

# Influence of bacterial communities based on 454-pyrosequencing on the survival of *Escherichia coli* O157:H7 in soils

Jincai Ma<sup>1,3</sup>, Abasiofiok M. Ibekwe<sup>1</sup>, Ching-Hong Yang<sup>2</sup> & David E. Crowley<sup>3</sup>

<sup>1</sup>USDA-ARS U. S. Salinity Laboratory, Riverside, CA, USA; <sup>2</sup>Department of Biological Sciences, University of Wisconsin, Milwaukee, WI, USA; and

<sup>3</sup>Department of Environmental Sciences, University of California, Riverside, CA, USA

**Correspondence:** Abasiofiok M. Ibekwe, USDA-ARS-U. S. Salinity Laboratory, 450 W. Big Springs Rd, Riverside, CA 92507, USA. Tel.: +1 951 369 4828; fax: +1 951 342 4964; e-mail: Mark.Ibekwe@ars.usda.gov

Received 28 September 2012; revised 14 January 2013; accepted 23 January 2013. Final version published online 26 February 2013.

DOI: 10.1111/1574-6941.12083

Editor: Julian Marchesi

## Keywords

contamination; fresh produce; organic; conventional.

## Abstract

Shiga toxin-producing *Escherichia coli* O157:H7 has been implicated in many foodborne illnesses. In this study, survival of *E. coli* O157:H7 in 32 soils from California (CA) and Arizona (AZ) was investigated. Our goal was to correlate the survival time of *E. coli* O157:H7 in soils with 16S rRNA pyrosequencing based bacterial community composition. Kohonen self-organizing map of survival and associated soil chemical, physical and biological variables using artificial neural network analysis showed that survival of *E. coli* O157:H7 in soils was negatively correlated with salinity (EC), but positively correlated with total nitrogen (TN) and water soluble organic carbon (WSOC). Bacterial diversity as determined by the Shannon diversity index had no significant ( $P = 0.635$ ) effect on *ttd*, but individual bacterial phyla had different effects. The survival of *E. coli* O157:H7 was positively correlated with the abundances of *Actinobacteria* ( $P < 0.001$ ) and *Acidobacteria* ( $P < 0.05$ ), and negatively correlated with those of *Proteobacteria* and *Bacteroidetes* ( $P < 0.05$ ). Our data showed that specific groups of bacteria correlate with the persistence of *E. coli* O157:H7 in soils thus opening new ways to study the influence of certain bacterial phyla on persistence of this pathogen and other related pathogens in complex environments.

## Introduction

Numerous studies have been carried out to investigate the fate of *Escherichia coli* O157:H7 in different agricultural systems with major focus on the survival of *E. coli* O157:H7 in soils and manure-amended soils (Jiang *et al.*, 2002; Franz *et al.*, 2008; Semenov *et al.*, 2008). One of the most recent review articles (van Elsas *et al.*, 2011) on survival of *E. coli* O157:H7 showed that temperature, soil structure, and microbial communities were the most important factors affecting survival (van Elsas *et al.*, 2011). These authors showed from their previous studies (van Elsas *et al.*, 2007) that the survival of *E. coli* O157:H7 was inversely proportional to the diversity of the microbial community established through differential fumigation and regrowth activities. Using polymerase chain reaction–denaturing gradient gel electrophoresis (PCR-DGGE), Franz *et al.* (2008) and Semenov *et al.*

(2008) showed that the changes in microbial diversities and community compositions coincided with an enhancement of the survival rate of the invading pathogen. The basic explanation presented by van Elsas *et al.* (2011) was that lowering of the complexity of the soil microbiota probably resulted in a reduction in functional redundancy, which enhances the chances for the introduced microorganisms to occupy a niche in the system and persist as a member of the microbial community. They concluded that some organisms could be direct competitors by occupying the same niche, whereas others might have antagonistic or predatory activities. However, these authors lamented the lack of data on impact of different functional bacterial groups on survival of *E. coli* O157:H7.

Depending on the soil physical, chemical, and biological characteristics and environmental factors (e.g. temperature and precipitation), the survival time of *E. coli* O157:H7 in soils varies from 1 week to 6 months, and

even longer in some extreme cases (Maule, 2000; Mubiru *et al.*, 2000; Ibekwe *et al.*, 2007, 2010; Ibekwe & Ma, 2011; Ma *et al.*, 2011). It would be interesting to investigate the role of specific environmental variables, particularly functional bacterial groups, on the overall survival of *E. coli* O157:H7 using traditional microbial ecology approaches (e.g. PCR-DGGE), next-generation sequencing techniques (e.g. 454-pyrosequencing and Illumina), and microarray protocol (e.g. PhyloChip). The low sequencing depth of the traditional PCR cloning approach when compared with the vast genetic diversity present in soils hindered a comprehensive characterization of the microbial community structure. In this aspect, the current community analyses typically represent a mere snapshot of the dominant members, with little information on taxa with medium to low abundances. High-throughput sequencing is a promising method, as it provides enough sequencing depth to cover the complex microbial communities (Shendure & Ji, 2008). For this reasons, we believe that the correlation of different bacterial phyla based on deep sequencing may provide some insights into some of the bacterial groups that may affect the persistence of *E. coli* O157:H7 in soil.

In this study, the survival of Shiga toxin-producing *E. coli* O157:H7 strain EDL933 in 32 soils from three major leafy green-producing areas of CA and AZ was investigated. The majority of *E. coli* O157:H7 outbreaks have been associated with Salinas Valley area in CA, where most of the summer produce is grown compared to Yuma, AZ, and Imperial Valley, CA, where winter produce is grown. Our work expands on the above studies using 454-pyrosequencing-based technology to address the role of different bacterial phyla on the survival of this pathogen. In the current study, the survival data were analyzed using the Weibull model (Mafart *et al.*, 2002; Franz *et al.*, 2008) and artificial neural network analysis using Kohonen self-organizing map of *E. coli* survival, and the data were correlated with soil bacterial community composition and soil properties.

## Materials and methods

### Soil collection and characterization

Soil samples were collected from three major fresh produce-growing areas: Salinas Valley area CA, Imperial Valley, southern CA, and Yuma Valley, AZ. The sampling Global Positioning System (GPS) coordinates, mean annual temperature (MAT), and mean annual precipitation (MAP), as well as other environmental parameters of sampling sites, were previously described (Ma *et al.*, 2012a, b; Table S1). The major fresh produce growing on these locations during sampling were lettuce and spinach.

Each sample (0–15 cm) was a composite of several individual soil cores taken at 5-m intervals. Equal numbers of samples from conventional farming practice and organic farming practice were collected from each area (32 soils, 16 organically, 16 conventionally managed soils). Yuma and Imperial Valley shared similar weather conditions, including MAT and MAP, due to the closeness of the two locations in comparison with Salinas Valley, CA. Soil pH was between 6.7 and 8.0. Bagged samples were taken to the laboratory under ice. Vegetation, roots, and stones were removed and the soil sieved (< 2 mm). Soil physical, chemical, and some biological characteristics are as reported (Ma *et al.*, 2012a, b). Total organic carbon (OC, %) and total nitrogen (TN, %) were determined using Flash 2000 NC Analyzers (Thermo Scientific, Lakewood, NJ). Soil microbial biomass carbon (MBC, mg kg<sup>-1</sup>) was extracted by the chloroform-fumigation–extraction method (Vance *et al.*, 1987). Salinity measured as electrical conductivity (EC, dS m<sup>-1</sup>) of each soil was obtained by determining the conductivity of soil water extract (water to soil ratio, 1 : 1, vol:wt) using a conductivity meter (Oakton conductivity meter, Oakton, Houston, TX).

