Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

Rachel L. Usher

April 26, 2020

Date

Characterizing relationships between bacterial indicators of foodborne pathogens on Mexican produce in the pre-harvest and post-harvest environment

By

Rachel L. Usher
Master of Public Health

Department of Epidemiology

_____

Juan S. Leon, PhD, MPH
Committee Chair

Characterizing relationships between bacterial indicators of foodborne pathogens on Mexican produce in the pre-harvest and post-harvest environment

By

Rachel L. Usher

B.S.
University of Georgia
2016

Thesis Committee Chair: Juan S. Leon, PhD, MPH

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2020

Abstract

Characterizing relationships between bacterial indicators of foodborne pathogens on Mexican produce in the pre-harvest and post-harvest environment

By Rachel L. Usher

In recent decades, the United States has experienced a rise in the number of foodborne disease outbreaks impacting millions of Americans every year. The majority of these new outbreaks originate from contaminated produce. Testing for foodborne pathogens such as pathogenic *E. coli* and *Salmonella* in agricultural samples is expensive and time consuming, so indicator bacteria of fecal contamination are commonly used to detect poor sanitation on agricultural products.

There are many potential sources of produce contamination in the agricultural environment including irrigation water, soil, farming equipment, and farmworker hands. Currently, there is little understanding of the fitness and relationships between indicator bacteria and pathogenic enteric bacteria on these surfaces and substrates in the agricultural environment. This study aims to characterize the relations between different bacterial indicators to see if these relationships are modified by sample type at different processing stages from Mexican farms distributing different types of produce.

Between 2011 and 2012, samples were collected from 11 farms producing three types of produce (cantaloupe, tomato, jalapeños), on different associated sample types (produce, soil, irrigation water, hand-rinse), before and after harvest. Each sample was analyzed for the presence and concentration of three bacterial indicators: *E. coli*, coliforms, and *Enterococcus*. In an analysis of the principal component 1 scores we found significant differences in the mean score values between produce samples in the pre-harvest ($p < 0.005$) and post-harvest ($p < 0.001$) environment in addition to significant differences between hand-rinse scores ($p < 0.05$). In all of these comparisons, the cantaloupe associated scores were significantly larger than tomato and jalapeño. We also found significant differences in the indicator bacteria relationships in the discriminant analysis among cantaloupe samples compared to the other produce types. Overall, relationships between bacterial indicators vary significantly by associated sample and produce type indicating physical characteristics of produce, such as the rough skin of a cantaloupe, may create a more ideal habitat for colonization by indicator bacteria than the waxy produce surfaces of jalapeño and tomato. This study will contribute much needed initial research characterizing the relative abundance and association of indicator bacteria within the agricultural environment.

Characterizing relationships between bacterial indicators of foodborne pathogens on Mexican produce in the pre-harvest and post-harvest environment

By

Rachel L. Usher

B.S.
University of Georgia
2016

Thesis Committee Chair: Juan S. Leon, PhD, MPH

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2020

Acknowledgements

I would like to thank Dr. Juan S. Leon for his insight and mentorship throughout the development of this thesis. His steadfast guidance, feedback, and encouragement through this process immeasurably grew my abilities and confidence as a researcher. I would also like to thank Dr. Jessica Prince Guerra for her advice on statiscal analyses and programming during the data analysis portion of this project.

Table of Contents

Literature Review

The number of produce-related foodborne outbreaks in the United States has risen steadily over the past few decades (1, 2). It is currently estimated that 9.4 million Americans get sick from a foodborne illness every year (3). This marks a shift from animal products historically associated with foodborne illness, such as eggs, dairy and meat, to a rise in the number of foodborne illnesses caused by produce (4). This trend is attributed to per capita increases in the amount of fresh produce consumed and the globalization of the food supply (5). The CDC reports that between 2009 and 2015, all 50 states and Puerto Rico reported foodborne outbreaks totaling 145 deaths directly attributed to foodborne illness (6). In an analysis of produce-associated outbreaks between 1973 and 1997, salad, lettuce, juice, melon, sprouts and berries were the most common sources of foodborne illness (1).

A produce-associated foodborne outbreak is defined as, "the occurrence of two or more cases of the same illness in which epidemiologic investigation implicates the same uncooked fruit, vegetable, salad or juice" (1). There are a wide variety of produce-associated bacteria, viruses and fungi implicated in foodborne outbreaks—including strains of *Salmonella*, *Escherichia coli*, *Listeria monoctyogenes*, *Vibrio,* norovirus, Hepatitis A and *Cyclospora* (6). Among all foodborne pathogens in the United States, one study found that *Salmonella* outbreaks have resulted in the most hospitalizations (35%) and deaths (28%) (3). The remainder of this literature review will focus on bacterial foodborne pathogens.

Enteric bacterial pathogens and indicator bacteria

Salmonella is one of the most researched and well-described bacterial species. Belonging to the *Enterobacteriaceae* family, Salmonellae bacteria are gram-negative, rod-shaped and motile. Historically, each serotype of this bacteria was classified as its own species, but now, it is widely accepted that there is a single species, *Salmonella enterica* (7). Serotypes of *S. enterica* are capable of infecting both humans and animals; the serotypes associated with disease in humans generally belong to subgroup 1 (8). *Salmonella* can cause a variety of diseases including enteric fever, bacteremia, enterocolitis and focal infections with enterocolitis being the most common. Symptoms can include diarrhea, fever, discomfort and muscle aches (7). After symptoms disappear, shedding in feces can continue for 8 weeks in adults and up to 20 weeks in children (9). While humans and wild animals can be reservoirs for *Salmonella*, animals, especially farm animals, can be chronic carriers of the bacteria (10). *Salmonella* is highly adaptable in a variety of aerobic environments once it is shed from a carrier's feces, making it particularly successful as a foodborne pathogen (11). This bacteria was the second leading cause of foodborne disease outbreaks and resulted in the most hospitalizations and deaths between 2001 and 2007 compared to 31 major foodborne pathogens (3).

Escherichia coli is a naturally occurring bacterium found in gastrointestinal tract of both animals and humans. It belongs to the genus *Escherichia* that is primarily composed of motile gram-negative bacilli. Specific strains of *E. coli* can cause gastrointestinal disease with different strains classified according to their virulence properties (12). Two groups are named according to the toxins they produce. *E. coli* that produce Shiga toxins, similar to the toxin made by *Shigella dysenteriae*, are called Stx-producing *E. coli* (STEC). *E. coli* that produce verocytotoxins, that act on Vero cells, are

called VT-producing *E. coli* (VTEC) (13). Among *E. coli* strains, STEC and VTEC are the only pathogenic group with a zoonotic reservoir in cattle. Once spread to humans, these pathogenic *E. coli* strains can cause severe diarrhea, hemolytic uremic syndrome (HUS), and death especially among the young and immunocompromised (10). *E. coli* is an enteric pathogen, but this bacterium has been particularly successful as a foodborne pathogen with its ability to withstand a variety of temperatures, pH and nutrient regimes in environments outside the gastrointestinal tract of humans (11).

Widespread testing for foodborne bacteria including pathogenic *E. coli* and *Salmonella* in environmental samples can be time consuming and challenging (14). Instead of testing for pathogens directly, microbial hygiene standards for agriculture and water have historically tested for other indicator bacterial species, such as total coliforms, enterococci and non-pathogenic *E. coli*, as indicators of fecal contamination. The idea is that these indicator bacteria act as index organisms to signal poor sanitary quality (15) and likely co-contamination with enteric pathogens. However, there is increasing uncertainty if these indicator species should continue to be used as a proxy for enteric pathogen contamination (16-20) due to variability in the sizes of microbial indicators and pathogens and abilities to tolerate an array of environmental stressors (14).

Produce contamination in the agricultural environment

During growth and harvest, there are a variety of sources and mechanisms that can contaminate produce. However, there are certain steps in the food production process where contamination is most likely. The locations of most frequent produce contamination include the growing period in the field, the harvesting and initial processing and in final preparation for consumption (21).

It is suggested that enteric pathogens can be transmitted via irrigation water, soil, equipment, hands, wild animals, and domestic animals (22-28) coming into contact with the produce surface. There are some studies that suggest that contamination of produce can even occur via invasion of the root system of the plant (29). In the agricultural environment, enteric pathogens must endure fluctuating physiochemical conditions before they are able to reach another host. Favoring wet conditions and limited UV radiation, studies have demonstrated the persistence of emteric pathogens such as *E. coli* and *Salmonella* in soils (30), manure (31) and water (32, 33) for extended periods of time.

The harvesting process can be a major source of produce contamination because the product is handled by workers who can transmit pathogens on their hands (34, 35). One study found that the majority of farmworker hands sampled tested positive for four bacterial indicators including *E. coli*, coliforms, enterococcus, and coliphages. The origin of this contamination is difficult to identify. Some studies supporting hand to produce contaminantion and vice versa. One intervention study that supplied hand hygiene treatment to workers harvesting jalapeños found that there was significantly decreased contamination on farmworker hands after sanitation. However, after continuing to harvest jalapeños for 30 minutes after the hygiene treatment, the microbial contamination on the hands of individuals receiving the hygiene treatment was virtually indistinguishable from the contamination type and quantity found on harvesters who never received the hygiene treatment (34).

Enteric pathogens are most successful at invading broken or weakened surfaces of leaves or fruiting bodies of plants where there are available nutrients and water (36).

