a variety of culture-dependent and molecular techniques. However, studies addressing the dynamic nature of these systems are very scarce. Here we have characterized the temporal variation during 1 year of the microbiota of five ponds with increasing salinity (from 18% to >40%), by means of CARD-FISH and DGGE. Microbial community structure was statistically correlated with several environmental parameters, including ionic composition and meteorological factors, indicating that the microbial community was dynamic as specific phylotypes appeared only at certain times of the year. In addition to total salinity, microbial composition was strongly influenced by temperature and specific ionic composition. Remarkably, DGGE analyses unveiled the presence of most phylotypes previously detected in hypersaline systems using metagenomics and other molecular techniques, such as the very abundant *Haloquadratum* and *Salinibacter* representatives or the recently described low GC *Actinobacteria* and *Nanohaloarchaeota*. In addition, an uncultured group of *Bacteroidetes* was present along the whole range of salinity. Database searches indicated a previously unrecognized widespread distribution of this phylotype. Single-cell genome analysis of five members of this group suggested a set of metabolic characteristics that could provide competitive advantages in hypersaline environments, such as polymer degradation capabilities, the presence of retinal-binding light-activated proton pumps and arsenate reduction potential. In addition, the fairly high metagenomic fragment recruitment obtained for these single cells in both the intermediate and hypersaline ponds further confirm the DGGE data and point to the generalist lifestyle of this new *Bacteroidetes* group.

The ISME Journal (2015) 9, 16–31; doi:10.1038/ismej.2014.95; published online 13 June 2014

Introduction

Multi-pond solar salterns consist of a series of interconnected ponds with increasing salinity, from seawater level to sodium chloride saturation (Benlloch *et al.*, 2002; Pašić *et al.*, 2005; Oh *et al.*, 2010; Boujelben *et al.*, 2012). In continuously operated salterns, such as the Santa Pola (Spain) facilities studied here, seawater is pumped through a set of shallow ponds where sequential precipitation of calcium carbonate, calcium sulfate

and sodium chloride occurs. This salinity gradient is accompanied by a succession of different groups of microorganisms, from marine to extremely halophilic microbiota, each adapted to a salinity range.

Saltern ponds have been traditionally considered 'steady-state' systems because of the stability of both salinity and microbial composition of the different ponds (Guixa-Boixereu *et al.*, 1996; Ghai *et al.*, 2011). Although it is generally assumed that the composition of saltern microbial communities is 'fairly stable over time even when subjected to environmental perturbations' (Gasol *et al.*, 2004; Ghai *et al.*, 2011), most studies do not include different time points and provide only snapshots of the microbial diversity without addressing *in situ*

Correspondence: J Antón, Department of Physiology, Genetics, and Microbiology, University of Alicante, Alicante 03080, Spain. E-mail: anton@ua.es

Received 7 October 2013; revised 8 May 2014; accepted 10 May 2014; published online 13 June 2014

temporal dynamics. However, some studies hint at a more dynamic picture of these systems. A metagenomic analysis of the San Diego salterns community (Rodríguez-Brito *et al.*, 2010) using samples harvested at various time points separated by 1 day to >1 year showed that most abundant microbial 'species' persisted, but their relative abundances and microdiversity changed over time. Notwithstanding, this study only included two time points (01 July 2004 and 28 November 2005) for ponds of medium high salinities (above 18%), and thus its conclusions cannot be extrapolated to wider temporal ranges for this kind of environments. More recently, a study of the Tunisian solar salterns of Sfax (Boujelben *et al.*, 2012) also found a temporal succession of phylotypes in medium salinity (around 20%) ponds, with some specific bacterial sequences appearing only at some time points along the year.

In order to gain insight into the changes experienced by microbial communities in saltern ponds we have undertaken a comprehensive seasonal study of five ponds of medium and high salinity (from 18% to 40% total salts) from Santa Pola solar salterns. The prokaryotic community of these saltern ponds has been studied for decades using several approaches, depending on the available tools, from classical cultivation to culture-independent techniques. The lower-medium (15–25%) salinity ponds are inhabited by members of the *Proteobacteria*, *Cyanobacteria*, *Bacteroidetes* or haloarchaea-like *Halosarcina* or *Halorubrum*, whereas crystallizers (ponds where NaCl precipitates) are dominated by high numbers of the archaeon *Haloquadratum walsbyi* and extremely halophilic *Bacteroidetes*, such as *Salinibacter ruber*, that account for up to 75% and 18% of the total community, respectively (Antón *et al.*, 1999; Benlloch *et al.* 2002; Burns *et al.* 2004; Estrada *et al.*, 2004; Ghai *et al.*, 2011; Narasingarao *et al.*, 2012; Zhaxybayeva *et al.*, 2013). Recently, the use of Next-Generation Sequencing (NGS) metagenomics and single-cell genomics has led to the discovery of new abundant groups (Ghai *et al.*, 2011; Narasingarao *et al.*, 2012), such as a group of low GC *Actinobacteria* in low and intermediate salinities, and a new archaeal lineage of *Nanohaloarchaea* in crystallizers.

Our results indicate that saltern microbiota is dynamic and within a given pond microbial composition can change considerably along the year. In addition, total salinity is not the only factor structuring microbial assemblages in hypersaline environments as ponds of the same salinity harbored different microbial communities. Some microbial phylotypes were strongly associated with specific environmental conditions while others were very widely distributed both spatially and temporally. Especially noteworthy was the case of an uncultured *Bacteroidetes* phylotype that was present along the whole salinity gradient and could thus constitute a generalist strategist. Here, we have

further characterized this bacterial group by means of database search and single-cell genome analyses.

Materials and methods

Sample collection

Five contiguous ponds of Bras del Port salterns (Santa Pola, Spain, 38° 12' N, 0° 36' W) were analyzed in this study (Figure 1): two medium concentrators (CM1 and CM2), a brine concentrator (CCAB), and two crystallizers (CR30 and CR41). Samples were collected on the following dates in 2006: 23 January (JN), 7 March (MR), 26 April (AP), 6 June (J1), 27 June (J2), 25 July (JL), 5 September (SP), 2 October (OC) and 28 November (NV).

Environmental parameters

Temperature, pH and salinity were measured *in situ* while solar radiation and rainfall data were provided by the meteorological station of Bras del Port. Chemical analyses (sulfate, chloride, carbonate, phosphate, sodium, potassium, calcium, magnesium and ammonium) were carried out by atomic absorption spectrometry (Unicam SOLAAR 969, Unicam Ltd., Cambridge, UK) at the Department of Agrochemistry and Environment (University Miguel Hernández de Elche, Spain).

CARD-FISH

Archaeal and bacterial counts were measured following an optimized CARD-FISH protocol modified from Pernthaler *et al.* (2004). Samples were fixed with formaldehyde (7% final concentration; Sigma, St Louis, MO, USA) at 4 °C for 16 h, diluted



Figure 1 Aerial photograph of Bras del Port saltern ponds, Santa Pola, Alicante. Water is pumped from the sea to preconcentrators, medium concentrators, a brine concentrator and finally to the crystallizers. The ponds analyzed in this study are framed with dotted white lines: CM1 and CM2 (medium concentrators), CCAB (brine concentrator) and CR30 and CR41 (crystallizers). Arrows indicate the water circuit.

(1/1000, 1/500 and 1/200) in 1X phosphate-buffered saline (8 g NaCl, 1.44 g Na₂HPO₄, 2 g KCl, 2 g KH₂PO₄, pH 7.0) and filtered through 0.2- μ m pore diameter GTTP Isopore filters (Millipore, Madrid, Spain). The filters were stored at -20°C until used. Cell permeabilization was carried out with proteinase K (150 $\mu\text{g ml}^{-1}$ final concentration; Promega, Madrid, Spain) in Tris-HCl 20 mM pH 8.0 at 37°C for 5 min, followed by incubation with lysozyme (3 mg ml^{-1} final concentration; Sigma) in TE buffer pH 8.0 at 37°C for 60 min. Peroxidase inactivation was performed in 0.1 M HCl at room temperature for 1 min. Probes Eub338-I-III and Arc915 were used for *Bacteria* (Amann *et al.*, 1995; Daims *et al.*, 1999) and *Archaea* (Massana *et al.*, 1997), respectively. Probe Non338 was used to check for non-specific hybridization (Wallner *et al.*, 1993). Hybridization was performed in filters for 150 min at 35°C in an isotonic humid chamber. For signal amplification, filters were incubated with fluorescein-tyramides conjugates at 37°C for 30 min. Finally, filters were stained with DAPI (4',6-diamidino-2-phenylindole-dihydrochloride; 1 $\mu\text{g ml}^{-1}$; Sigma) for 1 min and observed using a Leica DMLA microscope (Leica Microsystems, Barcelona, Spain).

Nucleic acids extraction

Samples (500 ml) were centrifuged in a Beckman Coulter Avanti J-25centrifuge (Beckman Instruments, GmbH, Munich, Germany), with a JLA-16 rotor at 38 000 g , for 15 min at 4°C . Cell pellets were resuspended in 600 μl of extraction buffer (100 mM Tris-HCl pH 8.0, 100 mM EDTA, pH 8.0) in 2 ml phenol-resistant tubes. Lysozyme (3 mg ml^{-1} final concentration) was added, and samples were incubated for 15 min at 37°C with agitation (180 r.p.m.). Then proteinase K (150 $\mu\text{g ml}^{-1}$ final concentration) and sodium dodecyl sulfate 10% (6 $\mu\text{g ml}^{-1}$ final concentration) were added, and the samples were incubated for 30 min at 37°C (180 r.p.m.); 120 μl of 5 M NaCl and 120 μl of hexadecyltrimethylammonium bromide (10% CTAB, 0.7 M NaCl) were added to the samples and incubated at 65°C for 10 min, followed by three cycles of freezing in liquid nitrogen and thawing in a water bath at 65°C . Finally, 900 μl of phenol:chloroform:isoamyl alcohol (25:24:1) were added to each sample and then centrifuged for 5 min at 16 000 g (4°C). After centrifugation, the aqueous phases were transferred to clean tubes, and the phenol extraction was repeated until the interphase between the aqueous and organic phases was completely clean (two or three times approximately). Nucleic acids were precipitated by adding 0.1 volumes of sodium acetate 3 M (pH 4.8) and 0.6 volumes of cold isopropanol (-20°C) and centrifuged for 30 min at 16 000 g (4°C). Nucleic acids were then washed with 70% ethanol and centrifuged for 15 min at 16 000 g (4°C). Pellets were dried for 30 min to evaporate residual ethanol and resuspended in 50–100 μl of milli-Q

water. The extracted nucleic acids were analyzed by electrophoresis in 0.8% agarose gels (LE, FMC Bioproducts, Rockland, ME, USA) and visualized under ultraviolet light after staining with ethidium bromide (100 $\mu\text{g ml}^{-1}$).

