Science of the Total Environment 563–564 (2016) 199–209

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

# Bacterial diversity and composition in major fresh produce growing soils affected by physiochemical properties and geographic locations

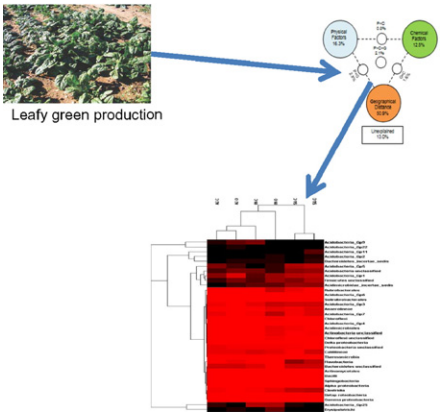
Jincai Ma [a,b], A. Mark Ibekwe [b,*], Ching-Hong Yang [c], David E. Crowley [d]

[a] Key Laboratory of Groundwater Resources and Environment, Ministry of Education, Jilin University, Changchun 130021, China
[b] USDA-ARS U. S. Salinity Laboratory, Riverside, CA 92507, United States
[c] Department of Biological Sciences, University of Wisconsin, Milwaukee, WI 53211, United States
[d] Department of Environmental Sciences, University of California, Riverside, CA 92521, United States

## HIGHLIGHTS

- Geographic distance was the most significant factor affecting microbial composition.
- Physical and chemical properties significantly impacted microbial communities.
- Higher numbers of OTUs were observed in organic soils than in convention soils

## GRAPHICAL ABSTRACT




## ARTICLE INFO

*Article history:*
Received 8 January 2016
Received in revised form 14 April 2016
Accepted 15 April 2016
Available online 30 April 2016

Editor: D. Barcelo

*Keywords:*
Bacterial ecology
Pyrosequencing
Artificial neural network

## ABSTRACT

Microbial diversity of agricultural soils has been well documented, but information on leafy green producing soils is limited. In this study, we investigated microbial diversity and community structures in 32 (16 organic, 16 conventionally managed soils) from California (CA) and Arizona (AZ) using pyrosequencing, and identified factors affecting bacterial composition. Results of detrended correspondence analysis (DCA) and dissimilarity analysis showed that bacterial community structures of conventionally managed soils were similar to that of organically managed soils; while the bacterial community structures in soils from Salinas, California were different ($P < 0.05$) from those in soils from Yuma, Arizona and Imperial Valley, California. Canonical correspondence analysis (CCA) and artificial neural network (ANN) analysis of bacterial community structures and soil variables showed that electrical conductivity (EC), clay content, water-holding capacity (WHC), pH, total nitrogen (TN), and organic carbon (OC) significantly ($P < 0.05$) correlated with microbial communities. CCA based variation partitioning analysis (VPA) showed that soil physical properties (clay, EC, and WHC), soil chemical variables (pH, TN, and OC) and sampling location explained 16.3%, 12.5%, and 50.9%, respectively, of total variations in bacterial community structure, leaving 13% of the total variation unexplained. Our current study showed that bacterial community composition and diversity in major fresh produce growing soils from California and Arizona is a function of soil physiochemical characteristics and geographic distances of sampling sites.

Published by Elsevier B.V.

* Corresponding author at: USDA-ARS-U. S. Salinity Laboratory, 450 W. Big Springs Rd, Riverside, CA 92507, United States.
*E-mail address:* Mark.Ibekwe@ars.usda.gov (A.M. Ibekwe).

http://dx.doi.org/10.1016/j.scitotenv.2016.04.122
0048-9697/Published by Elsevier B.V.

## 1. Introduction

Microorganisms play crucial roles in regulating biogeochemical cycling of nutrients and ecosystem function (Torsvik, et al., 2002, Bardgett, et al., 2008, Gardi, et al., 2009 and van der Heijden and Wagg, 2013). Soil microbial communities may vary greatly in their diversity and composition because of differences in soil biotic and abiotic factors that select for different taxonomic groups (Baker, et al., 2009). In addition, soil management and geographic distance may have a major influence in structuring bacterial communities (Dequiedt, et al., 2009 and Drenovsky, et al., 2010). Soil microbial diversity also may determine a soil's resilience and ability to respond to changes in local geographic characteristics and agricultural practices, e.g. organic and conventional management (Kennedy, 1999); therefore, it is essential to investigate microbial community structures to evaluate soil quality (Bending, et al., 2000 and van der Heijden and Wagg, 2013). Analyses of microbial community diversity and structure in farmland soil, wetland soil, forest soils, and even in soils from extreme environments have frequently been reported (Vance, et al., 1987, Liu, et al., 2007, Lueke, et al., 2010 and Paula, et al., 2014). However, information on bacterial ecology of fresh produce growing soils and the extent to which they are shaped by various environmental factors is limited.

Soil microbial ecology has been extensively studied using traditional techniques, among which the most widely used are polymerase chain reaction (PCR)-based techniques, including 16s rRNA gene clone libraries, PCR- denature gradient gel electrophoresis (DGGE) (Kozdroj and van Elsas, 2001; Ibekwe et al., 2001; Ibekwe et al., 2002), and PCR- terminal restriction fragment length polymorphism analyses (Hartmann and Widmer, 2006). Phospholipid fatty acid (PLFA) profiling was another frequently used method in studying microbial ecology in soils (Wu, et al., 2009; Ibekwe and Kennedy 1998; Ibekwe and Kennedy 1999). In recent years, next generation sequencing (NGS) strategies have become the main-stream methods in soil microbial ecology (Roesch, et al., 2007). Compared to earlier microbial ecology techniques, NGS protocols have many advantages, e.g. higher throughput, better coverage, and greater resolution, which enable researchers to identify and study microbial groups with relatively lower abundance (<1%) (Shendure and Ji, 2008). Previous work has demonstrated that bar-coded pyrosequencing can be used to analyze relatively large numbers of individual samples, and survey the bacterial community in each sample to an extent that would be difficult using classical microbial ecology techniques (Dethlefsen, et al., 2008; Jones, et al., 2009).

In the United States nearly 70% of commercial fresh produce, typically lettuce, cabbage, and spinach, are grown in two regions that rotate seasonally. In summer seasons, the fresh produce is grown mainly by farmers in the Salinas Valley located in the central California coast, while during winter seasons farms in Imperial and Yuma Valleys located in southern California, and southwest Arizona are the major suppliers. Monitoring microbial community composition and diversity in soils from the leafy green producing regions is necessary to achieve a sustainable leafy green production, since microbial diversity is a representative indicator reflecting the overall health and quality of agricultural soil (Gardi, et al., 2009).

It is generally accepted that soil microbial communities with a higher biodiversity are more stable and resistant to environment disturbance than those with lower diversity (Girvan, et al., 2005 and van Diepeningen, et al., 2006). Therefore, a soil with reduced biological complexity might be more susceptible to invasive microbes, such as *Escherichia coli* O157:H7 and *Salmonella*, because such soil ecosystems offer an enhanced opportunity for those microorganisms to persist (Semenov, et al., 2008 and van Elsas, et al., 2011). Indeed, leafy greens from major fresh produce growing soils in California and Arizona have previously been infected with zoonotic pathogens that resulted in nationwide foodborne outbreaks (Cooley, et al., 2007 and Taylor, et al., 2013). Recent research also showed that the indigenous microbial community structure was correlated with the persistence of human pathogens, e.g. *E. coli* O157:H7 and *Salmonella*, which were introduced into the agricultural soils via different pathways (Gorski, et al., 2011, Ma, et al., 2013; Ma, et al., 2014). Therefore, a thorough understanding of indigenous microbial ecology is needed not only for the evaluation of the health and quality of agricultural soils, but also for assessment of the health risk of human pathogens associated with soils and the leafy greens grown in those soils.

In the current study, we selected the three major leafy green producing areas in California and Arizona as study sites, and investigated the bacterial community structures in soils from those sites using next generation sequencing technique. The objectives of this research were to: 1) understand the composition of bacterial communities in major leafy green soils, 2) compare the bacterial community structure in conventionally and organically managed soils, and 3) identify factors affecting bacterial community structures in these soils.

## 2. Material and methods

### 2.1. Soil sample collection and characterization

A total of 32 soil samples were collected from farms (plots) of three major fresh produce growing areas during the fall of 2010: 12 from Yuma, Arizona, 12 from Imperial Valley, southern California, and 8 from Salinas Valley, northern California (Ma, et al., 2012a). Equal numbers of organically and conventionally managed soil samples were collected at each farm. Soil selections were made on the basis of grower interviews and on-farm field examinations to ensure that all soil-forming factors, except management, were the same for each field pair. Each field pair consisted of two side-by-side fields, one organic and one conventional. Fields chosen in each pair had the same microclimate, soil profile, soil type, soil classification, and produce. The organic plots in Salinas and Imperial Valleys were certified organic (USDA) for >8 years, based on personal communications with the local extension specialist, and the Yuma Valley organic field was certified for >10 years, thus providing sufficient time for the organic farming practices to influence soil properties. The Yuma Valley plots were managed by the University of Arizona, and all the organic fields from the three regions relied only on organically certified fertilizers, compost, and rotation with legumes. The conventional farms also had been managed conventionally and included the use of inorganic and organic fertilizers and synthetic pesticides.