These conditions occur unintentionally (via damaged produce) or intentionally (via preparing pre-cut produce) during the processing of produce which further supports the growth of pathogenic bacteria. Golden Delicious apples with broken skin inoculated pathogenic and non-pathogenic strains of *E. coli* showed exponential increase in growth of all strains over the 144-hour study period (37). *Salmonella* increased by 5 $\log_{10}$ over a week on cut honeydew surfaces held at 20-degrees Celsius (38). These studies demonstrate that enteropathogens are capable colonizers of favorable produce environments.

Survival strategies of bacteria on the plant surface

Bacteria have demonstrated the ability to survive by exploiting favorable niches of the plant environment. These niches are particularly important for microbes when the majority of the plant environment offers limited nutrients, scarce water and direct UV radiation. These unfavorable conditions, specifically limited access nutrients, have been shown to limit bacterial growth on surfaces of leaves in the field (39). While certain parts of the plants surface are nutrient poor such as the waxy top-surfaces of leaves, whole-cell bacterial biosensors have shown that plant sugars including sucrose, fructose and glucose are abundant in leaf oases, veins, and trichomes (40, 41). There is evidence that motile bacteria will preferably live in these areas of the plants with available nutrients while simultaneously avoiding areas with high UV radiation (42). To remain in favorable environments, pathogenic bacteria have been shown to attach to plant cells in the phyllosphere with the same mechanisms they use to attach to human epithelial cells in the gut. For example, recent studies examining virulence factors of different strains of *E. coli* and *Salmonella* have shown that curli, long aggregative fimbriae facilitate binding to

cells in alfalfa sprouts (43, 44) in addition to mouse models (45, 46). These examples show how bacterial enteropathogens use existing motility and attachment strategies to increase their fitness and survival on the surfaces of produce.

Biofilms are a powerful medium to sustain enteric pathogens in the agricultural environment by providing protection and resources. Biofilms are aggregate complex structures that contain a matrix of bacteria and fungi (47). On the surfaces of plants, bacteria contained within biofilms are better able to buffer physical changes in the phyllosphere with the ability to withstand desiccation and nutrient poor conditions. These clumps of microorganisms tend to aggregate in favorable plant niches including leaf veins and glandular trichosomes (48). In a review examining how biofilms interact with the surfaces of plants, it was found that 10% to 40% of the total bacteria population was found within biofilms (49). With a wide variety of bacterial species clustered within biofilms, the exchange of genetic material among species is likely. It has been suggested that biofilms could serve as a medium for human enteric pathogens to increase their resistance by picking up favorable genes from epiphytic plant bacteria (50). Overall, the relationships between bacterial species within favorable ecological niches on plants are complex and poorly described.

Competitive and cooperative action between bacteria on plant surfaces

All surfaces in the agricultural environment are host to an ongoing and evolving interaction between a diversity of microbial species. The perpetual need for resources requires that bacteria use a variety of tools to optimize survival by accumulating nutrients, maintaining their access to favorable environments, and, in some situations, producing toxins with antimicrobial properties to actively inhibit the growth of other

species (51). In a survey of the microbiology literature, Hibbing et al. 2010 assessed the broad understanding of bacterial relationships gained through laboratory and theoretical studies (51). This review concluded that the literature lacks studies examining bacterial relationships in natural environments outside of tightly regulated laboratory conditions. In our opinion, this conclusion applies to the food safety literature with scarce information characterizing relationships, antagonistic or cooperative, between enteric pathogens and/or fecal indicator species in the agricultural environment. The literature most closely related to this topic highlights the relative abundance of enteropathogenic and epiphytic plant bacteria within favorable plant ecological niches (29, 52).

With limitations on resources, bacteria frequently engage in competitive behavior to acquire nutrients and support growth in a variety of ecological niches. This type of behavior is demonstrated between epiphytic bacteria and enteric pathogens on an array of produce surfaces. In a study where lettuce was inoculated with *S. Newport*, *E. coli* O157:H7 and the common epiphytic bacteria *Enterobacter asburiae*, the growth of the two enteric pathogens was reduced 10-fold (52). Similarly, wounded apples inoculated with *E. coli* O157:H7, in addition to *Pseudomonas syringiae*, a common plant bacterium, reduced the growth of *E. coli* (37). While these two previous studies examined bacterial relationships during a limited window of the plant lifecycle, Cooley et al. 2003 built a unique study to examine the fate of both *S. entericia* serovar Newport and *E. coli* O157:H7, in addition to the plant bacteria *Enterobacter asburiae*, over the lifetime of the thale cress plant. This study tracked the population of each bacterial species from germination to plant maturity of thale cress in a hydroponic environment conventionally used in food production. This study saw the proliferation of both species of enteric

pathogens when they were introduced to the root system of the young plants eventually resulting in the contamination of the entire plant when grown in the sterile conditions with 100% humidity. While the amount of enterobacteria was reduced as the plants matured, both pathogens were still detected on the plant 21 days after inoculation. In subsequent tests, *Enterobacter asburiae* was introduced to the environment in addition to *S. entericia* serovar Newport and *E. coli* O157:H7. The growth and success of all three bacterial species was tracked over the growth cycle of the thale cress. In this second set of experiments, the growth of the two enteric species was reduced by 10-fold compared to the initial tests without *Enterobacter asburiae* present (29). Taken together, these three studies suggest that naturally occurring epiphytic bacteria can outcompete enteropathogenic bacteria in plant ecological niches acting as a potential biocontrol agent against enteric pathogens in the agricultural environment.

With few examples in the surveyed literature, there was one study that did assess the relative fitness among pathogenic and indicator bacteria in the same environment. Barak et al. 2002 assessed the motility and propagation of *S. Newport* and *E. coli* O157:H7 on alfalfa sprouts when co-inoculated with other species of epiphytic bacteria. Barak et al. 2002 found that *S. Newport* attached to alfalfa sprouts as easily as a variety of co-inoculated epiphytic bacteria and did so significantly more than *E. coli* O157:H7. In a repeated test with other nonpathogenic *E. coli* serotypes isolated from cabbages, the nonpathogenic *E. coli* attached to sprouts as well as other co-inoculated epiphytic bacteria. It was hypothesized that the biological attachment mechanisms differences between *E. coli* serovars is responsible for the different outcomes (53). This study

demonstrates the variable levels of fitness among pathogenic bacteria and indicator bacteria compared to epiphytic bacteria in a controlled environment.

While the ecological principle of competition is well described and researched in a variety of biological systems, there is an increasing interest in describing cooperative relationships between species. Bacterial models are particularly useful for this type of study due to their short generation time and abundance in a wide array of environmental conditions. Types of cooperative bacterial behavior could involve advanced signaling where bacteria interact with the environment around them via cohesive multicellular behaviors (51). While underlying mechanisms driving symbioses are not always discussed in the food safety literature, there are examples of cooperative relationships between bacteria in the microbiome of produce.

Accurately characterized by Hibbing et al. 2010 in their assessment of the microbial literature discussed earlier. There are scarce on examples of studies examining the interactions between bacteria in environments outside of the laboratory. This lack of in-situ research was indentified in a review of microbial ecology literature (51). However, a few studies have examined the interactions between enteric bacteria on plant models in the laboratory. In a study of the fitness of *Salmonella enterica* serovar Thompson on cilantro, the proliferation and growth of *S. enterica* was compared to commonly found plant bacteria *Pantoae agglomerans* and *Pseudomonas chlororaphis*. Visualizations of the bacteria using CLSM found that *S. enterica* and *P. agglomerans* were forming heterogenous aggregations on leaf veins suggesting commensal relationships. However, *S. enterica* was ultimately not as successful as the two native plant bacteria to survive on the cilantro surface (54). Another study examined the survival and growth of *E. coli*

O157:H7 on lettuce in combination with two common epiphytic bacteria *Wausteria paucula* and *Enterobacter asburiae*. The survival of *E. coli* increased six-fold in combination with *W. paucula*. The benefits of commensalism between the two bacterial species were only observed on lettuce foliage and not in the rhizosphere nor exudate. It is possible that the common epiphytic bacteria modified the environment of the plant surface by increasing the availability of nutrients or providing production of a polysaccharide matrix to prevent desiccation of the environment *E. coli* O157:H7 was also inhabiting (52). These studies provide some idea of the multifaced relationships between bacterial species that sometimes work together to exploit existing niches and potentially chemically interact with the plant surface to create favorable conditions within the ecology of the plant.

There are also numerous opportunities for proliferation and symbiosis of bacterial species after the harvesting of produce. In an examination of the success of *Salmonella* in the post-harvest environment, produce with soft rot were twice as more likely to test positive for suspected *Salmonella.* These results were confirmed in the lab when fresh potato, carrot and pepper samples were inoculated with *Salmonella typhimurium*. *Erwina carotovora* is a bacterium that causes soft-rot disease in post-harvest produce. Samples co-inoculated with *Salmonella* and *Erwina carotovora* saw a 10-fold increase in *Salmonella* compared to disks with *Salmonella* alone (55). These examples of bacterial symbioses from the post-harvest environment are similar to those in the pre-harvest environment—bacteria on plant surfaces can benefit from the activity of other species of bacteria through the release of previously unavailable nutrients and creation of new favorable niches for growth.

Purpose and significance of the current study

Currently, there is a need to understand the ecology of bacteria in the agricultural community that ultimately cause foodborne illness. With substantive documentation and understanding of the growth and proliferation of enteric pathogens within host species, there is less knowledge about how they persist and interact in the agricultural environment. Instead of testing for pathogenic enteric bacteria directly in environmental samples, it is standard to test for fecal indicator bacteria as a proxy for likely co-contamination with the bacterial species that cause foodborne illness. There are a variety of bacterial contamination routes throughout the food production process that can originate from contaminated water, organisms, and farmworker hands. To survive the wide ranging physical and chemical conditions in the field and harvesting environments, enteric bacteria have employed a variety of strategies to seek out favorable conditions on the plant surface. The relationships between enteric pathogenic and indicator bacteria are poorly described. From the studies that do exist, we understand that they engage in both competitive and cooperative strategies with naturally occurring plant bacteria to create favorable ecological niches and novel sources of nutrients.

