Canadian Journal of Microbiology
Revue canadienne de microbiologie

Factors affecting the occurrence of *Escherichia coli* O157 contamination in irrigation ponds on produce farms in the Suwannee River Watershed

Journal:	*Canadian Journal of Microbiology*
Manuscript ID:	cjm-2012-0599.R1
Manuscript Type:	Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Gu, Ganyu; University of Florida, Plant Pathology and Emerging Pathogens Institute Luo, Zhiyao; University of Florida, Food Science and Human Nutrition Cevallos-Cevallos, Juan; University of Florida, Plant Pathology and Emerging Pathogens Institute Adams, Paige; University of Georgia, Biological & Agricultural Engineering Vellidis, George; University of Georgia, Biological & Agricultural Engineering Wright, Anita; University of Florida, Food Science and Human Nutrition van Bruggen, Ariena; University of Florida, Plant Pathology and Emerging Pathogens Institute
Keyword:	Escherichia coli O157, irrigation ponds, fecal indicators, biserial correlation, discriminant analysis

Factors affecting the occurrence of *Escherichia coli* O157 contamination in irrigation ponds on produce farms in the Suwannee River Watershed

Ganyu Gu , Zhiyao Luo, Juan M. Cevallos-Cevallos, Paige Adams, George Vellidis , Anita Wright and Ariena H. C. van Bruggen

G. Gu and A.H.C. van Bruggen. Emerging Pathogens Institute and Department of Plant Pathology, University of Florida, Gainesville, FL 32611, United States.

Z. Luo and A. Wright. Department of Food Science and Human Nutrition, University of Florida, Gainesville, FL 32611, United States.

J.M. Cevallos-Cevallos. Emerging Pathogens Institute and Department of Plant Pathology, University of Florida, Gainesville, FL 32611, United States; Centro de Investigaciones Biotecnológicas del Ecuador (CIBE). Escuela Superior Politécnica del Litoral (ESPOL). Campus Gustavo Galindo, Guayaquil, Ecuador.

P. Adams and G. Vellidis. Biological & Agricultural Engineering Department, University of Georgia, Tifton, GA 31793, United States.

Corresponding autor: Ganyu Gu (e-mail: ggu@ufl.edu); Ariena H. C. van Bruggen (e-mail: ahcvanbruggen@ufl.edu)

Running title: *Escherichia coli* O157 occurrence in irrigation ponds

Abstract Outbreaks of enteritis caused by *Escherichia coli* O157 associated with fresh produce have resulted in questions about the safety of irrigation water; however, associated risks have not been systematically evaluated. In this study, the occurrence and distribution of the human pathogen *E. coli* O157 from vegetable irrigation ponds within the Suwannee River Watershed in Georgia were investigated and the relationship to environmental factors was analyzed. Surface and subsurface water samples were collected monthly from 10 vegetable irrigation ponds from March, 2011 to February, 2012. *E. coli* O157 was isolated from enriched filtrates on CHROMagar™ and Sorbitol MacConkey agar media and confirmed by an agglutination test. Presence of virulence genes *stx1*, *stx2* and *eae* was tested by PCR. In addition, 27 environmental variables of the sampled ponds were measured. Denaturing gradient gel electrophoresis was conducted for the analysis of bacterial communities in the water samples. Biserial correlation coefficients were calculated to evaluate the correlations between the environmental factors and the occurrence of *E. coli* O157. Stepwise and canonical discriminant analyses were used to determine the factors that were associated with the presence and absence of *E. coli* O157 in water samples. All 10 ponds were positive for *E. coli* O157 some of the time, mainly in summer and fall of 2011. The temporal distribution of this bacterium differed among the 10 ponds. Temperature, rainfall, populations of fecal coliform, and culturable bacteria were positively correlated with the occurrence of *E. coli* O157 ($P < 0.05$), while the total nitrogen concentration, oxidation-reduction potential and dissolved oxygen concentration were negatively correlated with the occurrence of this pathogen ($P < 0.05$). Temperature and rainfall were the most important factors contributing to the discrimination between samples with and without *E. coli* O157, followed by bacterial diversity and culturable bacteria population density. Bacterial numbers and diversity, including fecal coliforms and *E. coli* O157, increased after rainfall (and

47 possibly run-off from pond margins) in periods with relatively high temperatures, suggesting that
48 prevention of run-off may be important to minimize the risk of enteric pathogens in irrigation
49 ponds.
50
51 *Key words*: *Escherichia coli* O157, irrigation ponds, fecal indicators, biserial correlation,
52 discriminant analysis

Introduction

Enterohemorrhagic *Escherichia coli* (EHEC), an important zoonotic pathogen, is attracting global attention. *E. coli* O157 is typically responsible for most infections (80%) among the Shiga toxin producing bacteria in the United States and in some European countries (Muniesa et al. 2006; Paton and Paton 1998). The pathogenicity of *E. coli* O157 is associated with a number of virulence factors, in particular Shiga toxins 1 and 2 (encoded by *stx1* and *stx2* genes), and intimin (encoded by the *eae* gene).

Contamination of water bodies with fecal matter or inadequately treated water bodies have led to many outbreaks of *E. coli* O157 in the past, which showed that water contamination with *E. coli* O157 is a major issue of concern and needs more attention and research to control the outbreaks occurring worldwide (Bruneau et al. 2004; Hrudey et al. 2003; Olsen et al. 2002). In addition, irrigation water may play an important role in contaminating soils and vegetables with *E. coli* O157 (Solomon et al. 2002). *E. coli* O157:H7 from irrigation sources can adhere to plants, enter into plants, and survive for long periods of time (Franz and van Bruggen 2008; Franz et al. 2007; Semenov et al. 2009; Solomon et al. 2002). The 2006 multistate *E. coli* O157 outbreak associated with spinach was reported to be related to contaminated irrigation water (Gelting et al. 2011). However, associated risks of irrigation water for the contamination of *Escherichia coli* O157 have not been systematically evaluated.

Fecal pollution, including *E. coli* O157, is traditionally evaluated by detecting fecal indicator bacteria like fecal coliforms and generic *E. coli* (USEPA 2002). However, some studies provided evidence that these indicators may not be adequate to assess the health risks of surface water (Ahmed et al. 2009). Chigor et al. (2010) reported that some environmental factors,

including water turbidity, and concentrations of nitrate, phosphate and chloride were positively correlated to the population of fecal coliforms in a river used for fresh produce irrigation in Nigeria (Chigor et al. 2010). Rain events may promote pathogen transport from forested and grassed buffer zones into farm ponds (Gaertner et al. 2009). Survival of *E. coli* O157 strains in the environment depends on physiological differences in terms of the ability to use carbon compounds under aerobic versus semi-anaerobic conditions (Franz et al. 2011). Thus, the risk of survival of this pathogen is related to the availability of organic carbon and nitrogen sources as affected by microbial competition (Franz et al. 2008b; Franz et al. 2011), as well as average temperatures and their daily amplitudes (Semenov et al. 2007) and moisture conditions (Semenov et al. 2010). In addition a negative correlation between enteropathogen populations and microbial diversity has been reported for manure, soil and plants (Klerks et al., 2007; van Elsas et al., 2012; van overBeek et al., 2010).

The Suwannee Watershed, including all the creeks, streams, rivers and springs that feed the Suwannee River, is an ideal location to investigate the risk of transmission of *E. coli* to vegetables, because waterways and irrigation ponds are interspersed with irrigated vegetable production areas. This watershed extends from South Central Georgia to Northwest Central Florida and drains 9,950 square miles (http://www.epa.gov/owow/showcase/suwaneeriver/location.html). Fruit and vegetable crops in this region are mainly irrigated via drip systems, overhead sprinklers, or pivots that are used daily during periods of high evapotranspiration (May-October). The water source is primarily from surface water captured in small man-made reservoirs. Many ponds are supplemented with ground water, pumped from the Floridian Aquifer, to maintain volumes during periods without stream flow. Many growers are testing their water sources for generic *E. coli*, depending on the

requirements of the buyers of their produce. These requirements vary from being very lax to a strict requirement for a 5-day geometric mean of *E. coli* for each irrigation source. However, the relationship between these test results and the presence of virulent *E. coli* has not been established (Shelton et al. 2011).

