Journal of Applied Microbiology ISSN 1364-5072

ORIGINAL ARTICLE

The effectiveness of treating irrigation water using ultraviolet radiation or sulphuric acid fertilizer for reducing generic *Escherichia coli* on fresh produce—a controlled intervention trial

W. Beauvais[1,2,*], A.K. Englishbey[3,4,*], C.M. Marconi[5,6], U. Cholula[7,8], A.M. Belias[9], M. Wemette[1], J. Usaga[9,10], J.J. Churey[9], R.W. Worobo[9], J. Enciso[7], J.R. Anciso[11], K. Nightingale[3] and R. Ivanek[1]

1 Department of Population Medicine and Diagnostic Sciences, College of Veterinary Medicine, Cornell University, Ithaca, NY, USA
2 Comparative Pathobiology Department, College of Veterinary Medicine, Purdue University, West Lafayette, IN, USA
3 Animal and Food Sciences, Texas Tech University, Lubbock, TX, USA
4 Hygiena LLC, New Castle, DE, USA
5 Texas A&M AgriLife-Weslaco, Weslaco, TX, USA
6 College of Agriculture and Life Sciences - IR-4 Project, North Carolina State University, Raleigh, NC, USA
7 Biological and Agricultural Engineering, College of Agriculture and Life Sciences, Texas A&M University, College Station, TX, USA
8 Department of Agriculture, Veterinary and Rangeland Sciences, University of Nevada, Reno, NV, USA
9 Department of Food Science, College of Agriculture and Life Sciences, Cornell University, Ithaca, NY, USA
10 National Center for Food Science and Technology, University of Costa Rica, Ciudad Universitaria Rodrigo Facio, San Jose, Costa Rica
11 Horticultural Sciences, Texas A&M AgriLife Extension Service, Weslaco, TX, USA

Keywords
agriculture, food safety, microbial contamination, ultraviolet applications, water.

Correspondence
Wendy Beauvais, Comparative Pathobiology Department, College of Veterinary Medicine, Purdue University, VPTH Building 125, 725, Harrison St, West Lafayette, IN, 47906, USA.
E-mail: wbeauvai@purdue.edu

*Wendy Beauvais and April K. Englishbey contributed equally to this work. Author order reflects the lead contribution to final statistical analyses and manuscript, and the corresponding authorship.

2020/2257: received 19 October 2020, revised 6 January 2021 and accepted 18 January 2021

doi:10.1111/jam.15011

Abstract

Aims: The aims of this study were to: (i) estimate the effectiveness of ultraviolet radiation (UV) and sulphuric acid-based fertilizer (SA), at reducing levels of generic *Escherichia coli* in surface irrigation water and on produce and surface soil in open produce fields; and (ii) describe the population dynamics of generic *E. coli* in produce fields.
Methods and Results: Spinach and cantaloupe plots were randomly assigned to control, UV or SA treatment groups. Irrigation water was inoculated with Rifampicin-resistant *E. coli* prior to treatment. More than 75% of UV- and SA-treated tank water samples had counts below the detection limit, compared to a mean count of 3·3 Log_{10} CFU per ml before treatment. Levels of Rifampicin-resistant *E. coli* in soil and produce both increased and decreased over 10–15 days after irrigation, depending on the plot and time-period.
Conclusions: UV and SA treatments effectively reduce the levels of *E. coli* in surface irrigation water. Their effectiveness at reducing contamination on produce was dependent on environmental conditions. Applying wait-times after irrigation and prior to harvest is not a reliable means of mitigating against contaminated produce.
Significance and Impact of the Study: The results are of timely importance for the agricultural industry as new FSMA guidelines require producers to demonstrate a low microbial load in irrigation water or allow producers to apply a wait-time to mitigate the risk of contaminated produce.

Introduction

Outbreaks of foodborne illness are frequently attributed to bacterial contamination of fresh produce, and irrigation water has been identified as an important vehicle for produce contamination (Lynch *et al.* 2009). The use of irrigation water with high counts of generic *Escherichia coli* (geometric mean ≥126 colony forming units

(CFU) 100 ml^{-1} and statistical threshold value ≥410 CFU per 100 ml) is restricted by the Food and Drug Administration Food Safety Modernization Act (FSMA) introduced in 2011 (FDA 2013; FDA 2014). Approximately 58% of irrigation water in the USA originates from surface water (Kenny *et al*. 2009) which is particularly vulnerable to faecal contamination from wildlife, livestock and other sources. Therefore, there is an urgent need for safe and effective treatments of surface irrigation water to reduce human exposure to faecal pathogens via consumption of fresh produce.

Existing types of treatment strategies to reduce microbial loads in surface waters include chlorine-based treatments (e.g. calcium or sodium hypochlorite and chlorine dioxide), filtration and ultraviolet (UV) radiation (Jones *et al*. 2014; Allende and Monaghan 2015; Chang 2015; López-Gálvez *et al*. 2017; FAO and WHO 2019). However, chlorine-based treatments leave disinfectant by-products on the produce, are inactivated by organic matter, and can negatively impact soil and produce quality (Allende and Monaghan 2015; Chang 2015; López-Gálvez *et al*. 2017). While filtration does not leave behind by-products, filters can become clogged by the particulates present in the majority of surface water sources (Allende and Monaghan 2015). Peracetic acid is less corrosive than chlorine-based disinfectants and is not inactivated by organic matter, however it is expensive (Chang 2015). Generally, acids may exert an antimicrobial effect via inhibition of enzymes, membrane function, nutrient transport or metabolic activity (Chang and Fang 2007). Sulphuric acid (H_2SO_4) has a wide range of applications in agriculture and has been used to eliminate *E. coli* O157:H7 from alfalfa seeds prior to sprouting (Pandrangi *et al*. 2006). Sulphuric acid washes or sprays are often applied to the surface of meat or poultry products to reduce the level of micro-organisms and to prevent microbial growth (FDA 2011). Sulphuric acid is regulated as a pesticide under the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) (USEPA 1993). It is exempt from the requirement of a tolerance for residues when used in accordance with Good Agricultural Practices (GAPs) as a pH control agent in pesticide formulations applied to growing crops or to raw agricultural commodities after harvest (HSDB 2010). Sulphuric acid-based fertilizers with 27–55% sulphuric acid have been commonly used to remove biofilm buildup in irrigation drip pipes and as water amendments in agriculture to reduce pH of calcareous soils and high bicarbonate waters (Gregory 2001; Enciso-Medina *et al*. 2011). To the best of our knowledge, there are no published data on the effectiveness of sulphuric acid-based fertilizers to reduce the microbial load of enteropathogens in surface irrigation water.

Ultraviolet radiation has been shown to be effective at reducing the microbial load in unfiltered surface water under laboratory conditions. Samples of surface water were inoculated with *E. coli* O157:H7, *Salmonella* and *Listeria monocytogenes* and the inactivation of pathogens observed following UV radiation was ≥99·9% (up to 7 log_{10} CFU per ml inactivation in highly contaminated pond water) (Jones *et al*. 2014). Banach *et al*. (2021) reported that UV radiation was effective at reducing the load of *E. coli* in irrigation water, however, when the authors field-tested UV treatment of water that was used to irrigate lettuce and endive in a semi-open high tunnel, *E. coli* counts on produce were both low and comparable in control and UV treatment groups. UV radiation has not to our knowledge been tested for treatment of surface irrigation water in a controlled trial using open produce fields. Also, there are concerns that highly turbid water does not allow UV radiation to penetrate well (Chase *et al*. 2019). Therefore, while there are several water treatment methods currently available, additional methods are needed that have been shown to be effective under open field conditions and that do not have a lasting environmental impact.

The proposed FSMA regulations allow producers to mitigate the risk of contaminated produce by applying a wait-time of up to 4 days between irrigation and harvest, assuming a 0·5 Log_{10} reduction per day, if irrigation water is found to be contaminated (defined as >126 CFU per ml or 90th percentile >410 CFU per ml generic *E. coli*) (FDA 2013; FDA 2014). Therefore, there is an urgent need to validate that this assumption mitigates risk.

There are relatively few data on the population dynamics of *E. coli* and other micro-organisms, including pathogens, on produce fields, under field conditions. A recent literature review (Snellman *et al*. 2014) reported that mean decay rates of *E. coli* O157:H7 in field trial studies ranged between 0·4 and 1·64 Log_{10} CFU per day in the initial days following contamination and subsequent decreases were observed (Wood *et al*. 2010; Fonseca *et al*. 2011; Moyne *et al*. 2011; Bezanson *et al*. 2012). Several other studies describe the decline of *E. coli* on lettuce grown in greenhouses or growth chambers over the course of up to 5 days post-inoculation (Tyagi 2014; Linden *et al*. 2016; Jang and Matthews 2018). However, one study found that *E. coli* grew on field-grown romaine lettuce following irrigation (Chase *et al*. 2019). Another study found nonpathogenic *E. coli* on spinach grown in a growth chamber 28 days post-inoculation, and *E. coli* O157:H7 was detected in the soil in which spinach plants were grown at 28 days post-inoculation, although it was not detected on the spinach beyond 7 days (Patel *et al*. 2010). A laboratory experiment showed the presence of

viable *E. coli* cells within spinach plants 20 days after the plants were inoculated with *E. coli* via syringe infiltration (Wright *et al.* 2017). Under storage conditions, *E. coli* has been shown to grow on shredded iceberg lettuce and cantaloupe rinds at temperatures of 22 and 25°C respectively (del Rosario and Beuchat 1995; Chang and Fang 2007).

The objectives of this study were to: (i) estimate the effectiveness of two treatments, UV radiation and sulphuric acid-based fertilizer, at reducing the levels of generic *E. coli* in (a) surface irrigation water (both immediately after the treatment and in the field) and (b) soil surface and harvested produce (i.e. spinach and cantaloupe) from open produce fields irrigated with the treated water and (ii) describe the dynamics of generic *E. coli* in the soil and on the produce post irrigation with inoculated water.

Materials and methods

The three controlled field trials were conducted at Texas A&M AgriLife Research and Extension Center in Weslaco, TX including two spinach (*Spinacia oleracea* L.) trials: November 2016–February 2017 and November 2017–February 2018 and a cantaloupe (*Cucumis melo* var. *cantalupensis)* trial in April–June 2017. A preliminary trial on cantaloupe was conducted in summer 2016, however, these data are not presented because methods were changed after that trial.

In each of the three trials, two replicates of three plots were irrigated via a drip system and two replicates of three plots via a furrow system (3 plots × 2 irrigation types × 2 replicates × 3 trials = 36 plots in total). The three plots for each irrigation system and replicate were randomly allocated to three treatment groups: control (no treatment of inoculated irrigation water); UV treatment (inoculated irrigation water treated with UV radiation) and SA treatment (inoculated irrigation water treated with a sulphuric acid-based fertilizer). On each plot, either four or five irrigations with inoculated water were performed and produce were harvested over 10–15 days following the 4th and/or 5th irrigations (depending on sample type and produce type). We refer to the period of time from one irrigation with inoculated water to the next within the same growing season as one irrigation cycle. The plants were irrigated every 1–3 weeks as needed, as would be typically done by growers.