The site locations were mapped based on their global positioning system (GPS) coordinates (longitude and latitude) location information was recorded for each sample. Among the three sampling sites, Yuma and Imperial Valley shared much more similar weather conditions due to the closeness of the two locations in comparison to the third sampling sites; Salinas Valley located in northern California, which typically has a higher mean annual temperature (MAT) of 14.4 °C, and higher mean annual precipitation (MAP) of 328 mm. Fresh produce grown in the soils at the time the samples were taken were spinach and lettuce. After the plant residue on soil surface was removed, soil samples were collected from each farm located in above mentioned sampling areas using a stainless steel shovel. Each sample (0–15 cm) was a composite of up to 5 individual soil cores taken at 5 m intervals. Soil samples were sieved (2 mm), bagged, and stored at 4 °C in the dark until use. Soil properties characterized (Ma, et al., 2012a) included clay, silt, and sand contents, pH, electrical conductivity (EC), bulk density, water content, water-holding capacity (WHC), total organic carbon (OC), and total nitrogen (T-N) (Klute, 1996). Soil microbial biomass carbon (MBC) was determined by the chloroform-fumigation-extraction method (Vance, et al., 1987). The texture and physiochemical properties of the soils are as previously described (Ma, et al., 2012a).

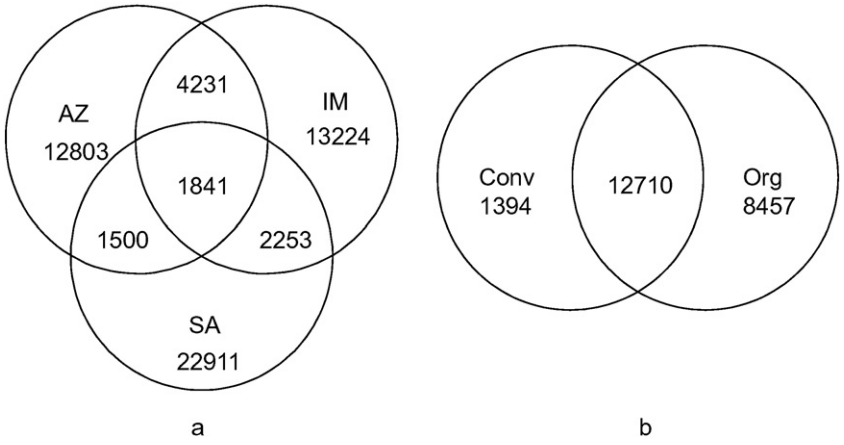



**Fig. 1.** Venn diagrams showing the shared species and unique species in soils collected from Yuma (AZ), Imperial Valley (IM), and Salinas Valley (SA) (Fig.1a) and in soils subjected to conventional (Conv) and organically (Org) managements (Fig. 1b).

### 2.2. Soil DNA Extraction, Pyrosequencing and sequence data analysis

Community DNA was extracted from 32 leafy green-producing soils using a Power Soil Extraction Kit (MO BIO Laboratories, CA) with the bead-beating protocol supplied by the manufacturer. The quality and concentration of the soil DNA were assessed using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, DE). The overall size of the soil DNA was checked by running an aliquot of soil DNA on a 1.0% agarose gel. The soil DNA samples (15.0 μl) were then submitted to Research and Testing Laboratories (Lubbock, TX) for PCR optimization and pyrosequencing analysis. Bacterial tag-encoded FLX amplicon pyrosequencing were carried out as previously described (Acosta-Martinez, et al., 2008 and Dowd, et al., 2008). The 16S universal Eubacterial primers 530F (5′-GTG CCA GCM GCN GCG G) and 1100R (5′-GGG TTN CGN TCG TTG) were used for amplifying the ~600 bp hypervariable region of 16S rRNA genes. Primer and PCR optimizations were done at the Research and Testing Laboratories (Lubbock, TX) according to protocols described previously (Acosta-Martinez, et al., 2008). All FLX-related procedures were performed following Genome Sequencer FLX System manufacturer's instructions (Roche, NJ, USA). Thus, moderate diversity pyrosequencing analysis (≥3000 reads per sample) was performed at the Research and Testing Laboratory (Lubbock, TX, USA). Tags which did not have 100% homology to the original sample tag designation were filtered from the data set. Sequences which were <200 bp after quality trimming also were not considered.

Bacterial pyrosequencing population data were further analyzed by performing multiple sequence alignment techniques using the dist.seqs function in MOTHUR, version 1.9.1 (Schloss, et al., 2009). All the raw reads were treated with the Pyrosequencing Pipeline Initial Process (Cole, et al., 2009) of the Ribosomal Database Project (RDP), (1) to sort those exactly matching the specific barcodes into different samples, (2) to trim off the adapters, barcodes and primers using the default parameters, and (3) to remove sequences containing ambiguous 'N' (Claesson, et al., 2009). Given that a number of diversity and richness estimators tend to suffer from sample size bias (Champely and Chessel, 2002), we "re-sampled" our sequence libraries so that they contained

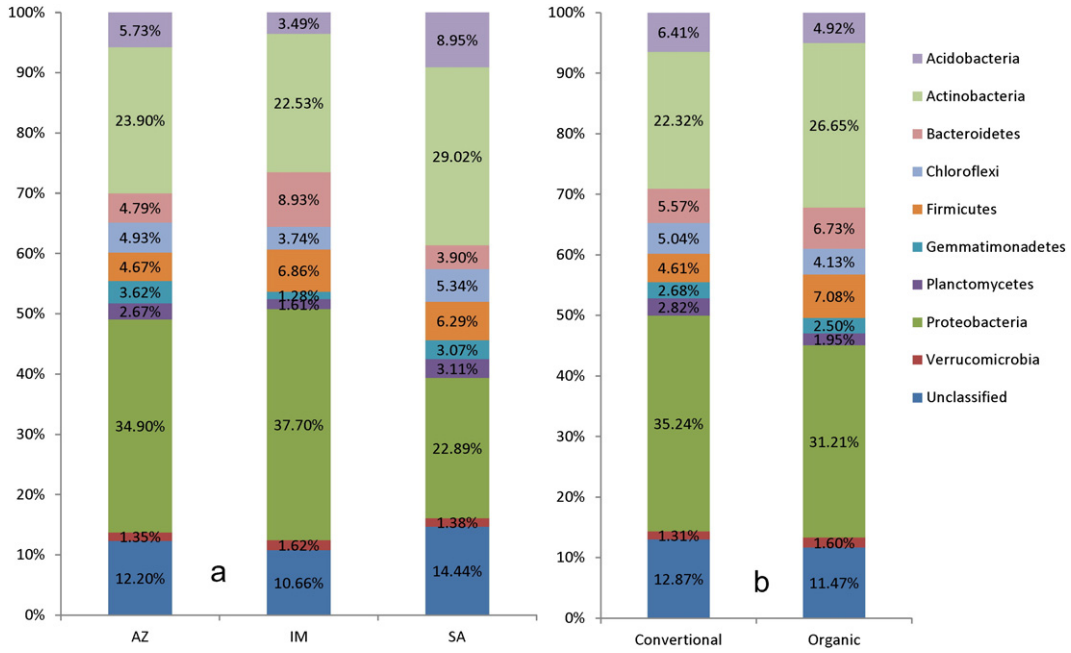



**Fig. 2.** Bar charts of dominant bacterial phyla in soils from Yuma (AZ), Imperial Valley (IM), and Salinas Valley (SA) (Fig. 2a), and in soils subjected to conventional (C) and organic (O) managements (Fig. 2b).

similar numbers of sequences. The subsample function in MOTHUR was used to randomly select a subsample of sequences from each library and these equally sized. The reduced data sets were used in all subsequent analyses, including detrended correspondence analysis (DCA), canonical correspondence analysis (CCA), variation partition analysis (VPA), Venn diagram analysis, and dissimilarity indices analysis.

Following chimera detection, and the re-sampling of the larger sequence libraries, the RDP Classifier function was used to assign identities to the bacterial pyrotag sequence data (Wang and Qian, 2009). MOTHUR was used to align the re-sampled data set and create an all-sample distance matrix, as well as assign sequences to operational taxonomic units (OTU = 97% similarity, using the h-cluster function), calculate diversity indices and richness estimates, and determine the degree of overlap shared among the soil communities. Overlap was calculated using the Yue-Clayton similarity estimator ($\theta_{YC}$), a metric that is scored on a scale of 0 to 1, where 0 represents complete dissimilarity and 1 represents identity (Yue and Clayton, 2005 and Schloss, et al., 2009). When comparing any given set of communities, $\theta_{YC}$ considers the distribution of OTUs between the communities, as well as their relative abundances.