The objectives of this study are to assess the prevalence of *E. coli* O157 in the irrigation ponds located in the Suwannee river watershed, evaluate the relationship between *E. coli* O157 occurrence and environmental factors, and investigate if current fecal indicators are adequate for the prediction of the presence and virulence of *E. coli* isolates.

Materials and Methods

Irrigation ponds and water sampling

The study was conducted from March 2011 to February 2012 within the Suwannee River watershed. The study area lies in the heart of a principal agricultural production area in the state Georgia, where vegetable acreage with supplemental irrigation is increasing rapidly. Ten irrigation ponds that serve as water sources for vegetable and field crops spread throughout this area were selected for the detection of *E. coli* O157 and measurement of environmental factors. The sizes and depths of the 10 ponds varied from 12,000 to 93,000 m^2, and 3 to more than 10 meters deep, respectively. The distances among the ponds range from 30 km to 200 km. In this study, two water samples (10 L) were collected monthly from each pond close to the pumps at two depths: near the irrigation water intake level and at 50 cm below the surface. Collected water samples were stored on ice in the field and transported to lab refrigerators for analysis. Totally 240 water samples were collected from these 10 irrigation ponds during the 12 months for *E. coli* O157 detection and measurement of environmental factors.

Isolation and identification of *E. coli* O157

Three 500 mL aliquots of each water sample were vacuum filtered through a 0.8 µm sterile isopore membrane and then a 0.45 µm sterile nitrocellulose membrane (Millipore Corporation, Billerica, MA). Both membranes were incubated without shaking at 37°C in 50 mL tubes containing 20 mL of Luria Bertani (LB) broth. After incubation for 18-20h, 1 mL enriched culture was transferred to 19 mL modified trypticase soy broth (TSB) containing 20 mg/L novobiocin (Jokinen et al. 2010) and incubated at 42 °C overnight. A bacterial suspension from each tube was first streaked on CHROMagar™ O157 with 2.5 mg/L potassium tellurite (CHROMagar, Paris, France) and incubated at 37°C for 24 h. CHROMagar™ O157 is a selective culture medium for isolation and enumeration of *E. coli* O157. Suspect colonies (color mauve) were then streaked on sorbitol MacConkey agar plates containing 50 mg/L cefixime and 2.5 mg/L tellurite and incubated at 37°C for 24 h. Because of the pathogen's inability to ferment sorbitol, *E. coli* O157 colonies are colorless in sorbitol MacConkey agar. Three colonies were randomly selected on each plate if more colonies that looked like *E. coli* O157 were present per plate. Positive colonies on sorbitol MacConkey agar plates were stored at -80 °C and tested for agglutination using O157 BD Difco™ *E. coli* Antisera (Difco, Detroit, MI) for confirmation. Streaking of colonies and agglutination tests were conducted in triplicate.

Virulence gene assays

Genomic DNA was extracted from *E. coli* O157 isolates using illustra™ bacteria genomicPrep Mini Spin Kit (GE Healthcare, Buckinghamshire, UK). The presence of virulence genes stx_1, stx_2 and *eae* were determined by PCR as reported before (Sanchez et al. 2010). *E. coli* O157 strains ATCC 43895 and ATCC 43888 were used as positive and negative controls, respectively.

The presence of these genes in *E. coli* O157 isolates serves as an indication of the potential virulence of the strains obtained.

Fecal indicator analysis

Water samples were screened for presence and absence of *E. coli* in Quanty-Tray cells and positive samples enumerated as MPN cells per 100 mL (Edberg et al. 1990). Two dilutions of each sample were assayed for MPN of generic *E. coli* (IDEXX Laboratories, Inc., Westbrook, ME). Fecal coliform bacteria (CFU/100 mL) analysis of two dilutions prepared from each sample were performed using membrane filtration techniques as described before (Stuart et al. 1977).

Chemical and physical parameters of irrigation water

During sampling from irrigation ponds, pond water was tested for temperature (Temp, ^{o}C), pH, specific conductance (SpCond, mS/cm), dissolved oxygen percentage (DO, %), dissolved oxygen concentration (DO Conc, mg/L), dissolved oxygen charge (DO Charge), turbidity (nephlometric turbidity units, NTU), and oxidation reduction potential (ORP, mV) with an YSI® 6600 Multiparameter Sonde (Yellow Springs, OH). All water samples were analyzed for the concentrations of nitrate (mg/L) and the soluble reactive portion of total phosphorus (Ortho-P, mg/L) and total suspended solids (TSS, mg/L) using standard analytical techniques in the lab (APHA 1998). Samples were also analyzed for dissolved organic carbon (DOC, mg/L) and total nitrogen (TN, mg/L) contents using TOC-VCPH/CPN and TNM-1 analyzers, respectively (Shimadzu Scientific Instruments, Kyoto, Japan). All tests were performed in duplicate.

Weather information collection and analysis

Temperature (°C) and rainfall data (mm) of the closest weather stations to the 10 irrigation ponds were collected from the Georgia automated environmental monitoring network (http://www.georgiaweather.net/). The distance from irrigation ponds to their closest weather stations ranged from 10-40 km. Average temperature and total rainfall from weather stations nearby the ponds in the second but last week (L2 avg Tm and L2 tot rain) and last week before sampling (L avg Tm and L tot rain) and between samplings (Average Tm and Total rainfall) were calculated based on the daily weather information.

Copiotrophic and culturable bacterial populations

Water samples were dilution-plated on carbon-rich and carbon-poor media (Senechkin et al., 2010) in triplicate and incubated for 2 or 15 days, respectively, to enumerate the populations of copiotrophic (Copio, CFU/100 mL on carbon-rich media) and culturable (Culturable, CFU/100 mL on carbon-poor media) bacteria, respectively, as reported elsewhere (Franz et al. 2008a).

DNA extraction of water samples

200 mL of each water sample was vacuum filtered through a 0.22 µm sterile nitrocellulose membrane (Millipore Corporation, Billerica, MA). In case of membrane clogging during filtration, multiple membranes were used. DNA of bacteria captured by the membranes was extracted using a PowerWater® DNA Isolation Kit (MOBIO Laboratories, Inc., Carlsbad, CA) and stored at -80°C until use.

Polymerase chain reaction (PCR) and denaturing gradient gel electrophoresis (DGGE)

To investigate a possible relationship between the occurrence of *E. coli* O157 and bacterial communities in the irrigation water, the Shannon-Wiener diversity index (Shannon diversity), Simpson diversity index (Simpson diversity), the species richness and evenness of bacterial

communities were determined as described before (Klerks et al. 2007). Briefly, extracted DNA was amplified by PCR using primers for16S rDNA genes of eubacteria with primers R1378 and U968 (Klerks et al. 2007) and DGGE was performed using a gradient maker and a Manostat® Carter Pumpdrive (Thermo Fisher Scientific Inc., Barrington, IL) and electrophoresis in 0.5x Tris-acetate-EDTA buffer for 16 h at 100 V at a constant temperature of 60°C. After electrophoresis, the gels were stained for 30 min with SYBR Gold nucleic acid gel stain (Life Technologies, Inc., Carlsbad, CA) and bands were visualized by the Molecular Imager® Gel Doc™ XR System (Bio-Rad Laboratories, Inc.) and standardized by referring to the DGGE marker. The DGGE banding patterns were analyzed using Gelcompar II software (version 1.61; Applied Maths, Woluwe, Belgium) for gel comparison. Each gel contained four marker lanes as standards and background corrections were performed before identification of bands with 5% significance threshold. Correspondence of operational taxonomic units (OTUs or bands) between different samples was performed with 1% dynamic range settings. DGGE analysis was done in duplicate on separate gels. The Shannon diversity, Simpson diversity, species richness and evenness of bacterial communities were calculated as the means of the two replicas.