Description of experimental station

The study was conducted at Texas A&M AgriLife Research and Extension Center in Weslaco, TX (longitude 26″9'N, latitude 97″57'W). This region has a semi-arid climate and an average annual rainfall of 558 mm. The soil at the site was a sandy clay loam composition (fine-loamy, mixed, hyperthermic Typic Calciustolls, pH 8·1).

The study plots were prepared in 2·04 m (80 inch) wide raised beds for cantaloupes and 1·02 m beds (40 inch) for the spinach, with a bed length of 48·8 m (160 feet), with each plot consisting of two raised beds with a furrow canal in between. All plots (including furrow plots) had drip tape buried 5 cm deep in the centre of the bed.

Plants and growth conditions

Fertilization was applied before planting following the standard recommended rate of 100-75-75 NPK. Pre-emergent herbicide and maintenance chemicals were applied to control noxious weeds, pests and diseases. Seeds were mechanically planted with MaterMacc Precision Vacuum Planter.

For the cantaloupe trials a locally grown variety (Primo) from Syngenta Seeds, Inc. (Hopkins, MS)/Rogers Brand was planted at a 3·2-cm depth and a spacing of 35·6 cm. Two rows of plants were planted per bed.

A locally grown variety of spinach (Viceroy) from Champion Seed Company (McAllen, TX) was sown at a 1·3-cm depth with a separation between plants of 6 cm, in three rows per bed.

Bacterial strain preparation

A marked *E. coli* strain was prepared in order to distinguish between the *E. coli* originating from the tank inoculum and other *E. coli* contaminants in water, soil and on produce. A Rifampicin-resistant derivative of *E. coli* (ATCC® 25922™) (RifR *E. coli*) was grown overnight in Brain Heart Infusion (BHI) broth at 37°C for 24 ± 2 h to achieve an approximate concentration of 9 $\log_{10}$ CFU per ml.

Description of water and irrigation systems

The irrigation water on the experimental station comes from an open pond which is sourced via canals from the Rio Grande River. Irrigation water had an average electrical conductivity of 0·13 S m^{-1} and was filtered using sand media filters. Water was pumped from the pond into holding tanks, where it was inoculated with RifR *E. coli* at a concentration of approximately 3·3 Log_{10} CFU per ml. If used for a treatment plot (UV or SA), water was treated prior to being pumped into the plots via the furrow or drip system. The drip tubing has a nominal discharge of 0·75 l h^{-1} per emitter and each emitter is spaced every 30 cm (Streamline; Netafim, Riverside, CA).

Description of water treatments

The UV treatment was applied to irrigation water using an experimental UV processor (Headwater Foods/FPE, Rochester, NY). The radiation intensity of the six UV lamps in the UV processor was 48 mW cm$-^{2}$. Sensors embedded in the UV processor collected data for calculation of the delivered UV energy exposure during the UV treatment of water. Specifically, data were recorded for each batch (batch number, the start and end date and time of the batch, the reason treatment was terminated (e.g. by the operator, due to a flow fault, due to a power failure, etc.). Additionally, within each batch a log was created every 30 s that included a time stamp, power density in μW cm-^{2}, and water flow in gallons h^{-1}. For each time stamp, the logged sensor reading (P) and flow reading (F) were used to calculate the delivered energy exposure (E) in mJ cm-^{2}. Based on the volume of the tube (1·8 gallons), conversion of hour units to seconds and conversion from μJ cm-^{2} to mJ cm-^{2}, the conversion constant was $K = 1{\cdot}8 \times 3600/1000 = 6{\cdot}480$. Therefore, the delivered UV energy exposure (expressed in mJ cm-^{2}) for each time stamp was calculated as $E = 6{\cdot}480 \times P/F$.

The achieved median fluence level was 36 mJ cm-^{2} for drip plots and 11 mJ cm-^{2} for furrow plots, due to differing water flow rates in the different irrigation systems.

The SA-based fertilizer treatment involved the use of fertilizer monourea with 12% sulphur (Nphuric 9-0-0-12 acid). Monourea was applied by directly injecting it into surface water prior to field irrigation applications using a fertilizer injection pump (Dosatron^{TM}) at the injection rate of 1 : 500. The fertilizer concentration per unit of irrigation water applied, expressed as gallons of fertilizer per gallons of water, was 0·0005 (1 gallon of fertilizer per 2000 gallons of water) for the furrow and 0·002 (0·65 gallon of fertilizer per 325 gallons of water) for the drip irrigation system in all trials. Considering the size of experimental plots and the amount of irrigation water, the fertilizer rate applied per irrigation per unit area was 10·2 gallons acre^{-1} for the furrow and 6·6 gallons acre^{-1} for the drip irrigation system in the cantaloupe trial. As per industry standards, spinach beds were half the width of the cantaloupe beds (1·02 *vs* 2·04 m), and thus in spinach trials experimental plots received twice the amount of the fertilizer per unit area in both irrigation systems.

Sample collection

Tank water samples of 100 ml were collected in triplicate from each tank pre-treatment and post-treatment into individual sterile cups. The pre-treatment samples served also as a verification of the concentration of RifR *E. coli* in water tanks.

A 100 ml container of irrigation water from the field was collected towards the end of irrigation at the beginning, middle and end of each plot. For the drip irrigation system, existing valves were used for collection at the beginning and end of the plots, however, at the middle of the plot a thin irrigation tube was inserted into the drip tape in only one of the two beds per plot. For the furrow irrigation system, samples were collected at the beginning, middle and end of the furrow canal between the left and right bed of each plot.

Produce samples were collected on various days during the 4th irrigation for cantaloupes as well as the 4th and 5th irrigations for spinach, with sampling days being selected to balance needs of statistical analysis and logistical feasibility. The days for cantaloupe sample collection for furrow irrigation were day 0, 1, 2, 3, 6, 8, 10 and 13 after irrigation; for drip irrigation, samples were collected on days 0, 1, 2, 5, 7, 9 and 12 after irrigation. Spinach samples were collected for furrow irrigation on days 0, 1, 2, 3, 6, 8, 10 and 15 after irrigation and for drip irrigation on days 0, 1, 2, 5, 7, 9 and 14 after irrigation. In all trials, day 0 denotes samples collected on the day of irrigation, immediately prior to irrigation. Both furrow and drip irrigation started at 8:00 AM and finished at approximately 3:00 PM. Sampling on the subsequent days was conducted between 8:00 AM and 12:00 PM.

Cantaloupe vines often spread from the raised bed to the canal so that some cantaloupe fruits are harvested from the canal. One cantaloupe, at market grade maturity, was collected at the beginning, middle and end of the beds, alternating between cantaloupes located in the canal, the raised bed to the left of the canal and the raised bed to the right of the canal and placed in a sterile bag. Cantaloupes were placed on a stainless-steel work surface that was cleaned between samples with paper towels, then sprayed with 70% ethanol and flamed to sterilize. The same procedure was performed on the cutting knives. Also, to prevent cross-contamination, a sheet of aluminium foil was used as a cutting board and replaced with every new sample. After quartering and removing the cantaloupe flesh, each quartered rind was cut vertically and horizontally, yielding four pieces per quarter, which were then cut into approximately 1-inch wide strips. The strips were placed inside a labelled Whirl Pak bag, and the bag was rolled down to remove air.

Approximately 25 ± 5 g of spinach leaves were collected from the beginning, middle and end of beds. At each sampling location (i.e. beginning, middle and end of a plot) two leaves per plant were collected from two plants from the rows closest to the canal and from one plant from the middle rows in both left and right beds. Therefore, each sample consisted of 12 leaves (6 plants × 2 leaves) collected and directly placed inside a

labelled Whirl Pak bag, and rolled down to remove air. Harvest was completed after the 4th irrigation, plants were allowed to regrow, and the second cut was harvested after the 5th irrigation. To guarantee the spinach regrowth from the same location would be harvested after the 5th irrigation, four linear feet were measured in the beginning, middle and end of each plot; a marker (wood stake) with the sampling day was placed at the beginning and end of the two linear feet and the other two feet were left as buffer between sampling days. Two leaves per plant were collected for our samples as described above and then the marked two linear feet was harvested and disposed of the leaves, leaving the remaining two linear feet undisturbed.

Boot sock samples of surface soil were collected only during the 4th irrigation. The days of soil sample collection for furrow irrigation were days 0 (just before irrigation), and 1, 3, 6 and 10 after irrigation; and for drip irrigation: days 0, 1, 2, 5 and 9. All samples *P* were collected in a 12-inch by 12-inch area from the vicinity of the collected produce sample using a boot sock soaked in buffered peptone water (BPW; ThermoFisher Scientific, Waltham, MA) and placed into a sterile plastic container. The produce-soil pairs were recorded for analysis.

Sample storage and transport

After processing, samples were immediately placed in a cooler container with icepacks, and shipped to the Texas Tech University International Center for Food Industry Excellence for microbiological analyses within 48 h of sample collection. To verify that the cold chain was maintained, upon arrival to the laboratory, sample containers were verified to have frozen icepacks. An infrared thermometer was used to spot check throughout the sample container (bottom, middle and top) to ensure temperature was representative of a refrigerated sample.

Laboratory tests

At each irrigation cycle, six samples (800 ml each) of irrigation water were collected before inoculation and delivered to Cornell University within 24 h from collection. The water was stirred thoroughly, and then pH (HI 2211 pH/ORP meter; Hanna, Woonsocket, RI) and turbidity (2100P portable turbidimeter; Hach, Loveland, CO) measurements were recorded. All turbidity values were recorded in nephelometric turbidity units (NTU).

Water samples (100 ml) were transferred into a sterile Whirl-Pak bag and combined with 400 ml BPW to achieve a 1 : 5 dilution. The solution was homogenized by hand shaking for 30 s. Spinach samples were weighed and adjusted by aseptically removing portions of spinach leaves, to confirm a final sample weight of 25 ± 2 g. A 1 : 10 dilution was performed in the same Whirl-Pak bag by combining the 25 g of spinach with 225 ml of BPW. The solution was homogenized by hand massaging for 1 min. Four hundred millilitre of BPW was added to each cantaloupe rind bag. Rinds were homogenized by hand massaging for 2 min. Boot sock samples (irrigated soil surface samples) were aseptically removed from the container and placed in a sterile Whirl-Pak bag followed by addition of 90 ml of BPW and homogenized in a stomacher for 30 s at 200 rev min^{-1}.