### Bacterial strains

An *E. coli* O157:H7 strain EDL933 was used in this study (ATCC 43895), as described previously (Ma *et al.*, 2011). *Escherichia coli* O157:H7 wild type was tagged with nalidixic acid in addition to rifampicin resistance to facilitate the enumeration, and its growth curve in LB (Luria-Bertani) broth and survival in soils were found to be identical to those of the nontagged wild-type strain (Ma *et al.*, 2011).

### *Escherichia coli* O157:H7 EDL933 survival in soils

Cells were pregrown overnight in LB medium, harvested by centrifugation (12 000 g) for 10 min, and washed with phosphate buffer (10 mM, pH 7.2) three times. The cell pellets were finally resuspended in sterile deionized water and used as inoculum. The cells were added in soils to a final density of about  $0.5 \times 10^7$  CFU per gram soil dry weight (gdw<sup>-1</sup>) according to a method slightly adapted from Franz *et al.* (2008). About 500 g of the inoculated soil was transferred to a plastic bag that was closed but, which had some holes at the top to allow air exchange as previous described (Ma *et al.*, 2011). The same amount of noninoculated soil was also placed into another plastic bag, which was used as a noninoculated control, with deionized water being added instead of cell suspension in triplicate plastic bags. The plastic bags were weighed and incubated at 22 °C. Moisture content of the soil sample

was adjusted to 50% of WHC, and water concentration was maintained constantly during the course of an experiment by adding deionized water weekly to balance the water loss by evaporation. Weight loss of the control samples was used as a reference measure of evaporated moisture.

The inoculated soils were sampled at days 0, 3, 6, 10, 14, 20, 27, 34, 40, and 48 to determine the survival of *E. coli* O157:H7 over time. At each point, three samples (about 1.0 g dry weight equivalent) from each triplicate plastic bag were removed from the middle of the soil sample and put into preweighed dilution tubes. The tubes containing soil samples were weighed to calculate the exact mass of soil sample. A 5.0 mL of 0.1% peptone buffer (Lab M, Lancashire, UK) was added to the test tube containing the soil sample, and the soil was thoroughly mixed with the buffer by inverting the tube several times and then vortexed for  $2 \times 20$  s; the resulting soil paste (cell suspension) was then subjected to 10-fold serial dilutions. Fifty microlitre of the two highest dilutions was plated in duplicate on SMAC (sorbitol MacConkey) agar supplemented with BCIG (5-bromo-4-chloro-3-indoxyl- $\beta$ -D-glucuronide) and appropriate antibiotics (rifampicin,  $100 \mu\text{g mL}^{-1}$ , and nalidixic acid,  $25 \mu\text{g mL}^{-1}$ ) for enumeration. The number of dilutions was determined based on preliminary tests. The inoculated SMAC agar plates were incubated at  $37^\circ\text{C}$  for 16 h, and the results were expressed as log colony-forming units (CFU) per gram soil dry weight ( $\text{gdw}^{-1}$ ). The detection limit of the plating method was *c.* 100 CFU  $\text{gdw}^{-1}$ . Our preliminary experiments showed that the average cell recovery rate of the method was from 90% to 110% of the theoretical value.

### Survival data modeling

Survival of *E. coli* O157:H7 was analyzed as previously described (Ma *et al.*, 2012a, b). In brief, the experimental data were fitted to the Weibull survival model proposed by Mafart *et al.* (2002) using GInaFiT version 1.5 developed by Dr. Annemie Geeraerd at Katholieke Universiteit, Leuven, Belgium (Geeraerd *et al.*, 2005). When using the Weibull survival model, the deactivation kinetics of the *E. coli* O157:H7 EDL933 population follows a Weibull distribution pattern (Franz *et al.*, 2008). The size of the surviving population can be calculated using equation 1,

$$\log(N_t) = \log(N_0) - (t/\delta)^p \quad (1)$$

where  $N_t$  is the number of survivors;  $N_0$  is the inoculum size ( $\text{CFU gdw}^{-1}$ );  $t$  is the time (day);  $p$  is the shape parameter; when  $p > 1$ , a convex curve is observed; when  $p < 1$ , a concave curve is observed; and when  $p = 1$ , a

linear curve is observed. The scale parameter,  $\delta$ , represents the time (day) needed for first decimal reduction. A very important and useful parameter,  $t_{td}$  (time needed to reach detection limit, 100 CFU  $\text{gdw}^{-1}$ ), can also be calculated when using GInaFiT to fit the experimental survival data.

### 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, Carlsbad, 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, Wilmington, 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  $\mu\text{L}$ ) were then submitted to Research and Testing Laboratories (Lubbock, TX) for PCR optimization and pyrosequencing analysis. Bacterial tag-encoded FLX amplicon pyrosequencing was carried out as previously described (Acosta-Martínez *et al.*, 2008; Dowd *et al.*, 2008; Acosta-Martínez *et al.*, 2010). 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 *c.* 600-bp hypervariable region of 16S rRNA genes. Primer and PCR optimizations were carried out at the Research and Testing Laboratories (Lubbock) according to the protocol described previously (Acosta-Martínez *et al.*, 2010; Gontcharova *et al.*, 2010; Nonnenmann *et al.*, 2010). All FLX-related procedures were performed following Genome Sequencer FLX System manufacturers instructions (Roche, Basel, Switzerland). Thus, moderate diversity pyrosequencing analysis ( $\geq 3000$  reads per sample) was performed at the Research and Testing Laboratory (Lubbock). Tags that did not have 100% homology to the original sample tag designation were not considered, as they might be suspect in quality. Sequences that were  $< 200$  bp after quality trimming 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 (Magurran, 2004), we 'resampled' our sequence libraries so that they contained similar numbers of sequences. The sub.sample function in MOTHUR was used to randomly select a subsample of sequences from each library, and these equally sized, reduced data sets were used in all subsequent analyses of our sequence libraries.

Following chimera detection and the resampling of the larger sequence libraries, the RDP Classifier function was used to assign identities to the bacterial pyrotag sequence data (Wang *et al.*, 2007). MOTHUR was used to align the resampled data set and create an all-sample distance matrix, as well as assign sequences to operational taxonomic units (OTU = 97% similarity, using the hcluster 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 & Clayton, 2005; 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.

## Statistical analysis

Linear regression analysis of the correlation of *ttd* with abundance of bacterial phyla, subclasses of *Proteobacteria* (*Alpha*-, *Beta*-, *Delta*- and *Gammaproteobacteria*), water-soluble organic carbon (WSOC) in soil water extracts, and plotting of survival parameters ( $N_0$ ,  $\delta$ , and  $p$ ) of *E. coli* O157:H7 in different group of soils was carried out using SYSTAT 12 (Systat Software, Chicago, IL).

Artificial neural network (ANN) analysis was conducted using Synapse program (Peltarion Inc., Stockholm, Sweden). Model training for *E. coli* survival data (*ttd*), soil physical, chemical, and biological variables used 500 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 the survival of *E. coli* O157:H7 in soils.

