The ISME Journal (2010) 4, 399–407

www.nature.com/ismej

ORIGINAL ARTICLE

Comparative community genomics in the Dead Sea: an increasingly extreme environment

Idan Bodaker[1,9], Itai Sharon[1,2,9], Marcelino T Suzuki[3], Roi Feingersch[1], Michael Shmoish[4], Ekaterina Andreishcheva[5], Mitchell L Sogin[5], Mira Rosenberg[1], Michael E Maguire[6], Shimshon Belkin[7], Aharon Oren[7,8] and Oded Béjà[1]

[1]*Faculty of Biology, Technion—Israel Institute of Technology, Haifa, Israel;* [2]*Faculty of Computer Science, Technion—Israel Institute of Technology, Haifa, Israel;* [3]*CNRS and Université Pierre et Marie Curie (Paris 6), Observatoire Océanologique de Banyuls, Ave du Fontaulé, Banyuls-sur-Mer, France;* [4]*Bioinformatics Knowledge Unit, Lorry I. Lokey Interdisciplinary Center for Life Sciences and Engineering,Technion—Israel Institute of Technology, Haifa, Israel;* [5]*Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, Woods Hole, MA, USA;* [6]*Department of Pharmacology, Case Western Reserve University, Cleveland, OH, USA;* [7]*Department of Plant and Environmental Sciences, Alexander Silberman Institute of Life Sciences, Hebrew University of Jerusalem, Jerusalem, Israel and* [8]*Moshe Shilo Minerva Center for Marine Biogeochemistry, Hebrew University of Jerusalem, Jerusalem, Israel*

Owing to the extreme salinity (~10 times saltier than the oceans), near toxic magnesium levels (~2.0 M Mg^{2+}), the dominance of divalent cations, acidic pH (6.0) and high-absorbed radiation flux rates, the Dead Sea represents a unique and harsh ecosystem. Measures of microbial presence (microscopy, pigments and lipids) indicate that during rare bloom events after exceptionally rainy seasons, the microbial communities can reach high densities. However, most of the time, when the Dead Sea level is declining and halite is precipitating from the water column, it is difficult to reliably measure the presence of microorganisms and their activities. Although a number of halophilic *Archaea* have been previously isolated from the Dead Sea, polar lipid analyses of biomass collected during Dead Sea blooms suggested that these isolates were not the major components of the microbial community of these blooms. In this study, in an effort to characterize the perennial microbial community of the Dead Sea and compare it with bloom assemblages, we performed metagenomic analyses of concentrated biomass from hundreds of liters of brine and of microbial material from the last massive Dead Sea bloom. The difference between the two conditions was reflected in community composition and diversity, in which the bloom was different and less diverse from the residual brine population. The distributional patterns of microbial genes suggested Dead Sea community trends in mono- and divalent cation metabolisms as well as in transposable elements. This may indicate possible mechanisms and pathways enabling these microbes to survive in such a harsh environment.
The ISME Journal (2010) **4,** 399–407; doi:10.1038/ismej.2009.141; published online 24 December 2009
Subject Category: integrated genomics and post-genomics approaches in microbial ecology
Keywords: hypersaline, salterns, halophiles, Halobacterium, Haloquadratum, metagenomics, halophiles

Introduction

The Dead Sea, located on the border between Israel, Jordan and the Palestinian Authority (area approximately 632 km^2, maximum depth of 300 m), presents fascinating challenges to the understanding of biological processes and the limits of life in one of the most extreme environments on earth (Boetius and Joye, 2009). Its waters currently contain approximately 347 g l^{-1} of salts, and have a highly unusual ionic composition: divalent cations (1.98 M Mg^{2+}, 0.47 M Ca^{2+}) dominate over monovalent cations (1.54 M Na^+, 0.21 M K^+). The anions are 99% Cl^- (6.48 M) and 1% Br^- (0.08 M) (values for the year 2007 (Bodaker *et al.*, 2009)). Sulfate concentrations are low (0.004 M), and the brine has a pH of approximately 6.0. The water budget of the Dead Sea is negative, and during the past decade it has dropped by an average of 1 m per year. The water is supersaturated with respect to Na^+, and massive amounts of halite (NaCl) precipitate to the bottom.

Currently, the conditions of the Dead Sea are too extreme to support high growth of microorganisms. Although many halophilic and halotolerant microorganisms are known to inhabit NaCl-dominated ('thalassohaline') brines up to halite saturation, only

Correspondence: A Oren or O Béjà, Faculty of Biology, Technion, Israel Institute of Technology, Technion City, Haifa, 32000, Israel. E-mail: orena@cc.huji.ac.il or beja@tx.technion.ac.il
This paper is dedicated to the memory of Moti Gonen, the skipper of the RV Taglit.
[9]These authors contributed equally to this work.
Received 12 October 2009; revised 19 November 2009; accepted 22 November 2009; published online 24 December 2009

a few organisms can live in the Dead Sea's 'athalassohaline' environment, dominated by divalent cations. Systematic monitoring since 1980 (Oren, 1999) has shown that microbial blooms consisting of the unicellular green alga *Dunaliella* and extremely halophilic *Archaea* of the *Halobacteriaceae* family develop only after significant dilution of the upper meters of the water column after exceptionally rainy winters. Two such events have been documented: in 1980, and even more dramatically in 1992, when the upper 5 m of the water column was diluted up to 70%, supporting a density of Archaea up to 3.5×10^7 ml^{-1}, coloring the Dead Sea water red (Oren and Gurevich, 1995). A number of species of halophilic *Archaea* have been isolated from the Dead Sea, including *Haloferax volcanii* (Mullakhanbhai and Larsen, 1975), *Haloarcula marismortui* (Oren *et al*., 1990), *Halorubrum sodomense* (Oren, 1983) and *Halobaculum gomorrense* (Oren *et al*., 1995). However, polar lipid analyses of biomass collected during Dead Sea blooms suggested that these isolates were not the major components of the microbial community at the time (Oren and Gurevich, 1993). From 1996 onward, *Dunaliella* cells were no longer observed in the water, and the numbers of prokaryotes remained below 5×10^5 ml^{-1}.