Following sample processing and homogenization, a 1 ml aliquot was serially diluted, and 100 µl of the undiluted sample (water) or appropriate dilutions were directly spiral plated in duplicate onto MacConkey agar (ThermoFisher Scientific) supplemented with 100 µg ml^{-1} of Rifampicin. All plates were incubated at 37°C for 24 ± 2 h, before enumerating Rif-R generic *E. coli*. Minimum detection limits were 10 CFU per ml of water sampled, 900 CFU per bootsock of soil sampled (i.e. 10 CFU per ml of sample diluent), 100 CFU per gram of spinach sampled (i.e. 100 CFU per ml of sample diluent) and 4000 CFU per rind of cantaloupe sampled (i.e. 10 CFU per ml of cantaloupe rind diluent).

The original Whirl-Pak homogenates from all samples were incubated at 35°C for 24 ± 2 h followed by screening for *Salmonella* using the BAX®PCR system (Hygiena, Camarillo, CA) and BAX® System Standard assay for *Salmonella* (Hygiena). If the BAX PCR system indicated a potential positive sample, a modified version of the FDA Bacteriological Analytical Manual procedure was performed as described briefly below. A 1-ml aliquot of the potential positive sample was added to a tube containing 9 ml Rappaport Vassiliadis (RV; ThermoFisher Scientific) broth and a tube containing 9 ml Tetrathionate (TT; ThermoFisher Scientific) broth. Selective RV and TT enrichments were incubated at 42 and 35°C respectively for 24 ± 2 h. Following incubation, a 100-µl aliquot of RV and TT enrichments were plated onto Chromagar *Salmonella* plates (Becton, Dickson and Company—BD, Franklin Lakes, NJ) and incubated at 37°C for 24 ± 2 h. Up to six, well-isolated, typical presumptive *Salmonella* colonies (round and mauve in colour), collectively from RV and TT Chromagar *Salmonella* plates, were confirmed using a molecular method (PCR reaction that detects the *Salmonella* specific portion of the *invA* gene) previously described and validated by Nucera *et al.* (2006).

Sequencing

Genomic DNA was extracted from overnight cultures grown in brain-heart infusion broth (BHI) (Becton Dickinson), using the GenElute Bacterial Genomic DNA kit

(Sigma, St Louis, MO). Genomic DNA extractions were quantified using a fluorometer (Qubit 2.0; Thermo Fisher Scientific, Waltham, MA), and the concentration was adjusted at 0·25 ng μl^{-1}. Libraries were prepared using 1 ng of gDNA using the Nextera XT kit with Nextera indexes set A (Illumina, San Diego, CA) following the manufacturer's recommendations. The quality of library preparations was evaluated on a 2200 TapeStation instrument (Agilent Technologies, Santa Clara, CA) and the individual library concentrations were adjusted to 8 nmol l^{-1}. Finally, all strains in a single run were pooled together in equal fractions and 4 pg of the pooled libraries were sequenced using the sequencing by synthesis technology using a 300 bp-read cartridge on a MiSeq sequencer (Illumina). Whole genome sequence raw data were assembled using the SPAdes pipeline, and annotated using both RAST and Prokka (Aziz *et al*. 2008; Bankevich *et al*. 2012; Seemann 2014). Bioinformatic analyses were performed using pipelines from the Center for Genomic Epidemiology website (http://www.genomicepidemiology.org) and various in-house scripts and protocols. Assembly files were uploaded and *Salmonella* serotypes were assigned using the SeqSero pipeline.

Additional data collected

For each date on which samples were taken, hourly and daily data on the minimum, maximum and mean temperature; minimum, maximum and mean relative humidity; total solar radiation; mean wind speed; maximum wind gust; mean soil temperature (1 inch below soil surface); total precipitation; mean leaf wetness; mean quantum radiation; mean UV radiation; and total evapotranspiration were gathered from the following weather stations: Center weather station in Weslaco, Texas (Lat: 26°9′16·83″N, Long: 97°57′39·50″W) and Annex weather station in Mercedes, Texas (Lat: 26°9′44·47″N, Long: 97°56′26·90″W). These weather stations are run by Texas A&M AgriLife Extension and data are available online at: http://southtexasweather.tamu.edu/

The raw weather data were screened for missing data and implausible values. Any missing or implausible values from the Weslaco, Texas station were replaced with the values from the Mercedes, Texas station. The hourly data were converted into daily data. In addition to information about weather variables on a particular day, additional weather data were collected for each weather variable for 1, 2, 3, 4 and 5 days prior to each sampling date.

Statistical analysis

All statistical analyses were performed in R (R Core Team 2017). The following R packages were used: 'lme4' (logistic and linear regression analyses) (Bates *et al*. 2015) 'rpart' (regression trees) (Therneau and Atkinson 2018), 'reshape' (to manipulate data tables) (Wickham 2007) and 'forestplot' (to produce forest plots) (Gordon and Lumley 2017).

All laboratory results, sample descriptors and additional data were compiled into a single spreadsheet and all variables were summarized separately by sample type using plots and summary statistics.

Samples with no RifR *E. coli* detected were assigned a value of zero while samples with any RifR *E. coli* counted were assigned a value of one for the purposes of the binary outcome analyses (i.e. above or below the detection limit). Also, samples with no RifR *E. coli* detected were assigned a value of the minimum detection limit for a particular sample type for the purposes of the continuous-outcome analyses (i.e. Log_{10} CFU per sample and Log_{10} CFU per ml).

Data were initially analysed by sample type (tank water, field water, soil surface and produce, i.e. spinach and cantaloupe), combining data from all three trials. Thus, the results were respectively expressed per 100 ml of tank water, per 100 ml of field water, per 1 bootsock soil surface, per 25 g of spinach as well as per 1 cantaloupe rind (because only the rinds of the fruits underwent testing (not the pulp)). The effects of treatments and time elapsed since irrigation were the primary predictor variables of interest. The outcomes of interest were: (i) the proportion of samples that had positive counts of RifR *E. coli* (i.e. counts above the detection limit) and (ii) the concentration of RifR *E. coli* (in log_{10} CFU per sample and log_{10} CFU per ml). Associations between these predictor variables and outcomes of interest were investigated using plots; Wilcoxon rank sum test was used to compare log_{10} CFU per ml of RifR *E. coli* among treatment groups; logistic regression was used to screen univariate associations between predictor variables and the detection of RifR *E. coli* (binary variable: above or below the detection limit). In addition, possible confounders and effect modifiers were identified from the dataset and the relationship between each of them and the outcomes of interest were also investigated. For screening of weather variables, random effects logistic regression was used to account for autocorrelation of weather variables by date, by including sampling date as a random effect in each univariate model.

Due to unbalanced number of samples from the cantaloupe and two spinach trials, and based on preliminary analyses, it was decided to analyse soil and produce samples from spinach and cantaloupe trials separately (i.e. four separate analyses). Tank and field water samples were each analysed by combining data from all three trials. Multivariable analysis was conducted separately for

each sample type (tank water, field water, soil and produce). Random effects logistic regression was used to assess the treatment effects and the effects of time since irrigation simultaneously. A backwards stepwise procedure was followed. Any predictor or potentially confounding/effect modifying variable with a univariable *P* value <0·20 was included in the initial model as a fixed effect. Variables were dropped from the model according to p-value, and retained if they had a multivariable *P* value of <0·05. Treatment group, days since irrigation, irrigation type and year were forced into the model as random effects if they were not already included as fixed effects. Interaction terms between each co-variate and each predictor variable of interest were added one at a time, and retained if the interaction term had a *P*-value of <0·05. The resulting model (model A, developed separately for each of the two sample types (produce and soil) and the two produce types (spinach and cantaloupe)) was used to report the effects of treatment and time since irrigation.

To investigate if weather could explain the differences in trend in prevalence over time (soil and produce samples), weather variables were then added into model A as follows. Univariate analyses of the associations between weather variables and detection of RifR *E. coli* were conducted separately by sample type (soil and produce) and produce type (spinach and cantaloupe), by conducting logistic regression with detection of RifR *E. coli* as the outcome variable and each weather variable as the predictor variable (as a continuous variable). Weather variables with a univariate *P* value of <0·2 were subjected to Principle Component Analysis (PCA) as follows. The data were standardized by subtracting the mean and dividing by the standard deviation. The number of meaningful components to retain was determined based on two criteria. According to criteria 1, in decreasing order of variance accounted for, only the first components accounting for up to 90% of the total variance were retained. According to criteria 2, the rotated pattern had to demonstrate 'simple structure'. Here, loading is a correlation coefficient between a variable and its principle component, while 'simple structure' means that most variables have relatively high factor loadings on only one component and near zero loadings on the other components.

The results of PCA were used as follows. All weather variables in the final principal components were included in multivariable logistic regression models. The weather variables were first added into each model A individually. All variables with a *P*-value <0·05 when added into model A individually were then added into model A together and sequentially dropped from the model if the multivariable *P* value was >0·05, in decreasing order of *P* value to select the final model (Model B).

Classification and regression trees

Classification and regression tree (CART) methodology was used to further explore potential interactions between explanatory variables and to aid interpretation of findings. A single regression tree was determined for each sample type (tank water after treatment, water from the field, soil, spinach and cantaloupe) using the $\log_{10}$ counts of RifR *E. coli* per ml of diluent as the outcome variable. Zeros were imputed for counts below the detection limit. The optimal regression tree was selected by minimizing the cross-validated error from 10-fold cross-validation. In these trees, the splitting criterion is based on ANOVA and the summary statistic describing a node is the mean of the node. All explanatory variables significant at the *P*-value of 5% at the univariable level were used in the CART analysis. Unlike the random effects generalized linear models, the CART methodology cannot account for clustering in data; however, subsampling as part of the performed 10-fold cross-validation is expected to randomly break the clusters and alleviate the effect of autocorrelation on the findings.

Data availability

All outcome and descriptor variable data for all samples are available in a .csv file in Supporting Information (Table S1) as well as at: https://github.com/IvanekLab/surface-water-treatments.

Results

Samples collected

Over the course of the three trials, 502 tank water samples were collected (three pre-treatment samples and three post-treatment samples from 84 tank-loads in total—two missing samples due to human error). In total, 503 samples of irrigation water from the field were collected during irrigations (one missing sample due to human error). In total, 360 soil samples (after the 4th irrigation in all three trials), 234 cantaloupe samples (after the 4th irrigation in the cantaloupe trial) and 1079 spinach samples (after 4th and 5th irrigation in both spinach trials) were taken (Table 1).

Descriptive analysis

Laboratory tests

Turbidity of the tank water collected before inoculation with RifR *E. coli* ranged from 11·9 to 58·9 nephelometric turbidity units (NTU) and pH ranged from 6·35 to 8·19.

Of the 2768 samples collected in total, 26 were confirmed positive (1%) for *Salmonella* (Table 2) and 705 (26%) samples were positive for RifR *E. coli* (Table 3).

Table 1 Total number of samples collected

	Spinach				Cantaloupe		
	2016–7		2017–8		2017		
	Furrow*	Drip*	Furrow*	Drip*	Furrow*	Drip*	Total
Sample type							
Tank water (BT)[†]	45	45	44[d]	45	36	36	251
Tank water (AT)	44[d]	45	45	45	36	36	251
Field water[§]	90	90	90	90	72	71[‡]	503
Soil	60[¶]	60**	60[¶]	60**	60[¶]	60**	360
Produce	288[††]	251[‡‡]	288[††]	252[‡‡]	126[§]	108[¶¶]	1313
Total	527	491	527	492	330	311	2768

BT, before treatment; AT, after treatment.

