Journal of Microbiology (2015) Vol. 53, No. 10, pp. 678–685
DOI 10.1007/s12275-015-5080-x
Copyright © 2015, The Microbiological Society of Korea

Illumina-based analysis of bacterial diversity related to halophytes *Salicornia europaea* and *Sueada aralocaspica*

Ying-wu Shi[1], Kai Lou[1], Chun Li[1], Lei Wang[2], Zhen-yong Zhao[2], Shuai Zhao[2], and Chang-yan Tian[2*]

[1]*Institute of Microbiology, Xinjiang Academy of Agricultural Sciences, Xinjiang Laboratory of Special Environmental Microbiology, Urumqi 830091, Xinjiang, P. R. China*
[2]*State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, Xinjiang, P. R. China*

(Received Feb 11, 2015 / Revised Jun 17, 2015 / Accepted Aug 10, 2015)

We used Illumina-based 16S rRNA V3 amplicon pyrosequencing to investigate the community structure of soil bacteria from the rhizosphere surrounding *Salicornia europaea,* and endophytic bacteria living in *Salicornia europaea* plants and *Sueada aralocaspica* seeds growing at the Fukang Desert Ecosystem Observation and Experimental Station (FDEOES) in Xinjiang Province, China, using an Illumina genome analyzer. A total of 89.23 M effective sequences of the 16S rRNA gene V3 region were obtained from the two halophyte species. These sequences revealed a number of operational taxonomic units (OTUs) in the halophytes. There were between 22–2,206 OTUs in the halophyte plant sample, at the 3% cutoff level, and a sequencing depth of 30,000 sequences. We identified 25 different phyla, 39 classes and 141 genera from the resulting 134,435 sequences. The most dominant phylum in all the samples was Proteobacteria (41.61%–99.26%; average, 43.30%). The other large phyla were Firmicutes (0%–7.19%; average, 1.15%), Bacteroidetes (0%–1.64%; average, 0.44%) and Actinobacteria (0%–0.46%; average, 0.24%). This result suggested that the diversity of bacteria is abundant in the rhizosphere soil, while the diversity of bacteria was poor within *Salicornia europaea* plant samples. To the extent of our knowledge, this study is the first to characterize and compare the endophytic bacteria found within different halophytic plant species roots using PCR-based Illumina pyrosequencing method.

Keywords: halophyte, endophytic bacteria, diversity, Illumina, 16S rRNA gene

Introduction

Endophytic bacteria are defined as bacteria that colonize the internal tissue of the plant and do not harm the host. Endophytic bacteria have become an important research area in China. He (2009), in a study of hundreds of species of endophytic bacteria, not only explored the endogenous microbial physiological groups, but also found some new strain classifications. Microorganisms can in some cases increase their germination rate, strengthen the colonization ability of plants in harsh conditions, and can promote the growth of the plant (Chanway, 1997; Bent and Chanway, 1999).

Most information about endophytic bacterial diversity has been obtained using culture-dependent approaches. However, the unknown growth requirements of many bacteria and the presence of cells in a viable but non-cultivable state (Tholozan *et al*., 1999) means that the bacteria identified using conventional cultivation techniques constitute <1% of the bacterial species present (Amann *et al*., 1995). Culture-dependent biodiversity studies of the endophytic community are somewhat limited. 16S ribosomal RNA gene (rRNA) sequence-based methods can provide additional information for analyzing the community composition of endophytic bacteria. This approach has been successfully used for bacterial community analysis in a number of different environments, including soil ecosystems (Dunbar *et al*., 1999), marine (Cottrell and Kirchman, 2000) and rhizosphere environments (Smalla *et al*., 2001), in food (Cocolin *et al*., 2002), and in the human intestine (Kibe *et al*., 2005), in order to overcome the limitations of culture-dependent approaches.

Metagenomic sequencing represents a powerful alternative to 16S rDNA sequencing when analyzing complex microbial communities (Riesenfeld *et al*., 2004; Tringe and Rubin, 2005; von Mering *et al*., 2007). Illumina HiSeq 2000 has fewer errors than 454 sequencing (Smith *et al*., 2008) and it could provide a higher phylogenetic resolution than 454 based approaches (Spear *et al*., 2008). The advantage of Illumina is that it provides 30 times more reads and this enables us to perform in-depth sequencing of hundreds of samples in one run at a fraction of the costs, making it an excellent tool for microbial diversity studies.

Salicornia europaea and *Sueada aralocaspica* are halophytes and represent typical indicator plants in saline-alkaline soils. Their wild plant resources have great ecological, economic and social benefits. Halophyte benefits include enhanced growth and reproduction, and the development of resistance towards abiotic and biotic stresses (Piccoli *et al*., 2011; Choudhary, 2012). *Salicornia europaea* is a true halophyte, and the Na^+ concentration in its leaves can reach 3% (Zhao and Li, 1999), *Sueada aralocaspica* is an endemic species in the central Asian desert, but is only found around Xinjiang in China (Liu *et al*., 2009). Its fruit shows dichotocarpism. These two halophytes are important genetic resources and are used to study halophilic mechanisms in euhalophytes. However, there

*For correspondence. E-mail: tianchy@ms.xjb.ac.cn; Tel.: +86-991-7885301; Fax: +86-991-7885320

has been little research into moderately halophilic endophytic bacteria within these two halophytes when they are growing on the north slope of Tianshan Mountain, China.

