

World J Microbiol Biotechnol (2013) 29:2325–2334
DOI 10.1007/s11274-013-1399-9

ORIGINAL PAPER

A meta-analysis of the publicly available bacterial and archaeal sequence diversity in saline soils

Bin Ma · Jun Gong

Received: 19 April 2013 / Accepted: 3 June 2013 / Published online: 12 June 2013
© Springer Science+Business Media Dordrecht 2013

Abstract An integrated view of bacterial and archaeal diversity in saline soil habitats is essential for understanding the biological and ecological processes and exploiting potential of microbial resources from such environments. This study examined the collective bacterial and archaeal diversity in saline soils using a meta-analysis approach. All available 16S rDNA sequences recovered from saline soils were retrieved from publicly available databases and subjected to phylogenetic and statistical analyses. A total of 9,043 bacterial and 1,039 archaeal sequences, each longer than 250 bp, were examined. The bacterial sequences were assigned into 5,784 operational taxonomic units (OTUs, based on $\geq 97\%$ sequence identity), representing 24 known bacterial phyla, with Proteobacteria (44.9%), Actinobacteria (12.3%), Firmicutes (10.4%), Acidobacteria (9.0%), Bacteroidetes (6.8%), and Chloroflexi (5.9%) being predominant. *Lysobacter* (12.8%) was the dominant bacterial genus in saline soils, followed by *Sphingomonas* (4.5%), *Halomonas* (2.5%), and *Gemmatimonas* (2.5%). Archaeal sequences were assigned to 602 OTUs, primarily from the phyla Euryarchaeota (88.7%) and Crenarchaeota (11.3%).

Halorubrum and *Thermofilum* were the dominant archaeal genera in saline soils. Rarefaction analysis indicated that less than 25% of bacterial diversity, and approximately 50% of archaeal diversity, in saline soil habitats has been sampled. This analysis of the global bacterial and archaeal diversity in saline soil habitats can guide future studies to further examine the microbial diversity of saline soils.

Keywords Archaea · Bacteria · Halophilic · Microbial diversity · Saline soil

Introduction

Saline soils are increasingly abundant as a consequence of irrigation and desertification processes (Rengasamy 2006). Increase in salinity greatly affects soil microbial and biochemical properties, and causes shifts in the microbial community structure towards bacteria and archaea. According to a recent meta-analysis, the global microbial community composition is influenced more by salinity than by extremes of temperature, pH, or other physical and chemical factors (Lozupone and Knight 2007). Accordingly, the diversity of halophilic bacteria and archaea in saline soils is expected to be quite different from that in non-saline soils. The study of microbial diversity in saline soils is significant for understanding the ecological functions, saline adaption mechanisms and biotechnical potentials of microorganisms.

Halophilic and halotolerant bacteria and archaea are essential for the biogeochemical processes in saline soils (Dendooven et al. 2010; Blum et al. 2012), hence for the utilization of saline soils. Halophilic bacteria and archaea thrive in saline soils because they possess enzymes endowed with unique structural features and catalytic

Electronic supplementary material The online version of this article (doi:10.1007/s11274-013-1399-9) contains supplementary material, which is available to authorized users.

B. Ma · J. Gong (✉)
Laboratory of Environmental Microbiology, Yantai Institute of Coastal Zone Research, Chinese Academy of Sciences, Yantai, China
e-mail: jgong@yic.ac.cn

B. Ma · J. Gong
Network Laboratory for Coastal Microbiology, Yantai, China

power, enabling them to sustain metabolic and physiological processes under high salt conditions (Munawar and Engel 2012). Specific enzymes from soil halophiles are perceived to be potentially useful for a variety of applications, including the production of important biomolecules (Liszka et al. 2012) and the remediation of pollutants in saline conditions (Dastgheib et al. 2011). Soil halophiles are an excellent source for exploring novel enzymes possessing the inherent ability to function in high salt conditions (Singh et al. 2012).

Bacteria and archaea have been widely studied with respect to their biodiversity in saline soils. Initial studies employed traditional culture-dependent methods and resulted in the discovery of a large number of new bacterial and archaeal taxa (Lefebvre et al. 2006). The primary methods in current use are DNA-based molecular biology approaches such as cloning and sequencing of 16S rRNA genes, denaturing gradient gel electrophoresis (DGGE), terminal restriction fragment length polymorphism (T-RFLP) and quantitative PCR (Lefebvre et al. 2006; Cui et al. 2009; Blum et al. 2012). Cloning and sequencing of 16S rRNA genes has been generally favored over other methods because it allows the detection both of known and unknown microbes present in saline soils.

Our knowledge of the diversity of bacteria and archaea in saline and hypersaline soils is, however, still limited and fragmented. In most cases only a few soil samples of a single specific saline system, or a small number of cloned sequences or DGGE bands, have been investigated and analyzed. Some studies have focused only one particular microbial group such as Archaea (Valenzuela-Encinas et al. 2009; Youssef et al. 2012) or Actinobacteria (Singh et al. 2012). Numerous sequences recovered from saline soils are available from the GenBank, but many of these are largely overlooked because they are not published in the literature. Recently developed high-throughput sequencing technologies produces large datasets of sequence reads. However, unlike the longer sequences generated with Sanger sequencing, the length of these reads is too short to classify the sequences to genus level (Nelson et al. 2011). Currently, there is no consensus on the size or nature of the microbial diversity generally found in saline soils.

In this study, we performed a meta-analysis of all publicly available 16S rRNA gene sequences identified from saline and hypersaline soils in order to provide an updated overview of the bacterial and archaeal diversity in saline soil environments. Estimates of the current coverage of the halophilic bacterial and archaeal diversity already identified in saline soils were made, and particular gaps in the knowledge and understanding of the microbial populations present in saline soil habitats were identified.