*Field samples (field water, soil and produce) were collected from two replicated plots for each of three treatment groups (Control, SA treatment and UV treatment). Tank water samples were collected from separate tanks that were used for each treatment group and irrigation type, except the same tank was used for control and SA treatment groups on drip plots.

[†]Tank water samples all collected on the day of irrigation.

[‡]1 missing sample due to mis-collection in field and/or not being shipped on time for analyses.

[§]All collected immediately after irrigation.

[¶]Samples collected on day 0 (just before irrigation), and on days 1, 3, 6 and 10 after irrigation.

**Samples collected on day 0 (just before irrigation), and on days 1, 2, 5 and 9 after irrigation.

[††]Samples collected on day 0 (just before irrigation), and on days 1, 2, 3, 6, 8, 10 and 15 after irrigation.

[‡‡]Samples collected on day 0 (just before irrigation), and on days 1, 2, 5, 7, 9 and 14 after irrigation.

[§§]Samples collected on day 0 (just before irrigation), and on days 1, 2, 3, 6, 8 and 10 after irrigation.

[¶¶]Samples collected on day 0 (just before irrigation), and on days 1, 2, 5, 7 and 9 after irrigation.

During the spinach trials (conducted during winter months), mean daily temperature ranged from 11 to 26°C, mean daily relative humidity ranged from 42 to >99% and mean daily wind speed from 0·7 to 6·1 m s^{-1}. During the cantaloupe trial (conducted in summer), mean daily temperature ranged from 24 to 32°C, mean daily relative humidity from 52 to 84% and mean daily wind speed from 1·8 to 3·7 m s^{-1}.

Univariate and multivariable analysis

Salmonella-*positive samples*

Due to the low number of *Salmonella*-positive samples, they were not subjected to statistical analyses of associations but are presented by trial and sample type in Table 2. There were no positive samples among samples taken during the cantaloupe trial. In addition, none of the soil samples or any produce samples taken prior to irrigation were positive. The following serotypes were identified: *S*. Typhimurium, *S*. Montevideo, *S*. Javiana, *S*. Branderup, *S*. Baildon, *S*. II 912 and *S*. Saintpaul.

UV treatment effects

The counts of RifR *E. coli* in UV and control treatment group samples before treatment were very similar; in the UV treatment group, the median count was 3·3 log10 CFU per ml (interquartile range 3·3–3·4 log10 CFU per ml) and in the control group the median count was also 3·3 log10 CFU per ml (interquartile range 3·3–3·4 Log_{10} CFU per ml) (Table 3). UV treatment reduced the percentage of tank water samples that were positive for RifR *E. coli* (above the detection limit of 10 CFU per ml), from 100 to 15%. If we conservatively assume that samples with no counts detected had a count at the minimum detection limit (1 Log_{10} CFU per ml), there was a difference of 2·3 Log_{10} CFU per ml between the median counts in UV-treated samples and control samples. Conversely, if we assume all samples with no counts detected had zero CFU, there is an approximate difference between the medians of 3·3 Log_{10} CFU per ml. It is possible that the treatment could achieve an even higher reduction with higher contamination levels, because the inoculated water samples did not have counts exceeding 3·3 Log_{10} CFU per ml.

The reduction in the number of positive samples was also apparent in water samples collected from the field during irrigation; only 11% of field water samples on plots irrigated with UV-treated water were positive for RifR *E. coli* (above the detection limit of 10 CFU per ml) compared with 99% of samples on control plots. The effect of UV treatment on the proportion of soil and produce samples varied by produce type and irrigation cycle (Table 3; Fig. 1).

Table 2 Number (and % of total tested) [serotypes isolated*] of samples confirmed positive for *Salmonella*

	Spinach Trial 2016–7			Spinach Trial 2017–8		
	Control treatment group	SA treatment group	UV treatment group	Control treatment group	SA treatment group	UV treatment group
Tank water before treatment	1 (30) [Br]	1 (30) [S]	1 (30) [Br]	1 (30) [T]	0 (29)	3 (30) [S,Ba,T]
Tank water after treatment	0 (30)	0 (29)	0 (30)	2 (30) [S,T]	0 (30)	0 (30)
Field water	6 (60) [Br, S]	1 (60) [Br]	0 (60)	2 (60) [S,T]	5 (60) [S,Ba,T]	0 (60)
Spinach	1 (156) [M]	2 (156) [I9,Br,J]	0 (156)	0 (156)	0 (156)	0 (156)

**Salmonella* serotypes isolated: Br = *S*. Braenderup; S = *S*. Saintpaul; Ba = *S*. Baildon; T = *S*. Typhimurium; M = *S*. Montevideo; I9 = *S*. II 912; J = *S*. Javiana.

Table 3 Effects of UV and SA treatments on proportion of samples positive to RifR *Escherichia coli* and CFU per sample

Trial and irrigation (irr) numbers	Sample type	Treatment group	% positive (total number of samples)	*P* value (binary outcome)	Minimum, 25th, 50th, 75th percentiles, maximum $\log_{10}$ CFU per sample*,†	*P* value (continuous outcome)‡
All trials and irr numbers	Tank water after treatment	Control	100% (84)	Baseline	3·0, 5·3, 5·3, 5·4, 5·4	Baseline
		SA	24% (83)	$<1 \times 10^{-4}$§	3·0, 3·0, 3·0, 3·0, 5·4	$<1 \times 10^{-4}$
		UV	15% (84)	$<1 \times 10^{-4}$§	3·0, 3·0, 3·0, 3·0, 5·4	$<1 \times 10^{-4}$
All trials and irr numbers	Field water	Control	99% (168)	Baseline	3·0, 5·3, 5·3, 5·4, 5·4	Baseline
		SA	30% (168)	$<1 \times 10^{-4}$§	3·0, 3·0, 3·0, 4·3, 5·4	$<1 \times 10^{-4}$
		UV	11% (167)	$<1 \times 10^{-4}$§	3·0, 3·0, 3·0, 3·0, 5·4	$<1 \times 10^{-4}$
Spinach Years 1 and 2: irr 4	Soil	Control	13% (64)	Baseline	3·0, 3·0, 3·0, 3·0, 5·0	Baseline
		SA	13% (64)	1·0¶	3·0, 3·0, 3·0, 3·0, 4·1	0·34
		UV	16% (64)	0·02¶	3·0, 3·0, 3·0, 3·0, 4·2	0·74
Spinach Years 1 and 2: irr 4	Spinach	Control	1% (156)	Baseline	3·4, 3·4, 3·4, 3·4, 3·7	Baseline
		SA	4% (156)	0·17¶	3·4, 3·4, 3·4, 3·4, 5·5	0·15
		UV	1% (156)	1·0¶	3·4, 3·4, 3·4, 3·4, 3·7	1·0
Spinach Years 1 and 2: irr 5	Spinach	Control	4% (156)	Baseline	3·4, 3·4, 3·4, 3·4, 5·8	Baseline
		SA	1% (156)	0·04¶	3·4, 3·4, 3·4, 3·4, 5·7	0·03
		UV	1% (156)	0·07¶	3·4, 3·4, 3·4, 3·4, 5·8	0·09
Cantaloupe Trial: irr 4	Soil	Control	22% (32)	Baseline	3·0, 3·0, 3·0, 3·0, 4·8	Baseline
		SA	16% (32)	0·48**	3·0, 3·0, 3·0, 3·0, 5·3	0·58
		UV	38% (32)	0·13**	3·0, 3·0, 3·0, 4·1, 5·1	0·24
Cantaloupe Trial: irr 4	Cantaloupe	Control	26% (66)	Baseline	3·6, 3·6, 3·6, 3·6, 6·0	Baseline
		SA	17% (66)	0·18**	3·6, 3·6, 3·6, 3·6, 5·5	0·23
		UV	11% (66)	0·02**	3·6, 3·6, 3·6, 3·6, 5·3	0·07

*Samples with no RifR detected were assigned a value equal to the minimum detection limit).

†Water samples: 100 ml; soil samples: 1 bootsock; spinach samples: 25 g; cantaloupe samples: 1 rind.

‡*P* value from Wilcoxon rank sum test.

§Wald test *P* value from mixed effect logistic regression model with trial*irrigation type*tank as a random effect and treatment group as a fixed effect.

¶Wald test *P* value from mixed effect logistic regression model with irrigation type*year as a random effect and treatment group and day as fixed effects.

**Wald test *P* value from mixed effect logistic regression model with irrigation type as a random effect and treatment group and day as fixed effects.

SA Treatment effects

The counts of RifR *E. coli* in SA and control treatment group samples before treatment were also very similar; in the SA treatment group, the median count was 3·3 Log_{10} CFU per ml (interquartile range 3·3–3·4 $\log_{10}$ CFU per ml) and in the control group the median count was also 3·3 $\log_{10}$ CFU per ml (interquartile range 3·3–3·4 Log_{10} CFU per ml) (Table 3). SA treatment also resulted in a

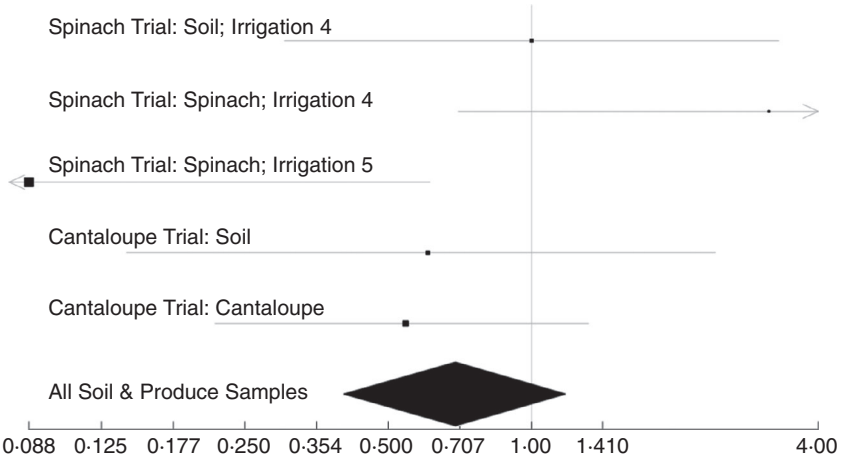

Figure 1 Forest plot to show the odds ratios and confidence intervals for the effects of sulphuric acid-based fertilizer treatment on the odds of samples being positive for Rifampicin-resistant (RifR) *Escherichia coli*, by produce type trial, sample type and irrigation number. Symbols show odds ratios and their sizes are inversely related to the standard error. Lines show confidence intervals and arrows indicate that they extend beyond displayed axis. Line at $x = 1$ indicates an odds ratio of 1. Odds ratios to the left of $x = 1$ indicate treatment is associated with a reduction in the logodds of detection of RifR *E. coli* and odds ratios to the right of $x = 1$ indicate treatment is associated with an increase in the logodds of detection of RifR *E. coli*.

significant reduction in the percentage of tank water samples that were positive for RifR *E. coli* (above the detection limit of 10 CFU per ml) from 98 to 24% (Table 3). If we conservatively assume that samples with no counts detected had a count at the minimum detection limit (1 $\log_{10}$ CFU per ml), there was a difference of 2·3 $\log_{10}$ CFU per ml between median counts in SA-treated samples and control samples. Conversely, if we assume all samples with no counts detected had zero CFU per ml, there is an approximate difference between the medians of 3·3 $\log_{10}$ CFU per ml. It is possible that the treatment could have achieved an even higher reduction with higher contamination levels.