### 2.3. Statistical analysis

DCA, CCA, VPA, Venn diagram analysis, and dissimilarity indices analysis were performed using R package v3.1.0. Linear regression analysis between soil structure (sand and clay content) and bacterial community structure (Shannon diversity index and total OTU count) was conducted using SYSTAT 12 (Systat Software, Chicago, IL).

Artificial neural network (ANN) analysis was conducted using the program Synapse (Peltarion Inc. Stockholm, Sweden). Model training for soil physical, chemical and biological variables used 5000 cycles, which was optimized by examining inflection points that minimized error for both training and validation data sets. Bacterial abundance data at the phylum level and the abundance data of subclasses of *Proteobacteria* were also used in the artificial neural network analysis to see the complete picture of the effects of all the physical, chemical, and biological factors on bacterial community composition and diversity in soils. All the statistical analysis was performed based on the reduced data sets with a total of 50,848 sequences.

## 3. Results

### 3.1. Bacterial community composition

Unique OTUs and shared OTUs in soils from Yuma, Arizona, Imperial Valley and Salinas Valley California are shown in Fig. 1A. Soils from Yuma, Arizona and Imperial Valley, CA shared >6000 OTUs, while Arizona soils and Imperial Valley soils only shared about 3300 and 4000 OTUs with soils from Salinas, CA. There were 1814 OTUs shared by all three locations. Fig. 1B showed that 12,710 OTUs were shared by conventionally and organically managed soils. Furthermore, higher numbers of OTUs were observed in organic soils than in conventionally managed soils. The relative abundance of the dominant phyla in soils from Arizona, Imperial Valley and Salinas, respectively are presented in Fig. 2A. In addition, Fig. 2B displays the relative abundance of major phyla in organic and conventionally managed soils. The results showed that the most abundant phyla were *Proteobacteria* followed by *Actinobacteria*, *Acidobacteria* and *Firmicutes*. Interestingly, the unclassified species accounted for >10% across all soils regardless of their collection sites and management strategies. Abundances of *Actinobacteria* and *Acidobacteria* in soils from Salinas Valley were relatively higher than those in the soils from the other two sites ($P < 0.05$, data not shown). There were no major differences observed in bacteria phyla abundances in conventional and organic soils.

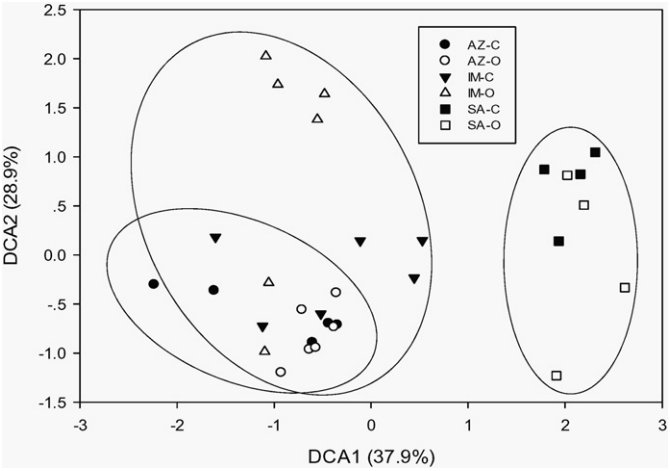



**Fig. 3.** Detrended corresponding analysis (DCA) of bacterial community structures in soils from Yuma (AZ), Imperial Valley (IM), and Salinas Valley (SA). Both conventionally (C) and organically (O) managed soils were included.

### 3.2. Comparison of soil properties and bacterial diversity in soils subjected to different soil management

Selected soil characteristics and bacterial diversity indices of organic and conventionally managed soils from each sampling sited are shown in Table 1. EC was significantly different ($P < 0.05$) between organic and conventionally managed soils from Yuma, Arizona, while no major differences were found in organic and conventional soils from Imperial and Salinas Valleys, CA. The other soil parameters, namely pH, clay, TN and OC, were similar between organic and conventional soil from each sampling site. The bacterial diversity indices, including Shannon indices and coverage, were not significantly different between organic and conventionally managed soils from each sampling location (Table 1).

Further analysis of the data by detrended corresponding analysis (DCA) showed that Yuma, Arizona soils and Imperial Valley soils were largely clustered together and well separated from the Salinas soils by DCA1. In addition, organic soils and conventional soils from each sampling site were not clearly separated by either DCA1 or DCA2. DCA1

**Table 1**
Soil properties and bacterial diversity of soils from Yuma, Arizona, Imperial Valley, California, and Salinas Valley, California. Both conventionally managed and organically managed soils were collected from each site.

| | | Soil property | | | | | | Bacterial diversity | |
|---|---|---|---|---|---|---|---|---|---|
| | | pH | EC ($dS\ m^{-1}$) | clay (%) | WHC (%) | TN (%) | OC (%) | Shannon index ($H$) | coverage |
| Yuma, AZ | Conv | 7.9 ± 0.0 | **1.95** ± 0.13 | 41.7 ± 2.4 | 48.0 ± 2.7 | 0.08 ± 0.03 | 1.82 ± 0.06 | 6.048 ± 1.474 | 0.923 ± 0.149 |
| | Org | 8.0 ± 0.1 | **1.22** ± 0.04 | 41.8 ± 2.0 | 48.1 ± 3.9 | 0.06 ± 0.01 | 1.89 ± 0.06 | 6.695 ± 0.801 | 0.916 ± 0.128 |
| Imperial Valley, CA | Conv | 7.7 ± 0.1 | 0.83 ± 0.98 | 33.9 ± 16.1 | 40.0 ± 6.4 | 0.06 ± 0.02 | 1.79 ± 0.35 | 6.253 ± 1.221 | 0.928 ± 0.120 |
| | Org | 7.8 ± 0.2 | 1.72 ± 0.78 | 31.4 ± 12.2 | 41.5 ± 9.5 | 0.08 ± 0.03 | 1.84 ± 0.37 | 6.668 ± 0.780 | 0.914 ± 0.077 |
| Salinas, CA | Conv | 7.7.5 ± 0.4 | 0.42 ± 0.18 | 16.1 ± 12.4 | 33.0 ± 18.9 | 0.11 ± 0.07 | 1.32 ± 0.78 | 7.013 ± 0.203 | 0.866 ± 0.057 |
| | Org | 7.3 ± 0.5 | 0.54 ± 0.23 | 21.0 ± 15.3 | 47.1 ± 21.5 | 0.17 ± 0.06 | 2.14 ± 1.00 | 7.299 ± 0.385 | 0.850 ± 0.115 |

EC, electrical conductivity; WHC, water holding capacity; TN, total nitrogen; OC, organic carbon. Bolded numbers indicate values were significant at 0.05 level.

and DCA2 together explained about 67% of total variation of bacterial community structure across all soils (Fig. 3). Further hierarchical clustering analysis (Fig. 4) based on class level abundance data confirmed the trend found in Fig. 2 and Fig. 3, which revealed that soils from Salinas clustered together and were well separated from Arizona and Imperial Valley. Organic and conventionally managed soils from the same sampling site clustered together indicating that there were no significant differences in bacterial community structure in those soils. Dissimilarity analysis (Table 2) revealed that bacterial community structures significantly differed from each other based only on sampling sites, while soil management showed no effect on overall bacterial community structure across all soils.

In order to identify which factors influenced bacterial community structure in soils, CCA was performed. CCA results showed that soil properties, including EC, pH, WHC, sand fraction, TN, and organic matter contents were major parameters affecting soil bacterial community structures (Fig. 5). Furthermore, partial CCA and variation partitioning analysis revealed that soils physical properties and soil chemical properties explained 16.3% and 12.5% of total variation in bacterial

**Table 2**
Dissimilarity analyses of bacterial community structures in soils from different sites and in soils subjected to different soil managements. The distance metrics used were, Bray, Horn, and Euclidean. Three distance indices were calculated, including mrpp, adonis and anosim. Bolded numbers indicate *P* values were significant at 0.05 level.

| | | mrpp | | Adonis | | Anosim | |
|---|---|---|---|---|---|---|---|
| | | *δ* | *P* | *F* | *P* | *R* | *P* |
| Sites | Bray | 0.810 | **0.001** | 3.221 | **0.001** | 0.404 | **0.001** |
| | Horn | 0.680 | **0.001** | 5.591 | **0.001** | 0.387 | **0.001** |
| | Euclidean | 624.7 | **0.001** | 2.551 | **0.005** | 0.086 | **0.048** |
| Soil management | Bray | 0.868 | 0.150 | 1.184 | 0.169 | 0.030 | 0.200 |
| | Horn | 0.787 | 0.075 | 1.628 | 0.091 | 0.060 | 0.078 |
| | Euclidean | 664.3 | 0.143 | 1.825 | 0.078 | 0.007 | 0.282 |