In this study, we report on the use of multitag Illumina-based 16S rRNA gene sequences to characterize the bacterial community structure in halophytes. Illumina sequencing, using the 16S rRNA gene as the biomarker, was conducted to examine the bacterial diversity of those halophyte samples, and to get a broader overview of the endophytic bacteria community's structural diversity in halophytes.

Materials and Methods

Plant sampling and plant surface sterilization

Individual plants and the root soils surrounding *Salicornia europaea* and freshly matured fruits of *Sueada aralocaspica* were collected from a natural population of plants (44° 14′ N; 87° 44′ E; 445 m above sea level) growing in saline desert soils near the Fukang Desert Ecosystem Observation and Experimental Station (FDEOES), Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, in Xinjiang Province, China, on 7 October 2012. The *Sueada aralocaspica* seeds were taken from at least 200 plants and allowed to dry naturally for 10 days at ambient room conditions. The seeds were separated from the dried plant material and sorted into brown and black seeds. Then each type of *Sueada aralocaspica* seed and the plant and rhizosphere soil samples from *Salicornia europaea* were individually pooled, thoroughly mixed and stored dry two weeks at 4°C in a closed cotton bag until used in the experiments. Plant surface sterilization was conducted as previously described (Shi *et al.*, 2009).

DNA extraction, PCR, and pyrosequencing

About 1.2 g of the surface-sterilized plants was frozen using liquid nitrogen and ground to a fine powder. DNA extraction was conducted as previously described (Zhou *et al.*, 1996; Sun *et al.*, 2008). L-V3 (5′-CCTACGGGAGGCAGCAG-3′) and R-V3 (5′-TTACCGCGGCTGCTGGC-3′) primers were chosen to amplify the V3 hypervariable region of the 16S rRNA gene (Dethlefsen *et al.*, 2008). The 50 μl PCR reaction mixture contained 100 ng of DNA extract, 1×Taq reaction buffer, 20 pmol of each primer, 200 μM each dNTP and 1.5 U of Taq DNA polymerase (Sangong Biotech). After initial denaturation at 95°C for 4 min, each thermal cycling consisted of denaturation at 94°C for 1 min, annealing at 55°C for 60 sec, and elongation at 72°C for 1.5 min. At the end of 30 cycles, the final extension step was at 72°C for 8 min.

Amplicons were sent for pyrosequencing on the Illumina platform at BGI-Shenzhen. An Illumina HiSeq 2000 was used to sequence the samples. We constructed a paired-end library with an insert size of ~350 base pairs for every sample. Adaptor contamination and low-quality reads were discarded from the raw reads and the remaining reads were filtered to eliminate sugar beet host DNA that was based on the 16S rDNA V3 reference.

Phylogenetic assignment, alignment, and clustering of sequence reads

After elimination of the attachment sequence and the low complexity sequence, carry-on overlap splicing to the single sample, processing of the elimination primer sequence enabled us to compare each unique tag against the RefV3 database using BLASTN (Huse *et al.*, 2008). All tags output only e value of 1e-5 that is less than the comparison result; if the result is greater than 50, only the output of the previous 50 is done prior to selecting the best classification results as species annotation information. There was no comparison with NA (Sequences that could not be classified into any known group) database sequence annotation. Comparison of multiple annotation information sequence is according to the mode principle annotation; if the 66% alignment results both support the same species classification, it is considered to belong to the classified species (Huse *et al.*, 2008).

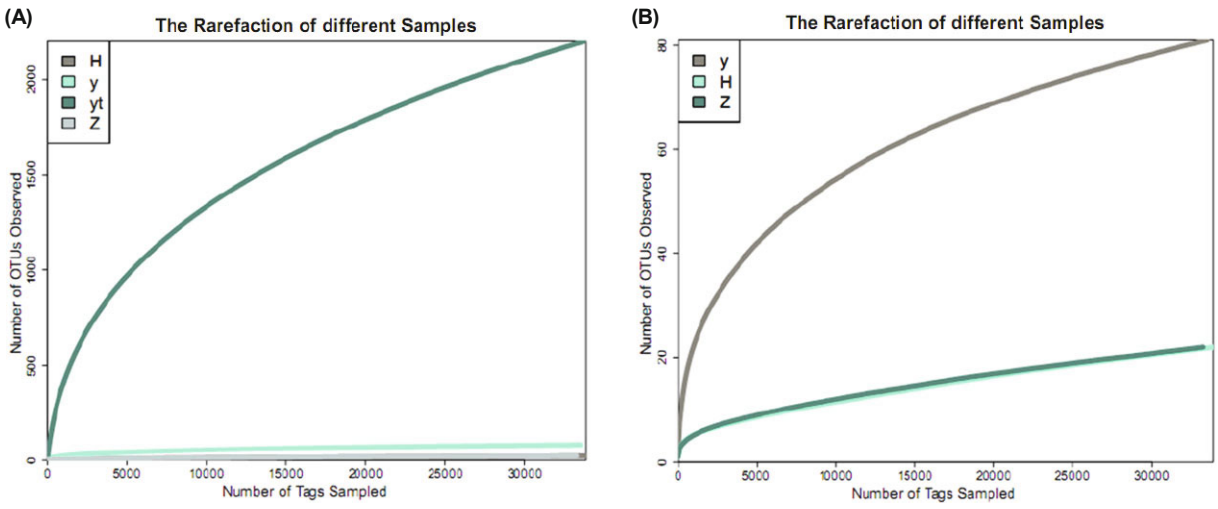

Fig. 1. Rarefaction analysis of the different samples. Rarefaction curves of OTUs clustered at 97% phylotype similarity level. (A) samples H, y, yt, Z; (B) samples y, H, Z.