PCR amplification and DGGE analysis of 16S rRNA genes

Archaeal and bacterial 16S rRNA genes were partially amplified with forward primers 341f-GC and 344f-GC, respectively, and the reverse universal primer 907R (Muyzer *et al.*, 1996; Schäfer *et al.*, 2001). Each PCR reaction contained MgCl₂ 1.5 mM, Tris-HCl 10 mM pH 9.0, KCl 50 mM, 200 μM of each dNTP, 0.2 M of each primer, 1 U Taq DNA polymerase recombinant (Invitrogen, Prat de Llobregat, Spain) and approximately 100 ng of template DNA. The PCR reactions were performed in a PTC-100 (Peltier-Effect Cycling). The PCR program for *Bacteria* was: 94°C for 5 min, 65°C for 1 min, 72°C for 3 min and 9 touchdown cycles of: 94°C for 1 min, 65°C (with a decreasing of 1°C in each cycle) for 1 min, 72°C for 3 min, followed by 20 cycles of: 94°C for 1 min, 55°C for 1 min, 72°C for 3 min. The PCR program for *Archaea* was: 94°C for 5 min and 30 cycles of: 94°C for 30 s, 56°C for 45 s, 72°C for 2 min. To minimize double-band formation, both programs included a final extension of 30 min at 72°C (Janse *et al.*, 2004). A reconditioning PCR was performed to avoid heteroduplex formation using a 10-fold dilution of the original PCR product (Thompson *et al.*, 2002). Reconditioning PCR conditions were: 94°C for 3 min, 3 cycles of 94°C for 15 s, 55°C for 30 s, 72°C for 2 min and finally one cycle of 7 min at 72°C . PCR products were concentrated and purified with Gel Band Purification Kit (GE Healthcare, Madrid, Spain), eluted in 20–30 μl milli-Q water and quantified using NanoDrop ND-1000 spectrophotometer (Thermo, Wilmington, MA, USA).

DGGE was performed by using the D-Code System (Bio-Rad, Madrid, Spain). PCR products were loaded onto 6% (w/v) polyacrylamide (acrylamide:bis-acrylamide gel stock solution 37.5:1; Bio-Rad), with 45–65% denaturing gradient (where 100% of denaturant contains 7 M urea and 40% formamide) in 1X TAE buffer (40 mM Tris, pH 8.0; 20 mM acetic acid; 1 mM EDTA) and subjected to 18 h of electrophoresis at 60°C and 70 V. After electrophoresis, the gels were stained with SYBR Green (1:10 000; Fluka, Madrid, Spain) for 15 min, washed in TAE 1X for 30 min and visualized using a computer image analyzer Typhoon 9410 (Amersham Biosciences, Cerdanyola del Valles, Spain). Bands were cut, transferred to sterile 1.5-ml tubes with 20 μl milli-Q water and incubated overnight at 4°C . Two microliters (approximately 100 ng) of each band were then re-amplified with the same primer pairs. Purity of amplification products was checked by DGGE and, finally, products yielding a single band were sequenced in the Technical Services Unit of

the University of Alicante, using a GeneAmp PCR System 2400 and an ABI PRISM 310 Genetic Analyzer (Applied Biosystems, Alcobendas, Spain). PCR products were sequenced with primer 907R in an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems). For an initial affiliation, sequences were checked at the GenBank nucleotide database using the BLAST software at the National Centre of Biotechnology Information website (<http://www.ncbi.nlm.nih.gov/>) (Altschul *et al.*, 1997). The tree reconstruction was performed by using the ARB software package (Ludwig *et al.*, 2004), with the SILVA REF 111 database (Pruesse *et al.*, 2007). The tree reconstruction was made using the neighbor-joining algorithm using a 40% conservational filter as implemented in the ARB package. To stabilize the tree, a selection of about 700 almost complete high-quality sequences representing a wide range of phyla within *Bacteria* and *Archaea* were included. For the final tree, only sequences of *Chlorobi* were kept as outgroup. Partial sequences were aligned against the same database using the SINA aligner, and the aligned sequences were inserted in the preexisting neighbor-joining tree by using the parsimony tool, all implemented in the ARB package. The DGGE band sequences obtained here have been deposited in the GenBank database (accession nos. FJ648083-FJ648198 and KF679127-KF679133).

SAG (single amplified genome) genomic sequencing, assembly and annotation

A 50- μ l water sample taken from CR30 in June 2011 was used for single-cell genome analysis. At that time, the pond had a salinity of 37.2% and harbored 1.74×10^7 cells ml^{-1} . Single-cell sorting was performed at the Single Cell Genomics Center (Bigelow Laboratory, Bigelow, ME, USA) as described elsewhere (Martínez-García *et al.*, 2012a, b). In brief, prokaryotic community was DNA-stained with SYTO-9, and single cells with high nucleic acid content were sorted with a MoFlo (Beckman Coulter) flow cytometer. Then, single cells deposited in 384-well plates were lysed with KOH, and their genomes were amplified by a multiple displacement amplification (Raghunathan *et al.*, 2005). These generated SAGs were PCR screened for 16S rRNA genes by using the primer sets 27F (5'-AG RGTTYGATYMTGGCTCAG-3') and 907R (5'-CCGT CAATTCMTTTRAGTTT-3') for *Bacteria* and Arc-344F (5'-ACGGGGYGCAGCAGCGCGCA-3') and Arch-915R (5'-GTGCTCCCCGCCAATTCCT-3') for *Archaea*. Sixteen S rRNA gene sequences of *Bacteroidetes* SAGs have been deposited in the GenBank database (accession nos. KF679134–KF679143). Finally, five SAGs belonging to the uncultured BC3 group were selected, and their genomes were sequenced by using paired-end Hi-Seq 2000 Illumina Technology (Solexa, San Diego, CA, USA) according to the manufacturer's protocol at the Center for Genomic Regulation (Barcelona, Spain).

Reads were 96 bases long and had an average insert size of 550 bases. Raw Illumina reads, which have been deposited in SRA (BioSample accessions: SAMN02434673, SAMN02434674, SAMN02434675, SAMN02434676, SAMN02434677) were preprocessed before assembly to trim either at the first undetermined base or at the first base having a PHRED quality below 20. The pairs having at least one read shorter than 30 (or 20 for SAG AB578_D06) bases after trimming were excluded from the assembly process (see Supplementary Table S1). SOAPdenovo2 (Luo *et al.*, 2012) was used to assemble paired-end reads into contigs. In order to ensure the minimum number of contigs with the highest N50 and N90 values, different combinations of kmer sizes were tested (range 21–41, in steps of 2) and merging rounds (from 0 to 3) allowing up to three mismatches during the merging process. Final assembly statistics are shown in Supplementary Table S2.

Genomic annotation was performed at the US Department of Energy Joint Genome Institute (JGI-DOE) with the pipeline annotation Prodigal (Hyatt *et al.*, 2010). These sequence data annotations were produced by the JGI-DOE (<http://www.jgi.doe.gov/>) in collaboration with the user community, and the corresponding SAGs were deposited and made accessible at the Integrated Microbial Genomes of JGI with the following IMG genome Taxon ID numbers: AB577-O10 (2522125006), AB578-D06 (2522125009), AB578-K19 (2522125007), AB578-L11 (2522125008) and AB577-P23 (3300002078). Analyses of average nucleotide identities values were performed by using JSpecies (Richter and Rosselló-Móra, 2009).

Metagenomic analyses

The basic approach of Rusch *et al.* (2007) was used to estimate the abundance of SAGs in the previously described metagenomes SS19 and CR30 (Ghai *et al.*, 2011). BLAST+ v2.2.22 was used to recruit metagenome sequences to each SAG assembly using the following parameters: –evalue 0.0001 –perc_identity 60 –outfmt 6. Then, BLAST output was parsed according to the selected cutoff identities, and the percentage of unique reads for each metagenome–genome pair was determined at 100, 95, 90, 80 and 70% identity thresholds. Finally, these hits were normalized according to the respective genome and metagenome sizes. To determine the presence of the previously characterized genes involved in arsenic resistance (Morgante and González-Pastor, personal communication) in the CR30 metagenome and the resulting *Bacteroidetes* SAGs assembly, a custom database was built with these genes, and BLAST searches were performed with an e-value threshold of 0.00001 by using the stand-alone blast version BLAST+ v2.2.22.

Multivariate analysis

Dimension reduction of environmental parameters (salinity, ionic composition, PAR, temperature, solar radiation and hours of light) was carried out by principal component analysis using the SPSS software (IBM, Chicago, IL, USA) (Business Editors/High-Tech Writers, 2002). Relationships between environmental and biological variables (cell numbers and Shannon index diversity, calculated using the PAST program (Hammer *et al.*, 2001)) were analyzed by the Redundancy Analysis (RDA) lineal ordination method. The Canonical Correspondence Analysis (CCA) unimodal ordination method was used to correlate environmental variables with prokaryotic phylotypes, while environmental and biological parameters were correlated using RDA. In both cases, Monte Carlo tests with 499 permutations were carried out to ensure the significance of canonical axes. CCA and RDA were carried out using the CANOCO 4.5 software package (Microcomputer Power, Ithaca, NY, USA) with the tool CANODRAW for biplot visualization (ter Braak and Smilauer, 2002).

Shannon index (H) was calculated according to the following formula where p_i is the proportion of phylotype 'i' relative to the total number of phylotypes:

$$H = - \sum_{i=1}^N p_i \ln p_i$$

Analysis of Similarity (ANOSIM) test was used to compare the variation in phylotypes abundance and composition among temporal samples of each pond. In addition, Similarity Percentage tool was used to identify the contribution of phylotypes to similarities between different samples from each pond (Supplementary Table S6). ANOSIM analysis was carried out with the Bray-Curtis distance and 999 permutations by means of the PRIMER 6.0 software (Plymouth Marine Laboratories, Plymouth, UK) computer program (Clarke, 1993).