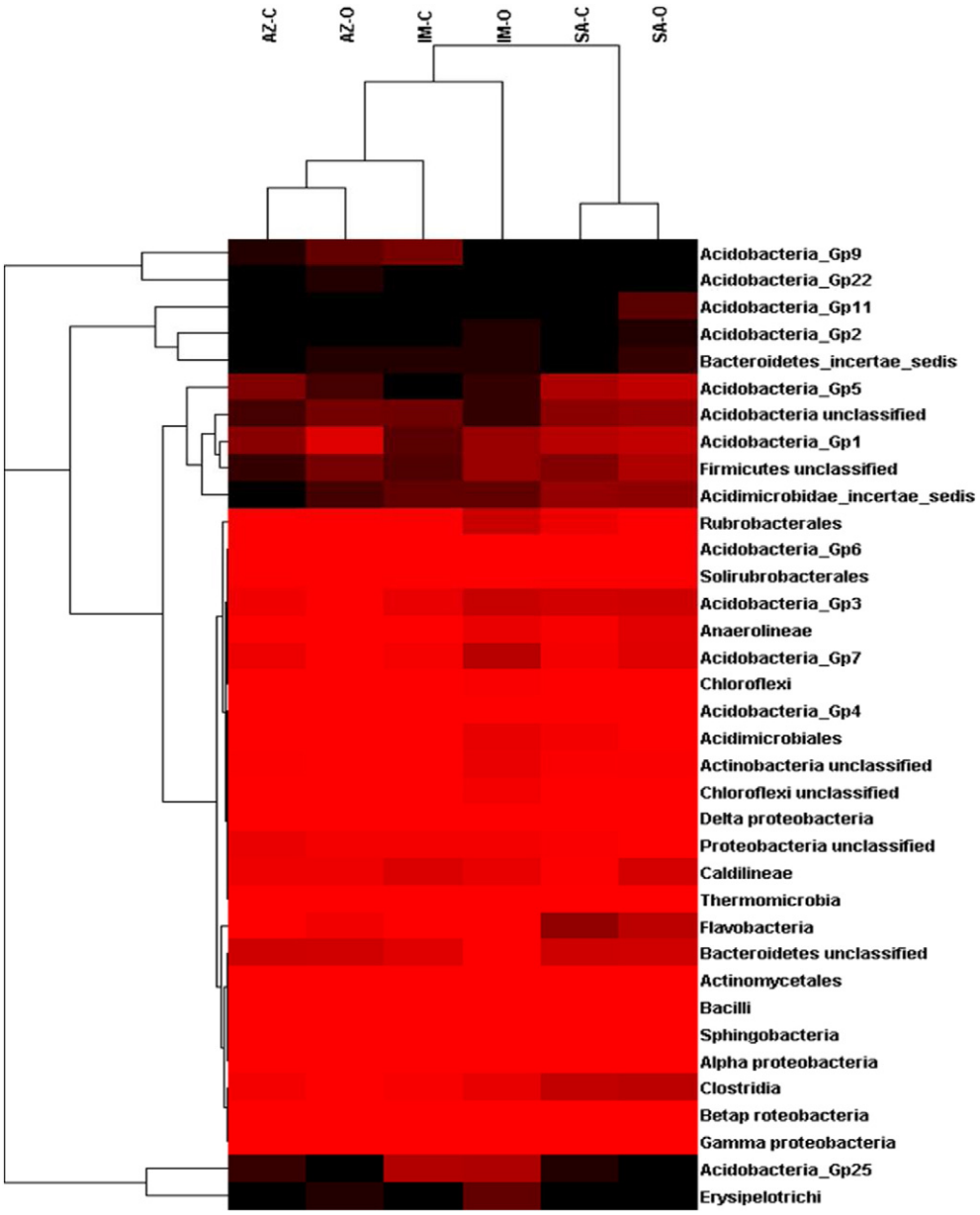



**Fig. 4.** Clustering analysis of bacterial community structure at class level in soils from Yuma (AZ), Imperial Valley (IM), and Salinas Valley (SA). Both conventionally (C) and organically (O) managed soils were included. The bacterial class names are presented. Red indicates abundance above 0, whereas black indicates signal intensities equals to 0. Brighter red coloring indicates higher abundance.

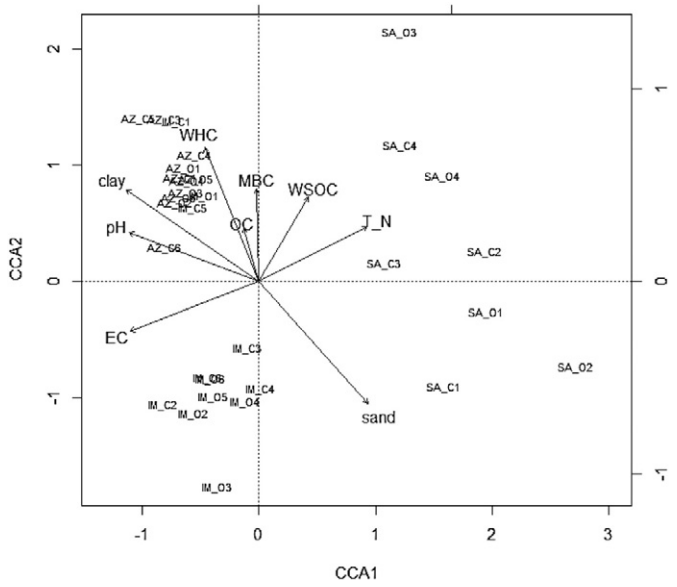



**Fig. 5.** Canonical correspondence analyses (CCA) of bacterial community structure and soil variables including EC salinity (dS $m^{-1}$), water holding capacity (WHC, %), sand content (%), clay content (%), total nitrogen (TN, %); organic carbon (OC %); water soluble organic carbon in soil water extract (WSOC, mg $kg^{-1}$); microbial biomass carbon (MBC mg $kg^{-1}$), and pH. The organically managed (O) and conventionally managed (C) soil samples from Yuma, Arizona (AZ), Imperial Valley, California (IM), and Salinas Valley, California (SA) were shown.

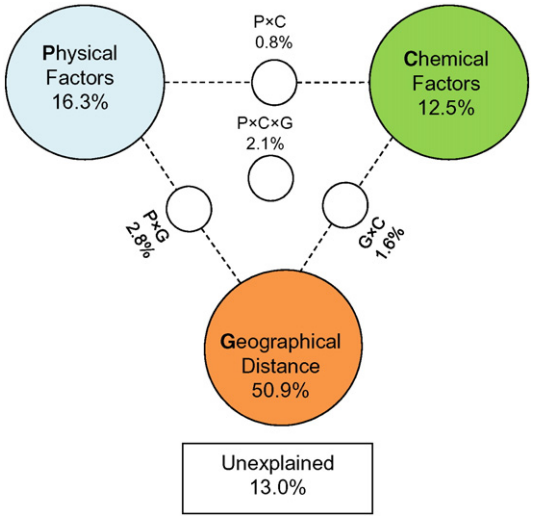



**Fig. 6.** Variation partitioning analysis (VPA) of microbial community structure explained (%) by soil physics (P), soil chemistry (C), and geographic distance (G).

community structure, while location explained >50% of overall variation of bacterial community, leaving 13% of total variation unexplained (Fig. 6).

### 3.3. Artificial neural network analysis

Artificial neural network analysis was applied to correlate soil properties with the bacterial community composition and diversity. By comparing image patterns of soil physiochemical properties, soil management, and soil bacterial ecology data, the network analysis could be trained to predict relationships between different factors, e.g. whether soil management exerted influence on abundance of bacterial phylum. The results of Kohonen self organizing map showed that pH, clay fraction, EC, WSOC, and total N were major factors that predicted bacterial community structures (Fig. 7 A&B). A sensitivity diagram (Fig. 7B) for the microbial community composition shows that EC, pH, WHC, and OC were major factors influencing community composition with EC as the strongest. Predictive models showed the relationships between the environmental variables and changes in the microbial community biomass and bacterial community taxon composition (Fig. 8a–c). For the abundances of subclasses of *Proteobacteria*, an elevated pH may result in the increase in *Alphaproteobacteria* abundances and decrease of the abundances of *Delta-* and *Gammaproteobacteria* (Fig. 8d-f).