Table 1. Alpha diversity of endophytic and rhizosphere bacteria in samples

Sample	Threshold	Number of OTUs	Number of OTUs				
			chao1	ACE	Shannon	npShannon	Simpson
y	0.03	81	111	104	1.027	1.033	0.552
yt	0.03	2,206	3,374	3,974	5.945	6.016	0.009
H	0.03	22	39	77	0.727	0.729	0.507
Z	0.03	22	44	66	0.706	0.709	0.523

Diversity estimations, CA and PCA

Mothur was used to assign operational taxonomic units (OTUs) to 16S rRNA at cutoff level of 3% (Schloss *et al.*, 2009). Rarefaction analysis was performed by mothur. The intra-sample rarefaction curves were generated using a re-sampling without replacement approach. Rarefaction curves provide a way of comparing the richness observed in different samples. We calculated the alpha diversity for each sample based on the OTU results. Alpha diversity is the analysis of species diversity in a single sample, including chao1, ACE values, and Shannon and the Simpson indices (Paul and Josephine, 2004).

The optimal classification level for species abundances in the samples allowed us to calculate the distance between samples and to undertake clustering analysis (CA) of samples. The optimal classification level, that is, the taxonomic classification level of the bacterial sequences as high as possible, the number of Tag assigned to the taxonomy as many as possible. Distance was calculated as the Minkowski distance since the bootstrap value was 1000, Cluster analysis was undertaken by the R-pvclust software package (http://www.is.titech.ac.jp/~shimo/prog/pvclust/), and the accuracy evaluation of the cluster provided two types of p-values: Approximately Unbiased (AU) and Bootstrap Probability (BP), where the AU was found to be closer to the non-biased assessment indicators. Principal component analysis (PCA) was based on optimal classification levels and relative abundance of the species, and species that stemmed from the differences between samples were found. The heatmap figure and Venn diagrams were produced by R packages heatmap (Kolde, 2012) and VennDiagram programmes (Chen, 2012), respectively.

Results

Richness and diversity of bacteria phylotypes

Four samples were sequenced using a single lane of a paired-end Illumina HiSeq 2000 run, resulting in 89.23 Mb reads and 594,868 tags of 150 nucleotides. After the elimination attachment sequence and low complexity sequence, carry-on overlap splicing to the single sample and processing of elimination primer sequence produced 134,435 bar-coded V3 16S rRNA sequences, with OTUs ranging from 22 to 2206. The rarefaction curves tended to approach the saturation plateau in the four sample types (Fig. 1). The total number of detected OTUs at 97% sequence similarity was much lower in all four sample types, up to 14 OTUs, with different phylogenetic OTUs ranging from 22 to 2186 (Table 1). The alpha diversity estimation demonstrated that endophytic bacteria diversity was abundant in the four sample types (Table 1).

All sequences were classified from phylum to genus by the Mothur program using the default setting. We identified 24 different phyla, 38 classes and 110 families in the samples. The four libraries showed very dissimilar 16S rRNA profiles even in phylum level distributions (Fig. 2). Three different phyla out of the 24 total phylotypes were common to all four libraries, which accounted for 11.07%, 0.63%, and 15.25% of the total reads in the H, y, yt, and Z libraries, respectively. Only one phyla contained mean OTU abundances that were more than 5% in each sample, but they jointly accounted for 96.38% of the total reads. Proteobacteria was the most abundant division (Fig. 2), accounting for approximately 42.07% to 99.26% of the reads across all samples, whereas the members from Bacteroidetes, Gemmatimonadetes, and Verruco-

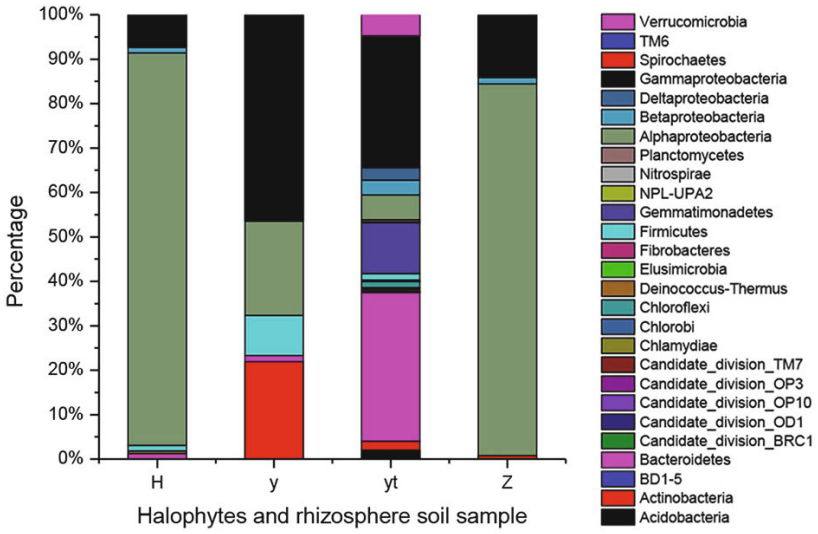

Fig. 2. Abundances of different phyla and classes in Proteobacteria in the four halophytes and rhizosphere soil samples. The abundance is presented in terms of percentage in total effective bacterial sequences in a sample, classified using BLASTN RefV3 database at a confidence threshold of 66%.