Materials and methods

Sample collection and fosmid library construction
Fosmid libraries were constructed from samples collected in the Dead Sea during a massive bloom in 1992 and in non-bloom conditions in 2007. The 1992 bloom sample was collected in 1992 at the Ein Gedi 310 station (31°25′N, 31°29′E), the deepest point at the center of the Dead Sea and had been cryopreserved since then. The 2007 sample was collected on March 8, 2007 at the Ein Gedi 100 station located 4 km east of Ein Gedi, at the location of the moored meteorological station (31° 25′ N, 35° 26′ E), in which the depth of the Dead Sea is approximately 100 m. Approximately, 800 l of Dead Sea surface water were mixed with 1/10 volumes of prefiltered double-distilled freshwater (790 l of brine with 10 l freshwater) and concentrated by tangential flow filtration with a Pellicon 2 unit (Millipore, Billerica, MA, USA) equipped with a C screen 0.22 μm Durapore 0.5 m^2 cartridge. Freshwater addition was a prerequisite to avoid salts crystallizing inside the 0.22 μm cartridge. Cells were collected by centrifugation of the retentate (4 °C, 38 900 × *g*, 0.5 h) and the pellet was embedded in agarose plugs (Béjà *et al*., 2000). The 1992 bloom sample was collected from surface waters and cells from ~5 l were concentrated by centrifugation and the reddish cell pellet was stored in −80 °C. For plug preparation, the 1992 bloom pellet was thawed on ice in July 2008 and the cells were embedded in agarose plugs. DNA was extracted from the agarose plugs using lysozyme and proteinase K treatments ((Béjà *et al*., 2000); detailed description could be found in (Sabehi and Béjà, 2007)) from both samples. Before cloning, DNA from both the 2007 and the crypreserved 1992 plugs were run in agarose gels and were shown to be intact (data not shown). The DNA was cloned into the pCC1fos vector (Epicentre, Madison, WI, USA) according to manufacturer's instructions to construct two fosmid libraries. Ten 96-well plates were taken for random end sequencing from both fosmid ends for each library.

Sequence preparation
Out of the 1920 BAC-end sequences performed on each library, 1773 (1992 library) and 1759 (2007 library) were used for further analyses after quality control (low-quality sequence ends were removed from all sequences from the 1992 bloom and 2007 data sets using Phred (Ewing and Green, 1998; Ewing *et al*., 1998).) and the removal of duplicates. On average, the read length was 734 or 844 bp for the 1992 or 2007 libraries, respectively. In addition, all sequences were scanned for possible vector contamination by BLASTing, the clone ends against the sequence libraries and removing all 100% hits. Multiple copies of clones coming from the same cultures were detected and removed.

Archaeal 16S rRNA amplification and construction of 16S rRNA tree
PCR 16S ribosomal RNA (rRNA) gene amplification was carried out using the 16S rRNA archaeon-specific primers Ar20F (5′-TTCCGGTTGATCCYGC CRG-3′) and Ar958R (5′-YCCGGCGTTGAMTCCA ATT-3′) (Hallam *et al*., 2003). The 20-μl amplification reaction mixtures contained 1 μl of template DNA, 2 μl of 10 × optibuffer, 0.4 μl each of 10 μM forward and reverse primer, 0.3 μl of BIO-X-ACT short DNA polymerase (Bioline, Berlin, Germany), 1 μl of 50 mM co-factor ($MgCl_2$) and 2 μl of 10 mM stock deoxynucleoside triphosphate mixture. Amplifications were carried out with the following profile: 94 °C for 2 min, then 29 cycles of 94 °C for 30 s, 54 °C for 30 s and 68 °C for 40 s, followed by a final extension at 68 °C for 10 min. Forty partial (ca. 920 bp) 16S rRNA sequences from the Dead Sea were verified for chimeras using BELLEROPHON (Huber *et al*., 2004) and a single putatively chimeric sequence was excluded from further analyses. The reminder sequences, a sequence from a saltern in Croatia (DQ889331) and the sequences of *Halomicrobium katesii* AI-5 (EF333994) and *Halobaculum gomorrense* DSM9297 (sequenced in this study) were imported into the SSUref_96 tree from the Silva project (Pruesse *et al*., 2007) using ARB (Ludwig *et al*., 2004). Sequences were automatically aligned, and the alignments manually verified against a set of sequences composed of: (1) All sequences of strains of officially named species (Deutsche Sammlung von Mikroorganismen und Zellkulturen, www.dmsz.org) in the *Halobacteriaceae*

family (International Committee on Systematics of Prokaryotes, www.the-icsp.org); (2) Previously published sequences longer than 1300 bp from uncultured microorganisms labeled as 'Halobacteriales' in the silva project tree tree_SSURef _1200_svl_96; and (3) *Methanospirillum hungatei* JF1 as the outgroup. A total of 280 sequences were used to construct a neighbor-joining phylogenetic tree using the F84 substitution model (Kishino and Hasegawa, 1989; Felsenstein and Churchill, 1996). This analysis used alignment positions that were selected based on a filter (mask) that included all positions common to more than 50% of the sequences, and excluded any position with ambiguities in any of the sequences. A total of 714 sequence positions were used to construct a tree and to calculate bootstrap values based on 100 random resamplings using PHYLIP (Felsenstein, 2005).

Metabolic potential
Metabolic potential was determined with the IMG/M (Integrated Microbial Genome with Microbiome Sample) (Markowitz *et al*., 2008a) of the Joint Genome Institute server (http://img.jgi.doe.gov/m). Frequency estimations of the different protein families were computed based on the statistical framework from (Sharon *et al*., 2009). Briefly, the frequency F_p of a protein family P in a sample D is given by

$$F_{p=}\frac{\hat{C}_p}{\sum_{Q\in D}\hat{C}_Q}$$

where $\hat{C}_p$ is an estimator for the number of occurrences of P in the sample, $\hat{C}_P$. Given that P is observed on R_p reads, $\hat{C}_P$ is given by

$$\hat{C}_p = \frac{R_p}{2\alpha(L(P)+g-2\mathrm{T})}$$

where *L(P)* is *P*'s length, g is the average read length, T is the minimum segment of P required for the gene to be discovered, and α is the sample coverage. It is to be noted that α is eliminated when F_P is computed.