### 3.4. Linear regression analysis of soil texture and Shannon diversity index and total OTU count

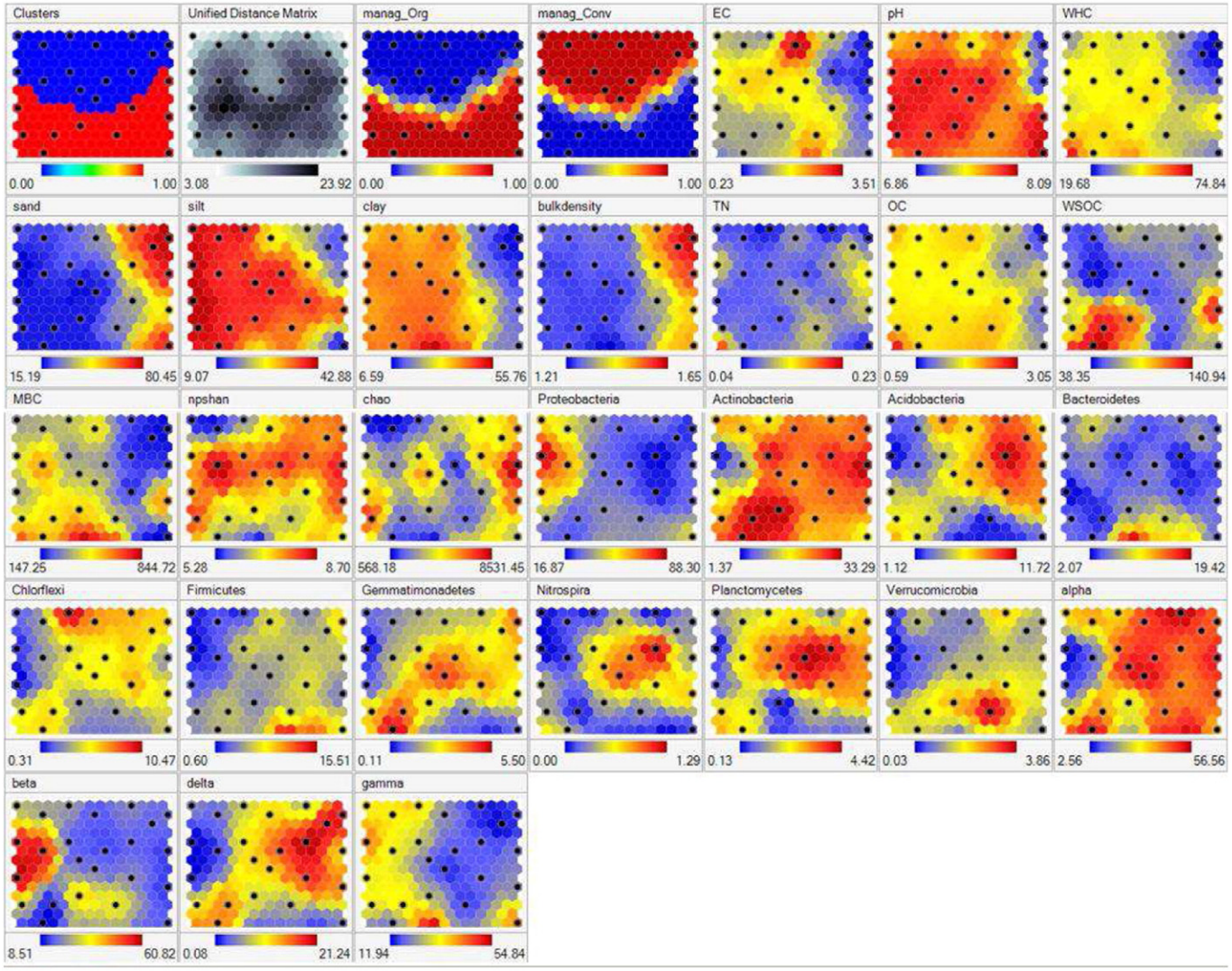
Linear regression analysis showed that there were strong correlations ($P < 0.01$) between soil texture (sand and clay fraction) and bacterial community structure (Shannon diversity and total OTU count). Sand content positively correlated with Shannon diversity and total OTU count, while clay displayed a reverse trend, i.e. negatively correlated with Shannon diversity index and total OTU count (Fig. 9).

## 4. Discussion

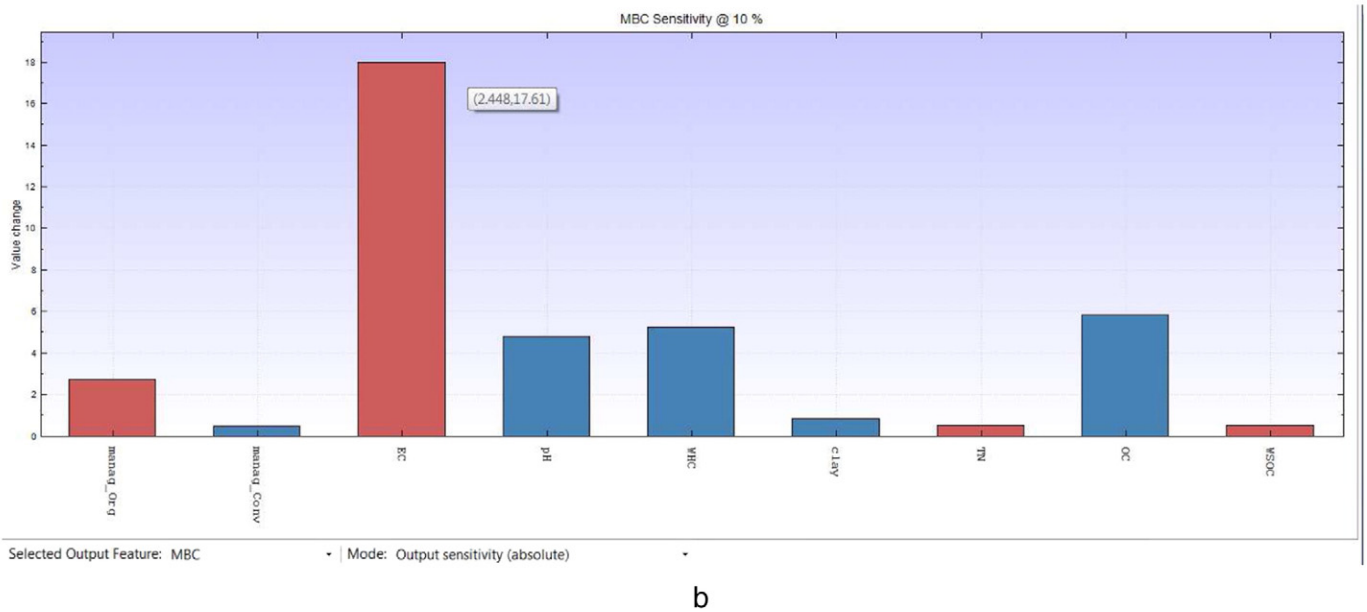
We observed that pH was one of the major factors influencing bacterial community structures in major leafy green producing areas of California and Arizona. Our result is consistent with previous studies where pH was the best predictor of bacterial composition and diversity (Fierer and Jackson, 2006, Lauber, et al., 2009, Rousk, et al., 2010, Griffiths, et al., 2011and Xiong, et al., 2012). Soil pH may indirectly affect bacterial community structure by changing other soil physiochemical properties, including nutrient availability, cationic metal solubility, organic carbon characteristics, soil moisture regime, and electrical conductivity, which might exert a more direct influence on bacterial community structure. Soil pH might also directly stress and select for different soil bacteria taxa. Soil microbes that are more sensitive to pH change might die off faster than those more tolerant of pH changes. The optimal intracellular pH levels of many soil bacteria are close to 7.0 (Langenheder, et al., 2003). Extreme pH values may impose a significant stress to certain taxa while others may have higher tolerance (Lauber, et al., 2009). Overall, our conclusion that pH may affect the bacterial community is largely in agreement with previous reports (e.g. Fierer and Jackson, 2006 and Lauber, et al., 2009), however, it should be noted that the pH range in our current study was very narrow, and such results should be interpreted with caution.

Previous investigations of EC effects on microbial community composition (Casamayor, et al., 2002, Langenheder, et al., 2003, Henriques, et al., 2006 and Wu, et al., 2006) have shown that EC is a major factor controlling microbial abundance, diversity, composition, and functions. In contrast to northern California, southern California and Arizona areas are dominated by a drier climate resulting in an increased EC due to decreased mean annual precipitation and increased mean annual temperature and use of higher salinity water for irrigation. There are two strategies that microbial communities adapt to during changes of EC; replacement of taxa that are more tolerant of the increased EC, as well

**Fig. 7.** (A): Kohonen self-organizing map of microbial community structure and associated soil variables for agricultural soils from major fresh produce producing soils. Org and Con represent organically and conventionally managed soils, respectively. EC salinity (dS $m^{-1}$), water holding capacity (WHC, %), silt and clay fraction (%), total nitrogen (TN, %); organic carbon (OC %); water soluble organic carbon in soil water extract (WSOC, mg $kg^{-1}$); microbial biomass carbon (MBC mg $kg^{-1}$); npshan represent Shannon and Chao biodiversity index, respectively. Alpha, beta, delta, and gamma indicate the relative abundances (%) of the subclasses of proteobacteria; (B) Predicted effects of soil variables on microbial community composition by artificial neural network modeling. ANN model 95% confidence interval ± 110 mg $kg^{-1}$ soil.