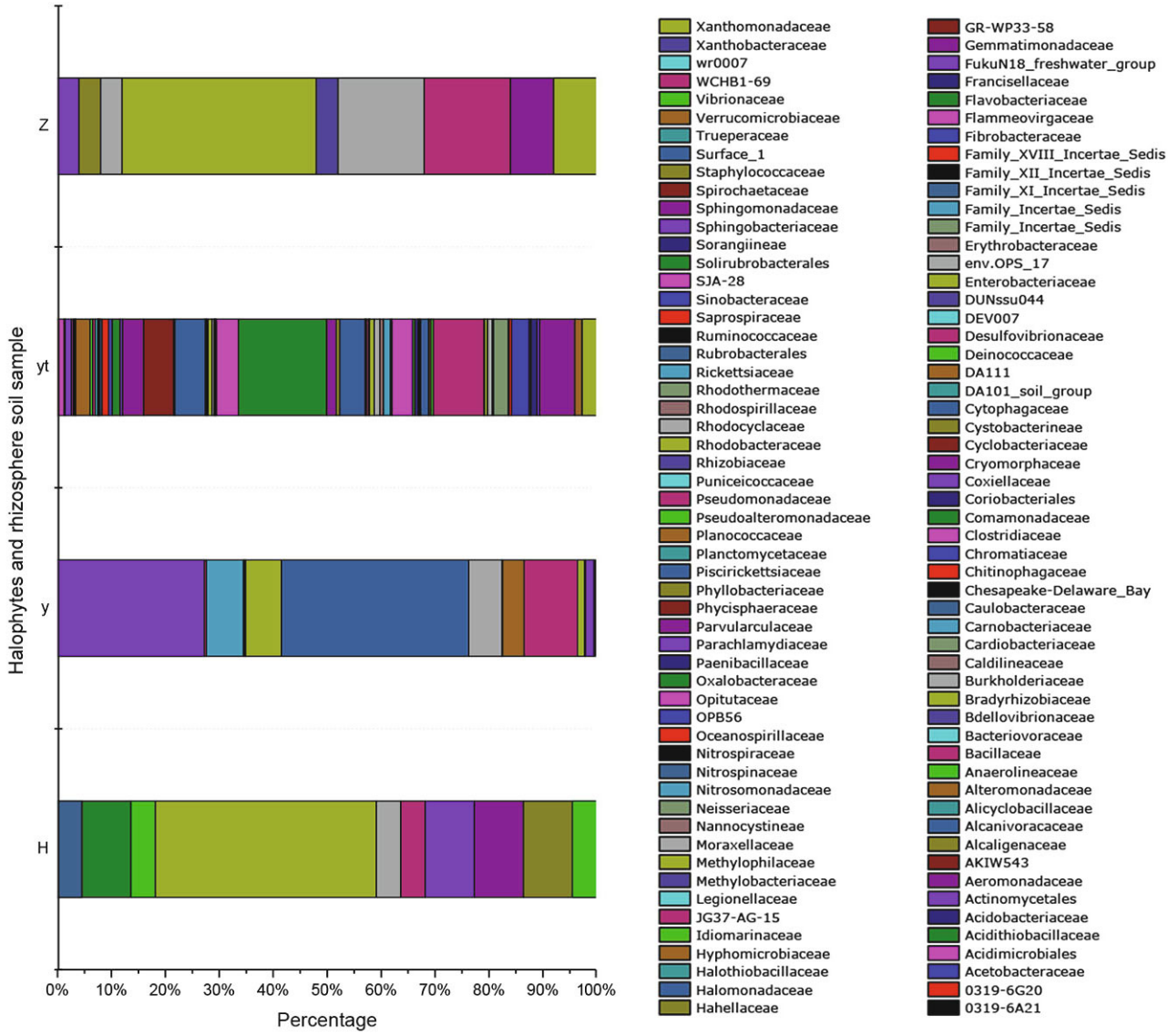

Fig. 3. Abundances of different orderes in bacteria in the four halophytes and rhizosphere soil samples. The abundance is presented in terms of percentage in total effective bacterial sequences in a sample, classified using BLASTN RefV3 database at a confidence threshold of 66%.

microbia made up 31.60%, 10.81%, and 4.46% of all the libraries, respectively. Sorangiineae (1.05%), Chitinophagaceae (1.12%), Enterobacteriaceae (1.20%), Nitrosomonadaceae (1.20%), Verrucomicrobiaceae (1.21%), Piscirickettsiaceae (1.32%), Comamonadaceae (1.37%), Moraxellaceae (1.39%), Gemmatimonadaceae(1.62%), Alteromonadaceae (2.54%), Xanthomonadaceae (2.57%), Rhodothermaceae (2.60%), Actinomycetales (2.78%), Sinobacteraceae (3.00%), Opitutaceae (3.52%), Cryomorphaceae (3.63%), Flammeovirgaceae (3.77%), Cyclobacteriaceae (5.23%), Cytophagaceae (5.29%), Sphingomonadaceae (6.07%), Halomonadaceae (6.59%), Pseudomonadaceae (9.40%), and Flavobacteriaceae (15.30%) were the major bacterial orders recovered, which together composed over 83.76% of the microbial community (Figs. 3 and 4).