Methods

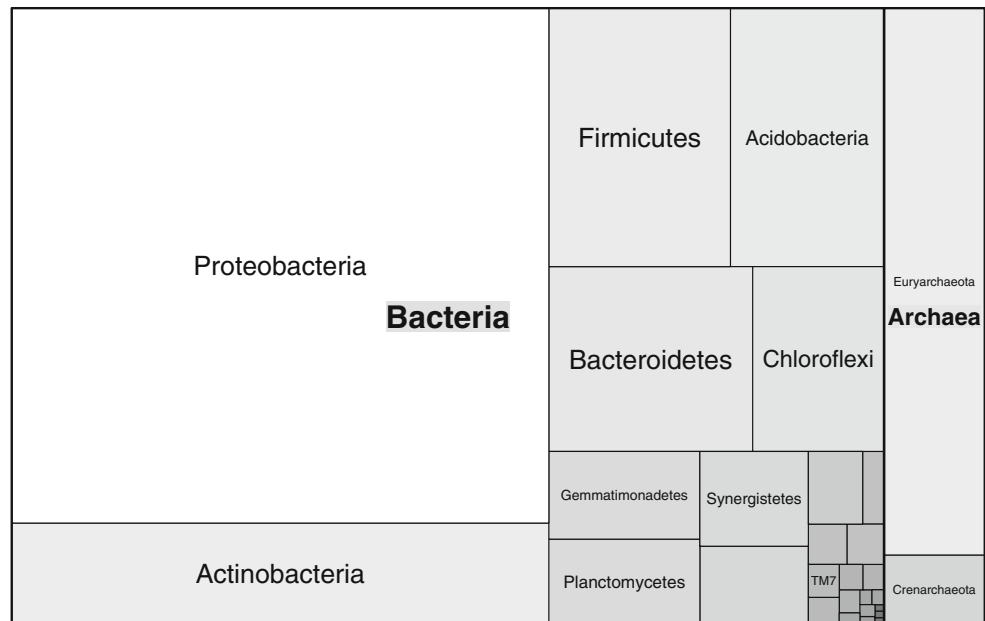
Sequence data collection

Initial sequence sets were obtained on 10 May 2012 from the GenBank (<http://www.ncbi.nlm.nih.gov>) and RDP (Release 10, <http://rdp.cme.msu.edu>) databases using the search terms ‘saline’ OR ‘hypersaline’ AND ‘soil’ AND ‘16S’. Non-16S rRNA sequences from GenBank were removed by checking the names of the sequences. All 16S rRNA gene sequences from the two databases were merged and duplicate sequences, identified based on accession numbers, were removed. Sequences with vector nucleotides or chimeras were checked by Mallard (<http://www.cf.ac.uk/biosi/research/biosoft/>). The 16S rRNA gene sequences of *Escherichia coli* (accession number: U00096) and *Methanothermobacter thermoautotrophicus* (accession number: AE000666) were selected as reference sequences for bacteria and archaea, respectively. Sequences shorter than 250 bp were removed from the dataset in order to avoid uncertainties when comparing and classifying short sequences that have little or no sequence overlap. The remaining sequences comprised the redacted composite dataset used in this study.

Phylogenetic analysis

Sequences were sorted by length and aligned with Kalign (Lassmann and Sonnhammer 2005). The sequences were classified into taxonomic ranks using the RDP Classifier with default settings (Cole et al. 2005). Treemaps based on the output classifications from the RDP Classifier were constructed using the treemap packages in R. Based on the classifications, the dataset was divided into the following groups: Archaea, Bacteria, Proteobacteria, Actinobacteria, Firmicutes, Acidobacteria, Bacteroidetes, Chloroflexi, and the collected “minor phyla” of bacteria that comprised sequences not assigned to any of the aforementioned phyla. Distances matrices of aligned sequences were computed within ARB using Jukes-Cantor correction (Cohen and Townsend 2009). Individual distance matrices were analyzed using Mothur (Schloss et al. 2009) to cluster OTUs, generate rarefaction curves, and determine the nonparametric ACE and Chao1 richness estimates. The genetic distance ≤ 0.03 was used to define species-level OTUs. The distance cut-offs were set at 0.05, 0.10, 0.15, and 0.2, which were assumed for taxonomic ranks genus, family, class/order and phylum, respectively. All the estimated asymptotes of the rarefaction curves were determined through R package monomol (<https://github.com/binma/monomol>). The coverage percentages were calculated as described by Nelson et al. (2011).

Fig. 1 Treemap of observed bacterial and archaeal taxa shown in their hierarchical order. Treemap showing taxonomic ranking of all taxa for all retrieved sequences. The size of *each box* is proportional to the number of sequences assigned to that taxon with respect to the entire dataset. The placement of *boxes* is arbitrary with respect to *boxes* within the same taxonomic rank and does not correspond to any form of phylogeny or relatedness



Accession numbers

The accession numbers for all sequences analyzed in this study are available on request from the corresponding author. The sequences are currently maintained in an in-house ARB database of 16S rRNA gene sequences for saline soils. A copy of this database and the sequence alignment are also available on request from the corresponding author.

Results

Data summary

A total of 10,082 sequences longer than 250 bp were retrieved from the two databases. Among these, only 3,342 sequences have been described in a total of 95 publications. The lengths of the remaining sequences were mostly about 1,500 bp and, less frequently, approximately 400 bp (Fig. S1). The 9,043 bacterial and 1,039 archaeal sequences were assigned to 5,784 and 602 OTUs, respectively (Fig. 1). The most abundant bacterial OTU contained 406 sequences, whereas the most abundant archaeal OTU contained 44 sequences. Based on assignment using the Ribosomal Database Project (RDP) Classifier, approximately 90 % of bacterial sequences were from six phyla, namely Proteobacteria, Actinobacteria, Firmicutes, Acidobacteria,

Bacteroidetes, and Chloroflexi. The others were from 18 other phyla, and 24 sequences were unclassified at the phylum level. For the archaea, all sequences were classified into two well-known phyla: nearly 89 % (922 sequences) belong to Euryarchaeota, and the remaining 11 % (117 sequences) belong to Crenarchaeota.

Bacteria

Proteobacteria

Proteobacteria appeared to be the most common bacterial taxon in saline soils. In our dataset, 4,647 sequences were derived from members of this phylum, accounting for about 46 % of all bacteria. There were 2,598 proteobacterial OTUs, with a Simpson diversity index of 0.0094. The majority (96.1 %) of the proteobacterial sequences were assigned to 326 known genera. Among the six proteobacterial classes represented in the dataset, Gammaproteobacteria were dominant, accounting for 53.7 % of the proteobacterial sequences; the proportions of alpha-, delta- and betaproteobacterial sequences were 29.3, 12.2, 4.6 %, respectively; Epsilon- and Zetaproteobacteria were represented by only 4 and 2 sequences, respectively (Fig. S2).