The effect was also apparent in water samples collected from the field during irrigation; 30% of field water samples on plots irrigated with SA-treated water were positive for RifR *E. coli* (above the detection limit of 1 $\log_{10}$ CFU per ml) compared with 99% of samples on control plots. The effect of SA treatment on the proportion of soil and produce samples was only statistically significant following the 5th irrigation during the spinach trials (Table 3; Fig. 2).

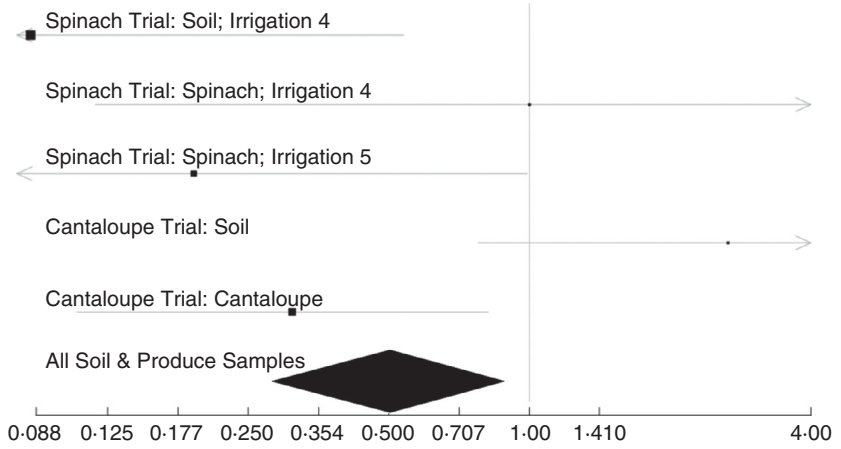

Figure 2 Forest plot to show the odds ratios and confidence intervals for the effects of UV treatment on the odds of samples being positive for Rifampicin-resistant (RifR) *Escherichia coli*, by produce type trial, sample type and irrigation number. Symbols show odds ratios and their sizes are inversely related to the standard error. Lines show confidence intervals and arrows indicate that they extend beyond displayed axis. Line at $x = 1$ indicates an odds ratio of 1. Odds ratios to the left of $x = 1$ indicate treatment is associated with a reduction in the logodds of detection of RifR *E. coli* and odds ratios to the right of $x = 1$ indicate treatment is associated with an increase in the logodds of detection of RifR *E. coli*.

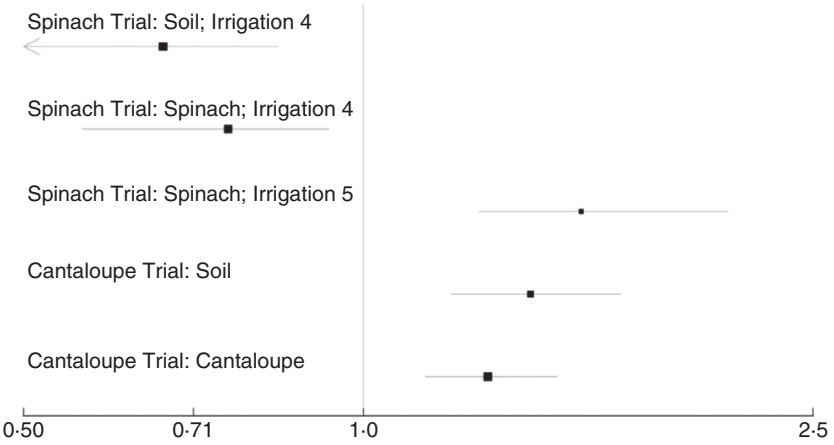

Figure 3 Forest plot to show the odds ratios and confidence intervals for the effects of time since irrigation on the odds of samples being positive for Rifampicin-resistant (RifR) *Escherichia coli*, by produce type trial, sample type and irrigation number. Symbols show odds ratios and their sizes are inversely related to the standard error. Lines show confidence intervals and arrows indicate that they extend beyond displayed axis. Line at $x = 1$ indicates an odds ratio of 1. Odds ratios to the left of $x = 1$ indicate time is associated with a reduction in the logodds of detection of RifR *E. coli* and odds ratios to the right of $x = 1$ indicate time is associated with an increase in the logodds of detection of RifR *E. coli*.

Effects of time since irrigation

The trends in Log_{10} CFU per sample over time since irrigation varied considerably among soil and produce samples, with net growth (an increase in *both* the proportion of samples in which RifR *E. coli* was detected ('positive samples') *and* an increase in Log_{10} CFU per sample among positive samples) occurring during the cantaloupe trial and following the 5th irrigation (regrowth) during spinach trials, and net decay (a decrease in the proportion of samples in which RifR *E. coli* was detected ('positive samples') *and* a decrease in Log_{10} CFU per sample among positive samples) occurring following the 4th irrigation (1st cut) during spinach trials (Fig. 3).

Effect of irrigation type

After SA treatment, RifR *E. coli* prevalence was much higher in furrow tank samples (46%) than drip tank samples (2%). A mixed effect regression model was fit to SA-treated samples only, which included tank as a random effect and irrigation type as a fixed effect. There was a high level of variability in prevalence between tanks (standard deviation 6 on the log odds scale) but irrigation type remained marginally significant ($P = 0{\cdot}06$). Repeated analysis using UV-treated samples showed no effect of irrigation type ($P = 0{\cdot}88$) but a high level of variability in prevalence between tanks (standard deviation 8·3 on the log odds scale). Irrigation type was not however a significant predictor of RifR *E. coli* status in soil or produce samples.

Correlation between soil and produce sample pairs

There was statistical evidence of a weak correlation between Log_{10} Rif *E. coli* counts in soil and spinach samples collected from the same location during year 2 (Spearman's rank correlation; rho = 0·34; $P = 0{\cdot}0001$). There was no evidence of a correlation between Log_{10} Rif *E. coli* counts in soil and spinach samples in year 1 or between Log_{10} Rif *E. coli* counts in soil and cantaloupe samples (Spearman's rank correlation; rho < 0·15; $P > 0{\cdot}10$).

Weather

Because of interaction between the time since sampling and irrigation number, results were reported separately by irrigation number. Based on the final multivariable models, in the spinach trials, higher minimum temperature five days prior to sampling was associated with a decrease in the logodds of detecting RifR *E. coli* in spinach (Table 4; $P = 0{\cdot}02$) and higher minimum temperature one day prior to sampling was associated with a decrease in the logodds of detecting RifR *E. coli* in soil (Table 4; $P = 0{\cdot}03$). Time-series plots revealed that following irrigation 4 in years 1 and 2, temperatures increased over the duration of the irrigation cycle while the logodds1 of detecting RifR *E. coli* in both soil and spinach decreased over the duration of the trial. However following irrigation 5 in year 2, temperatures were in a similar range to those of irrigation 4, but there was an increase in the logodds of detection of RifR *E. coli* over time. Moreover in irrigation 5 in year 1, no RifR *E. coli* was detected on spinach. In the cantaloupe trial, higher mean temperature three days prior to sampling was associated with a decrease in the logodds of detection of RifR *E. coli* in cantaloupe (Table 4; $P = 0{\cdot}004$) whereas a higher precipitation four days prior to sampling was associated with a decrease

Table 4 Effects of time weather variables on proportion of samples positive to RifR *Escherichia coli* (models B)

Trial and irrigation (irr) numbers	Sample type	Variable	Odds ratio	95% Confidence interval	*P* value*
Spinach Years 1 and 2: irr 4	Soil	Minimum temperature (C) 1 day prior to sampling	0·72	0·51–0·95	0·03
		Days since irrigation	0·77	0·58–0·97	0·03
		Control treatment group	Baseline	Baseline	Baseline
		UV treatment group	0·07	0·004–0·5	0·02
		SA treatment group	1·0	0·29–3·5	1·0
Spinach Years 1 and 2: irr 4	Spinach	Minimum temperature (C) 5 days prior to sampling	0·94	0·84–1·04	0·19
		Days since irrigation	0·85	0·61–1·08	0·24
		Control treatment group	Baseline	Baseline	Baseline
		UV treatment group	1·0	0·12–8·6	0·17
		SA treatment group	3·2	0·70–22·2	1·0
Spinach Years 1 and 2: irr 5	Spinach	Minimum temperature (C) 5 days prior to sampling	0·70	0·40–1·0	0·08
		Days since irrigation	1·6	1·3–2·4	0·0005
		Control treatment group	Baseline	Baseline	Baseline
		UV treatment group	0·18	0·02–1·0	0·07
		SA treatment group	0·08	0·004–0·59	0·03
Cantaloupe Trial: irr 4	Soil	Precipitation (cm) 4 days prior to sampling	0·38	0·20–0·65	0·001
		Days since irrigation	2·6	1·7–4·5	<0·001
		Control treatment group	Baseline	Baseline	Baseline
		UV treatment group	3·2	0·83–13·8	0·10
		SA treatment group	0·58	0·13–2·5	0·47
Cantaloupe Trial: irr 4	Cantaloupe	Mean temperature (C) 3 days prior to sampling	0·64	0·47–0·86	0·004
		Days since irrigation	1·1	0·96–1·3	0·16
		Control treatment group	Baseline	Baseline	Baseline
		UV treatment group	2·9	0·09–0·78	0·02
		SA treatment group	0·52	0·20–1·3	0·17

*Wald test *P* value from final mixed effect logistic regression model (model B). All weather variables were included as continuous variables.

in the logodds of detection of RifR *E. coli* in the soil (Table 4; $P = 0{\cdot}001$). Time-series plots revealed that there was a 2-day period of high rainfall (up to 4 cm) which was correlated with increased logodds of detection of RifR *E. coli* 4–5 days later, however, in the multivariable model that included days since irrigation, the apparent direction of the effect of precipitation was reversed.