Similarity analysis of the four sample types

As shown in Fig. 5, cluster analysis (CA), based on abundances of orders, revealed that bacterial communities in the four samples could be clustered into three groups: (1) Group I contained the *Salicornia europaea* samples; (2) Group II contained the two *Sueada aralocaspica* samples and (3) Group III contained rhizosphere soil from around *Salicornia europaea*. This grouping pattern was similar at the family and the genus levels.

Principal coordinates analysis (PCoA) was also conducted to evaluate similarities of different samples using three different approaches. These were RDP Classifier taxa, OTUs and UniFrac. For the first two approaches, taxa or OTUs were regarded as equally related, whereas UniFrac incorporates the degree of divergence in the phylogenetic tree of OTUs into PCoA (Hamady *et al.*, 2010; Qian *et al.*, 2010).

Fig. 4. Bacterial distribution of the 110 orderes among the four samples. Double hierarchical dendrogram shows the bacterial distribution. The bacterial phylogenetic tree was calculated using the neighbor-joining method and the relationship among samples was determined by Bray-Curtis distance and the complete clustering method. The heatmap plot depicts the relative percentage of each bacterial family (variables clustering on the vertical-axis) within each sample (horizon-axis clustering). The relative values for bacterial family are indicated by color intensity with the legend indicated at the top right corner.

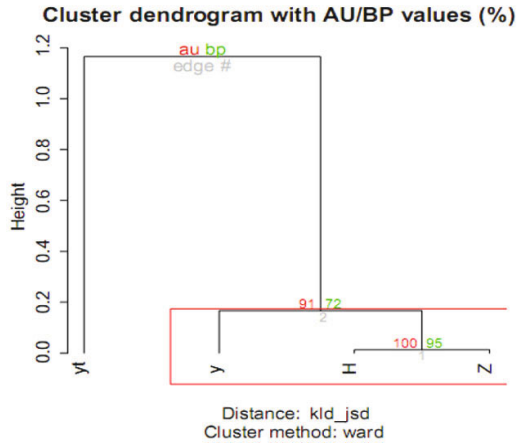

Fig. 5. Clusters that contain the essential 16S rRNA genes of endophytic bacteria in the 4 halophytes and soil samples. The calculation method of the distance to scale. Cluster analysis using R software is provided by hclust packet.

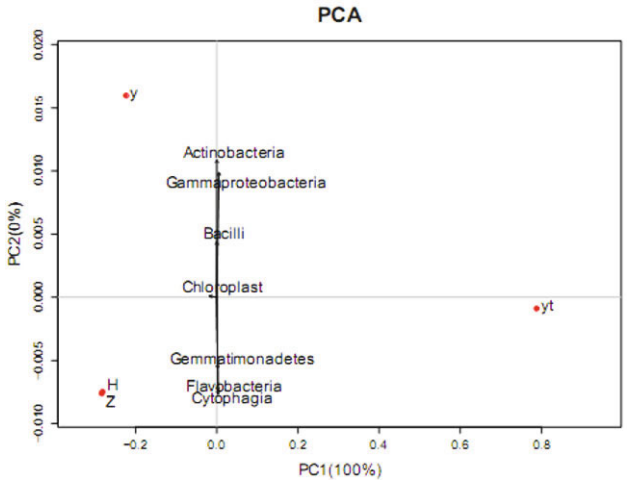

Fig. 6. Principal component analysis with similarity as instrumental variables, based on the abundance of 13 classes by the Illumina reads, was carried out with 3 plant samples and 1 soil sample from Changji. Two first components (PC1 and PC2) were plotted and represented 100% of whole inertia.

The PCoA analysis results are shown in Fig. 6 (UniFrac at 3% cutoff).

We hypothesized that there would be significant differences among the bacteria from different samples. As demonstrated by the CA and PCoA, the samples were very similar to others from the same geographical area, possibly due to their unique compositions based on area differences. There were also similarities between the plants.

Relationships between endophytic bacterial communities in the halophytes and halophyte-associated rhizosphere bacterial communities

The total number of unique OTUs detected throughout the study was 2,181. The greatest number of OTUs for any given sample was detected in the halophyte rhizosphere soil (Fig. 7). Endophytic bacterial diversity was reduced within the *Sueada aralocaspica* brown (22 OTUs) and black seeds (22 OTUs). There was greater variation in the rhizosphere soil taken from around *Salicornia europaea*, *Salicornia europaea* plants, and the *Sueada aralocaspica* seed comparisons (Fig. 7). There were 2206, 81, 22, and 22 OTUs exclusive to the soil bacteria from the rhizosphere around *Salicornia europaea*, the endophytic bacteria in *Salicornia europaea* plants and the brown and black seeds of *Sueada aralocaspica*, respectively (Fig. 7). Fourteen OTUs were common to all four sample types, which are Pseudomonadaceae, *Pseudomonas*, Alphaproteobacteria, Betaproteobacteria, Gammaproteobacteria, Acinetobacter, Moraxellaceae, Raoultella, *Escherichia*, Enterobacteriaceae, unclassified Enterobacteriaceae, Burkholderiales, Sphingomonadales, Sphingomonadaceae.

