CPS 2015 RFP
FINAL PROJECT REPORT

Project Title
Identification of novel indicator organisms to determine the risks of fecal contamination of irrigation waters

Project Period
January 1, 2016 – December 31, 2017 (extended to January 31, 2018)

Principal Investigator
Kelly R. Bright
The University of Arizona
Dept. of Soil, Water and Environmental Science
Tucson, AZ 85721
T: 520-621-6910
E: bright@email.arizona.edu

Co-Principal Investigator
Luisa A. Ikner
The University of Arizona
Dept. of Soil, Water and Environmental Science
Tucson, AZ 85721
T: 520-971-4163
E: Ikner@email.arizona.edu

Objectives

1. *To examine irrigation water samples to determine the levels of indicator (i.e., generic E. coli) and pathogenic (i.e., STEC and Salmonella) species by existing cultural and molecular methods.*
2. *To examine irrigation water samples to determine the levels of viruses known to be found in high concentrations in fecally-contaminated waters (i.e., pepper mild mottle virus, Aichiviruses, enteroviruses).*
3. *To determine the composition (presence and relative abundance) of the entire prokaryotic (bacterial) and eukaryotic (e.g., protozoan, fungal, algal) microbial communities found in irrigation water samples.*
4. *To identify groups of organisms or specific species whose presence correlate well (both presence/absence and relative abundance) with the occurrence of foodborne pathogens in irrigation waters.*

Funding for this project provided by the Center for Produce Safety through:
CDFA SCBGP grant# SCB15059

FINAL REPORT

Abstract

The methods used to detect *E. coli* were developed for drinking water and are known to produce high levels of false-positive and false-negative results when used for irrigation waters. Therefore, growers are required to make decisions about water quality/safety based on inaccurate tests. In the current project, the entire microbial compositions (the microbiome) of irrigation water samples from two different fresh produce growing regions in Arizona over a two-year period were determined using next-generation sequencing (NGS) microbiome analysis. Samples were collected from areas with historically higher and lower levels of contamination based on the results of previous studies. In addition, samples were collected from canals with various characteristics (e.g., main canals and lateral canals, cement-lined and non-lined canals, urban and rural canals, primary irrigation canals and return flow canals). Samples were also collected both during the winter leafy green growing season and during the summer to assess if there were any seasonal influences on the microbial water quality.

The microbial communities between samples that were positive and negative for the presence of pathogens were compared and analyzed to identify groups of organisms or specific species whose presence correlated well (both presence/absence and relative abundance) with the occurrence of foodborne pathogens. The presence of pathogens or shifts in the microbial communities were also compared to other water quality characteristics and potential impacts (e.g., rainfall, human and animal influences). The preliminary results of this work suggest that there are both significant regional and seasonal differences in the microbial populations found in irrigation waters. It is therefore important that the “normal” composition or “baseline” level of microorganisms is determined for each region before undertaking any NGS evaluation of the effects of various contamination inputs. Nevertheless, once such baseline evaluations have been conducted, NGS can provide valuable information on any changes such inputs have on the microbial water quality. For instance, when samples were analyzed from areas with known impacts, such as those with historically higher levels of fecal contamination or from return flow (drainage) canals that collect run-off from agricultural fields, shifts in the predominant microbial populations were observed.

This work should result in an approach that will provide region- and/or season-specific identification of species or microbial population shifts, which may be used to identify changes in the microbial quality of specific irrigation water systems. This information could allow members of the fresh produce industry to make risk-based assessments of water quality and help to determine when it is safe to irrigate fields.

Background

Bacterial indicators of fecal contamination, such as generic *Escherichia coli*, appear to be inadequate for estimating the occurrence and levels of microbial pathogens in irrigation waters. The standard methods used by the produce industry to detect *E. coli* in irrigation waters utilize techniques developed for drinking water, which have not undergone extensive scientific scrutiny; as such, these methods may not be wholly appropriate for determining the quality of irrigation waters. For instance, using standard methods, such as the Colilert® MPN test (which is used by the produce industry to determine irrigation water quality), *E. coli* may be found in ≥87.6% of samples. Nevertheless, based on the results of another study (Rock CM, Bright KR, Gerba CP, 2012 CPS project funded through CDFA SCBGP grant #SCB11069), standard methods such as the Colilert® MPN and the MI agar and m-ColiBlue24® membrane filtration

methods have high false-positive rates (from 51% to 71%). False-positive results have large implications of over-estimating the actual concentration of *E. coli* in a water sample. This could lead to costly interventions such as non-use of water, delayed irrigation, and product rejection.