The goal of this exploratory research is to evaluate the relationships between different ecological niches of foodborne bacterial indicators found on United States and Mexican farms that supply fresh produce to the United States. This will be accomplished by examining where bacterial indicators are most commonly found in the harvest and post-harvest environment on different sample types (produce, water, hands and soil) and processing stages (before and after harvest). It is currently unknown if particular ecological niches within the agricultural environment are favorable to co-contamination

by different combinations of bacterial indicators. With a remarkable diversity of bacterial species engaging in complex relationships over wide ranging agricultural environments, it is hypothesized that environmental conditions could impact the success of different bacterial species. To explore these bacterial relationships, the relationships between bacterial indicators will be assessed by sample type and processing stage. Identifying clustering relationships, if any, between fecal bacterial indicators can inform the understanding of the ecology of bacteria within the agricultural environment.

The literature describes a diversity of cooperative and antagonistic relationship between epiphytic and enteropathogenic bacteria. However, there is a lack of studies examining microbial relationships in-situ in the agricultural environment. There is also limited information characterizing how relationships between bacterial species change on different surfaces such as produce, soil, water and hands. This study will contribute much needed initial research characterizing the relative abundance and association of indicator bacteria on different surfaces in the harvest and post-harvest environment.

Data and Methods

Study Sites

The data used in this analysis represents a part of a larger research study analyzing fecal associated pathogen contamination of produce farms in northeast Mexico near the U.S. border (Nueva Leon, Coahuila, and Tamaulipas) between 2011 and 2012 (34, 56). This agricultural area of Mexico was selected due to large quantities of produce types of interest, ideal sampling plans, and a primary selling market in the United States. The produce of interest in this study included tomatoes (*Solanum lycopersicum*), jalapeños (*Capsicum annuum*), and cantaloupes (*Cucumis melo* var. *cantalupensis*). For each produce type, the Mexican state produce associations and the state Secretariat for Agriculture selected 3 to 5 farms to be sampled. A total of 11 farms were included in this study. Seven farms exclusively grew one type of produce; this included one jalapeño farm, one tomato farm, and five cantaloupe farms. Four farms grew both jalapeños peppers and tomatoes. This protocol for sampling was reviewed and approved by three Institutional Review Boards (IRB) located at La Universidad Autónoma de Nuevo León (UANL), North Carolina State University (NCSU), and Emory University (Emory IRB: 00035460, Appendix A).

Dr. Juan Leon at Emory University provided data access privileges. This analysis incorporated information from both direct rinse samples of the produce and affiliated samples found in the environment of the produce sample (ex. Irrigation water, soil, and hand-rinse). To clarify any confusion of language, this analysis will use the word "produce" in two ways: 1) as a general reference to farmed vegetables and fruit, and 2) the grouping of produce and produce associated hand, soil and irrigation samples. The actual names of the produce (ex. tomatoes jalapeño peppers and cantaloupes) will herein

be used to discuss produce itself. Specific sample types within produce categories will be discussed with hyphenated names (ex. cantaloupe-associated irrigation water or jalapeño pepper-associated soil).

Study Design

Sample Collection

Samples were collected according to the methods described in Bartz et al. 2017 (34). While the larger study gathered samples from multiple points in the production process, including before and after harvesting, through distribution in the packing shed, these analyses focused exclusively on the matched samples collected before and after harvest. In the "before harvest" environment, three types of samples were collected including produce rinse samples, irrigation water samples, and soil samples. Three subsamples comprised one produce rinse sample (n = 79, 1500mL) gathered in 0.1% peptone water from tomatoes (n = 23), jalapeño peppers (n = 19) and cantaloupes (n = 37). After sanitizing irrigation hoses with 200-ppm hypochlorite and allowing the water to run for 30-seconds, irrigation water samples (n = 72, 4500 mL) were collected from the field. Directly before harvest, soil samples (n = 80, 255 mL) by sampled produce were also gathered. In the "after harvest" environment, two types of samples were collected including produce rinse samples and hand rinse samples. Produce rinse samples were collected in the same manner as described previously from tomatoes (n = 23), jalapeño peppers (n = 19) and cantaloupes (n = 38) in 0.1% peptone water. Hand rinse samples (n = 80) represented a composite of hand rinses from three workers after harvesting produce each in 750 mL of 0.1% peptone water.

Microbial Indicator Testing

Samples were analyzed for microbial indicators using the protocols described by Heredia et al. 2016 (56). Within 24 hours of collection from the agricultural environment,

each sample type was sent to UANL for testing for three types of bacterial indicators including general *E. coli*, coliforms and *Enterococcus*. Individual replicate samples were placed in media so that the number of colony-forming units (CFU) could be counted for each bacterial indicator. Using the U.S. Food and Drug Administration (FDA) protocol (57), CFU data in combination with sample volume was used to estimate the measure of indicator growth. Any growth of an indicator on a plate resulted in a sample being marked positive for that indicator.

Statistical Analyses

Initial data processing was completed in R Studio 8.41 (R Studio, Boston, MA) of the GCIIFinalIndicatorData_FixedSourceCrop.csv dataset to subset the data to only include observations from the pre-harvest and harvest environment, remove unnecessary columns for analysis and generate the arithmetic mean $\log_{10}$-transform indicator bacteria concentrations due to highly skewed distributions. A total of 391 samples were used comprised of sample types including irrigation water, produce washes, soil, and hand-rinses. All further statistical analyses were completed using JMP Pro 15 (SAS Institute Inc., Cary, NC).

Descriptive statistics included the percentage of positive samples and the average concentration of indicators stratified by produce type (tomato, jalapeño peppers and cantaloupe), before and after harvest, sample type (produce wash, soil, irrigation water, and hand-rinse), and indicator type (*E. coli*, coliforms, and *Enterococcus*). The concentration of indicators by sample type was represented by a geometric mean and 95 percent confidence interval since the concentrations of $\log_{10}$-transformed indicator bacteria were still non-normal despite the transformation. To compare indicator

contamination amongst similar sample types and identify the contribution of different indicators to the overall contamination level, a principal component analysis was used to generate principal component scores for each sample in the pre-harvest and post-harvest environment. For both time points, only the first component was significant (with an eigenvalue greater than 1). Thus, subsequent statistical significance tests were run only including component 1 scores. A Kruskal-Wallis ANOVA test was used to compare component 1 scores within sample types in the pre-harvest and post-harvest environment due to significant outliers among the component 1 scores. With the criterion for the Kruskal-Wallis ANOVA test met, this test was used to compare the component 1 scores amongst similar sample types (pre-harvest: irrigation water, produce and soil; post-harvest: hand-rinse and produce) stratified by the associated produce (tomato, jalapeño pepper, and cantaloupe). Dunn's nonparametric comparison was used as an ad hoc test to identify significantly different principal component 1 scores amongst the associated produce sample type. To determine the similarities and differences among the microbial ecology of different sample types stratified by the associated produce, discriminant analysis was used to generate score summaries and canonical plots. The score summaries provided information on the percent misclassification of samples; this number indicates how effectively the model was able to identify different sample types (e.g. irrigation water versus soil versus hands associated with jalapeño pepper). The canonical plots provided a visual explanation of the differences and similarities between sample types.

Results

Descriptive Statistics

The four sample types collected were tested for three bacterial indicators that included coliforms (N= 391), *E. coli* (N=391) and *Enterococcus* (N=391). With a total of 391 samples included in this analysis, the majority of the samples were collected in the pre-harvest environment (N=231) compared to the post-harvest environment (N=160). Cantaloupe-associated samples, including produce, soil, irrigation water, and hand-rinse, had the greatest sample size (N=189) followed by tomato-associated samples (N=113) and jalapeño-associated samples (N=89).

Examining the quantity and concentration of indicators

Before we examined the relationships between different bacterial indicators, it was important to assess where bacterial indicators were most commonly found in the pre-harvest and post-harvest environment. To directly compare the presence and concentration of indicators between produce types, sample types, and study time periods, we measured the presence and absence of the three indicators to calculate a proportion of positive samples (Table 1). We found that the majority of all samples tested positive for coliforms (97%) and *Enterococcus* (86%) but not for *E. coli* (28%). This finding was true regardless if the data was stratified by sample type, produce type or sample collection time—the majority of the samples tested positive for coliforms and *Enterococcus* but not for *E. coli*. In conclusion, we determined that coliforms and *Enterococcus* were more prevalent in samples than *E. coli*.

In addition to calculating the proportion of samples positive for different bacterial indicators, we also were interested in comparing the relative concentration of different indicators. To assess the concentration of indicators between produce types, sample types and study time periods, we calculated the arithmetic mean for the $\log_{10}$-transformed

variables (Table 1). We found that in addition to being the most commonly found bacterial indicator, coliforms were the most concentrated indicator overall with the highest concentration in post-harvest hand rinses. *E. coli* was the least concentrated indicator overall with the highest prevalence also found in prevalence in post-harvest hand rinses. Comparing the relative concentration of all three bacterial indicators amongst sample types, pre-harvest irrigation water contained the lowest concentration for all three indicators. The standard deviation for these arithmetic means was relatively small indicating sufficient accuracy in these measurements. In conclusion, coliforms and *Enterococcus* were found in higher concentrations across all sample types.