Classification and regression trees

The optimal regression trees for tank water after treatment and field water were similar (Figs 4 and 5). Water from plots treated with either SA or UV treatments had a lower predicted mean $\log_{10}$ CFU per ml of RifR *E. coli* than water from control plots. Within samples from SA- or UV-treated plots, treatments were apparently more effective on drip plots, compared to furrow, regardless of treatment type. Within furrow plots, UV treatment was apparently more effective than SA treatment. The only exception to this was during the cantaloupe trial, tank water samples treated with SA had a predicted mean count of zero $\log_{10}$ CFU per ml.

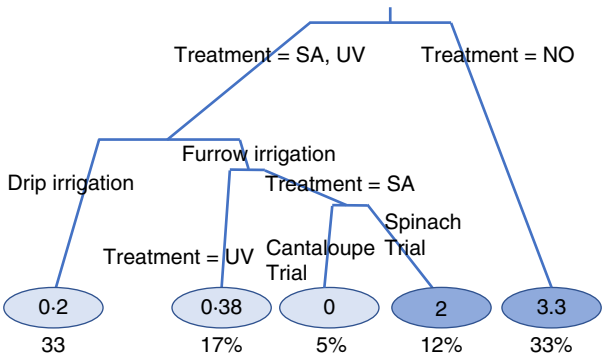

Figure 4 Regression tree for tank water samples after treatment. Total number of samples = 502. Notations: UV = UV treatment group plots; SA = sulphuric acid fertilizer treatment group plots; NO = control group plots; drip = plots irrigated by drip irrigation system; furrow = plots irrigated by furrow irrigation system; CantaloupeTrial = plots in the cantaloupe trial; SpinachTrial = plots in the spinach trials. Variables closer to the root of the tree are the stronger predictors of the mean level of $\log_{10}$ CFU per ml of rifampicin-resistant *Escherichia coli* (numbers given in the node at the base of each branch). % values at the base of each node are the percentage of samples in each branch.

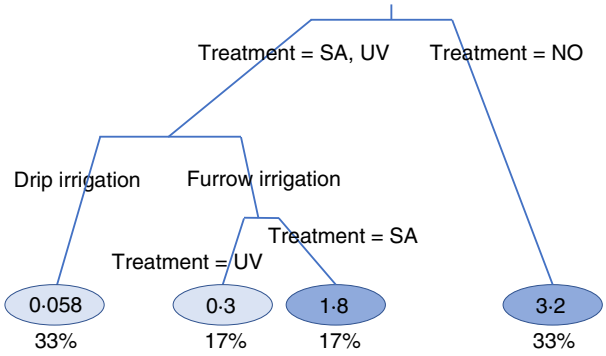

Figure 5 Regression tree for field water samples. Total number of samples = 503. Notations: UV = UV treatment group plots; SA = sulphuric acid fertilizer treatment group plots; NO = control group plots; drip irrigation = plots irrigated by drip irrigation system; furrow irrigation = plots irrigated by furrow irrigation system. Variables closer to the root of the tree are the stronger predictors of the mean level of Log_{10} CFU per ml of Rifampicin-resistant *Escherichia coli* (numbers given in the node at the base of each branch). % values at the base of each node are the percentage of samples in each branch.

The optimal regression tree for soil (Fig. 6) suggests that counts of RifR were higher if three days before sampling there was a higher total solar radiation (>20 MJ m^{-2}), particularly on all furrow plots or on UV-treated drip plots. On sampling days that were 3 days after a day with total solar radiation of <20 MJ m^{-2}, RifR *E. coli* counts were predicted to be higher on days when the minimum relative humidity had been >64%, 5 days prior, particularly on control or SA-treated plots.

The optimal regression tree for spinach (Fig. 7) suggests that mean temperature on the day of sampling was the strongest predictor of log_{10} CFU per ml of RifR *E. coli*. At mean daily temperatures <26°C, the predicted mean count was 0·032 log_{10} CFU per ml of diluent (equivalent to 0·032 log_{10} CFU per gram of spinach) regardless of treatment group. At mean daily temperatures >26°C, control group samples had a predicted mean count of 2·1 log_{10} CFU per ml of diluent and SA or UV treatment groups samples had a predicted count of 0·54 log_{10} CFU per ml of diluent, however, samples collected on these days accounted for only 4% of all spinach samples.

The optimal regression tree for cantaloupe (Fig. 8) suggests that total precipitation four days prior to sampling was the strongest predictor of log_{10} CFU per ml of RifR *E. coli*. On sampling days that were four days after total precipitation of >2·3 cm, the predicted mean count was 1·1 Log_{10} CFU per ml of diluent regardless of treatment group. On the remaining sampling days, predicted counts were 0·25 log_{10} CFU per ml of diluent regardless of treatment group.

Discussion

A recent review summarized current knowledge regarding the UV treatment and solar-driven disinfection as strategies to reduce contamination of irrigation water in produce growing (Banach and Fels-Klerx 2020). UV treatment of irrigation water involves no use of chemicals and is therefore a particularly attractive environmentally friendly option that has no impact on soil pH. However, initial investment for equipment is costly, and requires technical expertise. To the best of our knowledge, this was the first study that tested a field

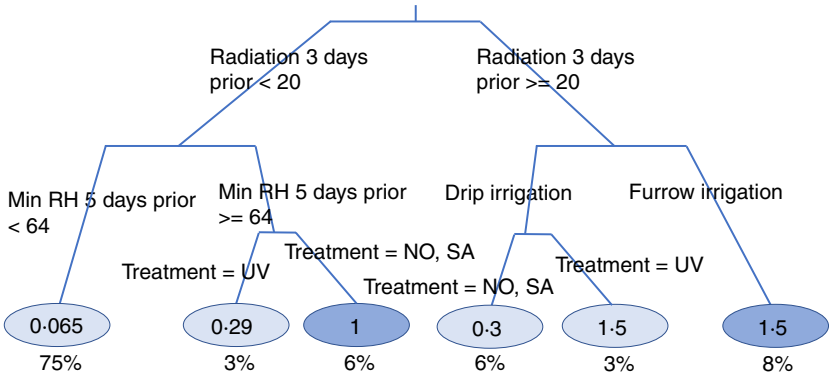

Figure 6 Regression tree for soil samples. Samples collected prior to irrigation excluded. Total number of samples = 288. Notations: Radiation 3 days prior = Total daily solar radiation in mJ m^{-2} 3 days prior to sampling; Min RH 5 days prior = Minimum relative humidity 5 days prior to sampling (%); UV = UV treatment group plots; SA = sulphuric acid fertilizer treatment group plots; NO = control group plots; drip irrigation = plots irrigated by drip irrigation system; furrow irrigation = plots irrigated by furrow irrigation system. Variables closer to the root of the tree are the stronger predictors of the mean level of rifampicin-resistant (RifR) *Escherichia coli* in log_{10} CFU per ml of diluent (standard volume of diluent used per bootsock of soil). The node at the base of each branch shows the mean RifR *E. coli* in log_{10} CFU per ml. % values at the base of each node are the percentage of samples in each branch.

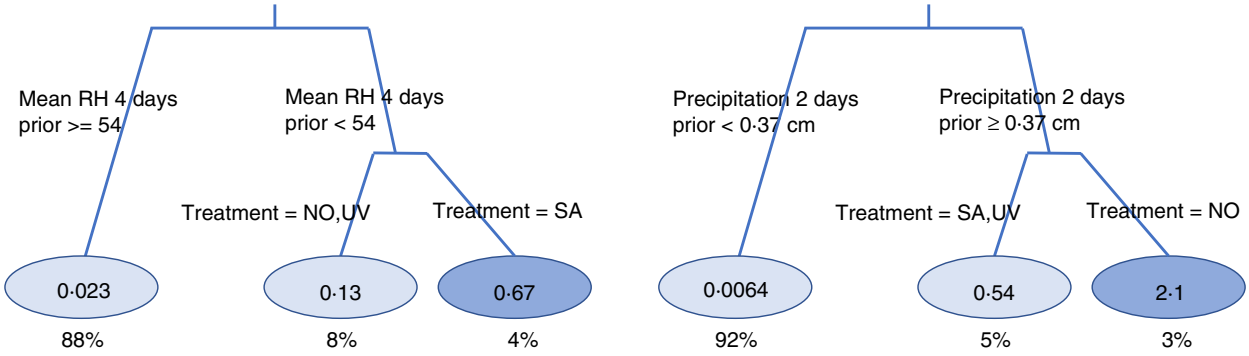

Figure 7 Regression trees for spinach samples. Left: samples taken after 4th irrigation (first cut). Total number of samples = 539. Right: samples taken after 5th irrigation (second cut). Total number of samples = 540. Samples collected prior to irrigation excluded. Notations: Mean RH 4 days prior = mean daily relative humidity 4 days prior to sampling (%); UV = UV treatment group plots; SA = Sulphuric Acid fertilizer treatment group plots; NO = control group plots; Precipitation 2 days prior = total daily precipitation two days prior to sampling (cm). Variables closer to the root of the tree are the stronger predictors of the mean level of Rifampicin-resistant *E. coli* in Log_{10} CFU per ml of diluent for spinach (given in the node at the base of each branch). % values at the base of each node are the percentage of samples in each branch.

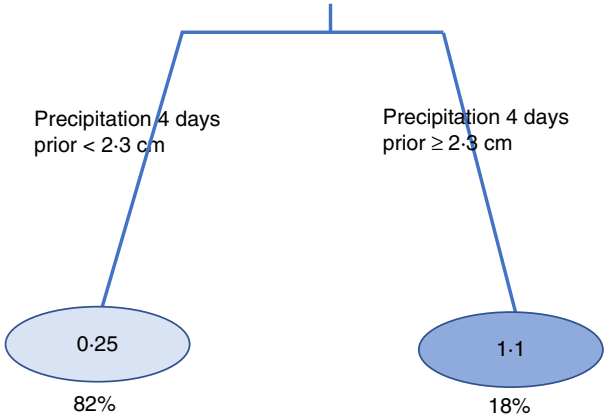

Figure 8 Regression tree for cantaloupe samples. Samples collected prior to irrigation excluded. Total number of samples = 198. Notations: Precipitation 4 days prior = Total daily precipitation in cm 4 days prior to sampling; Rifampicin-resistant *Escherichia coli* in Log_{10} CFU per ml of diluent for cantaloupe is given in the node at the base of each branch. % values at the base of each node are the percentage of samples in each branch.

application of the UV treatment in a controlled trial in open produce fields, and the results confirm its effectiveness at reducing the bacterial load, that had been previously shown under laboratory conditions (Jones *et al.* 2014) and in a semi-open high tunnel (Banach *et al.* 2021).

Previous studies have shown the ability of *E. coli* to photorepair following UV damage, when samples are exposed to sunlight (Bohrerova and Linden 2007). Despite this, we found some evidence of a reduction in the proportion of RifR *E. coli* contaminated soil and produce samples, comparing UV-treated plots with control plots in both the spinach and cantaloupe trials.