The gammaproteobacterial sequences represented 106 genera and 1,182 OTUs. Interestingly, 46.3 % (1,264 sequences) of gammaproteobacterial sequences representing 326 OTUs could be assigned to the genus *Lysobacter* of

the family *Xanthomonadaceae*. The second most abundant genus was *Halomonas* (family *Halomonadaceae*) with 269 sequences clustered into 146 OTUs. About 5.3 % of gammaproteobacterial sequences were unclassified at the genus level. There were 120 genera recognized from the 1361 alphaproteobacterial sequences, of which 404 (29.7 %) sequences representing 169 OTUs were from the genus *Sphingomonas* of the family *Sphingomonadaceae*, and 88 (6.4 %) sequences representing 35 OTUs were from the genus *Altererythrobacter* of the family *Erthrobacteraceae*. Approximately 2.6 % of alphaproteobacterial sequences were unclassified at the genus level. The 565 deltaproteobacterial sequences belonged to 444 OTUs representing 46 genera. Only one deltaproteobacterial sequence was unclassified at the genus level. The sequences belonging to Betaproteobacteria were assigned to 51 genera and 171 OTUs. Of the betaproteobacterial sequences, 10.2 and 9.7 % were assigned to the genera *Nitrosospira* and *Thiobacter*, respectively. Four Epsilonproteobacteria sequences were from the genera *Dehalospirillum*, *Sulfurovum*, and *Nitratiruptor*, and two zetaproteobacterial sequences were from the genus *Mariprofundus*.

Actinobacteria

The Actinobacteria was the phylum with the second largest number of representatives in our dataset with 913 sequences, 712 OTUs and a Simpson diversity index of 0.0012. About 53.5 % of the actinobacterial sequences belonged to the order Actinomycetales of the subclass Actinobacteridae, representing 99 genera and 382 OTUs (Fig S3). The most frequently observed families in the order Actinomycetales were *Micrococcaceae* (147 sequences), *Propionibacteraceae* (73 sequences), and *Streptomycetaceae* (59 sequences). The genus *Arthrobacter* (51 sequences) of the family *Micrococcaceae*, and the genus *Streptomyces* (45 sequences) of the family *Streptomycetaceae*, were the two best represented genera in the order Actinomycetales.

The numbers of sequences assigned to subclasses Acidimicrobidae, Nitrospirales, and Rubrobacteridae were 197, 131, and 96, respectively. The genera *Iamia* (81 sequences) and *Aciditerrimonas* (80 sequences) represented the most abundant taxa within the subclass Acidimicrobidae. The subclass Nitrospirales included 69 sequences of the genus *Nitrospirillum* and 62 sequences of the genus *Euzebya*. There were four genera of the subclass Rubrobacteridae with 79 OTUs, including 38 sequences of the genus *Rubrobacter*, 40 sequences of the genus *Conexibacter*, 16 sequences of the genus *Solirubrobacter*, and 2 sequences of the genus *Thermoleophilum*. One sequence, related to the genus *Asaccharobacter*, was the only representative of the subclass Coriobacteridae.

Firmicutes

The phylum with the third largest number of sequences (791) was Firmicutes. The number of identified OTUs was 599, with a Simpson diversity index of 0.0021. The majority of sequences were assigned to the classes Bacilli (72.1 % representing 406 OTUs) and Clostridia (27.7 % representing 191 OTUs) (Fig. S4). A total of 117 genera were recognized, including 60 belonging to the class Bacilli and 55 belonging to the class Clostridia. The 209 sequences of the genus *Bacillus* (26.4 % of Firmicutes sequences) appeared to be the most abundant within the phylum Firmicutes; no other genus accounted for more than 5 % of Firmicutes sequences. In the class Clostridia, 44 sequences were identified as *Thermanaeromonas*.

Acidobacteria

There were 664 sequences assigned into phylum Acidobacteria, representing 522 OTUs and with a Simpson diversity index of 0.0014. Apart from four sequences representing the class Holophagae, the majority of acidobacterial sequences were classified into 14 groups (Fig. S5). The most abundant was group Gp6 (69.2 %) with 336 OTUs followed by group Gp4 (21.8 %) with 119 OTUs.

Bacteroidetes

The dataset for the phylum Bacteroidetes included 634 sequences, with 393 OTUs and a Simpson diversity index of 0.0027. Although 17 sequences were unclassified at the class level, three classes were represented, namely Sphingobacteria (75.1 %) with 262 OTUs, Flavobacteria (16.7 %) with 69 OTUs, and Bacteroidia (8.9 %) with 45 OTUs (Fig. S6). There were 29 genera belonging to the class Sphingobacteria, the most abundant of which were *Gracilimonas* (25.9 %) and *Salinibacter* (23.0 %). There were far fewer sequences of genera such as *Salinimicrobium* (3.6 %) and *Owenweeksia* (2.8 %) in the class Flavobacteria and *Anaerophaga* (3.3 %) in the class Bacteroidia.

Chloroflexi

The 407 sequences belonging to phylum Chloroflexi were clustered into 339 OTUs, with a Simpson diversity index of 0.002. Over 60 % (248 sequences) of the sequences in this phylum were from the class Anaerolineae (Fig. S7), the two genera that were best represented being *Longilinea* (30.5 %) and *Bellilinea* (19.4 %). The second predominant

class was *Caldilineae*, all sequences of which were assigned to the genus *Caldilinea*.

Minor phyla

In addition to the six phyla described above, 18 phyla represented by 987 sequences were observed in the dataset. Of these 18 minor phyla, only Gemmatimonadetes (2.45 %), Planctomycetes (2.38 %), Synergistetes (1.90 %), and Cyanobacteria (1.57 %) represented more than 1 % of all the bacterial sequences. These four accounted for 78 % of all minor phyla sequences.

All 222 sequences of the phylum Gemmatimonadetes were assigned to the genus *Gemmatimonas*. The 215 sequences of the phylum Planctomycetes represented six genera and 161 OTUs, among which *Phycisphaera* with 71 OTUs (45.6 %) and *Planctomyces* with 25 OTUs (26.1 %) were dominant. The phylum Synergistetes (171 sequences) was the third most abundant among the minor phyla, and the majority of sequences (63.0 %) in this phylum were assigned to genus *Pyramidobacter*. Of the 142 cyanobacterial sequences, 42.7 and 23.1 % were associated with Cyanobacteria groups GpVII and GpIX, respectively.

Archaea

Euryarchaeota

The phylum Euryarchaeota accounted for 88.7 % of archaeal sequences that were clustered into 568 OTUs, with a Simpson diversity index of 0.0022. The majority (72.5 %) of euryarchaeotal sequences were assigned to the class Halobacteria with 361 OTUs and 33 genera (Fig. S8). *Halorubrum* of the class Halobacteria accounted for more halobacterial sequences (10.1 %) than any other genera. The class Methanomicrobia was the second most abundant class in the phylum Euryarchaeota with 156 sequences clustered into 122 OTUs. There were 80 and 75 sequences assigned to the orders Methanomicrobiales and Methanosarcinales, respectively. The genus *Methanolacinia* of the order Methanomicrobiales represented 48 sequences. The genus *Methanosaeta* of the order Methanosarcinales had 62 sequences. The third most abundant class was Thermoplasmata, which was clustered into 48 OTUs. Apart from one sequence of the genus *Ferroplasma*, all other sequences of the class Thermoplasmata (58) were assigned to the genus *Thermogymnomonas*. There were 35 sequences belonging to the class Methanobacteria, including 17 sequences of the genus *Methanobacterium*, 16 sequences of the genus *Methanospaera*, and 2 sequences of the genus *Methanothermobacter*.