Clusters
0.00 1.00
Unified Distance Matrix
3.08 23.92
manag_Org
0.00 1.00
manag_Conv
0.00 1.00
EC
0.23 3.51
pH
6.86 8.09
WHC
19.68 74.84
sand
15.19 80.45
silt
9.07 42.88
clay
6.59 55.76
bulkdensity
1.21 1.65
TN
0.04 0.23
OC
0.59 3.05
WSOC
38.35 140.94
MBC
147.25 844.72
npshan
5.28 8.70
chao
568.18 8531.45
Proteobacteria
16.87 88.30
Actinobacteria
1.37 33.29
Acidobacteria
1.12 11.72
Bacteroidetes
2.07 19.42
Chlorflexi
0.31 10.47
Firmicutes
0.60 15.51
Gemmatimonadetes
0.11 5.50
Nitrospira
0.00 1.29
Planctomycetes
0.13 4.42
Verrucomicrobia
0.03 3.86
alpha
2.56 56.56
beta
8.51 60.82
delta
0.08 21.24
gamma
11.94 54.84

a


MBC Sensitivity @ 10 %
(2.448,17.61)
Value change
0
2
4
6
8
10
12
14
16
18
manag_Org
manag_Conv
EC
pH
WHC
clay
TN
OC
WSOC
Selected Output Feature: MBC
Mode: Output sensitivity (absolute)

b

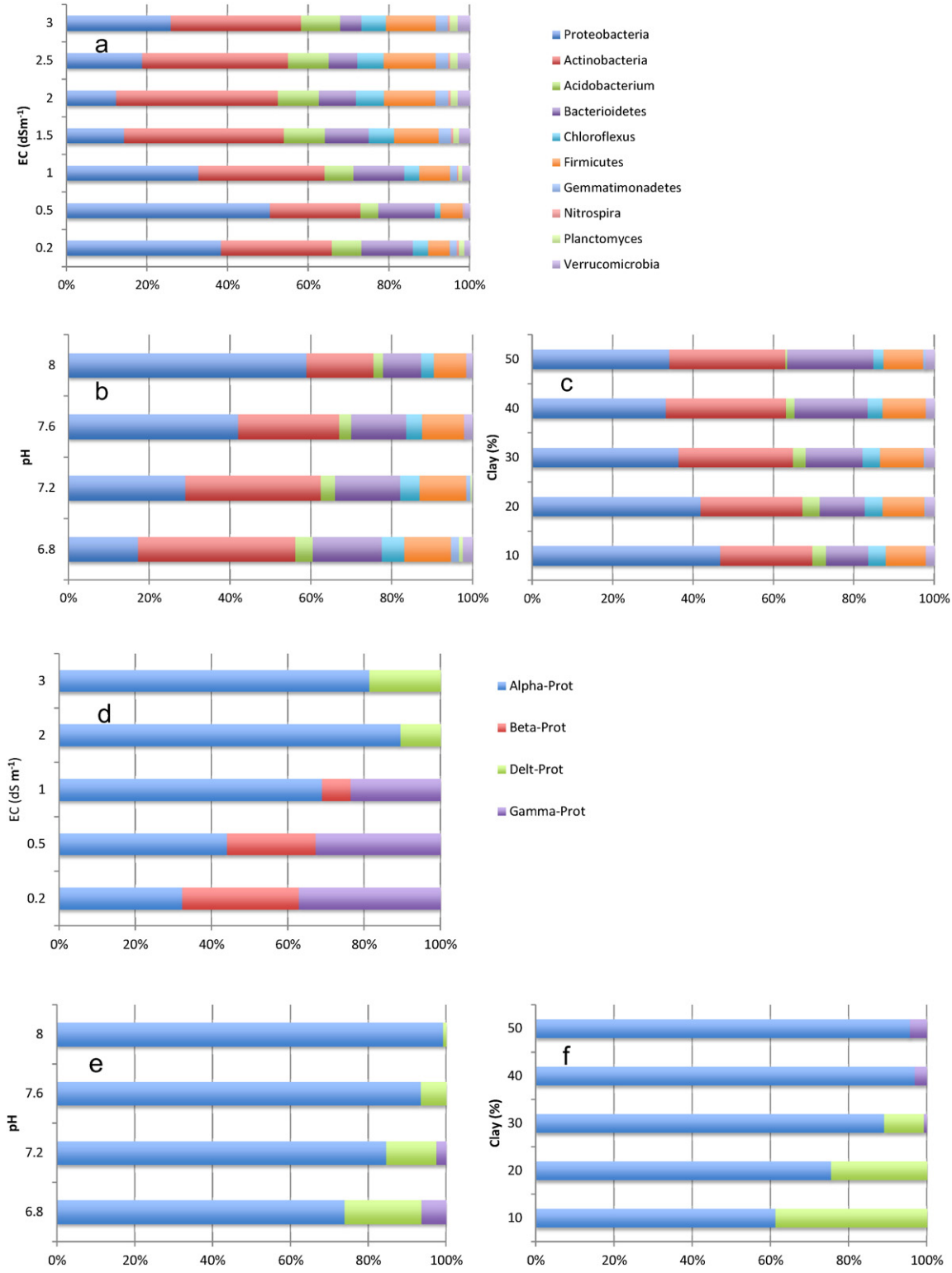



**Fig. 8.** Artificial neural network model predictions of the effects of EC, clay, and pH on changes of overall bacterial community structure (Fig. 8a–c) and proteobacterial community structure (Fig.8d–f). EC (dS $m^{-1}$).

as slow adaptation to the increased EC of the same taxa (Wu, et al., 2006).

In the current study, it was found that soil texture class was one of the factors influencing bacterial community structures across all the soil tested. The phenomenon could be explained by the interaction between particle size fractions as discussed in a previous publication (Chau, et al., 2011). In most soils, the silt and clay particles determine the water holding capacity, and more importantly, provide binding sites for organic carbon and trace elements required by bacteria to flourish. In addition, the clay and silt particle layers offer protection from desiccation, gas diffusion, toxic exogenous compounds, and predation by protozoa (Ranjard and Richaume, 2001). Spatial isolation hypothesis (Zhou, et al., 2002) might be applied here to further explain the close correlation between soil sand content and Shannon diversity index

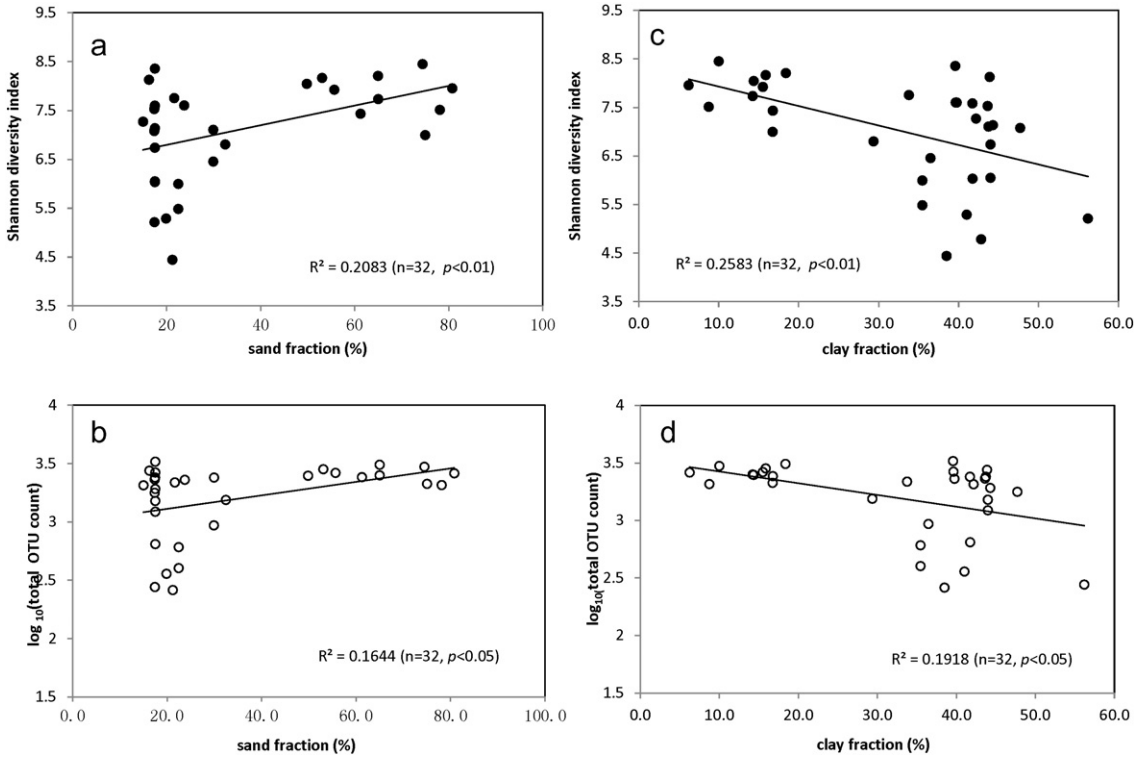



**Fig. 9.** Linear regression analyses of sand content and Shannon biodiversity diversity index (Fig.9a) and total OTU observed (Fig. 9b); and linear regression analysis of clay content and Shannon biodiversity diversity index (Fig.9c) and total OTU observed (Fig. 9d).

and total observed OTU counts. According to this theory, a coarse soil exhibit larger pores, and water is held in pore corners as isolated water films, which provide opportunities for increased bacterial diversity (Zhou, et al., 2002). Therefore, compared to a soil with high clay content, increased sand content may result in a higher number of isolated hydrated microhabitats. On the other hand, high clay content is associated with greater aggregate stability, increased protection of particulate and dissolved organic matter, and creation of protected microsites that are not accessible to predatory nematodes and protozoa. Indeed, on a microscale, the heterogeneous distribution of soil microbes is mainly determined by soil structure, porosity, and organic carbon content (Ranjard, et al., 2010). Other studies of bacterial communities in soils and in microcosms showed that hydraulically-induced spatial isolation in drier soils results in a higher diversity compared to wetter, hydraulically connected soils (Zhou, et al., 2002, Treves, et al., 2003). A larger number of hydrated isolated microhabitats available for bacterial colonization may promote the coexistence of multiple microbial species in close proximity without directly competing for nutrients (Zhou, et al., 2002, Treves, et al., 2003 and Carson, et al., 2010). In addition to the spatial isolation hypothesis, the effects of organic matter existed in sandy soils cannot be neglected, since organic matter is a major factor affecting the capability of a soils to hold moisture which is crucial for microorganisms to thrive in soils (Li, et al. 2004).