The patterns of the rarefaction curves for the halophyte samples were very different from those of the rhizosphere soils around *Salicornia europaea* (Fig. 2). Bacterial diversity was much higher in the rhizosphere soil compared to the *Salicornia europaea* sample. In the soil samples, OTU abundance (2,206) was particularly large at high levels of dissimilarity while in the *Salicornia europaea* samples the diversity (81 OTUs) was low at high levels of dissimilarity. This result suggested that the rhizosphere soil was very phylum rich while the *Salicornia europaea* samples were phylum poor.

The species shared among these communities have been illustrated using a Venn diagram, which compares the relationships among these communities in detail. The results show that the number of species shared between the rhizosphere soil bacteria and the *Salicornia europaea* endophytic bacteria communities was 71 (Fig. 7), i.e., 3.33% of the OTUs in the rhizosphere soil bacteria were present in the *Salicornia europaea* endophytic bacteria. The most abundant OTUs shared by the two groups were *Dietzia* (0.1% and 0.04% of the yt and y reads, respectively) and *Nocardioides* (0.01% and 0.03% of the yt and y reads, respectively).

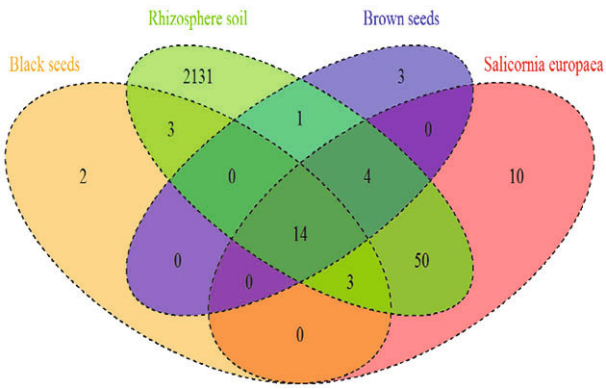

Fig. 7. Four-way Venn diagram of unique Number of tags detected in the four halophytes and rhizosphere soil samples. Black seeds, black seeds of *Sueada aralocaspica*. *Salicornia europaea*, the root, stem and leaf of *Salicornia europaea*. Rhizosphere soil, rhizosphere soil of *Salicornia europaea*. Brown seeds, brown seeds of *Sueada aralocaspica*.

Discussion

This study investigated endophytic bacteria diversity in halophytes found on the north slope of Tianshan Mountain. Illumina sequencing of the V3 region of bacterial 16S rRNA genes and metagenomic library analysis enabled us to observe the community structure of endophytic bacteria in halophytes found on the north slope of Tianshan Mountain. To the best of our knowledge, this is the first time PCR-based Illumina sequencing technology has been used to investigate endophytic bacterial diversity in halophytes. This study represents a proof of principle that short-read sequencing can be used to characterize complex microbiomes.

The capacity of the Illumina platform to generate enormous data sets is undoubtedly an advantage (Lazarevic *et al.*, 2009; Gloor *et al.*, 2010). However, if low-abundance phenotype discovery and accurate measurements of alpha diversity are desired, the errors must be managed effectively, otherwise community characterization is only useful at a coarse level. In this study, assembly was accomplished by the use of overlapping paired end reads, and a modified single linkage clustering protocol was applied at 97% sequence identity. Additionally, problems resulting from the sensitivity of the technology could be avoided by direct multiplex PCR amplifications of the environmental samples, as outlined in this protocol. Although biases are likely to be present in these results, particularly during the DNA extraction and PCR amplification procedures, these biases are unlikely to have missed so many taxa that our estimate of the number of OTUs is wrong.

The numbers of OTUs, Chao 1 and ACE at the 3% cutoff level suggested that the unique sequence approached the total number of sequences at the same sequencing depth and that the richness values varied by 1–2 times among the four samples. The rarefaction curve also tended to plateau, which indicated that the library was large enough to reflect the endophytic bacterial diversity of halophytes.

The most dominant phylum in all the samples was the Proteobacteria, which was consistent with other studies (Hallmann *et al.*, 1997; Moore *et al.*, 2006; Sun *et al.*, 2008). The other large groups were Actinobacteria, Carnobacteriaceae, Enterobacteriaceae, Halomonadaceae, Moraxellaceae, Planococcaceae, Pseudomonadaceae, Rhodobacteraceae, and Sphingobacteriaceae. This indicated that endophytic bacterial diversity was abundant in halophytes, and there might be some specificity between the endophytes and their host plants. However, the number of Proteobacteria classes identified varied between the four samples. In some studies on the endophytes using culture-independent approaches, Gammaproteobacteria dominated, with the genus *Pseudomonas* comprising 42% of the total isolate collection (Moore *et al.*, 2006). It has been previously observed that in many cases, Pseudomonas are abundant in both the soil environment (Samish *et al.*, 1963; Spiers *et al.*, 2000) and inside the plant (Gardner *et al.*, 1982; Tanprasert and Reed, 1997; Rademaker *et al.*, 1998; Huse *et al.*, 2008). It has been suggested that endophytic bacteria colonize plants primarily through the root network via natural and artificial wound sites, root hairs and at epidermal junctions (Sprent and de Faria, 1988; Pan *et al.*, 1997).