Due to differences in the trend in CFU per sample over time, which varied between trials and irrigation numbers, data from different trials were analysed separately. In addition, the number of positive samples on produce (particularly on spinach) was relatively low, which could be partly due to the minimum limit of detection (100 CFU per gram of spinach). This resulted in relatively wide confidence intervals for the treatment effect of UV and as a result, we cannot rule out a relatively strong effect of UV treatment (the lower bounds of confidence interval for odds ratios were as low as 0·02).

Sulphuric acid-fertilizer treatment has the advantage of being readily available and relatively inexpensive. However, use of acid-fertilizer may have unintended side effects on soil, such as affecting the soil chemistry or even microbiome, that may limit its use in places with acidic soils. To our knowledge, this study was the first to demonstrate the effectiveness of SA treatment at reducing the microbial load in irrigation water. There were relatively wide confidence intervals for the treatment effects of SA, which allows for the possibility that the effect of SA treatment could be relatively strong.

We cannot rule out that some of the RifR *E. coli* detected in water, soil or on produce could have been naturally occurring, as opposed to originating from the inoculation of irrigation water. If this was the case, it is expected that it would have affected all water tanks and plots (treatment and control) equally, meaning that we could have under-estimated the effectiveness of the treatments at reducing the presence of *E. coli* originating from irrigation water.

Conversely, samples with low levels of RifR *E. coli* contamination could have been misclassified as negative due to the minimum detection limits (10 CFU per ml of water; 100 CFU per gram of spinach; 900 CFU per bootsock of soil and 4000 CFU cantaloupe rind^{-1}), which could explain relatively low proportions of positive soil and produce samples, including in control plots. This could have also caused an under-estimation of the effectiveness of treatments at reducing the presence of RifR *E. coli* in soil and on produce.

Wind speed has previously been identified as a predictor for the presence of *E. coli* on farms (Chase *et al.* 2019). Although wind speed was not found to be a significant predictor in our study, we cannot rule out that possibility that there was some spread of RifR *E. coli* between control and treated plots, mediated by wind or by fomites, which could cause an under-estimation of the treatment effect. Reasonable precautions were taken to try to limit cross-contamination among plots. For each trial and year, new plots were used, thus avoiding residual contamination overtime between trials. For every two beds of the experimental plot, there were two 40-inch empty beds, to limit contamination between treatment groups and border effects. It is also possible that naturally occurring RifR *E. coli* could have caused an under-estimation of the apparent effect of UV or SA treatment.

In addition, although water was always treated according to treatment group, some irrigations were carried out without inoculation or sampling taking place. It is possible that water availability from irrigation could increase growth of RifR *E. coli* that are already present from the previous irrigation, perhaps at low levels, however, there was only one such irrigation performed between the 4th and 5th irrigations. The effectiveness of UV treatment at reducing contamination of produce may be dependent on environmental conditions post irrigation, which can potentially facilitate growth of even small numbers of bacteria that may have been resistant to the UV treatment. Alternatively, bacteria could recover from UV treatment.

One factor that may affect the effectiveness of UV treatment is the turbidity of water, since highly turbid water does not allow UV radiation to penetrate well (Chase *et al.* 2019). Jones *et al.* (2014) found that high turbidity was associated with a reduced effectiveness of UV treatment in water samples, however, in our study we did not find a statistically significant association between turbidity and the microbial counts after treatment. The reason for the discrepancy is unclear.

SA was apparently more effective in drip irrigated plots than furrow irrigated plots, which could be explained by higher achieved concentration of the fertilizer in drip irrigation, and fertilizer and water being more effectively mixed for the drip system due to higher water pressures. Our study was conducted on an experimental plot with alkaline soil (pH 8·1). It is possible that the effectiveness of SA may vary according to the chemical nature of the irrigation water and/or soil. Indeed, sulphuric acid fertilizer is used to correct the pH of alkaline soils (Chase *et al.* 2019).

The apparent effectiveness of SA and UV treatments could be sensitive to the storage and transport conditions of the samples in our study. Samples in our study were kept below 10°C during storage and transport. Under such conditions, previous authors have found no significant difference between *E. coli* counts after 6 h and after 24 h (Bhullar *et al.* 2019). When we compared water samples from control and UV-treated plots that were shipped on the day of sampling compared to the following day, we observed no statistically significance in the logodds of a sample being positive for RifR *E. coli*. However, we found that SA-treated water samples that were shipped on the day of collection were less likely to be positive for RifR *E. coli* than those that were shipped on the previous day, although in a mixed effect regression model that accounted for tank as a random effect, the effect was no longer significant ($P = 0{\cdot}50$). Additionally, in almost every case, those shipped on the day of sampling were from drip irrigation systems and those shipped on the day after sampling were from furrow irrigation systems. Therefore, it is not possible to determine if it was the irrigation type or the transport delay that caused the univariate effect, although there was no statistically significant effect of transport delay on samples from control or UV-treated plots, suggesting that an interaction between SA treatment and the irrigation system was a more likely explanation.

Sulphuric acid-based fertilizer appeared to be less effective at eliminating *Salmonella* than UV treatment, however, there was not sufficient statistical power to determine if this association was statistically significant or not. Nevertheless, SA-based fertilizer may not be an effective control in areas with frequent *Salmonella* contamination problems and/or alkaline soils (because of counteractive effects of the soil to the acidifying properties of the SA-fertilizer).

The trend over time was variable within our study despite relatively similar conditions for each trial and each irrigation cycle. In summary, net growth occurring during the cantaloupe trial, and following the 5th irrigation during the spinach trial where the second cut of spinach (regrowth) was harvested, but net decay occurring following the 4th irrigation where the first cut of spinach was harvested. The differences were not fully explained by weather, although temperature and precipitation did explain some of the variation. The results of both the

regression and CART analyses suggest that weather did play a role. Rainfall has previously been associated with contamination of spinach with *E. coli* (Park *et al*. 2014; Park *et al*. 2015), relative humidity has been strongly associated with the die-off pattern of *E. coli* and attenuated *Salmonella* in a multi-region study (Belias *et al*. 2020), and soil moisture was strongly associated with persistence of *E. coli* in soil (Gutierrez-Rodriguez *et al*. 2011).

However, despite similar temperatures in irrigations 4 and 5 of spinach, there were different trends in growth/decay. This could be due to the confounding effect of an unmeasured variable or a complex interaction between variables that is not captured by the model. In addition, spinach plants had a different morphology in the 5th irrigation compared to the 4th, because they represented a second cut (i.e. regrowth), which might have contributed to the observed decrease in contamination over time after irrigation 4 and an increase after irrigation 5. This would suggest that the first cut of spinach crop may be considered microbiologically safer than the second cut and is in agreement with an observational study conducted in the same region that showed a positive association between the odds of spinach contamination and the time since planting (Park *et al*. 2013). Spinach has also been shown to internalize *E. coli* when sprayed with contaminated irrigation water after cutting (Davey *et al*. 2013).

Previous studies have found that produce is more likely to become contaminated via irrigation water when furrow irrigation is used, compared to drip irrigation (Stine *et al*. 2005). Our study did not replicate this finding, but this could have been due to the relatively low number of positive produce samples.

The findings are of direct relevance to FSMA guidelines which allows producers to mitigate against contaminated surface irrigation water by applying a wait-time post last irrigation before harvest. The results of our field study confirm that both UV and SA treatments were effective at reducing *E. coli* levels from approximately 3 $\log_{10}$ CFU per ml to <1 $\log_{10}$ CFU per ml. It should be noted that the FSMA guidelines use a threshold of <1·26 $\log_{10}$ CFU per ml, which was below the minimum detection limit of our field study. It is therefore possible that some of the samples considered below the detection limit in our study would not have met the FMSA criteria. Since these treatments are likely to be effective at reducing counts of bacteria other than *E. coli*, some viruses and even oomycetes (Hanes *et al*. 2002; Hijnen *et al*. 2006; Jones *et al*. 2014) of concern to human health, as well as plant pathogens transmitted via irrigation water (Hong and Moorman 2005), there may be additional benefits to growers of using these treatments.

Critically, our results suggest that there is no safe wait-time between use of contaminated irrigation water and harvest, as *E. coli* can grow on produce in some field conditions. Del Rosario and Beuchat (1995) showed that *E. coli* can grow on the surface of stored cantaloupe rinds at 25°C, which is within the range of daily mean temperatures during this cantaloupe trial. *E. coli* has also been shown to grow within spinach plants and in soil, under experimental conditions; however, this has not been shown under field conditions previously to our knowledge (Wood *et al*. 2010; Fonseca *et al*. 2011; Moyne *et al*. 2011; Bezanson *et al*. 2012; Snellman *et al*. 2014).

In conclusion, UV and SA treatments are viable and effective methods of treating surface irrigation water. Applying wait-times after irrigation and prior to harvest is not always a reliable means of mitigating against contaminated produce. The results are based on a single produce-growing region. Additional studies with different locations and water and soil properties are warranted before wide-spread application of the UV and SA-fertilizer treatment methods evaluated in this study.

Acknowledgements

The authors thank Eduardo Serna for help with the setup of the controlled trials and logistics of sample collection at Texas A&M AgriLife - Weslaco. This work was supported by Agriculture and Food Research Initiative-Foundational Program grant 2016-67017-24421, program code A1331 (Improving Food safety), to R.I. from the U.S. Department of Agriculture National Institute of Food and Agriculture (USDA-NIFA; https://nifa.usda.gov/). The work was also partially supported by the Center for Produce Food Safety (CPS; https://www.centerforproducesafety.org/), award number 2018CP504, to R.I. Any opinions, findings, and conclusions expressed in this material are those of the authors and do not necessarily reflect the views of the USDA-NIFA nor CPS. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

RI, JA, JE, KKN and RW conceived and planned the study. CM and UC carried out the field experiments under direction of JA and JE. AKE carried out the microbial laboratory testing of all water, produce and soil samples, overseen by KKN. WB carried out the statistical analysis with contributions for different parts from AKE, AB and MW, overseen by RI. WB and AKE took the lead in writing the manuscript. All authors provided critical feedback and helped shape the research, analysis and manuscript.

Conflict of Interest

No conflict of interest declared.

References

Allende, A. and Monaghan, J. (2015) Irrigation water quality for leafy crops: a perspective of risks and potential solutions. *Int J Environ Res Public Health* **12**, 7457–7477.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S. *et al*. (2008) The RAST Server: rapid annotations using subsystems technology. *BMC Genom* **9**, 75.

Banach, J.l., Hoffmans, Y., Appelman, W., van Bokhorst-van de Veen, H. and van Asselt, E. (2021) Application of water disinfection technologies for agricultural waters. *Agr Water Manage* **244**, 106527–https://doi.org/10.1016/j.agwat.2020.106527.

Banach, J.L. and Van der Fels-Klerx, H.J. (2020) Microbiological reduction strategies of irrigation water for fresh produce. *J Food Prot* **83**, 1072–1087.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I. *et al*. (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* **19**, 455–477.

Bates, D., Maechler, M., Bolker, B. and Walker, S. (2015) Fitting linear mixed-effects models using lme4. *J Stat Softw* **67**, 1–48. https://doi.org/10.18637/jss.v067.i01.