Statistical analysis

The Chi square (χ^2) test was conducted to compare the occurrence of *E. coli* O157 in surface and subsurface water samples. Space (pond) and time (season) effects on the occurrence of *E. coli* O157 were assessed by logistic regression analysis. Unpaired t-tests were performed to compare log-transformed population densities of generic *E. coli* and fecal coliform bacteria in the water samples with those of *E. coli* O157 carrying virulence genes and the samples without *E. coli* O157 collected in the same months. Data were tested for normality after log-transformation. Biserial correlation coefficients were calculated to evaluate the correlations between the 27

environmental factors (Table 1) and the occurrence of *E. coli* O157 isolated from irrigation ponds (Portney and Watkins 1993). Biserial analysis is recommended for assessing the correlation between a continuous variable and a binomial (dichotomous) one (Portney and Watkins 1993). In this case the presence/absence of *E. coli* O157 was the dichotomous variable, and the values of the parameters tested were continuous variables.

For discriminant analysis, the factors to be considered should be normally distributed and not be strongly correlated with each other (Afifi and Clark 1984). Moreover, the number of factors should be considerably smaller than the number of observations. Therefore, all data were first standardized and normalized and all normalized factors were subjected to Pearson's correlation analysis, so that groups of factors could be distinguished that 'belonged together' (were correlated). Grouping was based on relatively high Pearson's correlation coefficient values ($r>0.7$) between factors, and one factor in each group was selected to be included in the discriminant analyses (to avoid highly correlated factors in the analyses). Stepwise discriminant and canonical discriminant analyses were performed based on the measurements of 18 selected environmental factors after standardization and normalization in SAS. The inclusion criterion for discriminant analysis was $P=0.15$ (Afifi and Clark 1984). Classification values were assigned to all *E. coli* O157 positive water samples and the same number of randomly selected negative samples (because the number of samples in the different classes should be about equal for the analysis). Stepwise discriminant analysis was used to identify variables that contributed most to the classification. Canonical discriminant analysis was used to determine the magnitude and direction of the association of individual variables with the classification variable. The same discriminant analyses were performed based on the relative intensities of 65 OTUs of DGGE

data to group the 240 water samples according to the predicted presence or absence of *E. coli* O157.

Chi square tests were conducted in Microsoft Excel. All other statistical analyses (logistic regression, t-tests, correlation and discriminant analyses) were performed using SAS (SAS release 9.2, SAS Institute Inc., Cary, NC).

Results

Spatial and seasonal distribution of *E. coli* O157 in irrigation water

All ten irrigation ponds assessed in the southern Georgia region of the Suwannee River Watershed were positive for *E. coli* O157 in water samples at some time point during this survey (Fig. 1A). The overall percentages of positive samples for surface (11.7%) and subsurface (17.5%) water samples were not significantly different (χ^2 =1.639, P = 1), so the data for the prevalence of *E. coli* O157 in both water samples from each pond per month were lumped in this study with an overall prevalence of 14.6%. Although there were no statistically significant spatial (pond) differences within the watershed with respect to the occurrence of *E. coli* O157 (P = 0.1025), pond 6 was most frequently positive (50% of the sampling times), while ponds 2, 4, 5, 7 and 8 were most frequently negative (83.3%) for *E. coli* O157.

There were significant seasonal differences in the detection of *E. coli* O157 (P < 0.001, Fig. 1A). The pathogen was isolated every month from May to December of 2011 from some irrigation ponds. The overall percentages of *E. coli* O157 occurrence in summer (June to August, 51.7%) and fall (September to November, 31.1%) were significantly higher than those in spring (March to May, 10.3%) and winter (December to February, 6.9%).

Out of all colonies isolated from the water samples during this study, only 44 were positive in CHROMagar™ O157 and sorbitol MacConkey, as well as for the agglutination tests. From the 44 isolated *E. coli* O157 colonies, three isolates (from pond 10 in August and pond 10 in June and August) were positive for stx_2 gene, one isolate (from pond 4 in November) was positive for stx_1 gene, and one isolate (from pond 2 in July) carried all three stx_1, stx_2 and *eae* virulence genes as indicated by conventional PCR. Pond 2 was adjacent to a tomato field, and discarded tomato fruits were noticed floating in the pond when sampling in July. The populations of generic *E. coli* and fecal coliform bacteria in the water samples with or without *E. coli* O157 carrying virulence genes were not significantly different (generic *E. coli*, P = 0.569; fecal coliforms, P = 0.311).

Relationships among environmental factors and their correlations to the occurrence of *E. coli* O157 in irrigation ponds

Overall, 27 environmental factors were evaluated as potential indicators for the risk of *E. coli* O157 contamination (Table 1, Table S1). Based on biserial correlation analysis, Average Tm, Total rainfall, culturable bacteria, copio bacteria and fecal coliforms were positively correlated to *E. coli* O157 occurrence in samples from the ten irrigation ponds. However, ORP, TN, Nitrate and DO concentration were negatively correlated with the occurrence. The positive correlation between Simpson diversity and *E. coli* O157 occurrence was almost significant (P = 0.0891). The temporal variation in the average values of the main environmental factors that correlated with *E. coli* O157 was similar or opposite to the temporal variation in *E. coli* O157 occurrence (Fig. 1 B-D, Table S2).

Environmental factors, which were highly positively correlated to each other (r > 0.7), were assigned to one factor group (Table S3). The factors in the first column of the table, which

had the highest correlation to *E. coli* O157 occurrence in each factor group, as well as independent factors (that were not correlated to other environmental factors) were selected for further discriminant analysis.

Correlations between OTUs of DGGE and the occurrence of *E. coli* O157 in irrigation ponds

In total, 65 OTUs (bands) were generated through DGGE analysis of the bacterial communities in the 240 water samples. Based on biserial correlation analysis, the intensities of 12 bands were significantly correlated to *E. coli* O157 occurrence ($P < 0.05$), and the correlations of another 6 bands were possibly significant ($0.1 > P > 0.05$) (Table S4). None of these bands were highly correlated ($r > 0.7$) to each other.

Environmental factors and bacterial OTUs affecting the distinction between the presence and absence of *E. coli* O157 in irrigation ponds

Discriminant analysis on 58 water samples positive for *E. coli* O157 (surface and subsurface water samples were considered as replicate samples from the same pond because averages of both samples were statistically indistinguishable) and the same number of randomly selected negative water samples resulted in significant separation of water samples with and without *E. coli* O157. The separation of samples into two *E. coli* O157 occurrence classes, based on the environmental factors (Fig. 2) or OTUs of DGGE (Fig. 3), are illustrated in a plot of canonical variable one versus the variable with the highest correlation to occurrence of the pathogen. Canonical variable one is the linear combination that maximized the correlation with *E. coli* O157 occurrence.

Stepwise discriminant analysis for the presence or absence of *E. coli* O157 in each water sample identified six environmental factors (Table 2) and 26 DGGE OTUs (data not shown) that contributed to the classification ($P < 0.05$). The selected OTUs by stepwise discriminant analysis that affected the classification of *E. coli* O157 presence included all 12 significantly correlated bands (Table S4) and other bands with relatively high correlation coefficients. The selected environmental factors and OTUs did not change when different sets of random negative samples were used for the analysis. Because the use of 12 microbial species for estimating the presence of *E. coli* O157 is not practical, no attempt to sequence the bands with significant correlations was made.

Discussion

E. coli O157 was detected in the irrigation ponds located in the Suwannee River Watershed with an overall prevalence of 14.6%, which indicated that irrigation water could be a contamination source of this pathogen for fresh vegetables and fruits in the field. Although there are no records linking *E. coli* O157 infections to contaminated vegetables or fruits in this region, there were several outbreaks of *E. coli* O157 due to contaminated water sources like swimming pools. Since the studied pond water serves as irrigation for vegetable crops like tomatoes and lettuce that are likely eaten raw, the epidemiological importance of *E. coli* O157 in this water source cannot be overlooked. The concern here is heightened by recent studies about the internalization of foodborne pathogens into the edible parts of plants (Franz et al. 2007; Gu et al. 2011). Enteric pathogens coming from feces of wild animals could have reached the irrigation ponds, potentially exposing the plants to *E. coli* O157.

The virulence of *E. coli* O157 has mainly been associated with the production of Shiga toxin genes (*stx1*and *stx2*) and the intimin gene (*eae*), although the presence of *stx1* and/or *stx2* does not by itself prove that the strains are pathogenic. In this study, five isolates were positive for at least one of these virulence genes, but only one isolate from pond 2 with discarded tomato fruits carried all three virulence genes. These isolates may pose serious threats for human health when the virulent genes are expressed in sufficient quantities. However, it should be realized that there are many different genotypes of *E. coli* O157, including small differences in the toxin genes with several *stx2* subtypes, that may differ in their pathogenicity towards humans (Franz et al. 2012). Conversely, other serotypes not screened in this study might be associated with EHEC disease; therefore, the presence of potential pathogens may be underestimated.