Organic management practices commonly lead to increased soil microbial biomass, increased microbial activity, and increased microbial species richness and diversity when compared to conventional farming (Mader, et al., 2002 and van Diepeningen, et al., 2006). While in our current study, the difference in bacterial community structures were not statistically significant at $P < 0.05$, which is consistent with a previous study where the same trend was also found between organically and conventionally managed agricultural soils (Semenov, et al., 2008; Ma et al., 2012b). Similarly, no differences were found between organic and conventional farming agricultural soils in bacterial community richness and diversity (Lopes, et al., 2011). Data presented in the current study as well as others suggest that differences in agricultural management may not be well reflected in bacterial diversity indices, but in overall bacterial community structure (Hartmann and Widmer, 2006 and Lopes, et al., 2011). The high similarity in soil properties between organic and conventional soils (Table 1) may also explain that no major differences in bacterial community structure were found in the current study. In addition, relatively small numbers of OTUs were found to be unique to organic and conventional soils, respectively (Fig. 1). Among those unique OTUs, indicator species may be identified that could better reflect the changes due to different farming practices.

Our data showed that geographic location could explain the majority of microbial community structure variations. This is in line with previous reports showing longer distance soils share fewer similarities indicating distantly sampled soils share fewer species in common (Fulthorpe, et al., 2008). Recent studies investigating the biogeography of bacteria across larger scale concluded that bacterial community structures do vary with the increase of geographical distance (Griffiths, et al., 2011 and Lopes, et al., 2011). Based on the data published to date, it is clear that geographic location is a major factor shaping microbial community structure. In our current study, we focused on agricultural soils, which also confirm the same result. Organic and conventional management may exert some influence on bacterial community structure, but the overall effect may not be big enough compared to the changes brought by geographic distances. It should be noted that due to the geographical locations, the soils may experience different climatic conditions, e.g. mean annual precipitation and mean annual temperature, which in turn may change the soil physiochemical properties. Therefore, the contribution to variation in bacterial community structure by soil physiochemical properties might be overestimated. During our study, 13% of total variation in bacterial community structure was left unexplained, and this could be due to unmeasured environmental variables such as oxygen content, redox potentials, water content, and trace elements, which may play an influential role in shaping microbial communities in leafy green producing soils. Moreover, sampling effects

and ecologically neutral processes of diversification may also contribute to the unexplained portion of microbial community variation (Ramette and Tiedje, 2007). Some key physical and chemical properties of our soils were able to predict the influence of some of these factors in our soils (Fig. 7 and Fig.8).

## 5. Conclusions

In summary, this study showed that both soil physiochemical properties and geographic location have a major influence on bacterial community structures in leafy green production soils from California and Arizona. A better understanding of diversity and composition of microbial communities and their controlling factors may enable farmers to evaluate the soil quality and adjust their soil management strategies to boost microbial diversity and abundance toward a healthy and sustainable soil ecology system.

## Acknowledgements

This research was supported by CSREES NIFA Agreement (No. 2008-35201-18709) and the 206 Manure and Byproduct Utilization Project of the USDA-ARS. The research was also partly financed by National Natural Science Foundation of China (No. 41571304). We thank Drs Jorge Fonseca of the University of Arizona Yuma, Mark Trent, UC-Davis, Imperial Agricultural Experiment Station, and James McCreight of USDA-ARS Salinas CA for providing soil samples for this study. We also thank Damon Baptista for technical help. Mention of trademark or propriety products in this manuscript does not constitute a guarantee or warranty of the property by the USDA and does not imply its approval to the exclusion of other products that may also be suitable.