Results and discussion

Description of samples: spatial and temporal 'environments'

Samples were taken from five hypersaline (Figure 1) ponds at nine time points during the year 2006, yielding a total of 45 water samples. For each of them, salinity, pH, photosynthetically active radiation, ionic composition (four cations and five anions) and temperature were measured; radiation, rainfall and hours of light were also recorded for every sampling time (Supplementary Table S3). These sixteen 'environmental' parameters will be referred to as the independent variables. Total cell counts, *Bacteria* and *Archaea* numbers and diversity (Shannon indices) will be considered as the dependent variables (Supplementary Table S4). In order to reduce the number of independent variables and

make them amenable for subsequent analyses, a principal component analysis was carried out and three components (C1–C3), which explained 82% of environmental data variance, were obtained (Supplementary Table S5). Component 1 was mainly determined by salinity, pH and calcium concentration. Hours of light, solar radiation and rainfall dominated C2, while C3 was defined by sodium, chloride and magnesium concentrations. When the 45 samples were arranged based on their three component values, they were grouped according to their pond of origin (medium concentrators, CM1 and CM2; brine concentrator, CCAB; and crystallizers, CR30 and CR41) and the sampling season. The corresponding spatial and temporal 'environments' are marked with discontinuous lines in Supplementary Figure S1.

Archaeal and bacterial diversity

Bacteria and *Archaea* assemblages were analyzed by DGGE (Supplementary Figure S2), which yielded a specific band pattern for each sample. Bands were recovered from each gel, sequenced and the corresponding partial 16S rRNA gene sequences grouped into phylotypes (that is, a group of partial 16S rRNA gene sequences sharing $\geq 97\%$ similarity). We take this value as a conservative boundary that would embrace species or very closely related species of the same genus according to Stackebrandt and Goebel (1994). A total of 45 phylotypes were detected (25 bacterial and 20 archaeal, Tables 1 and 2, respectively) whose distribution changed throughout the year in the different ponds (Figure 2). For labeling purposes, these groups of sequences were named as the closest cultured genera, although it must be noticed that this does not necessarily mean that the phylotypes belonged to these genera.

As in previous studies (Benlloch *et al.*, 2002), *Actinobacteria*, *Gammaproteobacteria* related to *Alkalilimnicola* and *Nitrococcus* and *Cyanobacteria* were detected in lower and medium salinity ponds (CMs and CCAB) together with *Euryarchaea* related to the genera *Halosarcina*, *Halococcus* and *Halorubrum*. The relative recovery frequency of these groups decreased along the salinity gradient, at the same time that the proportion of phylotypes related to *Bacteroidetes*, in particular with *S. ruber*, and *Archaea*, such as *Hqr. walsbyi*, increased over 25% salinity (Figure 2, and Supplementary Table S6, in which a list of the phylotypes found in each pond at all sampling times is shown). These two groups are frequently dominant in close-to-saturation ponds and will be discussed in more detail later. In addition, phylotypes related to *Nanohaloarchaea* were detected in the different ponds. As mentioned above, this new group had been discovered by using 'modern' molecular ecology techniques, such as NGS metagenomics and single-cell genomics (Ghai *et al.*, 2011; Narasingarao *et al.*, 2012). The detection

Table 1 Bacteria phylotypes retrieved during the course of this study

Phylotypes	Uncultured closest relative ^a Cultured closest relative ^a	Similarity	Ponds ^b	Months	Salinity
Phylum	ACTINOBACTERIA (AB)				
Genus	<i>Candidatus Flaviluna</i>				
AB1	Uncultured marine <i>Bacteria</i> clone MAS85 (AM747379) (1)	96%	CM2	ANNUAL (except JN)	20–30%
	<i>Candidatus Flaviluna lacus partial</i> , strain MWH-Creno3D3 (AM939567.1)	94%			
Genus	<i>Candidatus Rhodoluna</i>				
AB2	Uncultured <i>Actinobacteria</i> , clone lagoaC01 (EF598921.1) (3)	98%	CM1	JL-NV	20–30%
	<i>Candidatus Rhodoluna planktonica</i> , strain MWH-Dar1 (AJ565415.1)	96%			
Phylum	CYANOBACTERIA (CD)				
Genus	<i>Dactylococopsis</i>				
CB1	<i>Dactylococopsis</i> sp. strain PCC 8305 (AJ000711.1)	99%	CCAB	ANNUAL	20–40%
Genus	<i>Euhalothece</i>				
CB2	<i>Euhalothece</i> sp. strain MPI 95AH10 (AJ000709.1)	97%	CCAB	ANNUAL (except JN)	20–40%
Genus	<i>Cyanothece</i>				
CB3	<i>Cyanothece</i> sp. 115 (DQ243690.1)	79%	CR41	SP, OC, and NV	20–40%
Phylum	DELTAPROTEOBACTERIA (DP)				
Genus	<i>Desulfuromonas</i>				
DP1	Uncultured <i>Proteobacteria</i> , clone At12OctA12 (AY862790.1)	94%	CM2	J2, JL and SP	20–30%
	<i>Desulfuromonas alkaliphilus</i> strain Z-0531 (DQ309326.1)	86%			
Genus	<i>Desulfosarcina</i>				
DP2	Uncultured <i>Bacteria</i> clone P9_LJL_0205_030 (EU677285.1)	99%	CCAB	JN	20–30%
	<i>Desulfosarcina</i> sp. clone SB4_53 (AY177791.1)	96%			
Phylum	GAMMAPROTEOBACTERIA (GP)				
Genus	<i>Alkalilimnicola</i>				
GP1	Uncultured <i>Bacteria</i> clone 185ZF11 (CU467460.1)	99%	CM1	ANNUAL	19–30%
	<i>Alkalilimnicola ehrlichei</i> MLHE-1 (CP000453.1)	94%			
Genus	<i>Alkalispirillum</i>				
GP2	Uncultured <i>Bacteria</i> clone 185ZE06 (CU467472.1)	99%	CM1	ANNUAL	18–40%
	<i>Alkalispirillum</i> sp. Z 7008 (DQ217920.1)	99%	CM2 CCAB	JN JN	
Genus	<i>Nitrococcus</i>				
GP3	Uncultured <i>Bacteria</i> clone 185ZE06 (CU467472.1)	4%	CM2	JL-NV	20–30%
	<i>Nitrococcus mobilis</i> (ATCC 25380)	88%			
Genus	<i>Halospina</i>				
GP4	<i>Halospina denitrificans</i> strain 30–40%D 1-3 (DQ072719.1)	98%	CM2	ANNUAL	18–30%
Genus	<i>Halovibrio</i>				
GP5	Uncultured <i>Bacteria</i> clone SN144 (EU735693.1)	83	CR41	JN	20–30%
	<i>Halovibrio denitrificans</i> strain C4-2 (EU868859.1)				
GP6	<i>Pseudomonas</i> sp., clone 2B4 (AY987840.1)	95%	CR41	JN, MR	20–30%
	<i>Halovibrio denitrificans</i> , strain C4-2 (EU868859.1)	94%			
Phylum	BACTEROIDETES (BC)				
Genus	<i>Psychroflexus</i>				
BC1	Uncultured <i>Bacteria</i> , clone E4aA01 (DQ103640.1)	98%	CM2 CCAB CR41	ANNUAL	20–40%
	<i>Psychroflexus</i> sp. sp.gap-d-25 (DQ530461.1)	98%			
Genus	<i>Owenweeksia</i>				
BC2	<i>Owenweeksia hongkongensis</i> (DQ103638.1)	98%	CM2	JN	18%
Genus	<i>Cytophaga</i>				
BC3	Uncultured <i>Bacteria</i> , clone LCKS000B34 (EF201719.1)	98%	All except CM1	ANNUAL	18–> 40%
	<i>Cytophaga</i> sp. Dex80-37 (AB239011.1)	87%			
Genus	<i>Salinibacter</i>				
SR1	<i>Salinibacter ruber</i> DSM 13855 (CP000159.1)	99%	CMs CCAB and CRs	ANNUAL (except JN) ANNUAL	20–> 40%

Table 1 (Continued)

Phylotypes	Uncultured closest relative ^a Cultured closest relative ^a	Similarity	Ponds ^b	Months	Salinity
SR2	Uncultured <i>Bacteria</i> , clone SFC1D061 (AM981340.1)	98–99%	CCAB	ANNUAL (except JN)	30–>40%
	<i>Salinibacter ruber</i> DSM 13855 (CP000159.1)	93%	CRs	ANNUAL	
SR3	Uncultured <i>Bacteria</i> , clone SFA1H111 (CU467389.1)	99%	CM2	AB—JL	30–>40%
	<i>Salinibacter ruber</i> DSM 13855 (CP000159.1)	92%	CCAB	ANNUAL	
SR4	Uncultured <i>Bacteria</i> , clone SFC1D061 (AM981343.1)	78%	CM2	AB—JL	20–30%
	<i>Salinibacter ruber</i> DSM 13855 (CP000159.1)	71%			
SR5	Uncultured <i>Bacteria</i> , clone SFA1H111 (CU467389.1)	93%	CM2	AB—NV	20–30%
	<i>Salinibacter ruber</i> DSM 13855 (CP000159.1)	88%			
SR6	<i>Salinibacter sp.</i> clone SFD1E121 (FN393500.1)	93%	CR41	ANNUAL	20–>40%
	<i>Salinibacter ruber</i> DSM 13855 (CP000159.1)	89%			
SR7	Uncultured <i>Bacteria</i> , clone SFA1C121 (CU467418.1)	99%	CR41	SP, OC and NV	20–>40%
	<i>Salinibacter ruber</i> DSM 13855 (CP000159.1)	95%			
Genus		<i>Salisaeta</i>			
SL1	<i>Salinibacter sp.</i> clone salH12 (EF598986.1)	98%	CR30	SP and OC	20–>40%
	<i>Salisaeta longa</i> strain S4-4 (EU426570.2)	93%	CR41	SP, OC and NV	
SL2	Uncultured <i>Bacteria</i> , clone XB27 (FM211011.1)	96%	CM1	SP	20–30%
	<i>Salisaeta longa</i> strain S4-4 (EU426570.2)	93,5%			

^aPubmed access number.^bPonds: Medium concentrators CMs: CM1 and CM2; Brine concentrator: CCAB; and Crystallizers: CRs: CR30 and CR41.

of nanohaloarchaeal phylotypes in our samples reinforces the use of the 16S rRNA gene DGGE approach despite the fact that some authors considered it to be obsolete and very limited for environmental studies (Legault *et al.*, 2006), as it can only retrieve phylotypes that account for >0.4% of the community (Casamayor *et al.*, 2000). Indeed, around 92% of the phylotypes detected in June 2008 in CR30 using a NGS metagenomic approach (Ghai *et al.*, 2011) have been detected at different time points in the different samples studied here. In addition, many phylotypes described in this work were not present in the NGS study.