The significantly higher occurrence of *E. coli* O157 during the summer and fall could be partly attributable to run-off of soil by rain into the irrigation ponds and also to the higher multiplication rates at the higher temperatures during those seasons (Semenov et al. 2007; Semenov et al. 2009). Irrigation of the fruit and vegetable crops in this region were mainly applied from May to October to compensate for the high evapotranspiration, which is the period with high risks of contamination. If this period cannot be avoided for the planting of vegetable crops, like fall tomato crops, more intensive monitoring of the pathogen should be conducted.

Populations of fecal coliforms, but not of generic *E. coli*, were correlated to the occurrence of *E. coli* O157, and might be a suitable indicator for the prediction of this pathogen in irrigation ponds. However, there was a poor correlation between fecal indicators and potentially virulent *E. coli* O157 isolated in this study as well as in previous studies (Ahmed et al. 2009). Therefore, testing for fecal indicators alone may not be adequate to assess the microbiological quality of surface water and potential health risks. In addition, the presence of

343 virulence genes in *E. coli* O157 isolates was not related to the occurrence of *E. coli* O157. For
344 example, pond 6 was most frequently positive (50%) for *E. coli* O157 occurrence, but no
345 virulence genes were amplified from the bacterial colonies isolated from this pond. Thus, the use
346 of the tested indicators alone might not be appropriate for estimating the occurrence of virulent
347 *E. coli* O157.

348 Correlation and discriminant analyses indicated that the presence of *E. coli* O157 in water
349 samples from irrigation ponds was positively associated with average temperature, total rainfall,
350 populations of culturable, copiotrophic and fecal coliform bacteria, which were higher in *E. coli*
351 O157 positive water samples, while the TN, ORP and DO concentration were negatively
352 correlated to the occurrence of this pathogen. Higher temperatures correlated with higher
353 presence of the pathogen. This might be due to a combination of factors including the closer-to-
354 optimal temperatures for growth of the pathogen as well as populations of culturable,
355 copiotrophic and fecal coliform bacteria; and an increase in wildlife (and their droppings) during
356 warmer seasons. Temperature, total rainfall, Simpson diversity and evenness of bacterial
357 communities, culturable bacterial population and phosphorous concentration were consistently
358 selected by stepwise discriminant analysis for the classification of the presence of *E. coli* O157
359 (Table 2). The negative correlation between TN and the presence of *E. coli* O157 could be
360 explained by the relative infrequent samplings (once a month), while microbial dynamics have a
361 time constant of less than a day. Eutrophication by run-off from the first rainfall event and
362 discarded produce in early Spring probably led to a temporary increase in TN, immediately
363 followed by an increase in copiotrophic bacteria and a decline in TN; the increase in coliforms
364 and *E. coli* O157 occurred in early and midsummer, after TN had already declined. The
365 difference in the speed of microbial dynamics and the sampling interval could have been the

366 reason why TN was not selected by stepwise discriminant analysis to explain the occurrence of
367 *E. coli* O157. Nevertheless, the significant positive correlation between culturable bacterial
368 populations and the pathogen and the selection of phosphorus concentration as a classification
369 factor indicate a potential relation between *E. coli* O157 occurrence and eutrophication.

370 To minimize the risk of contamination of irrigation water with *E. coli* O157, it is
371 therefore important to prevent eutrophication by adding waste materials into the ponds. Under
372 eutrophied conditions, the predictability of the survival of *E. coli* O157 proved to be worse than
373 in more oligotrophic conditions (Semenov et al. 2011). In addition, electrolyzed oxidizing water
374 with higher oxidation-reduction potential values could be used for the inactivation of *E. coli*
375 O157 (Kim et al. 2000; Stevenson et al. 2004), because this pathogen survives better in anaerobic
376 conditions (Semenov et al. 2011). Therefore, increasing the pond water ORP by enhancing the
377 flow rate and biofiltration capacity, or pumping air into irrigation ponds to increase dissolved
378 oxygen concentration, especially in the seasons with higher temperature and rainfall, would be
379 suggested to reduce the risks of the contamination with *E. coli* O157 during irrigation.

380 The presence and absence of *E. coli* O157 was clearly distinguished by discriminant
381 analysis based on 26 selected OTUs generated from DGGE analysis. Among these 26 OTUs,
382 there were 18 OTUs for which the relative band intensities were correlated to the occurrence of
383 *E. coli*. The diversity and evenness of bacterial communities were selected as classification
384 factors by stepwise discriminant analysis. In addition, the presence or absence of *E. coli* O157
385 was also related to the composition of the bacterial communities in these irrigation ponds.
386 Further studies will be conducted to sequence the selected OTUs that were significantly
387 correlated to the presence or absence of *E. coli* O157 to identify potential biomarkers or
388 biocontrol agents for the pathogen. In addition, other methods, like microbial source tracking,

could be conducted in the future using extracted genomic DNAs to identify the sources of fecal contamination of the pond water (Dorai-Raj et al., 2012). There were no fences around the 10 irrigation ponds sampled in this study, and natural vegetation around the ponds varies in different planting seasons. Wild animals could be the main source of contamination with *E. coli* O157 in these ponds. Various wild animals live in this watershed, including deer, hogs, raccoons, squirrels, marsh rabbits, opossums, birds, armadillos, as well as various reptile and amphibian species. All of these animals except armadillos were reported to be carriers of *E. coli* (Jokinen et al. 2011; Leotta et al. 2006; Mohapatra et al. 2007). The seasonality of *E. coli* O157 in irrigation ponds with high frequency in summer and fall is similar to the prevalence of this pathogen on the hides and in the feces of cattle (Berry and Wells 2010), which may indicate that the contamination source is likely to be animals that fed on plants, like ruminants.

Results obtained in this study may lead to hypotheses about the factors that promote or suppress *E. coli* O157 in irrigation ponds. Further studies will be conducted for more detailed genotyping of the *E. coli* O157 isolates. In this project, real-time PCR based on the extracted genomic DNAs using the specific primers (Ahmed et al., 2009) were conducted to test the prevalence and population of *E. coli* O157 in pond water samples in addition to the enrichment method. However, no positive results were detected for any of the 10 ponds during the whole year (data not presented) probably due to the low population density of this bacterium in the pond water. Further studies using a more sensitive method would be useful to quantify the populations of STEC in irrigation ponds.

Supplementary Materials

411 **Table S1.** Measured variables (means and standard deviations of ten irrigation ponds over a
412 period of 12 months) for water samples that were positive or negative for *E. coli* O157.

413 **Table S2.** Mean values and standard deviations of highly correlated environmental factors from
414 March, 2011 to February, 2012 in 10 irrigation ponds.

415 **Table S3.** Groups of environmental factors with high positive correlations to each other ($r > 0.7$).

416 **Table S4.** Operational taxonomic units (OTUs) of DGGE data with significant correlations to the
417 occurrence of *E. coli* O157 isolated from irrigation ponds ($P<0.1$).

418

419 ## Acknowledgements

420 We appreciate the inspiration of Dr. Paige Adams during the experimental design and her tireless
421 efforts with water sample collection and evaluation, and are very sad that she was not able to see
422 the results of our work due to her untimely death. We are also grateful to Susie Richardson,
423 Herman Henry and Debbie Coker for field and laboratory assistance. This research was funded
424 by Dr. Douglas Archer, IFAS, University of Florida and by the Center for Produce Safety at UC
425 Davis, CA.

References

Afifi, A.A., and Clark, V. 1984. Computer-aided multivariate analysis. Van Nostrand Beinhold, New York.

Ahmed, W., Sawant, S., Huygens, F., Goonetilleke, A., and Gardner, T. 2009. Prevalence and occurrence of zoonotic bacterial pathogens in surface waters determined by quantitative PCR. Water Res. **43**(19): 4918-4928. doi:10.1016/j.watres.2009.03.041.

APHA. 1998. Standard methods for the examination of water and waste water, 20th Edition. American Public Health Association, Washington DC.