Cluster of Orthologous Group (COG) assignment was performed using the IMG/M system (Markowitz *et al*., 2008b). The COG assignment of the nine environments was tabulated and is presented in Supplementary Table S1. We used the Expander 4.0.2 for clustering (Shamir *et al*., 2005). From all COGs, we chose the 1000 most variable, performed a standardization (mean 0, variance 1 transformation) and hierarchical clustering with complete linkage (Figure 6).

Library-to-library similarities
Measurement of library-to-library similarities was made after BLASTx querying of all environments against each other and themselves using the following parameters (-p tblastx -e 1e-2 -F F). This very low expectation cutoff value was chosen because the different environments compared (three hypersaline environments and nine different marine samples) were very different from each other. The data set was computed according to (Rusch *et al*., 2007) with a cutoff value of 80% and only for fragments of more than 200 bp.

Species recruitment
The (Figure 3) two Dead Sea (1992 bloom and 2007) as well as the Spanish Salterns (Legault *et al*., 2006) data sets were blasted (-p blastn –r 2 –q -3 –F F –e 1e-2) against the genomes of *Haloarcula marismortui* ATCC 43049, *Halobacterium salinarum* R1, *Halobacterium* sp. NRC-1, *Haloferax volcanii* DS2, *Haloquadratum walsbyi* DSM 16790, *Natronomonas pharaonis* DSM 2160 and *Salinibacter ruber* DSM 13855. For each read with at least one hit of which alignment was longer than 400 bp with percent identity higher than 60% the best hit was plotted on Figure 3.

Nucleotide sequence accession numbers
Fosmid end sequences were deposited in GenBank under accession #s GQ444463-GQ446233 (1992) and GQ446234-GQ447987 (2007), and 16S rRNA gene sequences GQ861381-GQ861393 (1992) and GQ861355-GQ861378 & GQ861380 (2007).

Results and discussion

We embarked on a detailed metagenomic study of two different Dead Sea environments using the random bidirectional DNA sequencing of fosmid clones approach that was recently developed and used by DeLong *et al*. to characterize microbial communities in the ALOHA station (DeLong *et al*., 2006). We analyzed uncultivated microbial population structures and predicted metabolic activities in samples from the year 2007 and from the 1992 microbial bloom material that had been cryopreserved for nearly 16 years. The DNA of the cryopreserved sample appeared to be intact, however, the influence of such a long cryopreservation on the integrity of the bloom population verified could not be ruled out. The cell morphology of the 2007 population was pleomorphic, not different from the morphology observed with the 1992 bloom (see Figure 2 in Oren and Gurevich, 1993). Total Dead Sea community DNA from 2007 was purified from ~800 lof concentrated brine. Technically, Dead Sea brine microbes could not be concentrated because of massive salt precipitation inside the tangential flow device. Sample concentration was therefore achieved only after previous dilution with 1/10 volume of prefiltered freshwater, which lowered salt concentrations to below the precipitation point. The DNA was cloned into fosmids and both ends of clones were sequenced.

A PCR-based survey with archaeal-specific 16S rRNA primers was performed directly on Dead Sea samples as well as on fosmid clones. Not surprisingly,

the analysis revealed a diverse community of halophiles; however, most rRNAs had <93% similarity to any rRNA sequence currently in the GenBank (either from cultured organisms or from environmental clones, Figure 1). Of the different rRNA phylotypes recovered from the fosmids, three were somewhat similar to *Haloplanus natans* (96% in the 2007 sample), *Halorhabdus utahensis* (93% in the 2007 sample) and *Halobacterium noricense* (95% in the 1992 bloom), while the rest were <89–92% identical to any known halophilic isolate or environmental 16S rRNA sequence. A similar trend was observed with rRNAs found at the fosmid ends (both 16S and 23S rRNA, some with <88%