In addition, the Food Safety Modernization Act (FSMA) Produce Safety Rule standards of ≤126 *E. coli* per 100 ml for a geometric mean of 20 samples (a rolling dataset including the most recent 20 samples collected) or ≤410 *E. coli* per 100 ml as the statistical threshold value (STV) (90% of samples much fall below this level) have been used as arbitrary thresholds of risk for pathogen levels in irrigation water, for which there is little scientific basis. Based on our previous sampling data, *E. coli* was detected at numbers above this threshold in only 2.6% of samples, whereas *E. coli* and *Salmonella* species could be detected via other cultural methods with greater frequencies (8.1% and 12.4%, respectively) (Bright, unpublished). There is therefore a need to evaluate this threshold to see if it can be correlated with the presence of foodborne pathogens. It is also important to identify novel indicator organisms that are present in water when microbial pathogens are present or to detect a shift in the microbial composition for their use as potential future methods to replace the use of generic *E. coli*.

Our project goal was to improve the safety of irrigated food crops by identifying more appropriate indicator organisms or groups of organisms to determine the risks of contamination of crops by foodborne pathogens from irrigation water. Our hypothesis was that next-generation DNA sequencing (NGS) could be used as an inexpensive and powerful method to identify the composition (presence and relative abundance) of the entire prokaryotic (bacterial) and eukaryotic (e.g., protozoan) microbial communities found in irrigation waters. Such information could then be used to correlate the presence of foodborne pathogens with either shifts in the community makeup (from that of un-impacted, non-contaminated waters) or with the presence of specific species (i.e., potential indicator organisms). Such high-throughput sequencing parallelizes the sequencing process, producing thousands or millions of sequences at once, which reduces the costs and the time required. Two such NGS platforms are the Illumina MiSeq and HiSeq platforms, which allow for up to millions of reads of 250 base-pair (bp) sequences. These platforms allow for bi-directional sequencing, providing greater accuracy.

Research Methods and Results

During the current project, 99 irrigation water samples were collected in triplicate (for a total of 297 samples) from canals with various characteristics (e.g., main canals and lateral canals, cement-lined and non-lined canals, urban and rural canals) from two different fresh produce growing regions in Arizona. The growing season for leafy greens in these regions takes place from approximately October through early April each year. Half of the irrigation water samples were collected during the growing season and half during the summer, over two years. Data collected from previous studies regarding the occurrence of fecal contamination and the presence of pathogens in the irrigation water canals in these regions were studied to ensure that areas with historically low contamination and higher contamination levels were included. As a result of this examination of the historical data, samples from a second region were also collected as part of this study though not initially included in the original proposal; irrigation waters from this second region are typically of lower microbial quality than irrigation waters in the originally proposed region.

Each triplicate “sample” consisted of the following grab samples: a 3-liter sample, a 2-liter sample, and a 1.1-liter sample. Water quality parameters, such as the temperature, pH, electro-conductivity, and the amount of total dissolved solids, were recorded in the field. Other environmental parameters, such as the air temperature and relative humidity, also were

recorded in the field at the time of sample collection. The samples (on ice) were brought back to the University of Arizona for processing.