## References

Acosta-Martinez, V., Dowd, S., Sun, Y., Allen, V., 2008. Tag-encoded pyrosequencing analysis of bacterial diversity in a single soil type as affected by management and land use. Soil Biol. Biochem. 40, 2762–2770.
Baker, K.L., Langenheder, S., Nicol, G.W., Ricketts, D., Killham, K., Campbell, C.D., Prosser, J.I., 2009. Environmental and spatial characterisation of bacterial community composition in soil to inform sampling strategies. Soil Biol. Biochem. 41, 2292–2298.
Bardgett, R.D., Freeman, C., Ostle, N.J., 2008. Microbial contributions to climate change through carbon cycle feedbacks. ISME J. 2, 805–814.
Bending, G.D., Putland, C., Rayns, F., 2000. Changes in microbial community metabolism and labile organic matter fractions as early indicators of the impact of management on soil biological quality. Biol. Fertil. Soils 31, 78–84.
Carson, J.K., Gonzalez-Quinones, V., Murphy, D.V., Hinz, C., Shaw, J.A., Gleeson, D.B., 2010. Low pore connectivity increases bacterial diversity in soil. Appl. Environ. Microbiol. 76, 3936–3942.
Casamayor, E.O., Massana, R., Benlloch, S., et al., 2002. Changes in archaeal, bacterial and eukaryal assemblages along a salinity gradient by comparison of genetic fingerprinting methods in a multipond solar saltern. Environ. Microbiol. 4, 338–348.
Champely, S., Chessel, D., 2002. Measuring biological diversity using Euclidean metrics. Environ. Ecol. Stat. 9, 167–177.
Chau, J.F., Bagtzoglou, A.C., Willig, M.R., 2011. The effect of soil texture on richness and diversity of bacterial communities. Environ. Forensic 12, 333–341.
Claesson, M.,J., O'Sullivan, O., Wang, Q., et al., 2009. Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. PLoS One 4, e6669.
Cole, J.R., Wang, Q., Cardenas, E., et al., 2009. The ribosomal database project: improved alignments and new tools for rRNA analysis. Nucleic Acids Res. 37, D141–D145.
Cooley, M., Carychao, D., Crawford-Miksza, L., et al., 2007. Incidence and tracking of *Escherichia coli* O157:H7 in a major produce production region in California. PLoS One 2, e1159.
Dequiedt, S., Thioulouse, J., Jolivet, C., et al., 2009. Biogeographical patterns of soil bacterial communities. Environ. Microbiol. Rep. 1, 251–255.
Dethlefsen, L., Huse, S., Sogin, M.L., 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, 2383–2400.
Dowd, S.F., Sun, Y., Wolcott, R.D., Domingo, A., Carroll, J.A., 2008. Bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP) for microbiome studies: bacterial diversity in the ileum of newly weaned salmonella-infected pigs. Foodborne Path Dis. 5, 459–472.
Drenovsky, R.E., Steenwerth, K.L., Jackson, L.E., Scow, K.M., 2010. Land use and climatic factors structure regional patterns in soil microbial communities. Glob. Ecol. Biogeogr. 19, 27–39.
Fierer, N., Jackson, R.B., 2006. The diversity and biogeography of soil bacterial communities. PNAS 103, 626–631.
Fulthorpe, R.R., Roesch, L.F.W., Riva, A., Triplett, E.W., 2008. Distantly sampled soils carry few species in common. ISME J. 2, 901–910.
Gardi, C., Montanarella, L., Arrouays, D., et al., 2009. Soil biodiversity monitoring in Europe: ongoing activities and challenges. Eur. J. Soil Sci. 60, 807–819.
Girvan, M.S., Campbell, C.D., Killham, K., Prosser, J.I., Glover, L.A., 2005. Bacterial diversity promotes community stability and functional resilience after perturbation. Environ. Microbiol. **7**, 301–313.
Gorski, L., Parker, C.T., Liang, A., et al., 2011. Prevalence, distribution, and diversity of *Salmonella enterica* in a major produce region of California. Appl. Environ. Microbiol. 77, 2734–2748.
Griffiths, R.I., Thomson, B.C., James, P., Bell, T., Bailey, M., Whiteley, A.S., 2011. The bacterial biogeography of British soils. Environ. Microbiol. 13, 1642–1654.
Hartmann, M., Widmer, F., 2006. Community structure analyses are more sensitive to differences in soil bacterial communities than anonymous diversity indices. Appl. Environ. Microbiol. 72, 7804–7812.
Henriques, I.S., Alves, A., Tacao, M., Almeida, A., Cunha, A., Correia, A., 2006. Seasonal and spatial variability of free-living bacterial community composition along an estuarine gradient (Ria de Aveiro, Portugal). Estuar. Coast. Shelf Sci. 68, 139–148.
Ibekwe, A.M., Kennedy, A.C., 1998. Phospholipid fatty acid and carbon utilization patterns for analysis of microbial community structure under field and greenhouse conditions. FEMS Microbiol. Ecol. 26, 151–163.
Ibekwe, A.M., Kennedy, A.C., 1999. Fatty acid methyl ester (FAME) profile as a tool to investigate community structure of two agricultural soils. Plant Soil 206, 151–161.
Ibekwe, A.M., Papiernik, S.K., Gan, J., Yates, S.R., Yang, C.-H., Crowley, D.E., 2001. Impact of fumigants on soil microbial community. Appl. Environ. Microbiol. 67, 3245–3257.
Ibekwe, A.M., Kennedy, A.C., Frohne, P.S., Papiernik, S.K., Yang, C.-H., Crowley, D.E., 2002. Microbial diversity along a transect of agronomic zones. FEMS Microbiol. Ecol. 39, 186–191.
Jones, R.T., Robeson, M.S., Lauber, C.L., Hamady, M., Knight, R., Fierer, N., 2009. A comprehensive survey of soil acidobacterial diversity using pyrosequencing and clone library analyses. ISME J. 3, 442–453.
Kennedy, A.C., 1999. Bacterial diversity in agroecosystems. Agric. Ecosyst. Environ. 74, 65–76.
Klute, A., 1996. Methods of Soil Analysis: Physical and Mineralogical Methods. American Society of Agronomy, Madison WI, p. 1188.
Kozdroj, J., van Elsas, J.D., 2001. Structural diversity of microbial communities in arable soils of a heavily industrialised area determined by PCR-DGGE fingerprinting and FAME profiling. Appl. Soil Ecol. 17, 31–42.
Langenheder, S., Kisand, V., Wikner, J., Tranvik, L.J., 2003. Salinity as a structuring factor for the composition and performance of bacterioplankton degrading riverine DOC. FEMS Microbiol. Ecol. 45, 189–202.
Lauber, C.L., Hamady, M., Knight, R., Fierer, N., 2009. Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. Appl. Environ. Microbiol. 75, 5111–5120.
Li, H., Parent, L.E., Karam, A., Tremblay, C., 2004. Potential of sphagnum peat for improving soil organic matter, water holding capacity, bulk density and potato yield in a sandy soil. Plant Soil 265, 355–365.
Liu, B., Tu, C., Hu, S., Gumpertz, M., Ristaino, J.B., 2007. Effect of organic, sustainable, and conventional management strategies in grower fields on soil physical, chemical, and biological factors and the incidence of Southern blight. Appl. Soil Ecol. 37, 202–214.
Lopes, A.R., Faria, C., Prieto-Fernandez, A., Trasar-Cepeda, C., Manaia, C.M., Nunes, O.C., 2011. Comparative study of the microbial diversity of bulk paddy soil of two rice fields subjected to organic and conventional farming. Soil Biol. Biochem. 43, 115–125.
Lueke, C., Krause, S., Cavigiolo, S., Greppi, D., Lupotto, E., Frenzel, P., 2010. Biogeography of wetland rice methanotrophs. Environ. Microbiol. 12, 862–872.
Ma, J., Ibekwe, A.M., Leddy, M., Yang, C.-H., Crowley, D.E., 2012a. Assimilable organic carbon (AOC) in soil water extracts using *Vibrio harveyi* BB721 and its implication for microbial biomass. PLoS One 7 (5), e28519. http://dx.doi.org/10.1371/journal.pone.0028519.
Ma, J., Ibekwe, A.M., Crowley, D.E., Yang, C.-H., 2012b. Persistence of *Escherichia coli* O157:H7 in major leafy green producing soils. Environ. Sci. Technol. 46, 12154–12161.
Ma, J., Ibekwe, A.M., Yang, C.-H., Crowley, D.E., 2013. Influence of bacterial communities based on 454-pyrosequencing on the survival of *Escherichia coli* O157:H7 in soils. FEMS Microbiol. Ecol. 84, 542–554.
Ma, J., Ibekwe, A.M., Crowley, D.E., Yang, C.-H., 2014. Persistence of *Escherichia coli* O157 and non-O157 strains in agricultural soils. Sci. Total Environ. 490, 822–829.
Mader, P., Fliessbach, A., Dubois, D., Gunst, L., Fried, P., Niggli, U., 2002. Soil fertility and biodiversity in organic farming. Science 296, 1694–1697.
Paula, F.S., Rodrigues, J.L.M., Zhou, J.Z., et al., 2014. Land use change alters functional gene diversity, composition and abundance in Amazon forest soil microbial communities. Mol. Ecol. 23, 2988–2999.
Ramette, A., Tiedje, J.M., 2007. Multiscale responses of microbial life to spatial distance and environmental heterogeneity in a patchy ecosystem. PNAS 104, 2761–2766.
Ranjard, L., Richaume, A.S., 2001. Quantitative and qualitative microscale distribution of bacteria in soil. Res. Microbiol. 152, 707–716.
Ranjard, L., Dequiedt, S., Jolivet, C., et al., 2010. Biogeography of soil microbial communities: a review and a description of the ongoing French national initiative. Agron. Sustain. Dev. 30, 359–365.
Roesch, L.F., Fulthorpe, R.R., Riva, A., et al., 2007. Pyrosequencing enumerates and contrasts soil microbial diversity. ISME J. 1, 283–290.
Rousk, J., Baath, E., Brookes, P.C., et al., 2010. Soil bacterial and fungal communities across a pH gradient in an arable soil. ISME J 4, 1340–1351.
Schloss, P.D., Westcott, S.L., Ryabin, T., et al., 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol. 75, 7537–7541.
Semenov, A.V., Franz, E., van Overbeek, L., Termorshuizen, A.J., van Bruggen, A.H.C., 2008. Estimating the stability of *Escherichia coli* O157: H7 survival in manure-amended soils with different management histories. Environ. Microbiol. 10, 1450–1459.

Shendure, J., Ji, H.L., 2008. Next-generation DNA sequencing. Nature Biotech. 26, 1135–1145.
Taylor, E.V., Nguyen, T.A., Machesky, K.D., et al., 2013. Multistate outbreak of *Escherichia coli* O145 infections associated with romaine lettuce consumption. J. Food Prot. 76, 939–944.
Torsvik, V., Ovreas, L., Thingstad, T.F., 2002. Prokaryotic diversity — magnitude, dynamics, and controlling factors. Science 296, 1064–1066.
Treves, D.S., Xia, B., Zhou, J., Tiedje, J.M., 2003. A two-species test of the hypothesis that spatial isolation influences microbial diversity in soil. Microb. Ecol. 45, 20–28.
van der Heijden, M.G.A., Wagg, C., 2013. Soil microbial diversity and agro-ecosystem functioning. Plant Soil 363, 1–5.
van Diepeningen, A.D., de Vos, O.J., Korthals, G.W., van Bruggen, A.H.C., 2006. Effects of organic versus conventional management on chemical and biological parameters in agricultural soils. Appl. Soil Ecol. 31, 120–135.
van Elsas, J.D., Semenov, A.V., Costa, R., Trevors, J.T., 2011. Survival of *Escherichia coli* in the environment: fundamental and public health aspects. ISME J. 5, 173–183.
Vance, E.D., Brookes, P.C., Jenkinson, D.S., 1987. An extraction method for measuring soil microbial biomass C. Soil Biol. Biochem. 19 (5), 703–707.
Wang, Y., Qian, P.Y., 2009. Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. PLoS One http://dx.doi.org/10.1371/journal.pone.0007401.
Wu, Q.L., Zwart, G., Schauer, M., Kamst-van Agterveld, M.P., Hahn, M.W., 2006. Bacterioplankton community composition along a salinity gradient of sixteen high-mountain lakes located on the Tibetan Plateau, China. Appl. Environ. Microbiol. 72, 5478–5485.
Wu, Y., Ma, B., Zhou, L., Wang, H., Xu, J., Kemmitt, S., Brookes, P.C., 2009. Changes in the soil microbial community structure with latitude in eastern China, based on phospholipid fatty acid analysis. Appl. Soil Ecol. 43, 234–240.
Xiong, J.B., Liu, Y.Q., Lin, X.G., et al., 2012. Geographic distance and pH drive bacterial distribution in alkaline lake sediments across Tibetan Plateau. Environ. Microbiol. 14, 2457–2466.
Yue, J.C., Clayton, M.K., 2005. A similarity measure based on species proportions. Com Stat-Theory Methods. 34, 2123–2131.
Zhou, J.Z., Xia, B.C., Treves, D.S., et al., 2002. Spatial and resource factors influencing high microbial diversity in soil. Appl. Environ. Microbiol. 68, 326–334.