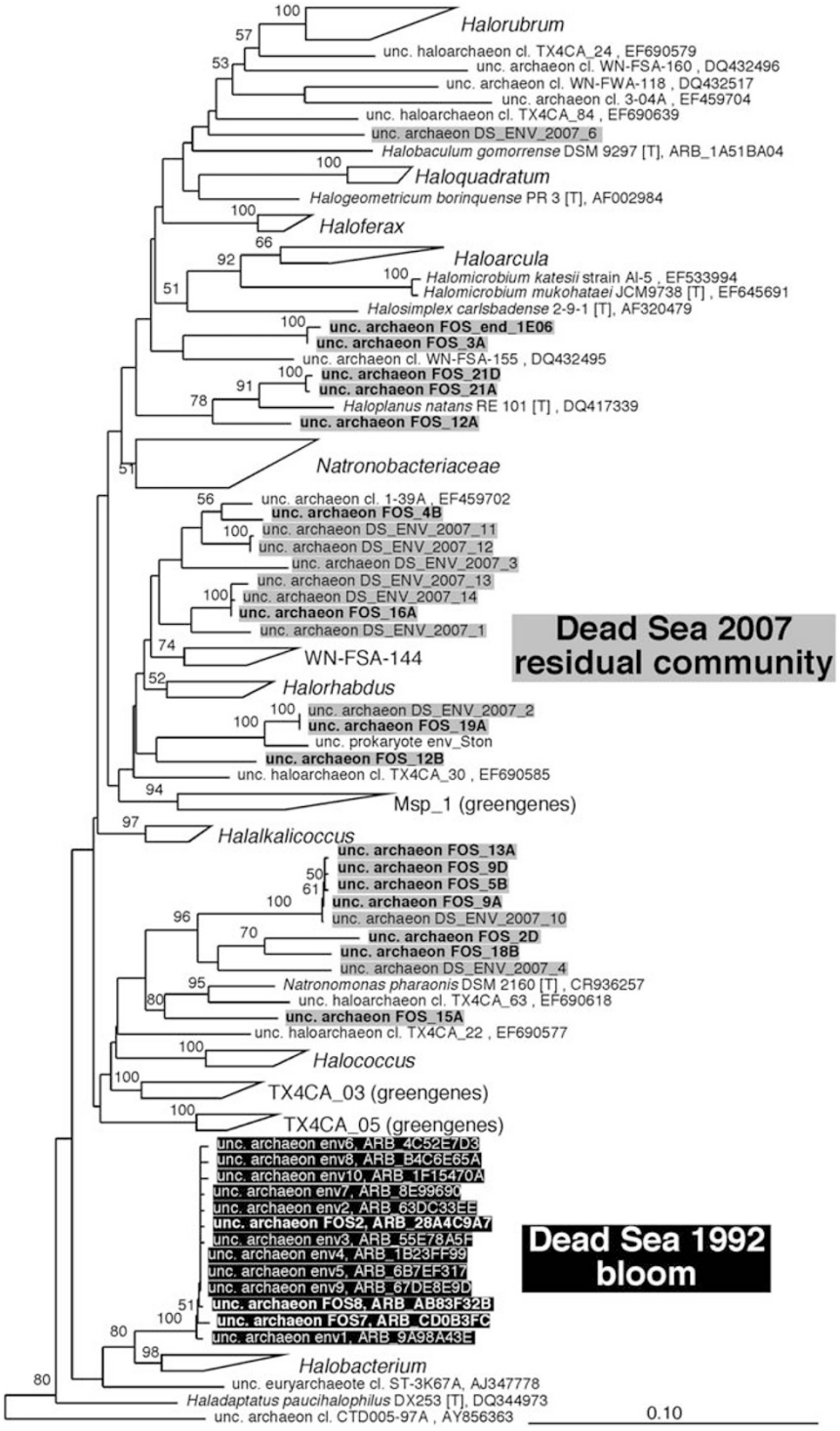

Figure 1 Phylogenetic tree of archaeal Dead Sea environmental and fosmid 16S rRNAs. *Methanosprillum hungatei* JF-1, CP000254 was used as an outgroup and is not shown for the sake of clarity. rRNAs sequences from fosmids are shown in boldface.

similarity to known halophiles). These results indicate that the community that adapted to survive in the brine is unique to the Dead Sea environment, and is likely adapted to survive under the harsh conditions of the Dead Sea. Interestingly, although rRNAs retrieved from the 2007 sample were affiliated with different haloarchaeal groups, the rRNAs from the bloom were found only as one cluster somewhat related to *Halobacterium salinarum*, but also without other closer published sequences. Morphologically, the Archaea in the bloom were of a flat, pleomorphic type, and not long rods as all *Halobacterium* species. Moreover, analysis of the polar lipid composition of the bloom community did not show the characteristic glycolipids of the genus *Halobacterium* (Oren and Gurevich, 1993). The dominant type of archaeon thus belonged to a new genus with 93% 16S rRNA gene similarity to *Halobacterium salinarum*.

Analysis of bacterial 16S rRNA obtained from the environment showed sequences belonging to *Firmicutes* (93–99% similarity) and *Fusobacteria* (89% similarity). However, we failed to detect the presence of bacterial rRNA in our fosmid library, which might indicate the absence or very low quantity of *Bacteria* in the Dead Sea samples.

We used a pyro-tag sequencing strategy that combines the use of amplicons of the V6 hypervariable region of the 16S rRNA as proxies for the presence of individual phylotypes (operational taxonomic units) with massively parallel sequencing (Sogin *et al*., 2006). Our goal was to provide assessments of archaeal diversity, evenness and community structure at a resolution two to three orders of magnitude greater than what we achieved by cloning and capillary sequencing of longer 16S rRNA amplicons. Indeed, as hinted by the long 16S rRNA amplicons analysis, the 1992 bloom was composed of several dominant *Halobacteriaceae* phylotypes that are not present in the 2007 community (Figure 2) as well as contained a larger number of moderate-to-high abundance taxa, and significantly higher ACE and Chao1 values (Table 1).

Recruitment of randomly sequenced fosmid ends against the non redundant GenBank database (BLASTn (Altschul *et al*., 1997)) showed that the majority of ends from the Dead Sea 1992 bloom, as well as from the 2007 residual community, had best hits at approximately 75–80% identity (Figure 3). This is in sharp contrast to the 95–100% identity observed earlier with fosmid ends retrieved under bloom conditions in a Spanish saltern crystallizer pond in which the population was dominated by square Archaea based on similar criteria (Legault *et al*., 2006). This suggests that many of the Archaea present in the Dead Sea have no cultured counterparts in collections.

Taxonomic bins of protein homologs found at the fosmid ends (BLASTx high-scoring sequence pair against the non-redundant GenBank database using the Joint Genome Institute blast tool confirmed the

Table 1 Dead Sea samples, Chao and ACE richness estimates

	Cluster distance			
	0.03		*0.06*	
Sample year	*ACE*	*Chao*	*ACE*	*Chao*
1992	440	431	182	180
2007	715	725	312	334
Ratio	1.625	1.682	1.714	1.855