Upon arrival in the laboratory, the 1.1-liter sample was evaluated for turbidity (cloudiness) as well as bacterial indicators and pathogens in separate tests. The Colilert® MPN method was used to test for total coliforms and generic *Escherichia coli* in the samples. In addition, these samples were tested for the presence of *Salmonella* species and *E. coli* after the concentration of one liter using membrane filtration. The filter was then placed into an enrichment broth and incubated overnight to increase the bacterial numbers. This broth was then used to inoculate a series of selective media for the isolation and identification of *Salmonella* species and *E. coli*. For *Salmonella* species isolation, this procedure included tetrathionate broth, followed by Rappaport Vassiliadis broth, then followed by streaking for isolation onto xylose lysine deoxycholate (XLD) agar plates. Presumptive *Salmonella* colonies (black centers) were sub-cultured onto CHROMagar™ Salmonella plates for further selection. Presumptive colonies on CHROMagar Salmonella plates (mauve color) were tested using API 20E Biochemical test strips for confirmation of the species identification. For *E. coli* isolation, the procedure included the inoculation of EC broth from the original enriched culture and incubation at 44.5°C for 24 hours (or at 48 hours if negative after 24 hours). The tubes that exhibited both bacterial growth (turbidity) and the production of gas (bubbles collected in an inverted Durham tube) were subcultured onto eosin methylene blue (EMB) agar plates, with overnight incubation at 44.5°C. Presumptive *E. coli* colonies, (darkly pigmented, with or without a metallic sheen) were tested using API 20E biochemical test strips for confirmation of the species identification. The isolates that were confirmed as either *Salmonella* or *E. coli* using API 20E biochemical strips were tested using conventional polymerase chain reaction (PCR) for virulence genes (e.g., *stx1* and *stx2* Shiga toxin genes and the *eae* gene for *E. coli* and the *invA* gene for *Salmonella*).

A total of 8 of the 297 water samples (2.7%) were found to be positive for *Salmonella* species. These samples represented 6 of the 99 triplicate samples (6.1%). For all but one of the triplicate samples, *Salmonella* was detected in only one of the three enriched sample replicates. This finding suggests that the levels of *Salmonella* were very low in these samples. This conclusion is supported by the results of a recent study conducted by our research group in which we found *Salmonella* levels in irrigation waters from these two regions to be quite low (usually from 0.1 to 1.0 CFU per 100 ml) (unpublished data); these levels would not be detected without an enrichment step. Depending upon the serovar, the infectious dose for *Salmonella* usually ranges between 10^5 and 10^{10} (100,000 to 10,000,000,000) bacteria, though this can be lower for the elderly, the very young, and immunocompromised individuals (Kothary and Babu 2001), or in foods that protect *Salmonella* from the antimicrobial effect of stomach acids (Gama et al. 2012). Based on the low levels found in irrigation waters in this region and the even lower levels that would likely make it to the edible portions of food crops via irrigation, these levels pose an extremely low risk to consumers of fresh produce in these regions.

Although none of the 297 irrigation water samples tested positive for the *E. coli* Shiga toxin genes (*stx1* and *stx2*), 13 samples (4.4%) did test positive for the presence of the intimin attachment and effacing virulence gene (*eae*). These represented 10 of the 99 triplicate samples (10.1%). Identification of the *eae* gene without the presence of *stx1* or *stx2* has also been noted previously by several other researchers [Trevor Suslow, presentation at 2017 CPS Annual Research Symposium; Michele Jay-Russell, CPS #2011-RR1 "Investigation of potential reservoirs of Shiga toxin–producing *E. coli* and *Salmonella* in produce production areas of Arizona and Mexico (rapid response)"]. The presence of this virulence gene suggests the potential for these isolates to be pathogenic; however, as with the positive *Salmonella* samples, only one or two of the three replicates was positive, suggesting very low levels of the potentially pathogenic *E. coli* strain.

All 297 samples also were examined for the presence of heterotrophic plate count (HPC) bacteria, which allows for a measurement of the general microbial quality of the irrigation waters. Heterotrophs are bacteria that require organic carbon for growth. Bacteria detected through HPC tests generally include those that are part of the natural (typically non-pathogenic) microbiota of water. HPC values do not directly relate to health risk, but increases in HPC levels can sometimes be associated with contamination. Serial dilutions of the irrigation water samples were performed and 0.1-ml volumes of each dilution were inoculated onto duplicate R2A plates using the spread plate technique. After incubation at 30°C for five days, colony forming units (CFU) on the plates were enumerated to determine the number of HPC bacteria per 100 ml of irrigation water. All of the samples had detectable HPC bacteria. The counts ranged from 4.0×10^3 to 3.2×10^7 CFU per 100 ml, with an arithmetic average of 1.56×10^6 CFU/100 ml and a geometric mean of 5.64×10^5 CFU/100 ml.