## Results

### Soil characterization

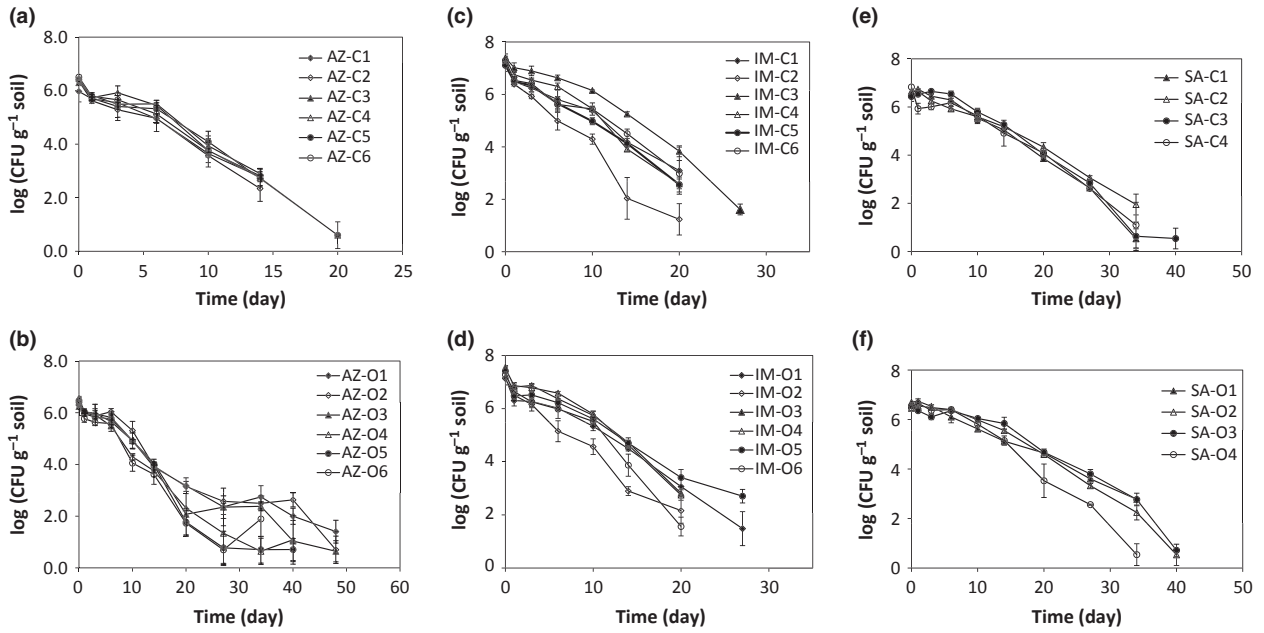
A total of 32 soils (16 organic, 16 conventionally managed soils) were collected from three major fresh produce-

growing areas in CA and AZ. Some physical and chemical properties are presented in Table S1 as previously reported by Ma *et al.*, 2012a, b. Yuma and Imperial Valley shared similar weather conditions, including MAT and MAP, due to the closeness of the two locations in comparison with Salinas Valley. Soil pH was between 6.7 and 8.0. The salinity of conventionally managed soils from Yuma was significantly higher ( $P < 0.01$ ) than that of organically managed soils. Also, soil salinity from Salinas Valley was significantly lower ( $P < 0.01$ ) than that from Yuma and Imperial Valley.

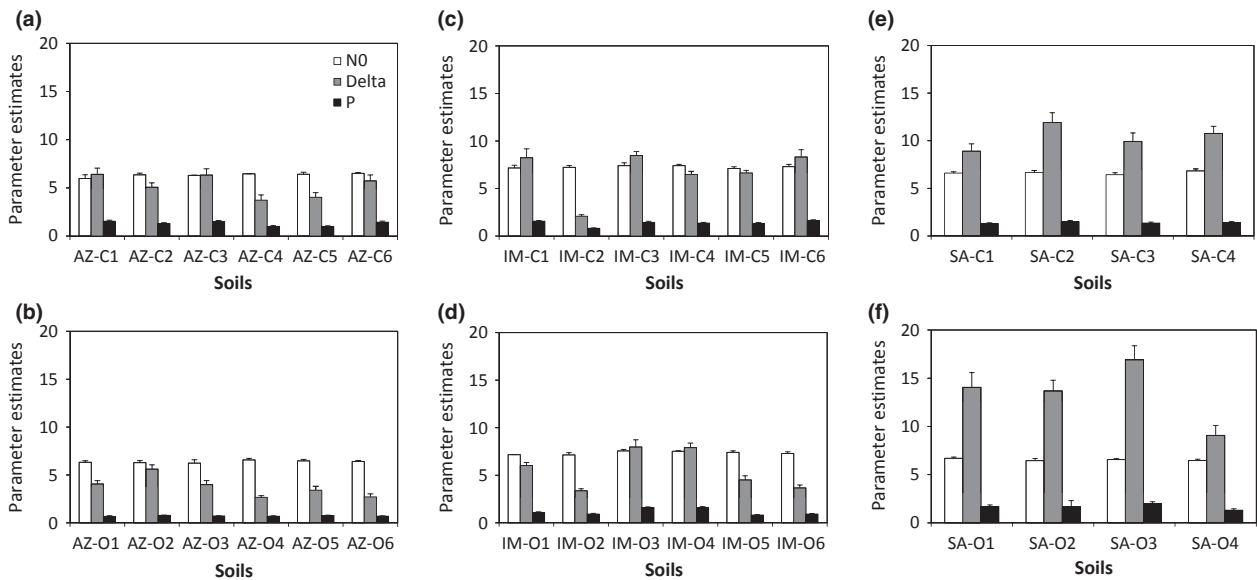
### Survival of *E. coli* O157:H7 in soils

The *E. coli* O157:H7 strain 933 was inoculated in soils from the three fresh produce-growing regions to test their survival at 22 °C (Fig. 1). The results (Fig. 1a; conventionally managed soils from Yuma, AZ) showed that there were no significant differences in deactivation profiles among the six soils. However, in organically managed soils, there were variations in deactivation profiles among the six soils resulting in significant differences in deactivation among sites (Fig. 1b). *Escherichia coli* O157:H7 survived significantly longer in organic soils (26 days on the average) before reaching the detection limit (*ttd*) of 100 CFU g<sup>-1</sup> soil from Yuma (AZ) than in conventional soil (15 days) from the same location. In organically managed soil (Fig. 1b), there was a sharp decline of cell population within the first 2 weeks postinoculation, followed by a steady decrease until cell concentration reached or dropped below detection limit. Survival data (Fig. 1c and d) showed that *E. coli* O157:H7 survived for about 20.5 days in soils from Imperial Valley and 35 days before reaching the detection limit (*ttd*) in soils from Salinas (Fig. 1e and f). No significant differences in survival patterns (*ttd*) were observed in organic and conventional soils from Imperial Valley and Salinas Valley.

Survival of *E. coli* O157:H7 in organic and conventionally managed soils was modeled by fitting the experimental data to the Weibull equation (Fig. 2). Modeling parameters showed inoculum size ( $N_0$ ), time needed for first log reduction, delta ( $\delta$ ), and the shape parameter ( $p$ ) were calculated from Eqn.(1) and used to explain the inactivation kinetics. More variations were observed in  $\delta$  values from different soils (Fig. 2). The initial sharp decrease in cell numbers is attributed to the faster decline of the initial population as indicated with smaller  $\delta$  (Fig. 2b) and as seen in some of the organically managed soils from Yuma, AZ, in comparison with the conventional soils that did not experience the initial sharp decrease. The shape parameter,  $p$ , was not different in the survival profiles of *E. coli*



**Fig. 1.** Survival profile of *Escherichia coli* O157:H7 in conventionally managed (C) and organically managed (O) soils from Yuma (a and b), Arizona (AZ), Imperial Valley (c and d), California (IM), and Salinas (e and f), California (SA). Standard errors are means from three composite samples.



**Fig. 2.** Survival parameters when *Escherichia coli* O157:H7 was inoculated into organically managed (O) and conventionally managed (C) soils from Yuma (a and b), Arizona (AZ), Imperial Valley (c and d), California (IM), and Salinas (e and f), California (SA). Survival parameters resulted from the modeling of the raw survival data using Weibull deactivation model, including predicted inoculum size ( $N_0$ , log CFU  $g^{-1}$  soil), first decimal reduction time ( $\delta$ , day), and shape parameter ( $p$ ). Standard errors are means from three composite samples.