Berry, E.D., and Wells, J.E. 2010. *Escherichia coli* O157:H7 recent advances in research on occurrence, transmission, and control in cattle and the production environment. Adv. Food Nutr. Res. **60**: 67-117. doi:10.1016/s1043-4526(10)60004-6.

Bruneau, A., Rodrigue, H., Ismael, J., Dion, R., and Allard, R. 2004. Outbreak of *E. coli* O157:H7 associated with bathing at a public beach in the Montreal-Centre region. Canada communicable disease report = Releve des maladies transmissibles au Canada **30**(15): 133-136.

Chigor, V.N., Umoh, V.J., and Smith, S.I. 2010. Occurrence of *Escherichia coli* O157 in a river used for fresh produce irrigation in Nigeria. Afr. J. Biotechnol. **9**(2): 178-182.

Dorai-Raj, S., O'Grady, J., Cormican, M. and Colleran, E. 2012. Identification of host-specific Bacteroidales 16S rDNA sequences from human sewage and ruminant feces. J. Basic Microbiol., **52**(3): 277–284. doi: 10.1002/jobm.201100184

Edberg, S.C., Allen, M.J., Smith, D.B., and Kriz, N.J. 1990. Enumeration of total coliforms and *Escherichia coli* from source water by the defind substrate technology. Appl. Environ. Microb. **56**(2): 366-369.

Franz, E., Semenov, A.V., Termorshuizen, A.J., de Vos, O.J., Bokhorst, J.G., and van Bruggen, A.H. 2008a. Manure-amended soil characteristics affecting the survival of *E. coli* O157:H7 in 36 Dutch soils. Environ. Microbiol. **10**(2): 313-327. doi:EMI1453.

Franz, E., Semenov, A.V., and van Bruggen, A.H. 2008b. Modelling the contamination of lettuce with *Escherichia coli* O157:H7 from manure-amended soil and the effect of intervention strategies. J. Appl. Microbiol. **105**(5): 1569-1584. doi:JAM3915.

Franz, E., and van Bruggen, A.H. 2008. Ecology of *E. coli* O157:H7 and *Salmonella enterica* in the primary vegetable production chain. Crit. Rev. Microbiol. **34**(3-4): 143-161. doi:901940522.

Franz, E., van Hoek, A.H.A.M., Bouw, E., and Aarts, H.J.M. 2011. Variability of *Escherichia coli* O157:H7 strain survival in manure-amended soil in relation to strain origin, virulence profile, and carbon nutrition profile. Appl. Environ. Microb. **77**(22): 8088-8096. doi:10.1128/aem.00745-11.

Franz, E., van Hoek, A.H.A.M., van der Wal, F.J., de Boer, A., Zwartkruis-Nahuis, A., van der Zwaluw, K., Aarts, H.J.M., and Heuvelink, A.E. 2012. Genetic features differentiating bovine, food, and human isolates of Shiga toxin-producing *Escherichia coli* O157:H7 in the Netherlands. J. Clin. Microbiol. **50**(3): 772-780. doi:10.1128/jcm.05964-11.

Franz, E., Visser, A.A., Van Diepeningen, A.D., Klerks, M.M., Termorshuizen, A.J., and van Bruggen, A.H. 2007. Quantification of contamination of lettuce by GFP-expressing *Escherichia coli* O157:H7 and *Salmonella enterica* serovar Typhimurium. Food Microbiol. **24**(1): 106-112. doi:S0740-0020(06)00072-4.

Gaertner, J.P., Garres, T., Becker, J.C., Jimenez, M.L., Forstner, M.R.J., and Hahn, D. 2009. Temporal analyses of salmonellae in a headwater spring ecosystem reveals the effects of precipitation and runoff events. J. Water Health. **7**(1): 115-121. doi:Doi 10.2166/Wh.2009.138.

Gelting, R.J., Baloch, M.A., Zarate-Bermudez, M.A., and Selman, C. 2011. Irrigation water issues potentially related to the 2006 multistate *E. coli* O157:H7 outbreak associated with spinach. Agr. Water. Manage. **98**(9): 1395-1402. doi:10.1016/j.agwat.2011.04.004.

Gu, G., Hu, J., Cevallos-Cevallos, J.M., Richardson, S.M., Bartz, J.A., and van Bruggen, A.H. 2011. Internal colonization of *Salmonella enterica* serovar Typhimurium in tomato plants. PLoS One. **6**(11): e27340. doi:10.1371/journal.pone.0027340.

Hrudey, S.E., Payment, P., Huck, P.M., Gillham, R.W., and Hrudey, E.J. 2003. A fatal waterborne disease epidemic in Walkerton, Ontario: comparison with other waterborne outbreaks in the developed world. Water. Sci. Technol. **47**(3): 7-14.

Jokinen, C., Edge, T.A., Ho, S., Koning, W., Laing, C., Mauro, W., Medeiros, D., Miller, J., Robertson, W., Taboada, E., Thomas, J.E., Topp, E., Ziebell, K., and Gannon, V.P.J. 2011. Molecular subtypes of *Campylobacter* spp., *Salmonella enterica*, and *Escherichia coli* O157:H7 isolated from faecal and surface water samples in the Oldman River watershed, Alberta, Canada. Water Res. **45**(3): 1247-1257. doi:10.1016/j.watres.2010.10.001.

Jokinen, C.C., Schreier, H., Mauro, W., Taboada, E., Isaac-Renton, J.L., Topp, E., Edge, T., Thomas, J.E., and Gannon, V.P.J. 2010. The occurrence and sources of *Campylobacter* spp., *Salmonella enterica* and *Escherichia coli* O157:H7 in the Salmon River, British Columbia, Canada. J. Water Health. **8**(2): 374-386. doi:10.2166/wh.2009.076.

Kim, C., Hung, Y.C., and Brackett, R.E. 2000. Roles of oxidation-reduction potential in electrolyzed oxidizing and chemically modified water for the inactivation of food-related pathogens. J. Food Protect. **63**(1): 19-24.

Klerks, M.M., Franz, E., van Gent-Pelzer, M., Zijlstra, C., and van Bruggen, A.H. 2007. Differential interaction of *Salmonella enterica* serovars with lettuce cultivars and plant-microbe factors influencing the colonization efficiency. Isme J. **1**(7): 620-631. doi:ismej200782.

Leotta, G.A., Deza, N., Origlia, J., Toma, C., Chinen, I., Miliwebsky, E., Iyoda, S., Sosa-Estani, S., and Rivas, M. 2006. Detection and characterization of Shiga toxin-producing *Escherichia coli* in captive non-domestic mammals. Vet. Microbiol. **118**(1-2): 151-157. doi:10.1016/j.vetmic.2006.07.006.

Mohapatra, B.R., Broersma, K., and Mazumder, A. 2007. Comparison of five rep-PCR genomic fingerprinting methods for differentiation of fecal *Escherichia coli* from humans, poultry and wild birds. Fems Microbiol. Lett. **277**(1): 98-106. doi:10.1111/j.1574-6968.2007.00948.x.

Muniesa, M., Jofre, J., Garcia-Aljaro, C., and Blanch, A.R. 2006. Occurrence of *Escherichia coli* O157:H7 and other enterohemorrhagic *Escherichia coli* in the environment. Environ. Sci. Technol. **40**(23): 7141-7149.

Olsen, S.J., Miller, G., Breuer, T., Kennedy, M., Higgins, C., Walford, J., McKee, G., Fox, K., Bibb, W., and Mead, P. 2002. A waterborne outbreak of *Escherichia coli* O157 : H7 infections and hemolytic uremic syndrome: Implications for rural water systems. Emerg. Infect. Dis. **8**(4): 370-375.

Paton, A.W., and Paton, J.C. 1998. Detection and characterization of shiga toxigenic *Escherichia coli* by using multiplex PCR assays for *stx(1)*, *stx(2)*, *eaeA*, enterohemorrhagic E-coli *hlyA*, *rfb*(O111), and *rfb*(O157). J. Clin. Microbiol. **36**(2): 598-602.

Portney, L.R., and Watkins, M.P. 1993. Foundation of clinical research, applications to practice. Appleton & Lange, East Norwark, Connecticut.