During the long process of evolution, symbiotic relationships have formed between endophytic bacteria and host plants. Genetic diversity of endophytic bacteria is determined by environmental conditions and host genotype. Soil organic matter content and composition are key factors that affect microbial biomass, community composition and biological activity. Microbial diversity and soil types, in particular, are closely related to soil organic matter content. Endophytic bacteria the two types of seeds of *Sueada aralocaspica* were studied in the four sample types investigated in this study. The results show clear differences in number and species of endophytic bacteria between the black seeds and the brown seeds. One of the reasons for these differences may be genotype variability among the plants investigated. Adams and Kloepper (Adams and Kloepper, 2002) studied the impact of cotton plant genotype on the inherent bacterial population of their seeds, seedling stems and root tissues. They found that the distinct genetic, morphological and physiological characteristics of individual cotton cultivars led to differences in the endophytic bacterial community structure among the cultivars. In this study, the two halophyte species had strong agronomic traits. Hence, we can infer that the differences in nutritional structure among the genotypes of the halophytes are probably one of the key reasons for the differences seen in the endophytic bacterial communities.

In summary, endophytic bacterial diversity was high in the halophytes found on the north slope of Xinjiang. The distribution of endophytic bacteria showed obvious dynamic changes. The host genotypes and growth stages are the most important factors to affect the endophytic bacteria community. These active endophytes and halophytes have a close relationship during the growth of the plants and future studies should address endophytic bacterial functional diversity in halophytes. Comprehensive knowledge of the endophytic bacteria and their various roles will lay the foundation for further research and applications.

Acknowledgements

This work was supported by the National 863 High Technology R&D Program (NO.2012AA101404-6) and the State Key Laboratory of Desert and Oasis Ecology (Y371162).

References

Adams, P.D. and Kloepper, J.W. 2002. Effect of host genotype on indigenous bacterial endophytes of cotton (*Gossypium hirsutum* L.). *Plant Soil* **240**, 181–189.

Amann, R.I., Ludwig, W., and Schleifer, K.H. 1995. Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiol. Rev.* **59**, 143–169.

Bent, E. and Chanway, C.P. 1998. The growth-promoting effects of a bacterial endophyte on lodgepole pine are partially inhibited by the presence of other rhizobacteria. *Can. J. Microbiol.* **44**, 980–988.

Chanway, C.P. 1997. Inoculation of tree roots with plant growth promoting soil bacteria: an emerging technology for re-forestation. *Forest. Sci.* **43**, 99–112.

Chen, H.B. 2012. Venn Diagram: Generate high-resolution Venn and Euler plots. R package version 113.

Choudhary, D.K. 2012. Microbial rescue to plant under habitat-

imposed abiotic and biotic stresses. *Appl. Microbiol. Biotechnol.* **96**, 1137–1155.

Cocolin, L., Rantsiou, K., Iacumin, L., Cantoni, C., and Comi, G. 2002. Direct identification in food samples of *Listeria* spp. and *Listeria monocytogenes* by molecular methods. *Appl. Environ. Microbiol.* **68**, 6273–6282.

Cottrell, M.T. and Kirchman, D.L. 2000. Community composition of marine bacterioplankton determined by 16S rRNA gene clone libraries and fluorescence *in situ* hybridization. *Appl. Environ. Microbiol.* **66**, 5116–5122.

Dethlefsen, L., Huse, S., Sogin, M.L., and Relman, D.A. 2008. The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biol.* **6**, e280.

Dunbar, J., Takala, S., Barns, S.M., Davis, J.A., and Kuske, C.R. 1999. Levels of bacterial community diversity in four arid soils compared by cultivation and 16S rRNA gene cloning. *Appl. Environ. Microbiol.* **65**, 1662–1669.

Gardner, J.M., Feldman, A.W., and Zablotowicz, M. 1982. Identity and behaviour of xylem-residing bacteria in rough lemon roots of Florida citrus trees. *Appl. Environ. Microbiol.* **43**, 1335–1342.

Gloor, G.B., Hummelen, R., Macklaim, J.M., Dickson, R.J., Fernandes, A.D., MacPhee, R., and Reid, G. 2010. Microbiome profiling by Illumina sequencing of combinatorial sequence-tagged PCR products. *PLoS One* **5**, e15406.

Hallmann, J., Quadt-Hallmann, A., Mahafee, W.F., and Kloepper, J.W. 1997. Bacterial endophytes in agricultural crops. *Can. J. Microbiol.* **43**, 895–914.

Hamady, M., Lozupone, C., and Knight, R. 2010. Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data. *ISME J.* **4**, 17–27.

He, R.Q. 2009. Plant endophyte becomes one of hot topics in the current microbiological study in China. *Microbiology* **36**, 1.

Huse, S.M., Dethlefsen, L., Huber, J.A., Mark, W.D., Relman, D.A., and Sogin, M.L. 2008. Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS Genet.* **4**, e1000255.

Kibe, R., Sakamoto, M., Yokota, H., Ishikawa, H., Aiba, Y., Koga, Y., and Benno, Y. 2005. Movement and fixation of intestinal microbiota after administration of human feces to germfree mice. *Appl. Environ. Microbiol.* **71**, 3171–3178.

Kolde, R. 2012. Pheatmap: Pretty Heatmaps. R package version 061.

Lazarevic, V., Whiteson. K., Huse, S., Hernandez, D., Farinelli, L., Osterås, M., Schrenzel, J., and François, P. 2009. Metagenomic study of the oral microbiota by Illumina high-throughput sequencing. *J. Microbiol. Methods* **79**, 266–271.

Liu, Y.F., Wei, Y., and Yan, C. 2009. Germination characteristics and ecological adaptation of dimorphic seeds of *Sueada aralocaspica*. *Acta Ecologica Sinica* **29**, 6609–6614.