A total of seven phylotypes related to the genus *Salinibacter* (see Table 1 and Figure 2) were recovered in this study. In 1998, the presence of extremely halophilic bacteria (later described as '*Salinibacter*') was discovered in the crystallizer CR30 (Antón *et al.*, 2000). However, at that time, only two phylotypes were detected, and only one of them (named EHB-1, from 'extremely halophilic bacterium' 1) has been brought to pure culture and classified as *S. ruber* (Antón *et al.*, 2002). Currently, all *Salinibacter sp.* isolates recovered from different hypersaline environments around the world (Antón *et al.*, 2008, 2013) correspond to EHB-1. Here, in samples taken 8 years after the first description of *Salinibacter* in these salterns, a total of seven different phylotypes related to this genus have been detected. Only one of them (SR1) had a similarity >97% with *S. ruber* isolates, providing a new example of the biases imposed by cultivation even in 'culturable' species. Other *Salinibacter*-related phylotypes (see Table 1) presented the highest identities with environmental clones from medium

salinity ponds in Tunisian salterns (Baati *et al.*, 2008, 2010, 2011). Remarkably, some *Salinibacter* phylotypes, like SR4 and SR7, were detected only at specific temporal points from CM2 and CR41, respectively, underscoring the dynamic nature of the microbial communities in the salterns.

Although highly divergent *Salinibacter*-related phylotypes have been frequently retrieved from hypersaline environments (Mutlu *et al.*, 2008), most *Haloquadratum* sequences previously detected were very closely related to each other. For instance, *Hqr. walsbyi* cultured strains have <1.8% divergence in their 16S rRNA gene sequences (Legault *et al.*, 2006) while environmental phylotypes recovered from highly distant saltern ponds in Spain (Bolhuis *et al.*, 2006; Legault *et al.*, 2006), Israel (Soerensen *et al.*, 2005), Turkey (Mutlu *et al.*, 2008), Tunisia (Baati *et al.*, 2008), Peru (Maturrano *et al.*, 2006) and Australia (Oh *et al.*, 2010) normally present <1% divergence. In contrast, some *Haloquadratum* sequences recovered in this work (Table 2 and Figure 2) showed a higher divergence (up to 5%) in the analyzed 16S rRNA gene fragment. In particular, phylotype HQR3 was detected only at some time points in low and medium salinity ponds while HQR7 was detected only in the crystallizer CR41 that had not been sampled previously. These examples, again, underscore the relevance of taking samples at different time points and from different crystallizers.

A main outcome of this study is that, despite the fact that two analyzed crystallizers displayed very similar salinities (Figure 2, Tables 1 and 2), remarkable differences in their microbial community structures were found (28.73% dissimilarity as calculated with ANOSIM, Supplementary

Table 2 Archaea phylotypes retrieved during the course of this study

Phylotypes	Uncultured closest relative ^a Cultured closest relative ^a	Similarity	Ponds ^b	Months	Salinity
Genus		<i>Haloquadratum</i>			
HQR1	<i>Haloquadratum walsbyi</i> DSM 16790 (AM180088.1)	98%	CMs CCAB CR30 CR41	ANNUAL (except JN) ANNUAL ANNUAL ANNUAL (except JN)	20–> 40%
HQR2	Uncultured <i>Haloarchaea</i> clone Cry7_clone27 (GQ374968.1) <i>Haloquadratum walsbyi</i> C23 (NR_028207.1)	99% 92%	CMs CCAB CR30 CR41	ANNUAL (except JN) ANNUAL (except JN) AB-NV ANNUAL J2-NV	20–> 40%
HQR3	Uncultured <i>Haloarchaea</i> , clone Cry7_clone20 (GQ374960.1) <i>Haloquadratum walsbyi</i> DSM 16790 (AM180088.1)	95% 95%	CM2 CCAB	MR-JL and NV MR-J2	20–40%
HQR4	<i>Haloquadratum walsbyi</i> DSM 16790 (AM180088.1) <i>Haloquadratum walsbyi</i> DSM 16790 (AM180088.1)	98–99%	CCAB CR41	ANNUAL ANNUAL (except JN)	20–> 40%
HQR5		99%	CCAB CR30	ANNUAL MR, SP and NV	20–> 40%
HQR6	<i>Haloquadratum walsbyi</i> DSM 16790 (AM180088.1)	99%	CCAB CR30	EN-J1 ANNUAL(except SP and OC)	20–> 40%
HQR7	Uncultured <i>Archaea</i> , clone 186ZH05 (CU467175.1) <i>Haloquadratum walsbyi</i> DSM 16790 (AM180088.1)	96% 96%	CR41 CR41	MR-J2 JN-J2	20–> 40%
Genus		<i>Halorubrum</i>			
HRR1	Uncultured <i>Archaea</i> clone Baj_clon46 (GQ374995.1) <i>Halorubrum orientale</i> strain EJ-52T (AM235786.1)	99–100%	CMs CCAB CR30-CR41	ANNUAL (except JN from CM2)	20–> 40%
HRR2	Uncultured <i>Archaea</i> clone Cry7_clone60 (GQ375002.1) <i>Halorubrum</i> sp. DV427 (FJ492047.1)	98% 92%	CCAB	ANNUAL	20–> 40%
HRR3	Uncultured <i>Archaea</i> clone SFH1B111 (FN391236.2) <i>Halorubrum</i> sp strain DW6 (EF194098.1)	94% 92%	CM1	JN	19%
HRR4	<i>Halorubrum</i> sp. clone SFH1A082 (FN391273.2) <i>Halorubrum</i> sp. strain EN-2 (AM268115.1)	85% 92%	CM1	JL	20–30%
HRR5	Uncultured <i>Archaea</i> clone Baj_clone46 (GQ374995.1) <i>Halorubrum orientale</i> , strain EJ-52T (AM235786.1)	94% 94%	CM2	ANNUAL	20–30%
HRR6	Uncultured <i>Archaea</i> clone Cry7_clone81 <i>Halorubrum alkaliphilum</i> strain DZ-1 (NR_028206.)	99% 98%	CR30	ANNUAL	20–> 40%
HRR7	Uncultured <i>Haloarchaea</i> clone SFG1H031 (AM947489.1) <i>Halorubrum</i> sp. EN-2 (AM268115.1)	85% 85%	CM1	JL, SP and OC	20–30%
Genus		<i>Halorhabdus</i>			
HRD	Uncultured <i>Haloarchaea</i> clone 186ZF04 (CU467268.1) <i>Halorhabdus utahensis</i> AX-2 (NR_028172.1)	99% 91%	CR30	ANNUAL	20–> 40%
Genus		<i>Halosarcina</i>			
HSN	Uncultured <i>Haloarchaea</i> clone 3a-PML 22% pond (AF477932.1) <i>Halosarcina pallida</i> (EF055454.1)	99% 94%	CMs CCAB	ANNUAL J2-NV	18–40%
Genus		<i>Halococcus</i>			
HCC	<i>Halococcus hamelinensis</i> (NR_028168.1)	92%	CMs CCAB CR30 CR41	ANNUAL ANNUAL J2- JL SP-NV	18–> 40%
Genus		<i>Natronomonas</i>			
NMN	<i>Natronomonas</i> sp. DV462A (FJ492048.1)	93%	CM1	De J1 a JL	20–30%
NMN2	<i>Natronomonas moolapensis</i> 8.8.11 strain CSW8.8.11 (AY498645.2)	93%	CM2	ANNUAL (-JN)	20–30%
NMN3	<i>Natronomonas</i> sp. DV462A (FJ492048.1)	94%	CCAB	JN-JL	30–40%
Lineage		<i>Nanohaloarchaea</i>			
NSA1	<i>Candidatus</i> Nanosalinarum (Clade J0756AB)	97–99%	CM1 CM2 CCAB CRs	MR, AB and JL JL and NV JL and SP ANNUAL	20–> 40%
NSA2			CM1	MR, AB and JL	
NSR1	<i>Candidatus</i> Nanosalina (Clade J07AB43)	97–99%	CM1 CM2 CCAB CRs	AP, J1 and J2 J1, J2, JL and NV J2-NV AP-NV	20–> 40%
NSR2			CM1 CM2	AP, J1 and J2 J1, J2, JL and NV	
NSR3			CCAB	J2-NV	20–> 40%
NSR4			CRs	AP-NV	
NSR5			CRs	AP-NV	

^aPubmed access number.

^bPonds: Medium concentrators CMs: CM1 and CM2; Brine concentrator: CCAB; and Crystallizers: CRs: CR30 and CR41.

Table S7). This fact is intriguing if we consider that both ponds are fed with the same water from CCAB. Although a common portion of CCAB community remained in both crystallizers, there were also

crystallizer-specific phylotypes, such as *Halorhabdus*-related phylotypes that were detected at all sampling times in CR30 or HQR7 *Haloquadratum* and *Salinibacter* phylotype SR6 present only in