Assessing relationships between indicators

Beyond examining the presentence and concentration of bacterial indicators, this study sought to learn more about the relationships between bacterial indicators in the agricultural environment. To help us identify which bacterial indicators were the most important in the explaining the variance in contamination in the pre and post-harvest environment, we conducted a principal component analysis (PCA) on the concentrations of the three bacterial indicators stratified by before and after harvest. The goal of the principal component analysis is to reduce the dimensions of correlated variables into a smaller set of newly created variables called principal components that explain most of the variation of the original variables (58). In this application of PCA, there was only one significant newly created variable, called principal component 1 (PC1), that was significant with an assigned eigenvalue greater than 1. Components 2 and 3 were not significant with eigenvalues less than 1 and were not included in subsequent analyses. PC1 could be described as a variable characterizing total contamination by including all three bacterial indicators. This principal component 1 (PC1) explained 75% of the total

variance in the mode in the pre-harvest environment (eigenvalue=2.24) and 65% of the variance in the post-harvest environment (eigenvalue=1.94). In other words, PC1 was accurately able to describe the variance of contamination 65% - 75% of the time. Each eigenvalue is characterized by multiple eigenvectors, ranging from -1 to 1, to determine if the original dimensions work in combined action or contrast to produce that principal component. Here, all three bacterial indicators played a relatively equal role in contributing to component 1 of the PCA with eigenvalues greater than or equal to 0.50 in the pre-harvest environment and post-harvest environment. All three bacterial indicators played an important role in explaining the variance of contamination among all samples in the pre-harvest and post-harvest environment (Table 2).

While the output from the PCA tests contributed to our understanding of the relative variance explained by different indicators overall, this analysis did not allow for comparisons amongst produce stratified by sample types. To identify statistically significantly different principal component scores amongst the same sample type (e.g. irrigation water, produce, soil and hand-rinse samples by jalapeño, tomato and cantaloupe), we conducted a Kruskal-Wallis ANOVA test that allowed for the non-normal distributions of the principal component 1 (PC1) scores (Figure 1). Sample types with statistically significant Kruskal-Wallis ANOVA results were then assesed using non-parametric Wilcoxon comparisons for each pair to identify the pairs with significantly different principal component scores. We found significant differences between produce PC1 scores in the pre-harvest environment between cantaloupe and the other two produce types. Cantaloupe produce samples showed significantly higher PC1 scores and significantly lower soil sample PC1 scores compared to jalapeño and tomato

(Figure 1A). There were no significant differences between PC1 scores for irrigation water samples (Figure 1A). We also found significant differences between produce PC1 scores in the post-harvest environment between cantaloupe and the other two produce types. Both cantaloupe produce samples and hand-rinse samples demonstrated significantly higher PC1 scores compared to the other two produce types (Figure 1B). In conclusion, we showed that the contamination profile of cantaloupe is significantly different than jalapeño and tomato in both the pre-harvest and post-harvest environment.

While principal component analysis and Kruskal-Wallis ANOVA allowed comparison between produce categories stratified by sample type, it did not allow us to assess the similarities and differences in microbial ecology between sample types. To determine if sample types could be distinguished by their microbial ecology, we used discriminant analysis (DA) to generate canonical plots stratified by produce type and time period of the $\log_{10}$-transformed indicator concentrations (Figure 2). The circles displayed in the canonical plots represent the independent variable generated by the DA. Outer circles in the canonical plots represent the 95% confidence interval of the independent variable created by the DA. Cirlces that fail to overlap indicate statistically significant differences between sample types. Irrigation water and soil samples were the most similar in the pre-harvest environment amongst tomato, jalapeño and cantaloupe with overlapping circles (Figure 2A-C). Cantaloupe (Figure 2C), but not tomato and jalapeño (Figure 2A-2B), produce samples were significantly different than irrigation and soil samples. In the post-harvest environment, produce and hand-rinse samples for tomato and jalapeño, were significantly different (Figure 2D-E). The length and directionality of the bacterial indicator vectors in the canonical plots indicate the importance of these

indicators in defining the microbial ecology of particular samples. For tomato and jalapeño samples in the post-harvest environment, *E. coli* and *Enterococcus* were the most important in defining the microbial makeup of the hand-rinse samples (Figure 2D-E). This contrasts with cantaloupe samples in the post-harvest environment that are defined by *E. coli* and coliforms (Figure 2F). Similar to inconsistencies in the DA results from the pre-harvest environment between cantaloupe compared to tomato and jalapeño, there were also anomalies among cantaloupe DA results in the post-harvest environment. With significant overlap between produce and hand-rinse cantaloupe samples, the model poorly distinguished the differences between these two sample types resulting in no significant results (Figure 2F). Additionally, the vector patterns were different in length and directionality for the cantaloupe samples. In conclusion, tomato and jalapeño sample types showed relatively similar bacterial indicator relationships, with homologus canonical structure in 95% confidence intervals overlap and vector magnitude and directionality, in the pre-harvest and post-harvest environment. The canonical structure of cantaloupe samples was different than tomato and jalapeño samples in the pre-harvest and post-harvest environment.

The canonical plots visually highlighted the similarities and differences in the relationships between the bacterial indicators. To further investigate the accuracy of the discriminant analysis model in distinguishing different sample types, we also produced score summaries from the discriminant analyses (Table 3). In the pre-harvest environment, the highest percent misclassification was seen between jalapeño samples (37%), followed by tomato samples (24%) and cantaloupe samples (23%). Likely, the lower percent misclassification score with cantaloupe samples can be attributed to the

significant differences between produce samples versus irrigation water and soil (Figure 2C). In the post harvest environment, the model differentiated tomato (15%) and jalapeño (18%) produce and hand-rinse samples with greater accuracy. The model did worst in accurately classifying the differences between produce and hand-rinse for cantaloupe in the post-harvest environment with 46% of the observations misclassified. A high percent misclassification score in this application suggests similarities between the microbial ecology of cantaloupe produce and hand-rinse samples. The score summaries further confirmed the similarities and differences visually assessed in the canonical plots. In conclusion, this output confirms that the model sometimes performed poorly by failing to identify sample types according to their bacterial indicators.

Discussion

The goal of this study was to evaluate the relationships between foodborne bacterial indicators. Specifically, we investigated whether these relationships were modified by sample type (produce, soil, hands, water) and/or processing stages (before and after harvest) on samples from different produce types from farms in Mexico that supply fresh produce to the United States. Our results showed that each of the bacterial indicators contributed relatively equally to explaining the variance in the PCA models. We also found that the contamination profile of cantaloupe is significantly different than different than jalapeño and tomato in both the pre-harvest and post-harvest. Lastly, the canonical structure of cantaloupe samples was different than tomato and jalapeño samples in the pre-harvest and post-harvest environment potentially indicating differences in bacterial indicator relationships.

Previous studies have characterized relationships between foodborne and epiphytic bacteria on plant models (52, 54). However, these studies were conducted in a laboratory setting and used one plant model. Neither of these studies examined the relationships between different species of indicator bacteria. This study is unique because we collected samples from three produce types (tomato, jalapeño and cantaloupe), from four different sample types (produce, hand rinse, soil sample and irrigation water), and at multiple points in the agricultural process (before and after harvest) from eleven different farms. To summarize, the overall study design of this research is novel in comparison to previous research because of the large sample size and variety of samples types collected.

Contribution of bacterial indicators to the contamination profile

In a principal component analysis (PCA) that included concentrations of all three bacterial indicators, all indicators contributed relatively equally in explaining variance of contamination among all samples in the pre-harvest and post-harvest environment. Each eigenvalue is characterized by multiple eigenvectors, ranging from -1 to 1, to determine if the original dimensions work in combined action or contrast to produce that principal component. It is possible that because all three indicators eigenvectors were positive and of similar magnitude (e.g. working together to produce the contamination profile of PC1) because *E. coli*, *Enterococcus* and coliforms can originate from the same contamination source in feces (59). While similar PCA analyses of foodborne bacterial indicators were not available in the literature, prior analyses of surface water and irrigation water contaminants have found that there are positive relationships between *E. coli* and coliforms in addition to *E. coli* and enterococcus (17, 20, 60). Other literature suggests that different indicator bacteria could originate from different contamination sources (61), so it is also plausible that our PCA results represent equal exposure of samples to a variety of contaminants from a variety of sources. Ongoing research is needed to assess if these indicators and current detection methods are the best methods of determining foodborne pathogen contamination (14, 62).

Contamination modified by sample type and associated produce type

When the principal component analysis results were stratified by sample type and associated produce, there were significant differences between cantaloupe as compared to jalapeño and tomato. The biggest differences in the mean PC1 scores in the Kruskal-Wallis test were seen between produce samples in the pre-harvest and post-harvest environment with significantly higher scores for cantaloupe (Figure 1). While the

literature does not have similar statistical analyses of bacterial indicators to compare to, it is possible that the differences in the contamination profile between produce types is due to physical differences in the produce skin surface between jalapeño and tomato as compared to cantaloupe. In previous studies of epiphytic bacterial species on plants, direct UV radiation and limited nutrients and water limited the growth of bacteria (39). It is possible that the produce surface of cantaloupe provides sufficient and sustainable access to moisture and nutirents over a period of time to support the growth of bacterial indicators due to the rough nature of the rind compared to the smooth surface of jalapeño and tomato. The first quantitative analysis to look at differences in fitness between bacteria compared bacteria naturally found on leaves (*Pseudomonas syringae*) and foodborne pathogens (*E. coli*, *Salmonella*) (63). This study found that foodborne bacterial species only performed as well as plant bacteria during wet and low-light conditions. Additionally, the success of bacterial growth and proliferation on leaves varied greatly between the different produce species included in the study. Bacterial populations were the greatest on uneven trichromatic leaves found on bean, tomato and cucumber plants while populations were the lowest on the waxy leaves of corn, oat and pea (63). This study did not examine the fitness of bacteria on the produce surface, but it is plausible that the uneven and rough surface of the cantaloupe would harbor bacteria similar to the trichromatic leaves included in this study.