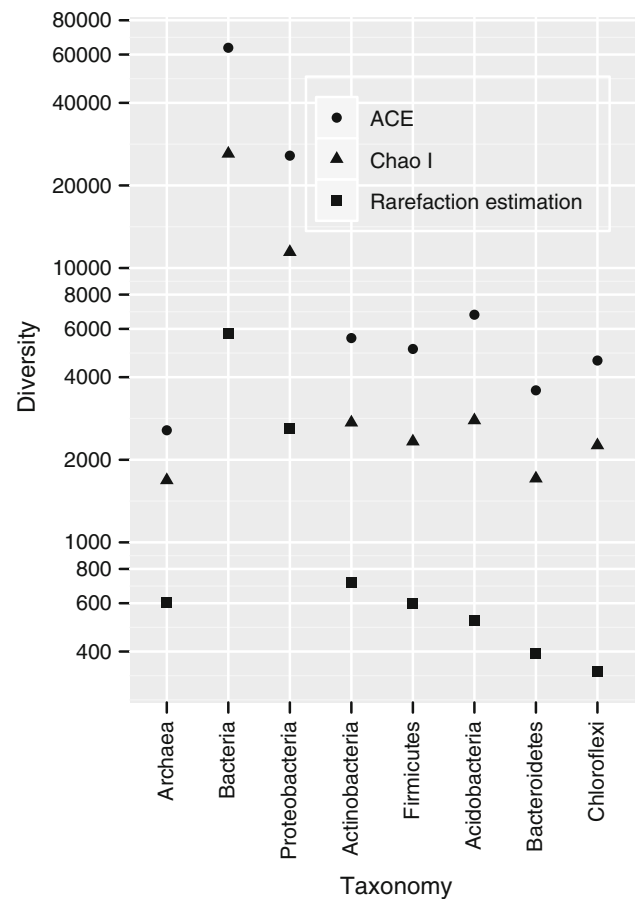


Fig. 2 Diversity statistics for Archaea, Bacteria, and the major phyla. Values were calculated using a 0.03 dissimilarity cut-off

Crenarchaeota

There were 117 sequences (11.3 % of archaeal sequences) assigned to the phylum Crenarchaeota, which were clustered into 34 OTUs, with a Simpson diversity index of 0.0019. All the crenarchaeotal sequences were assigned to the class Thermoprotei, including 92 belonging to the order Thermoproteales, 16 to the order Feravidicoccales, 5 to the order Desulfurococcales, and 4 to the order Acidilobales (Fig. S9). Of the 92 Thermoproteales sequences, 90 represented the genus *Thermofilum*.

Diversity estimates

Bacterial diversity

All the abundance-based coverage estimation (ACE) and Chao1 values were greater than the corresponding asymptotic estimates based on the monomolecular model (monomol). Like the monomol estimates, the ACE and Chao1 also estimate the maximum species richness for an OTU definition. However, the values of Chao1 and ACE

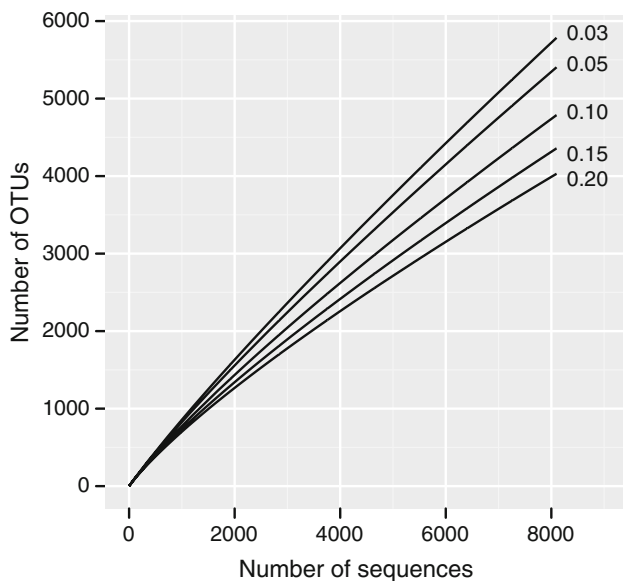


Fig. 3 Rarefaction curve for the Bacteria with different dissimilarity cut-offs

for OTUs were five and ten times greater than monomol estimates, respectively (Fig. 2). The values of Chao1 and ACE of OTUs for Proteobacteria, Actinobacteria, Firmicutes, and Bacteroidetes were five and ten times, respectively, greater than monomol estimates, but six and fourteen times, respectively, greater than rarefaction estimates for Acidobacteria and Chloroflexi.

These findings indicate that compared to the total level of bacterial diversity in saline soils, the coverage of present dataset was extremely low. Rarefaction analysis of Bacteria showed that for all taxonomic ranks, the rarefaction curves were all projecting upward with none reaching a horizontal asymptote (Fig. 3). Among all taxonomic ranks of Bacteria, it was estimated that less than 25 % of the diversity was recovered. The coverage rates of Proteobacteria were less at species and genus levels, but greater at higher taxonomic ranks, than that of the bacterial dataset as a whole (Fig. 4). The coverage rates of Acidobacteria and Bacteroidetes at phylum level were lower than at class level. The coverage rates of Chloroflexi at species and genera levels were significantly lower than those of other phyla.

Archaeal diversity

The coverage of archaeal diversity was significantly greater than bacteria, but still low compared to estimated richness. The estimated ACE and Chao1 values were four and three times greater respectively than the values of monomol estimates (Fig. 2). The discrepancy between the different diversity estimates for Archaea was less than that of

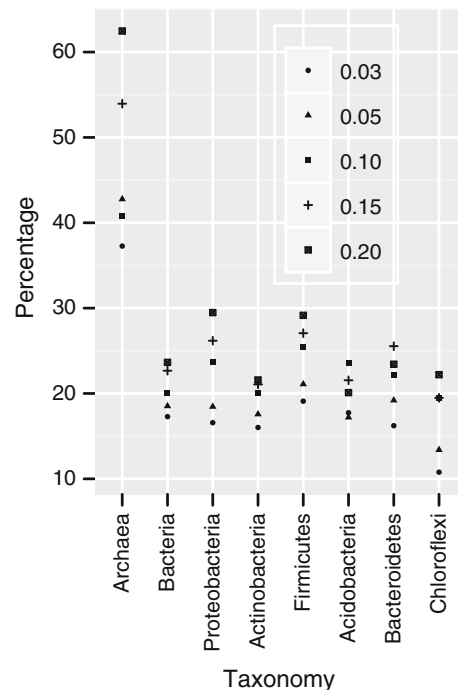


Fig. 4 Coverage of OTUs for Archaea, Bacteria, and the major phyla. The coverage was the percentage of OTUs in monomol estimation with 0.03 dissimilarity cut-off

Bacteria. Like the bacteria, the rarefaction curves for all taxonomic levels of archaeans failed to reach to the horizontal asymptote, but the coverage of Archaea was greater than 50 % both at class and phyla levels, and was approximately 40 % at lower taxonomic levels (Fig. 5).