Virus analyses:

The collected 3-liter grab samples were concentrated for viruses by passage through stacked pre-filter (pore size of 20 to 25 µm) and negatively charged HA (90 mm in diameter, pore size of 45 µm) filter membranes. Following elution from the HA filters (which results in the primary concentration of the samples from 3 liters down to 10 ml), the samples also underwent a secondary concentration step in which the volume was further reduced to ~0.65 ml via a centrifugal ultrafiltration step. These secondary concentrates were frozen and stored at −80°C pending further evaluation. The sample concentrates were then purified and the nucleic acids extracted using MO BIO PowerViral™ Environmental RNA/DNA Isolation Kits. In a previous study by our research group, this kit was found to be more efficient for the extraction and purification of viral nucleic acids from environmental water concentrates than two other widely used commercial kits; it was also better at removing environmental inhibitors (Iker et al. 2013).

In the detection of viruses by quantitative PCR (qPCR), environmental inhibitory substances can reduce viral nucleic acid extraction efficiency and interfere with cDNA synthesis and/or polymerase activity. This can strongly affect the molecular detection and quantification of viruses in environmental samples. A known concentration of murine norovirus (MNV) was therefore added to each sample prior to the nucleic acid extraction step. The amount of MNV remaining in the sample following extraction and purification was assayed in separate qPCR reactions as a process control to determine the extraction efficiency for each sample. A known concentration of adenovirus 2 (AdV2) was also added to each sample following nucleic acid extraction and assayed in separate qPCR assays as a qPCR inhibition control to determine if there were any substances that were inhibitory for the qPCR in any of the samples.

All 297 samples were then assayed via qPCR for the presence of viral indicators and pathogens. Assays for Aichivirus (AiV), pepper mild mottle virus (PMMV), and enteroviruses were performed with a LightCycler® 480 Real-Time PCR Instrument II (Roche Diagnostics, Mannheim, Germany) using previously established protocols and primer/probe sets. Aichivirus has been detected in high numbers in both raw and treated sewage in the U.S., Japan, Venezuela, and Tunisia, and in shellfish in France. Although not a human pathogen, pepper mild mottle virus is the most abundant virus in fecally-contaminated waters worldwide; enteroviruses (a group of human enteric viruses) are also commonly found in contaminated waters. PMMV was detected in 76 of the 297 samples (25.6%), representing at least one of the three replicate samples in 60 of the 99 (61.0%) triplicate samples. Neither AiV nor the enteroviruses were detected in any of the irrigation water samples.

Next-generation sequencing analyses:

The collected 2-liter grab samples were concentrated for bacteria and protozoans by passage through stacked pre-filter (pore size of 20 to 25 µm) and negatively charged HA filter (90 mm in diameter, pore size of 45 µm) membranes. The filters were placed in sterile petri dishes sealed with parafilm and stored at −80°C pending further evaluation.

For the recovery of microorganisms, the filters were each placed in separate sterile 50-ml conical tubes so that the top side of the filter was facing inwards. The bacteria and protozoans on the HA filters were then eluted from the filters in 10 ml of buffered peptone water with agitation for one minute on a vortexer using a 50-ml tube adapter. The filters were removed from the tubes and discarded, and then the eluting solution was pelleted via centrifugation (10 minutes at ~3,500 ×*g* rpm at room temperature) using a benchtop centrifuge. The supernatant was carefully removed and discarded and the pellet containing microorganisms was re-suspended in 400 µl of sterile physiological saline (0.85% NaCl at pH 7.0). The DNA from the bacteria and protozoans in 200 µl of this suspension was extracted using a bead-based homogenizer (i.e., bead-beater) and purified using a PowerLyzer® PowerSoil® DNA Isolation Kit. This kit is also efficient at the removal of environmental inhibitors. The extracts from the 297 samples were sent to a commercial laboratory to be sequenced using NGS technologies.

This was accomplished through the use of sequencing of genes such as the 16S rRNA gene, which is highly conserved among all bacteria, and the 18S rRNA gene, which is highly conserved among all eukaryotes (e.g., protozoans, fungi, algae). Small differences (e.g., 3% different) in such a conserved gene can then be used to separate microorganisms at the species or near-species level. These defined differences are expressed as operational taxonomic units (OTUs). The DNA in each sample was amplified using PCR assays for both the 16S (to identify bacterial species) and 18S (to identify protozoan species) rRNA genes, and then sequencing reactions were performed using the PCR products using the Illumina HiSeq 2500 RR platform. Approximately 300,000 paired-end 2x250 bp reads were performed per sample. The Caporaso V4 515f – 926r primer sets used by the Earth Microbiome Project were used for the 16S analyses. For general eukaryotic 18S identification, the universal TAReukF and TAReukR primer set that amplifies a ~380 bp region of the eukaryotic 18S rDNA was used. The following steps were performed for the resulting sequences generated:

1. Denoising – corrects errors in reads
2. Chimera Checking – formation of chimeric sequences occurs when an aborted sequence extension is misidentified as a primer and is extended incorrectly in subsequent PCR cycles
3. FASTQ File Generation – text based formatted data files that store the nucleotide sequences generated by the sequencer and their corresponding quality scores encoded as ASCII characters
4. Quality Checking and FASTA Formatted Sequence/Quality File Generation – for use in subsequent analyses
5. Sequence Clustering – places sequences into clusters using an OTU selection program
6. Taxonomic Identification – phylogenetic trees are constructed and sequences are compared to several databases for classification
7. Diversity Analysis

Of the 297 individual samples for bacterial analyses (16S rRNA gene), 16 did not amplify using the primer sets. This result was likely due to PCR inhibitors that were naturally found in the irrigation water samples. Nevertheless, for each sampling location/time, at least one of the triplicate samples was amplified in all but two instances. In other words, of the 99 samples collected in triplicate over the course of the project, sequencing information was available for 97. For the eukaryotic analyses (18S rRNA gene), successful PCR amplification was observed in

far fewer samples. This result was likely primarily due to fewer eukaryotes (e.g., protozoans, fungi, algae) being present in the samples rather than simply inhibition of the PCR (though all of the samples exhibiting inhibition in the bacterial analyses also had limited or no amplification in the 18S analyses). Amplification was observed in only 134 of the 297 individual samples for 18S; sequencing information was available for a total of 62 of the 99 triplicate samples.

For each amplified sample, a Krona diversity chart was generated. An example chart is shown in **Figure 1** in the Appendix. Based on this information, the most abundant organisms per sample were determined. Preliminary analyses identified 21 basic genetic “profiles” for the irrigation waters included in the study. These were based on the top five most abundant organisms present for each set of triplicate samples, and did not consider the entire makeup and diversity of the samples. Nevertheless, some interesting trends were observed. For instance, more diversity was found in Arizona’s Region 2, which historically has had higher levels of contamination than Region 1. Profile 5 was found in 33.3% of the triplicate samples from Region 2; profile 2 was found in 20.0% of the samples. Seven other profiles were each found in approximately 6.7% of samples. For Region 1, nine profiles were also observed; however, they exhibited less overall diversity (found in 38.0%, 25.4%, 18.3%, 8.5%, 2.8%, 2.8%, 1.4%, 1.4%, and 1.4% of samples, respectively). For this region with historically lower contamination levels, a few profiles were found the vast majority of the time (see **Table 1**). This result suggests that with more contamination inputs, greater microbial diversity is observed. In addition, the most commonly found profiles differed greatly between the two regions in Arizona that were included in the study. For instance, profile 1 was found in 38% of samples in Region 1, yet was only found in 6.7% of samples in Region 2. Profile 5, which was found in 33.3% of samples in Region 2, was only found in 1.4% of samples in Region 1 (Table 1).

Large seasonal differences also were observed in both regions (Table 1). For instance, 75.0% of the samples collected during the winter in Region 1 were from profile 1; this same profile was not observed in any samples from this region collected during the summer. For Region 1, 48.6% of the samples collected during the summer were from profile 2, but only 2.8% collected during the winter were from profile 2. For Region 2, the most abundant profile found in the winter was profile 2 (28.6%); the most abundant profile found in the summer was profile 5 (50.0%).

In addition to these regional differences in microbial populations, large differences were observed between primary irrigation canals and return flow (drainage) canals in Region 1 (Table 1). With return flows, the excess water used to irrigate one field is collected (e.g., in drainage canals in this Region) and the water is subsequently used to irrigate fields downstream. Such return flows are known to have lower microbial and chemical water quality as they pick up many contaminants on their way through the fields. As with the areas from Region 2 with historically higher contamination levels than Region 1, the water samples collected from these return flow canals also had greater microbial diversity than the waters collected from the primary irrigation canals. In addition, the microbial profiles from the return flow canals exhibited very little overlap in microbial profiles with the waters collected from primary irrigation canals in either region.