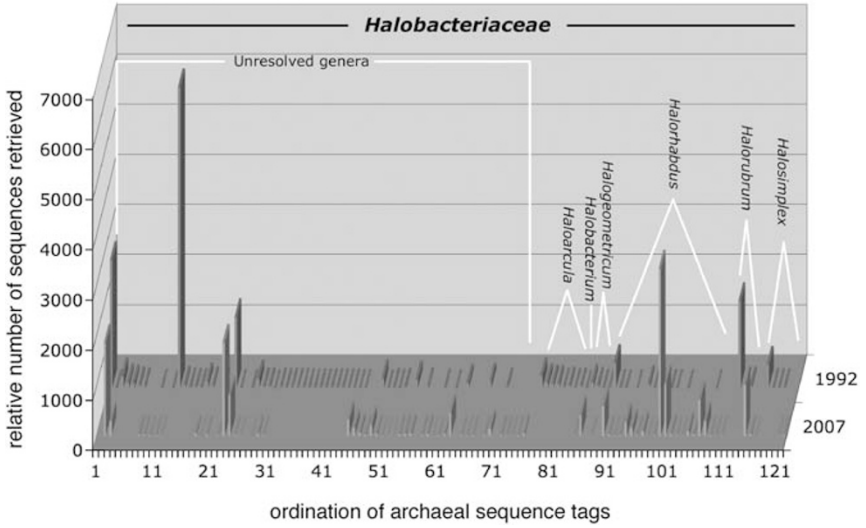

Figure 2 Pyro-tag sequencing of Dead Sea archaeal V6 region. Histogram of archaeal pyro-tag sequences from the 1992 (red) and 2007 (blue) Dead Sea samples. The *X* axis represents the ordination of archaeal sequence tags that occurred at least 15 times summed across both samples. After assigning taxonomy to the genus level by the GAST algorithm, pyro-tags were first sorted alphabetically by a taxonomic name then numerically by increasing GAST distance from the nearest V6 reference sequence. The major taxonomic designations are labeled. The color reproduction of this figure is available on the html full text version of the manuscript.

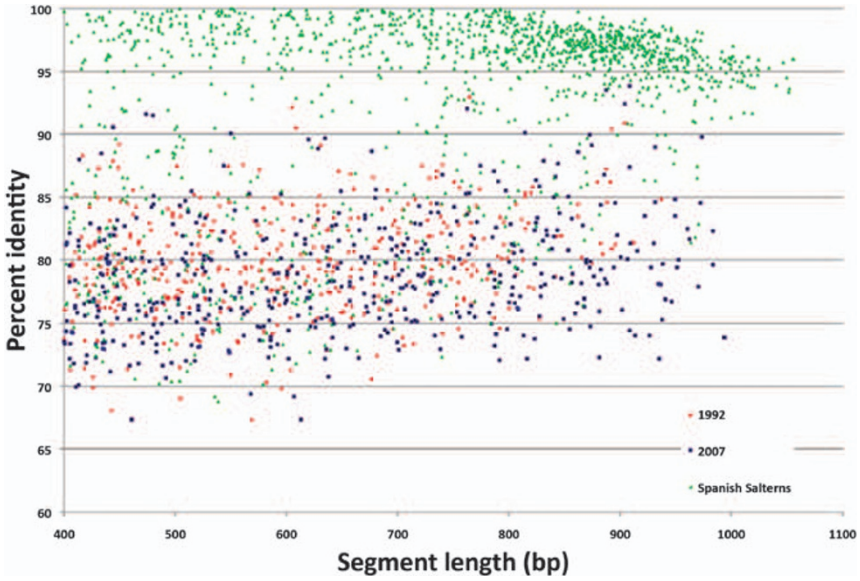

Figure 3 Recruitment of different hypersaline data sets against fully sequenced genomes. Each dot represents the most significant alignment of a read queried against one of the following genomes: *Haloarcula marismortui* ATCC 43049, *Halobacterium salinarum* R1, *Halobacterium* sp. NRC-1, *Haloferax volcanii* DS2, *Haloquadratum walsbyi* DSM 16790, *Natronomonas pharaonis* DSM 2160 and *Salinibacter ruber* DSM 13855 (an extremely halophilic member of the *Bacteria*). The *X* axis represents the length of alignment (min. 400 bp); the *Y* axis represents percent identity (min 60%). As can be observed, the Spanish salterns data set yields relatively high and long alignments (mostly against *Haloquadratum walsbyi*, data not shown) while most sequences of both Dead Sea data sets do not have close fully sequenced relatives.

rRNA-based observations of archaeal dominance (~95% of assigned genes). Archaea distribution suggested that the residual community of the Dead Sea in 2007 was composed of different haloarchaeal groups with no dominance of a specific group (Figure 4). Under bloom conditions, however, *Halobacterium*-like high-scoring sequence pairs reached 50% from archaeal assigned hits. This is also in contrast to the observations made under bloom conditions in the Spanish saltern crystallizer pond (Legault *et al.*, 2006) in which the population was dominated by environmental '*Haloquadratum walsbyi'*. Small numbers of bacterial high-scoring sequence pairs were also observed in the Dead Sea metagenome, adding up to 5–7% of the protein homologs. These could be the result of a small yet significant bacterial presence in the Dead Sea residual community or could alternatively be coming from bacterial-like sequences embedded in archaeal genomes (Nelson *et al.*, 1999; Worning *et al.*, 2000).

To find Dead Sea communities' enriched traits, fosmid end sequences were compared against the COG database (Tatusov *et al.*, 1997; Tatusov *et al.*, 2003) in comparison with fosmid or Bacteria Artificial Chromosome (BAC) libraries from the Spanish saltern data set (Legault *et al.*, 2006) and from different marine environments (DeLong *et al.*, 2006; Martín-Cuadrado *et al.*, 2007; Feingersch *et al.*, 2010). Interestingly, although the Spanish sample was composed of a different archaeal repertoire and had, on average, a much different percent G+C content (GC%) content (Figure 5), it