Discussion

Saline soils occur in all continents and under almost all climatic conditions. The nature and properties of these soils are also diverse such that they require specific approaches for their reclamation and management in order to maintain their long term productivity. Bacteria and archaea play essential roles in biogeochemical processes in saline soils. However, a synthesis of information on halophilic bacteria and archaea in saline soils is lacking.

This meta-analysis study was based on publicly available 16S rRNA gene sequences recovered from saline soil habitats. The sequence data obtained from the Genbank and RDP databases were analyzed regardless of their previously assigned taxonomic information. Saline soils have received increasing attention because of the shortage of arable land and the demands for ecological restoration. It is therefore necessary to gain a better understanding of the diversity and ecological function of microorganisms in

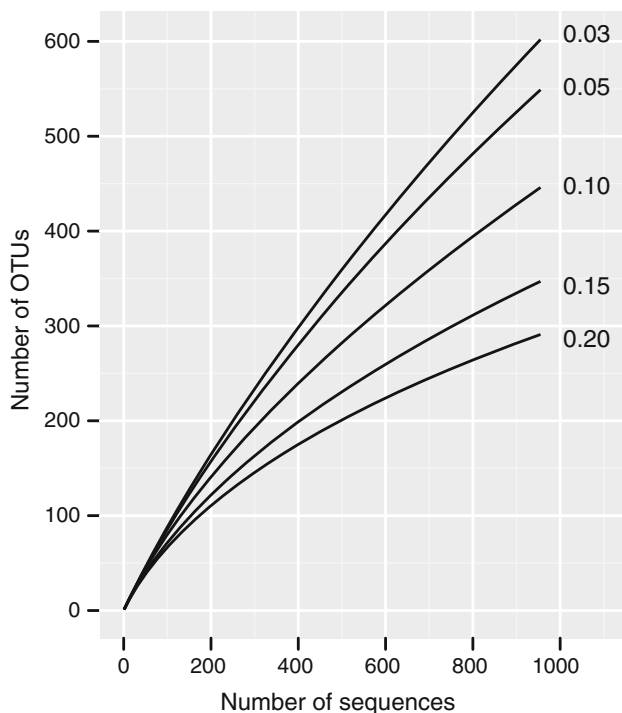


Fig. 5 Rarefaction curve for the Archaea with different dissimilarity cut-offs

saline soil habitats. These results represent a global overview of the microbial diversity of saline soils.

Bacteria

The present dataset, which was generated from the Genbank and RDP databases, was largely dominated by Proteobacteria with more than half of bacterial sequences belonging to this group. The dominant class within the Proteobacteria was Gammaproteobacteria, followed by Alphaproteobacteria and Deltaproteobacteria. In studies of soils from the former lake Texcoco, Valenzuela-Encinas et al. (2009) showed that the dominant class of Proteobacteria in both high (158 dS m^{-1}) and low (0.65 dS m^{-1}) saline soils was Gammaproteobacteria, whereas in medium (56 dS m^{-1}) saline soils the dominant class was Alphaproteobacteria. Furthermore, Deltaproteobacteria was present in both high and low saline soils, but absent in medium saline soils (Valenzuela-Encinas et al. 2009).

The dominant genera in the phylum Proteobacteria from the present dataset have been widely detected in saline habitats. The genus *Lysobacter*, to which approximately half the sequences of the class Gammaproteobacteria belonged, has previously been reported as the most abundant in hyper-saline and hyper-alkaline environments (Brito-Echeverria et al. 2011). The genus *Sphingomonas* (class Alphaproteobacteria) also appears to be widely distributed in saline soils (Farias et al. 2011).

Desulfovermiculus halophilus (class Deltaproteobacteria) is an extremely halophilic bacterium and is capable of growing at a salt concentration of 346 g L^{-1} and pH 9.8 (Beliakova et al. 2006). Members of the genus *Nitrosospira* (class Betaproteobacteria) were reported to be representative ammonia-oxidizing bacteria (AOB) in saline soil habitats (Mohamed et al. 2010).

Approximately 10 % of sequences in the present dataset belonged to the phylum Actinobacteria and it has previously been noted that many of the halophilic bacteria isolated from saline soils belong to this phylum (Sardinha et al. 2003; Valenzuela-Encinas et al. 2012). In one study, however, it was found that although Actinobacteria was represented in both medium and low saline soils, it was absent in extremely high saline soils (Valenzuela-Encinas et al. 2009). The actinobacterial sequences were assigned to the subclasses Actinobacteridae, Acidimicrobidae, Nitriliruptoridae, and Rubrobacteridae. More than half of actinobacterial sequences belonged to the order Actinomycetes, which is characterized by its mycelial growth, and with specific mechanisms of adapting to extreme environments, particularly saline and alkaline habitats (Singh et al. 2012). Actinomycetes can grow on media containing 3.5 % NaCl (Maldonado et al. 2005). Less than 20 % of sequences belonging to Actinomycetes could be assigned to known genera, which probably reflects the lack of previous studies of such organisms from saline soils. The sequences assigned to cultivable Actinomycetes in the present dataset belonged to two genera, *Streptomyces* and *Arthrobacter*. Members of the genus *Arthrobacter* have previously been reported as halotolerant bacteria from saline habitats (Plotnikova et al. 2011). The genus *Iamia* of the order Acidimicrobidae, which was the second most frequently encountered order in the phylum Actinobacteria in the present study, exists widely in saline habitats such as seawater and coastal environments (Plotnikova et al. 2011; Tkavc et al. 2011). Another dominant genus *Aciditerrimonas* (order Acidimicrobidae) was isolated from a sulfataric field, where the salinity is usually higher than in other habitats. In the present dataset, the dominant genera of the order Nitriliruptoridae were *Nitriliruptor* and *Euzebya*. *Nitriliruptor* is a deep-branching haloalkaliphilic actinobacterium isolated from soda lakes and is capable of growth on aliphatic nitriles (Sorokin et al. 2009). The genus *Euzebya* is a deep-branching marine actinobacterium isolated from sea cucumbers (Kurahashi et al. 2010). Consequently, *Euzebya* is expected to adapt to saline habitats. The genus *Rubrobacter* (order Rubrobacteridae) is capable in bioremediating organic and radioactive pollutants, but it is rarely reported in saline habitats (Jurado et al. 2012). The genera *Conexibacter* and *Solirubrobacter* (order Rubrobacteridae) have been isolated from cultivated field soils, but seldom from saline soils (An et al. 2011; Seki et al. 2012).