O157 in all soils (Fig. 2), with almost all the  $p$  values being  $\geq 1$ , except that the  $p$  values calculated from survival data in AZ organically managed soils were smaller than 1.

### Pyrosequencing data analysis

From the pyrosequencing analysis, a total of 98 730 sequence tags and 18 368 OTUs were obtained from the

32 soils after normalization, with the highest OTUs (4153) in organically managed soils from Salinas Valley area in CA. On the average, the number of OTUs was higher in all the organically managed soils than that in the conventionally managed soils (Table 1). The alpha diversity indices (Shannon, Simpson, and CHAO) of bacterial communities in each soil are listed in Table 1. *Actinobacteria*, *Proteobacteria*, *Acidobacteria*, and *Bacteroidetes* were the dominant phyla among the bacterial communities in soils, and these four phyla accounted for about 75% of the total bacterial composition (Fig. 3a). Further analysis showed that *Alpha*- and *Gamma*-subclasses of *Proteobacteria* dominated (*c.* 65%) the composition of this phylum, while the other two subclasses, *Beta*- and *Deltaproteobacteria* only accounted for about 33% of the total *Proteobacteria* composition (Fig. 3b).

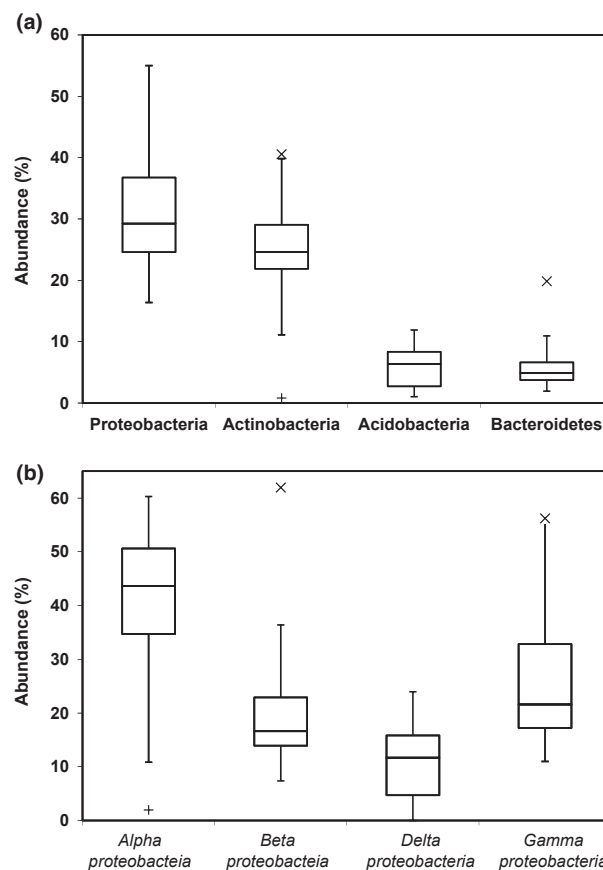
### Role of different bacterial phyla on the survival of *E. coli* O157:H7 in soils

Bacterial diversity had no significant effect ( $P = 0.63$ ) on the survival of *E. coli* O157:H7 in soils (Fig. 4). However, further regression analysis revealed that individual phyla correlated differently with the survival of *E. coli* O157:H7 in soil (Fig. 5). It was found that *Actinobacteria* (Fig. 5a) and *Acidobacteria* (Fig. 5b) significantly ( $P < 0.001$  and  $P < 0.05$ , respectively) correlated with higher survival of *E. coli* O157:H7, while the other two major phyla, *Proteobacteria* (Fig. 5c) and *Bacteroidetes* (Fig. 5d) significantly ( $P < 0.05$ ) correlated with the suppression of survival of *E. coli* O157:H7 in our soils. Further correlation analysis showed that in organically managed soils, *Actinobacteria*, *Acidobacteria*, and *Deltaproteobacteria* correlated with higher survival ( $P < 0.05$ ) of *E. coli* O157:H7, while in conventionally managed soils, only *Acidobacteria* and *Deltaproteobacteria* displayed a significant positive correlation ( $P < 0.05$ ) in the survival of *E. coli* O157:H7 (Table 2). Interestingly, *Firmicutes* was found to positively ( $P < 0.05$ ) correlate with the overall survival of *E. coli* O157:H7 only in conventional soils (Table 2). It was also found that WSOC concentrations in soil water extracts

were highly correlated with the abundance of *Actinobacteria* in the total bacterial communities (Fig. 5e).

### Artificial neural network analysis of survival time, soil properties, and bacterial diversity

The Kohonen self-organizing map (KSOM) patterns of *ttt* with those of soil management, soil physical, chemical, and biological variables showed that the longer survival (clustered in red color; Fig. 6a) was largely correlated

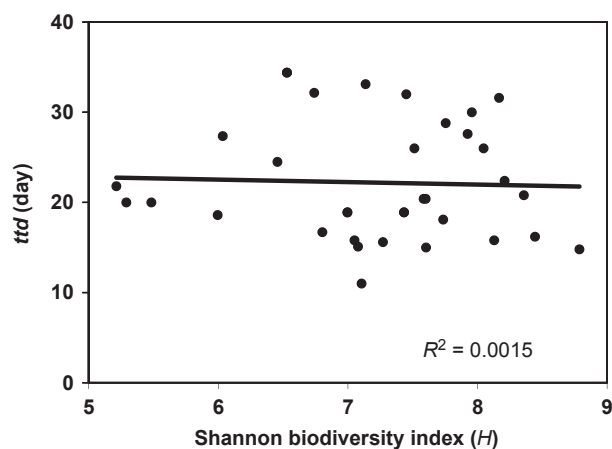


**Fig. 3.** Abundance of dominated phyla revealed by 454-pyrosequencing targeting 16S rRNA genes. (a) Dominant bacterial phyla; (b) relative abundance for subclasses of *Proteobacteria*.

**Table 1.** Alpha diversity indices (97%) based on 454-pyrosequencing data from organic and conventionally managed soils from the three fresh produce-growing areas of CA and AZ

Soils	Seq-Tags	OTUs	Chao	Inverse Simpson	Shannon diversity	Simpson	Coverage
AZ-conv	16455	2526	4371.946	35.95508	6.048294	0.027812	0.923488
AZ-org	16455	2776	5017.781	187.7758	6.695063	0.005325	0.915588
IM-conv	16455	2315	4289	108.9919	6.25369	0.009175	0.928532
IM-org	16455	2781	5241.569	171.504	6.667548	0.005831	0.914069
SA-conv	16455	3817	7677.545	136.9915	7.013107	0.0073	0.866363
SA-org	16455	4153	8766.529	349.2988	7.298596	0.002863	0.849772

Conv, conventionally managed soils; org, organically managed soil.



**Fig. 4.** Relationship between survival of *Escherichia coli* O157:H7 (*ttf*) and bacterial diversity index (*H*) using the Shannon index of diversity.

with high levels of WSOC, high levels of TN, low levels of salinity, high abundance of *Actinobacteria* and *Acidobacteria*, and low abundance of *Proteobacteria* and *Firmicutes*. A sensitivity diagram (Fig. 6b) for the *ttf* shows that both WSOC and pH were major factors influencing the survival of *E. coli* O157:H7. The two surface plots (Fig. 6c and d) show data for WSOC and EC vs. *ttf*, and for pH and WSOC vs. *ttf*. These showed that the best *ttf* is associated with low salinity, low pH, and high WSOC. Further artificial neural network modeling results (Fig. 7) showed that for a given soil, an increase in TN and clay content and a decrease in salinity correlated with higher *ttf*. Salinity of conventional soils from Yuma was significantly higher ( $P < 0.01$ ) than that of organic soils. Also, soil salinity from Salinas Valley was significantly lower ( $P < 0.01$ ) than that from Yuma and Imperial Valley. Our field data showed that increasing salinity content in

soil corresponded to significant linear decrease in the survival of *E. coli* O157:H7. Both regression analysis and artificial neural network analysis of survival data are in agreement on most of the environmental factors that correlate significantly with the survival of *E. coli* O157:H7 in the studied soils.