Because of the lack of amplification in numerous samples during the initial PCR step for 16S and 18S rRNA gene detection, troubleshooting was undertaken that resulted in a delay of the final sequencing work. The project team received a 1-month no-cost extension from CPS to extend the project term to January 31, 2018. The sequences that were expected to be completed by late January 2018 were instead not generated until the last week of February 2018. These sequencing analyses require the use of computers with a large amount of processing power (e.g., a “super computer”) and additional time (days to weeks depending on the depth of the analyses). The detailed sequencing analyses were performed by the team after the end of the project term—as of July 2018 these analyses did not yield findings of interest to report.

References cited:
Gama JA, Abby SS, Vieira-Silva S, Dionisio F, Rocha EP. (2012) Immune subversion and quorum-sensing shape the variation in infectious dose among bacterial pathogens. PLoS Pathogens 8(2): e1002503. doi: 10.1371/journal.ppat.1002503.

Iker BC, Bright KR, Pepper IL, Gerba CP, Kitajima M. (2013) Evaluation of commercial kits for the extraction and purification of viral nucleic acids from environmental and fecal samples. Journal of Virological Methods 191(1): 24-30.

Kothary MH, Babu US. (2001) Infective dose of foodborne pathogens in volunteers: A review. Journal of Food Safety 21: 49-73.

Outcomes and Accomplishments

A total of 99 irrigation water samples were collected in triplicate (three replicate samples for a total of 297 samples) from two different fresh produce growing regions in Arizona over a two-year period. Canals with various characteristics (e.g., main canals and lateral canals, cement-lined and non-lined canals, urban and rural canals) were included in the study. In addition, seasonal differences in the microbial water quality were examined by collecting half of the samples during the winter leafy greens growing season and the summer non-growing season. Data collected from previous studies regarding the occurrence of fecal contamination and the presence of pathogens in the irrigation water canals in these regions were studied to ensure that areas with historically lower contamination and higher contamination levels were included. And finally, the microbial water quality of primary irrigation canals was compared to that of water samples collected from return flow (drainage) canals, which included run-off from agricultural fields. In some areas, these return flows are used to irrigate subsequent food crop fields downstream.

The irrigation water samples were examined for a number of microbial, chemical, and physical water quality parameters as well as the canal characteristics (e.g., main canal vs. lateral canal vs. drainage canal, cement-lined vs. unlined, urban vs. rural): the water temperature, pH, turbidity, electro-conductivity, salinity, and total dissolved solids; the numbers of total coliforms and *E. coli* (both generic and those with Shiga toxin–producing *E. coli* [STEC] virulence genes); the numbers of heterotrophic plate count (HPC) bacteria; the presence of *Salmonella* species through enrichment/selection; and the levels of viral pathogens (enteroviruses and Aichiviruses) and viral fecal indicator species (pepper mild mottle virus). **Next-generation sequencing (NGS) microbiome analyses were conducted on the 297 irrigation water samples for both the 16S rRNA gene (bacteria) and the 18S rRNA gene (eukaryotic organisms such as protozoans, fungi, and algae). This likely constitutes the largest collection of metagenomics information ever collected for irrigation waters.** These sequences can be analyzed in a number of different ways in the future to provide information even beyond the scope of the current project. All of the water quality parameters and information on the presence/quantities of microbial indicators/pathogens will be used as meta-data to compare with the microbial populations (both presence and relative abundances) found in each of the samples via NGS. It is the hope of the investigators that specific species (i.e., novel indicator organisms) or shifts in the microbial populations (such as those that have already been observed) can be correlated with changes in the microbial quality of irrigation waters and can provide meaningful information regarding the potential safety of the water for use in the irrigation of food crops.

Summary of Findings and Recommendations

Although the NGS microbiome analyses are ongoing, the preliminary analysis of the sequencing data provides information regarding both significant regional and seasonal differences between irrigation canals in Arizona. In addition, systems/canals with known higher levels of contamination (e.g., those with historically higher levels of contamination/possible contamination inputs and return flow canals that include run-off from agricultural fields) exhibit more microbial diversity than those with historically lower levels of contamination/contamination inputs. In addition, the bacterial population profiles of the most abundant organisms present vary greatly between such areas.