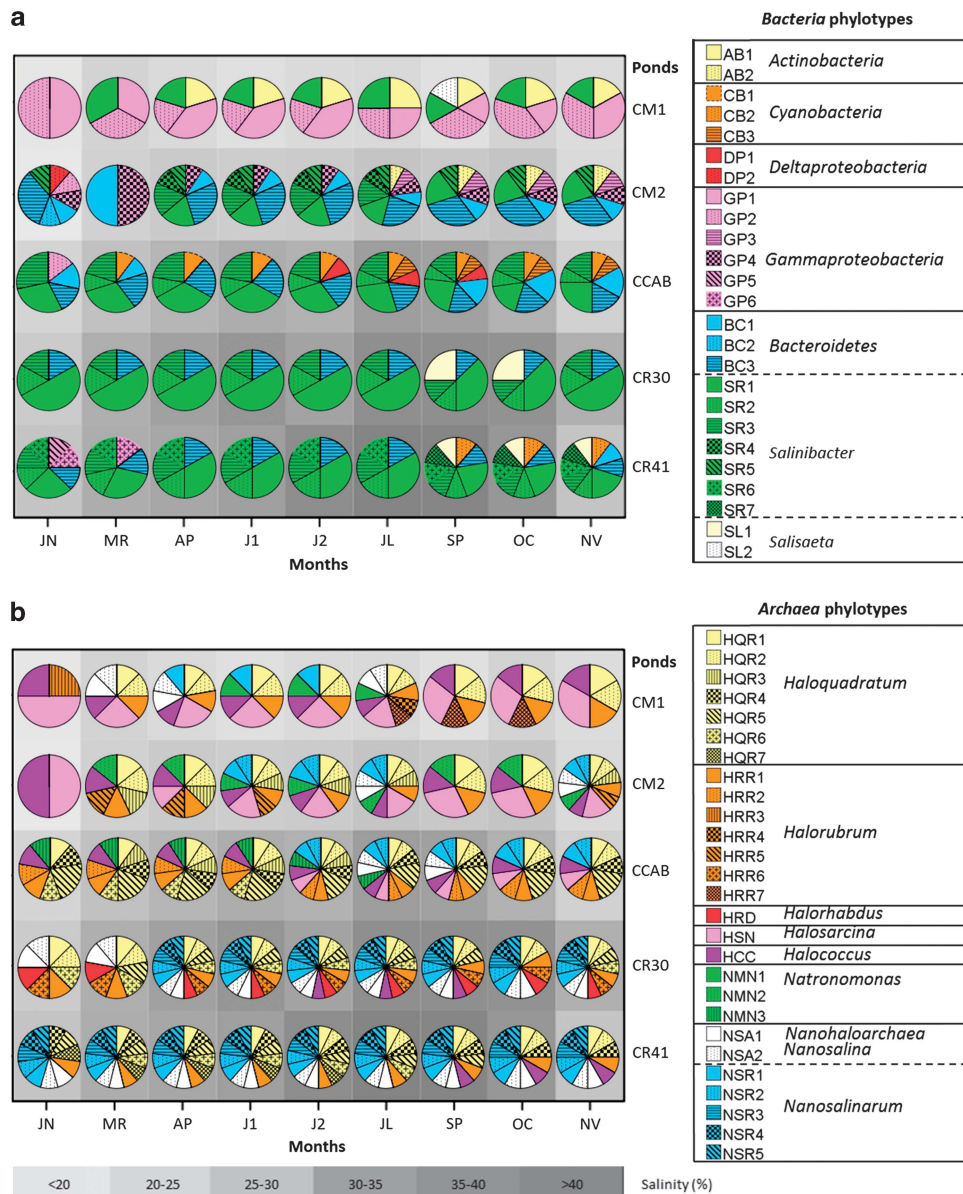


Figure 2 Temporal distribution of relative abundance of bacterial (a) and archaeal (b) phylotypes in the five analyzed ponds (medium concentrators CM1 and CM2, brine concentrator CCAB, and crystallizers CR30 and CR41). A complete description of the phylotypes is provided in Tables 1 and 2.

CR41 samples (HQR7 from January to June and SR6 in all samples). In addition, CR41 harbored *Cyanobacteria* from September to November, which is noteworthy as normally this group of prokaryotes is absent from close-to-saturation hypersaline waters being the green alga *Dunaliella* the only photoautotroph (Estrada *et al.*, 2004; Gasol *et al.*, 2004).

Prokaryotic community dynamics and their relation to environmental parameters

Total microbial counts were in the range previously observed in solar salterns (Guixa-Boixereu *et al.*, 1996; Antón *et al.* 1999, 2000; Benlloch *et al.* 2002; Oren, 2002; Estrada *et al.*, 2004; Burns *et al.*, 2007;

Mutlu *et al.* 2008; Eulloumi *et al.*, 2009) ranging from 6.3×10^6 to 4.79×10^7 cells ml^{-1} in CR30 (OC) and CCAB (JL), respectively (Figure 3). Although throughout the concentration process from seawater to saturated brines there is an overall increase of cell density, at concentrations $>20\%$ this trend is not always observed (Guixa-Boixereu *et al.*, 1996; Boujelben *et al.*, 2012), as shown here. Total counts changed in the different ponds throughout the year, with lower salinity ponds being more stable than crystallizers. As expected, the highest counts in each pond were reached in summer samples while samples taken in colder months harbored less dense communities. This is in agreement with the low microbial numbers (around 5×10^6 cells ml^{-1})

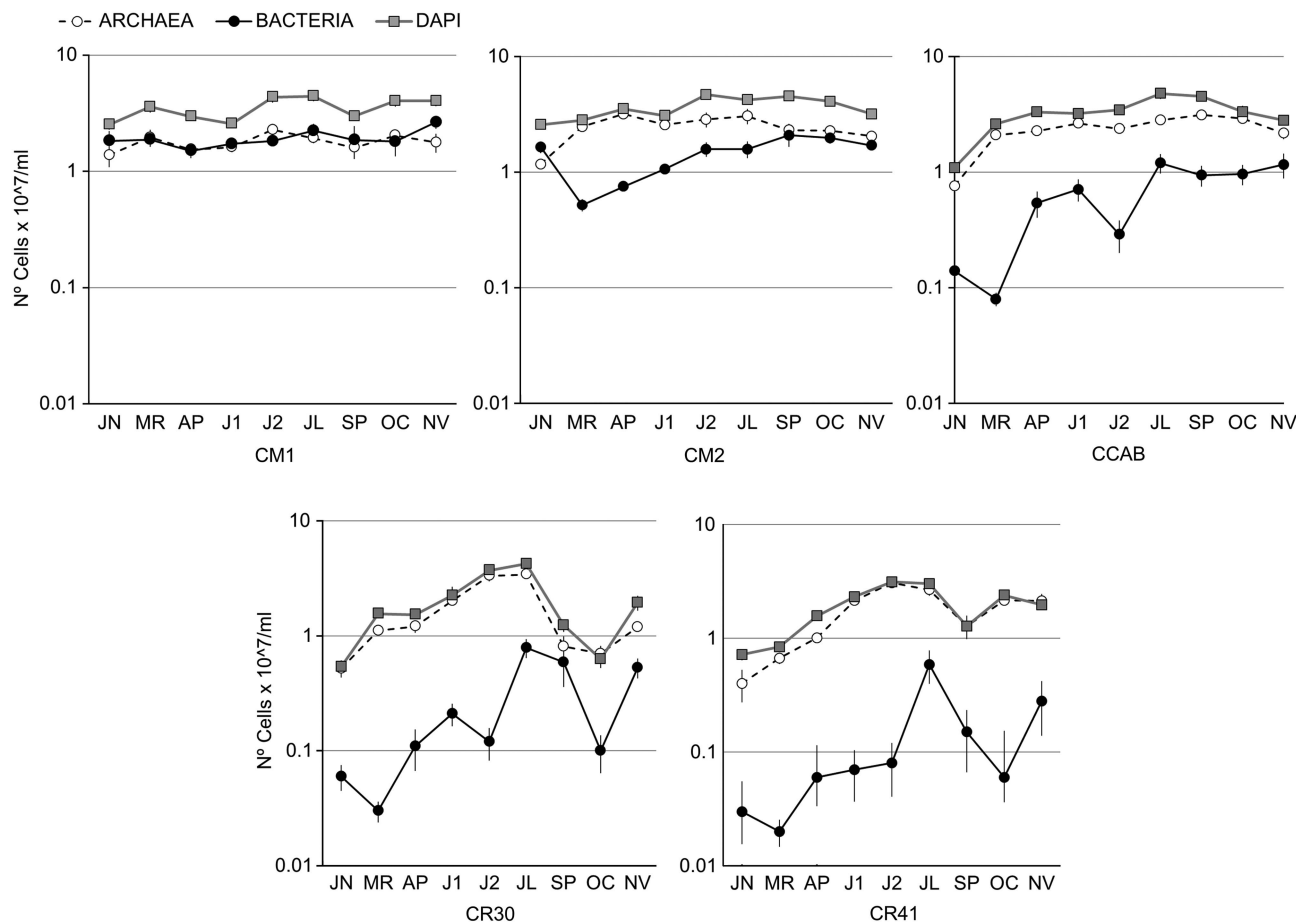


Figure 3 Temporal variation of total counts (grey symbols), *Archaea* (white symbols) and *Bacteria* (black symbols) in the different ponds.

detected in Andean Maras salterns (Maturrano *et al.*, 2006), with an annual average temperature < 20 °C, which present the lowest cell concentration reported from crystallizer ponds.

Bacterial and archaeal numbers (determined by CARD-FISH) also changed with salinity and in every pond throughout the year. As expected for hypersaline systems (Antón *et al.*, 2000, Maturrano *et al.*, 2006; Mutlu *et al.*, 2008) *Bacteria* were less abundant than *Archaea* in almost all the samples, except in a few CM samples (Figure 3).

Along with the changes in cell numbers, there was also a temporal succession of different phylotypes in each pond (Figure 2). Changes in phylotype composition mainly occurred during winter–spring and summer–autumn transitions and were more drastic in CMs and CCAB than in crystallizers (in fact, CR30 was the steadiest pond). Microbial diversity in the environment can be explained in the frame of the ‘core-seed bank’ model proposed by Pedrós-Alió (2006), according to which ‘core’ (most abundant) microorganisms would be maintained through active growth, while minority species (‘seed’ bank) would be below the detection threshold of molecular techniques, such as DGGE. However, apart from

differences among environments, temporal variations should also be taken into account as, depending on the time of sampling, a given phylotype could be part of either the core or the seed bank. For instance, that would be the case of *Salinibacter* SR4 phylotype that was only detected from AP to JL samples (CM2), while was below the DGGE detection limit (that is, part of the seed bank) the rest of the year.

Relationships among environmental parameters and prokaryotic numbers and diversity were investigated by redundant detrended analysis (RDA; Figure 4; see Supplementary Tables S8–11). In this analysis, RDA1 and RDA2 ordination axes explained 87.5% total variance data. Salinity was positively correlated with most ions and negatively with calcium, although ammonium concentration was independent of salinity. Numbers of *Bacteria* and total cell counts showed a negative dependence with salinity, while *Archaea* was rather independent of this variable (Supplementary Tables S12 and S13). This may seem contradictory with previous results (Guixa-Boixereu *et al.*, 1996; Antón *et al.*, 2000; Pedrós-Alió, 2005) that indicated that archaeal numbers increased with salinity. This discrepancy

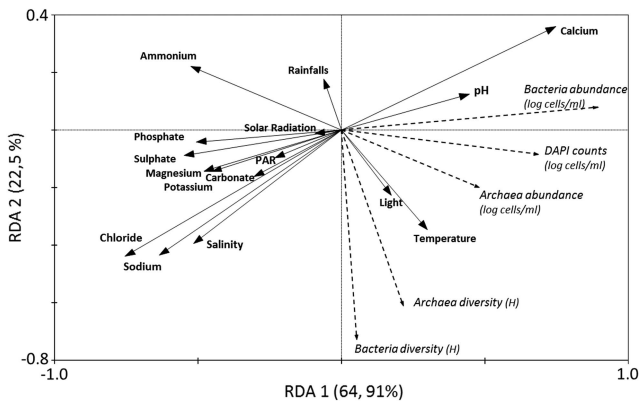


Figure 4 Redundance detrended analysis (RDA) canonical ordination biplot of environmental and biological parameters. Environmental parameters are represented by black arrows and biological parameters by dotted arrows. DAPI, *Archaea* and *Bacteria* abundances are represented by the logarithm of the number of cells per milliliter and the diversity of *Archaea* and *Bacteria* by Shannon diversity indices (*H*). RDA yielded two synthetic canonical axes (RDA1 and RDA2) that explained 89.2% of data variance (Supplementary Table S7).

is explained, because the correlation between archaeal numbers and salinity is only circumscribed to some months. However, this correlation disappears when the whole data set is considered. In addition, if ponds of lower salinity (that is, from marine to around 15%) had been included in the analysis, correlation between salinity and archaeal number would probably be more apparent. Both *Archaea* and *Bacteria* diversity were highly influenced by salinity and by meteorological factors (Supplementary Table S8). In particular, archaeal diversity was strongly correlated with hours of light and temperature. Temperature had also been found to have a strong influence on the prokaryotic community structure in ~20% salinity ponds from Sfax solar salterns (Boujelben *et al.*, 2012).