Subsequent studies have examined the survival of a variety of foodborne pathogens on the surfaces of produce. In a laboratory-based experiment, Stine et al. 2005 found that all bacterial and viral foodborne pathogens included in the study (except *S. enterica*) persisted the longest on the surface of cantaloupe compared to lettuce and bell

pepper (64). The lenticles on a cantaloupe surface, or the netting found on the rind, have been hypothesized to provide protection for microorganisms. These lenticles provide an adequate surface for microbial attachment (65) and resist disinfection from common antimicrobial treatments (66). In comparison, the waxy surface of vegetables like bell pepper can entrap bacteria due to the complex structure of cuticular wax found on many vegetables (67). These studies support our findings of higher bacterial loads on cantaloupe surfaces in the pre-harvest and post-harvest environment. In comparison, the waxy coating on some vegetables like jalapeños and tomatoes provided a less ideal colonization surface. These results in combination with prior studies suggest that rougher plant surfaces offer greater protection and increase the likelihood of bacterial contamination as compared to waxy plant surfaces.

Relationship between bacterial indicators

Cantaloupe results suggest significantly different indicator bacteria relationships compared to tomato and jalapeño. While the discriminant analysis canonical structure was relatively similar for irrigation water, produce and soil associated with tomato and jalapeño (Figure 2A, 2B), the canonical structure was different for cantaloupe (Figure 2C). Specifically, cantaloupe produce samples DA results were statistically significantly different than cantaloupe associated irrigation water and soil; this result was not seen in jalapeño and tomato DA results. It is likely that the relationships between bacterial indicators were significantly different on the cantaloupe produce samples due to the enhanced protection and attachment sites on the cantaloupe produce surface discussed earlier (65, 66). Prior research on plant surface biofilm tells us that bacteria within biofilms can better withstand changes in the environment by maintaining a more constant

moisture and nutrient level in favorable plant niches (47). Our results support the idea that the relationships between indicator bacteria on cantaloupe could be significantly different due to the relative proliferation of biofilms on the surface of cantaloupe as compared to jalapeños and tomatoes. Due to multiple multi-state salmonella outbreaks connected to cantaloupe produce since the 1990's, multiple studies have been conducted to learn more about the interaction between foodborne pathogens and cantaloupe surfaces. Many of these studies focused efficacy of different sanitation methods to kill foodborne pathogens on cantaloupe rinds (68-70). These studies found proliferation of *Salmonella* spp. in addition to and *Listeria monocytogenes* V7 and *E. coli* 0157:H7 in the biofilms of cantaloupes after lab-based inoculation. These foodborne pathogens inside biofilms were subsequently resistant to common agricultural antimicrobials such as lauroyl arginate ethyl (69) and sodium hypochlorite (70). None of these studies included bacterial indicators such as *Enterococcus* or coliforms in their methods, but it is likely these bacteria would find protection in the biofilms of the fibrous cantaloupe rind similar to previously tested foodborne pathogens.

Unsurprisingly, cantaloupe indicator bacteria relationships were also different in the post-harvest environment with great similarities between produce and hand-rinse samples (Figure 2F). This similarity stands in contrast to tomato and jalapeño samples that showed significant differences between produce and hand-rinse samples (Figure 2D, 2E). One lab-based experiment that tracked the rate of colonization of two strains of *Salmonella* on a cantaloupe rind postinoculation. This study found that initial *Salmonella* biofilm formation occurred after two hours and *Salmonella* embedding in the surface occurred after 24 hours regardless of temperature (68). It is possible that the relationships

between indicator bacteria on produce and hands were similar due to rapid colonization and proliferation of bacterial indicators passed from hands to produce or vice versa. Prior experimental studies support the hypothesis that contaminants can be readily spread between produce and agricultural worker hands (34). Transmission to hands was more probable with cantaloupe samples due to bacterial biofilm proliferation resulting in similarities in the relationships between indicators on the produce and hands.

Strengths and limitations

This study is a novel contribution to the literature because it assessed the presence and concentration of three bacterial indicators including *E. coli*, *Enterococcus*, and coliforms: 1) at different time points in the agricultural process before and after harvest, 2) on different sample types including irrigation water, produce, soil and hand rinse, 3) with different associated produce types including tomato, jalapeño, and cantaloupe. Unlike the vast majority of other studies, data for this study was collected at farms and not in a laboratory setting. With many studies characterizing the behavior and fitness of bacteria in controlled laboratory studies, in-field studies are a critically important for verifying real world application of lab-generated predictions (51). Novel statistical methods were used to evaluate the contamination profile and relationships between bacterial indicator species using principal component analysis and discriminant analysis. These analyses provide more nuance than tracking bacterial concentrations over time that is typically seen in the literature with similar studies (54). There are limitations with this dataset and analyses performed. Only three bacterial indicators were included due to sample size restrictionsions. Samples were collected between 2011 and 2012 and not standardized to collection in one seaon; this may have impacted study results with

evidence that seasonality plays an important role in bacterial indicator abundance in environmental samples (71) Herbs, root vegetables and berries were not included in this study. Thus, the external validity of our results is limited. A greater variety of produce types with different physical characteristics and growing environments would be recommended for future study.

Implications

- Similarities in the physical characteristics of produce may create ecological niches on plant surfaces that have similar microbial ecologies. While waxy produce surfaces might be less ideal for colonization, fibrous and rough produce surfaces like cantaloupe rind create ideal conditions for microbial colonization growth in biofilms. Prior studies indicate that these biofilms are resistant to traditional sterilization methods with antimicrobials suggesting different produce types may require different sanitation strategies.
- Significant cantaloupe contamination in the pre-harvest environment carried over to the post-harvest environment with similar contamination found on farmworker hands. This provides evidence that contamination of the physical produce item is carried through the agricultural environment as a pathway for contaminating farmworker hands and vice versa.

References

1. Sivapalasingam S, Friedman CR, Cohen L, et al. Fresh produce: a growing cause of outbreaks of foodborne illness in the United States, 1973 through 1997. *Journal of Food Protection* 2004;67(10):2342-53.
2. Tauxe R, Kruse H, Hedberg C, et al. Microbial hazards and emerging issues associated with produce a preliminary report to the national advisory committee on microbiologic criteria for foods. *Journal of Food Protection* 1997;60(11):1400-8.
3. Scallan E, Hoekstra RM, Angulo FJ, et al. Foodborne illness acquired in the United States—major pathogens. *Emerging Infectious Diseases* 2011;17(1):7.
4. Painter JA, Hoekstra RM, Ayers T, et al. Attribution of foodborne illnesses, hospitalizations, and deaths to food commodities by using outbreak data, United States, 1998–2008. *Emerging Infectious Diseases* 2013;19(3):407.
5. Pollack SL. Consumer demand for fruit and vegetables: the US example. *Changing Structure of Global Food Consumption and Trade* 2001;6:49-54.
6. Dewey-Mattia D, Manikonda K, Hall AJ, et al. Surveillance for foodborne disease outbreaks—United States, 2009–2015. *MMWR Surveillance Summaries* 2018;67(10):1.
7. Darwin KH, Miller VL. Molecular Basis of the Interaction of *Salmonella* with the Intestinal Mucosa. *Clinical Microbiology Reviews* 1999;12(3):405-28.
8. Hook E. *Salmonella* species (including typhoid fever). *Principles and Practice of Infectious Diseases* 1990;1700-13.

9. H.F. Gomez GGC. *Textbook of Pediatric Infectious Disease: Salmonella*. 4 ed. Philadelphia, PA: The W.B. Saunders Co.; 1998:1331-4.

10. *Control of Communicable Diseases Manual*. 19 ed. Washington, DC: American Public Health Association; 2008:164-66.

11. Juneja VK, Sofos JN. *Pathogens and Toxins in Foods: Challenges and Interventions*. ASM Press; 2010:108-118.

12. Nataro JP, Kaper JB. Diarrheagenic *escherichia coli*. *Clinical Microbiology Reviews* 1998;11(1):142-201.

13. Melton-Celsa AR. Shiga toxin (Stx) classification, structure, and function. *Microbiology Spectrum* 2014;2(2).

14. Savichtcheva O, Okabe S. Alternative indicators of fecal pollution: relations with pathogens and conventional indicators, current methodologies for direct pathogen monitoring and future application perspectives. *Water Research* 2006;40(13):2463-76.

15. Mossel D. Index and indicator organisms-a current assessment of their usefulness and significance. *Food Technology in Australia* 1978.

16. Baggi F, Demarta A, Peduzzi R. Persistence of viral pathogens and bacteriophages during sewage treatment: lack of correlation with indicator bacteria. *Research in Microbiology* 2001;152(8):743-51.

17. Holvoet K, Sampers I, Seynnaeve M, et al. Relationships among hygiene indicators and enteric pathogens in irrigation water, soil and lettuce and the impact of climatic conditions on contamination in the lettuce primary production. *International Journal of Food Microbiology* 2014;171:21-31.

18. Silva AM, Vieira H, Martins N, et al. Viral and bacterial contamination in recreational waters: a case study in the Lisbon bay area. *Journal of Applied Microbiology* 2010;108(3):1023-31.