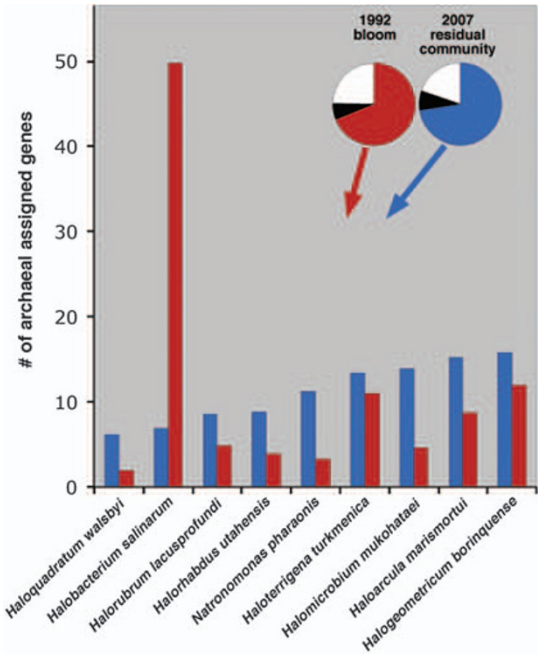

Figure 4 Taxon distribution of top Dead Sea high-scoring sequence pairs (HSPs). Top pies represent unassigned (white), bacterial (black) or archaeal (red for 1992 and blue for 2007) gene assignments. Columns represent archaeal assigned genes at the species level.

clustered with the Dead Sea samples based on COG categories and was somewhat more similar to the 2007 sample (Figure 6 and Supplementary Table S1). This indicates that the metabolic capacity

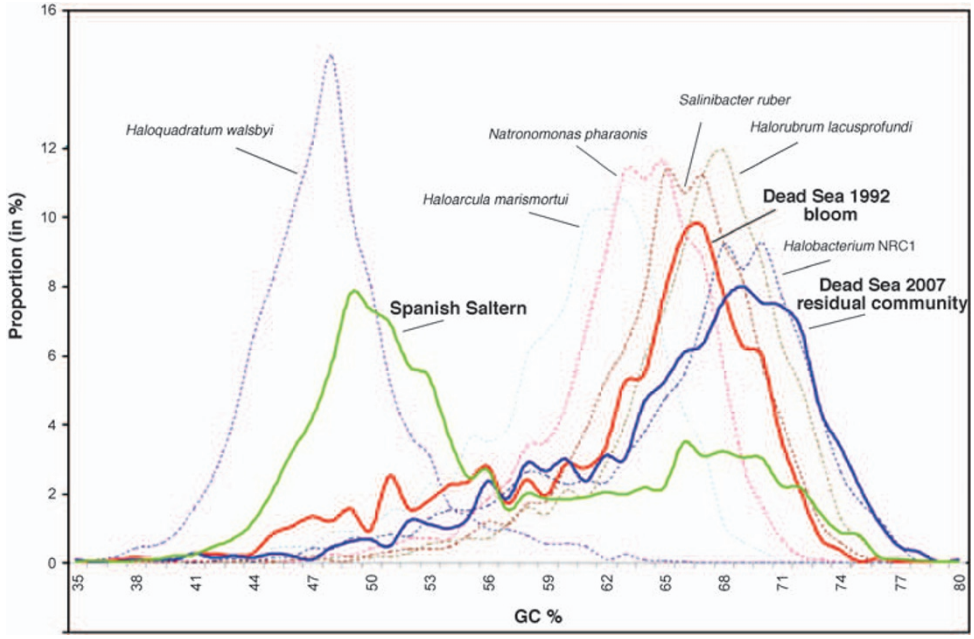

Figure 5 GC% distribution of the Dead Sea 1992 bloom and the 2007 residual community fosmid ends. Data for selected haloarchaeal and halobacterial genomes, as well as a Spanish saltern crystallizer pond fosmid ends, are shown for reference.

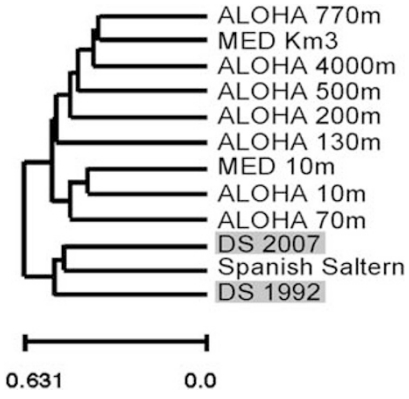

Figure 6 Hierarchical clustering of different environments based on metabolic potential of fosmid ends. The 1000 most variable Cluster of Orthologous Group (COG) categories from the Dead Sea (DS2007 and DS1992), Spanish saltern and nine marine environments (ALOHA station (DeLong *et al.*, 2006) and Mediterranean Sea surface (Feingersch *et al.*, 2010) and deep (Martín-Cuadrado *et al.*, 2007) water samples) were normalized, standardized and clustered using the Expander statistical package. Normalized COG assignments to the different environments are found in Supplementary Table S1.

required to survive under high salt conditions is shared between different high salt environments. Enrichment in COGs related to transposable elements (Supplementary Figure S1), intein and homing endonucleases, Na^+/H^+ and K^+/H^+ antiporters, as well as divalent cation (Mg^{2+} and Co^{2+}) channels (Figure 7a) was observed in the Dead Sea environments. Enrichment in transposases was also observed earlier in the Spanish saltern environment (Legault *et al.*, 2006) and is consistent with the high number of transposable elements observed in halophile genomes (Charlebois and Doolittle, 1989; Grogan, 2007).

The enrichment in putative archaeal (see phylogenetic tree in Figure 7b) CorA magnesium channels (up to 11-fold for COG 0598; see Figure 7a) in the highly magnesium-enriched Dead Sea environment points to a potential, yet unknown, resistance mechanism used by Dead Sea halophiles. Although CorA (COG 0598) is usually associated with Mg^{2+} influx activity, when exposed to high extracellular Mg^{2+} concentrations, CorA can mediate Mg^{2+} efflux (Gibson *et al.*, 1991; Papp-Wallace and Maguire, 2008). Interestingly, the $Mg^{2+}/Co^{2+}/Ni^{2+}$ channel MgtE (COG 2239; Figure 7a) was completely absent from the Dead Sea libraries.