The different phylotypes found in the analyzed ponds were arranged by canonical correspondence analysis, CCA, according to their optimum environmental factors (Figure 5). CCA axes 1 and 2 explained 54.3% total variance data (Supplementary Table S11). As in the RDA analysis, the first canonical axis (CCA1) was highly correlated with salinity and ion concentrations (except ammonium) and the second one (CCA2) with meteorological factors. One of the main outcomes of this analysis was that phylotypes belonging to the same genus could have very different environmental optima (for instance, this is the case of SR1 and SR6, both related to *Salinibacter* sp. but located in very different regions of CCA space; see Figure 5). In addition, ammonium concentration, which was less correlated with the rest of ions, influenced the distribution of some phylotypes, such as SR7 or HQR7 (related to *S. ruber* and *Hqr. walsbyi*, respectively). *Nanohaloarchaea*-related phylotypes showed a relatively narrow range of environmentally optimum conditions (circled in Figure 5) with

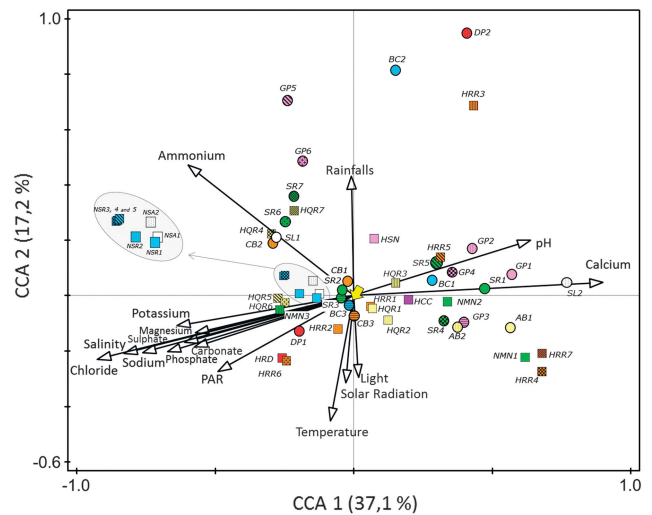


Figure 5 Canonical correspondence analysis (CCA) ordination biplot of phylotypes (symbols) and environmental parameters (arrows). Zone A (right) is defined by low salinity, high calcium concentration and high pH values; zone B (center) represents medium values (and non-determining environmental factor) and zone C (left) high salinity, low calcium concentration, and low pH. Archaeal (squares) and bacterial (circles) phylotypes are named as in Tables 1 and 2. Symbol colors are as in Figure 2. The yellow arrow points to BC3 phylotypes, which were further analyzed by SGC. *Nanosalina* (white) and *Nanosalinarum* phylotypes (blue squares) are enclosed within a grey line.

Candidatus Nanosalinarum phylotypes (except NRS1) being more ‘halophilic’ than *Candidatus Nanosalina* phylotypes.

Phylotypes located in the center of the CCA biplot were found under a wide range of environmental conditions and could thus correspond to generalists. Among them, especially worth mentioning is the phylotype BC3 that was detected from 18% to 41% salinities in different samples taken throughout the whole year (see Figure 2). This phylotype (Figure 6) belongs to a clade merely formed by environmental clones of uncultured *Bacteroidetes*. The closest cultured bacterium, *Cytophaga* sp. Dex80-37, was only very distantly related to BC3 (87% sequence identity) and affiliates differently within the *Flavobacteriaceae/Bacteriales/Sphingobacteriaceae* branch (Figure 6). Database searches indicated that sequences very closely related (ore even identical) to BC3 were retrieved in the NGS study of the Santa Pola salterns mentioned above (Ghai *et al.*, 2011), as well as in other studies from hypersaline environments worldwide, such as in Lake Chaka (China), Chula Vista salterns (California, USA), Lake Tyrrel (Australia) and Lake Aran-o-Bidgol (Iran) (Jiang *et al.*, 2007; Wang *et al.*, 2011; Podell *et al.*, 2013). However, in none of these studies there is an explicit reference to this group nor is its ubiquity acknowledged. This lag between the first report of environmental sequences and the recognition of their ecological relevance is not uncommon, and it happened, for instance, with nanohaloarchaeal sequences, which were first reported in 1999

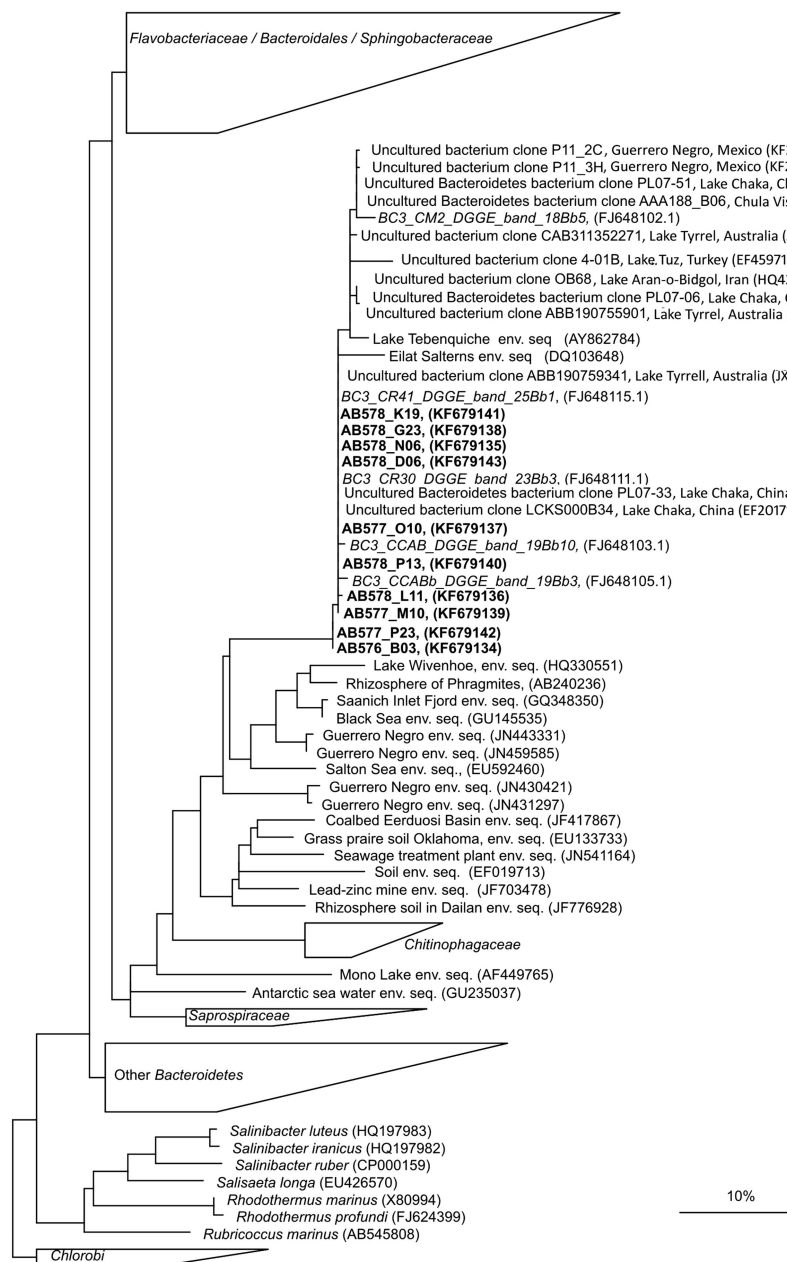


Figure 6 Phylogenetic reconstruction based on a neighbor joining analysis of a subset of almost complete sequences belonging to the phylum *Bacteroidetes*. DGGE band sequences corresponding to phylotype BC3 are in italics and 16S rRNA corresponding to the SAGS obtained from a CR30 sample are in bold. Members of *Chlorobi* were included as outgroup. Partial sequences of DGGE and clones were added to the tree using the parsimony tool implemented in ARB (Ludwig *et al.*, 2004). The bar indicates 1% sequence divergence.

(Grant *et al.*, 1999), but their abundance and world-wide distribution were not clearly stated until more than 10 years later (Narasingarao *et al.*, 2012).

Characterization of the phylotype BC3 by means of single-cell genomics

Despite its apparent abundance and wide distribution, the phylotype BC3 has not been cultured yet. For this reason, we decided to carry out a culture-independent approach to gain some insight into its metabolism by using high-throughput single-cell

sequencing. For that purpose, we performed fluorescence-activated cell sorting in combination with whole-genome amplification to recover the genetic information for these prokaryotes in the crystallizer CR30. Results of the 16S rRNA gene PCR screening and sequencing for the generated SAGs showed that a total of 10 SAGs belonging to the *Bacteroidetes* BC3 phylotype were recovered (Figure 6), 5 of which were selected for further genome sequencing. The resulting genome assembly from the low-coverage Illumina sequencing for each SAGs was ≈ 0.4 Mpb (Supplementary Tables S1 and S2), recovering

approximately 10% of the genome (considering that the mean of *Bacteroidetes* genomes is 4 ± 1.7 Mpb; data obtained from IMG prokaryote database, dated 15 December 2012). The GC content was significantly lower (47%) than other sequenced extremely halophilic *Bacteroidetes*, such as *S. ruber* (66% GC). Comparison of average nucleotide identities (Supplementary Figure S3) among the sequenced SAGs indicated that all of them might belong to the same species as the average nucleotide identity threshold for species circumscription had been observed to be between 94% and 96% (Richter and Rosselló-Móra, 2009). This result must be taken cautiously as fragments accounting for at least 20% of the genomes of the strains under comparison are needed for stable average nucleotide identity calculations (Richter and Rosselló-Móra, 2009). However, except for L11, we are confident with regard to the results, given that the amount of aligned genome stretches for average nucleotide identity calculations were relatively high and ranged between 12% and 26% of the total sequenced DNA. The hypothesis that all cells belonged to the same species is reinforced by their almost identical 16S rRNA gene sequences.