19. Tian P, Yang D, Shan L, et al. Concurrent Detection of Human Norovirus and Bacterial Pathogens in Water Samples from an Agricultural Region in Central California Coast. *Frontiers in Microbiology* 2017;8.

20. Wilkes G, Edge T, Gannon V, et al. Seasonal relationships among indicator bacteria, pathogenic bacteria, *Cryptosporidium* oocysts, *Giardia* cysts, and hydrological indices for surface waters within an agricultural landscape. *Water Research* 2009;43(8):2209-23.

21. Lynch MF, Tauxe RV, Hedberg CW. The growing burden of foodborne outbreaks due to contaminated fresh produce: risks and opportunities. *Epidemiology and Infection* 2009;137(3):307-15.

22. Ailes EC, Leon JS, Jaykus L-A, et al. Microbial concentrations on fresh produce are affected by postharvest processing, importation, and season. *Journal of Food Protection* 2008;71(12):2389-97.

23. Barak JD, Schroeder BK. Interrelationships of food safety and plant pathology: the life cycle of human pathogens on plants. *Annual Review of Phytopathology* 2012;50:241-66.

24. Pagadala S, Marine SC, Micallef SA, et al. Assessment of region, farming system, irrigation source and sampling time as food safety risk factors for tomatoes. *International Journal of Food Microbiology* 2015;196:98-108.

25. Berry ED, Wells JE, Bono JL, et al. Effect of proximity to a cattle feedlot on Escherichia coli O157: H7 contamination of leafy greens and evaluation of the potential for airborne transmission. *Applied and Environmental Microbiology* 2015;81(3):1101-10.

26. Jacobsen CS, Bech TB. Soil survival of *Salmonella* and transfer to freshwater and fresh produce. *Food Research International* 2012;45(2):557-66.

27. Todd EC, Greig JD, Bartleson CA, et al. Outbreaks where food workers have been implicated in the spread of foodborne disease. Part 5. Sources of contamination and pathogen excretion from infected persons. *Journal of Food Protection* 2008;71(12):2582-95.

28. Johnston LM, Moe CL, Moll D, et al. The Epidemiology of Produce-Associated Outbreaks of Foodborne Disease. *Microbial Hazard Identification in Fresh Fruit and Vegetables* 2006:37-72.

29. Cooley MB, Miller WG, Mandrell RE. Colonization of *Arabidopsis thaliana* with *Salmonella enterica* and enterohemorrhagic *Escherichia coli* O157: H7 and competition by *Enterobacter asburiae*. *Applied and Environmental Microbiology* 2003;69(8):4915-26.

30. Islam M, Morgan J, Doyle MP, et al. Persistence of *Salmonella enterica* serovar Typhimurium on lettuce and parsley and in soils on which they were grown in fields treated with contaminated manure composts or irrigation water. *Foodborne Pathogens and Disease* 2004;1(1):27-35.

31. Nicholson FA, Groves SJ, Chambers BJ. Pathogen survival during livestock manure storage and following land application. *Bioresource Technology* 2005;96(2):135-43.

32. Chalmers R, Aird H, Bolton F. Waterborne *Escherichia coli* O157. *Journal of Applied Microbiology* 2000;88(S1):124S-32S.

33. Steele M, Odumeru J. Irrigation water as source of foodborne pathogens on fruit and vegetables. *Journal of Food Protection* 2004;67(12):2839-49.

34. Bartz FE, Lickness JS, Heredia N, et al. Contamination of fresh produce by microbial indicators on farms and in packing facilities: elucidation of environmental routes. *Applied and Environmental Microbiology* 2017;83(11):e02984-16.

35. León-Félix J, Martínez-Bustillos RA, Báez-Sañudo M, et al. Norovirus contamination of bell pepper from handling during harvesting and packing. *Food and Environmental Virology* 2010;2(4):211-7.

36. Harris LJ, Farber JN, Beuchat LR, et al. Outbreaks Associated with Fresh Produce: Incidence, Growth, and Survival of Pathogens in Fresh and Fresh-Cut Produce. *Comprehensive Reviews in Food Science and Food Safety* 2003;2(s1):78-141.

37. Janisiewicz W, Conway W, Brown M, et al. Fate of *Escherichia coli* O157: H7 on fresh-cut apple tissue and its potential for transmission by fruit flies. *Applied and Environmental Microbiology* 1999;65(1):1-5.

38. Leverentz B, Conway WS, Alavidze Z, et al. Examination of bacteriophage as a biocontrol method for *Salmonella* on fresh-cut fruit: a model study. *Journal of Food Protection* 2001;64(8):1116-21.

39. Wilson M, Lindow SE. Coexistence among epiphytic bacterial populations mediated through nutritional resource partitioning. *Applied and Environmental Microbiology* 1994;60(12):4468-77.

40. Leveau JH, Lindow SE. Appetite of an epiphyte: quantitative monitoring of bacterial sugar consumption in the phyllosphere. *Proceedings of the National Academy of Sciences* 2001;98(6):3446-53.

41. Miller WG, Brandl MT, Quiñones B, et al. Biological sensor for sucrose availability: relative sensitivities of various reporter genes. *Applied and Environmental Microbiology* 2001;67(3):1308-17.

42. Jacobs J, Carroll T, Sundin G. The role of pigmentation, ultraviolet radiation tolerance, and leaf colonization strategies in the epiphytic survival of phyllosphere bacteria. *Microbial Ecology* 2005;49(1):104-13.

43. Jeter C, Matthysse AG. Characterization of the binding of diarrheagenic strains of *E. coli* to plant surfaces and the role of curli in the interaction of the bacteria with alfalfa sprouts. *Molecular Plant-Microbe Interactions* 2005;18(11):1235-42.

44. Torres AG, Jeter C, Langley W, et al. Differential binding of *Escherichia coli* O157: H7 to alfalfa, human epithelial cells, and plastic is mediated by a variety of surface structures. *Applied and Environmental Microbiology* 2005;71(12):8008-15.

45. Sukupolvi S, Lorenz RG, Gordon JI, et al. Expression of thin aggregative fimbriae promotes interaction of *Salmonella typhimurium* SR-11 with mouse small intestinal epithelial cells. *Infection and Immunity* 1997;65(12):5320-5.

46. Uhlich GA, Keen JE, Elder RO. Variations in the csgD promoter of *Escherichia coli* O157: H7 associated with increased virulence in mice and increased invasion of HEp-2 cells. *Infection and Immunity* 2002;70(1):395-9.

47. Morris CE, Monier J-M, Jacques M-A. A technique to quantify the population size and composition of the biofilm component in communities of bacteria in the phyllosphere. *Applied and Environmental Microbiology* 1998;64(12):4789-95.

48. Monier J-M, Lindow SE. Aggregates of Resident Bacteria Facilitate Survival of Immigrant Bacteria on Leaf Surfaces. *Microbial Ecology* 2005;49(3):343-52.

49. Morris CE, Monier J-M. The ecological sigificance of biofilm formation plant-associated bacteria. *Annual Review of Phytopathology* 2003;41(1):429-53.

50. Fett WF. Naturally occurring biofilms on alfalfa and other types of sprouts. *Journal of Food Protection* 2000;63(5):625-32.

51. Hibbing ME, Fuqua C, Parsek MR, et al. Bacterial competition: surviving and thriving in the microbial jungle. *Nature Reviews Microbiology* 2009;8:15.

52. Cooley MB, Chao D, Mandrell RE. *Escherichia coli* O157: H7 survival and growth on lettuce is altered by the presence of epiphytic bacteria. *Journal of Food Protection* 2006;69(10):2329-35.

53. Barak J, Whitehand L, Charkowski A. Differences in attachment of *Salmonella enterica* serovars and *Escherichia coli* O157: H7 to alfalfa sprouts. *Applied and Environmental Microbiology* 2002;68(10):4758-63.

54. Brandl MT, Mandrell RE. Fitness of *Salmonella enterica* serovar Thompson in the Cilantro Phyllosphere. *Applied and Environmental Microbiology* 2002;68(7):3614-21.

55. Wells JM, Butterfield JE. *Salmonella* contamination associated with bacterial soft rot of fresh fruits and vegetables in the marketplace. *Plant Disease* 1997;81(8):867-72.

56. Heredia N, Caballero C, Cardenas C, et al. Microbial indicator profiling of fresh produce and environmental samples from farms and packing facilities in northern Mexico. *Journal of Food Protection* 2016;79(7):1197-209.

57. Maturin L, Peeler J. BAM aerobic plate count. *Bacteriological Analytical Manual US Food and Drug Administration* 2001.

58. Jolliffe I. Principal component analysis. *Technometrics* 2003;45(3):276.

59. Scott TM, Rose JB, Jenkins TM, et al. Microbial source tracking: current methodology and future directions. *Applied and Environmental Microbiology* 2002;68(12):5796-803.

60. Draper AD, Doores S, Gourama H, et al. Microbial survey of Pennsylvania surface water used for irrigating produce crops. *Journal of Food Protection* 2016;79(6):902-12.

61. Boehm AB, Sassoubre LM. *Enterococci* as indicators of environmental fecal contamination. *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*: Massachusetts Eye and Ear Infirmary, 2014.

62. Desmarais TR, Solo-Gabriele HM, Palmer CJ. Influence of soil on fecal indicator organisms in a tidally influenced subtropical environment. *Applied and Environmental Microbiology* 2002;68(3):1165-72.

63. O'Brien R, Lindow S. Effect of plant species and environmental conditions on epiphytic population sizes of *Pseudomonas syringae* and other bacteria. *Phytopathology* 1989;79(5):619-27.