Firmicutes, the third most abundant phylum in the present dataset, accounted for 8.8 % of all sequences. Among the sequences assigned to Firmicutes, the dominant genera were *Bacillus* and *Thermanaeromonas*. A huge number of bacteria belonging to the genus *Bacillus* have been isolated from saline and alkaline soil habitats (Ren and Zhou 2005). Isolated strains of halophilic *Bacillus* are an important resource for exploring halophilic enzymes and for remediating pollutants in saline environment (Liszka et al. 2012).

Members of the phylum Acidobacteria are physiologically diverse and ubiquitous in soils, but are under-represented in culture-based studies (Barns et al. 2007). In the present study, Acidobacteria was the fourth most abundant phylum, but most of the acidobacterial sequences were unclassified even at class level. Valenzuela-Encinas et al. (2009) showed that the Acidobacteria was the dominant phylum in medium-saline soils but absent in high and low saline soils. Acidobacteria was found to represent significant proportions of bacterial communities in saline sediments from Qinghai Lake (Dong et al. 2006) and mangrove sediments of Sundarban (Ghosh et al. 2010).

The phylum Bacteroidetes was the fifth most abundant phylum in the present dataset. The dominant genera included *Gracilimonas*, *Salinibacter*, *Salisaeta*, *Salinicrobium*, *Owenweeksia*, and *Anaerophaga*. The genus *Gracilimonas* has not previously been reported in saline environments, but the related form *Fodinibius salinus* is a moderately halophilic bacterium isolated from a salt mine (Wang et al. 2012). The majority of bacteria in the genera *Salinibacter* and *Salisaeta* have been detected in hypersaline environments (Vaisman and Oren 2009; Brito-Echeverria et al. 2011). Likewise, members of the genus *Salimicrobium*, are known from marine solar salterns (Yoon et al. 2009) and those of *Owenweeksia* are generally isolated from marine habitats (Lau et al. 2005). The genus *Anaerophaga* is a recently reported genus comprising moderately thermohalophilic, strictly anaerobic, fermentative bacteria (Gao et al. 2011).

Members of the phylum Chloroflexi are generally found in hypersaline environments, i.e., high saline soils and hypersaline wastewater. Most of sequences in the present dataset belonging to Chloroflexi were assigned to the class Anaerolineae, members of which are frequently detected in marine sediments. In the present study, the dominant genera of Anaerolineae were *Longilinea* and *Bellilinea*. However, the bacteria in these two genera have rarely been found in saline habitats. The predominant genus *Caldilinea* in the class Caldilineae has been found in hot springs and activated sludge, where the salinities are expected to be high (Kragelund et al. 2011). Two genera *Sphaerobactereinae* and *Thermomicrobium* (class Thermomicrobia)

have also been found in extremely thermal soils (Morino and Ito 2012).

Some genera of minor phyla were found in abundance in the present dataset. The genus *Gemmatimonas* (phylum Gemmatimonadetes) is known for having members with active roles in biogeochemical transformations in hypersaline soils (Zhang et al. 2003). The genera *Phycisphaera* and *Planctomyces* (phylum Planctomycetes) are symbionts of marine algae or sponges (Fukunaga et al. 2009; Jogler et al. 2011). Although more than 100 sequences were assigned to the genus *Pyramidobacter* in the present study, this genus has not previously been reported in saline habitats.

Archaea

More than 60 % of archaeal sequences were assigned to the class Halobacteria (phylum Eurarchaeota). Members of the class Halobacteria are extremely halophilic and haloalkaliphilic (Brito-Echeverria et al. 2011; Youssef et al. 2012). Members of the genus *Halorubrum* are ubiquitous halophilic, aerobic archaeans that inhabit saline and hypersaline soils (Papke et al. 2004). *Natronomonas*, has also been widely detected in alkaline-saline soil habitats (Valenzuela-Encinas et al. 2008). These are the two most commonly encountered genera of Halobacteria in the present study. The growth kinetics of *Natronomonas* has highly unusual temperature requirements for growth, suggesting that the physiology of *Natronomonas* is different with other halobacteria (Inoue et al. 2011). The members of the genus *Halomicrobium*, i.e., *Halomicrobium zhouii*, *H. katesii*, and *H. mukohataei*, are also halophilic (Cui et al. 2009). The genus *Halococcus* has been found in the rhizosphere of the halophyte *Halonemum strobilaceum* inhabiting hypersaline coastal areas of the Arabian Gulf, and is known to be an extremely halophilic, oil-utilizing microorganism (Al-Mailem et al. 2010). The genus *Thermofilum* (phylum Crenarchaeota) is generally detected from hyperthermophilic habitats, which are also saline or hypersaline environments (Flores et al. 2012).

Data analyses

One important issue in the present study is that the composite dataset cannot currently address the beta diversity of species in soils with different saline levels. The bacterial and archaeal communities vary in soils with different salt concentrations (Valenzuela-Encinas et al. 2009). Although the dataset contains sequences recovered from a wide array of studies, the beta diversity could not be determined because some studies submitted sequences to databases without published information on the soil properties.

Analysis of a composite dataset carries a few potential caveats and limitations that bear mentioning. The size of datasets generated from different studies varied over a large range. The studies with more sequences in a dataset have the effect of artificially inflating the representation of soils with some salt concentration compared to others. Furthermore, some studies focused on one or a few phylogenetic groups, again skewing the composite dataset by increasing sequences from these groups. For bacterial communities, haloalkaliphilic Actinobacteria have been widely studied (Siddikee et al. 2011). As such, while the global diversity defined by the dataset can represent the types of bacteria and archaea generally present in saline soils, the abundance of some OTUs may not directly reflect the actual abundance of the corresponding organism found in any given saline soils.

Acknowledgments The work was financially supported by the Main Direction Program of Knowledge Innovation of CAS (grant no. KSCX2-EW-G-12B), the Double Hundred Talent Plan of Yantai, and the One Hundred Talent Program of CAS awarded to JG. Thanks are due to Dr. Alan Warren for proofreading.