Metagenomic fragment recruitment data demonstrated that these sequenced SAGs were predominant in both the intermediate salinity pond of 19% (SS19) and the crystallizer CR30 (Ghai *et al.*, 2011) as the recruitment ratio was as high as the very abundant *S. ruber* (Figure 7). Indeed, in the SS19 pond, the recruitment ratio of BC3 was higher than *S. ruber*, which points to a more generalist lifestyle of this new group in terms of salinity range.

Genome annotation suggested that this hypersaline planktonic cluster contains unexpected metabolic capabilities. Arsenate respiration metabolic features were detected as a specific transporter to

pump out the highly toxic form arsenite, and the key gene regulator *arsR* involved in arsenate respiration (Murphy and Saltikov, 2009) were found in this group. In addition, the ATP-dependent chaperone ClpB, which has been demonstrated to confer resistance to arsenic (Morgante and González-Pastor, personal communication), was not only detected in the phylotype BC3 but was also fairly abundant in the metagenome of CR30, indicating that this chaperone could be a common mechanism for resistance to arsenic in saline systems together with the use of the enzymes 3-isopropylmalate dehydrogenase, 5'phosphoribosyl anthranilate isomerase and aspartate semialdehyde dehydrogenase (Supplementary Figure S4).

Like other *Bacteroidetes*, this group also possesses a complete repertoire of gliding-motility proteins that have been related to degradation of complex polysaccharides (Gomez-Pereira *et al.*, 2012). In addition, transporters for large biopolymers, glycosyl hydrolases and transferases were detected. For instance, SAGs AB578-K19 and AB578-D06 possessed a nucleotide-binding protein with the corresponding catalytic domain previously demonstrated to be involved in DNA uptake, which could be used either as a source of organic matter or for DNA transformation. Thus some traits of this group resembled those of other planktonic *Bacteroidetes* previously described to be involved in the remineralization of high-molecular-weight organic matter (Gomez-Pereira *et al.*, 2012). The putative DNA-uptake systems found in the BC3 group could be related to its use as P source, as hypersaline systems are considered to be P-limited (in this study, N:P rations based on ammonia and phosphate concentrations ranged from 277 to 2770). In addition, a beta-lactamase and an ABC-antimicrobial peptide transporter were present at least in two of the single cells. These genes could be involved in interactions of BC3 with their accompanying microbiota. Indeed, a recent metabolomic characterization of *S. ruber* strains, with which BC3 shares its habitat, has shown that this bacterium produces a plethora of different antimicrobial-like compounds (Antón *et al.*, 2013).

Furthermore, a nitroreductase enzyme involved in the utilization of complex aromatic and heterocyclic compounds with nitro-functional groups was detected. This enzyme has been reported to have potential applications in bioremediation, biocatalysis and biomedicine (Emptage *et al.*, 2009). In addition, the *cas* genes 2 and 6 and the corresponding CRISPR arrays (Makarova *et al.*, 2011), together with different components of the putative type 1 site-specific restriction modification systems (endonucleases and methyltransferases) involved in degrading foreign DNA, were found in some of the single cells, suggesting that this group might be under potential highly selective viral predatory pressure. This is not surprising given that hypersaline environments harbor the highest viral

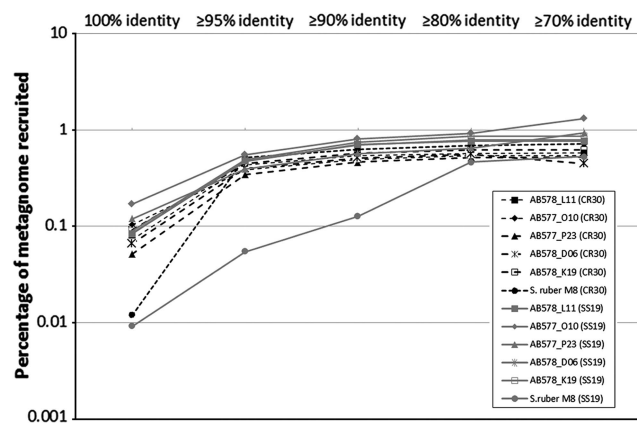


Figure 7 Metagenomic fragment recruitment of *Bacteroidetes* SAGs and the reference bacteroidetes species *S. ruber*. The fraction of recruited metagenome reads from the crystallizer pond CR30 and the medium-salinity pond SS19 was analyzed using a range of different DNA identity thresholds and normalized per Mbp genome size.

concentrations reported so far for aquatic systems (Santos *et al.*, 2012). Finally, a gene coding for the light-driven proton pump bacteriorhodopsin was present at least in one of the sequenced SAGs.

Conclusions

Here, we have shown temporal changes in prokaryotic assemblages in a given pond as well as differences between ponds of very similar salinities, underscoring the need for caution when drawing general conclusions based on local data. In addition, we have found some traits that are considered to be characteristic of hypersaline systems, such as the abundance of *Haloquadratum*-related phylotypes and the predominance of *Bacteroidetes* representatives. The comparison of samples taken at different time points and different ponds has also enabled the detection of a new *Bacteroidetes* phylotype across a wide range of salinity and environmental conditions, which led us to propose its generalist nature. Phylotype BC3, now known to be widely distributed in salterns worldwide, had not been detected in Santa Pola saltern 10 years ago with techniques of the same resolution as the ones used here (Benlloch *et al.*, 2002). Overall, our results show that solar saltern microbiota undergoes considerable changes throughout the year, and therefore the idea that hypersaline ponds are 'steady-state' systems should be reconsidered.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

This work was supported by the projects CGL2012-39627-C03-01 and 02 of the Spanish Ministry of Economy and Competitiveness, which were also co-financed with FEDER support from the European Union. TG group research is funded in part by a grant from the Spanish Ministry of Economy and Competitiveness (BIO2012-37161), a grant from the Qatar National Research Fund grant (NPRP 5-298-3-086) and a grant from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC (grant agreement no. ERC-2012-StG-310325). We thank the staff of the Bras del Port salterns for their help throughout the sampling campaign. We thank Heinz Himmelbauer and the CRG's ultra-sequencing unit for their support and Esther Rubio-Portillo for her help with statistics software.

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W *et al.* (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389–3402.

- Amann RI, Ludwig W, Schleifer KH. (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol* **59**: 143–169.
- Antón J, Llobet-Brossa E, Rodríguez-Valera F, Amann R. (1999). Fluorescence in situ hybridization analysis of the prokaryotic community inhabiting crystallizer ponds. *Environ Microbiol* **1**: 517–523.
- Antón J, Lucio M, Peña A, Cifuentes A, Brito-Echeverría J, Moritz F *et al.* (2013). High metabolomic microdiversity within co-occurring isolates of the extremely halophilic bacterium *Salinibacter ruber*. *PLoS One* **8**: 5.
- Antón J, Oren A, Benlloch S, Rodríguez-Valera F, Amann R, Rosselló-Mora R. (2002). *Salinibacter ruber* gen. nov., sp. nov., a novel, extremely halophilic member of the *Bacteria* from saltern crystallizer ponds. *Int J Syst Evol Microbiol* **52**: 485–491.
- Antón J, Peña A, Santos F, Martínez-García M, Schmitt-Kopplin P, Rosselló-Mora R. (2008). Distribution, abundance and diversity of the extremely halophilic bacterium *Salinibacter ruber*. *Saline Systems* **28**: 4–15.
- Antón J, Rosselló-Mora R, Rodríguez-Valera F, Amann R. (2000). Extremely halophilic bacteria in crystallizer ponds from solar salterns. *Appl Environ Microbiol* **66**: 3052–3057.
- Baati H, Guermazi S, Amdouni R, Gharsallah N, Sghir A, Ammar E. (2008). Prokaryotic diversity of a Tunisian multipond solar saltern. *Extremophiles* **12**: 505–518.
- Baati H, Guermazi S, Gharsallah N, Sghir A, Ammar E. (2010). Microbial community of salt crystals processed from Mediterranean seawater based on 16S rRNA analysis. *Can J Microbiol* **56**: 44.
- Baati H, Jarboui R, Gharsallah N, Sghir A, Ammar E. (2011). Molecular community analysis of magnesium-rich bittern brine recovered from a Tunisian solar saltern. *Can J Microbiol* **57**: 975.
- Benlloch S, López-López A, Casamayor E, Øvreås L, Goddard V, Daae FL *et al.* (2002). Prokaryotic genetic diversity throughout the salinity gradient of a coastal solar saltern. *Environ Microbiol* **4**: 349–360.
- Bolhuis H, Palm P, Wende A, Falb M, Rampp M, Rodríguez-Valera F *et al.* (2006). The genome of the square archaeon *Haloquadratum walsbyi*: life at the limits of water activity. *BMC Genomics* **7**: 169.
- Boujelben I, Gomariz M, Martínez-García M, Santos F, Peña A, López C *et al.* (2012). Spatial and seasonal prokaryotic community dynamics in ponds of increasing salinity of Sfax solar saltern in Tunisia. *Antonie Van Leeuwenhoek* **101**: 845–857.
- Burns DG, Camakaris HM, Janssen PH, Dyal-Smith ML. (2004). Cultivation of Walsby's square haloarchaeon. *FEMS Microbiol Lett* **238**: 469–473.
- Burns DG, Janssen PH, Itoh T, Kamekura M, Li Z, Jensen G *et al.* (2007). *Haloquadratum walsbyi* gen. nov., sp. nov., the square haloarchaeon of Walsby, isolated from saltern crystallizers in Australia and Spain. *Int J Syst Evol Microbiol* **57**: 387–392.
- Business Editors/High-Tech Writers (2002). SPSS enabling technologies partner program to provide embeddable analytic and predictive technologies to software and solutions providers. *Business Wire* **1**.
- Casamayor EO, Schafer H, Baneras L, Pedros-Alio C, Muyzer G. (2000). Identification of and spatio-temporal differences between microbial assemblages from two neighboring sulfurous lakes: comparison by microscopy and denaturing gradient gel electrophoresis. *Appl Environ Microbiol* **66**: 499–508.