Sanchez, S., Martinez, R., Garcia, A., Vidal, D., Blanco, J., Blanco, M., Blanco, J.E., Mora, A., Herrera-Leon, S., Echeita, A., Alonso, J.M., and Rey, J. 2010. Detection and characterisation of O157:H7 and non-O157 Shiga toxin-producing *Escherichia coli* in wild boars. Vet. Microbiol. **143**(2-4): 420-423. doi:10.1016/j.vetmic.2009.11.016.

Semenov, A.V., Franz, E., and van Bruggen, A.H.C. 2010. COLIWAVE a simulation model for survival of *E. coli* O157:H7 in dairy manure and manure-amended soil. Ecol. Model. **221**(4): 599-609. doi:10.1016/j.ecolmodel.2009.10.028.

Semenov, A.V., van Bruggen, A.H., van Overbeek, L., Termorshuizen, A.J., and Semenov, A.M. 2007. Influence of temperature fluctuations on *Escherichia coli* O157:H7 and *Salmonella enterica* serovar Typhimurium in cow manure. Fems Microbiol. Ecol. **60**(3): 419-428. doi:FEM306.

Semenov, A.V., van Overbeek, L., Termorshuizen, A.J., and van Bruggen, A.H. 2011. Influence of aerobic and anaerobic conditions on survival of *Escherichia coli* O157:H7 and *Salmonella enterica* serovar Typhimurium in Luria-Bertani broth, farm-yard manure and slurry. J. Environ. Manage. **92**(3): 780-787. doi:S0301-4797(10)00363-4.

Semenov, A.V., van Overbeek, L., and van Bruggen, A.H. 2009. Percolation and survival of *Escherichia coli* O157:H7 and *Salmonella enterica* serovar Typhimurium in soil amended with contaminated dairy manure or slurry. Appl. Environ. Microbiol. **75**(10): 3206-3215. doi:AEM.01791-08.

Senechkin, I. V., Speksnijder, A. G., Semenov, A. M., van Bruggen, A. H. C., and van Overbeek, L. S. 2010. Isolation and partial characterization of bacterial strains on low organic carbon medium from soils fertilized with different organic amendments. Microb. Ecol. **60**(4):829-839. doi: 10.1007/s00248-010-9670-1

Shelton, D.R., Karns, J.S., Coppock, C., Patel, J., Sharma, M., and Pachepsky, Y.A. 2011. Relationship between *eae* and *stx* Virulence Genes and *Escherichia coli* in an Agricultural Watershed: Implications for Irrigation Water Standards and Leafy Green Commodities. J. Food Prot. **74**(1): 18-23. doi:10.4315/0362-028x.jfp-10-241.

Solomon, E.B., Yaron, S., and Matthews, K.R. 2002. Transmission of *Escherichia coli* O157:H7 from contaminated manure and irrigation water to lettuce plant tissue and its subsequent internalization. Appl. Environ. Microb. **68**(1): 397-400. doi:10.1128/aem.68.1.397-400.2002.

Stevenson, S.M.L., Cook, S.R., Bach, S.J., and McAllister, T.A. 2004. Effects of water source, dilution, storage, and bacterial and fecal loads on the efficacy of electrolyzed oxidizing water for the control of *Escherichia coli* O157:H7. J. Food Prot. **67**(7): 1377-1383.

Stuart, D.G., McFeters, G.A., and Schillinger, J.E. 1977. Membrane filter technique for the quantification of stressed fecal coliforms in the aquatic environment. Appl. and Environ. Microb. **34**(1): 42-46.

USEPA. 2002. Improved enumeration methods for the recreational water quality indicators: *Enterococci* and *Escherichia coli.* EPA/821/R-97/004. U. S. Environmental Protection Agency, Washington, DC.

van Elsas, J. D., Chiurazzi, M., Mallon, C. A., Elhottova, D., Kristufek, V., and Salles, J. F. 2012. Microbial diversity determines the invasion of soil by a bacterial pathogen. Proc. Natl. Acad. Sci. **109**(4):1159-1164. doi: 10.1073/pnas.1109326109

van Overbeek, L. S., Franz, E., Semenov, A.V., de Vos, O. J., and van Bruggen, A. H. C. 2010. The effect of the native bacterial community structure on the predictability of *E. coli* O157:H7 survival in manure-amended soil. Lett. Appl. Microbiol. **50**(4):425-430. doi: 10.1111/j.1472-765X.2010.02817.x

Figure Legends

Fig. 1. Occurrence of *E. coli* O157 and dynamic changes in environmental factors. A, Occurrence of *E. coli* O157 from March, 2011 to February, 2012 in 10 irrigation ponds; B, dynamic changes in total rainfall and average temperature nearby the ponds between samplings; C, dynamic changes in total nitrogen and dissolved oxygen concentration of ponds at the time of sampling for *E. coli* O157; D, dynamic changes in oxidation reduction potential and fecal coliform populations of ponds at the time of sampling for *E. coli* O157. P1 - P10 are the 10 irrigation ponds located in the Suwannee River watershed.

Fig. 2. Grouping of *E. coli* O157 occurrence in water samples from 10 irrigation ponds located in the Suwannee River watershed by discriminant analysis of environmental factors based on canonical function one and standardized values of temperature. Small circles represent water samples without *E. coli* O157, and stars represent *E. coli* O157 positive water samples. *E. coli* O157 positive samples were mainly grouped in a big circle.

Fig. 3. Grouping of *E. coli* O157 occurrence in water samples from 10 irrigation ponds located in the Suwannee River watershed by discriminant analysis of DGGE data based on canonical function one and standardized values of band 5. Small circles represent water samples without *E. coli* O157, and stars represent *E. coli* O157 positive water samples. *E. coli* O157 positive samples were mainly grouped in a big rectangle.

624 **Table 1.** Correlation coefficients between environmental factors and the occurrence of *E. coli*

625 O157 isolated from irrigation ponds.

626

Environment factors	Biserial correlation coefficients to *E. coli* O157 occurrence	P value
L2 avg Tm[1]	0.424	0
Average Tm[2]	0.416	0
L avg Tm[3]	0.348	0
Temp[4]	0.324	0
Total rainfall[5]	0.249	0.0001
ORP[6]	-0.245	0.0001
TN[7]	-0.232	0.0003
Culturable[8]	0.223	0.0005
Fecal coliforms[9]	0.222	0.0005
L2 tot rain[10]	0.212	0.0009
Nitrate[11]	-0.161	0.0127
DO Conc[12]	-0.142	0.0282
Copio[13]	0.131	0.0432
DO Charge[14]	0.118	0.0682
Simpson diversity[15]*	0.110	0.0891
Ortho-P[16]	-0.104	0.1086
pH[17]	-0.104	0.1160
L tot rain[18]	0.081	0.2142
Shannon diversity[19]	-0.078	0.2303
Species richness[20]	0.055	0.3936
TSS[21]	0.055	0.3965
Generic *E. coli*[22]	0.054	0.4021
SpCond[23]	-0.033	0.6097
Evenness[24]	0.027	0.6798
Turbidity[25]	-0.018	0.7780
DOC[26]	0.014	0.8314
DO%[27]	-0.001	0.9897

627

[1] Average temperature in the pond region in the second to last week before sampling (°C); [2] average temperature in the pond region between samplings (°C); [3] average temperature in the pond region in the last week before sampling (°C); [4] temperature of the pond water during sampling (°C); [5] total rainfall of the pond region between samplings (mm); [6] oxidation reduction potential (mV); [7] total nitrogen (mg/L); [8] population of culturable bacteria (CFU/100ml); [9] population of fecal coliform bacteria (CFU/100ml); [10] total rainfall in the pond region in the second but last week before sampling (mm); [11] nitrate concentration (mg/L); [12] dissolved oxygen concentration (mg/L); [13] population of copiotrophic bacteria (CFU/100ml); [14] dissolved oxygen charge (mV); [15] Simpson diversity index of bacterial community; [16] soluble reactive portion of total phosphorus (mg/L); [17] pH value of water samples; [18] total rainfall in the pond region in the last week before sampling (mm); [19] Shannon-Wiener diversity index of bacterial community; [20] species richness of bacterial community; [21] total suspended solids (mg/L); [22] population of generic *E. coli* (CFU/100ml); [23] specific conductance (mS/cm); [24] evenness of bacterial community; [25] turbidity of water samples (NTU); [26] total organic carbon (mg/L); [27] dissolved oxygen percentage (%); * Simpson index is negatively correlated to the diversity of bacterial communities. Environment factors with significant correlations ($P < 0.05$) to *E. coli* O157 occurrence are shown in bold font.