Belias, A.M., Sbodio, A., Truchado, P., Weller, D., Pinzon, J., Skots, M. *et al*. (2020) Effect of weather on the die-off of *Escherichia coli* and attenuated *Salmonella* Typhimurium on preharvest leafy greens following irrigation with contaminated water. *Appl Environ Microbiol* https://doi.org/10.1128/AEM.00899-20.

Bezanson, G., Delaquis, P., Bach, S., McKellar, R., Topp, E., Gill, A., Blais, B. and Gilmour, M. (2012) Comparative examination of *Escherichia coli* O157:H7 survival on romaine lettuce and in soil at two independent experimental sites. *J Food Prot* **75**, 480–487.

Bhullar, M.S., Shaw, A., Hannan, J. and Andrews, S. (2019) Extending the holding time for agricultural water testing EPA method 1603 for produce growers. *Water (Switzerland)* **11** (10), 1–6. https://doi.org/10.3390/w11102020.

Bohrerova, Z. and Linden, K.G. (2007) Standardizing photoreactivation: comparison of DNA photorepair rate in *Escherichia coli* using four different fluorescent lamps. *Water Res* **41**, 2832–2838.

Chang, J.M. and Fang, T.J. (2007) Survival of *Escherichia coli* O157:H7 and *Salmonella enterica* serovars Typhimurium in iceberg lettuce and the antimicrobial effect of rice vinegar against *E. coli* O157:H7. *Food Microbiol* **24**, 745–751.

Chang, T. (2015) *Evaluation of multiple disinfection methods to mitigate contaminated irrigation water*. Master's Thesis, University of Tennessee.

Chase, J.A., Partyka, M.L., Bond, R.F. and Atwill, E.R. (2019) Environmental inactivation and irrigation-mediated regrowth of *Escherichia coli* O157:H7 on romaine lettuce when inoculated in a fecal slurry matrix. *PeerJ* 1–19.

Davey, L.E., Erickson, M.C., Webb, C.C., Di, J.C., Payton, A.S., Flitcroft, I.D., Phatak, S.C. and Doyle, M.P. (2013) Internalization of *Escherichia coli* O157:H7 following spraying of cut shoots when leafy greens are regrown for a second crop. *J Food Prot* **76**, 2052–2056.

Enciso-Medina, J., Multer, W.L. and Lamm, F.R. (2011) Management, maintenance, and water quality effects on the long-term performance of subsurface drip irrigation systems. *Appl Eng Agric* **27**, 969–978.

FAO and WHO (2019). Safety and quality of water used in food production and processing - meeting report. Microbiological Risk Asssessment Series no. 33. Rome.

FDA (U.S. Food and Drug Administration). (2011) *GRAS Notification for Sulfuric Acid and Sodium Sulfate Blend used as an Antimicrobial in Meat and Poultry Processing*. Silver Spring, MD: FDA.

FDA (U.S. Food and Drug Administration) (2013) *Standards for the Growing, Harvesting, Packing and Holding of Produce for Human Consumption, A Proposed Rule*. Silver Spring, MD: FDA.

FDA (U.S. Food and Drug Administration) (2014) Standards for the growing, harvesting, packing, and holding of produce for human consumption. Proposed rule; supplemental notice of proposed rulemaking. 2014 [Docket No. FDA–2011–N–0921].

Fonseca, J.M., Fallon, S.D., Sanchez, C.A. and Nolte, K.D. (2011) *Escherichia coli* survival in lettuce fields following its introduction through different irrigation systems. *J Appl Microbiol* **110**, 893–902.

Gordon, M. and Lumley, T. (2017) forestplot: Advanced Forest Plot Using 'grid' Graphics. R package version 1.7.2. https://CRAN.R-project.org/package=forestplot

Gregory, J. (2001) *Uses of Sulfuric Acid as a Water Amendment in Agriculture*. Presented at the International Irrigation Show, San Antonio, Texas, November 4, 2001.

Gutierrez-Rodriguez, E., Gundersen, A., Sbodio, A.O. and Suslow, T.V. (2011) Variable agronomic practices, cultivar, strain source and initial contamination dose differentially affect survival of *Escherichia coli* on spinach. *J Appl Microbiol* **112**, 109–118.

Hanes, D.E., Worobo, R.W., Orlandi, P.A., Burr, D.H., Miliotis, M.D., Robl, M.G., Bier, J.W., Arrowood, M.J. *et al*. (2002) Inactivation of *Cryptosporidium parvum* oocysts in fresh apple cider by UV irradiation. *Appl Environment Microbiol* **68**, 4168–4172. https://doi.org/10.1128/AEM.68.8.4168-4172.2002.

Hijnen, W.A.M., Beerendonk, E.F. and Medema, G.J. (2006) Inactivation credit of UV radiation for viruses, bacteria and protozoan (oo)cysts in water: a review. *Water Res* **40**, 3–22.

Hong, C.X. and Moorman, G.W. (2005) Plant pathogens in irrigation water: challenges and opportunities. *CRC Crit*

Rev Plant Sci **24**, 189–208. https://doi.org/10.1080/07352680591005838.

Jang, H. and Matthews, K.R. (2018) Survival and interaction of *Escherichia coli* O104:H4 on *Arabidopsis thaliana* and lettuce (*Lactuca sativa*) in comparison to *E. coli* O157:H7: Influence of plant defense response and bacterial capsular polysaccharide. *Food Res Int* **108**, 35–41.

Jones, L.A., Worobo, R.W. and Smart, C.D. (2014) UV light inactivation of human and plant pathogens in unfiltered surface irrigation water. *Appl Environ Microbiol* **80**, 849–854.

Kenny, J.F., Barber, N.L., Hutson, S.S., Linsey, K.S., Lovelace, J.K. and Maupin, M.A. (2009) Estimated use of water in the United States in 2005. *US Geol Surv Circ* **52**.

López-Gálvez, F., Sampers, I., Gil, M.I. and Allende, A. (2017) Modelling of *E. coli* inactivation by chlorine dioxide in irrigation water. *Agric Water Manag* **192**, 98–102.

Lynch, M.F., Tauxe, R.V. and Hedberg, C.W. (2009) The growing burden of foodborne outbreaks due to contaminated fresh produce: risks and opportunities. *Epidemiol Infect* **137**, 307–315.

Moyne, A.-L., Sudarshana, M.R., Blessington, T., Koike, S.T., Cahn, M.D. and Harris, L.J. (2011) Fate of *Escherichia coli* O157:H7 in field-inoculated lettuce. *Food Microbiol* **28**, 1417–1425.

Nucera, D.M., Maddox, C.W., Hoien-Dalen, P. and Weigel, R.M. (2006) Comparison of API 20E and invA PCR for identification of *Salmonella enterica* Isolates from swine production units. *J Clin Microbiol* **44**, 3388–3390.

Pandrangi, S., Elwell, M.W., Anantheswaran, R.C. and Laborde, L.F. (2006) Efficacy of sulfuric acid scarification and disinfectant treatments in eliminating *Escherichia coli* O157:H7 from alfalfa seeds prior to sprouting. *J Food Sci* **68**, 613–617.

Park, S., Navratil, S., Gregory, A., Bauer, A., Srinath, I., Jun, M., Szonyi, B., Nightingale, K. *et al*. (2013) Generic *Escherichia coli* contamination of spinach at the preharvest stage: effects of farm management and environmental factors. *Appl Environ Microbiol* **79**, 4347–4358.

Park, S., Navratil, S., Gregory, A., Bauer, A., Srinath, I., Szonyi, B., Nightingale, K., Anciso, J. *et al*. (2014) Farm management, environment, and weather factors jointly affect the probability of spinach contamination by generic *Escherichia coli* at the preharvest stage. *Appl Environ Microbiol* **80**, 2504–2515.

Park, S., Navratil, S., Gregory, A., Bauer, A., Srinath, I., Szonyi, B., Nightingale, K., Anciso, J. *et al*. (2015) Multifactorial effects of ambient temperature, precipitation, farm management, and environmental factors determine the level of generic *Escherichia coli* contamination on preharvested spinach. *Appl Environ Microbiol* **81**, 2635–2650.

Patel, J., Millner, P., Nou, X. and Sharma, M. (2010) Persistence of enterohaemorrhagic and nonpathogenic *E. coli* on spinach leaves and in rhizosphere soil. *J Appl Microbiol* **108**, 1789–1796.

R Core Team (2017) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

del Rosario, B.A. and Beuchat, L.R. (1995) Survival and growth of enterohemorrhagic *Escherichia coli* O157:H7 in cantaloupe and watermelon. *J Food Prot* **58**, 105–107.

Seemann, T. (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069.

Snellman, E., Marianne, F., Ravaliya, K., and Assar, S. (2014) *Memorandum to File-Review of Microbial Decay Constants Reported in Field Trials of Contaminated Produce, September 2014*. Federal document FDA-2011-N-0921-0992. Silver Spring, MD: U.S. Food and Drug Administration.

Stine, S.W., Song, I., Choi, C.Y. and Gerba, C.P. (2005) Application of microbial risk assessment to the development of standards for enteric pathogens in water used to irrigate fresh produce. *J Food Prot* **68**, 913–918.

Therneau, T. and Atkinson, B. (2018). rpart: Recursive Partitioning and Regression Trees. R package version 4.1-13. https://CRAN.R-project.org/package=rpart.

Tyagi, D. (2014) *Impact of pre-harvest environmental factors on the survival of enterohemorrhagic E. coli* and *Salmonella* on lettuce. North Dakota State University of Agriculture and Applied Science.

USEPA (U.S. Environmental Protection Agency). (1993) R.E.D. FACTS: Mineral Acids. EPA-738-F-93-025. Office of Prevention, Pesticides and Toxic Substances.

Van der Linden, I., Cottyn, B., Uyttendaele, M., Vlaemynck, G., Heyndrickx, M., Maes, M. and Holden, N. (2016) Microarray-based screening of differentially expressed genes of *E. coli* O157:H7 Sakai during preharvest survival on butterhead lettuce. *Agriculture*. https://doi.org/10.3390/agriculture6010006

Wickham, H. (2007) Reshaping data with the reshape package. *J Stat Softw* **21**(12).

Wood, J.D., Bezanson, G.S., Gordon, R.J. and Jamieson, R. (2010) Population dynamics of *Escherichia coli* inoculated by irrigation into the phyllosphere of spinach grown under commercial production conditions. *Int J Food Microbiol* **143**, 198–204.

Wright, K.M., Crozier, L., Marshall, J., Merget, B., Holmes, A. and Holden, N.J. (2017) Differences in internalization and growth of *Escherichia coli* O157:H7 within the apoplast of edible plants, spinach and lettuce, compared with the model species *Nicotiana benthamiana*. *Microb Biotechnol* **10**, 555–569.

Supporting Information

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

Table S1. Data corresponding to each sample in the randomized controlled trial.