## Discussion

In the current study, regression and artificial neural networks analyses generated similar trends on the relationships between *E. coli* O157 survival and environmental variables including soil properties and bacterial community structures. When the neural network-based models are adequately trained, it could be a good approach with more precise predictions (Karul & Soyupak, 2006). Our results showed that salinity was one of the significant abiotic factors in the determination of survival of *E. coli* O157 in soils studied, which is overall consistent with our recent study (Ma et al., 2012a, b). These authors noted that salinity of soils in southern California is higher due to the high evaporation and low precipitation (Rhoades, 1981). However, the effects of salinity on the survival of commensal *E. coli* strains have been investigated in seawater (Carlucci & Pramer, 1960; Anderson et al., 1979). A study on the survival of commensal *E. coli* in seawater (Shabala et al., 2009) showed that salinity may adversely influence the survival of *E. coli* by a general osmotic effect or by specific ion toxicity, which leads to a decrease in activity in some enzymes that are essential for cell metabolism, for example,  $K^+$  transport. Ma et al. (2012a, b) noted in a recent study that the influence of salinity on the survival of *E. coli* O157:H7 in soils, especially in conventionally managed soils, might also hold true due to the effects of salinity on cells. They suggested that soil salinity might interfere with ion transport,

**Table 2.** Correlation between survival time (*ttf*) with dominant bacterial communities from both organic and conventional soils

Group	Organic soil			Conventional soil		
	<i>R</i>	<i>R</i> <sup>2</sup>	<i>P</i>	<i>R</i>	<i>R</i> <sup>2</sup>	<i>P</i>
<i>Actinobacteria</i>	+0.61	0.379	0.011*	+0.40	0.163	0.135
<i>Acidobacteria</i>	+0.64	0.411	0.001***	+0.62	0.391	0.013*
<i>Bacteroidetes</i>	-0.64	0.415	0.007**	-0.44	0.189	0.105
<i>Proteobacteria</i>	-0.43	0.186	0.095	-0.48	0.232	0.069
<i>Alphaproteobacteria</i>	+0.18	0.033	0.502	+0.16	0.025	0.572
<i>Betaproteobacteria</i>	-0.39	0.152	0.137	-0.26	0.065	0.358
<i>Deltaproteobacteria</i>	+0.62	0.386	0.010**	+0.67	0.449	0.006**
<i>Gammaaproteobacteria</i>	-0.25	0.063	0.350	-0.47	0.220	0.078
<i>Chloroflexi</i>	+0.28	0.081	0.287	+0.37	0.137	0.175
<i>Firmicutes</i>	+0.11	0.010	0.706	+0.59	0.352	0.020*

\*, \*\*, and \*\*\* denotes significance at the 0.05, 0.01, and 0.001 test levels, respectively '+' indicates a positive correlation, and '-' indicates a negative correlation.

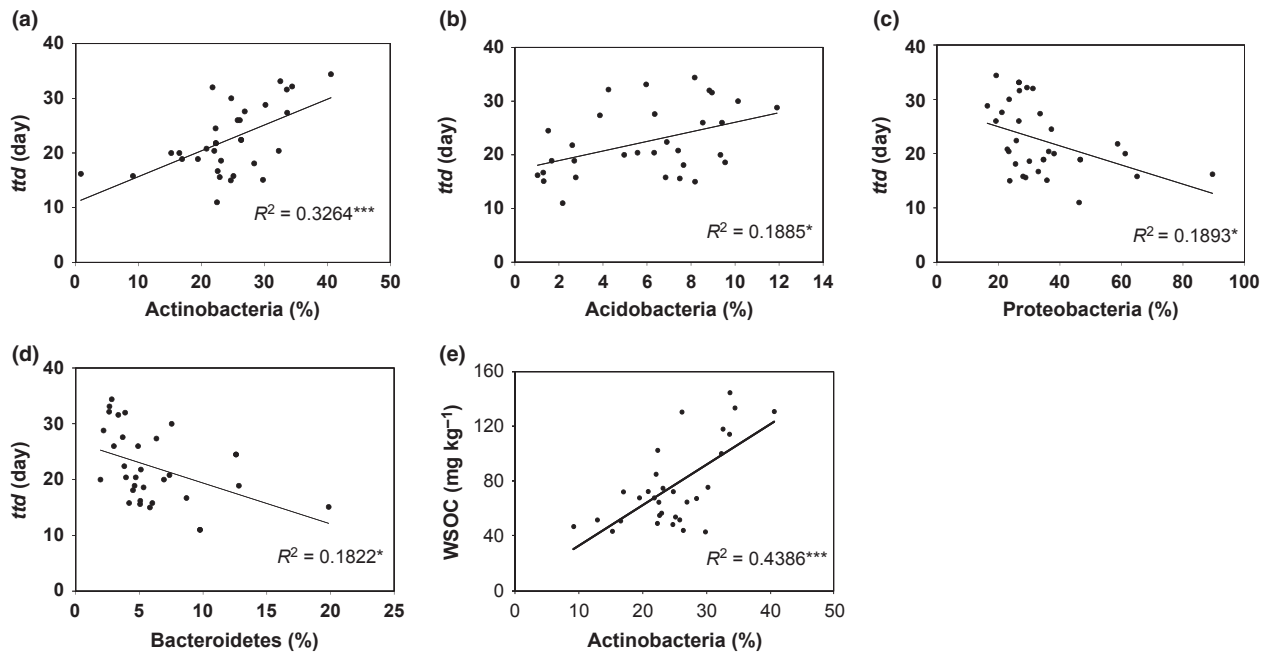


Fig. 5. Correlation of relative abundance of major bacterial phyla (a: *Actinobacteria*; b: *Acidobacteria*; c: *Proteobacteria*; and d: *Bacteroidetes*) with the survival of *E. coli* O157:H7 (*ttd*) in soils.