- Clarke KR. (1993). Non-parametric multivariate analyses of changes in community structure. *Aust J Ecol* **18**: 117–143.
- Daims H, Brühl A, Amann R, Schleifer KH, Wagner M. (1999). The domain-specific probe EUB338 is insufficient for the detection of all *Bacteria*: development and evaluation of a more comprehensive probe set. *Syst Appl Microbiol* **22**: 434–444.
- Elloumi J, Carrias JF, Ayadi H, Sime-Ngando T, Bouaïn A. (2009). Communities structure of the planktonic halophiles in the solar saltern of Sfax, Tunisia. *Estuar Coast Shelf Sci* **81**: 19–26.
- Emptage CD, Knox RJ, Danson MJ, Hough DW. (2009). Nitroreductase from *Bacillus licheniformis*: a stable enzyme for prodrug activation. *Biochem Pharmacol* **77**: 21–29.
- Estrada M, Henriksen P, Gasol JM, Casamayor EO, Pedrós-Alió C. (2004). Diversity of planktonic photoautotrophic microorganisms along a salinity gradient as depicted by microscopy, flow cytometry, pigment analysis and DNA-based methods. *FEMS Microbiol Ecol* **49**: 281–293.
- Gasol JM, Casamayor EO, Joint I, Garde K, Gustavson K, Benlloch S *et al.* (2004). Control of heterotrophic prokaryotic abundance and growth rate in hypersaline planktonic environments. *Aquat Microb Ecol* **34**: 193–206.
- Ghai R, Pašić L, Fernández AB, Martín-Cuadrado AB, Mizuno CM, McMahon KD *et al.* (2011). New abundant microbial groups in aquatic hypersaline environments. *Sci Rep Nature* **1**: 135.
- Gomez-Pereira PR, Schueler M, Fuchs BM, Bennke C, Teeling H, Waldmann J *et al.* (2012). Genomic content of uncultured *Bacteroidetes* from contrasting oceanic provinces in the North Atlantic Ocean. *Environ Microbiol* **14**: 52–66.
- Grant S, Grant WD, Jones BE, Kato C, Li L. (1999). Novel archaeal phylotypes from an East African alkaline saltern. *Extremophiles* **3**: 139–145.
- Guixa-Boixereu N, Calderón-Paz JI, Heldal M, Bratbak G, Pedrós-Alió C. (1996). Viral lysis and bacterivory as prokaryotic loss factors along a salinity gradient. *Aquat Microb Ecol* **11**: 215–227.
- Hammer Ø, Harper DAT, Ryan PD. (2001). PAST: paleontological statistics software package for education and data analysis. *Palaentol Electron* **4**: 1.
- Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**: 119.
- Janse I, Bok J, Zwart G. (2004). A simple remedy against artificial double bands in denaturing gradient gel electrophoresis. *J Microbiol Methods* **57**: 279–281.
- Jiang H, Dong H, Yu B, Liu X, Li Y, Ji S *et al.* (2007). Microbial response to salinity change in Lake Chaka, a hypersaline lake on Tibetan plateau. *Environ Microbiol* **9**: 2603–2621.
- Legault BA, Lopez-Lopez A, Alba-Casado JC, Doolittle WF, Bolhuis H, Rodriguez-Valera F *et al.* (2006). Environmental genomics of "Haloquadratum walsbyi" in a saltern crystallizer indicates a large pool of accessory genes in an otherwise coherent species. *BMC Genomics* **4**: 7–171.
- Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar *et al.* (2004). ARB: a software environment for sequence data. *Nucleic Acids Res* **32**: 1363–1371.
- Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J *et al.* (2012). SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience* **1**: 18.
- Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P *et al.* (2011). Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* **9**: 467–477.
- Martínez-García M, Brazel D, Swan B, Arnosti C, Chain P, Reitenga K *et al.* (2012b). Capturing single cell genomes of active polysaccharide degraders: an unexpected contribution of Verrucomicrobia. *PLoS One* **7**: e35314.
- Martínez-García M, Swan BK, Poulton NJ, Gomez ML, Masland D, Sieracki ME *et al.* (2012a). High-throughput single-cell sequencing identifies photoheterotrophs and chemoautotrophs in freshwater bacterioplankton. *ISME J* **6**: 113–123.
- Massana R, Murray AE, Preston CM, Delong EF. (1997). Vertical distribution and phylogenetic characterization of marine planktonic *Archaea* in the Santa Barbara Channel. *Appl Environ Microbiol* **63**: 50–56.
- Maturrano L, Santos F, Rossello-Mora R, Anton J. (2006). Microbial diversity in Maras Salterns, a hypersaline environment in the Peruvian Andes. *Appl Environ Microbiol* **72**: 3887–3895.
- Murphy JN, Saltikov CW. (2009). The ArsR repressor mediates arsenite-dependent regulation of arsenate respiration and detoxification operons of *Shewanella* sp. strain ANA-3. *J Bacteriol* **191**: 6722.
- Mutlu MB, Martínez-García M, Santos F, Peña A, Guven K, Antón J. (2008). Prokaryotic diversity in Tuz Lake, a hypersaline environment in Inland Turkey. *FEMS Microbiol Ecol* **65**: 474–483.
- Muyzer G, Hottenträger S, Teske A, Wawer C. (1996). Denaturing gradient gel electrophoresis of PCR-amplified 16S rDNA—a new molecular approach to analyse the genetic diversity of mixed microbial communities. In: Akkermans ADL, van Elsas JD, de Bruijn FJ (eds). *Molecular Microbial Ecology Manual*. Kluwer: Dordrecht, Netherland, pp 1–23.
- Narasimarao P, Podell S, Ugalde JA, Brochier-armanet C, Emerson JB, Brocks JJ *et al.* (2012). De novo metagenomic assembly reveals abundant novel major lineage of *Archaea* in hypersaline microbial communities. *ISME J* **6**: 81–93.
- Oh D, Porter K, Russ B, Burns D, Dyal-smith M. (2010). Diversity of Haloquadratum and other haloarchaea in three, geographically distant, Australian saltern crystallizer ponds. *Extremophiles* **14**: 161–169.
- Oren A. (2002). Molecular ecology of extremely halophilic *Archaea* and *Bacteria*. *FEMS Microbiol Ecol* **39**: 1–7.
- Pašić L, Bartual SG, Ulrih NP, Grabnar M, Velikonja BH. (2005). Diversity of halophilic archaea in the crystallizers of an Adriatic solar saltern. *FEMS Microbiol Ecol* **54**: 491–498.
- Pedrós-Alió C. (2005). Diversity of microbial communities: the case of solar salterns. In: Gunde-Cimerman N, Plemenitas A, Oren A (eds). *Adaptation to Life at High Salt Concentrations in Archaea, Bacteria, and Eukarya. Cellular Origins, Life in Extreme Habitats and Astrobiology (COLE)*. Kluwer Academic Publishers: Netherlands, pp 71–90.
- Pedrós-Alió C. (2006). Marine microbial diversity: can it be determined? *Trends Microbiol* **14**: 257–263.
- Pernthaler A, Pernthaler J, Amann R. (2004). Sensitive multicolor fluorescence in situ hybridization for the identification of environmental microorganisms. In:

- Kowalchuk GA, de Bruijn FJIM, Head IM, Akkermans AD, van Elsas JD (eds). *Molecular Microbial Ecology Manual*, 2nd edn Kluwer Academic Publishers: Dordrecht, The Netherlands; Boston, MA, USA; London, UK, pp 711–726.
- Podell S, Ugalde JA, Narasingarao P, Banfield JF, Heidelberg KB, Allen EE. (2013). Assembly-driven community genomics of a hypersaline microbial ecosystem. *PLoS One* **18**: e61692.
- Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J *et al.* (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* **35**: 7188–7196.
- Raghunathan A, Ferguson HR, Bornarth CJ, Song W, Driscoll M, Lasken RS. (2005). Genomic DNA amplification from a single bacterium. *Appl Environ Microbiol* **71**: 3342–3347.
- Richter M, Rosselló-Móra R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci USA* **106**: 19126.
- Rodriguez-Brito B, Li L, Wegley L, Furlan M, Angly F, Breitbart M *et al.* (2010). Viral and microbial community dynamics in four aquatic environments. *ISME J* **4**: 739–751.
- Rusch DB, Halpern AL, Sutton G, Heidelberg KB, Williamson S, Yooseph S *et al.* (2007). The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biol* **5**: e77.
- Santos F, Yarza P, Parro V, Meseguer I, Rosselló-Móra R, Antón J. (2012). Culture-independent approaches for studying viruses from hypersaline environments. *Appl Environ Microbiol* **78**: 1635.
- Schäfer HL, Bernard L, Courties C, Lebaron F, Servais P *et al.* (2001). Microbial community dynamics in Mediterranean nutrient-enriched seawater mesocosms: changes in the genetic diversity of bacterial populations. *FEMS Microbiol Ecol* **34**: 243–253.
- Soerensen KB, Canfield DE, Teske AP, Oren A. (2005). Community composition of a hypersaline endoevaporitic microbial mat. *Appl Environ Microbiol* **71**: 7352–7365.
- Stackebrandt E, Goebel BM. (1994). A place for DNA \pm DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int J Syst Bacteriol* **44**: 846–849.
- ter Braak CJF, Smilauer P. (2002). *CANOCO Reference Manual & CanoDraw for Windows User's Guide: Software for Canonical Community Ordination (version 4.5)*. Microcomputer Power: Ithaca, New York, USA, p 500.
- Thompson JR, Marcelino LA, Polz MF. (2002). Heteroduplexes in mixed-template amplifications: formation, consequence and elimination by 'reconditioning PCR'. *Nucleic Acids Res.* **30**: 2083–2088.
- Wallner G, Amann R, Beisker W. (1993). Optimizing fluorescent in situ hybridization with rRNA-targeted oligonucleotide probes for flow cytometric identification of microorganisms. *Cytometry* **14**: 136–143.
- Wang J, Yang D, Zhang Y, Shen J, van der Gast C, Hahn MW *et al.* (2011). Do patterns of bacterial diversity along salinity gradients differ from those observed for macroorganisms. *PLoS One* **6**: e27597.
- Zhaxybayeva O, Stepanauskas R, Mohan NR, Papke RT. (2013). Cell sorting analysis of geographically separated hypersaline environments. *Extremophiles* **17**: 265–275.

Supplementary Information accompanies this paper on The ISME Journal website (<http://www.nature.com/ismej>)