Mg^{2+} is a chaotropic ion, and as such is toxic at high concentrations. In the absence of compensating (kosmotropic) ions such as Na^+, a Mg^{2+} concentration of 2.3 M appears to be the upper limit for life (Hallsworth *et al.*, 2007). In the Dead Sea, the Mg^{2+} concentration rapidly approaches this upper value when the Na^+ content of the water decreases. Little is known regarding the intracellular Mg^{2+} and Ca^{2+} concentrations of halophilic Archaea in the Dead Sea. Experiments to estimate such concentrations in the Dead Sea isolate *Hrr. sodomense* grown in up to 0.95 M Mg^{2+} and 0.5 M Ca^{2+} yielded values that were $<40\%$ of the concentrations in their surrounding medium, and those values were probably overestimations because of methodological problems (Oren, 1986). In any case, it is clear that efficient removal of divalent cations is an essential function for microorganisms in the Dead Sea, and therefore the enrichment in putative CorA magnesium channels in the active community of 1992 may well point

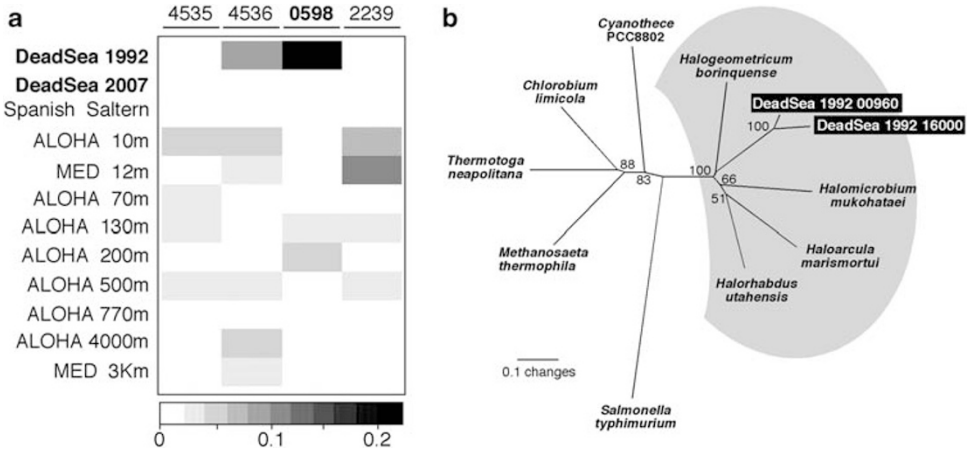

Figure 7 Frequencies and phylogenetic analysis of Mg^{2+} transport related COGs from the Dead Sea. (**a**) Frequencies of different Mg^{2+} transport related COGs from the Dead Sea fosmid ends in comparison with the Spanish saltern crystallizer, and nine different marine environments are presented in the grayscale gradient. The intensity of gray shading is proportional to the percentage of each COG category from all COG hits at a specific environment. Normalized numbers can be found in Supplementary Table S1. Numbers above the heat map represent COG numbers. (**b**) Phylogenetic analysis of COG0598 proteins (Mg^{2+} channel CorA protein) from the 1992 Dead Sea sample.

to a special adaptation of life in the high-magnesium waters of the Dead Sea.

Drying out of the Dead Sea with a further increase in divalent cation concentrations is expected to continue in the years to come. It will be interesting to repeat these metagenomics studies in the future (providing that microbial communities will still be found in the Dead Sea) to monitor further adaptations of life to the increasingly extreme conditions of the unique environment of the Dead Sea.

Acknowledgements

We thank the late M Gonen and the crew of the RV *Taglit* for their logistic support and technical assistance at sea, and N Kyrpides and VM Markowitz for their assistance and support with the IMG/M system. This work was supported by grants from the Bridging the Rift (BTR) Foundation (IB, SB, AO and OB) and the NASA Astrobiology Institute Cooperative Agreement, NASA NNA04CC04A (MLS).

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.

Béjà O, Suzuki MT, Koonin EV, Aravind L, Hadd A, Nguyen LP *et al.* (2000). Construction and analysis of bacterial artificial chromosome libraries from a marine microbial assemblage. *Environ Microbiol* **2**: 516–529.

Bodaker I, Béjà O, Sharon I, Feingersch R, Rosenberg M, Oren A *et al.* (2009). Archaeal diversity in the Dead Sea: microbial survival under increasingly harsh conditions. In: Oren A, Naftz DL, Palacios P, Wurtsbaugh WA (eds) *Saline Lakes Around the World: Unique Systems with Unique Values*. The SJ and Jessie E Quinney Natural Resources Research Library, College of Natural Resources, Utah State University: Salt Lake City, UT, pp 137–143.

Boetius A, Joye S. (2009). Ecology thriving in salt. *Science* **324**: 1523–1525.

Charlebois RL, Doolittle WF. (1989). Transposable elements and genome structure in Halobacteria. In: Berg D, Howe M (eds) *Mobile DNA*. ASM Press: Washington DC, pp 297–307.

DeLong EF, Preston CM, Mincer T, Rich V, Hallam SJ, Frigaard N-U *et al.* (2006). Community genomics among stratified microbial assemblages in the ocean's interior. *Science* **311**: 496–503.

Ewing B, Green P. (1998). Basecalling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* **8**: 186–194.

Ewing B, Hillier L, Wendl M, Green P. (1998). Basecalling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* **8**: 175–185.

Feingersch R, Suzuki MT, Shmoish M, Sharon I, Sabehi G, Partensky F *et al.* (2010). Microbial community genomics in eastern Mediterranean Sea surface waters. *ISME J* **4**: 78–87.

Felsenstein J. (2005). *PHYLIP (Phylogeny Inference Package)*. Department of Genome Sciences, University of Washington: Seattle.

Felsenstein J, Churchill G.A. (1996). A Hidden Markov Model approach to variation among sites in rate of evolution. *Mol Biol Evol* **13**: 93–104.

Gibson MM, Bagga DA, Miller CG, Maguire ME. (1991). Magnesium transport in *Salmonella typhimurium*: the influence of new mutations conferring Co2+ resistance on the CorA Mg2+ transport system. *Mol Microbiol* **5**: 2753–2762.