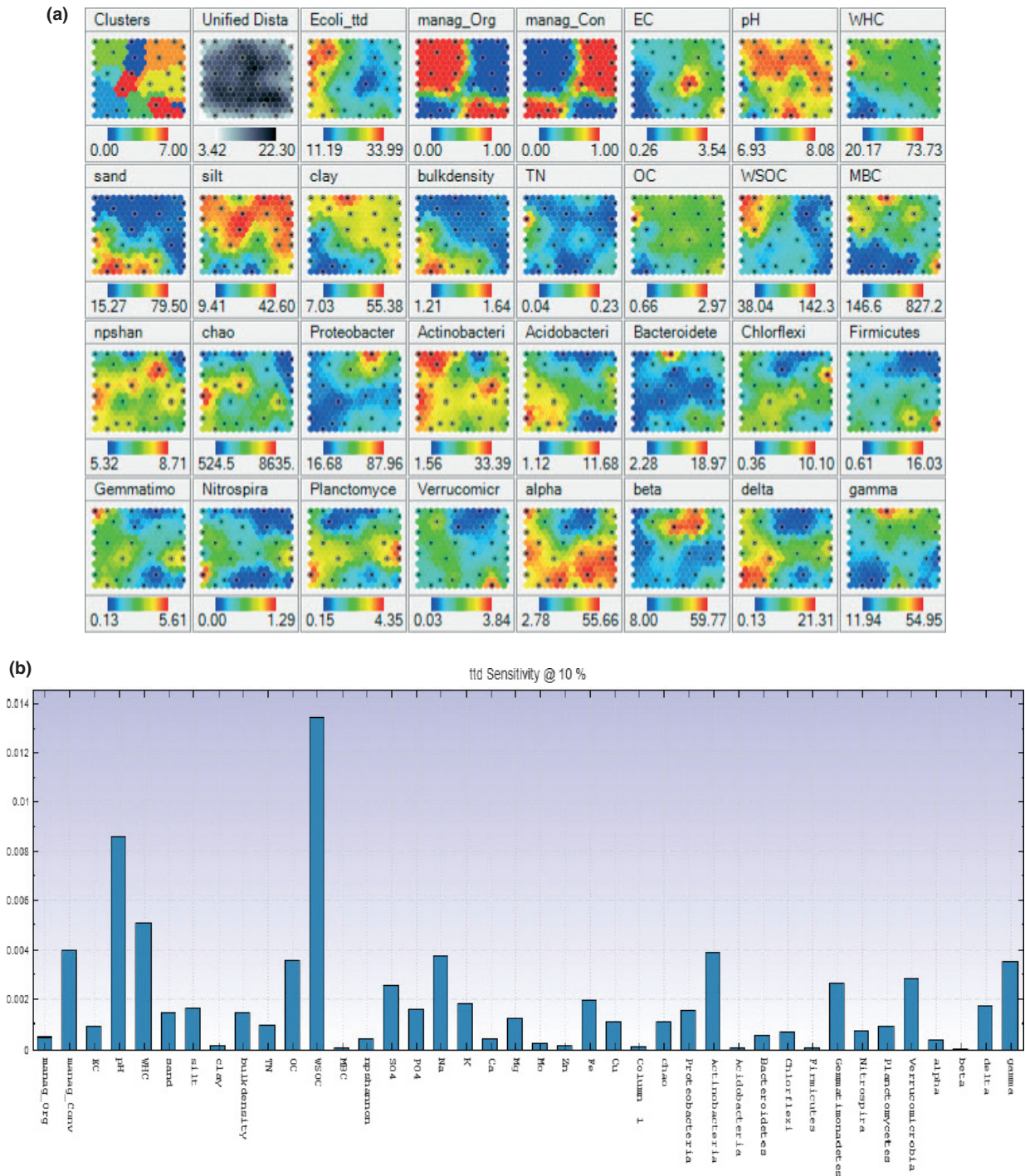
inhibit enzyme activity, and suppress synthesis of crucial proteins in *E. coli* O157:H7, and finally, result in a reduced survival capacity in soils.

Linear regression analysis (Fig. 4) showed that the survival of *E. coli* O157:H7 was not significantly affected by the Shannon diversity index determined based on pyrosequencing data analysis. Previous studies (Jiang *et al.*, 2002; Unc *et al.*, 2006; Semenov *et al.*, 2007) have shown that the total indigenous microbial communities have a negative effect on the survival of *E. coli* O157:H7 introduced into soils as a result of predation, substrate competition, and antagonism. A recent investigation (Franz *et al.*, 2008) showed that in organically managed soils, the survival of *E. coli* O157:H7 was negatively correlated with microbial diversity determined based on PCR-DGGE. The diversity and species richness of the indigenous microbial communities were important factors to determine the survival behavior of *E. coli* introduced into the soils (van Elsas *et al.*, 2007). It is widely accepted that microbial ecosystems with a high level of diversity (Trevors, 1998) are generally more stable and resistant to perturbation than those with a lower diversity (Tilman, 1997). Therefore, the soil systems with reduced biological complexity might be more susceptible to invasion by *E. coli* O157:H7, because those ecosystems offer enhanced opportunities for the pathogens to persist (Girvan *et al.*, 2005; van Elsas *et al.*, 2007; Semenov *et al.*, 2008; Ibekwe *et al.*, 2010; Ibekwe & Ma, 2011). Our study has advanced

the current state of science by exploring the effect of individual phyla or groups of microorganisms on the persistence of *E. coli* O157:H7. The use of 454 deep sequencing technology has aided us in this project to study the microbial ecology at a higher coverage and resolution. To the best of our knowledge, the current study might be the first work to investigate how the individual phyla influence the survival of *E. coli* O157:H7 in leafy green-production soils.

Our data showed that different groups of bacteria might have distinct impacts on the survival of *E. coli* O157:H7. *Actinobacteria* and *Acidobacteria* were highly correlated with the survival of *E. coli* O157 in the soils tested. *Actinobacteria* constitute one of the largest phyla among *Bacteria* and represent gram-positive bacteria with a high G+C content (Goodfellow & Williams, 1983). *Actinobacteria* are widely distributed in terrestrial ecosystems, especially in soils, with variable physiological and metabolic properties, which enable them to decompose and mineralize naturally occurring compounds such as cellulose and chitin, therefore playing a crucial role in organic matter turnover. In this study, the survival of *E. coli* O157:H7 was strongly correlated with both the abundance of *Actinobacteria* and the levels of soil WSOC. It should be noted that WSOC comprises physically protected carbon as well as carbon contained in soil macropores that are easily accessible to microorganisms, and WSOC was a major factor correlating positively with the





**Fig. 6.** Visual Inspection of correlations between variables using artificial neural networks. (a) Kohonen self-organizing map of *Escherichia coli* survival and associated soil chemical, physical, and biological variables. (a) Kohonen self-organizing map of *E. coli* O157:H7 survival (time to detection limit, *ttd*), associated soil properties, and abundance of bacterial at the phylum level. alpha, beta, delta, and gamma represent the relative abundance of *Alpha*-, *Beta*-, *Delta*- and *Gammaproteobacteria* among *Proteobacteria* community in soils, respectively. Scale bar in each KSOM map indicates the range of each soil physical, chemical, or biological variable. (b) Sensitivity analysis of output variable to selected input variables: Salinity or EC ( $\text{dS m}^{-1}$ ), water holding capacity (WHC, %), silt content (%), clay content (%), total nitrogen (TN, %); organic carbon (OC, %); water-soluble organic carbon in soil water extract (WSOC,  $\text{mg kg}^{-1}$ ); microbial biomass carbon (MBC  $\text{mg kg}^{-1}$ ); Shannon biodiversity index (*H*) etc. Two surface plots showing data for (c) WSOC and EC vs. *ttd* and for (d) pH and WSOC vs. *ttd*. These show that greater *ttd* is associated with low salinity, high pH, and high WSOC.