References

- Al-Mailem DM, Sorkhoh NA, Marafie M, Al-Awadhi H, Eliyas M, Radwan SS (2010) Oil phytoremediation potential of hypersaline coasts of the Arabian Gulf using rhizosphere technology. *Bioresour Technol* 101:5786–5792
- An DS, Wang L, Kim MS, Bae HM, Lee ST, Im WT (2011) *Solirubrobacter ginsenosidimitans* sp. nov., isolated from soil of a ginseng field. *Int J Syst Evol Microbiol* 61:2606–2609
- Barns SM, Cain EC, Sommerville L, Kuske CR (2007) Acidobacteria phylum sequences in uranium-contaminated subsurface sediments greatly expand the known diversity within the phylum. *Appl Environ Microbiol* 73:3113–3116
- Beliakova EV, Rozanova EP, Borzenkov IA, Turova TP, Pusheva MA, Lysenko AM, Kolganov TV (2006) The new facultatively chemolithoautotrophic, moderately halophilic, sulfate-reducing bacterium *Desulfovermicillus halophilus* gen. nov., sp. nov., isolated from an oil field. *Microbiology* 75:201–211
- Blum JS, Kulp TR, Han S, Lanoil B, Saltikov CW, Stolz JF, Miller LG, Oremland RS (2012) *Desulfohalophilus alkaliarsenatis* gen. nov., sp. nov., an extremely halophilic sulfate- and arsenate-respiring bacterium from Searles Lake, California. *Extremophiles* 16:727–742
- Brito-Echeverria J, Lucio M, Lopez-Lopez A, Anton J, Schmitt-Kopplin P, Rossello-Mora R (2011) Response to adverse conditions in two strains of the extremely halophilic species *Salinibacter ruber*. *Extremophiles* 15:379–389
- Cohen DL, Townsend RR (2009) Is there added value to adding ARB to ACE inhibitors in the management of CKD? *J Am Soc Nephrol* 20:1666–1668
- Cole JR, Chai B, Farris RJ, Wang Q, Kulam SA, McGarrell DM, Garrity GM, Tiedje JM (2005) The Ribosomal Database Project (RDP-II): sequences and tools for high-throughput rRNA analysis. *Nucleic Acids Res* 33:D294–D296
- Cui HL, Zhou PJ, Oren A, Liu SJ (2009) Intraspecific polymorphism of 16S rRNA genes in two halophilic archaeal genera, *Haloarcula* and *Halomicrobium*. *Extremophiles* 13:31–37
- Dastgheib SM, Amoozgar MA, Khajeh K, Ventosa A (2011) A halotolerant *Alcanivorax* sp. strain with potential application in saline soil remediation. *Appl Microbiol Biotechnol* 90:305–312
- Dendooven L, Alcantara-Hernandez RJ, Valenzuela-Encinas C, Luna-Guido M, Perez-Guevara F, Marsch R (2010) Dynamics of carbon and nitrogen in an extreme alkaline saline soil: a review. *Soil Biol Biochem* 42:865–877
- Dong H, Zhang G, Jiang H, Yu B, Chapman LR, Lucas CR, Fields MW (2006) Microbial diversity in sediments of saline Qinghai Lake, China: linking geochemical controls to microbial ecology. *Microb Ecol* 51:65–82
- Farias ME, Revale S, Mancini E, Ordonez O, Turjanski A, Cortez N, Vazquez MP (2011) Genome sequence of *Sphingomonas* sp. S17, isolated from an alkaline, hyperarsenic, and hypersaline volcano-associated lake at high altitude in the Argentinean Puna. *J Bacteriol* 193:3686–3687
- Flores GE, Shakya M, Meneghin J, Yang ZK, Seewald JS, Geoff Wheat C, Podar M, Reysenbach AL (2012) Inter-field variability in the microbial communities of hydrothermal vent deposits from a back-arc basin. *Geobiology* 10:333–346
- Fukunaga Y, Kurahashi M, Sakiyama Y, Ohuchi M, Yokota A, Harayama S (2009) *Phycisphaera mikurensis* gen. nov., sp. nov., isolated from a marine alga, and proposal of *Phycisphaeraceae* fam. nov., *Phycisphaerales* ord. nov. and *Phycisphaerae* classis nov. in the phylum Planctomycetes. *J Gen Appl Microbiol* 55:267–275
- Gao Z, Liu X, Ruan L (2011) Genome sequence of *Anaerophaga* sp. strain HS1, a novel, moderately thermophilic, strictly anaerobic bacterium isolated from hot spring sediment. *J Bacteriol* 193:5572
- Ghosh A, Dey N, Bera A, Tiwari A, Sathyaniranjan K, Chakrabarti K, Chattopadhyay D (2010) Culture independent molecular analysis of bacterial communities in the mangrove sediment of Sundarban, India. *Saline Syst* 6:1
- Inoue K, Itoh T, Ohkuma M, Kogure K (2011) *Halomarina oriensis* gen. nov., sp. nov., a halophilic archaeon isolated from a seawater aquarium. *Int J Syst Evol Microbiol* 61:942–946
- Jogler C, Glockner FO, Kolter R (2011) Characterization of *Planctomyces limnophilus* and development of genetic tools for its manipulation establish it as a model species for the phylum Planctomycetes. *Appl Environ Microbiol* 77:5826–5829
- Jurado V, Miller AZ, Alias-Villegas C, Laiz L, Saiz-Jimenez C (2012) *Rubrobacter bracarensis* sp. nov., a novel member of the genus *Rubrobacter* isolated from a biodeteriorated monument. *Syst Appl Microbiol* 35:306–309
- Kragelund C, Thomsen TR, Mielczarek AT, Nielsen PH (2011) Eikelboom's morphotype 0803 in activated sludge belongs to the genus *Caldilinea* in the phylum Chloroflexi. *FEMS Microbiol Ecol* 76:451–462
- Kurahashi M, Fukunaga Y, Sakiyama Y, Harayama S, Yokota A (2010) *Euzebya tangerina* gen. nov., sp. nov., a deeply branching marine actinobacterium isolated from the sea cucumber *Holothuria edulis*, and proposal of *Euzebyaceae* fam. nov., *Euzebyales* ord. nov. and *Nitriliruptoridae* subclassis nov. *Int J Syst Evol Microbiol* 60:2314–2319
- Lassmann T, Sonnhammer EL (2005) Kalign—an accurate and fast multiple sequence alignment algorithm. *BMC Bioinf* 6:298
- Lau KW, Ng CY, Ren J, Lau SC, Qian PY, Wong PK, Lau TC, Wu M (2005) *Owenweeksia hongkongensis* gen. nov., sp. nov., a novel marine bacterium of the phylum ‘Bacteroidetes’. *Int J Syst Evol Microbiol* 55:1051–1057
- Lefebvre O, Vasudevan N, Thanasekaran K, Moletta R, Godon JJ (2006) Microbial diversity in hypersaline wastewater: the example of tanneries. *Extremophiles* 10:505–513