Grogan DW. (2007). Mechsnisms of genome stability and evolution. In: Cavicchioli R (ed) *Archaea: Molecular and Cellular Biology*. ASM Press: Washington, DC, pp 120–138.

Hallam SJ, Girguis PR, Preston CM, Richardson PM, DeLong EF. (2003). Identification of methyl coenzyme M Reductase A (*mcrA*) genes associated with methane-oxidizing archaea. *Appl Environ Microbiol* **69**: 5483–5491.

Hallsworth JE, Yakimov MM, Golyshin PN, Gillion JL, D'Auria G, de Lima Alves F *et al.* (2007). Limits of life in $MgCl_2$-containing environments: chaotropicity defines the window. *Environ Microbiol* **9**: 801–813.

Huber T, Faulkner G, Hugenholtz P. (2004). Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics* **20**: 2317–2319.

Kishino H, Hasegawa M. (1989). Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in hominoidea. *J Mol Evol* **29**: 170–179.

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* **7**: 171.

Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar *et al.* (2004). ARB: a software environment for sequence data. *Nucleic Acid Res* **32**: 1363–1371.

Markowitz VM, Szeto E, Palaniappan K, Grechkin Y, Chu K, Chen IM *et al.* (2008a). The integrated microbial genomes (IMG) system in 2007: data content and analysis tool extensions. *Nucleic Acids Res* **36**: D528–D533.

Markowitz VM, Ivanova NN, Szeto E, Palaniappan K, Chu K, Dalevi D *et al.* (2008b). IMG/M: a data management and analysis system for metagenomes. *Nucleic Acids Res* **36**: D534–D538.

Martín-Cuadrado AB, López-García P, Alba JC, Moreira D, Monticelli L, Strittmatter A *et al.* (2007). Metagenomics of the deep Mediterranean, a warm bathypelagic habitat. *PLoS ONE* **2**: e914.

Mullakhanbhai MF, Larsen H. (1975). *Halobacterium volcanii* spec. nov., a Dead Sea halobacterium with a moderate salt requirement. *Arch Microbiol* **104**: 207–214.

Nelson KE, Clayton RA, Gill SR, Gwinn ML, Dodson RJ, Haft DH *et al.* (1999). Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of *Thermotoga maritima*. *Nature* **399**: 323–329.

Oren A. (1983). *Halobacterium sodomense* sp. nov., a Dead Sea halobacterium with an extremely high magnesium requirement. *Int J Syst Bacteriol* **33**: 381–386.

Oren A. (1986). Relationships of extremely halophilic bacteria towards divalent cations. In: Megusar F, Gantar M (eds) *The 4th ISME Conference, Perspectives in Microbial Ecology*. Slovene Society for Microbiology: Ljubljana, Slovenia, pp 52–58.

Oren A. (1999). Microbiological studies in the Dead Sea: future challenges toward the understanding of life at the limit of salt concentrations. *Hydrobiologia* **405**: 1–9.

Oren A, Gurevich P. (1993). Characterization of the dominant halophilic archaea in a bacterial bloom in the Dead Sea. *FEMS Microbiol Ecol* **12**: 249–256.

Oren A, Gurevich P. (1995). Dynamics of a bloom of halophilic archaea in the Dead Sea. *Hydrobiologia* **315**: 149–158.

Oren A, Gurevich P, Gemmell RT, Teske A. (1995). *Halobaculum gomorrense* gen. nov. sp. nov., a novel extremely halophilic archaeon from the Dead Sea. *Int J Syst Bacteriol* **45**: 747–754.

Oren A, Ginzburg M, Ginzburg BZ, Hochstein LI, Volcani BE. (1990). *Haloarcula marismortui* (Volcani) sp. nov., nom. rev., an extremely halophilic bacterium from the Dead Sea. *Int J Syst Bacteriol* **40**: 209–210.

Papp-Wallace KM, Maguire ME. (2008). Regulation of CorA Mg2+ channel function affects the virulence of *Salmonella enterica* serovar *typhimurium*. *J Bacteriol* **190**: 6509–6516.

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.

Rusch DB, Halpern AL, Heidelberg KB, Sutton G, Williamson SJ, Yooseph S *et al.* (2007). The sorcerer II global ocean sampling expedition: I, the northwest Atlantic through the eastern tropical Pacific. *PLoS Biol* **5**: e77.

Sabehi G, Béjà O. (2007). BAC libraries construction from marine microbial assemblages. In: Kowalchuk GA, de Bruijn FJ, Head IM, Akkermans AD, van Elsas JD (eds) *Molecular Microbial Ecology Manual*. Kluwer Academic Publishers: Dordrecht, The Netherlands, pp 1863–1879.

Shamir R, Maron-Katz A, Tanay A, Linhart C, Steinfeld I, Sharan R *et al.* (2005). EXPANDER-an integrative program suite for microarray data analysis. *BMC Bioinformatics* **6**: 232.

Sharon I, Pati A, Markowitz VM, Pinter RY. (2009). A statistical framework for the functional analysis of metagenomes. In: Batzoglou S (ed) *The 13th Annual International Conference on Research in Computational Molecular Biology (RECOMB)*. Springer: Berlin/Heidelberg; Tucson, AZ, pp 496–511.

Sogin ML, Morrison HG, Huber JA, Mark Welch D, Huse SM, Neal PR *et al.* (2006). Microbial diversity in the deep sea and the underexplored 'rare biosphere'. *Proc Natl Acad Sci USA* **103**: 12115–12120.

Tatusov RL, Koonin EV, Lipman DJ. (1997). A genomic perspective on protein families. *Science* **278**: 631–637.

Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV *et al.* (2003). The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* **4**: 41.

Worning P, Jensen LJ, Nelson KE, Brunak S, Usserya DW. (2000). Structural analysis of DNA sequence: evidence for lateral gene transfer in *Thermotoga maritima*. *Nucleic Acids Res* **28**: 706–709.

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