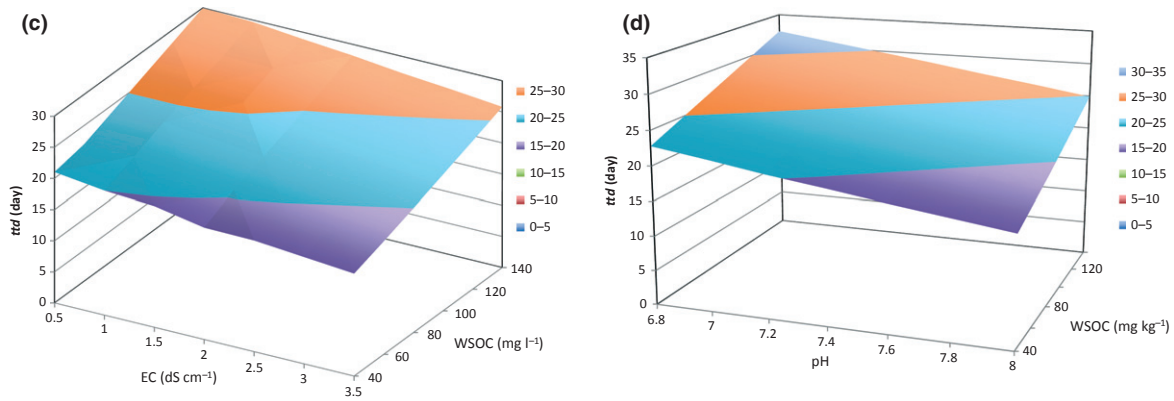
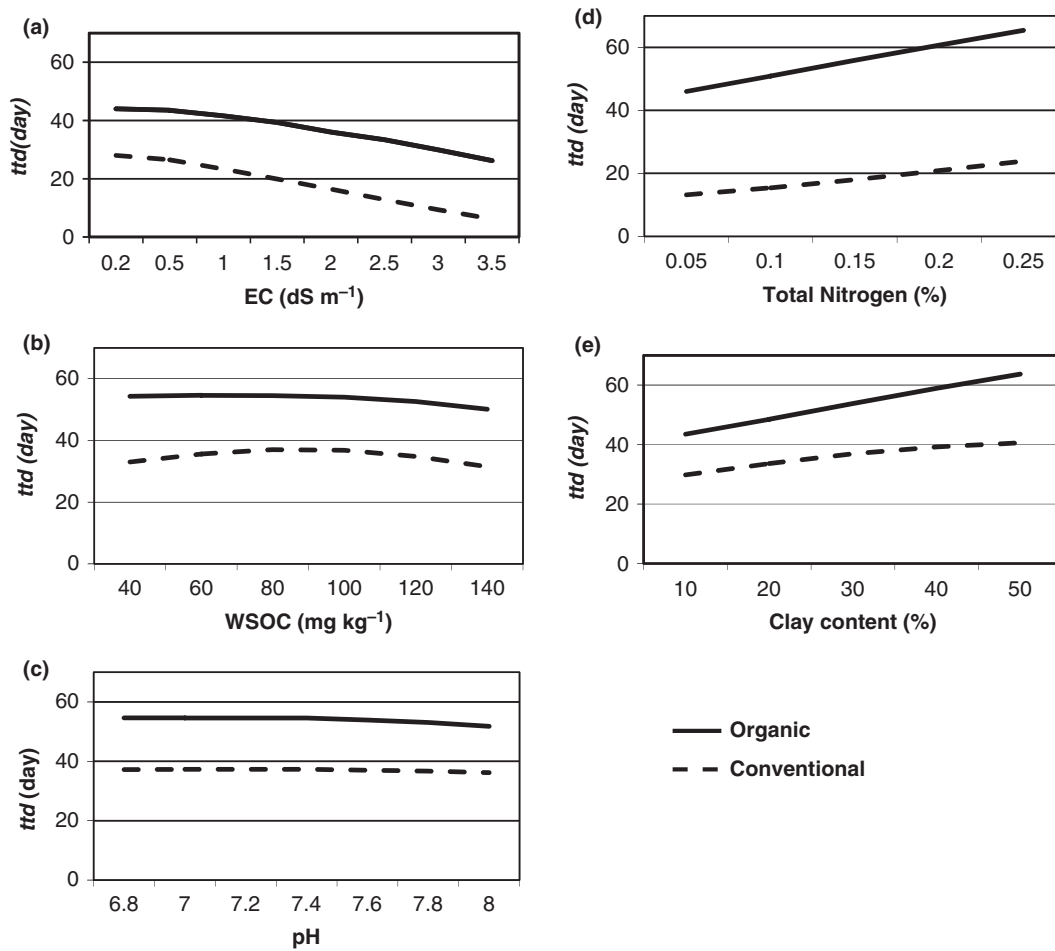


Fig. 6. Continued



**Fig. 7.** Predicted effects of soil variables, (a) electrical conductivity (EC), (b) water-soluble organic carbon (WSOC), (c) pH, (d) total nitrogen (TN), and (e) clay content (%), based on artificial neural network analysis, on the survival (time to reach detection limit (*ttd*)) of *Escherichia coli* O157:H7 in both organically and conventionally managed soils. The analysis is nonlinear, which allows better fit to the data exhibiting complex behavior.

survival of *E. coli* O157:H7 in soils as observed in the regression and artificial neural network analyses. This suggests that *Actinobacteria* may play a major role in improving the survival of *E. coli* O157:H7 in soils by degrading high-molecular-weight organic compounds into readily available organic carbon and energy sources for *E. coli* O157:H7 cells to use as carbon sources for growth and survival. *Acidobacteria* may also play a similar role in carbon turnover in soil as *Actinobacteria* (Hugenholtz et al., 1998; Barns et al., 1999; Eichorst et al., 2011). Data from whole-genome analysis of three strains suggested that at least some *Acidobacteria* are versatile heterotrophs capable of breaking down plant polymers (Ward et al., 2009). In contrast to *Actinobacteria* and *Acidobacteria*, *Proteobacteria* and *Bacteroidetes* displayed a negative correlation ( $P < 0.05$ ) on the survival of *E. coli* O157:H7 in soils. The suppressive effect caused by *Proteobacteria* on *E. coli* O157:H7 may be explained by competition for nutrients and habitats within the same niche. As a member of *Gammaproteobacteria*, *E. coli* O157:H7 shares many physiological and ecological similarities with other members of the *Proteobacteria* phylum, especially with the *Gamma*-subclass. Interestingly, it was found that not all of the *Proteobacteria* subclasses showed a negative correlation on the survival of *E. coli* O157:H7. Among the four major subclasses, *Beta*- and *Deltaproteobacteria* showed negative correlation, while *Alpha*- and *Deltaproteobacteria* showed positive correlation on the survival of the pathogen. In agreement with our findings, recent research (Westphal et al., 2011) showed that the survival of *E. coli* O157:H7 was suppressed by *Beta*- and *Deltaproteobacteria* in sand-based dairy livestock bedding. In the current study, it was also found that *Bacteroidetes* exhibited negative correlation ( $P < 0.05$ ) on the survival of *E. coli* O157:H7 in soils. *Bacteroidetes* are, for the most part, strict anaerobes often associated with the intestinal tracts of animals. *Bacteroidetes* are also known for their ability to degrade complex polymers. A recent study (Westphal et al., 2011) showed that some clones belonging to *Bacteroidetes* displayed some suppressive effect on the survival of *E. coli* O157:H7 in sand-based dairy livestock bedding, indicating that some specific species among *Bacteroidetes* might enhance the survival of *E. coli* O157:H7 in soils, although overall the whole phylum shows negative correlation on the persistence of the pathogen.

Because different groups of bacteria displayed distinct effects on the survival of *E. coli* O157:H7, the overall effect of microbiota on the persistence of *E. coli* O157 largely depends on the individual microbial group that dominates the microbial communities. If the suppressive groups dominate, the survival time of *E. coli* O157:H7 decreases, and if the supportive groups dominate, the survival of *E. coli* O157:H7 might be promoted.

## 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. 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 and Sam Henshaw for facilitating the soil sampling and transport to the laboratory. 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.
- Acosta-Martínez V, Dowd SE, Bell CW, Lascano R, Booker JD, Zobeck TM & Upchurch DR (2010) Microbial community composition as affected by dryland cropping systems and tillage in a semiarid sandy soil. *Diversity* **2**: 910–931.
- Anderson IC, Rhodes M & Kator H (1979) Sublethal stress in *Escherichia coli*: a function of salinity. *Appl Environ Microbiol* **38**: 1147–1152.
- Barns SM, Takala SL & Kuske CR (1999) Wide distribution and diversity of members of the bacterial kingdom *Acidobacterium* in the environment. *Appl Environ Microbiol* **65**: 1731–1737.
- Carlucci AF & Pramer D (1960) An evaluation of factors affecting the survival of *Escherichia coli* in sea water: II. salinity, pH, and nutrients. *Appl Microbiol* **8**: 247–250.
- Caesson M, O'Sullivan O, Wang Q, Nikkilä J, Marchesi JR, Smidt H, de Vos WM, Ross RP & O'Toole PW (2009) Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. *PLoS ONE* **4**: e6669.
- Cole JR, 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.
- Dowd SE, Sun Y, Wolcott RD & Carroll JA (2008) Bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP) for microbiome studies: bacterial diversity in the ileum of newly weaned Salmonella-infected pigs. *Foodborne Pathog Dis* **5**: 459–472.
- Eichorst SA, Kuske CR & Schmidt TM (2011) Influence of plant polymers on the distribution and cultivation of bacteria in the phylum *Acidobacteria*. *Appl Environ Microbiol* **77**: 586–596.