- Liszka M, Clark M, Schneider E, Clark DS (2012) Nature versus nurture: developing enzymes that function under extreme conditions. *Ann Rev Chem Biomol Eng* 3:77–102
- Lozupone CA, Knight R (2007) Global patterns in bacterial diversity. *Proc Natl Acad Sci USA* 104:11436–11440
- Maldonado LA, Fenical W, Jensen PR, Kauffman CA, Mincer TJ, Ward AC, Bull AT, Goodfellow M (2005) *Salinispora arenicola* gen. nov., sp. nov. and *Salinispora tropica* sp. nov., obligate marine actinomycetes belonging to the family *Micromonosporaceae*. *Int J Syst Evol Microbiol* 55:1759–1766
- Mohamed NM, Saito K, Tal Y, Hill RT (2010) Diversity of aerobic and anaerobic ammonia-oxidizing bacteria in marine sponges. *ISME J* 4:38–48
- Morino M, Ito M (2012) Functional expression of the multi-subunit type calcium/proton antiporter from *Thermomicrobium roseum*. *FEMS Microbiol Lett* 335:26–30
- Munawar N, Engel PC (2012) Overexpression in a non-native halophilic host and biotechnological potential of NAD⁺-dependent glutamate dehydrogenase from *Halobacterium salinarum* strain NRC-36014. *Extremophiles* 16:463–476
- Nelson MC, Morrison M, Yu ZT (2011) A meta-analysis of the microbial diversity observed in anaerobic digesters. *Bioresour Technol* 102:3730–3739
- Papke RT, Koenig JE, Rodriguez-Valera F, Doolittle WF (2004) Frequent recombination in a saltern population of *Halorubrum*. *Science* 306:1928–1929
- Plotnikova EG, Yastrebova OV, Anan'ina LN, Dorofeeva LV, Lysanskaya VY, Demakov VA (2011) Halotolerant bacteria of the genus *Arthrobacter* degrading polycyclic aromatic hydrocarbons. *Russ J Ecol* 42:502–509
- Ren PG, Zhou PJ (2005) *Tenuibacillus multivorans* gen. nov., sp. nov., a moderately halophilic bacterium isolated from saline soil in Xin-Jiang, China. *Int J Syst Evol Microbiol* 55:95–99
- Rengasamy P (2006) World salinization with emphasis on Australia. *J Exp Bot* 57:1017–1023
- Sardinha M, Muller T, Schmeisky H, Joergensen RG (2003) Microbial performance in soils along a salinity gradient under acidic conditions. *Appl Soil Ecol* 23:237–244
- Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 75:7537–7541
- Seki T, Matsumoto A, Shimada R, Inahashi Y, Omura S, Takahashi Y (2012) *Conexibacter arvalis* sp. nov., isolated from a cultivated field soil sample. *Int J Syst Evol Microbiol* 62:2400–2404
- Siddikee MA, Tipayno SC, Kim K, Chung JB, Sa T (2011) Influence of varying degree of salinity-sodicity stress on enzyme activities and bacterial populations of coastal soils of Yellow Sea, South Korea. *J Microbiol Biotechnol* 21:341–346
- Singh SP, Raval VH, Purohit MK, Thumar JT, Gohel SD, Pandey S, Akbari VG, Rawal CM (2012) Haloalkaliphilic bacteria and actinobacteria from the saline habitats: new opportunities for biocatalysis and bioremediation. In: Satyanarayana T, Johri B, Anil P (eds) *Microorganisms in environmental management*. Springer, Netherlands, pp 415–429
- Sorokin DY, van Pelt S, Tourova TP, Evtushenko LI (2009) *Nitriliruptor alkaliphilus* gen. nov., sp. nov., a deep-lineage haloalkaliphilic actinobacterium from soda lakes capable of growth on aliphatic nitriles, and proposal of *Nitriliruptoraceae* fam. nov. and *Nitriliruptorales* ord. nov. *Int J Syst Evol Microbiol* 59:248–253
- Tkavc R, Gostincar C, Turk M, Visscher PT, Oren A, Gunde-Cimerman N (2011) Bacterial communities in the 'petola' microbial mat from the Secovlje salterns (Slovenia). *FEMS Microbiol Ecol* 75:48–62
- Vaisman N, Oren A (2009) *Salisaeta longa* gen. nov., sp. nov., a red, halophilic member of the Bacteroidetes. *Int J Syst Evol Microbiol* 59:2571–2574
- Valenzuela-Encinas C, Neria-Gonzalez I, Alcantara-Hernandez RJ, Enriquez-Aragon JA, Estrada-Alvarado I, Hernandez-Rodriguez C, Dendooven L, Marsch R (2008) Phylogenetic analysis of the archaeal community in an alkaline-saline soil of the former lake Texcoco (Mexico). *Extremophiles* 12:247–254
- Valenzuela-Encinas C, Neria-Gonzalez I, Alcantara-Hernandez RJ, Estrada-Alvarado I, Zavala-Diaz de la Serna FJ, Dendooven L, Marsch R (2009) Changes in the bacterial populations of the highly alkaline saline soil of the former lake Texcoco (Mexico) following flooding. *Extremophiles* 13:609–621
- Valenzuela-Encinas C, Alcantara-Hernandez RJ, Estrada-Alvarado I, de la Serna FJZD, Dendooven L, Marsch R (2012) The archaeal diversity and population in a drained alkaline saline soil of the former Lake Texcoco (Mexico). *Geomicrobiol J* 29:18–22
- Wang YX, Liu JH, Xiao W, Zhang XX, Li YQ, Lai YH, Ji KY, Wen ML, Cui XL (2012) *Fodinibius salinus* gen. nov., sp. nov., a moderately halophilic bacterium isolated from a salt mine. *Int J Syst Evol Microbiol* 62:390–396
- Yoon JH, Kang SJ, Oh KH, Oh TK (2009) *Salimicrobium flavidum* sp. nov., isolated from a marine solar saltern. *Int J Syst Evol Microbiol* 59:2839–2842
- Youssef NH, Ashlock-Savage KN, Elshahed MS (2012) Phylogenetic diversities and community structure of members of the extremely halophilic Archaea (order Halobacteriales) in multiple saline sediment habitats. *Appl Environ Microbiol* 78:1332–1344
- Zhang H, Sekiguchi Y, Hanada S, Hugenholtz P, Kim H, Kamagata Y, Nakamura K (2003) *Gemmatimonas aurantiaca* gen. nov., sp. nov., a gram-negative, aerobic, polyphosphate-accumulating micro-organism, the first cultured representative of the new bacterial phylum Gemmatimonadetes phyl. nov. *Int J Syst Evol Microbiol* 53:1155–1163