64. Stine SW, Song I, Choi CY, et al. Effect of relative humidity on preharvest survival of bacterial and viral pathogens on the surface of cantaloupe, lettuce, and bell peppers. *Journal of Food Protection* 2005;68(7):1352-8.

65. Ukuku DO, Fett WF. Relationship of cell surface charge and hydrophobicity to strength of attachment of bacteria to cantaloupe rind. *Journal of Food Protection* 2002;65(7):1093-9.

66. Gagliardi JV, Karns JS. Leaching of *Escherichia coli* O157:H7 in Diverse Soils under Various Agricultural Management Practices. *Applied and Environmental Microbiology* 2000;66(3):877-83.

67. Agrios G. *Plant Pathology*. San Diego, CA: Academic Press; 1997:66-73.

68. Annous BA, Solomon EB, Cooke PH, et al. Biofilm formation by *Salmonella* spp. on cantaloupe melons. *Journal of Food Safety* 2005;25(4):276-87.

69. Fu Y, Deering AJ, Bhunia AK, et al. Pathogen biofilm formation on cantaloupe surface and its impact on the antibacterial effect of lauroyl arginate ethyl. *Food Microbiology* 2017;64:139-44.

70. Fu Y, Deering AJ, Bhunia AK, et al. Biofilm of *Escherichia coli* O157: H7 on cantaloupe surface is resistant to lauroyl arginate ethyl and sodium hypochlorite. *International Journal of Food Microbiology* 2017;260:11-6.

71. Wilkes G, Brassard J, Edge TA, et al. Bacteria, viruses, and parasites in an intermittent stream protected from and exposed to pasturing cattle: Prevalence, densities, and quantitative microbial risk assessment. *Water Research* 2013;47(16):6244-57.

Figures and Tables

Table 1. Summary statistics

		Coliforms		***E. coli***		***Enterococcus***		**Mean microbial concentration (± SD)**		
	Sample unit	No. of samples	No. (%) positive	No. of samples	No. (%) positive	No. of samples	No. (%) positive	Coliforms (log_{10} CFU)	*E. coli* (log_{10} CFU)	*Enterococcus* (log_{10} CFU)
Tomato before harvest										
Irrigation Water	cfu/100 mL	21	21 (100%)	21	14 (67%)	21	19 (90%)	1.57 (± 0.78)	0.24 (± 0.86)	0.41 (± 0.71)
Produce	cfu/fruit	23	22 (96%)	23	3 (13%)	23	18 (78%)	4.97 (± 1.73)	0.14 (± 1.19)	3.35 (± 1.67)
Soil	cfu/gram	23	21 (91%)	23	3 (13%)	23	14 (61%)	2.37 (± 0.86)	0.48 (± 0.90)	1.50 (± 0.75)
Tomato after harvest										
Hand-rinse	cfu/hand	23	22 (96%)	23	5 (22%)	23	23 (100%)	5.69 (± 1.88)	1.70 (± 1.54)	6.29 (± 1.43)
Produce	cfu/fruit	23	23 (100%)	23	1 (4%)	23	19 (83%)	5.06 (± 1.77)	0.07 (± 1.23)	3.79 (± 1.59)
Jalapeño before harvest										
Irrigation Water	cfu/100 mL	13	11 (85%)	13	3 (23%)	13	12 (92%)	1.53 (± 1.34)	-0.47 (± 0.55)	0.65 (± 1.07)
Produce	cfu/fruit	19	18 (95%)	19	3 (16%)	19	14 (74%)	4.21 (± 2.45)	0.12 (± 1.18)	3.42 (± 1.83)
Soil	cfu/gram	19	18 (95%)	19	1 (5%)	19	10 (53%)	2.73 (± 1.22)	0.13 (± 1.05)	1.39 (± 0.56)
Jalapeño after harvest										
Hand-rinse	cfu/hand	19	18 (95%)	19	8 (42%)	19	19 (100%)	5.00 (± 2.36)	2.08 (± 1.56)	5.73 (± 1.42)
Produce	cfu/fruit	19	18 (95%)	19	1 (5%)	19	16 (84%)	3.92 (± 2.88)	0.04 (± 0.98)	3.80 (± 2.00)
Cantaloupe before harvest										
Irrigation Water	cfu/100 mL	38	35 (92%)	38	12 (32%)	38	29 (76%)	1.74 (± 1.51)	-0.17 (± 0.93)	0.53 (± 1.49)
Produce	cfu/fruit	37	37 (100%)	37	15 (41%)	37	37 (100%)	6.51 (± 1.01)	2.34 (± 1.86)	7.16 (± 1.53)
Soil	cfu/gram	38	38 (100%)	38	13 (34%)	38	31 (82%)	2.51 (± 1.21)	-0.58 (± 1.07)	1.30 (± 1.23)
Cantaloupe after harvest										
Hand-rinse	cfu/hand	38	38 (100%)	38	16 (42%)	38	38 (100%)	6.46 (± 1.51)	2.66 (± 1.89)	7.04 (± 1.71)
Produce	cfu/fruit	38	38 (100%)	38	11 (29%)	38	38 (100%)	6.18 (± 1.24)	2.23 (± 1.89)	7.11 (± 1.47)

Table 2. Principal component results

Before harvest

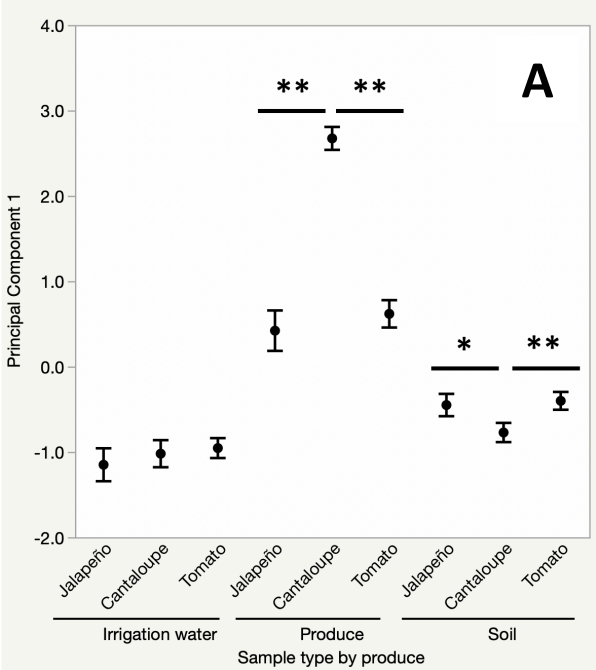
Component	Eigenvalue	Percent of variance explained	Coliforms (log_{10}) eigenvector	*E. coli* (log_{10}) eigenvector	*Enterococcus* (log_{10}) eigenvector
1	2.24	74.64	0.60	0.51	0.62
2	0.58	19.19	-0.42	0.86	-0.29
3	0.19	6.17	0.68	0.08	-0.73

After harvest

Component	Eigenvalue	Percent of variance explained	Coliforms (log_{10}) eigenvector	*E. coli* (log_{10}) eigenvector	*Enterococcus* (log_{10}) eigenvector
1	1.94	64.55	0.59	0.50	0.63
2	0.70	23.43	-0.50	0.84	-0.20
3	0.36	12.02	0.63	0.20	-0.75

Figure 1. Significantly different produce and hand-rinse cantaloupe PC1 scores

This figure displays mean principal component 1 (PC1) scores and their standard error bars. These PC1 scores are stratified by sample types (irrigation water, produce, soil, and handrinse) and subdivided by associated produce type (jalapeño, cantalouple, and tomato). The significantly different pairs (inditcated by a horizontal line) are highlighted with an asterisk (*), with one asterick representing a p-value < 0.05, two asterisks representing a p-value < 0.005, and three asterisks representing a p-value < 0.001. Figure 1A represents PCA analysis in the “before harvest” environment, and Figure 1B represents PCA analysis in the “after harvest” environment. Note that different sample types were analyzed in the “before harvest” (irrigation water, produce, and soil) versus “after harvest” (hand-rinse and produce) environment.


A
4.0
3.0
2.0
1.0
0.0
-1.0
-2.0
Principal Component 1
**
**
*
**
Jalapeño
Cantaloupe
Tomato
Jalapeño
Cantaloupe
Tomato
Jalapeño
Cantaloupe
Tomato
Irrigation water
Produce
Soil
Sample type by produce

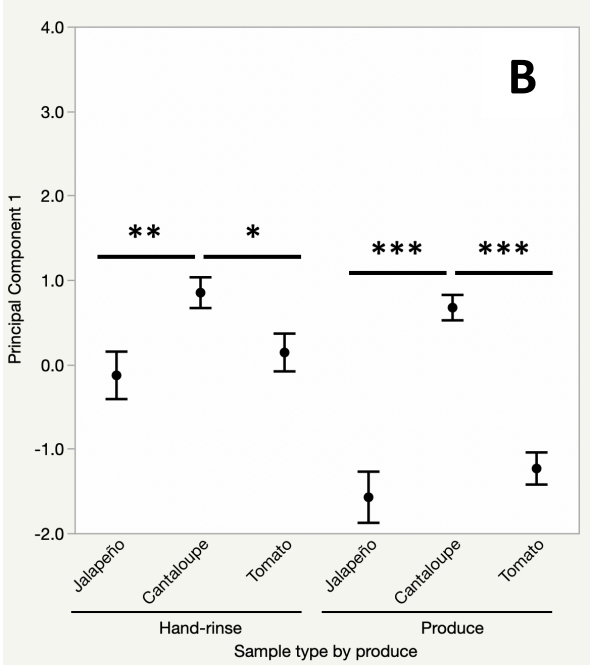
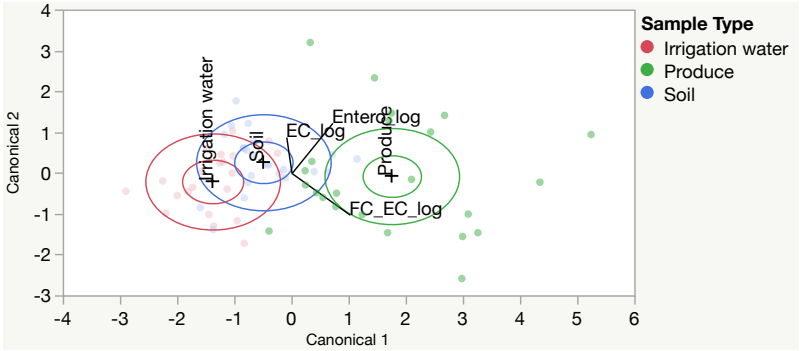