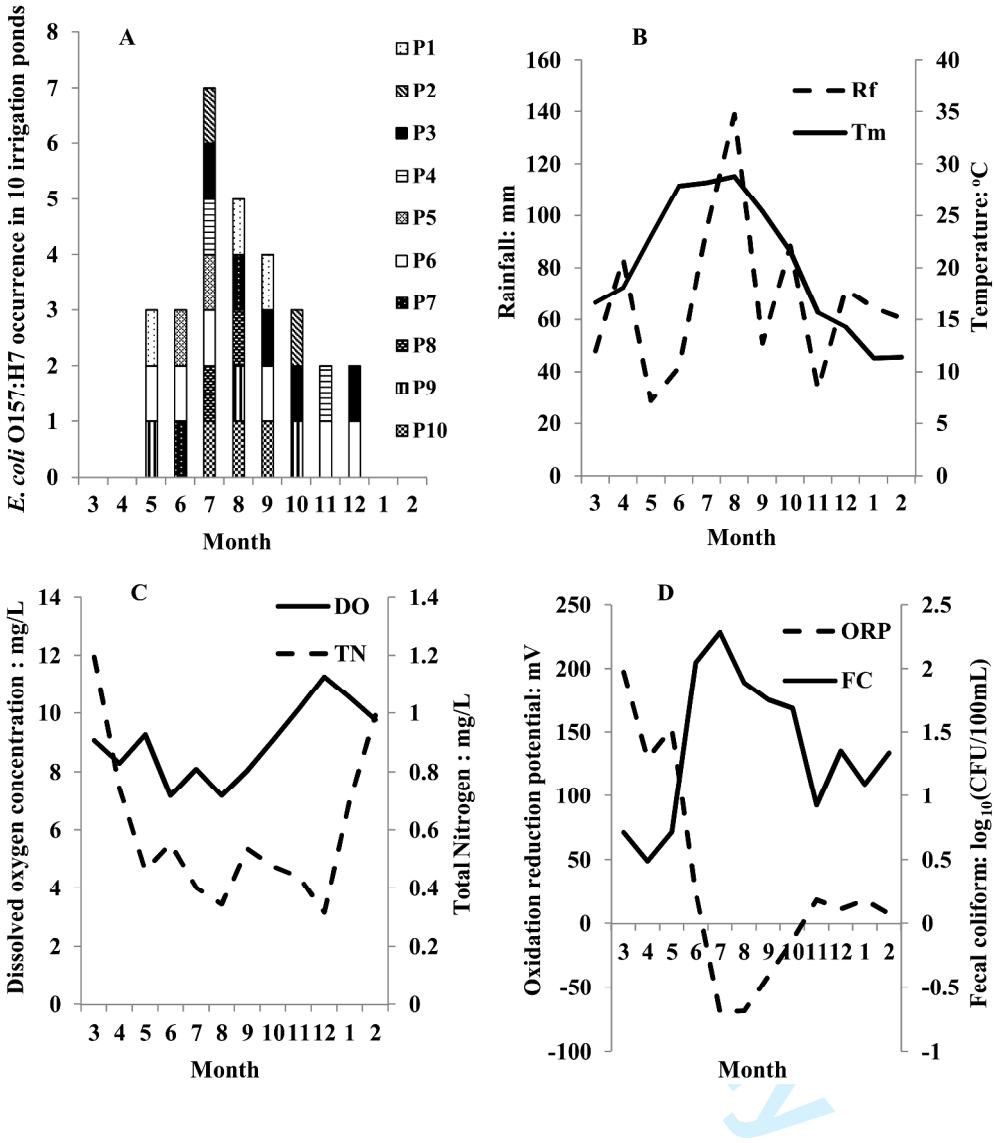
646 **Table 2.** Environmental factors that contributed significantly to classification of the occurrence
647 of *E. coli* O157 in irrigation ponds by stepwise and canonical discriminant analyses. All data
648 presented originate from canonical discriminant analysis; the order of the factors comes from the
649 stepwise discriminant analysis.

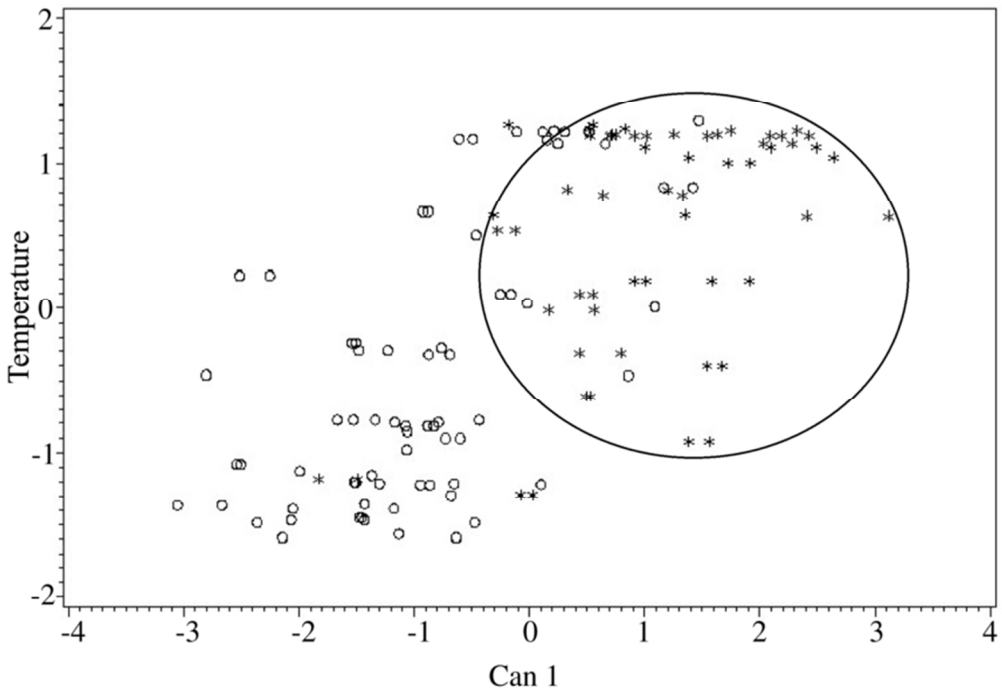
650

Factors[2]	Canonical function 1 [1]		*E. coli* O157 occurrence/ Mean values of factors	
	Standardized Coefficients[3]	Pooled within-class correlation[4]	Negative	Positive
L2 avg Tm (^{o}C)	0.869	0.789	18.65	25.05
Total rainfall (mm)	0.289	0.281	61.2	85.87
Simpson diversity (-)	0.463	0.461	0.0653	0.071
Evenness (-)	0.417	0.414	0.9444	0.9457
Culturable (CFU/100ml)	0.216	0.211	5759.6	11989
Ortho-P (mg/L)	0.224	0.223	0.0443	0.0677