Moore, F.P., Barac, T., Borremans, B., Oeyen, L., Vangronsveld, J., van der, Lelie, D., Campbell, C.D., and Moore, E.R. 2006. Endophytic bacterial diversity in poplar trees growing on a BTEX-contaminated site: the characterisation of isolates with potential to enhance phytoremediation. *Syst. Appl. Microbiol.* **29**, 539–556.

Pan, M.J., Rademan, S., Kuner, K., and Hastings, J.W. 1997. Ultrastructural studies on the colonisation of banana tissue and *Fusarium oxysporum* f. sp. *cubense* race 4 by the endophytic bacterium *Burkholderia cepacia*. *J. Phytopathol.* **145**, 479–486.

Paul, F.K. and Josephine, Y.A. 2004. Bacterial diversity in aquatic and other environments: what 16S rDNA libraries can tell us. *FEMS Microbiol. Ecol.* **47**, 161–177.

Piccoli, P., Travaglia, C., Cohen, A., Sosa, L., Cornejo, P., Masuelli, R., and Bottini, R. 2011. An endophytic bacterium isolated from roots of the halophyte *Prosopis strombulifera* produces ABA, IAA, gibberellins A(1) and A(3) and jasmonic acid in chemically-defined culture medium. *Plant Growth Regul.* **64**, 207–210.

Qian, P.Y., Wang, Y., Lee, O.O., Lau, S.C., Yang, J., Lafi, F.F., Al-Suwailem, A., and Wong, T.Y. 2010. Vertical stratification of microbial communities in the Red Sea revealed by 16S rDNA pyrosequencing. *ISME J.* **5**, 507–518.

Rademaker, J.L.W., Louws, F.J., and de Bruijn, F.J. 1998. Characterisation of the diversity of ecologically important microbes by rep-PCR genomic fingerprinting, pp. 1–27. Molecular microbial ecology manual, Kluwer Academic Publishers, Netherlands.

Riesenfeld, C.S., Schloss, P.D., and Handelsman, J. 2004. Metagenomics: genomic analysis of microbial communities. *Annu. Rev. Genet.* **38**, 525–552.

Samish, Z., Etinger-Tulczynska, R., and Bick, M. 1963. The microflora within the tissue of soft fruit and vegetables. *J. Food Sci.* **28**, 259–266.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., *et al.* 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **75**, 7537–7541.

Shi, Y.W., Lou, K., and Li, C. 2009. Promotion of plant growth by phytohormone producing endophytic microbes of sugar beet. *Biol. Fert. Soils* **45**, 645–653.

Smalla, K., Wieland, G., Buchner, A., Zock, A., Parzy, J., Kaiser, S., Roskot, N., Heuer, H., and Berg, G. 2001. Bulk and rhizosphere soil bacterial communities studied by denaturing gradient gel electrophoresis: plant-dependent enrichment and seasonal shifts revealed. *Appl. Environ. Microbiol.* **67**, 4742–4751.

Smith, D.R., Quinlan, A.R., Peckham, H.E., Makowsky, K., Tao, W., Woolf, B., Shen, L., Donahue, W.F., Tusneem, N., Stromberg, M.P., *et al.* 2008. Rapid whole-genome mutational profiling using next-generation sequencing technologies. *Genome Res.* **18**, 1638–1642.

Spear, G.T., Sikaroodi, M., Zariffard, M.R., Landay, A.L., French, A.L., and Gillevet, P.M. 2008. Comparison of the diversity of the vaginal microbiota in HIV-infected and HIV-uninfected women with or without bacterial vaginosis. *J. Infect. Dis.* **198**, 1131–1140.

Spiers, A.J., Buckling, A., and Raineythen, P.B. 2000. The causes of *Pseudomonas* diversity. *Microbiology* **146**, 2345–2350.

Sprent, J.I. and de Faria, S.M. 1988. Mechanisms of infection of plants by nitrogen fixing organisms. *Plant Soil* **110**, 157–165.

Sun, L., Qiu, F., Zhang, X., Dai, X., Dong, X., and Song, W. 2008. Endophytic bacterial diversity in rice (*Oryza sativa* L.) roots estimated by 16S rDNA sequence analysis. *Microb. Ecol.* **55**, 415–424.

Tanprasert, P. and Reed, B.M. 1997. Detection and identification of bacterial contaminants from strawberry runner explants. *In Vitro Cell. Dev. Biol. - Plant* **33**, 221–226.

Tholozan, J.L., Cappelier, J.M., Tissier, J.P., Delattre, G., and Federighi, M. 1999. Physiological characterization of viable-but-nonculturable *Campylobacter jejuni* cell. *Appl. Environ. Microbiol.* **65**, 1110–1116.

Tringe, S.G. and Rubin, E.M. 2005. Metagenomics: DNA sequencing of environmental samples. *Nature Rev. Genet.* **6**, 805–814.

von Mering, C., Hugenholtz, P., Raes, J., Tringe, S.G., Doerks, T., Jensen, L.J., Ward, N., and Bork, P. 2007. Quantitative phylogenetic assessment of microbial communities in diverse environments. *Science* **315**, 1126–1130.

Zhao, K.F. and Li, F.Z. 1999. China Halophyte, pp. 1–6. Science Press, Beijing, P. R. China.

Zhou, J., Bruns, M.A., and Tiedje, J.M. 1996. DNA recovery from soils of diverse composition. *Appl. Environ. Microbiol.* **62**, 316–322.