B
4.0
3.0
2.0
1.0
0.0
-1.0
-2.0
Principal Component 1
**
*
***
***
Jalapeño
Cantaloupe
Tomato
Jalapeño
Cantaloupe
Tomato
Hand-rinse
Produce
Sample type by produce

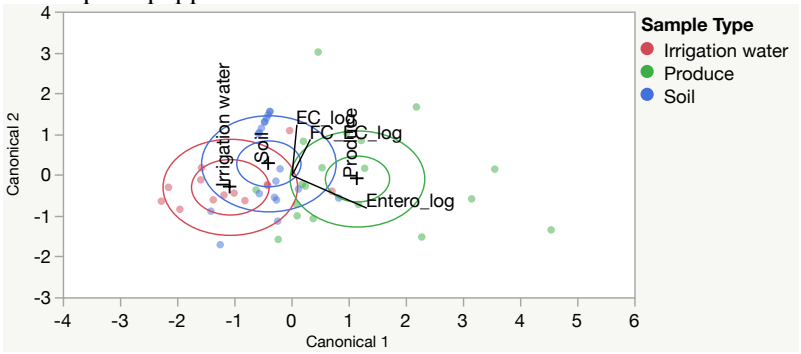
Figure 2. Discriminant analysis plots by produce type and time period

This figure presents the canconical plots from the discriminant analysis of the log_{10}-transformed indicator concentrations. Plots are stratified by associated produce type horizontally (tomato, jalapeño, and cantaloupe) and by sampling time vertically (before and after harvest). These plots facilitate visualization of the differences in microbial ecology between sample types. Statistically significant differences in mircrobial ecology are indicated when sample type circles do not overlap (with outer circles representing the 95% confidence interval and inner circles representing the area where 50% of the observations are found). The three vectors per figure portray the dependent variables (bacterial indicator concentrations) in the model; the length of the vector and directionality indicated the importance of that bacterial indicator in defining the microbial ecology of that particular sample type. Dots represent individual sample types according to their canonical 1 and 2 scores from the discriminant analysis.

A: Tomato “before harvest”

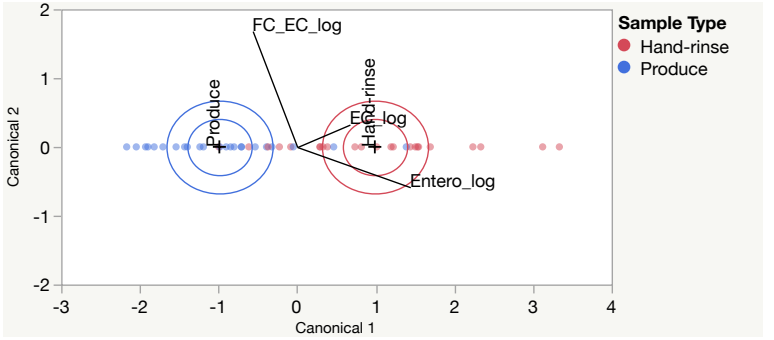


B: Jalapeño pepper “before harvest”

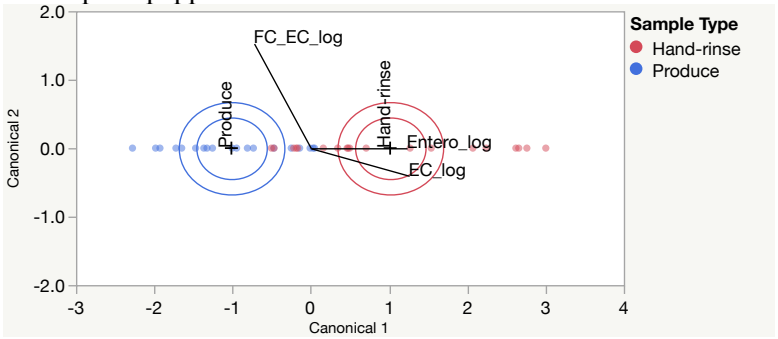


C: Cantaloupe “before harvest”

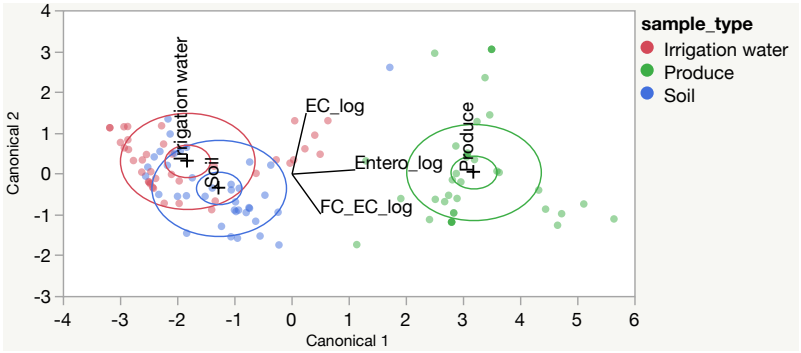
D: Tomato “after harvest”



E: Jalapeño pepper “after harvest”



F: Cantaloupe “after harvest”



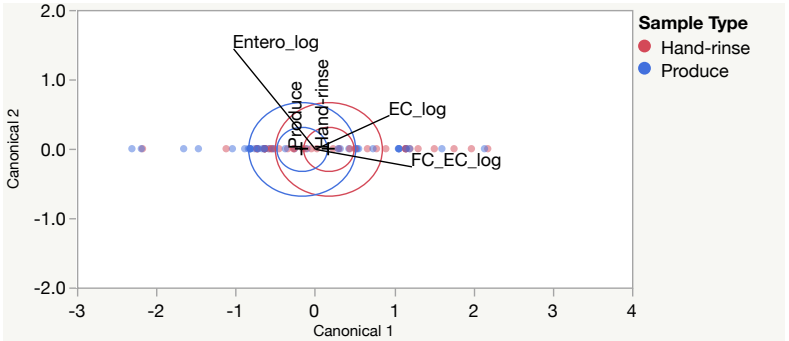


Table 3. Discriminant analysis score summaries

	Count of samples	Count misclassified	Percent misclassified
"Before Harvest"			
Tomato	67	16	23.88
Jalapeño	51	19	37.25
Cantaloupe	113	26	23.01
"After Harvest"			
Tomato	46	7	15.22
Jalapeño	38	7	18.42
Cantaloupe	76	35	46.05

Appendix A: IRB Approval

TO: Juan Leon, PhD
Principal Investigator
*SPH: Global Health

DATE: July 10, 2018

RE: **Continuing Review Expedited Approval**
CR8_IRB00035460

IRB00035460
Identification and Control of Microbiological Hazards in Imported Fresh Fruits and Vegetables: A Field Epidemiological and Intervention Study in Northern Mexico

Thank you for submitting a renewal application for this protocol. The Emory IRB reviewed it by the expedited process on **07/09/2018**, per 45 CFR 46.110, the Federal Register expeditable category [F7], and/or 21 CFR 56.110. This reapproval is effective from **07/09/2018** through **07/08/2019**. Thereafter, continuation of human subjects research activities requires the submission of another renewal application, which must be reviewed and approved by the IRB prior to the expiration date noted above. Please note carefully the following items with respect to this reapproval:

- Consent Documents
  - consentimiento_enjuaguemanos_11.22.2017_CLEAN.docx
  - Informacion-Encuesta Manipulador 23 MAR 2011.doc
  - Informacion-Encuesta-Productor-Manager 23 MAR 2011.docx
  - Oral Script for Written Consent_FarmManagerSurvey_Spanish_4.26.2011.doc
  - Oral Script for Written Consent_FarmManagerSurvey_ver4.26.2011_CLEAN.doc
  - OralScript_Hand Rinsing_ver11.22.2017_CLEAN.docx
- Protocol Document
  - CGProtocol_11.22.17CLEAN.docx

Any reportable events (e.g., unanticipated problems involving risk to subjects or others, noncompliance, breaches of confidentiality, HIPAA violations, protocol deviations) must be reported to the IRB according to our Policies & Procedures at www.irb.emory.edu, immediately, promptly, or periodically. Be sure to check the reporting guidance and contact us if you have questions. Terms and conditions of sponsors, if any, also apply to reporting.

Before implementing any change to this protocol (including but not limited to sample size, informed consent, and study design), you must submit an amendment request and secure IRB approval.

In future correspondence about this matter, please refer to the IRB file ID, name of the Principal Investigator, and study title. Thank you.

Sincerely,

Heather Yates
Analyst Assistant
This letter has been digitally signed

CC: Prince-Guerra Jessica *SPH: Global Health