- van Elsland JD, Hill P, Chronáková A, Grekova M, Topalova Y, Elhottová D & Kristůfek V (2007) Survival of genetically marked *Escherichia coli* O157:H7 in soil as affected by soil microbial community shifts. *ISME J* **1**: 204–214.
- van Elsland JD, Semenov AV, Costa R & Trevors JT (2011) Survival of *Escherichia coli* in the environment: fundamental and public health aspects. *ISME J* **5**: 173–183.
- Franz E, Semenov AV, Termorshuizen AJ, de Vos OJ, Bokhorst JG & van Bruggen AHC (2008) Manure-amended soil characteristics affecting the survival of *E. coli* O157:H7 in 36 Dutch soils. *Environ Microbiol* **10**: 313–327.
- Geeraerd AH, Valdramidis VP & van Impe JF (2005) GlnaFit, a freeware tool to assess non-log-linear microbial survivor curves. *Int J Food Microbiol* **102**: 95–105.
- Girvan MS, Campbell CD, Killham K, Prosser JI & Glover LA (2005) Bacterial diversity promotes community stability and functional resilience after perturbation. *Environ Microbiol* **7**: 301–313.
- Gontcharova V, Young E, Sun Y, Wolcott RD & Dowd SE (2010) A comparison of bacterial composition in diabetic ulcers and contralateral intact skin. *Open Microbiol J* **4**: 8–19.
- Goodfellow M & Williams ST (1983) Ecology of *Actinomycetes*. *Annu Rev Microbiol* **37**: 189–216.
- Hugenholtz P, Goebel BM & Pace NR (1998) Impact of culture independent studies on the emerging phylogenetic view of bacterial diversity. *J Bacteriol* **180**: 4765–4774.
- Ibekwe AM & Ma J (2011) Effects of fumigants on microbial diversity and persistence of *E. coli* O157:H7 in contrasting soil microcosms. *Sci Total Environ* **409**: 3740–3748.
- Ibekwe AM, Grieve CM & Yang C-H (2007) Survival of *Escherichia coli* O157:H7 in soil and on lettuce after soil fumigation. *Can J Microbiol* **53**: 623–635.
- Ibekwe AM, Papiernik SK, Grieve CM & Yang C-H (2010) Quantification of persistence of *Escherichia coli* O157:H7 in contrasting soils. *Int J Microbiol* **2011**: doi:10.1155/2011/421379.
- Jiang X, Morgan J & Doyle MP (2002) Fate of *Escherichia coli* O157:H7 in manure-amended soil. *Appl Environ Microbiol* **68**: 2605–2609.
- Karol C & Soyupak S (2006) A comparison between neural network based and multiple regression models for chlorophyll-a estimation. *Ecological Informatics: Scope, Techniques and Applications*, 2nd edn (Recknagel F, ed.), pp. 309–323. Springer-Verlag, Berlin.
- Ma J, Ibekwe AM, Yi X, Wang H, Yamazaki A, Crowley DE & Yang CH (2011) Persistence of *Escherichia coli* O157:H7 and its mutants in soils. *PLoS ONE* **6**: e23191.
- Ma J, Ibekwe AM, Crowley DE & Yang C-H (2012a) Survival of *Escherichia coli* O157:H7 in major leafy green producing soils. *Environ Sci Tech* **46**: 12154–12161.
- Ma J, Ibekwe AM, Leddy M, Yang C-H & Crowley DE (2012b) Assimilable organic carbon (AOC) in soil water extracts using *Vibrio harveyi* BB721 and its implication for microbial biomass. *PLoS ONE* **7**: e28519.
- Mafart P, Couvert O, Gaillard S & Leguerinel I (2002) On calculating sterility in thermal preservation methods: application of the Weibull frequency distribution model. *Int J Food Microbiol* **72**: 107–113.
- Magurran A (2004) *Measuring Biological Diversity*. John Wiley and Sons, Malden, MA.
- Maule A (2000) Survival of verocytotoxigenic *Escherichia coli* O157 in soil, water and on surfaces. *Symp Ser Soc Appl Microbiol* **29**: 71S–78S.
- Mubiru DN, Coyne MS & Grove JH (2000) Mortality of *Escherichia coli* O157:H7 in two soils with different physical and chemical properties. *J Environ Qual* **29**: 1821–1825.
- Nonnenmann MW, Bextine B, Dowd SE, Gilmore K & Levin JL (2010) Culture-independent characterization of bacteria and fungi in a poultry bioaerosol using pyrosequencing: A new approach. *J Occup Environ Hyg* **7**: 693–699.
- Rhoades JD (1981) Determining leaching fraction from field measurements of soil electrical conductivity. *Agric Water Manage* **3**: 205–215.
- Schloss PD, Westcott SL, 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 AV, van Bruggen AHC, van Overbeek L, Termorshuizen AJ & Semenov AM (2007) Influence of temperature fluctuations on *Escherichia coli* O157:H7 and *Salmonella enterica* serovar Typhimurium in cow manure. *FEMS Microbiol Ecol* **60**: 419–428.
- Semenov AV, Franz E, van Overbeek L, Termorshuizen AJ & van Bruggen AHC (2008) Estimating the stability of *Escherichia coli* O157:H7 survival in manure amended soils with different management histories. *Environ Microbiol* **10**: 1450–1459.
- Shabala L, Bowman J, Brown J, Ross T, McMeekin T & Shabala S (2009) Ion transport and osmotic adjustment in *Escherichia coli* in response to ionic and non-ionic osmotic. *Environ Microbiol* **11**: 137–148.
- Shendure J & Ji H (2008) Next-generation DNA sequencing. *Nat Biotechnol* **26**: 1135–1145.
- Tilman D (1997) Community invasibility, recruitment limitation, and grassland biodiversity. *Ecology* **78**: 81–92.
- Trevors JT (1998) Bacterial biodiversity in soil with an emphasis on chemically-contaminated soils. *Water Air Soil Pollut* **101**: 45–67.
- Unc A, Gardner J & Springthorpe S (2006) Recovery of *Escherichia coli* from soil after addition of organic wastes. *Appl Environ Microbiol* **72**: 2287–2289.
- Vance ED, Brookes PC & Jenkinson DS (1987) An extraction method for measuring soil microbial biomass C. *Soil Biol Biochem* **19**: 703–707.
- Wang Y & Qian P (2009) Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. *PLoS ONE* **4**: e7401.
- Ward NL, Challacombe JF, Janssen PH *et al.* (2009) Three genomes in the phylum Acidobacteria provide insight into their lifestyles in soils. *Appl Environ Microbiol* **74**: 2046–2056.

- Westphal A, Williams ML, Baysal-Gurel F, LeJeune JT & Gardene BBM (2011) General suppression of *Escherichia coli* O157:H7 in sand-based dairy livestock bedding. *Appl Environ Microbiol* **77**: 2113–2121.
- Yue JC & Clayton MK (2005) A similarity measure based on species proportions. *Comm Stat A Theor Meth* **34**: 2123–2131.

## Supporting Information

Additional Supporting Information may be found in the online version of this article:

**Table S1.** Soil properties.