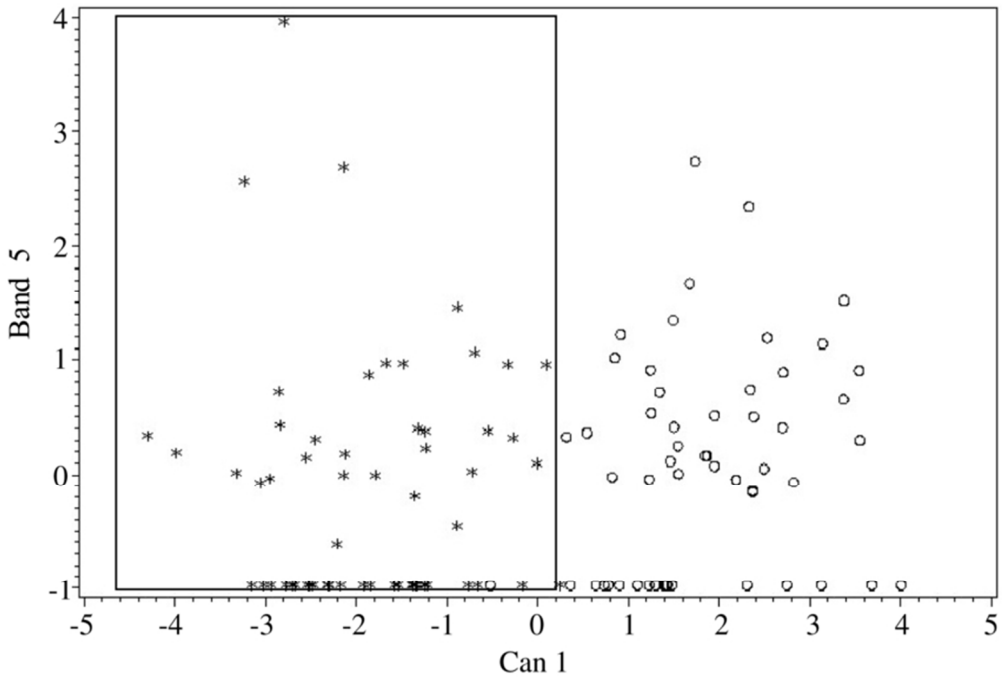
651

652 [1] Canonical function 1 is the linear combination of selected factors that best explain the
653 presence/absence grouping of *E. coli* O157; if there are only two groups, there is only one
654 canonical function.
655 [2] Environmental factors are listed in order of selection by stepwise discriminant analysis. All
656 variables listed contributed significantly to the discrimination.
657 [3] Standardized coefficients indicate the direction (positive or negative) and the degree to which
658 each variable contributes to the classification.
659 [4] Pooled within-class correlations between each selected variable and the canonical function;
660 variables with a correlation of 0.2 or more were considered important.


A
E. coli O157:H7 occurrence in 10 irrigation ponds
Month
P1
P2
P3
P4
P5
P6
P7
P8
P9
P10
B
Rainfall: mm
Temperature: °C
Rf
Tm
Month
C
Dissolved oxygen concentration : mg/L
Total Nitrogen : mg/L
DO
TN
Month
D
Oxidation reduction potential: mV
Fecal coliform: log10(CFU/100mL)
ORP
FC
Month



72x49mm (300 x 300 DPI)



72x48mm (300 x 300 DPI)

Table S1 Measured variables (means and standard deviations of ten irrigation ponds over a period of 12 months) for water samples that were positive or negative for *E. coli* O157:H7.

	Measured variables		*E. coli* O157:H7 positive samples		*E. coli* O157:H7 negative samples	
			mean	standard deviation *	mean	standard deviation *
Bacterial data	*E. coli*	MPN/100 mL	13.97	28.01	17.77	35.46
	Fecal coliforms	CFU/100mL	33.05	81.03	91.95	175.40
	Copiotrophic bacteria on C-rich agar	CFU/100mL	2572.20	4327.40	3931.00	4719.60
	Culturable bacteria on C-poor agar	CFU/100mL	5759.60	8975.90	11989.00	17713.00
	Shannon-Wiener Diversity Index	-	2.90	0.24	2.85	0.33
	Simpson Diversity Index	-	0.07	0.02	0.07	0.03
	Species Richness (S)	-	22.07	5.17	21.38	5.88
	Species evennness	-	0.94	0.02	0.95	0.02
Chemical characteristics	Total soluble solids (TSS)	mg/L	9.00	9.04	7.85	8.70
	Nitrate concentration	mg/L	0.16	0.26	0.07	0.17
	Ortho-phosphate concentration	mg/L	0.04	0.09	0.07	0.12
	Dissolved oxygen (DO)	%	104.90	55.70	104.80	32.80
	DO Concentration	mg/L	9.20	2.32	8.36	2.97
	DO Charge	mV	59.96	13.30	63.27	5.94
	Dissolved organic carbon (DOC)	mg/L	9.21	7.96	9.45	5.54
	Dissolved organic Nitrogen (DON)	mg/L	0.67	0.59	0.37	0.18
	pH	-	8.02	0.76	7.84	0.74
	Oxidative-reductive potential (ORP)	mV	43.45	91.10	-9.35	85.40
Physical characteristics	Temperature (water) at sampling	^{o}C	21.79	6.28	26.75	6.00
	Average air temperature (1 month)	^{o}C	18.98	5.25	24.41	4.52
	Lag (one week before sampling) of average temperature (L Avg Tm)	^{o}C	18.67	6.02	23.82	5.71
	Lag (two weeks) of average temperature (L2 avg Tm)	^{o}C	18.65	6.16	25.05	4.77
	Total rainfall (1 month)	mm	61.20	37.00	85.87	52.10
	Lag (one week) of total rainfall (L tot rain)	mm	21.59	23.18	26.06	25.48
	Lag (two weeks) of total rainfall (L2 tot rain)	mm	16.14	20.22	27.04	25.18
	Specific conductivity	mS/cm	0.21	0.24	0.19	0.23
	Turbidity	NTU	14.64	15.21	13.98	16.28

* The standard deviations reflect the variation over time rather than the variation due to sampling error.

Table S2 Mean values and standard deviations of highly correlated environmental factors from March, 2011 to February, 2012 in 10 irrigation ponds

	DO concentration (mg/L)		Total nitrogen concentration (mg/L)		Oxidative-reductive potential (mV)		Fecal coliform bacteria (Log_{10}(CFU/100 mL))	
month	Mean	Stand. dev.	Mean	Stand. dev.	Mean	Stand. dev.	Mean	Stand. dev.
Mar-11	9.08	1.46	1.19	1.03	197.41	75.37	0.72	0.82
Apr-11	8.28	0.72	0.74	0.45	129.16	14.29	0.49	0.68
May-11	9.28	1.45	0.46	0.18	152.94	31.11	0.72	0.95
Jun-11	7.17	1.40	0.55	0.41	24.48	114.13	2.05	2.19
Jul-11	8.09	1.99	0.40	0.16	-69.73	8.12	2.29	2.38
Aug-11	7.17	2.35	0.34	0.17	-68.16	20.12	1.89	2.16
Sep-11	8.05	3.52	0.54	0.56	-41.11	11.77	1.75	2.05
Oct-11	9.09	3.45	0.48	0.45	-12.59	14.10	1.69	1.87
Nov-11	10.14	10.14	0.44	0.44	18.39	18.39	0.93	0.93
Dec-11	11.28	1.91	0.32	0.32	11.37	20.89	1.35	1.32
Jan-12	10.54	1.14	0.70	0.54	18.36	17.42	1.09	1.43
Feb-12	9.77	2.15	0.99	0.57	7.56	18.68	1.34	1.62

Table S3. Groups of environmental factors with high positive correlations to each other (r > 0.7).

Group	Environmental factors			
1	**TSS***	Turbidity		
2	**DO Conc**	DO%		
3	**L2 avg Tm**	Temp	Average Tm	L avg Tm
4	**TN**	Nitrate		
5	**Simpson index**	Shannon index	Species richness	
6	**Culturable**	Copio		

* Factors in the first column have the highest correlations to *E. coli* O157 occurrence in each factor group, and are shown in bold.

Table S4. Operational taxonomic units (OTUs) of DGGE data with significant correlations to the occurrence of *E. coli* O157 isolated from irrigation ponds (P<0.1).

Relative band position (%)	Band No.	Biserial correlation coefficient to *E. coli* O157 occurrence	P-value
4.5	**b5**	0.183	0.005
46.2	**b36**	-0.182	0.005
57.1	**b43**	0.169	0.009
22.5	**b18**	-0.162	0.012
21.5	**b17**	-0.149	0.020
41.4	**b33**	0.145	0.025
10.7	**b10**	0.144	0.025
43	**b34**	0.142	0.028
64.7	**b48**	-0.137	0.035
60	**b45**	0.131	0.043
17.1	**b14**	-0.131	0.043
49.1	**b38**	0.127	0.050
0.2	b1	-0.122	0.059
1	b2	-0.121	0.062
79.8	b56	0.117	0.072
61.7	b46	0.112	0.083
50.5	b39	-0.109	0.089
37	b29	-0.108	0.094

DGGE OTUs (bands) with significant correlations (P < 0.05) to *E. coli* O157 occurrence are shown in bold font.