It is therefore apparent that the “normal” composition or “baseline” levels of microorganisms/ populations need to be determined for each region prior to undertaking any NGS evaluation of the effects of various contamination inputs. Nevertheless, once such baseline evaluations have been conducted, NGS can provide valuable information on any changes such inputs have on the microbial water quality. For instance, the baseline predominant organisms for a particular region could be determined and rapid tests developed for their identification (e.g., molecular assays such as PCR). These could be used for the routine monitoring of the irrigation water in the region. A change or shift in the populations could stimulate further testing, such as testing for specific pathogens or other indicator organisms aside from *E. coli* (e.g., specific indicators of fecal contamination), or indicate to the farmer that mitigation steps might be necessary (e.g., waiting to irrigate, allowing for die-off of pathogens between the final irrigation event and produce harvest, and/or treatment of the water source).

This work should result in an approach that will provide region- and/or season-specific identification of species or microbial population shifts, which may be used to identify changes in the microbial quality of irrigation waters. This could allow members of the fresh produce industry to make risk-based assessments of water quality and help to determine when it is safe to irrigate fields.

APPENDICES

Publications and Presentations

This work has previously been presented at the Center for Produce Safety's Annual Research Symposium in 2016 (lightning round talk and poster) and 2017 (full presentation). A presentation of the final results will be given at the 2018 CPS Research Symposium in Charlotte, NC.

A paper for submission in a peer-reviewed scientific publication is currently under preparation and should be submitted within the next six months following the conclusion of the sequencing analyses.

Budget Summary

Funding has been expended to date that includes approximately $52,039 in materials and supplies, $1,612 in travel expenses (gasoline, hotels, per diem), $171,500 in personnel costs, and $8,575 in indirect costs, for a total of ~$233,726 spent on the grant to date. The remaining funds will be spent on travel for the PI to attend the 2018 CPS Annual Research Symposium in Charlotte, NC, in June 2018.

Suggestion to CPS

It would be extremely helpful to allow a portion of investigator salaries (e.g., no more than 25%) to be paid from CPS grants for those faculty members whose salaries are not provided by their respective institution.

Table and Figure

Table 1. Summary of the microbial diversity found in primary irrigation canals and return flow canals from two regions in Arizona. The bacterial microbiome "profiles" are based upon the top five most abundant organisms found across the triplicate samples.

Bacterial Microbiome "Profile"	Region 1 Canals			Region 1 Return Flow Canals			Region 2 Canals		
	Winter	Summer	Total	Winter	Summer	Total	Winter	Summer	Total
1	75.0%	-	38.0%	-	14.3%	9.1%	14.3%	-	6.7%
2	2.8%	48.6%	25.4%	-	14.3%	9.1%	28.6%	12.5%	20.0%
3	11.1%	25.7%	18.3%	-	-	-	-	-	-
4	2.8%	2.9%	2.8%	-	-	-	14.3%	-	6.7%
5	2.8%	-	1.4%	-	-	-	14.3%	50.0%	33.3%
6	-	17.1%	8.5%	-	-	-	-	-	-
7	-	-	-	25.0%	-	9.1%	-	-	-
8	5.6%	-	2.8%	-	-	-	-	-	-
9	-	-	-	-	14.3%	9.1%	-	-	-
10	-	-	-	-	14.3%	9.1%	-	-	-
11	-	-	-	-	-	-	-	12.5%	6.7%
12	-	2.9%	1.4%	-	-	-	-	-	-
13	-	2.9%	1.4%	-	-	-	-	-	-
14	-	-	-	25.0%	-	9.1%	-	-	-
15	-	-	-	25.0%	28.6%	27.3%	-	-	-
16	-	-	-	-	-	-	14.3%	-	6.7%
17	-	-	-	-	-	-	14.3%	-	6.7%
18	-	-	-	-	-	-	-	12.5%	6.7%
19	-	-	-	-	-	-	-	12.5%	6.7%
20	-	-	-	-	14.3%	9.1%	-	-	-
21	-	-	-	25.0%	-	9.1%	-	-	-

Figure 1. Example Krona chart that shows the bacterial diversity in a single sample of irrigation water. The results for *Limnohabitans* species are highlighted in the figure (bottom left) and to the right.